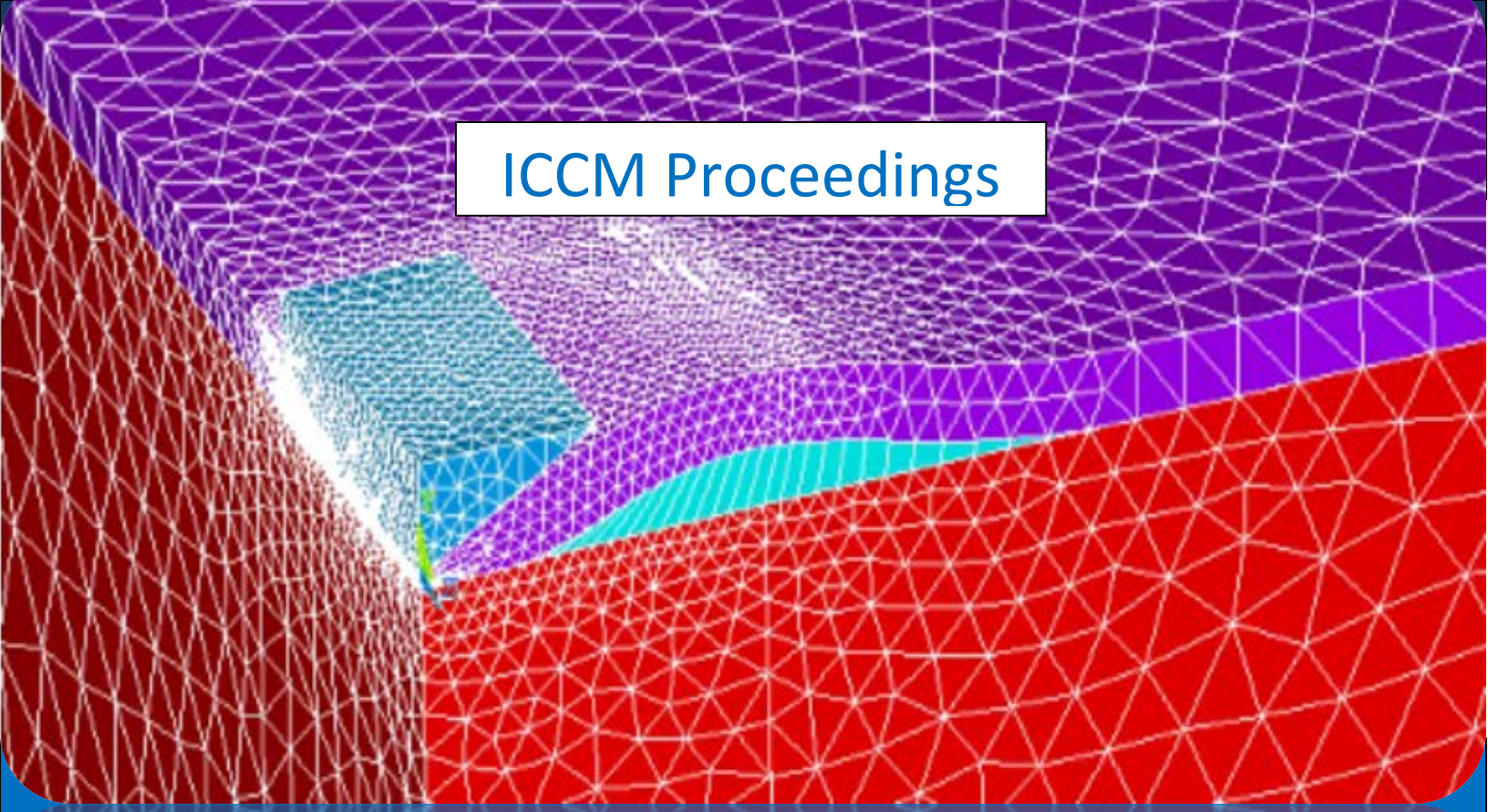


ICCM Proceedings

Proceedings of the International Conference on Computational Methods

(Vol.4, 2017)

8th ICCM, 25th-29th July 2017, Guilin, Guangxi, China

Editors: **G. R. Liu, Xu Han, Zirui Li**

ScienTech Publisher

ISSN 2374-3948 (online)

ICCM2017

Proceedings of the International Conference on Computational Methods (Vol. 4, 2017)

8th ICCM, 25th-29th July 2017, Guilin, Guangxi, China

Edited by

G. R. Liu
University of Cincinnati, USA

Xu Han
Hebei University of Technology, China

Zirui Li
Wenzhou University, China

Proceedings of the International Conference on Computational Methods, Vol.4, 2017

This volume contains full papers accepted by the 8th ICCM, held on 25th-29th July 2017, at Guilin, Guangxi, China.

First Edition: August 2017

International Standard Serial Number: ISSN 2374-3948 (online)

Papers in this Proceedings may be identically cited in the following manner: Author names, *Paper title, Proceedings at the 8th ICCM2017, 25th-29th July, 2017, Guilin, Guangxi, China*, Eds: G.R. Liu, Xu Han, and Zirui Li, Paper ID, ScienTech Publisher.

Published by
Scientech Publisher LLC, USA
http://www.sci-en-tech.com/

PREFACE

Dear Friends and Colleagues,

On behalf of the organizing committee and the co-chairs, we would like to welcome you to the 8th International Conference on Computational Methods (ICCM) in Guilin, China, from July 25th to 29th, 2017. ICCM provides a forum for scholars, researchers, industry practitioners, engineers, and graduate and undergraduate students to present novel work on computational methods, numerical modeling and simulation, and their applications in science and engineering.

Computational Modeling and Simulation are fundamental subjects in engineering and sciences. They can be applied to many of the primary engineering disciplines, including Aerospace, Biomedical, Civil, Chemical, Mechanical, and Materials Engineering among others. Computational Modelling and Simulation cover a broad range of research areas, from conventional structural and mechanical designs, failure analysis, dynamic and vibration analysis, and fluid mechanics to cutting-edge computational mechanics, nano-micro mechanics, multi-scale mechanics, coupled multi-physics problems and novel materials. This is reflected in the variety of disciplines featured in the conference topics.

ICCM was founded and first chaired by Professor Gui-Rong Liu in Singapore in 2014. This was followed by ICCM 2007 in Hiroshima, Japan, ICCM 2010 in Zhangjiajie, China, ICCM 2012 in Gold Coast, Australia, ICCM 2014 in Cambridge, UK, ICCM 2015, in Auckland, New Zealand, and ICCM 2016 in Berkeley, US.

ICCM 2017 in Guilin, China includes over 490 oral presentations in 63 technical sessions, including 4 Plenary Talks, 12 Thematic Plenary Talks, and 80 Keynotes.

ICCM is unique in showcasing the current developments and trends in the general topic of Computational Methods and their relationship to global priorities in science and engineering. The papers scheduled for presentation at ICCM address many urgent and grand challenges in modern engineering and sciences. All ICCM manuscripts were peer-reviewed by independent reviewers. Full papers will be published electronically in an e-Proceedings. Selected abstracts/papers may be invited to be expanded into a journal paper for publication in special issues of some international journals.

We would like to express our sincere gratitude for the contributions of all ICCM 2017 participants and presenters at this international event. The contributions from the International Scientific Committee, Mini-Symposium Organizers, and the expert reviewers and volunteers were invaluable. Special thanks also go to the organizing committee, particularly the supporting staff from Guilin University of Electronic Technology, Guilin Amity Meeting & Exhibition, and ScienTech Services for technical and administrative support, and clear and timely communication.

Finally, we would like to thank you for your valuable contributions to the ICCM 2017 conference. We are looking forward to your participation and continued engagement for the future ICCMs.

Professor GR Liu
Conference Honorary Chairman
University of Cincinnati, USA

Professor Xu Han
Conference Chairman
Hebei University of Technology, China

Professor Zirui Li
Secretary General
Wenzhou University, China

ORGANIZATION COMMITTEE

Conference Chairman
Xu Han (Hunan University / Hebei University of Technology, China)
Daoguo Yang (Guilin University of Electronic Technology, China)

Honorary Chairman
Guirong Liu (University of Cincinnati, USA)
Daining Fang (Peking University, China)

International Co-Chairs
Patrizia Trovalusci (Sapienza Universita Di Roma, Italy, Europe)
Boo Cheong Khoo (National University of Singapore, Singapore, Southeast Asia)
Chongmin Song (University of New South Wales, Australia)
Paulo Pimenta (Universidade de Sao Paulo, Brazil, South America)
Dia Zeidan (German Jordanian University, Jordan, Middle East)
Jagdish Prakash (University of Botswana, Botswana, Africa)

Local Co-Chairmen
Moubin Liu (Peking University)
Chao Jiang (Hunan University)
Xiaowei Gao (Dalian University of Technology)
Tony Sheu (National Taiwan University)
Jiawei Xiang (Wenzhou University)
Dean Hu (Hunan University)
Hongfu Qiang (Rocket Force University of Engineering)

Secretary Generals
Zirui Li (Wenzhou University)
Guiyong Zhang (Dalian University of Technology)

Treasurer
Gang Yang (Hunan University)

Secretaries
Yaoyao Mo (Guilin University of Electronic Technology)
Ying Liu (Hunan University)
An Guo (Hunan University)

International Scientific Advisory Committee

Remi Abgrall (Switherland)
Takayuki Aoki (Japan)
Fumihiro Ashida (Japan)
Santiago Badia (Spain)
Jorge Belinha (Portugal)
Peter Betsch (Germany)
Philipp Birken (Sweden)
Jan Blachut (UK)
Tinh Quoc Bui (Japan)
Zbigniew Buliński (Poland)
Song Cen (China)
Yuki Onishi (Japan)
Marc Oudjene (France)
Pawel Packo (Poland)
Qing Peng (USA)
Umberto Perego (Italy)
Paulo M. Pimenta (Brazil)
Leong Hien Poh (Singapore)
Alexander Popp (Germany)
Qinghua Qin (Australia)
Sinsin Jerry Quek (Singapore)
Roham Rafiee (Iran)

Bin Chen (China)
Chuin-Shan David Chen (Taiwan)
Haibo Chen (China)
Jeng-Tzong Chen (Taiwan)
Jianbing Chen (China)
Lei Chen (USA)
Shaohua Chen (China)
Songying Chen (China)
Gengdong Cheng (China)
Yuan Cheng (Singapore)
Yumin Cheng (China)
Francisco Chinesta (France)
Maenghyo Cho (South Korea)
Fangsen Cui (Singapore)
Haitao Cui (China)
Raj Das (New Zealand)
Linhong Deng (China)
Xiaoping Du (USA)
George Dulikravich (USA)
Chia-Ming Fan (Taiwan)
Zhuojia Fu (China)
Xiao-Wei Gao (China)
M.G.D. Geers (Netherlands)
Yuantong Gu (Australia)
Zhongwei Guan (UK)
Wanlin Guo (China)
Mohamed Hamdaoui (France)
Zhicheng He (China)
Shujuan Hou (China)
Yu Huang (China)
Dariusz Jacek Jakóbczak (Poland)
Chao Jiang (China)
Hiroshi Kanayama (Japan)
Guozheng Kang (China)
Yoon Young Kim (South Korea)
Ioannis Kougioumtzoglou (USA)
Canh Van Le (Vietnam)
Chin-Long Lee (New Zealand)
Ikjin Lee (South Korea)
Tae Hee Lee (South Korea)
Vitor Leitao (Portugal)
Hwa Liang Leo (Singapore)
Yee Tak Andrew Leung (Hong Kong)
Faxin Li (China)
Hua Li (Singapore)
Qing Li (Australia)
Alessandro Reali (Italy)
Janusz Rębielak (Poland)
Daya Reddy (South Africa)
Xiaodan Ren (China)
Erick I. Saavedra Flores (Chile)
Takahiro Saitoh (Japan)
Mikio Sakai (Japan)
Božidar Šarler (Slovenia)
Chang Shu (Singapore)
Piotr Sielicki (Poland)
Aram Soroushian (Iran)
George Stefanou (Greece)
Yu Su (China)
Yuichi Tadano (Japan)
B.C. Vincent Tan (Singapore)
Yuanqiang Tan (China)
Rong Tian (China)
Zhaofeng Tian (Australia)
Liyong Tong (Australia)
Nguyen-Thoi Trung (Vietnam)
Chia Cheng Tsai (Taiwan)
Ken-ichi Tsubota (Japan)
Yuri Vassilevski (Russia)
Decheng Wan (China)
Cheng Wang (China)
Dongdong Wang (China)
Hu Wang (China)
Jie Wang (China)
Jizeng Wang (China)
Lifeng Wang (China)
Wenquan Wang (China)
Xianqiao Wang (USA)
Yanfei Wang (China)
Yue-Sheng Wang (China)
Peijun Wei (China)
Hengan Wu (China)
Yang Xiang (Australia)
Zhihai Xiang (China)
Feng Xiao (Japan)
Xu Xu (China)
Takayuki Yamada (Japan)
Fu-Ling Yang (Taiwan)
Judy Yang (Taiwan)
Qingsheng Yang (China)
Richard Chunhui Yang (Australia)
Z. Yang (China)

Quanbing Eric Li (Hong Kong)
Wei Li (China)
Weiguo Li (China)
Xikui Li (China)
Yan Li (USA)
CW Lim (Hong Kong)
Kian Meng Lim (Singapore)
San-Yih Lin (Taiwan)
Jijun Liu (China)
Moubin Liu (China)
Yan Liu (China)
Yijun Liu (USA)
Yinghua Liu (China)
Zishun Liu (China)
Ping Lu (USA)
Weizhen Jane Lu (Hong Kong)
Zhen Luo (Australia)
Jabareen Mahmood (Israel)
Francesco Mammoliti (Italy)
Karol Miller (Australia)
Rafael Montenegro Armas (Spain)
Sundararajan Natarajan (India)
Fermín Navarrina (Spain)
Perumal Nithiarasu (UK)
Xiaodong Niu (China)
Xesús Nogueira (Spain)
Francesco Noto (Italy)
Masao Ogino (Japan)
Kenji Oguni (Japan)
Jianyao Yao (China)
Zhenhan Yao (China)
Hongling Ye (China)
Jingjie Yeo (Singapore)
Shinobu Yoshimura (Japan)
Zohar Yosibash (Israel)
Byeng Dong Youn (South Korea)
Chengxiang Rena Yu (Spain)
Mengyan Zang (China)
Sergey Zelepugin (Russia)
Kaiyang Zeng (Singapore)
Haifei Zhan (Australia)
Aman Zhang (China)
Chuanzeng Zhang (Germany)
Guiyong Zhang (China)
Jian Zhang (China)
Jianming Zhang (China)
Liangchi Zhang (Australia)
Lihai Zhang (Australia)
Nianmei Zhang (China)
Xiong Zhang (China)
Zhao Zhang (China)
Zhennan Zhang (China)
Zhiqian Zhang (Singapore)
Kejie Zhao (USA)
Liguo Zhao (UK)
Hong Zheng (China)
Yao Zheng (China)
Zhuo Zhuang (China)

CONTENTS

A discrete-continuum coupled finite element model to simulate all failure modes in fibre reinforced concrete

***H. Zhang, Y.J. Huang, †Z.J. Yang, S.L. Xu**

College of Civil Engineering and Architecture, Zhejiang University, 310058, China.

*Presenting author: huizhangzju@zju.edu.cn
†Corresponding author: zhjyang@zju.edu.cn

Abstract

Fibre reinforced concrete (FRC) exhibits complicated failure modes such as fibre breakage, mortar cracking and spalling, fibre-mortar interfacial debonding, depending on many material properties, geometric dimensions and loading conditions. Most existing numerical models are unable to reproduce the above failure modes that may occur simultaneously or sequentially, mainly due to the difficulty in generating finite element meshes with a large number of randomly-oriented fibres. Herein we develop a discrete-continuum coupled finite element model for FRC capable of effectively simulating all the major failure modes. The continuum damage-plasticity model is used to simulate damage and fracture behaviour of the mortar, while debonding of fibre-mortar interfaces is modelled by nonlinear cohesive interfacial elements. Unique techniques are identified to generate conforming meshes between fibres and the surrounding mortar so that randomly-oriented fibres are easily modelled. The model is validated by simulating single fibre pullout tests with different inclination angles and direct tensile tests of multiple randomly-distributed fibres.

Keywords: Fibre reinforced concrete; Cohesive crack model; Damaged plasticity model; Damage and fracture; Interfacial debonding; Finite element model

Introduction

Fibre reinforced concrete (FRC) is a cementitious composite material making use of fibres' high tensile strength and bridging capability to compensate the low fracture resistance of traditional concrete. At micro/meso-scales, an FRC material is a composite comprising mortar matrix, reinforcing fibres, fibre-mortar interfaces and pores. Its mechanical properties and failure behaviour are very complicated, depending on many factors such as the strength, stiffness and volume/weight fractions of the constituents, the length, shape and distribution of fibres, the size, number and distribution of pores, and inter-phasic interactions such as bond-slip between fibres and mortar. Depending on the relative values of these factors, various failure modes with distinguished characteristics can take place, such as fibre breakage, mortar cracking, crushing and spalling, and fibre-mortar interfacial debonding.

Although many finite element (FE) models with multiple fibres have so far been developed for numerical homogenization of elastic properties, either based on computer-generated fibres with random orientations [1] or more recently, converted directly from realistic images, such as from X-ray computed tomography scanning [2], nonlinear FE models with multiple fibres are still rare. For example, Cunha et al. [3] used the smeared crack model and Yu et al. [4] assumed elastic behaviour for the matrix and both modelled beam-bending tests with a single cohesive crack in the middle. The fibre-matrix interaction was indirectly modelled by equivalent tensile stress-strain constitutive laws of fibres transformed from load-slip curves of the single fibre pullout tests (SFPTs). However, the equivalent transformation remains empirical and arguable. In particular, for the fibre embedded length, which is a key parameter, Cunha et al. [3] used a quarter of fibre length and Yu et al. [4] used the total fibre length

assuming uniform stress distribution along the fibre. It is also difficult to interpret the resultant stress in a fibre, as it mixes the effects of elongation in itself and bond-slip on the fibre-matrix interface.

This study is aimed at developing an innovative, easy-to-implement, discrete-continuum coupled finite element model, which is capable of simulating all the possible failure mechanisms in FRC with a large number of fibres at the micro/meso-scales. In this model, all the material phases, i.e., the fibres, the matrix and the fibre-matrix interfaces are individually and explicitly represented. All the major failure mechanisms at micro/meso-scales can be simulated by the proposed model, depending on the relative material properties of different phases. Moreover, the ambiguous assumptions and reliance on the load-slip curves from macro-scale SFPTs, required by many existing models for interfacial bond-slip simulation, are now not needed. As realistic constitutive laws are assigned for each phase, all the results have clear physical interpretation, and parametric studies can be carried out to optimise the key parameters at the material level.

Methodology

Mesh Generation

Fibres are randomly distributed in the domain. Random numbers are generated and used to define the first fibre's centre point and its orientation. The two end points of the fibre are then calculated as the fibre length is known. The fibre will be shortened if it intersects with the domain boundary. The next fibre is then generated in the same way until the given volume fraction of fibres is reached.

A two-step scheme is devised to generate the mesh. In the first step (see Fig. 1a), all the fibres are set as Part A and the matrix as Part B in ABAQUS. A Boolean operation is then performed to merge the fibres into the matrix, which makes the fibre geometries as boundaries of the matrix so that the fibres share the same nodes as the matrix after meshing (i.e., conforming). In the second step, zero-thickness cohesive elements (COH2D4 in ABAQUS) are inserted between pairs of fibre and matrix elements to model the fibre-matrix interface, using a simple MATLAB code (Fig. 1b). The cohesive elements are arranged in a unique way that both the fibres and interfaces are deformed in the plane. Fig. 1c shows a special case when two fibres intersect with each other. At the intersection, three nodes with the same coordinates are used, i.e., N1 of fibre 1, N2 of fibre 2 and N3 of the matrix.

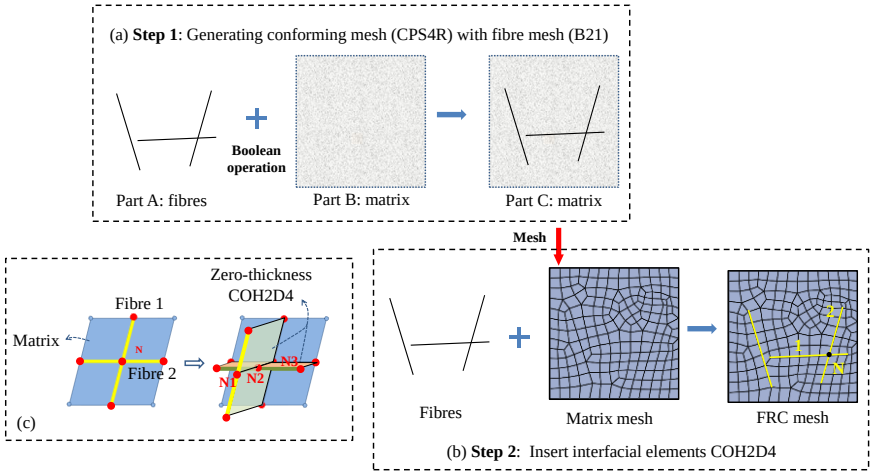


Figure 1. A two-step scheme to generate the FRC mesh

Constitutive Laws

The fibres are modelled using two-noded Timoshenko beam elements (B21 in ABAQUS) with elastic-plastic stress-strain laws with yielding strain hardening and rupture. The bending stiffness and tensile stiffness are calculated based on the cross-sectional area of the fibre as an input.

The mortar matrix is described by the concrete damage plasticity (CDP) model in ABAQUS. The CDP model has been widely used in static and dynamic damage and fracture modelling of concrete [5]-[7]. It has also been used to model the mortar matrix [8].This study assumes linear-elastic pre-peak compressive and tensile stress-strain relations for the matrix. The post-peak softening response in compression (σ_c-ε_c) is described using [9] and the tensile traction (σ_t)-crack opening displacement (w) by [10].

The fibre-matrix interfacial behaviour is simulated by zero-thickness cohesive elements (COH2D4 in ABAQUS) which assumes that there exist a normal traction t_n and a tangential traction (shear cohesion) t_s across the crack surfaces. The tractions decrease monotonically as functions of the corresponding relative displacements of the crack surfaces. In this study, the max normal traction t_n^0 is assumed as 10 times the max shear traction t_s^0 so that the interfacial shear slip governing debonding is dominant. Typical linear and exponential softening curves for t_s-δ_s can be used. The resilient feature of cohesive elements is that its formulation is based on the damage mechanics framework, within which the stiffness k_s upon unloading and reloading is degraded as δ_s increases, due to irreversibly progressive damage. The damage is characterized by a non-negative scalar index D representing the overall damage of the interface caused by all physical mechanisms. Apart from the damage evolution laws, a damage initiation law is also needed. A maximum stress law is used in this study, i.e., when t_s reaches t_s^0, D develops and the tangential shear stiffness k_s starts to degrade.

Benchmark: Modelling of Single Fibre Pullout Tests (SFPTs)

The SFPTs with three fibre inclination angles (θ = 0 °, 30 ° and 60 °) performed by [11] were simulated first using the develop methodology as a benchmark. The fibre has diameter d_f = 0.5 mm and embedment length L_e = 10 mm. The geometry, dimensions and boundary conditions are shown in Fig. 2. The fibre free end is constrained vertically and can only move horizontally. This is to simulate the stress condition of a fibre crossing a crack. The material properties are given in Table 1. The fracture energy for the matrix is G_f = 0.04 N/mm for all the simulations in this study. According to the experiment, the fibre-matrix interfacial bonding strength is $t_s^0 = P_{max}/(\pi d_f l_f)$ = 3.3 MPa, where the maximum pullout force P_{max} was estimated to be 52 N after averaging experimental results for θ = 0 °. The exponential damage evolution law was used for the cohesive elements.

For all the simulations in this study, the ABAQUS/Explicit solver is employed with total time 0.01 s to ensure the quasi-static loading condition for all the simulations in this paper.

Table 1. Material parameters for the SFPTs

	E (GPa)	ν –	ρ (kg/m^3)	f_c (MPa)	f_t (MPa)	f_y (MPa)	f_b (MPa)	ε_b –	L_f (mm)	D_f (mm)	t_s^0 (MPa)
Matrix	27	0.2	2100	36.5	3.0	–	–	–	–	–	–

Fibre	200	0.33	7800	–	–	635	847	0.1	10	0.5	–
Interface	10	0.3	2100	–	–	–	–	–	–		3.3

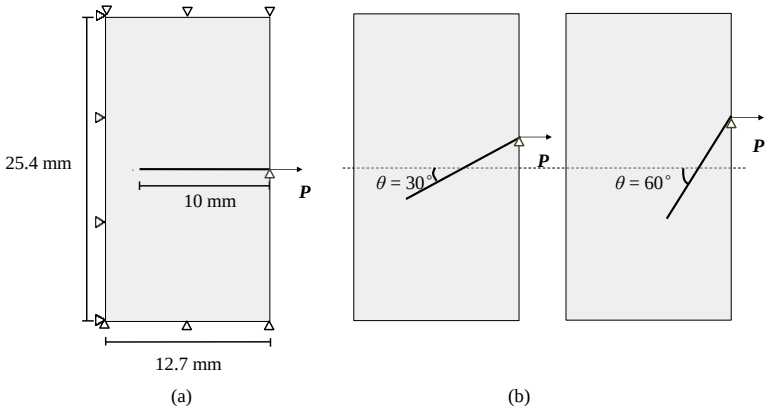


Figure 2. (a) In-plane dimensions and boundary conditions in the pullout tests; (b) fibres inclination angles of 30 °and 60 °

A mesh convergence study was carried out first for the specimen with θ = 30 ° using three meshes as shown in Fig. 3. The three meshes have 25, 50 and 75 divisions in the fibre across a zone with width of 4 mm, 8 mm and 12 mm, respectively. The predicted pullout force-slip (*P-s*) curves are plotted in Fig. 4. It can be seen that virtually identical results were obtained from the medium and fine meshes. Therefore, the medium mesh was used after considering the balance between efficiency and accuracy.

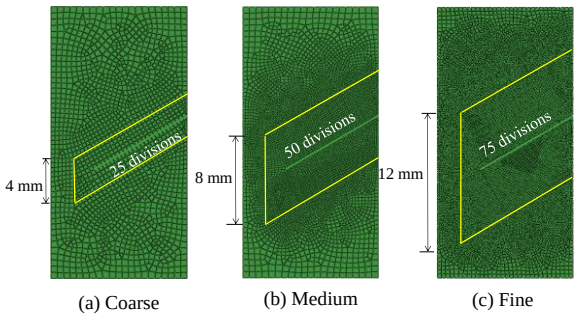


Figure 3. Three meshes for sensitivity study

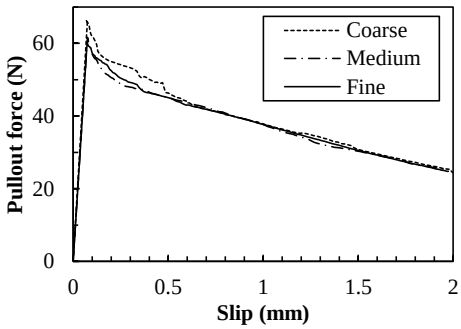


Figure 4. The pullout force slip (*P-s*) relations from three meshes (θ = 30 °)

The predicted P-s curves are compared with the experimental data with good agreement, as shown in Figs. 5a, 5b and 5c for θ = 0 °, 30 °and 60 °, respectively.

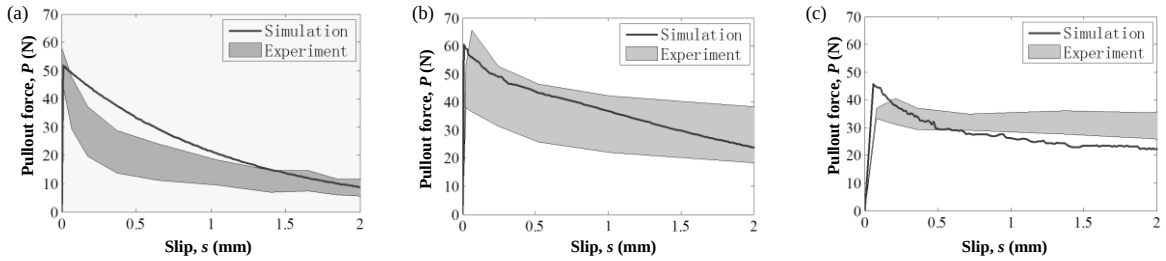


Figure 5. Pullout force-slip (*P*-*s*) relations: (a) θ = 0 °; (b) θ = 30 °; (c) θ = 60 °

Modelling Direct Tensile Tests of UHPFRC Specimens with Multi-Fibres

Direct tensile tests of UHPFRC specimens are simulated herein against the experiment of Yoo et al. [12] to demonstrate the capability of modelling multiple cracks. The material parameters listed in Table 2 are from the experiment. Ten random samples with V_f = 2.0% are modelled and one of them is shown in Fig. 6a with the dimensions and boundary conditions. Fig. 6b shows a local mesh in this sample with typical elemental size 0.5 mm.

Table 2. Material parameters for the UHPFRC specimen

	E (GPa)	ν –	ρ (kg/m^3)	f_c (MPa)	f_t (MPa)	f_y (MPa)	f_b (MPa)	ε_b –	L_f (mm)	D_f (mm)	t_s^0 (MPa)
Matrix	27	0.2	2100	150	9	–	–	–	–	–	–
Fibre	200	0.33	7800	–	–	2500	2800	0.1	13	0.2	–
Interface	10	0.3	2100	–	–	–	–	–	–		7.5

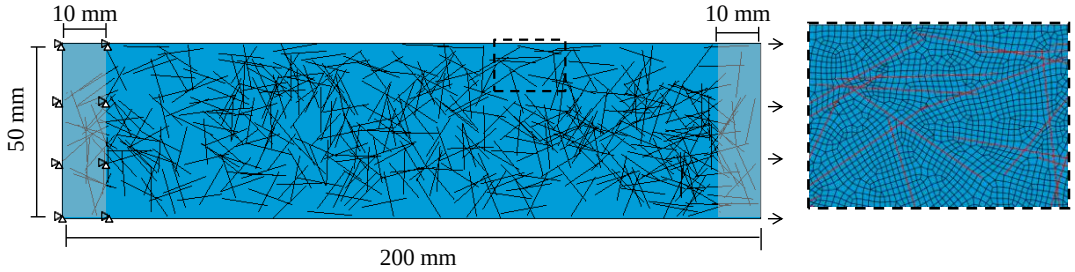


Figure 6. Direct tension of a UHPFRC specimen

The simulated stress-displacement (σ-d) curves from the ten samples are compared with the experimental data in Fig. 7. Excellent agreement can be seen. Again the scatter among the simulation results reflects the different fibre distribution in different samples. The damage and fracture process for one sample is illustrated in Fig. 8 at a few displacements. It can be seen that damage initiates mostly in the matrix areas with fewer fibres before the peak load (Fig. 8a). As the displacement increases, these damaged areas gradually develop into many distributed, mostly parallel cracks near the peak load (Fig. 8b), which is typical for the UHPFRC due to fibre bridging. A very complicated damage and crack pattern follows with strong interaction with fibre-matrix interfacial debonding (Fig. 8c). Eventually, the model fails with one or a few localized major cracks (Fig. 8d).

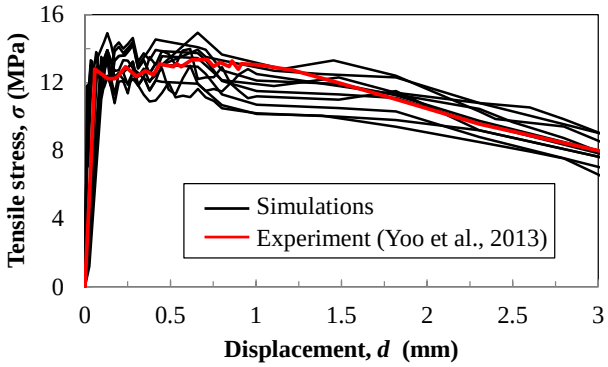


Figure 7. Tensile stress-displacement (*σ*-*d*) curves of SFRC under uniaxial tension for different samples comparing to the experiment result

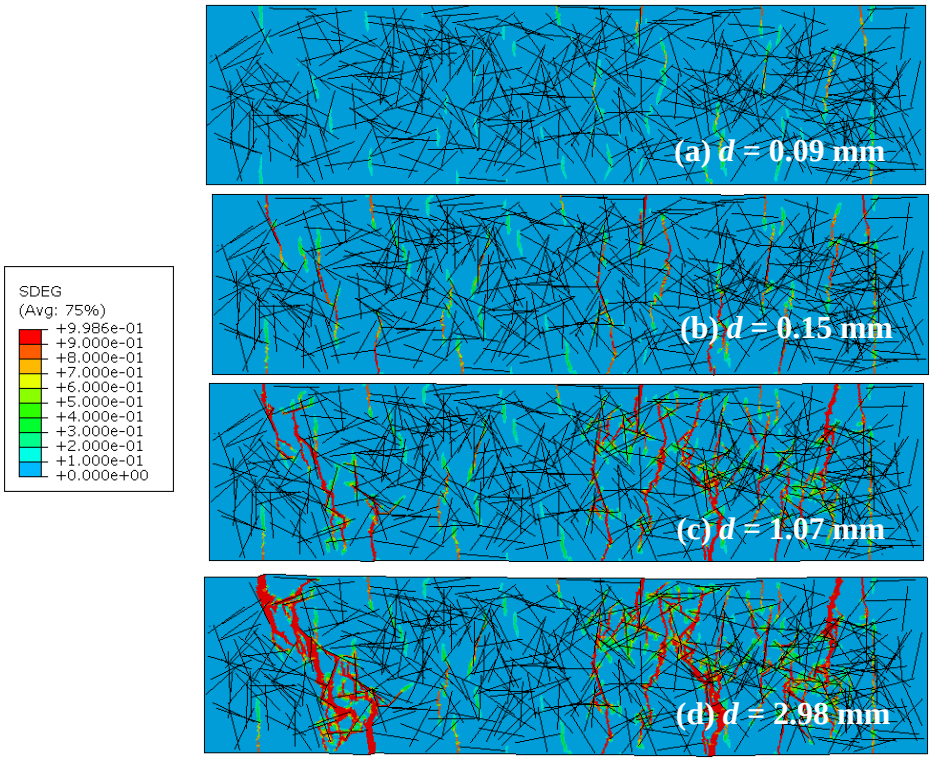


Figure 8. Cracking process of a UHPFRC sample

A cut-off view of Fig. 8d is further shown in Fig. 9 with matrix elements with DAMAGET ≥ 0.9 removed, highlighting the mechanisms such as fibre pullout (A), fibre bending (B, C) and fibre yielding (C) as well as matrix damage due to the snubbing effect of inclined fibres (A, B, C).

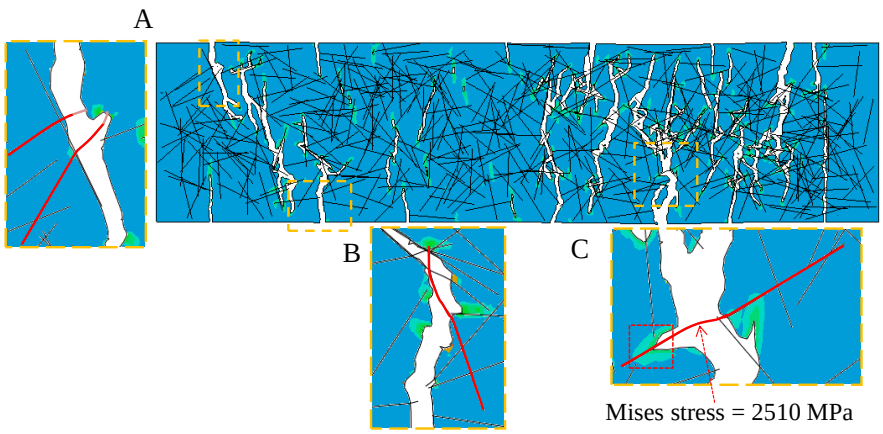


Figure 9. Cut-off view of the fractured sample in Fig. 8d

The failure patterns of two other samples are shown in Fig. 10. It can be seen that the very complicated multi-cracking behaviour of UHPFRC are well captured by the models. The very different failure patterns due to random distribution of fibres are also reported in the experiment [12].

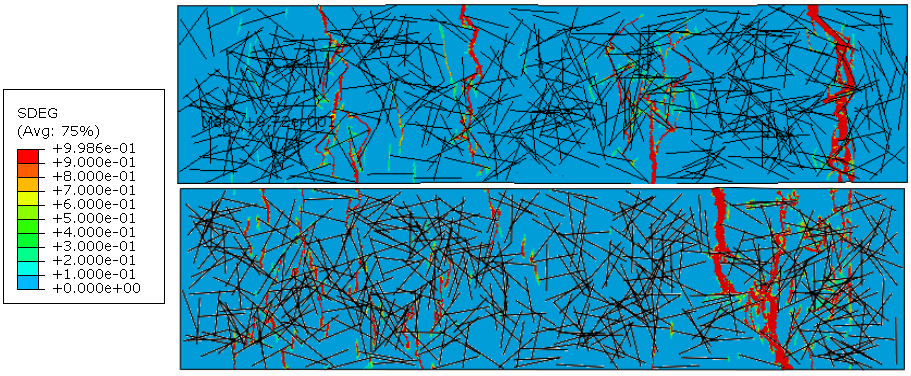


Figure 10. Failure patterns of two other UHPFRC samples

Conclusions

This study has developed an easy-to-implement but effective discrete-continuum coupled modelling approach to simulate complicated nonlinear damage and fracture behaviour of fibre reinforced concrete at the meso-scale. It has been demonstrated that, thanks to explicit, direct modelling of fibres, matrix and fibre-matrix interfaces, this approach is capable of effectively simulating all the possible failure mechanisms that may occur simultaneously or sequentially in one specimen, including fibre pullout, yielding and rupture, interfacial debonding, and matrix cracking and spalling (snubbing effect). This capability allows for parametric studies of key material and geometric factors such as the interfacial bonding strength, matrix tensile strength, fibre volume fraction, and fibre embedment length, so that "optimal" FRC materials can be designed at meso-scale for desired mechanical properties with minimum material cost by optimising these factors.

Acknowledgements

This research is funded by Zhejiang Provincial Natural Science Foundation of China (No. LR14E080002), the Fundamental Research Funds for the Central Universities (No. 172210173) and the National Natural Science Foundation of China (No. 51378462).

References

[1] Gal, E., and Kryvoruk, R. (2011) Meso-scale analysis of FRC using a two-step homogenization approach. *Computers & Structures* **89**, 921-929.

[2] Qsymah, A., Sharma, R., Yang, Z., Margetts, L., and Mummery, P. (2017) Micro X-ray computed tomography image-based two-scale homogenisation of ultra high performance fibre reinforced concrete. *Construction & Building Materials* **130**, 230-240.

[3] Cunha, V. M., Barros, J. A., and Sena-Cruz, J. M. (2012) A finite element model with discrete embedded elements for fibre reinforced composites. *Computers & Structures* **94**, 22-33.

[4] Yu, R. C., Cifuentes, H., Rivero, I., Ruiz, G., and Zhang, X. (2016) Dynamic fracture behaviour in fibre-reinforced cementitious composites. *Journal of the Mechanics & Physics of Solids* **93**, 135-152.

[5] Chen, G. M., Teng, J. G., Chen, J. F., and Rosenboom, O. A. (2010) Interaction between steel stirrups and shear-strengthening FRP strips in RC beams. *Journal of Composites for Construction* **14**, 498-509.

[6] Chen, G. M., Chen, J. F., and Teng, J. G. (2012) On the finite element modelling of RC beams shear-strengthened with FRP. *Construction & Building Materials* **32**, 13-26.

[7] Xu, N., Xu, B., Zeng, X., Jiang, Z., and Chen, J. M. (2010) Dynamic load-displacement behavior of RC shear walls under different loading rates: tests and simulations. *Advanced Materials Research* **163**, 1780-1785.

[8] Huang, Y. J., Yang, Z. J., Chen, X. W., and Liu, G. H. (2016). Monte Carlo simulations of meso-scale dynamic compressive behavior of concrete based on X-ray computed tomography images. *International Journal of Impact Engineering* **97**, 102-115.

[9] Guo, Z.H. Concrete strength and constitutive relation: principle and application. Beijing: China Architecture and Building Press; 2004 ISBN: 9787112062782

[10] Hordijk, D. A. (1992) Tensile and tensile fatigue behaviour of concrete; experiments, modelling and analyses. *Heron* **37,** 1–79.

[11] Leung, C. K. Y., and Shapiro, N. (1999) Optimal steel fiber strength for reinforcement of cementitious materials. *Journal of Materials in Civil Engineering* **11**, 116-123.

[12] Yoo, D. Y., Kang, S. T., Lee, J. H., and Yoon, Y. S. (2013) Effect of shrinkage reducing admixture on tensile and flexural behaviors of uhpfrc considering fiber distribution characteristics. *Cement & Concrete Research* **54**, 180–190.

In Vivo Magnetic Resonance Image-Based 3D Computational Models to Quantify Right Ventricle Morphological and Mechanical Characteristics for Healthy and Patients with Tetralogy of Fallot

Heng Zuo,[1] †*Dalin Tang[1,2] Chun Yang,[1,3] Zheyang Wu,[1] Xueying Huang,[1,4] Rahul H. Rathod,[5] Alexander Tang,[5] Kristen L. Billiar,[6,7] Tal Geva,[5]

[1] Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609
[2] School of Biological Science & Medical Engineering, Southeast University, Nanjing, 210096, China
[3] China Information Tech. Designing & Consulting Institute Co., Ltd., Beijing, 100048, China
[4] School of Mathematical Sciences, Xiamen University, Xiamen, Fujian 361005, China
[5] Department of Cardiology, Boston Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, MA 02115 USA
[6] Department of Biomedical Engineering, Worcester Polytechnic Institute, MA 01609, USA
[7] Department of Surgery, University of Massachusetts Medical School, Worcester, MA 01655

*† Presenting and corresponding author: Dalin Tang, Southeast University, Nanjing China; Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609, email: dtang@wpi.edu , fax: 508 831-5824

Abstract

Patients with repaired tetralogy of Fallot (TOF) account for the majority of cases with late onset right ventricle (RV) failure. It is a challenge to differentiate patient with better outcome after pulmonary valve replacement (PVR) from patients with worse outcome. Comparing TOF patients with healthy people may provide information to address this challenge.

Cardiac magnetic resonance (CMR) data were obtained from 16 TOF patients (8 male, median age, 42.75) and 6 healthy group (HG) volunteers (1 male, median age, 20.1). The patients with positive ejection fraction (EF) changes after PVR form the better-outcome patient group (BPG, n=5). The patients with negative EF changes is called the worse-outcome patient group (WPG, n=11). CMR-based patient-specific computational RV/LV models were constructed to obtain RV wall thickness (WT), volumes, curvature, and mechanical stress and strain for analysis.

At begin-of-ejection, BPG stress was very close to HG stress (54.7±38.4 kPa vs. 51.2±55.7 kPa, p=0.6889) while WPG stress was much higher than HG stress (94.3±89.2 kPa vs. 51.2±55.7 kPa, p=0.0418). BPG RV volume was 43.3% higher than HG RV volume while WPG RV volume was 108.1% higher than that from HG. BPG longitudinal curvature (L-cur) was 65.1% higher than HG L-cur, while WPG L-cur was 26.7% higher than HG L-cur. Circumferential curvature, RV strain and wall thickness did not provide much useful information.

BPG stress was shown to be close to HG stress and stress may be used as an indicator to differentiate BPG patients from WPG patients, with further validations.

Keywords: ventricle modeling, cardiac mechanics, magnetic resonance imaging, normal ventricle, right ventricle, tetralogy of Fallot.

1. Introduction

With the recent development of computational modelling and medical imaging technology, computer modeling and computer-aided procedures become more widely used in cardiac function analysis and patient-specific surgical design, replacing traditional empirical and often risky experimentation to examine the efficiency and suitability of various reconstructive cardiac procedures. Recent reviews are given in [1-4]. In this paper, patient-specific computational models based on cardiac magnetic resonance (CMR) imaging were used to quantify right ventricle morphological and mechanical characteristics for healthy and patients with tetralogy of Fallot (TOF). These information would form basis for further cardiac research and for potential clinical applications treat TOF patients.

Tetralogy of Fallot (TOF) is a congenital heart defect which involves four anatomical abnormalities of the heart: pulmonary infundibular stenosis, overriding aorta, ventricular septal defect and right ventricular hypertrophy. With the introduction of TOF repair surgery, survival of TOF patients has increased a lot starting from the 80s. Recently, the relevant reports show that long-term survival rate for repaired TOF patients decreased significantly after the first two decades of the initial repair [5]. In the third and fourth decade after initial surgery, lots of patients with repaired TOF present severe right ventricle (RV) dilation and dysfunction which is caused by the residual anatomic defects left by initial TOF repair. The defects, including pulmonary regurgitation and scarred myocardium from the ventriculotomy, lead to the late onset RV failure. Pulmonary valve replacement (PVR), which mainly addresses chronic pulmonary regurgitation, is one traditional surgical approach for repaired TOF patients with failing RV. Although the current PVR surgical approaches address pulmonary regurgitation issue, many patients do not experience an improvement in RV function and some show a decline after PVR [6-12].

In our previous publications, 3D computational RV/LV models were constructed based on cardiac magnetic resonance (CMR) imaging data for TOF patients to investigate and optimize PVR surgery. In [13,14], computational ventricular models were used in the comparison between regular PVR surgeries and PVR surgeries with RV remodeling, and PVR surgeries with RV remodeling were found to result in reduced stress/strain conditions in the patch area which may lead to improved recovery of RV function. In [15], computational ventricular models with different patch materials were constructed and solved to evaluate the effect of patch materials on RV function. In [16], computational ventricular models with contracting band were built to investigate the impact of band material stiffness variations, band length and active contraction. These results indicated that computational models were powerful in the investigation of PVR surgeries.

In this study, CMR-based computational RV/LV models were constructed for 6 healthy people and 16 TOF patients. The purposes of this study are: a) obtain RV morphological and mechanical parameters (circumferential and longitudinal curvature, RV stress and strain) for healthy people which are lacking in the current literatures; b) find the differences in morphological and mechanical stress/strain characteristics between TOF patients and healthy people and see if this will help to differentiate better outcome TOF patients from worse outcome TOF patients.

2. Data acquisition, models and methods

2.1 Data acquisition

This study was approved by the Boston Children's Hospital Committee on Clinical Investigation. CMR data were obtained from 22 people (9 male, 13 female; median age, 36.6 years; 16 with TOF, 6 healthy) previously enrolled in our RV surgical remodeling trial [17]. For the 16 TOF patients, CMR data before and 6 months after PVR were available for model construction and analysis. Based on their RV ejection fraction (EF) changes, the patients were categorized into two groups, the Better-Outcome Patient Group (BPG, n=5) which had positive RV EF changes (RV EF change: 3.94 ± 2.20) and Worse-Outcome Patient Group (WPG, n=11) which had negative RV EF changes (RV EF change: -8.88 ± 5.30, p-value: 0.00015). Demographic information, RV volumes, pressure conditions, and RV EF for the participants are summarized in Table 1.

Table 1. Demographic and CMR data for healthy volunteers and TOF patients.

Patient	Sex	Age (y)	Begin-Filling Pressure	Begin-Ejection Pressure	RV EDV (cm^3)	RV ESV (cm^3)	RV EF (%)	ΔEF (%)
Healthy Group								
H1	F	46.7	3.6	22	128.4	46.9	63	-
H2	M	23.6	5	27.9	226.6	105.4	53	-
H3	M	20.8	4.5	24	231.7	107.0	54	-
H4	M	19.4	3.9	23.8	213.5	94.2	56	-
H5	M	17.7	4.2	24.3	233.7	105.5	55	-
H6	M	6.7	4.3	24.8	67.6	28.2	58	-
Mean ± SD		22.5 ±13.2	4.25 ±0.48	24.5 ±1.93	183.6 ±69.4	81.2 ±34.6	56.5 ±3.62	-
Better-Outcome Patient Group								
P1	M	22.5	21.6	31.4	406.9	254.5	37.5	1.4
P2	F	42.0	10	45	323.3	177.8	45.0	4.0
P3	F	14.3	3	29	204.0	104.3	48.8	5.6
P4	F	15.3	2	15	193.7	105.1	45.7	6.6
P5	M	17.0	3	27	188.3	108.3	42.5	2.0
Mean ± SD		22.2 ±11.5	7.92 ±8.29	29.5 ±10.7	263.2 ±97.7	150.0 ±66.2	43.9 ±4.22	3.92 ±2.24
Worse-Outcome Patient Group								
P6	F	38.5	6	28	328.8	196.0	40.4	-3.4
P7	M	47.7	2	31	408.8	254.8	37.7	-2.6
P8	M	50.0	3	33	364.6	239.5	34.3	-2.9
P9	F	56.9	5	41	385.1	184.6	52.1	-18.0
P10	M	11.6	10	36	204.2	121.3	40.6	-8.4
P11	M	43.5	17	65	665.1	464.0	30.2	-15.2
P12	M	54.1	4	63	334.8	170.8	49.0	-7.0
P13	F	49.5	12	52	277.2	151.3	45.4	-5.0
P14	M	17.8	2	30	365.0	178.0	51.2	-9.5
P15	F	44.6	11	50	299.0	186.0	37.8	-12.3
P16	F	45.3	9	49	571.1	371.3	35.0	-13.4
Mean ± SD		41.8 ±14.4	7.36 ±4.82	43.5 ±13.2	382.2 ±131	228.9 ±102	41.2 ±7.27	-8.88 ±5.29

Abbreviations: F: Female; M: male; EDV: end-diastolic volume; ESV: end-systolic volume; EF: ejection fraction.

CMR acquisition procedures have been previously described [13-16,18] and were omitted here. Each CMR data set consists of 30 time steps per cardiac cycle, and each time step data has 9-14 equidistant slices covering ventricles in ventricular short axis from base to apex. Three-dimensional RV/LV geometry and computational meshes were constructed following the procedures described in [13-15]. Figure 1 shows one set of CMR images from a TOF patient before the PVR surgery with segmented contours and re-constructed 3D RV/LV geometries. Our two-layer model construction and fiber orientation information were also provided [2,19].

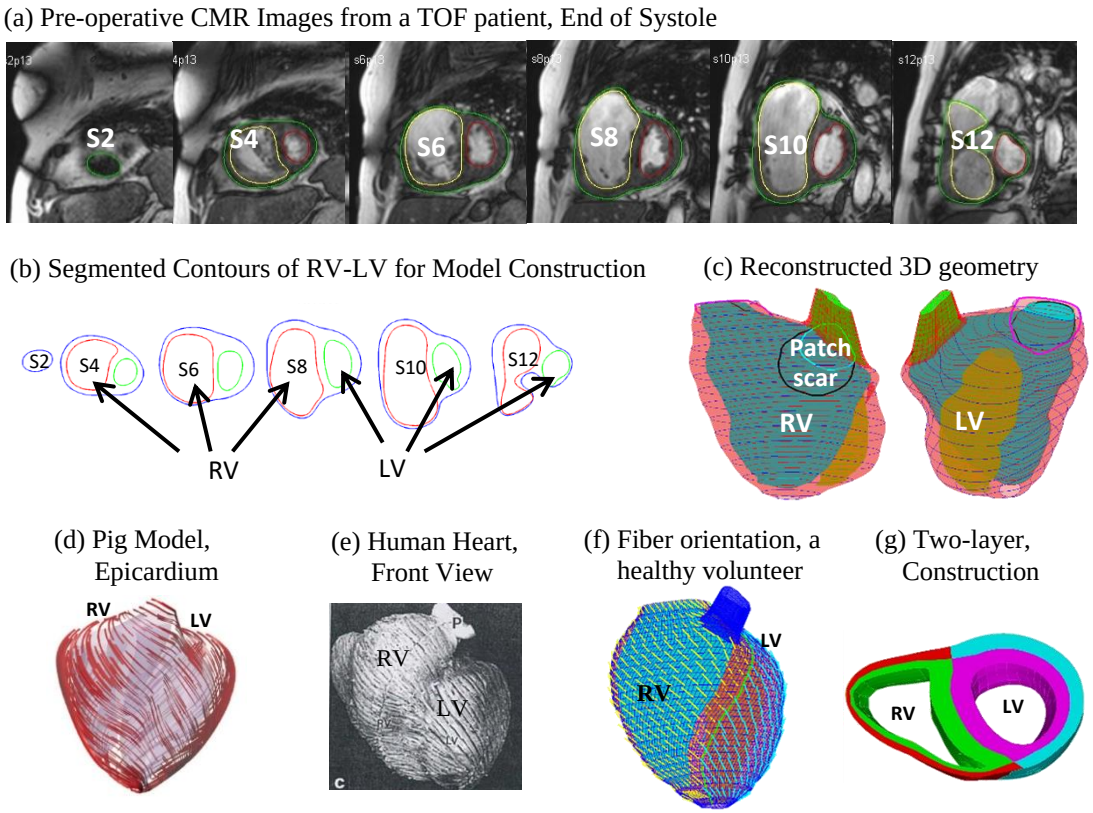


Figure 1. Illustration of model construction procedure using selected CMR image slices from a TOF patient. (a) Pre-operative CMR images of a TOF patient; (b) segmented contours; (c) reconstructed 3D geometry; (d-e) fiber orientation from a pig model [16] and a human heart [13]; (f) fiber orientation from one RV/LV model of a healthy volunteer; (g) two-layer construction.

2.2 The active anisotropic RV/LV models

The ventricular material was assumed to be hyperelastic, anisotropic, nearly-incompressible and homogeneous. Right Ventricular Outflow Tract (RVOT) material, patch and scar were assumed to be hyper-elastic, isotropic, nearly-incompressible and homogeneous. The governing equations for the structure models are:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ij}}{\partial x_j}, \; i = 1,2,3 \tag{1}$$

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial a_i} + \frac{\partial u_i}{\partial a_j} + \sum_l \frac{\partial u_l}{\partial a_i}\frac{\partial u_l}{\partial a_j}\right), \; i,j = 1,2,3 \tag{2}$$

Here $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\varepsilon}$ is Green's strain tensor, $\boldsymbol{u}$ is the displacement, and ρ is material density.

The normal stress on the outer RV/LV surface was assumed to be zero. On the inner RV/LV surfaces, the normal stress was assumed to be equal to the imposed RV/LV pressure conditions:

$$P|_{RV} = P_{RV}(t)\ ,\ P|_{LV} = P_{LV}(t) \tag{3}$$

The nonlinear Mooney-Rivlin model was used to describe the nonlinear anisotropic and isotropic material properties. The strain energy function for the isotropic modified Mooney-Rivlin model (used for patch, scar tissue and RVOT material) was given by Tang et al. [16-18]:

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1[\exp\big(D_2(I_1 - 3)\big) - 1] \tag{4}$$

$$I_1 = \sum C_{ii}\ ,\ I_2 = \tfrac{1}{2}[I_i^2 - C_{ij}C_{ij}] \tag{5}$$

where I_1 and I_2 are the first and second strain invariants, $C = \left[C_{ij}\right] = X^T X$ is the right Cauchy–Green deformation tensor, $\boldsymbol{X}$=$\left[X_{ij}\right] = [\partial x_i/\partial a_j]$ ((x_i) is current position, (a_i) is original position), and c_i and D_i are material parameters chosen to match experimental measurements [13,20].

The strain energy function for the anisotropic modified Mooney-Rivlin model was used for the ventricle tissue [14, 15]:

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1\left[\exp\big(D_2(I_1 - 3)\big) - 1\right] + K_1/(K_2)\exp[K_2(I_4 - 1)^2 - 1] \tag{6}$$

where $I_4 = C_{ij}(\mathbf{n}_\mathrm{f})_i(\mathbf{n}_\mathrm{f})_j$, C_{ij} is the Cauchy-Green deformation tensor, $\mathbf{n}_\mathrm{f}$ is the fiber direction, K_1 and K_2 are material constants. The anisotropic (transversely isotropic) strain-energy function with respect to the local fiber direction was given below [1].

$$W = \frac{C}{2}(e^Q - 1) \tag{7}$$

$$Q = b_1 E_{ff}^2 + b_2(E_{cc}^2 + E_{rr}^2 + E_{cr}^2 + E_{rc}^2) + b_3(E_{fc}^2 + E_{cf}^2 + E_{fr}^2 + E_{rf}^2) \tag{8}$$

where E_{ff} is fiber strain, E_{cc} is cross-fiber in-plane strain, E_{rr} is radial strain, and E_{cr}, E_{fr} *and* E_{fc} are the shear components in their respective coordinate planes, C, b_1, b_2, and b_3 are parameters to be chosen to fit experimental data. It should be noted that Equations (7)-(8) were used because it is desirable to use local coordinate system to identify material parameters which are independent of fiber directions.

Biaxial mechanical testing of human myocardium was performed in Billiar's lab and results were reported in our previous paper (see Figure 2) [21]. Active contraction and relaxation were modeled by material stiffening and softening. In our material model, parameter values c_1, D_1 and C in equations (6) and (7) were adjusted at every CMR time step to match CMR-measured RV volume data for each patient. Patient-specific stress-stretch curves derived from the modified Mooney-Rivlin models for one healthy volunteer at begin of filling and begin of ejection were given in Figure 2 (d). Fiber orientation was set the same way as in our previous papers (see Figure 1) [2,19,21].

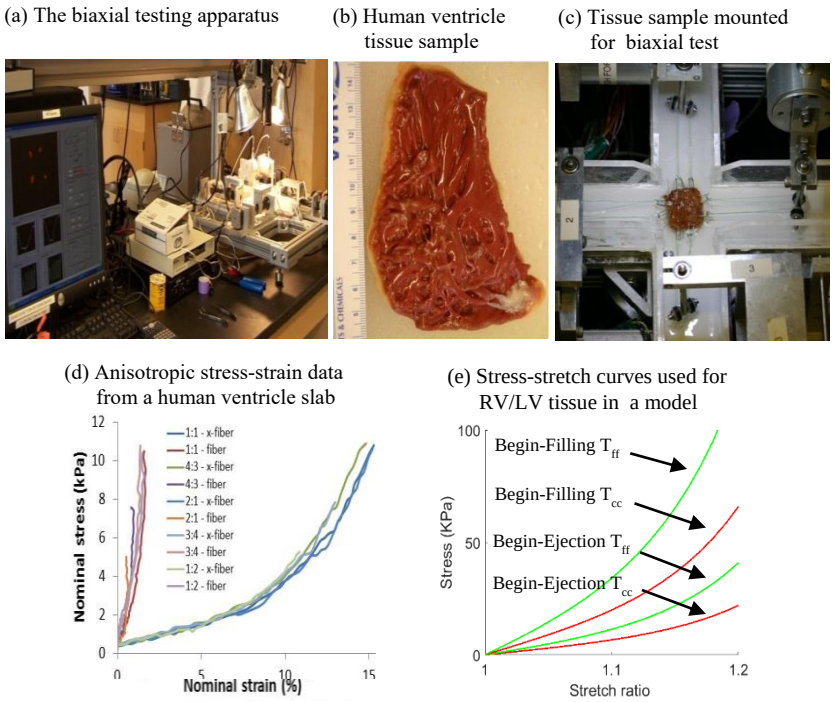


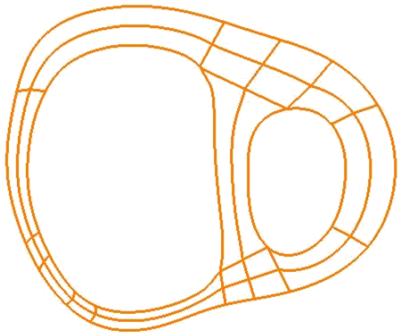
Figure 2. Biaxial mechanical testing and Stress-Stretch curves in the RV/LV model for a healthy volunteer with parameter values chosen to fit CMR data. (a) The biaxial testing apparatus in Dr. Billiar's lab; (b) a human right ventricle tissue sample; (c) tissue sample mounted for biaxial test; (d) anisotropic data from biaxial testing using human RV tissue sample; (e) stress-stretch curves from a healthy volunteer used in our RV/LV model. Model parameter values in Eq. (7)-(8): Begin-Filling (BF): C=27.06 kPa, b_1=8.7875; b_2=1.7005; b_3=0.7743; Begin-Ejection (BE): C=9.02 kPa, b_1=8.7875; b_2=1.7005; b_3=0.7743. T_{ff}: Stress in the fiber direction; T_{cc}: Stress in fiber circumferential direction.

2.3 Geometry-fitting Mesh Generation

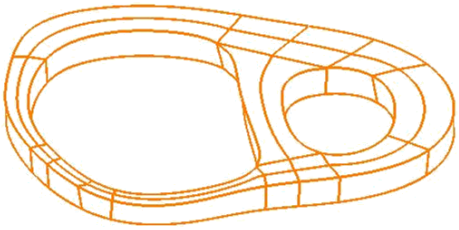
In our patient-specific ventricular models, ventricles have complex irregular geometries which are challenging for mesh generation. A geometry-fitting mesh generation technique was developed to generate mesh for our models. Figure 1(g) gives an illustration of RV/LV geometry between two slices. In each slice, points were firstly defined based on the results of MRI segmentation. Then, lines were defined to divide the slice into geometry-fitting areas (called "surfaces" in ADINA). The neighboring slices were stacked to form volumes. Using this technique, the 3D RV/LV domain was divided into many small "volumes" to curve-fit the irregular ventricular geometry with patch and scar as inclusions. Finally, meshes were generated in each small volume. 3D surfaces, volumes and computational mesh were made under ADINA computing environment. Figure 3 shows the mesh generation technique by using two neighboring slices. For the H1 model constructed in this paper, the finite element ADINA structure model had 20688 meshes. Mesh analysis was performed by decreasing mesh size by 10% (in each dimension) until solution differences were less than 2%. The mesh was then chosen for our simulations.

(a) Points in one slice

(b) Lines and surfaces in one slice



(c) Volumes between two neighboring slices



(d) Meshes between two neighboring slices

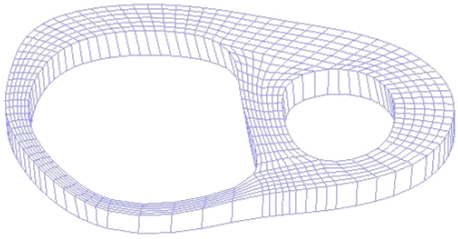


Figure 3. Geometry-fitting mesh generation processing. (a) Points defined in one slice, (b) Lines defined in one slice and lines divided the slice into geometry-fitting surfaces, (c) Volumes defined between two neighboring slices, (d) Generated meshes between two neighboring slices.

2.4 Pre-shrink process

Numerical simulation needs to start from an initial condition where the initial ventricular geometry, pressure and stress/strain conditions of a working heart were provided. Since stress conditions are too hard to be measured in vivo, our numerical simulations started from zero-load ventricular geometries with zero pressure and zero stress/strain distributions. Under the in vivo condition, the ventricles were pressurized and the zero-load ventricular geometries were not known. In our model construction process, a pre-shrink process was applied to the in vivo begin-filling ventricular geometries to generate the starting shape (zero-load ventricular geometries) for the computational simulation. The initial shrinkage for the inner ventricular surface was 2-3% and begin-filling pressure was applied so that the ventricles would regain its in vivo morphology. The ventricular out surface shrinkage was determined by conservation of mass so that the total ventricular wall mass was conserved. Without this pre-shrink process, the actual computing domain would be greater than the actual ventricle due to the initial expansion when pressure was applied.

2.5 Solution methods and morphological and stress/strain data for analysis

The RV/LV computational models (n=22) were constructed and solved by ADINA (ADINA R&D, Watertown, Mass) using finite elements and the New-Raphson iteration method. CMR-measured RV volume and pressure data were used to adjust model parameters so that model-predicted RV volume matched CMR-measure data.

Each ventricle model had 9-14 CMR slices. Every slice was divided into 4 quarters, each with equal inner wall circumferential length. Ventricular wall thickness (WT), circumferential

curvature (C-cur), longitudinal curvature (L-cur), maximal principle stress (Stress-P_1) and maximal principle strain (Strain- P_1) were calculated at all nodal points (100 points per slice, 25 points per quarter). Their average values over the 25 points in each quarter provided the "quarter" values of these parameters. Those values were collected for analysis. The formulas used for calculation of circumferential curvature (κ_c) at each point was

$$\kappa_c = \frac{x'y'' - x''y'}{(x'^2 + y'^2)^{3/2}} \tag{9}$$

The formulas used for calculation of longitudinal curvature (κ) at each point was

$$\kappa = \frac{\sqrt{(z''(t)y'(t) - y''(t)z'(t))^2 + (x''(t)z'(t) - z''(t)x'(t))^2 + (y''(t)x'(t) - x''(t)y'(t))^2}}{(x'^2(t) + y'^2(t) + z'^2(t))^{3/2}} \tag{10}$$

Details can be found from [21].

2.6 Statistical analysis

Continuous variables (RV volumes, WT, C-cur, L-cur, Stress-P_1 and Strain-P_1 values) were summarized as mean standard deviation or median (range). Unpaired Student *t* test was used to compare mean RV volumes between different groups. Due to the small size of data, the quarter mean values were used in the analysis of RV wall thickness, curvatures, Stress-P_1 and Strain-P_1. Similar to what we did in [21], the Linear Mixed-Effect Model (LMM) was used to take care of data dependence structure and compare quarter mean values of RV WT, C-cur and L-cur, Stress-P_1 and Strain-P_1 between different outcome groups.

3. Results

3.1 Results from Healthy Group

Table 2 summarized the average values of the geometrical and mechanical parameters from all the 6 computational models of healthy volunteers at begin of ejection. At the beginning of ejection, mean WT of healthy group (HG) was 0.51 cm. Average C-cur and L-cur from HG were 0.81 1/cm and 0.85 1/cm respectively. Mean RV volume of HG was 183.6 cm^3. Average HG Stress-P_1 and Strain-P_1 were 51.3 kPa and 0.51. These values from the healthy group would be used as the baseline in the following investigation.

Table 2. Summary of mean geometric and stress/strain parameter values at begin of ejection.

	WT (cm)	C-cur (1/cm)	L-cur (1/cm)	RV EDV (cm^3)	Stress-P_1 (kPa)	Strain-P_1
HG						
H1	0.35	1.15	0.87	125.2	70.2	0.63
H2	0.68	0.68	0.56	227.1	51.6	0.58
H3	0.64	0.77	0.68	226.7	36.3	0.40
H4	0.51	0.62	0.61	213.1	56.0	0.39
H5	0.57	0.57	1.27	232.6	45.9	0.57
H6	0.31	0.90	1.11	66.37	48.0	0.51
Mean ± SD	0.51 ±0.15	0.81 ±0.19	0.85 ±0.29	183.6 ±69.4	51.3 ±11.4	0.51 ±0.10
BPG						
P1	0.39	0.47	1.24	406.9	56.9	0.29
P2	0.47	0.43	0.96	323.3	82.4	0.44
P3	0.48	0.50	1.20	204.0	61.9	0.48
P4	0.42	0.53	1.84	193.7	33.5	0.46
P5	0.51	0.53	1.85	188.3	42.0	0.40
Mean ± SD	0.45 ±0.05	0.49 ±0.04	1.42 ±0.40	263.2 ±97.7	55.3 ±18.9	0.41 ±0.08
WPG						
P6	0.34	0.39	0.77	328.8	65.3	0.43
P7	0.65	0.37	1.01	408.8	41.0	0.33
P8	0.49	0.54	1.54	364.6	64.1	0.36
P9	0.48	0.42	0.91	385.1	172.1	0.66
P10	0.41	1.34	1.32	204.2	82.9	0.49
P11	0.80	0.36	0.59	665.1	82.4	0.23
P12	0.71	0.44	0.72	334.8	83.1	0.42
P13	0.45	0.46	0.97	277.2	191.7	0.66
P14	0.43	0.65	1.60	365.0	65.4	0.44
P15	0.46	0.44	1.23	299.0	154.3	0.51
P16	0.59	0.33	1.25	571.1	76.2	0.34
Mean ± SD	0.53 ±0.14	0.52 ±0.29	1.08 ±0.33	382.2 ±131.1	98.0 ±50.1	0.44 ±0.13

Abbreviations: WT: wall thickness; C-cur: circumferential curvature; L-cur: longitudinal curvature; RV: right ventricle.

3.2 Comparison of geometrical parameters: TOF patients have noticeable differences in RV volume, L-cur and C-cur from healthy group

Table 3 summarized and compared the average values of geometrical parameters (RV volume, wall thickness, L-cur and C-cur) between healthy group (HG) and patient group (PG = BPG + WPG). Bar plots of the average values are given in Figure 4 showing group differences. RV volume was the parameter with the most noticeable difference between HG and PG. At the beginning of ejection, average PG RV volume was 87.9% higher than that from HG (344.9±131.3 cm^3 vs. 183.6±69.4 cm^3, p=0.0102). At the beginning of filling, average RV

volume of PG was 151.5% higher than that from HG (204.2±97.9 cm^3 vs. 81.2±34.6 cm^3, p=0.0076). The high percentage difference at begin-filling was due to the fact that RV of PG contracted much less that HG.

C-cur and L-cur also showed large differences between HG and PG. At begin of ejection, mean PG C-cur was 35.8% lower than mean HG C-cur (0.52 ± 1.21 1/cm vs. 0.81 ± 1.05 1/cm, p=0.0237), and mean PG L-cur was 38.4% higher than mean HG L-cur (1.19 ± 1.21 1/cm vs. 0.86 ± 0.71 1/cm, p=0.0756). At begin of filling, average C-cur of PG was 22.9% lower than that from HG (0.64 ± 1.23 1/cm vs. 0.83 ± 0.51 1/cm, p=0.1519), and average L-cur of PG was 23.2% higher than that from HG (1.22 ± 1.22 1/cm vs. 0.99 ± 0.66 1/cm, p=0.2585).

It is worth noting that the ratio of L-cur over C-cur for PG at begin-ejection is 2.29, compared to 1.06 for HG. At begin of filling, the ratio of L-cur over C-cur for PG is 1.90, compared to 1.19 for HG. So PG average RV longitudinal curvature is 100% greater than PG average circumferential curvature, while L-curvature and C-curvature for HG were about equal.

RV WT did not show much differences between HG and PG. That is clear from both Fig. 4 and Table 3.

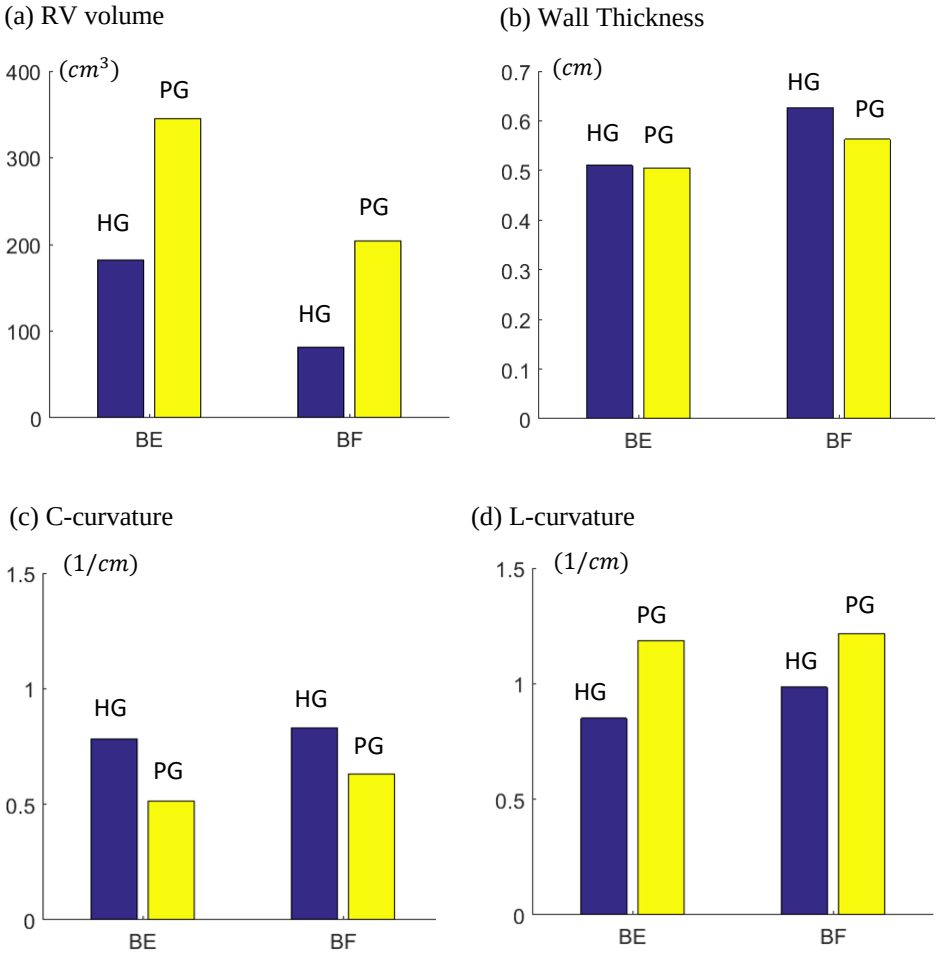


Figure 4. Bar plots comparing average RV volume, WT, C-cur, L-cur values from Healthy Group (HG) and Patient Group (HG) at Begin-Ejection (BE) and Begin-Filling (BF). Blue: HG; Yellow: PG.

Table 3. Comparison of RV volumes, geometric parameters, and stress/strain values between healthy group (HG) and patient group (PG=BPG+WPG) at begin of ejection and begin of filling.

	Begin of Ejection			Begin of Filling		
	PG	HG	P value	PG	HG	P value
RV volume (cm^3)	344.9±131.3	183.6±69.4	0.0102	204.2±97.9	81.2±34.6	0.0076
WT (cm)	0.51±0.24	0.51±0.30	0.9315	0.57±0.27	0.64±0.32	0.3616
C-cur (1/cm)	0.52±1.21	0.81±1.05	0.0237	0.64±1.23	0.83±0.51	0.1519
L-cur (1/cm)	1.19±1.21	0.86±0.71	0.0756	1.22±1.22	0.99±0.66	0.2585
Stress-P_1 (kPa)	82.2±79.4	51.2±55.7	0.1031	7.31±8.49	3.00±2.30	0.0831
Strain-P_1	0.43±0.19	0.51±0.17	0.1486	0.06±0.07	0.07±0.06	0.5376

Data is based on quarter mean values. Values are expressed as mean±standard deviation. Abbreviations as in Table 2.

3.3 Comparison of mechanical parameters: Stress-P_1 shows a large difference between TOF patient group and healthy group

Figure 5 gave stress and strain plots of one healthy volunteer and one TOF patient at Begin-Ejection and Begin-Filling respectively. Without patch and scar, Stress-P_1 and Strain-P_1 distributions of the healthy volunteer were more uniform than that from the TOF patient model near the patch area. Table 3 also summarized and compared RV maximum principal stress and strain (denoted by Stress-P_1 and Strain-P_1) between HG and PG. Figure 6 gave the bar plots of average stress and strain values, showing clear comparisons between healthy group and patient group.

At the beginning of ejection, average Stress-P_1 of PG was 60.5% higher than that from HG (82.2±79.4 kPa vs. 51.2±55.7 kPa, p=0.1031). At the beginning of filling, mean Stress-P_1 of PG was 143.7% higher than that from HG (7.31±8.49 kPa vs. 3.00±2.30 kPa, p=0.0831). The high percentage should be discounted because the overall stress values were small. At begin of ejection, average Strain-P_1 from HG was 18% higher than that from PG. Noticing that average Strain-P1 values from both HG and PG at begin-filling were about the same, higher strain from HG means that healthy ventricles had better contractibility, consistent with our expectations.

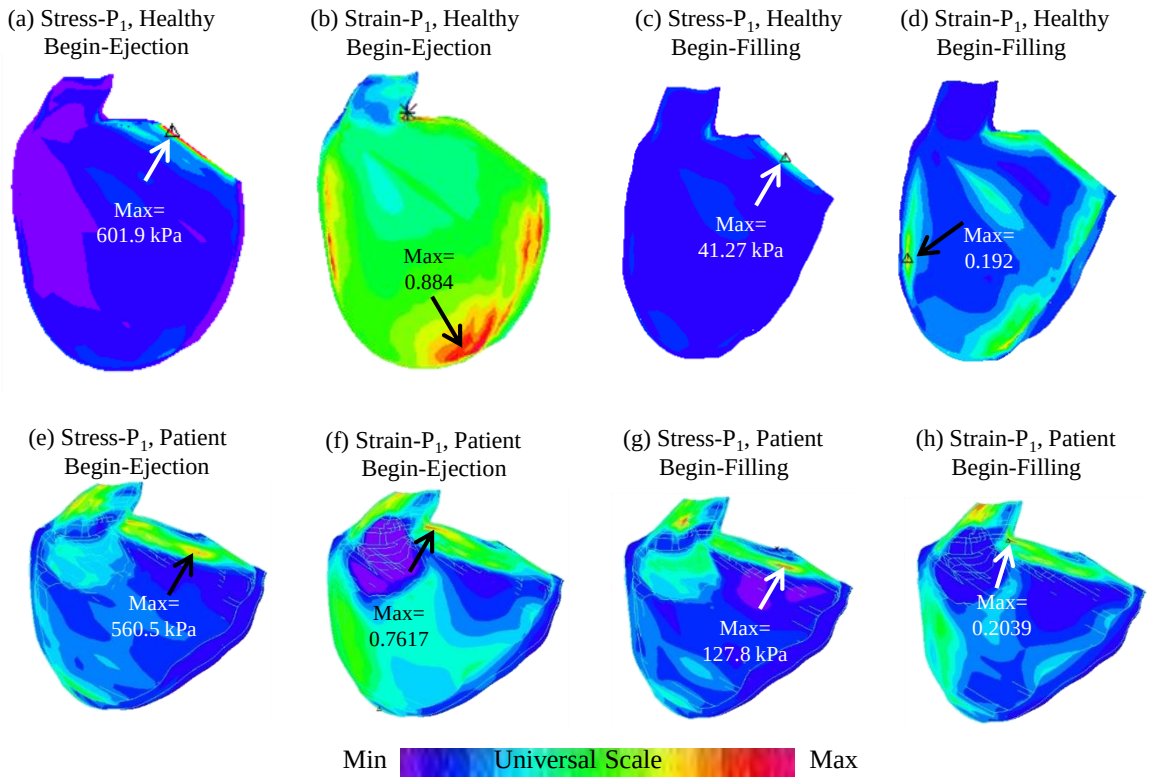


Figure 5. Stress and strain plots from one healthy volunteer (a)-(d) and one TOF patient (e)-(h) showing stress/strain distribution patterns.

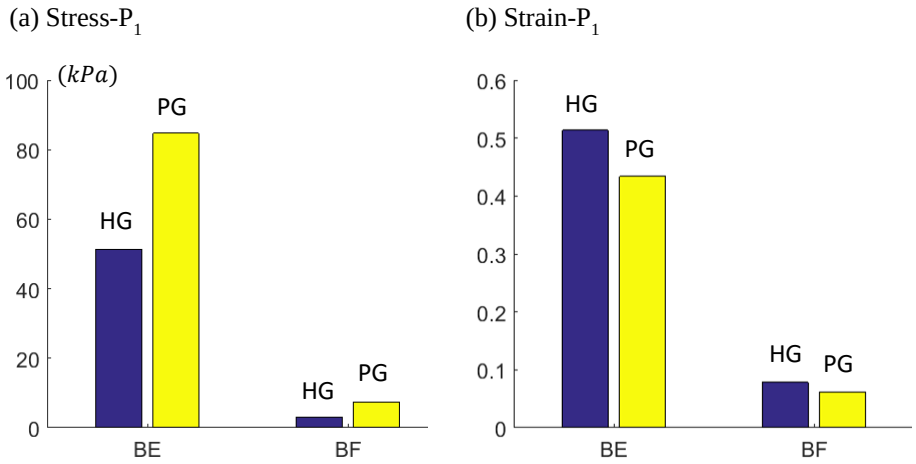


Figure 6. Bar plots comparing average Stress-P_1 and Strain-P_1 values from Healthy Group (HG) and Patient Group (HG) at Begin-Ejection (BE) and Begin-Filling (BF). Blue: HG; Yellow: PG.

3.4 HG may help differentiate BPG from WPG

Table 4 summarized and compared geometrical and mechanical parameter values of BPG and WPG to HG. Figure 7 gave the bar plots of average Stress-P_1, Strain-P_1, RV volume, C-cur, L-cur and WT at begin-ejection, showing the differences among the three groups. Table 4 and Figure 7 showed that differences in wall thickness, C-cur and Strain-P1 between BPG and WPG may not be very useful in differentiating BPG patients from WPG patients.

Stress-P_1 from BPG was found to be closer to that from HG, compared to Stress-P_1 of WPG. At the beginning of ejection, mean Stress-P_1 of BPG was only 6.8% higher than that from HG (54.7±38.4 kPa vs. 51.2±55.7 kPa, p=0.6889), and the difference was not significant; while average Stress-P_1 of WPG was 84.1% higher than that of HG (94.3±89.2 kPa vs. 51.2±55.7 kPa, p=0.0418), and the difference was significant. At the beginning of filling, average Stress-P_1 of BPG was 25% higher than that from HG (3.76±4.17 kPa vs. 3.00±2.30 kPa, p=0.5968), while average Stress-P_1 of WPG was 195.7% higher than that of HG (8.87±9.39 kPa vs. 3.00±2.30 kPa, p=0.0290). The results suggested that comparing patient's RV stress values with healthy RV stress values may help identify patients with possible better outcome.

Similarly, BPG RV volumes at Begin-Ejection were closer to HG RV volumes (263 cm^3 vs. 184 cm^3, 43% higher) compared to WPG volumes (382 cm^3 vs. 184 cm^3, 107% higher). BPG L-curvature was much greater than HG L-curvature at Begin-Ejection (1.42 vs. 0.86 1/cm, 65% higher) than WPG L-cur over HG (1.09 vs. 0.86 1/cm, 27% higher). Based on these results, RV volume and L-cur could be useful in identifying better-outcome patients.

Table 4. Comparison of geometric and stress/strain mean values between healthy group (HG) and patient groups (better-outcome patient group (BPG), worse-outcome patient group (WPG) at begin of ejection and begin of filling.

	Begin of Ejection (maximal volume and pressure)			**Begin of Filling (minimal volume and pressure)**		
	BPG	HG	P value	BPG	HG	P value
RV volume (cm^3)	263.2±97.7	183.6±69.4	0.1482	150.0±66.2	81.2±34.6	0.0534
WT (cm)	0.45±0.20	0.52±0.30	0.4441	0.50±0.21	0.64±0.32	0.1099
C-cur (1/cm)	0.49±0.26	0.81±1.05	0.0094	0.63±0.34	0.83±0.51	0.0082
L-cur (1/cm)	1.42±1.40	0.86±0.71	0.0263	1.58±1.56	0.99±0.66	0.0420
Stress-P_1 (kPa)	54.7±38.4	51.2±55.7	0.6889	3.76±4.17	3.00±2.30	0.5968
Strain-P_1	0.41±0.18	0.51±0.17	0.1042	0.03±0.02	0.07±0.06	0.1047
	WPG	HG	P value	WPG	HG	P value
RV volume (cm^3)	382.1±131.1	183.6±69.4	0.0038	228±102.4	81.2±34.6	0.0041
WT (cm)	0.53±0.26	0.52±0.30	0.8150	0.60±0.29	0.64±0.32	0.6508
C-cur (1/cm)	0.54±1.45	0.81±1.05	0.0709	0.64±1.46	0.83±0.51	0.2427
L-cur (1/cm)	1.09±1.11	0.86±0.71	0.2006	1.07±1.00	0.99±0.66	0.6194
Stress-P_1 (kPa)	94.3±89.2	51.2±55.7	0.0418	8.87±9.39	3.00±2.30	0.0290
Strain-P_1	0.43±0.20	0.51±0.17	0.2603	0.08±0.07	0.07±0.06	0.9860

Data is based on quarter mean values. Values are expressed as mean±standard deviation. Abbreviations as in Table 2.

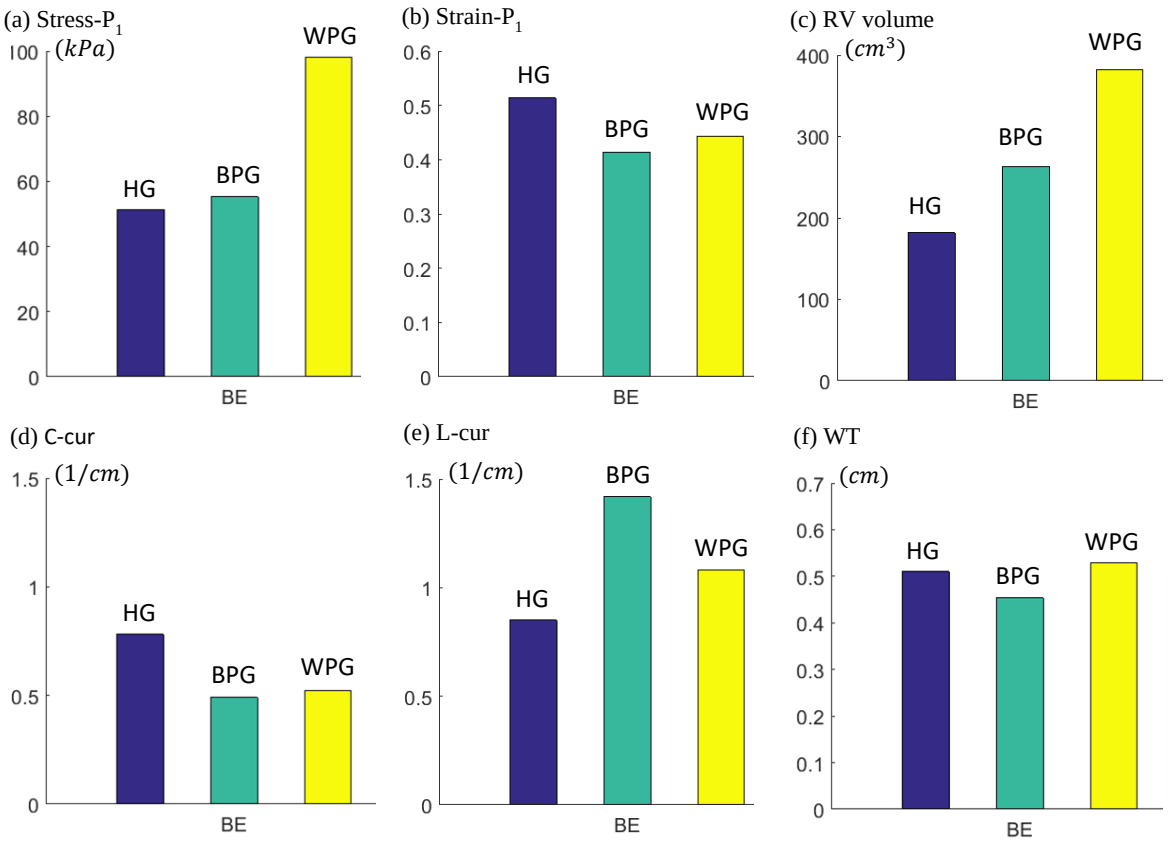


Figure 7. Bar plots comparing average Stress-P_1, Strain-P_1, RV volume, C-cur, L-cur and WT values from Healthy Group (HG), Better-outcome Patient Group (BPG) and Worse-outcome Patient Group (WPG) at Begin-Ejection (BE). Blue: HG; Green: BPG; Yellow: WPG.

4. Discussion

4.1 Modeling techniques for models based on in vivo data with complex geometry

It should be emphasized that the pre-shrink and mesh generation techniques presented in this paper is of general interest for models based on in vivo geometry and of complex structures. In vivo data of organs such as ventricles and arteries are under pressure and internal stress conditions. Most mechanical models require zero-stress geometry as their starting point for stress/strain calculations. Our pre-shrink pressure presented in this paper is a way to obtain the zero-load ventricle geometry as our model starting geometry. Without the shrinking process, as soon as pressure is added to the ventricle, the ventricle will be inflated and its volume will be greater than its in vivo size. This is a major difference between models based on in vivo data and models based on ex vivo data.

It should also be noted that we are using zero-load ventricle geometry in our models, which is not the same as zero-stress geometry. There should still be residual stress in the zero-load geometry. However, obtaining zero-stress geometry involves cutting-open the ventricle to release the residual stress, and then wrapping it up to obtain the residual stress. Not only the numerical procedure is extremely complex, we also do not have real data about how much the ventricle would open to perform the open-close process. Therefore, it should be understood that zero-load geometry was used as an approximation to the zero-stress geometry.

4.2 Motivation to construct models of healthy people

TOF patients have mixed results after PVR. It has remained challenging for the surgeons and clinicians differentiate patients with better outcome from those with worse outcome. This paper is trying to see if information from healthy people could be helpful in meeting that challenge. At the same time, general mechanical stress/strain and morphological information for healthy people will be good contributions since such data are still lacking in the current literature.

It should be explained that our purpose is not only looking for differences between TOF patients and healthy people. We were also trying to find methods and indicators which could help us to separate BPG from WPG by using HG information. As the main result of this paper, it was found that BPG Stress-P_1 and HG Stress-P_1 were close to each other. In fact, they were not statistically different. This indicates RV stress could be a biomarker to be used for possible prediction of post-PVR outcome. RV volume and longitudinal curvature could serve the same purpose in a similar way.

4.3 Limitations

One limitation of this study is the small sample size which results in limited statistical power. The reason for the small sample size is the extensive amount of the time required for constructing each computational model. Under the current status of computer technology, it takes approximately 1 month to generate one 3D patient-specific model. Thus, improving the model-building technique to make the process less labor-intensive and more clinically applicable will be a major effort of our future work.

Several improvements can be added to our current models for more accurate results: a) fluid-structure interactions can be added to obtain blood flow velocity and shear stress which can be also included in the investigation of predictors for good recovery after PRV; b) patient-specific and location-specific measurements of tissue mechanical properties (such as MRI with tagging) will be very desirable for improved accuracy of our models; c) inclusion of patient-specific fiber orientations; d) inclusion of pulmonary valve mechanics in the current model.

5. Conclusion

Our preliminary results indicated that RV stress from the better-outcome patient group was close to stress from the healthy group and could be used as a potential indicator to differentiate BPG patients from WPG patients, with further validations.

Funding:

This research was supported in part by National Heart, Lung and Blood Institute grants R01 HL089269 (PI del Nido, Tang, Geva), R01 HL63095 (PI del Nido) and 5P50HL074734 (PI: Geva). Tang's research was also supported in part by National Sciences Foundation of China grants 11672001, 81571691.

Conflict of Interest:

The authors declare that they have no conflict of interest.

Ethical approval:

This study was approved by the Boston Children's Hospital Committee on Clinical Investigation. The Boston Children's Hospital IRB approval number is: IRB-CRM09-04-0237.

References

1. McCulloch, A. D., Waldman, L., Rogers, J., Guccione, J. M. (1992) Large-scale finite element analysis of the beating heart. *Critical Reviews in Biomedical Engineering* **20,** 427-449.
2. Hunter, P. J., Pullan, A. J., Smaill, B. H. (2003) Modeling total heart function. *Annual Review of Biomedical Engineering* **5,** 147-177.
3. Pfeiffer, R. E., Tangney, R. J., Omens, H. J., McCulloch, A. D. (2014) Biomechanics of cardiac electromechanical coupling and mechanoelectric feedback, *Journal of Biomechanical Engineering* **136,** 021007.
4. Trayanova, A. N., Rice, J. J. (2011) Cardiac electromechanical models: from cell to organ, *Frontiers in Physiology* **2,** 00043.
5. Nollert, G., Fischlein, T., Bouterwek, S., Bohmer, C., Klinner, W., Reichart, B. (1997) Long-term survival in patients with repair of tetralogy of Fallot: 36-year follow-up of 490 survivors of the first year after surgical repair. *Journal of The American College of Cardiology* **30,** 1374-1383.
6. McKenzie, E. D., Khan, M. S., Dietzman, T. W., Guzm an-Pruneda, F. A., Samayoa, A. X., Liou, A., et al. (2014) Surgical pulmonary valve replacement: a benchmark for outcomes comparisons. *Journal of Thoracic and Cardiovascular Surgery* **148,** 1450-1453.
7. Waien, S. A., Liu, P. P., Ross, B. L., Williams, W. G., Webb, G. D., McLaughlin, P. R. (1992) Serial follow-up of adults with repaired tetralogy of Fallot. *Journal of The American College of Cardiology* **20,** 295-300.
8. Burchill, L. J., Wald, R. M., Harris, L., Colman, J. M., Silversides, C. K. (2011) Pulmonary valve replacement in adults with repaired tetralogy of Fallot. *Journal of Thoracic and Cardiovascular Surgery* **14,** 92-7.
9. Therrien, J., Siu, S. C., McLaughlin, P. R. (2000) Pulmonary valve replacement in adults late after repair of tetralogy of Fallot: are we operating too late? *Journal of The American College of Cardiology* **36,** 1670-1675.
10. Meijboom, F. J., Roos-Hesselink, J. W., McGhie, J. S., Spitaels, S. E., van Domburg, R. T., Utens, L. M., et al. (2008) Consequences of a selective approach toward pulmonary valve replacement in adult patients with tetralogy of Fallot and pulmonary regurgitation. *Journal of Thoracic and Cardiovascular Surgery* **135,** 50-55.
11. Vliegen, H. W., Van Straten, A., De Roos, A., Roest, A. A., Schoof, P. H., Zwinderman, A. H., et al. (2002) Magnetic resonance imaging to assess the hemodynamic effects of pulmonary valve replacement in adults late after repair of tetralogy of Fallot. *Circulation* **106,** 1703-1707.
12. Tweddell, J. S., Simpson, P., Li, S. H., Dunham-Ingle, J., Bartz, P. J., Earing, M. G., et al. (2012) Timing and technique of pulmonary valve replacement in the patient with tetralogy of Fallot. *Journal of Thoracic and Cardiovascular Surgery* **15,** 27-33.
13. Tang, D., Yang, C., Geva, T., del Nido, P. J. (2010) Image-based patient-specific ventricle models with fluid-structure interaction for cardiac function assessment and surgical design optimization. *Progress in Pediatric Cardiology* **30,** 51-62.
14. Tang, D., Yang, C., Geva, T., del Nido, P. J. (2008) Patient-specific MRI-based 3D FSI RV/LV/Patch models for pulmonary valve replacement surgery and patch optimization. *Journal of Biomechanical Engineering* **130,** 041010.
15. Tang, D., Yang, C., Geva, T., Gaudette, G., del Nido, P. J. (2011) Multi-physics MRI-based two-layer fluid-structure interaction anisotropic models of human right and left ventricles with different patch materials: cardiac function assessment and mechanical stress analysis. *Computers & Structures* **89,** 1059-1068.
16. Yang, C., Tang, D., Geva, T., Rathod, R., Yamauchi, H., Gooty, V., et al. (2013) Using contracting band to improve right ventricle ejection fraction for patients with repaired tetralogy of Fallot: a modeling study using patient-specific CMR-based 2-layer anisotropic models of human right and left ventricles. *Journal of Thoracic and Cardiovascular Surgery* **145,** 285-293.
17. Geva, T., Gauvreau, K., Powell, A. J., Cecchin, F., Rhodes, J., Geva, J., et al. (2010) Randomized trial of pulmonary valve replacement with and without right ventricular remodeling surgery. *Circulation* **122(11 Suppl),** S201-208.
18. Tang, D., Yang, C., Geva, T., del Nido, P.J. (2014) Right ventricular local longitudinal curvature as a marker and predictor for pulmonary valve replacement surgery outcome: an initial study based on preoperative and postoperative cardiac magnetic resonance data from patients with repaired tetralogy of Fallot. *Journal of Thoracic and Cardiovascular Surgery* **147,** 537-538.
19. Sanchez-Quintana, D., Anderson, R., Ho, S. Y. (1996) Ventricular myoarchitecture in tetralogy of Fallot. *Heart* **76,** 280–286.
20. Humphrey, J. D. (2002) Cardiovascular Solid Mechanics. *New York: Springer-Verlag*.
21. Tang, D., Yang, C., del Nido, P. J., Zuo, H., Rathod, R. H., Geva, T., etc. (2015) Mechanical stress is associated with right ventricular response to pulmonary valve replacement in patients with repaired tetralogy of Fallot mechanical stress, *Journal of Thoracic and Cardiovascular Surgery* **151,** 687-694.

Influence factors of teachers' pro-industry teaching demand adjust Industry 4.0

†* Chun-Mei Chou [1], Chien-Hua Shen [2], His-Chi Hsiao[3], Tsu- Chuan Shen [4] , Yu-Jen Tsen[5] ,Kuan-Fu Shen[6], Hsiang-Li Shen[7] and Jia- Ming Chen [8]

[1]Institute of Vocational and Technological Education National Yunlin University of Science & Technology, Yunlin, Taiwan
[2] Department of Business Administration Transworld Institute of Technology Yunlin , Taiwan
[3] Department of Business Administration ,Cheng Shin University Kaohsiung, Taiwan
[4] Sacred Hearts High School, Yunlin , Taiwan
[5] Center for General Education, TransWorld University
[6] Chien Hsin University of Science and Technology Department of Finance, Taoyuan, Taiwan
[7] Chien Hsin University of Science and Technology Secretariat Office, Taoyuan, Taiwan
[8] Institute of Vocational and Technological Education, National Yunlin University of Science & Technology, Yunlin, Taiwan
*Presenting / Corresponding author: choucm@yuntech.edu.tw

Abstract

This study examines 322 vocation teachers' pro-industry teaching demand and its influencing factors to serve as a school reference for adjust industry 4.0. The results show that teachers' industry development has a significant direct effect on vocation teachers' teaching pro-industry demand, and pro-industry self-efficacy has a significant effect on pro-industry teaching demand through industry development. The influence pattern and empirical data of pro-industry self-efficacy and industry development on pro-industry teaching demand has a good fit. This will serve as a reference for vocation schools' encouragement of teacher professional development and industry need.

Keywords: cognitive apprentice; teacher development; pro-industry teaching; industry 4.0

Introduction

It is affected by the aging of the young population in the developed countries. They are faced with the problem of reducing the labor force. In 2011, the United States launched the AMP project and the " CPS program ", Germany in 2012 to promote the" Industrial 4.0 "program to enhance the international manufacturing leadership, Japan in 2013 launched the "Japan Industrial Revitalization Plan" to enhance equipment research and development, South Korea in 2014 proposed" manufacturing innovation 3.0 " Strategy, to help small and medium manufacturing to establish the wisdom and optimization of production, China also issued in 2015 "Made in China 2025" program[1] [2].

Facing the problem of aging, low birthrate and lack of industrial workers, the coming of the industrial age of 4.0 will be an opportunity for Taiwan. The Taiwanese government has promoted the "Productivity 4.0" and intellectualized policies of Taiwan's manufacturing industry. It is important that Taiwan's industrial development and put forward its unique value proposition to keep the country and the industry competitive in the global competition[3] [4]. Therefore, teachers' pro-industry teaching need to understand the results of teaching strategies, teachers need to enhance the interpretation and demand connotation of industrial transformation in the process of industrial transformation[5] [6]. It is important that depending on teachers' self-reflection, professional knowledge and rich teaching experience to construct [7] [8].

The above research questions and their results are all based on the vocational training center of the teachers in the industry 4.0, as a follow-up as soon as possible to adjust the occupation class pre-service teachers It is also an urgent motive to study the content of the teaching specialized course in the teaching industry, and to establish the cognitive mechanism of the

professional education of the pre-occupation teachers. The purposes of this study are to address the 2 following issues.

1. There is no significant correlation between tertiary teachers' pro-industry self-efficacy, industry development and pro-industry teaching demand.
2. Influence models of tertiary teachers pro-industry self-efficacy, industry development, and pro-industry teaching demand fit the data collected by this study.

Methodology

Subjects

This study treats 338 teachers from vocation schools as the population, and adopts random sampling and cluster sampling for survey.

Measure

The research tool is a "Questionnaire of Influence Factors Vocation Teachers' Pro-industry teaching demand." The questionnaire includes industry development scale, pro-industry self-efficacy scale and pro-industry teaching demand scale [7] [8] [9] [10]. The "Questionnaire of Influence Factors Vocation Teachers' Pro-industry teaching demand" was reviewed by three experts for subject contents suitability to ensure the scale expert validation. Five vocation teachers were invited to answer the questionnaire to enhance the validity of the scales contents. In addition, Ten vocation schools were selected for a pre-test, and 126 teachers were selected as the pre-test objects in total. The scales used in this study are in self-assessment form, and a Likert 5-point scale is used as the scoring method. There are five levels of choices from "agree" to "do not agree" five equal portions of 5, 4, 3, 2 and 1 are given in this order. The higher the score an individual receives, the larger extent of agreement the individual has. The scales factors, number of questions reliability and validity are shown in Table1.

Table 1. An overview of factors, number of questions, reliability and validity for vocation teachers' industry development, pro-industry self-efficacy and pro-industry teaching demand scale

Factor name	No.	Cronbach α	Factor loading	Total reliability Cronbach α	Accumulated explained variance
Pro-industry self-efficacy scale					
Personal efficacy	4	.89	20.15%		
Teaching efficacy	5	.88	19.23%	.87	58.28%
Industry efficacy	4	.87	18.90%		
Industry development scale					
Industry competition	6	.86	19.98%		
Curriculum reform	4	.87	18.74%	.85	57.73%
Pro-industry teaching	4	.85	19.01%		
Pro-industry teaching demand scale					
Teaching practice	5	.87	20.15%	.86	57.55%
Industry practice	4	.86	19.22%		
Pro-industry in-service	4	.88	18.18%		

Data analysis

In processing the survey data used in this study, the collected questionnaires were coded, and Statistical Package for Social Science (SPSS version 12.0) and linear structural analysis (LISREL version 8.5) were used to verify the correlation among the factors of "industry development", "pro-industry self-efficacy" and " pro-industry teaching demand" variables and their effects in order to achieve the purpose of this study. In this study, the statistical test level $\alpha = 0.05$.

Results

The empirical results of vocation teachers' pro-industry teaching demand are shown in Figure 1, and are analyzed as follows:

The estimated value of the direct affecting parameter between industry development and pro-industry self-efficacy is 0.46 ($t = 7.28$, $p<.05$). This means that industry development has a significant effect on pro-industry self-efficacy. The estimated value of the direct affecting parameter between industry development and pro-industry teaching demand is 0.86 ($t = 6.58$, $p<.05$). This means that industry development does necessarily have a significant effect on pro-industry teaching demand. The estimated value of the direct affecting parameter between pro-industry self-efficacy and pro-industry teaching demand" is 0.54 ($t = 7.32$, $p<.05$). This means that pro-industry self-efficacy has a significant effect on pro-industry teaching demand. In summary, in this study of vocation teachers' pro-industry teaching demand and its influence pattern, pro-industry self-efficacy has a significant effect on industry development, but does not have a significant effect on pro-industry teaching demand. Industry development has a significant effect on pro-industry teaching demand.

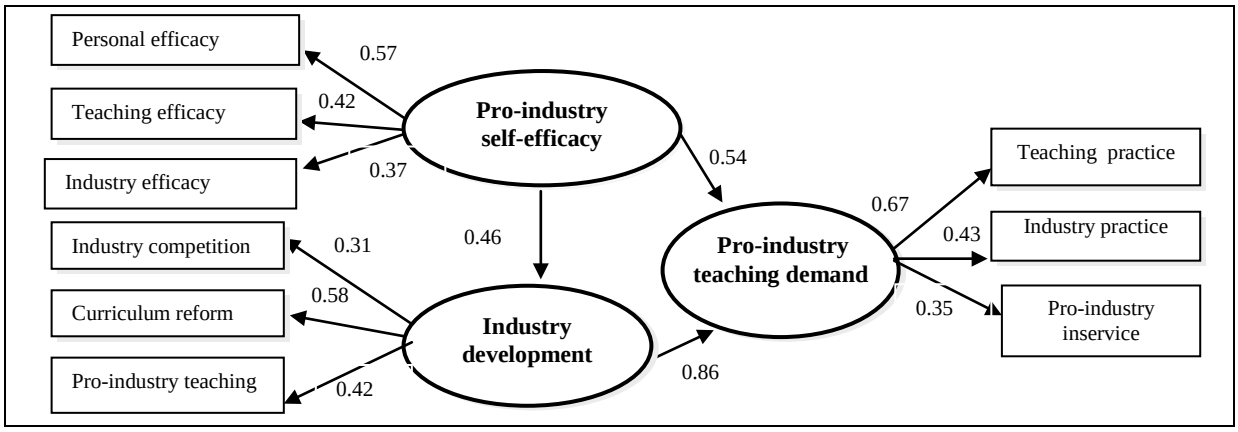


Figure 1. Path of influence factors of vocation teachers' pro-industry teaching demand.

Conclusion

Teachers' industry development has a significant direct effect on pro-industry teaching demand, and pro-industry self-efficacy has a significant effect on pro-industry teaching demand through industry development. The influence pattern and empirical data of pro-industry self-efficacy and industry development on pro-industry teaching demand has a good fit.

The influence effects of industry development and pro-industry teaching demand shows that for vocation teachers, the influence of pro-industry self-efficacy on pro-industry teaching demand comes mainly through their awareness of industry development. In addition, industry development has a direct and significant effect on pro-industry teaching demand. From the influence of industry development, pro-industry self-efficacy and pro-industry teaching demand, we can clearly see that compared with industry development has a greater influence on pro-industry teaching demand [11] [12] [13] [14].

Regarding the test results, according to the goodness of fit test standard by Hair et al, the model in this study has a good overall fit [16]. In the absolute fitness and incremental fitness tests, all indices meet the standard, and have the best fit. Most of the parsimonious fitness indices meet the test standard, and have a good fit. Overall, in the pro-industry teaching

demand and its influence model established in the study based on theories, both the model and the data have a good fit, and in the parameter estimation most of the estimated values are significant. This shows that all the indices of latent variables have their importance, and only the parameter value of pro-industry self-efficacy on pro-industry teaching demand is low. Overall, the empirical data have a good explanatory power [15] [16].

The results show that among all latent variables in the model, the direct influence of pro-industry self-efficacy on pro-industry teaching demand is not significant, indicating that the assumed influence of pro-industry self-efficacy on teachers' pro-industry teaching demand needs further testing; this is something worthy of a more in-depth study and validation in the future. Based on test results, although the overall result is acceptable, the model consistency level is not entirely satisfactory, and its industry development has a relatively low explanatory power for pro-industry teaching demand.

Acknowledgements

This paper was written while the authors were supported by a grant from the National Science Council, R.O.C. (MOST 105-2511-S-224-001-MY3)

References

[1] Bauer, W., Hämmerle, M., Schlund, S., & Vocke, C. (2015). Transforming to a hyper-connected society and economy –towards an "Industry 4.0". *Procedia Manufacturing, 3*, 417 – 424.

[2] Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588-606.

[3] Carter, A.G., Creedy, C.K., & Sidebotham, M. (2016). Efficacy of teaching methods used to develop critical thing in nursing and midwifery undergraduate students: A systematic review of the literature. *Nurse Education Today, 40*, 209-218.

[4] Chen, G. & Zhang, J. (2015). Study on training system and continuous improving mechanism for mechanical engineering. *The Open Mechanical Engineering Journal, 9*, 7-14.

[5] Chesnut, S.R. & Burley, H. (2015). Self-efficacy as a predictor of commitment to the tachign profession: A meta-analysis. *Education Research Review, 15*, 1-16.

[6] Faller, C., & Feldmüller, D. (2015). Industry 4.0 Learning Factory for regional SMEs. *Procedia CIRP, 32*, 88 – 91.

[7] Flores, I.M. (2015). Developing preservice teachers' self-efficacy through field-based science teaching practice with elementary students. *Research in Higher Education Journal, 27*,1-19.

[8] Frazzon, E.M., Hartmann, J., Makuschewitz, T., & Scholz-Reiter, b. (2013). To- wards socio-cyber-physical systems in production networks. Proce- dia CIRP, 7 49–54.

[9] Hascher, T. & Hagenauer, G. (2016). Openness to theory and its importance for pre-service teachers' self-efficacy, emotions, and classroom behqaviour in the teaching practicum. *International Journal of Educational Research, 77*, 15-25.

[10] Kilday, J.E., Lenser, M.L., & Miller, A.D. (2016). Considering students in techers' self-efficacy: Examination of a scale for student-oriented teaching. *Teaching and Teacher Education, 56*, 61-74.

[11] Lee, J., Bagheri, B., and Kao, H.A. "Research Letters A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems ". Manufacturing Letters 3 (2015): 18–23.

[12] Locke, T., & Johnston M. (2016). Developing an individual and collective self-efficacy scale for the teaching of writing in high schools. *Assessing Writing, 28*, 1-14.

[13] Schuh, G., Gartzen, T., Rodenhauser, T. & Marks, A. (2015). Pormoting work-based learning through Industry 4.0. *Procediia CIRP, 32*, 82-87.

[14] Sokolor, B. & Ivanov, D. (2015). Integrated scheduling gof material flows and information services industry 4.0 supply networks. *IFAC-Papers OnLine, 48*(3),1533-1538.

[15] Wang, S. Wan, J., Zhang, D., Li, D., & Zhang, C. (2016). Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. *Computer Networks, 101*(4), 158–168.

[16] Weiss, A., Huber, A., Minichberger, H., & Ikeda, M. (2016).First Application of Robot Teaching in an Existing Industry 4.0 Environment: Does It Really Work?*Societies* 2016, *6*(3), 20; doi:10.3390/soc6030020

Fast Analysis and Reanalysis for Structures with Nonlinear Supports

***Guanxin Huang, †Zhijun Yang and Xin Chen**

The Key Laboratory of Mechanical Equipment Manufacturing & Control Technology

Guangdong University of Technology, Guangzhou, 510006, China

†Corresponding author: yangzj@gdut.edu.cn

Abstract

A fast reanalysis method for structures with nonlinear supports is developed based on Indirect Factorization Updating (IFU) in this study. The famous Newton-Ralfson's method is employed to solve the nonlinear equation, therefore, the tangent stiffness matrix should be calculated and factorized repeatedly in the iterative process. The nonlinearity of the supports as well as structural modifications will lead to change of tangent stiffness matrix. In order to improve the efficiency of solving process, the IFU method is applied to deal with the change of tangent stiffness matrix. The numerical example shows that the proposed method is effective for structures with nonlinear supports.

Keywords: Reanalysis, Nonlinear supports, Indirect factorization updating

1 Introduction

Nonlinear reanalysis is one of the most challenging problem in reanalysis research area. Some achievements about nonlinear reanalysis have been gained in recent decades. Kirsch [1] developed a general reanalysis approach – Combined Approximation (CA), which can be used for nonlinear problems. Leu and Tsou [2] developed Kirsch's method for nonlinear dynamic analysis of framed structures. Akgün et al. [3] extended SMW (Sherman-Morrison-Woodbury) formulas to some nonlinear reanalysis problems. Deng et al. [4] developed a pseudoforce method for nonlinear analysis and reanalysis of structural systems. Hurtado [5] proposed a method based on Shanks transformation for both linear and nonlinear reanalysis problems. Materna et al. [6] proposed a nonlinear reanalysis method based on residual increment approximations.

Generally, nonlinear reanalysis methods are developed based on linear approaches. Because of the high requirement of accuracy, exact reanalysis methods are more suitable for nonlinear problems. Recently, Huang et al. [7] proposed an Indirect Factorization Updating (IFU), which is exact and suitable for structures with low-rank modifications. In this study, the IFU method is extended for reanalysis of structures with nonlinear supports.

2 Fast initial analysis for structures with nonlinear supports

A brief example for structures with nonlinear supports is shown in Fig. 1a. The equilibrium equation of the structure can be stated as

$$\mathbf{K}_0\mathbf{u} = \mathbf{F}(\mathbf{u}), \tag{1}$$

where, $\mathbf{K}_0$ is the stiffness matrix, $\mathbf{u}$ is the displacement vector, and $\mathbf{F}$ is the load vector, which depends on $\mathbf{u}$.

By using Newton-Ralfson's method, Eq. (1) can be solved by solving

$$\mathbf{K}_T^{(i)}\Delta\mathbf{u}^{(i)} = \Delta\mathbf{F}^{(i)}, \left(i = 0,1,\ldots\right), \tag{2}$$

where, $\mathbf{K}_T^{(i)}$ and $\Delta\mathbf{F}^{(i)}$ are tangent stiffness matrix and residual load vector of i-th iteration, respectively, which are calculated as

$$\mathbf{K}_T^{(i)} = \mathbf{K}_0 - \left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T, \tag{3}$$

and

$$\Delta \mathbf{F}^{(i)} = \mathbf{K}_0 \mathbf{u}^{(i-1)} - \mathbf{F}\left(\mathbf{u}^{(i-1)}\right). \tag{4}$$

Assume that the nonlinear supports are applied on several degree of freedoms (DOFs), which are numbered as d_i (i=1,2,…,s). In this case, only the d_i-th diagonal member of $\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T$ are non-zero. Therefore, $\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T$ can be expressed as

$$\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T = \begin{bmatrix} -\mathbf{p}_{d_1} & -\mathbf{p}_{d_2} & \cdots & -\mathbf{p}_{d_s} \end{bmatrix} \begin{bmatrix} \mathbf{e}_{d_1} & \mathbf{e}_{d_2} & \cdots & \mathbf{e}_{d_s} \end{bmatrix}^T = \mathbf{P}\mathbf{E}^T, \tag{5}$$

where, $\mathbf{p}_j$ indicates the *j*-th column of $\left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T$, and

$$\mathbf{e}_j = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}^T, \text{ (1 is the } j\text{-th member).} \tag{6}$$

Using SMW formula, Eq. (2) can be solved as

$$\Delta \mathbf{u}^{(i)} = \left(\mathbf{K}_0^{-1} - \mathbf{K}_0^{-1}\mathbf{P}\left(\mathbf{I} + \mathbf{E}^T\mathbf{K}_0^{-1}\mathbf{P}\right)^{-1}\mathbf{E}^T\mathbf{K}_0^{-1}\right)\Delta \mathbf{F}^{(i)}. \tag{7}$$

Therefore, only $\mathbf{K}_0^{-1}$ (or factorization of $\mathbf{K}_0$) need to be calculated before solving process, and Eq. (7) can be calculated very efficiently.

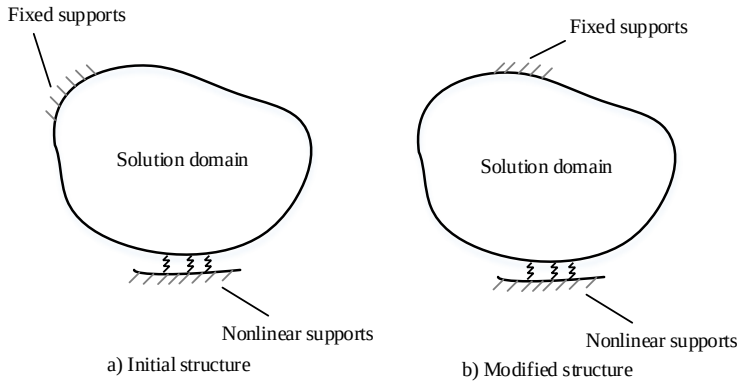


Fig. 1 A structure with nonlinear supports

3 Fast reanalysis for structures with nonlinear supports

Assume that a local modification (such as change of fixed supports as shown in Fig. 1b) is than applied on the structure, and the equilibrium equation becomes

$$\mathbf{Ku} = \mathbf{F}(\mathbf{u}), \tag{8}$$

and the tangent stiffness matrix in Eq. (2) becomes

$$\mathbf{K}_T^{(i)} = \mathbf{K} + \left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T. \tag{9}$$

Define

$$\mathbf{K} = \mathbf{K}_0 + \Delta\mathbf{K}, \tag{10}$$

$$\Delta\mathbf{K}_T = \Delta\mathbf{K} + \left(\frac{\partial \mathbf{F}}{\partial \mathbf{u}^{(i-1)}}\right)^T. \tag{11}$$

Equation (2) becomes

$$(\mathbf{K}_0 + \Delta\mathbf{K}_T)\Delta\mathbf{u}^{(i)} = \Delta\mathbf{F}^{(i)}. \tag{12}$$

In order to obtain an exact solution of Eq. (12) efficiently, the IFU method [7] is employed.

4 Numerical example

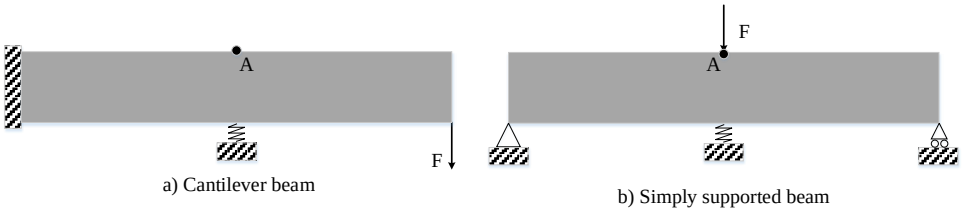
As shown in Fig. 2 is a tow-dimensional beam. Two different work condition is considered: cantilever beam as the initial structure and simply supported beam as the modified structure. A nonlinear support is applied on the middle-bottom of the beam as shown in Fig. 2. The law of the support is

$$f = -kx^3, \tag{13}$$

where, x is the deformation of the support, and

$$k = 1 \times 10^9 N / mm^3. \tag{14}$$

The material parameters are modulus of elasticity $E = 70 \times 10^3 MPa$, and Poisson's ratio $\upsilon = 0.3$. The forces in both Fig. 2a and 2b are linearly increased from 0 to 10N. The analysis results are shown in Fig. 3, and the comparisons of computational efficiency are shown in Table 1. Fig. 3 shows that the results of the SMW formula based initial analysis and the IFU based reanalysis are almost the same as the ones of the full analysis. From Table 1, it appears that the computational efficiency of the SMW formula based initial analysis and the IFU based reanalysis are higher than that of full analysis.

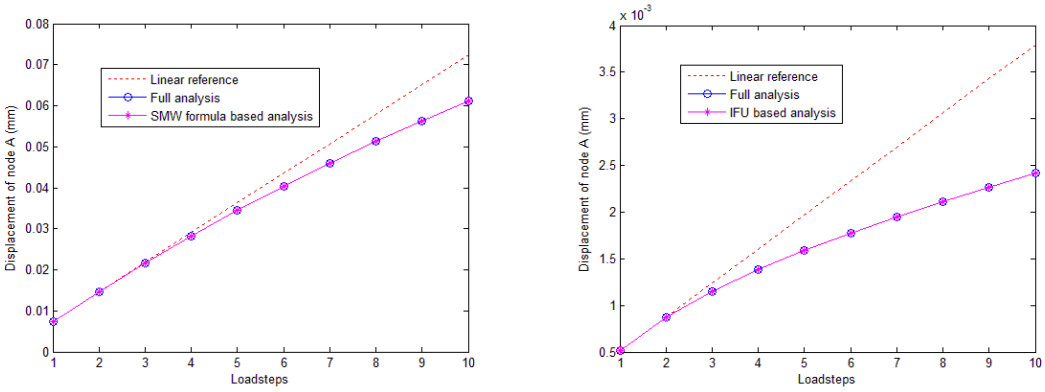


a) Cantilever beam b) Simply supported beam

Fig. 2 Models of the two-dimensional beam

Table 1 Comparison of computation efficiency

Models	Analysis methods	Computational cost (s)
Initial analysis	Full analysis	11.3176
	SMW formula based analysis	2.2826
Reanalysis	Full analysis	11.4454
	IFU based analysis	6.2782



a) Results of the cantilever beam b) Results of the simply supported beam

Fig. 3 Analysis results of the two-dimensional beam

5 Summary

This study developed a fast analysis and reanalysis method for structures with nonlinear supports. The SMW formula is applied in initial analysis, and the IFU method is adopted in reanalysis. The numerical example shows that the computational efficiency of the proposed fast analysis and reanalysis is high than that of full analysis, and exact solutions can still be guaranteed.

Acknowledgement

The research described in this paper was financially supported by Guangdong Natural Science Foundation (Grant No. 2015A030312008, 2016A030308016), Guangdong Science and Technology Plan (Grant No. 2015B010104006, 2015B010104008, 2013B010402014, 2015B010133005, and 2015A010104009), Science and Technology Program of Guangzhou (Grant No. 201510010281), National Key Basic Research Program of China (Grant No. 2011CB013104), National Natural Science Foundation of China (Grant No. U1134004, U1401240).

References

[1] U. Kirsh, Combined approximations – a general reanalysis approach for structural optimization, Structural and Multidisciplinary Optimization 20(2000) 97-106.
[2] L.J. Leu, C.H. Tsou, Application of a reduction method for reanalysis to nonlinear dynamic analysis of framed structures, Computational Mechanics 26(2000) 497-505.
[3] M.A. Akgün, J.H. Garcelon, R.T. Haftka, Fast exact linear and non-linear structural reanalysis and the Sherman-Morrison-Woodbury formulas, International Journal for Numerical Methods in Engineering 50(2001) 1587-1606.
[4] L. Deng, Michel Ghosn, Pseudoforce method for nonlinear analysis and reanalysis of structure systems, Journal of Structural Engineering 127(2001) 570-578.
[5] J.E. Hurtado, Reanalysis of linear and nonlinear structures using iterated Shanks transformation, Computer Methods in Applied Mechanics and Engineering, 191(2002) 4125-4229.
[6] D. Materna, V.K. Kalpakides, Nonlinear reanalysis for structural modifications based on residual increment approxiamtions, Computational Mechanics, 57(2016) 1-18.
[7] G. Huang, H Wang, G. Li, An exact reanalysis method for structures with local modifications, Structural and Multidisciplinary Optimization, 54(2016) 1-11.

Numerical prediction and analysis of motion response of high speed planning craft in regular waves

Hongjie Ling, Zhidong Wang
The school of naval architecture and ocean engineering, Jiangsu University of Science and Technology
Zhenjiang Jiangsu 212003

Abstract

To accurately estimate the hydrodynamic performance of planning craft in waves, the high-speed sailing of planning craft in regular head waves was numerically simulated with multiple degrees of freedom by performing sea keeping model experiments. In the simulation, the FINE/MARINE software was adopted and the body-fitted mesh technique was employed.

The numerical results were then compared with the experimental data to verify the accuracy of the simulation. It is found that the numerical simulation can accurately and efficiently simulate the motion attitude and the hydrodynamic characteristics of the planning craft in high speed navigation in waves. Based on the time domain and frequency domain analysis method of the motion responses of the planning craft, the analysis of the influence of the period of incident wave and the natural frequency on the response of planning craft has been finished, which provide guidance and reference for the design of planning craft.

Key words： planning craft; motion response; regular wave; body-fitted mesh; numerical simulation

0 Introduction

As an important part of the field of high performance crafts, the planning craft has attracted more and more attention and application because of the superior combat performance. Owing to the strong nonlinear characteristics, such as overtopping and slamming of planning crafts in high speed sailing in waves, the accurate prediction of planing craft hydrodynamic and motion performance has become the focus of attention of scholars at home and abroad. With the development of computer hardware, the solution of ship hydrodynamics and motion response based on RANS equation has become the research direction of many scholars recently.

Hydrodynamic calculation research of planning crafts surface began with the towing test in Langley pool in 1940s, carried out by the National Advisory Committee for Aeronautics (NACA) [1].A lot of subsequent studies have been carried out based on the experimental data. Among these studies, Savitsky[2]-[4] presents a series of empirical or semi-empirical formulas for calculating the resistance of planning crafts based on the test results. In recent years，The modern CFD technique, which is aimed to solve the Reynolds time-average equation RANS in real time, is applied to the accurate prediction of hydrodynamic performance of surface high speed craft. According to the application of the 2D+t theory and the fully nonlinear boundary element method，Hui Sun and Odd M. F[5] calculated the added mass, damping coefficient and restoring force coefficient of the planning craft and presented

the nonlinear time domain simulation of hydrodynamic and motion characteristics of planning craft, which shows that the hull hydrodynamic coefficients, pitch amplitude and the height of the center of gravity have a significant impact on the motion response of planning crafts. Based on experimental test, N.Santoro[6],etc have studied the hydrodynamic force and torque acting on the planning craft hull at high speed, focusing on the distribution of pressure at the bottom of the planning craft.

Su Yumin and Duan Wenyang[7]-[13] of Harbin Engineering University made a systematic study on the motion response of planning crafts in still water and regular waves by using self-programming and commercial software FLUENT, and then analyzed the hydrodynamic performance, motion performance and wake flow characteristics of planning crafts. Dong Wencai et al. [14] from Naval University of Engineering, studied the longitudinal motion in head sea regular waves of the deep V type planning craft, and analyzed the influence of wave factors on the longitudinal motion of planning craft.

This paper, based on FINE/MARINE software, carries out the numerical prediction of the longitudinal three degree of freedom motion response of the planning craft in the regular waves by utilizing six degree of freedom motion response module. In order to ensure the mesh quality of the planning craft in waves with large amplitude motion, the hydrodynamic characteristics, the motion response characteristics and the flow field distribution of the planning craft with different wave periods are quantitatively analyzed based on the body-fitted mesh technology, and the results are compared with the experimental results for verification.

1 Computational Model

1.1 Computational model and grid partition

According to the model of a certain type of planning craft, the SOLIDWORKS software is used to complete the 3-D modeling, and the HEXPRESS software is used for mesh generation. Then mesh encryption in the vicinity of hull and the free surface was conducted, among which the upper part is air domain while the lower part is water domain. Water domain is 10L * 4L * 3L, and the air domain is 10L * 4L * 1.5L. The craft bow is 3L away from the entrance and the stern is 6L away from the exit. In the computational domain, the standard wall function is used for the hull, the prescribed pressure (hydrostatic pressure) is used in upper boundary and the lower boundary, while the far field is adopted in the inlet boundary, the exit boundary and the bilateral boundary. Figure 1 is a schematic diagram of the computational domain and the hull surface grid division, and the main parameters of the planning craft model are shown in Table 1.

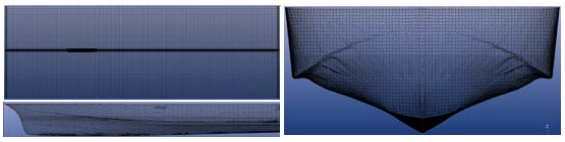


Figure 1. Diagrammatic sketch of the computational domain and the hull surface grid division

Table 1. Main parameters of planning craft model

Principal dimension	Parameter	Principal dimension	Parameter
Total length *L*/m	2.75	Length from center of gravity to stern l_g/m	1.048
Total width *B*/m	0.78	Craft weight *m*/kg	125.4
Draft *d*/m	0.17	Dead rise angle *β*/(°)	24.65
Longitudinal moment of inertia I_y /(kg•m^2)	53	Initial angle attack α/(°)	3

1.2 Calculation condition and numerical calculation method

Based on the FINE/Marine software, the parameter setting of the longitudinal three degree of freedom motion response prediction of the planning craft in the regular waves are as follows：3-D unsteady two-phase flow, k-ω(SST-Menter) (SST-Menter) turbulence model(the value of k and ω is related to Reynolds number); while the speed of the planning craft is given, invoking the motion of the six degree of freedom motion, release the motion of heave and pitch freedom; the velocity inlet is used to make waves, and the body-fitted mesh technique was adopted in forward, heave and pitch motion; the turbulence equation is discretized by central difference (AVLSMART) scheme; the momentum equation is discretized by central difference (AVLSMART) scheme; the free surface is discretized by the mixed free surface capture and reconstruction (BRICS) scheme; the pressure velocity coupling algorithm is used. Calculation conditions are shown in Table 2.

Table 2. Calculation condition table

Number	H(m)	T(s)	V(m/s)	m（kg）	Fr_{∇}	λ(m)	T_e(m)	λ/L
A1		1.63				4.13	0.65	1.50
A2		1.88				5.50	0.81	2.00
A3		2.10				6.88	0.96	2.50
A4	0.20	2.30	3.86	125.4	1.74	8.25	1.11	3.00
A5		2.48				9.63	1.24	3.50
A6		2.66				11.00	1.37	4.00
A7		2.97				13.75	1.62	5.00

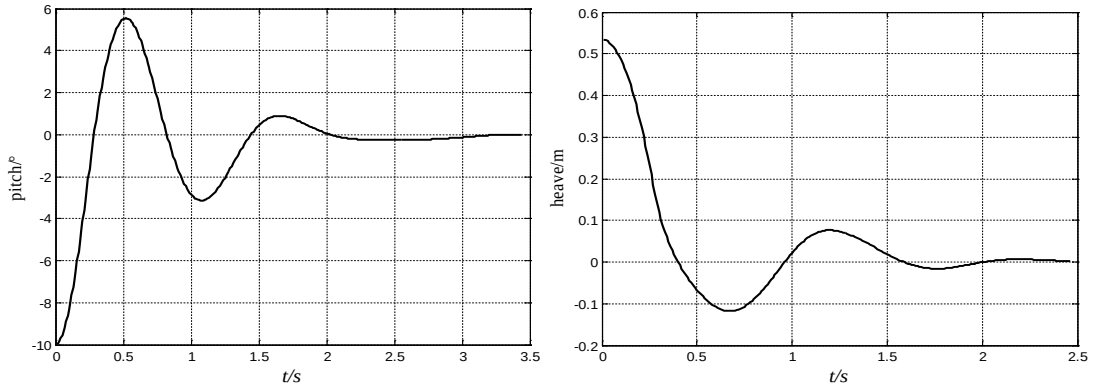
2 Calculation results and analysis

From the point of view of hydrodynamics, if the Volume Froude number $Fr\nabla \geq 1.0$, the craft belongs to high-speed ship, including high speed displacement ship and power lift ship. For high-speed displacement ships whose $Fr\nabla$ is between 1.0 and 3.0, the main supporting force is the static buoyancy； For hydrodynamic lift type ships whose $Fr\nabla$ is no less than 3.0, the main supporting force is the hydrodynamic lift. As a result of high speed fluid acting on the

hull surface, the splash phenomenon shows up at the interface of water and gas, and the influence of splash on the high speed ship in navigation should not be ignored. This paper studies the planning craft at the speed of 3.86m/s, Fr_{∇}=1.74.

2.1 Determination of the natural period of rolling, pitching and heaving

In the high speed navigation of planning craft, the composition of the motion response is complex. The natural period of three degrees of freedom, for which the restoring force exists, of the planning craft has an important influence on the motion response. As a result, the hydrostatic damping of the pitching and heaving of the planning craft is carried out to calculate the natural period of the planning craft. As can be seen from Figure 1, the natural period of pitching and heaving are respectively 1.08s and 1.0s.



（a）Pitch decay curve （b）Heave decay curve

Figure 1. Hydrostatic decay curve of planning boat

2.2 Calibration of incident wave elements

As a planning craft sails in the target wave environment, set a wave height observation point at the position of 2 times the length of craft in the forward direction of the planning craft, and real-time output wave height. Figure 2 shows the time history curves of different period (part time), where we can see that the error between incident wave height and target wave height is less than 5%, and the period is the same with the encounter period, indicating that the wave elements satisfy the calculation requirement.

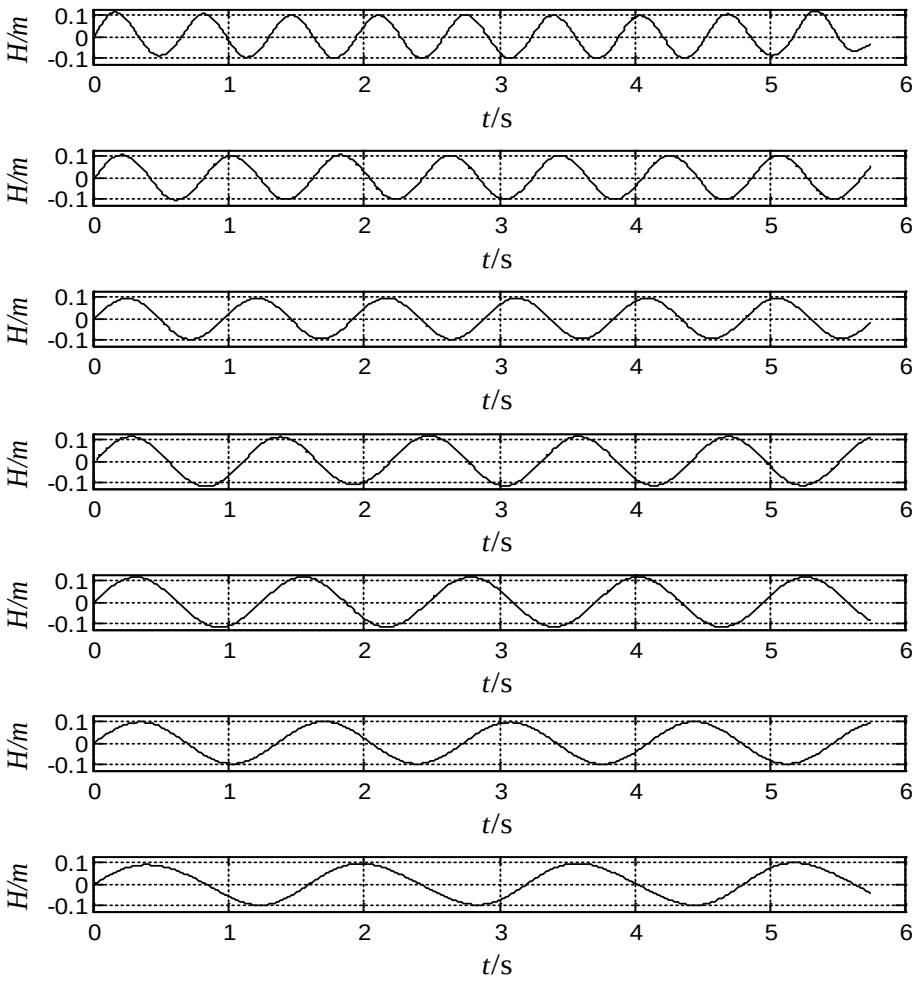


Figure 2. Wave duration curve of different period

2.3 Hydrodynamic performance of planning boat in waves

Figure 11 and Figure 12 show the variation curve of the resistance and dynamic lift of planning craft (for a period of time). It can be seen that the resistance and the dynamic lift force change with the periodic change of the wave when the planning craft is sailing in the waves. Figure 3 shows the load at the position where resistance and dynamic lift equilibrium (load is the average value of the duration curve). Figure 4 and figure 5 show the variation of the amplitude of the component of resistance and dynamic lift with wavelength.

Figure 3, figure 4 and figure 5 can be combined to reflect the change rule of the resistance and dynamic lift of the planning craft moves in waves with wavelength.

It can be seen from the figure that wavelength has little influence on the equilibrium position of the resistance (equivalent to the hydrostatic resistance), and the mainly affect its amplitude. Due to the high frequency characteristic of the two order slow drift force of waves, the amplitude of resistance decreases with the increase of wavelength.

When λ/L=1.5/2.0, resistance have remarkable high frequency characteristics, even shows quadruple-frequency phenomenon (refer to resistance spectrum in Figure 6 and Figure 8). The contribution of high frequency components to the resistance is large, especially when λ/ L = 1.5, the high frequency resistance accounts for about 45% of the total resistance. The main reason for this phenomenon is: the wavelength is equal to the craft length so that overtopping occurred when the planning craft is sailing at high speed (verified in flow field Figure 16 (a)), and the hull slamming is remarkable. The strong nonlinear characteristics appear in the flow field around the hull.

When λ/L>2.0, the resistance is mainly composed of wave frequency components, as the contribution of high frequency components to the total resistance is less than 5%. The resistance value of the planning craft changes periodically with the peak value and the valley value of the encounter wave. For the same wave height, the longer the wavelength is, the smaller the wave steepness is, and the smaller the resistance change caused by waves is, Therefore, the resistance change amplitude decreases with the increase of wavelength.

As can be seen from the figure, the effect of the wavelength on the equilibrium position of the dynamic lift (which is equivalent to the dynamic lift in calm water) is not significant. As Fr_{∇} =1.74, the static buoyancy played a major role in the balance of the hull weight while dynamic lift accounted for about 10%;

When λ/L=1.5/2.0/3.0, dynamic lift have significant high frequency characteristics, even shows quadruple-frequency phenomenon (refer to dynamic lift spectrum in Figure 7, Figure 9 and Figure 10). The contribution of high frequency components to the dynamic lift is large, especially when λ/ L = 1.5, the high frequency resistance accounts for about 55% of the total resistance. The main reason for this phenomenon is: the wavelength is equal to the craft length so that overtopping occurred when the planning craft is sailing at high speed (verified in flow field Figure 15 (a)), and the hull slamming is remarkable. The strong nonlinear characteristics appear in the flow field around the hull.

Different from the resistance characteristics, the dynamic lift appears triple frequency at λ/L=2.5, and there is a peak at the equilibrium position. As the encounter frequency of incident wave is close to the natural period of the pitching and heaving of the planning craft, the resonance phenomenon occurs.

When λ/L>2.5, the dynamic lift is mainly composed of wave frequency components while the contribution of high frequency components to dynamic lift is about 20%. The dynamic lift value of the planning craft changes periodically with the wave. For the same wave height, the longer the wavelength is, the higher the wave lift capacity increases with the increase of wavelength, So the dynamic lift amplitude increases with the increase of wavelength.

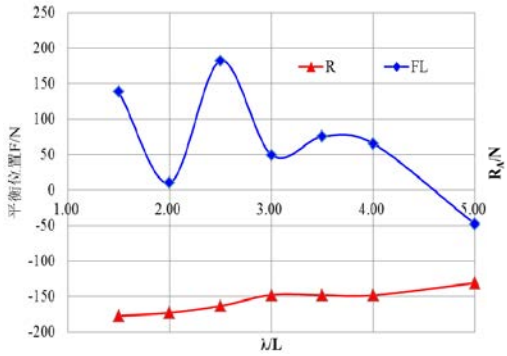


Figure 3. Load change curves with wavelength at resistance and dynamic lift equilibrium position

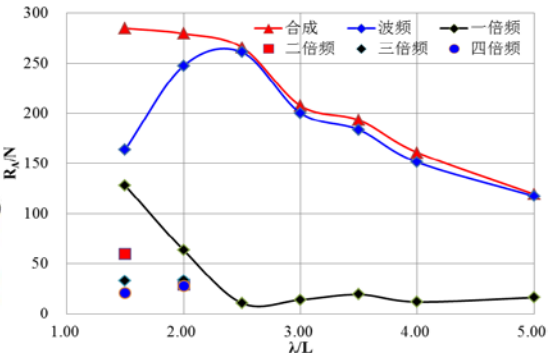


Figure 4. Resistance amplitude versus wavelength curve

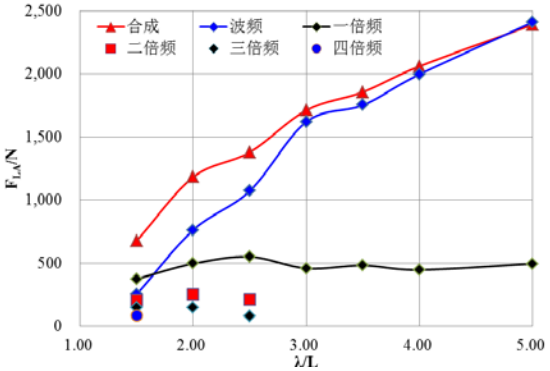


Figure 5. Dynamic lift amplitude versus wavelength curve

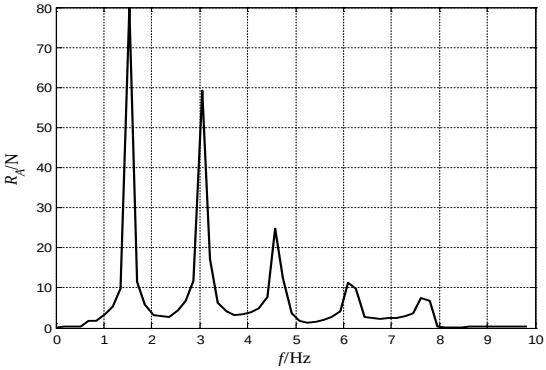


Figure 6. Resistance spectrum whenλ/L=1.5

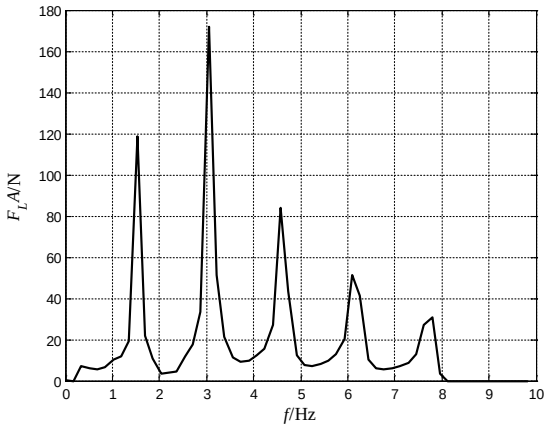


Figure 7. Dynamic lift spectrum whenλ/L=1.5

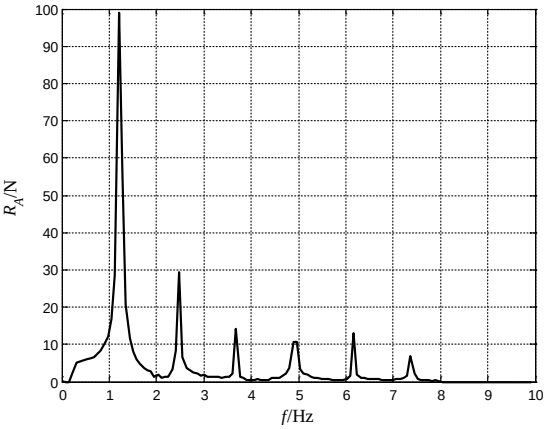


Figure 8. Resistance spectrum whenλ/L=2.0

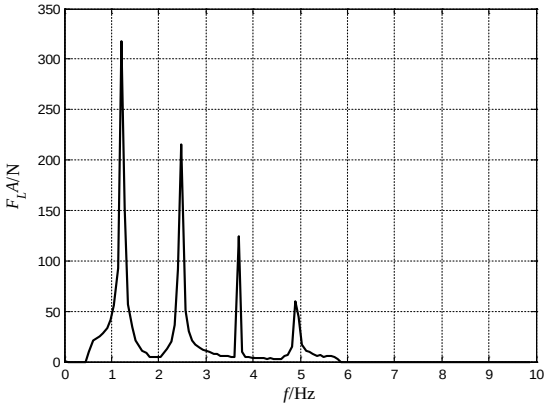


Figure 9. Dynamic lift spectrum whenλ/L=2.0

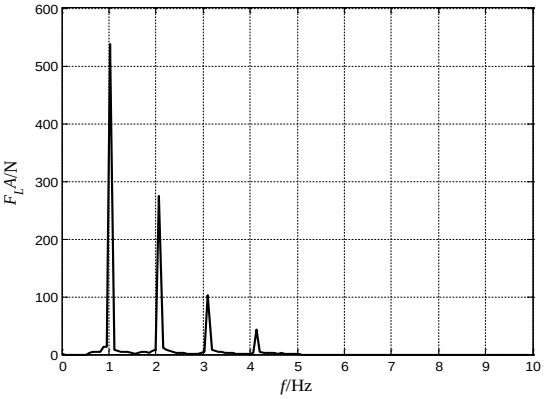
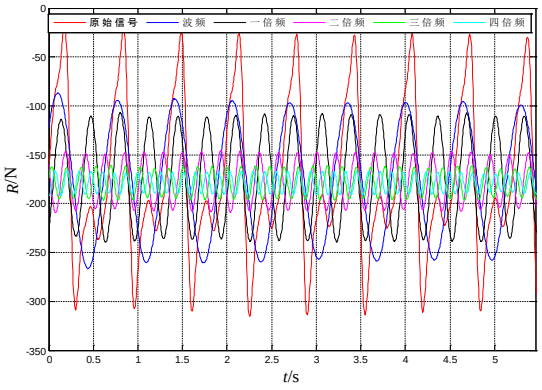
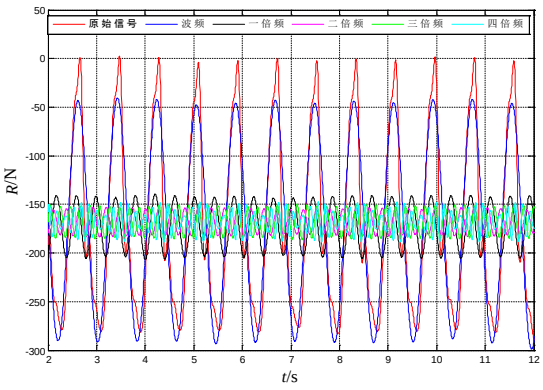


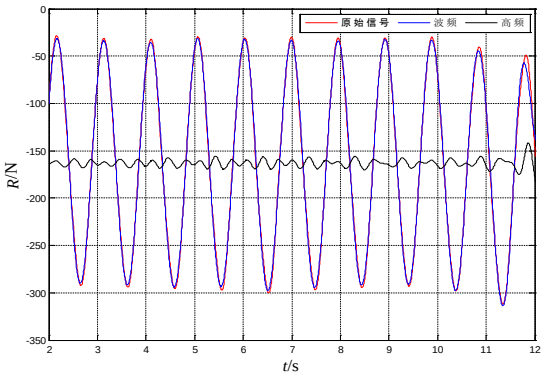
Figure 10. Dynamic lift spectrum when λ/L=2.5



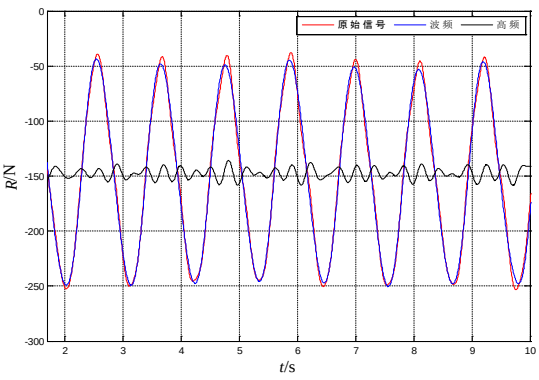
（a）λ/L=1.5



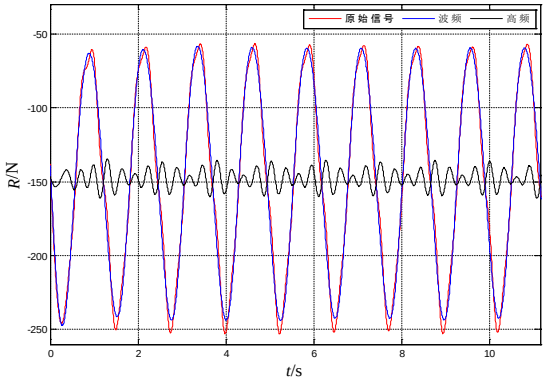
（b）λ/L=2.0



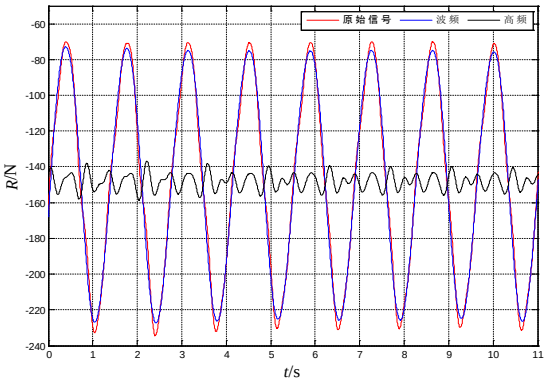
（c）λ/L=2.5



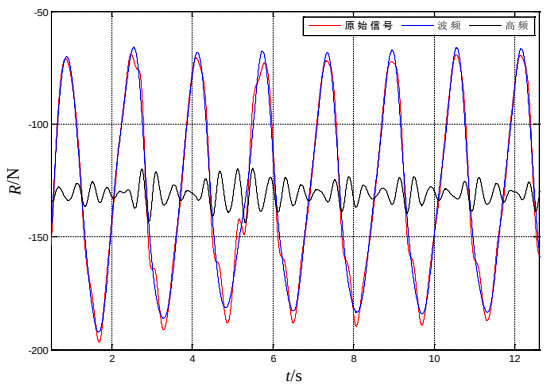
（d）λ/L=3.0



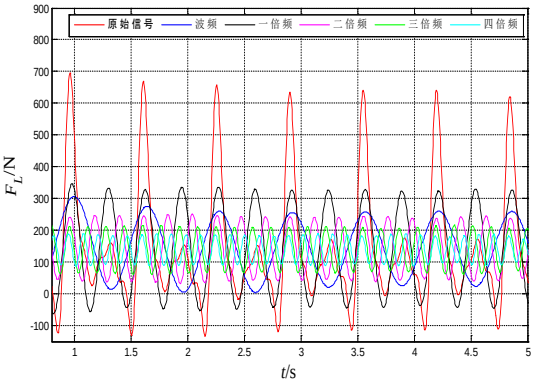
（e）λ/L=3.5



（f）λ/L=4.0

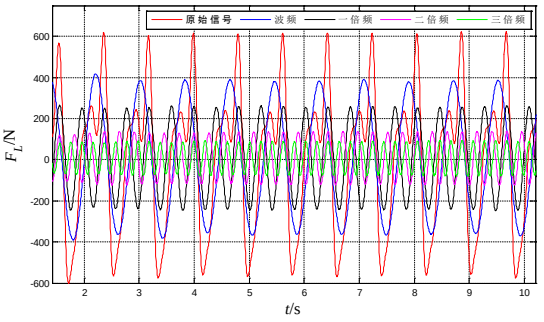


（g）λ/L=5.0

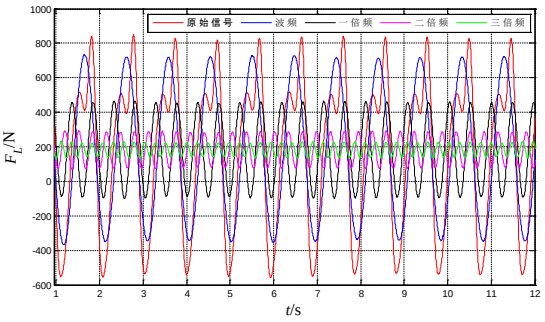


（a）λ/L=1.5

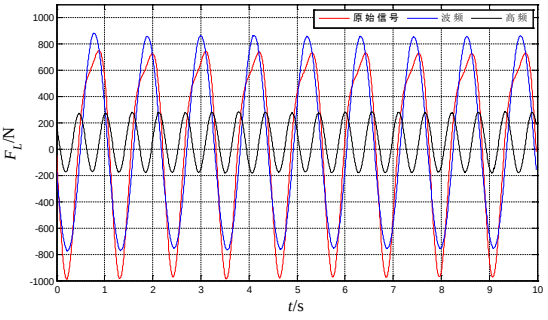
Figure 11. Duration curves of resistance components at different wavelengths



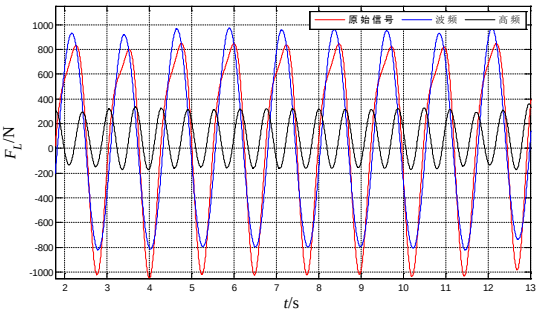
（b）λ/L=2.0



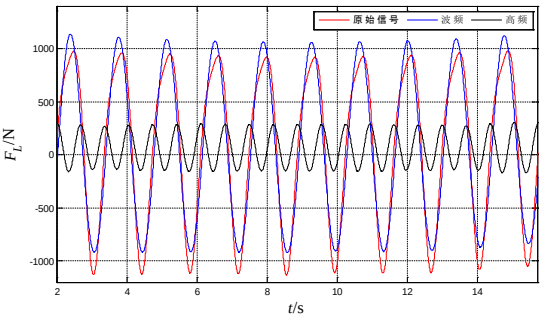
（c）λ/L=2.5



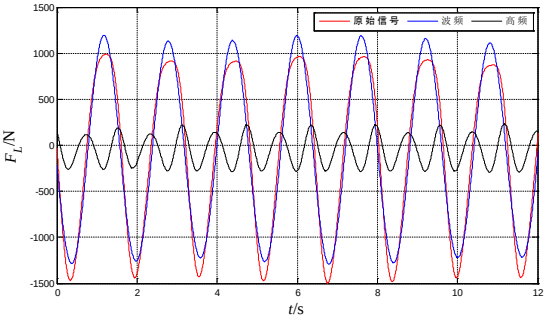
（d）λ/L=3.0



（e）λ/L=3.5



（f）λ/L=4.0



（g）λ/L=5.0

Figure 12. Duration curves of dynamic lift components at different wavelengths

2.4 Analysis of the motion response characteristics of planning boat in waves

Figure 15 shows the time history curve of heave and trim angle (a certain period of time). It can be seen from the figure that if λ/L=1.5/2.0/3.0, there is obvious high frequency components in the motion of heave and trim angle, and the high frequency component accounts for about 10%, which is consistent with the change characteristic of dynamic lift; when λ/L>2.5, the heave and the trim angle is mainly composed of wave frequency, with high frequency components accounting for less than 3%.

Figure 13 and figure 14 show the variation of the amplitude of the heave and the trim angle with the wavelength. And the numerical results are compared with the experimental data, which indicates they are in good agreement. The error is generally less than 5%. It can be seen from the figure that the amount of heave and trim angle increases with the wavelength and then decreases, and the response reaches the maximum at λ/L=3.0. The resonance phenomenon occurs when the frequency is close to the wave encounter period and the natural period of heave and pitch.

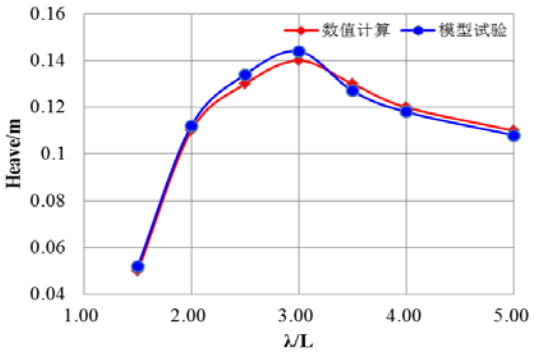


Figure 13. Comparison between the calculation results and experiment data of the heave

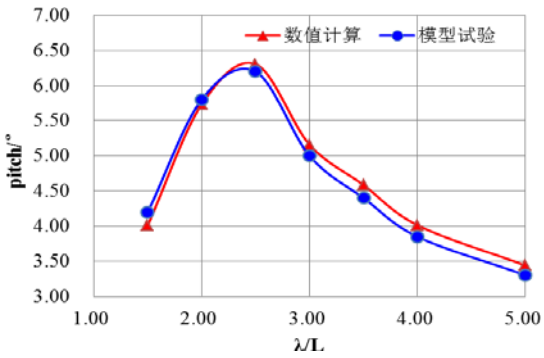
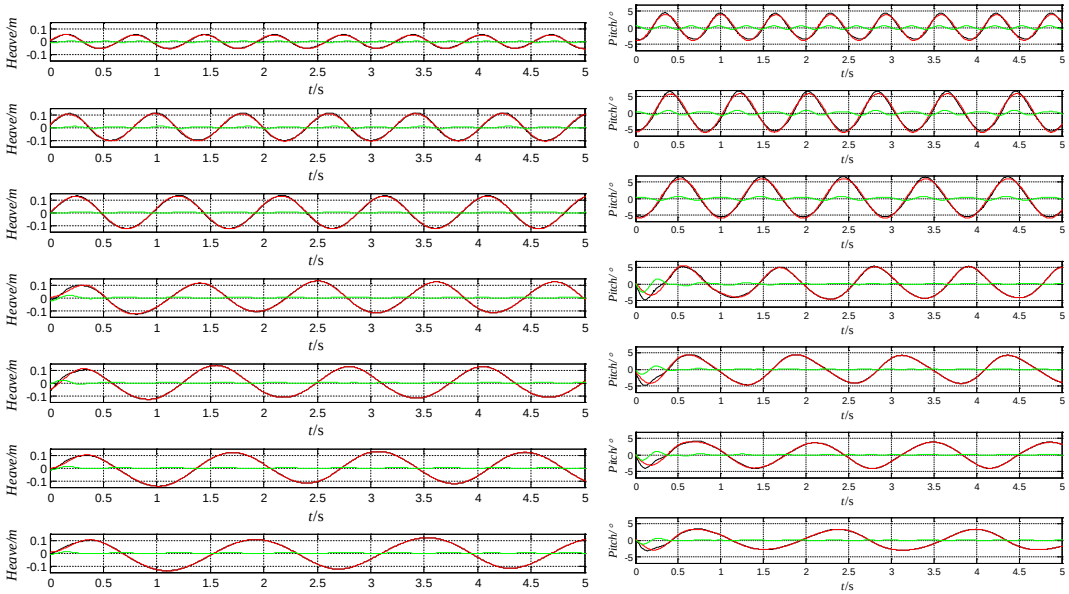


Figure 14. Comparison between calculation results and experiment data of the trim angle amplitude



（a）heave **（b）trim angle**

Figure 15. Duration curves of heave and trim angle at different wavelengths

2.5 Flow field distribution of planning boats in waves

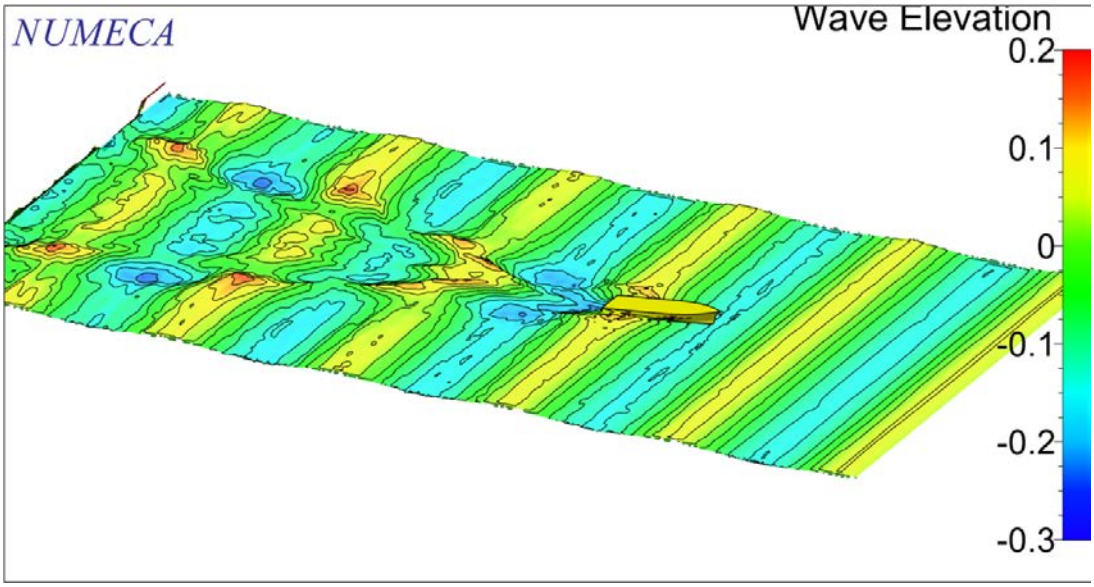
Figure 15 shows the wake field wave clouds map of planning crafts with different wavelength. From the figure we can obviously see that, when the planning craft is sailing in waves at high speed, the craft hull has a negative effect on the flow. The wake field has the following characteristics：

（1）The water flowing out from both sides of the hull, resulting in a large number of splash and spray.

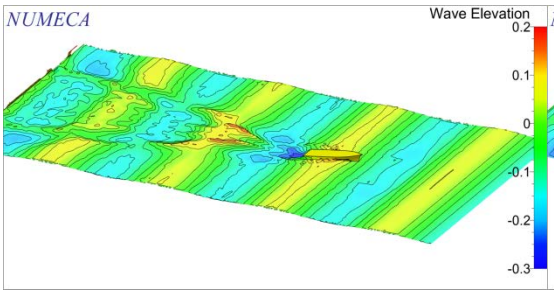
（2）A cavity forms at the trailing edge of the hull.

（3）Scattered wave and shear wave from a "chicken wake flow" at the hull tail edge, causing the superposition of incident wave and traveling wave system. As a result, the 3-D wave crest of the longitudinal profile of the planning craft is appeared in the wake field.

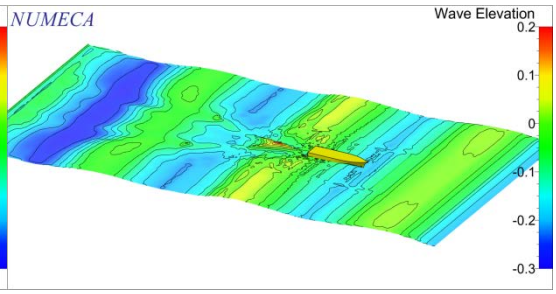
（4）When λ/L=1.5/2.0, overtopping phenomenon occurs.



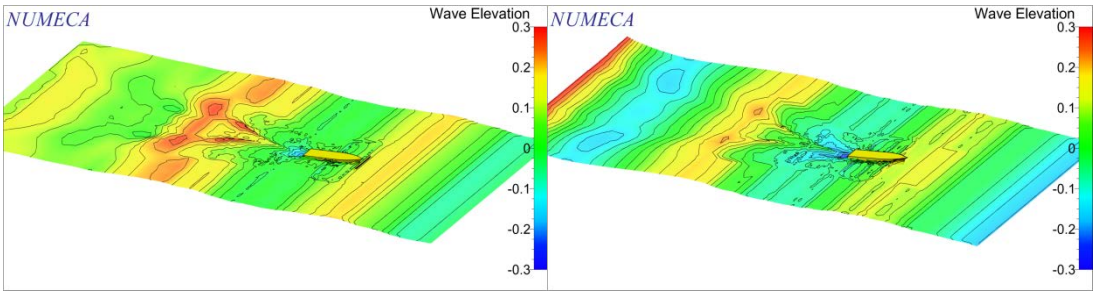
（a）λ/L=1.5



（b）λ/L=2.0

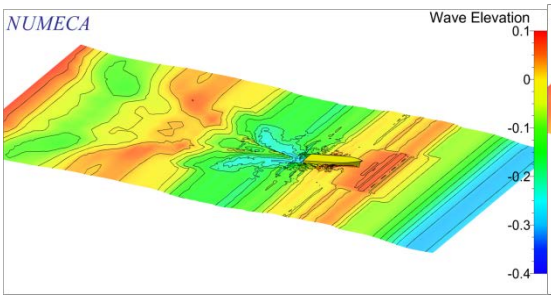


（c）λ/L=2.5

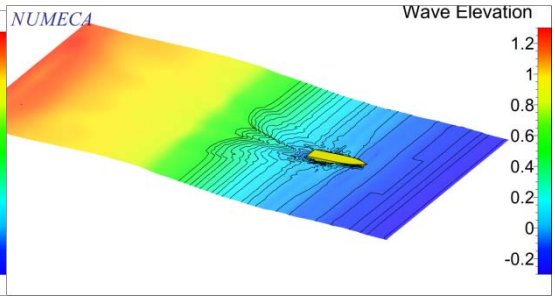


（d）λ/L=3.0

（e）λ/L=3.5



（f）λ/L=4.0



（g）λ/L=5.0

Figure 16. The waveform of the stern flow field

3 Conclusion

In this paper, based on the FINE/MARINE software, the numerical simulation of the longitudinal motion response of high speed planning craft, with the application of six degree of freedom motion response module, has been conducted. The numerical results are compared with the model tests and the following conclusions are obtained：

1）The results of numerical simulation of the motion response of high speed planning craft are in good agreement with the experimental values, with an error less than 5%.

2）When λ/L=1.5/2.0, overtopping and slamming happens on the planning craft. The composition of the load acting on the planning craft is complex, which has a great influence on the motion response.

3）When λ/L=1.5/2.0, The high frequency component has a great contribution to the resistance and the dynamic lift of the planning boat, which should be given more attention. When λ/L>2.5, it mainly consists of wave frequency component. The high frequency component is so little that can be ignored.

4）The Fine/Marine software can be used to deal well with the nonlinear problems such as overtopping and slamming.

Reference

[1] ChaInbliss D B，Body G M．The planing chamcteristics of two V-shaped prismatic surfaces having angles of deadrise of 20 and 40[R]. Washington D C: Langley Aeronautical Laboratory, NACA,1953:28-76.
[2] SAVITSKY D. Hydrodynamic design of planing hulls[J].Marine Technology ,1964,1(1):79-96.
[3] SAVITSKY D. Overview of planing hull forms[C]//InProc HPMV'92. Alexandria, Va; American Society of Naval Engineers，1992:1-14.

[4] SAVITSKY D. and BROWN P W.Procedures for hydrodynamic evaluation of planing hulls in smooth and rough water[J].Marine Technology,13(4),1976,381-400.
[5] Hui Sun and Odd M. Faltinsen.The influence of gravity on the performance of planning vessels in Calm Water[J]. J EngMath, 2007,58(1):91-107.
[6] SANTORO N, BEGOVIC E,etc. Experimental study of the hydrodynamic loads on high speed planingcraft[J].ProcediaEngineering, 2014:186-193.
[7] SU Yumin, Qingtong,etc. Numerical simulation of a planing vessel at high speed[J].Journal of Marine Sciences and Application.(2012) 11: 178-183.
[8] WANG Shuo, SU Yumin,etc. RANSE simulation of high-speed planing craft in regular waves[J]. Journal of Marine Sciences and Application, 2012,vol.11(4):447-452.
[9] Xin Zhu,Wenyang Duan,etc. The frequency domain solution for the motion simulation of the prismatic planing craft in regular head waves [J]. Journal of Harbin Engineering University.2012,11(33):1326-1333.
[10] WANG Shuo, SU Yumin,etc. Numerical study on longitudinal motions of a high-speed planing craft in regular waves [J]. Journal of Harbin Engineering University.2014,1(35):45-52.
[11] WANG Shuo, SU Yumin,etc. Numerical simulation of static and direct motion of planing craft [J]. Journal of South China University of Technology(Natural Science Edition),2014,4,(41):119-126.
[12] JIANG Yi,SUN Hanbing,etc. Numerical simulation on hydrodynamic performances of double-stepped planning boat. [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition),2015,4(43):74-78.
[13] Xin Zhu,Wenyang Duan,etc. A study on the extreme value of slamming pressure in planing boats [J]. Journal of Wuhan University of Technology(Transportation Science&Engineering),2013,12(37):1188-1192.
[14] Dong Wencai,Yue Guoqiang,Experimental study on longitudinal motion of deep V planing craft [J]. SHIP ENGINEERING,2004, 26 (2):14-16.

Numerical Simulation of Mars Exploration Rover Heat Shield Separation

†*G.W. Xu[1], Z. Liu[1], and Y.J. Yang[1]

[1] China Academy of Aerospace Aerodynamics, China.

*Presenting author: elexgw@163.com
†Corresponding author: elexgw@163.com

Abstract

A DDES method, which is a hybrid RANS-LES method, is proposed for numerical simulation of Mars Exploration Rover heat shield separation. The heat shield and back shell drag characteristic and aerodynamic interference effect between them are analyzed at different axial separation distances. The computational results are compared to the experimental results. And then security analysis of heat shield separation is discussed. The results indicate that the aerodynamic interference effect of heat shield and back shell is the most serious inducing a suction force while the axial distance is less than two heat shield diameters. And there will be a recontact risk after heat shield separating while the axial distance is less than five heat shield diameters. To ensure successful heat shield separation, the axial distance must be larger than five heat shield diameters.

Keywords: Numerical Simulation, Heat Shield Separation, Afterbody, Mars Exploration Rover, Hybrid RANS-LES Method, Drag Characteristic

1.Introduction

In order to safely touch down on Mars' surface, Entry Descent and Landing (EDL) methods are employed in the lander missions to Mars[1]-[2]. The implementation process include cruise stage separation, hypersonic entry, parachute deploy, heat shield separation, radar data collection starts, back shell separation, powered descent and sky crane flyaway etc. Thereinto, heat shield separation is the key stage which may incur risk easily.

After heat shield separation, what remains is the back shell with the lander attached to it and this combined body is referred to as afterbody. When air flow past the heat shield that is a blunt body, a wake zone will develop in the backside. And there will be a zone near blunt body base where the air flow direction is inverse to the incoming flow direction, which is called return-flow zone. While heat shield is jettisoned by separation spring, the whole afterbody will be located at the wake of heat shield and there will be an aerodynamically induced suction force pushes the heat shield back into afterbody. Therefore, to ensure successful heat shield separation, two step problems must be solved. First, the separation spring has to impart sufficient impulse to overcome this suction. Second, there has to be a sufficient ballistic coefficient difference between the heat shield and afterbody such that the heat shield descents faster. However, there are considerable uncertainty in Mars' atmospheric conditions, namely atmospheric density and winds. It is very important to do the research of heat shield separation.

There are much research work about Mars exploration in America[3]-[6], including the investigation of heat shield separation. Lang carried out wing tunnel test of Viking heat shield separation previously[7], and the test results were systematic analyzed by Behzad[8]-[9]. In China, there are also some research about Mars exploration, which are mostly review

article[10]-[11] or analysis of EDL technology[12]-[14], and none of research paper about heat shield separation has been seen.

The focus of this paper is numerical simulation of heat shield separation. Aerodynamic properties of heat shield and afterbody are discussed separately. And only the drag force characteristic is researched.

2.Heat Shield Separation Model

Mars Exploration Rover[15] is chose as calculation model , shown in Figure 1. Drag force of heat shield and afterbody at different separation distances are compared and analyzed. The separation distances include axial separation distance (Δx) and lateral separation distance (Δz), shown in Figure 2. For preliminary study, only the axial separation distance is investigated.

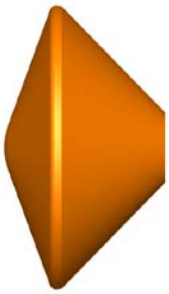


Figure 1. Model of Mars Exploration Rover

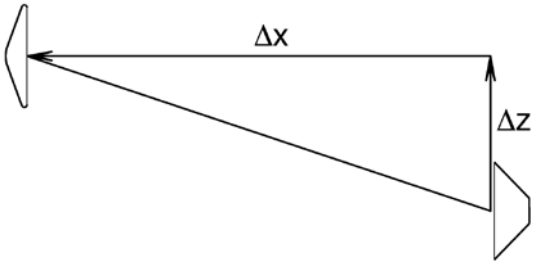


Figure 2. Separation distances between heat shield and afterbody

3.Calculation Method

RANS-LES hybrid method is adopted for numerical calculation[16], which integrates Reynolds Average Navier-Stokes (RANS) and Large Eddy Simulation (LES). The basic idea of RANS-LES hybrid method is simulating high-frequency small-scale motion in near wall area by using RANS method and low-frequency large-scale motion in separated flow area by using LES method. The classical Detached Eddy Simulation (DES)[17]-[18] is improved and Delayed Detached Eddy Simulation (DDES)[19] is gained. The method is constructed from RANS turbulence model equation. The detail of DDES method and control equation can be seen from reference [16] and reference [20].

In order to validate the reliability of above calculation method, Shenzhou capsule configuration is used for case check, wind tunnel test of which has been carried out in FD-12 wind tunnel at China Academy of Aerospace Aerodynamics. Test condition: Mach number equal to 0.9, Reynolds number per unit length equal to 1.8×10^7. Comparison of computational result and experimental result are presented in Table 1. The error is within 2%, indicating the veracity of calculation method.

Table 1. Comparison of computational result and experimental result

Results	*Axial force coefficient*
Experimental result	1.20
Computational result	1.22

4.Numerical Simulation Results

The heat shield separation height is 5 kilometers of Mars Orbital Laser Altimeter (MOLA)[21]. Table 2 presents Mars atmosphere parameters.

Table 2. Mars atmosphere parameters

Ma	*Height[km]*	$\rho_\infty[kg/m^3]$	$T_\infty[K]$
0.3	5	0.00762	222

Heat Shield Drag Force characteristics

Figure 3 shows variation of heat shield drag force coefficient with axial separation distance Δx (normalized by heat shield diameters D). As axial separation distance increasing, the drag force coefficient of heat shield augments rapidly and reaches maximal value at Δx=1D. Then it minishes rapidly for axial distance 1D<Δx<3D and reaches minimal value at Δx=3D. For axial distance 4D<Δx<10D, drag force coefficient of heat shield grows slowly and almost remain constant at Δx=10D, meaning heat shield is no longer affected by afterbody while Δx>10D.

The computational drag force coefficient of heat shield is compared with the experimental result in reference [8] as a function of axial separation distance up to ten diameters away, shown in Figure 4. The computational results match well with the experimental results.

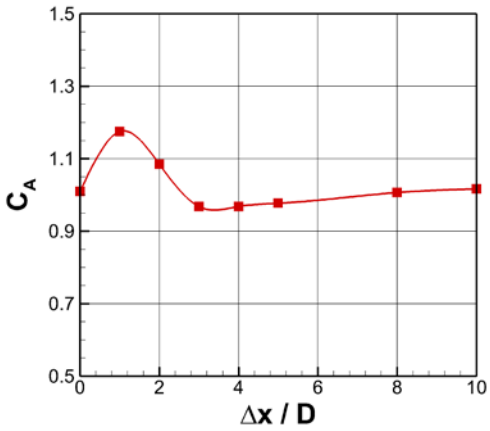


Figure 3. Variation of heat shield drag force coefficient with axial separation distance

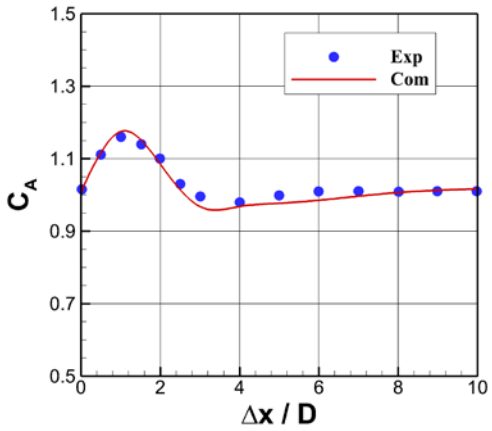


Figure 4. Comparison of heat shield computational result and experimental result

Back Shell Drag Force characteristics

Figure 5 shows variation of afterbody drag force coefficient with axial separation distance Δx. The drag force coefficient of afterbody is nearly zero at Δx=0, and the reason for this behavior is because the heat shield shadows afterbody and blocks the flow of incoming air. As axial separation distance increasing, the drag force coefficient of afterbody decreases rapidly and reaches minimal value at Δx=1D, and then increases for axial distance 1D<Δx<10D. Especially while Δx<2D, the drag force coefficient of afterbody is negative, which means the afterbody is being sucked forward because of formation of low pressure air flow in the volume between heat shield and afterbody.

The computational drag force coefficient of afterbody is also compared with the experimental result in reference [8] as a function of axial separation distance up to ten diameters away, shown in Figure 6. The comparison results are qualitative similar and quantitative different. The computational drag force coefficient reaches minimal value at Δx=1D while the experimental drag force coefficient reaches minimal value at Δx=2D.

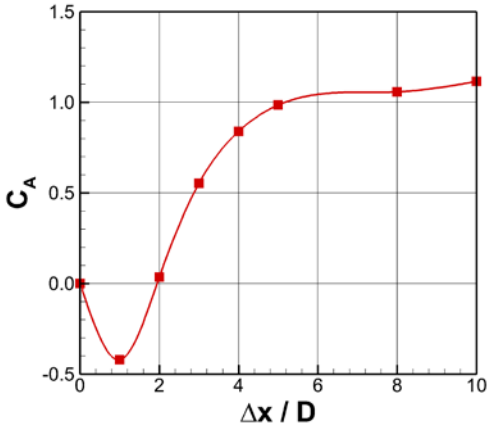


Figure 5. Variation of back shell drag force coefficient with axial separation distance

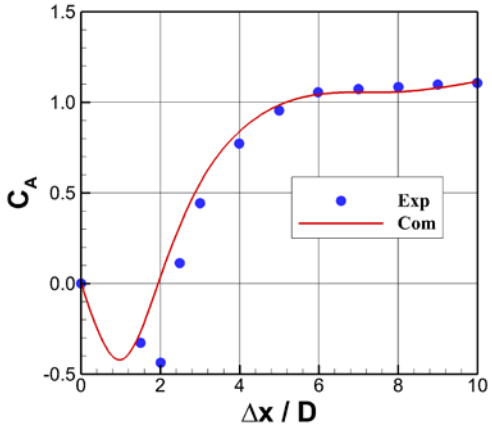


Figure 6. Comparison of back shell computational result and experimental result

Comparative Analysis

The drag coefficient curves for heat shield and afterbody are presented in Figure 7 as a function of axial separation distance.

The interference aerodynamics occurs when heat shield and afterbody are in close proximately to each other which influence their respective aerodynamics. This interaction effect decreases as their separation distance increases. When the axial separation distance is ten heat shield diameters away from afterbody, there is little aerodynamic interference.

For Δx<5D, the drag force coefficient of heat shield is bigger than afterbody, so that afterbody will descend faster than heat shield and there will be a risk of recontact. For 0<Δx<2D, the drag coefficient of afterbody becomes negative because of formation of low pressure air flow in the volume between heat shield and afterbody (see Fig.8). Conversely to the "suction" phenomenon on afterbody, the heat shield experiences increased drag in this region. The separation springs need to produce a sufficient impulse to overcome this "suction" force, otherwise the heat shield will simply slam back into afterbody. For Δx>5D, the drag force coefficient of heat shield is smaller than afterbody, so that heat shield will descend faster than afterbody and not recontact.

In an ideal scenario, as the heat shield separates, the spring system produces sufficient impulse to overcome the suction force and pushes the heat shield away from afterbody until Δx=5D. And then the ballistic coefficient of the heat shield will be just slightly higher than afterbody in order for separation to continue increasing with time while Δx>5D, and so that the heat shield will descend faster than afterbody and not recontact.

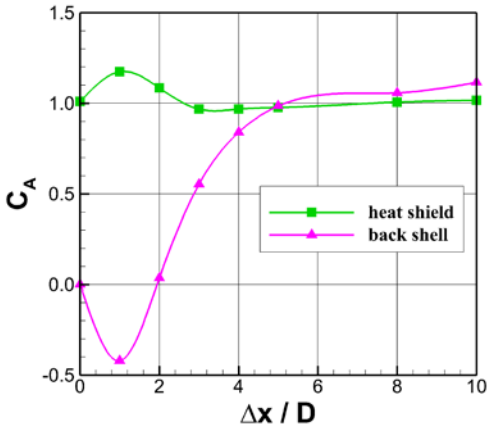


Figure 7. Comparison of heat shield and back shell drag force coefficient

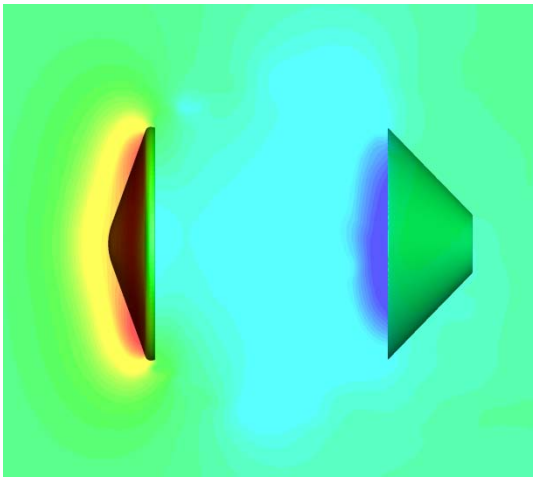


Figure 8. Flow field contour of axial separation distance Δx=1D

5.Conclusions

A RANS-LES method was constructed to numerical simulate heat shield separation of Mars Exploration Rover. This paper discusses how to assess the heat shield separation risk numerically and minimize the recontact risk. The outcome indicates that the computational results match well with the experimental results, which means the veracity of calculation method.

The drag force coefficient of heat shield is bigger than afterbody for $\Delta x<5D$, and there will be a recontact risk in this region. For $0<\Delta x<2D$, the interaction effect between heat shield and afterbody is most serious and the drag coefficient of afterbody becomes negative. For $\Delta x>5D$, the drag force coefficient of heat shield is smaller than afterbody, therefore the heat shield will descend faster than afterbody and not recontact.

References

[1] Vanov M. C., Blood E. M., Cook B. T., et al. (2011)*Entry, Des-cent and Landing Systems Analysis Study: Phase 2 Report on Mars Science Laboratory Improvement*, NASA/TM-2011-216988.
[2] Way D. W., Powell R. W., Chen A., et al. (2007)*Mars Science Laboratory: Entry, Descent, and Landing System Performance*, NASA Paper-1467.
[3] Karl T. E., Stephen J. A. (2003)*Computational Aeroheating Predictions for Mars Lander Configurations*, 36th AIAA Thermophysics Conference, Orlando,USA.
[4] Karl T. E., Artem A. D., Michael J. W., et al. (2007)*Aerothermodynamic Environments Definition for the Mars Science Laboratory Entry Capsule*, 45th AIAA Aerospace Sciences Meeting and ExhibitReno, Nevada, USA.
[5] Mark S., Artem A. D., Pieter B., et al. (2009)*Aerodynamic Challenges for the Mars Science Laboratory Entry, Descent and Landing*, 41st AIAA Thermophysics Conference, San Antonio, USA.
[6] Antonio V., Giuseppe P. (2009)*Aerodynamic Analysis of a Capsule Vehicle For a Manned Exploration Mission to Mars*, 16th AIAA/DLR/DGLR International Space Planes and Hypersonic Systems and Technologies Conference.
[7] Lang J. R., Sarver D. A. (1972)*Experimental aerodynamic Characteristics of the Viking Lander Capsule During Aeroshell Staging*, Martin Marietta Report TR-3720301.
[8] Behzad R., Prasun D., Robert M. (2011)*Mars Exploration Rover Heat Shield Recontact Analysis*, 21st AIAA Aerodynamic Decelerator Systems Technology Con-ference and Seminar, Dublin, Ireland.
[9] Behzad R., Eric M. Q. (2004)*Mars Exploration Rover Terminal Descent Mission Modeling and Simulation*, AAS-04-271.

[10] Gao B. (2009)Mars Exploration Entry Descent and Landing Technologies, *Spacecraft Recovery & Remote Sensing* **30**,1-9.
[11] Jia H., Rong W., Chen X., et al. (2010)Entry Descent and Landing System Analysis of Typical Sampling & Return Capsules , *Spacecraft Engineering* **19**,56-63.
[12] Jia H., Rong W. (2013)Review and Analysis of EDL Demonstrator Module of ExoMars 2016 Mission, *Spacecraft Engineering* **22**,109-115.
[13] Lu Y. Y., Rong W., Wu S. T. (2012)Introduction and New Technology of EDL System of MSL, *Spacecraft Engineering* **21**, 117-126.
[14] Peng Y. M., Li S., Man Y. Y., et al. (2010)New Progress of Mars Entry, Descent and Landing Technologies —Mars Science Laboratory Case Study, *Spacecraft Recovery & Remote Sensing* **31**, 7-14.
[15] Mark S., Cheatwood F. M., Desai P. N. (2005)*Static Aerody-namics of the Mars Exploration Rover Entry Capsule*, 43rd AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, USA.
[16] Liu Z., Yang Y. J., Zhou W. J., et al. (2014)Study of Unsteady Separation Flow Around Airfoil at High Angle of Attack Using Hybrid RANS-LES Method, *Acta Aeronautica et Astronautica Sinica* **35**, 372-380.
[17] Spalart P. R. (2009)Detached-Eddy Simulation, *Annual Re-view of Fluid Mechanics* **41**, 181-202.
[18] Strelets M. K. (2001)*Detached eddy simulation of massively separated flows*, AIAA 2001-0879.
[19] Shur M. L., Spalart P. R., Strelets M. K., et al. (2008)A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities, *International Journal of Heat and Fluid Flow* **29**, 1638-1649.
[20] Xu G. W., Zhou W. J., Chen B. Y., et al. (2014)Impact of Ablation on Aerodynamic Characteristics of Return Capsule, *Journal of Astronautics* **35**, 1114-1119.
[21] Picardi G., Biccari D., Melacci P., et al. (2005)*Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS): Subsurface Performances Evaluation*, ADA445262.

The Application of Data Mining in Health Care Informatics

***Hankun Hu[1], Fengjie Sun[2] and †Weidong Mao[2]**

[1]Zhongnan Hospital of Wuhan University, China
[2]School of Science & Technology, Georgia Gwinnett College, USA

*Presenting author: wb000400@whu.edu.cn
†Corresponding author: wmao@ggc.edu

Abstract

Health informatics is a multidisciplinary field that uses health information technology (HIT) to improve health care via any combination of higher quality, higher efficiency, and new opportunities. The availability of public health care information gives the opportunity for researchers to access and analyze the data. Data mining has been applied in health informatics in many area. Both environmental and genetic factors have roles in the development of some diseases. Complex diseases, such as Crohn's disease or Type II diabetes, are caused by a combination of environmental factors and mutations in multiple genes. Patients who have been diagnosed with such diseases cannot easily be treated. However, many diseases can be avoided if people at high risk change their living style. How to identify their susceptibility to diseases before symptoms are found and help them make informed decisions about their health becomes an important topic in health informatics. The susceptibility to complex diseases can be predicted through the analysis of the genetic data. With the development of DNA microarray technique, it is possible to access the human genetic information related to specific diseases. This paper uses a combinatorial method to analyze the genetic case-control data for Crohn's disease. A distance based cluster method has been applied to publicly available genotype data on Crohn's disease for epidemiological study and achieved a high accurate result.

Keywords: Health Informatics, Data Mining, Susceptibility Prediction

Introduction

Health informatics (also called health care informatics, healthcare informatics, medical informatics, nursing informatics, clinical informatics, or biomedical informatics) is informatics in health care. It is a multidisciplinary field that uses health information technology (HIT) to improve health care via any combination of higher quality, higher efficiency (spurring lower cost and thus greater availability), and new opportunities. The disciplines involved include information science, computer science, social science, behavioral science, management science, and others. It deals with the resources, devices, and methods required to optimize the acquisition, storage, retrieval, and use of information in health and biomedicine. Health informatics tools include amongst others computers, clinical guidelines, formal medical terminologies, and information and communication systems [1] [2]. It is applied to the areas of nursing, clinical medicine, pharmacy, public health, occupational therapy, physical therapy, biomedical research, and alternative medicine[3]. The availability of public health care information gives the opportunity for researchers to access and analyze the data.

Data mining has been applied in health informatics in many areas and one of those important area is the disease control and prevention. Complex diseases, such as Crohn's disease or Type II diabetes, are caused by a combination of environmental factors and mutations in multiple genes. Patients who have been diagnosed with such diseases cannot easily be treated. However, many diseases can be avoided if people at high risk change their living style. How to identify their susceptibility to diseases before symptoms are found and help them make informed decisions about their health becomes an important topic in health informatics. The

susceptibility to complex diseases can be predicted through the analysis of the genetic data. With the development of DNA microarray technique, it is possible to access the human genetic information related to specific diseases. Assessing the association between DNA variants and disease has been used widely to identify regions of the genome and candidate genes that contribute to disease [4]. 99.9% of one individual's DNA sequences are identical to that of another person. Over 80% of this 0.1% difference will be Single Nucleotide Polymorphisms (SNP) and they promise to significantly advance our ability to understand and treat human disease. A SNP is a single base substitution of one nucleotide with another. Each individual has many single nucleotide polymorphisms that together create a unique DNA pattern for that person. It is important to study SNPs because they represent genetic differences among human beings. Genome-wide association studies require knowledge about common genetic variations and the ability to genotype a sufficiently comprehensive set of variants in a large patient sample [5]. High-throughput SNP genotyping technologies make massive genotype data, with a large number of individuals, publicly available. Accessibility of genetic data makes genome-wide association studies for complex diseases possible.

It's important to search for informative SNPs among a huge number of SNPs. These informative SNPs are assumed to be associated with genetic diseases. Tag SNPs generated by the multiple linear regression based method [6] are good informative SNPs, but they are reconstruction-oriented instead of disease-oriented. Although the combinatorial search method for finding disease-associated multi-SNP combinations has a better result, the exhaustive search is still very slow.

The distance-based algorithm and cluster analysis have been used to solve the classification problem. In this algorithm, each item that is mapped to the same class may be thought of as more similar to the other items in that class than it is to the items found in other classes. Therefore, similarity measures may be used to identify the "alikeness" of different items in the database [7]. In our algorithm, the similarity is measured by the distance between the item and some neighbor clusters whose class label are previously known. The algorithm can be applied in our case-control study to predict an individual's susceptibility to Crohn's disease by comparing its genetic data with that of other individuals to find the similarity.

In this paper, we first address the disease susceptibility prediction problem [8] [9] [10]. This problem is to assess accumulated information targeted to predicting genotype susceptibility to complex diseases with significantly high accuracy and statistical power. Next, we introduce the cluster-based distance algorithm and its application in disease susceptibility prediction problem. We will also introduce the case tagging algorithm which is used to reduce the size of data and improve prediction results. The proposed method is applied to a publicly available data for Crohn's disease [11].

Disease Susceptibility Prediction Problem

A SNP is a single base substitution of one nucleotide with another. Both substitutions have to be observed in the general population at a frequency greater than 1%. An example of a SNP is individual "A" has a sequence GAACCT, while individual "B" has sequence GAGCCT, the polymorphism is an A/G. Each individual has many single nucleotide polymorphisms that together create a unique DNA pattern for that person. Haplotype is the set of adjacent SNP's are present on alleles in a block pattern. The genotype is the descriptor of the genome which is the set of physical DNA molecules inherited from the organism's parents. A pair of haplotype consists a single genotype.

SNP's are bi-allelic and can be referred as 0 if it's a majority and 1, otherwise. If both haplotypes are the same allele, then the corresponding genotype is homogeneous, can be represented as 0 or 1. If the two haplotypes are different, then the genotype is represented as 2.

The case-control sample populations consist of N individuals which are represented in genotype with M SNPs. Each SNP attains one of the three values 0, 1, or 2. The sample G is an (0, 1, 2)-valued $N \times M$ matrix, where each row corresponds to an individual, which is a sequence of 0, 1 and 2, each column corresponds to a SNP.

The disease susceptibility prediction problem can be formulated as follows:
Data sets have n genotypes. The input for a prediction algorithm includes:

(G1) Training genotype set $G = (g_i \mid i = 1..n)$;
(G2) Disease status $s(g_i) \in \{0,1\}$, indicating if g_i, is in case (1) or in control (0) , and
(G3) Testing genotype g_t without any disease status.

We will refer to the parts (G1-G2) of the input as the training set and to the part (G3) as the test case. The output of prediction algorithms is the disease status of the genotype g_t, i.e., $s(g_t)$.

Cluster-Based Distance Algorithm

To find an individual's disease status, we compute its similarity with other individuals whose disease status are already known. In the distance-based algorithm, each item that is mapped to the same class may be thought of as more similar to the other items in that class than it is to the items found in other classes. Therefore, similarity measures may be used to identify the "alikeness" of different items in the database. The idea of similarity measures can be abstracted and applied to more general classification problems. The difficulty lies in how the similarity measures are defined and applied to the items.

In this algorithm. We determine the similarity among sequences by computing their hamming distance. The hamming distance between two strings (in our case, two genotypes, each represents an individual) of equal length is the number of positions for which the corresponding symbols are different. For example, the hamming distance between genotype 1 (01021011) and genotype 2 (01021011) is 0, but the hamming distance between genotype 3 (21021210) and genotype 4 (01021011) is 3.

For the training data set, we build graph-based clusters for each class. In other words, we generate N clusters in case class and N clusters in control class given the threshold N. First we generate the graph G_{case} and $G_{control}$ based on the hamming distance among individual genotypes in case class and in control class, respectively. The Kruskal's algorithm is used to find the minimum spanning tree (MST) for G_{case} and $G_{control}$. To generate *N* clusters for Gcase, we need to remove the largest *N-1* edges. As a result, the inter-cluster distance is maximized and the intra-cluster distance is minimized. For these *2xN* clusters, we choose the genotype that has the minimum distance with all other genotype in the same cluster as the centroid of the cluster.

K nearest neighbors (KNN) is used as the classification scheme based on the use of distance measures. The KNN technique assumes that the entire training set includes not only the data in the set but also the desired classification for each item. In effect, the training data become the model. When a classification is to be made for a new item, its distance to each item in the training set must be determined. Only the K closest entries in the training set are considered further. The new item is then placed in the class that contains the most items from this set of K closest items. In our case, the hamming distance of the testing genotype to each centroid in the training set (including both case and control set) will be computed, then we find out the K closest genotypes which have smaller hamming distance than others. From the set of K centroid, if most of them are coming from the case group, then the testing genotype will be classified as case, otherwise, it will be classified as control. Obviously, it will be better if k is an odd number. Figure 1 shows how to classify the testing item when N is 4 and K is 3.

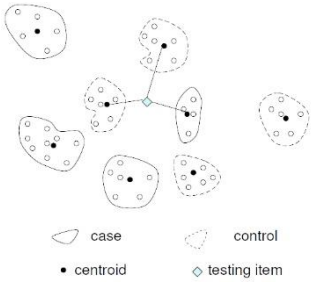


Figure 1. Classify the testing item, N=4, K=3

Results and Discussion

The genetic data is derived from the 616 kilo-base region of human Chromosome 5q31 that may contain a genetic variant responsible for Crohn's disease by genotyping 103 SNPs for 129 trios [11]. All offspring belong to the case population, while almost all parents belong to the control population. In the entire data, there are 144 case and 243 control individuals.

To measure the quality of prediction methods, we need to measure the deviation between the true disease status and the result of predicted susceptibility, which can be regarded as measurement error. We will present the basic measures used in epidemiology to quantify accuracy of our methods. The basic measures are sensitivity and specificity. Sensitivity is the proportion of persons who have the disease who are correctly identified as cases, while specificity is the proportion of people who do not have the disease who are correctly classified as controls. Sensitivity (accuracy in classification of case) and Specificity (accuracy in classification of controls) and Accuracy are calculated as follows in Eq. (1):

$$\begin{aligned} \text{Sensitivity} &= a\ (a + c) \\ \text{Specificity} &= d\ (b + d) \\ \text{Accuracy} &= (a + d)\ (a + b + c + d) \end{aligned} \tag{1}$$

a = True positive, people with the disease who test positive

b = False positive, people without the disease who test positive

c = False negative, people with the disease who test negative

d = True negative, people without the disease who test negative

Sensitivity is the ability to correctly detect a disease. Specificity is the ability to avoid calling normal as disease. Accuracy is the percent of the population that is correctly predicted.

We use K-fold cross validation method to measure the quality of the algorithm. In the K-fold cross validation, the data set is divided into K subsets, and the holdout method is repeated K times. Each time, one of the K subsets is used as the test set and the other K-1 subsets are put together to form a training set. In our experiment, we use 5-fold cross validation.

Table 1 is the experiment result. In this table, we compare the result when K is 1, 3, 5, 7 and the number of cluster N is 6, 10, 14, 18 and 22. The best result is as high as 100% for sensitivity, 100% for specificity, and 90% for accuracy, respectively.

Table 1. Experiment results

K	Measures	Number of Clusters (N)				
		6	**10**	**14**	**18**	**22**
1	Sensitivity	52	63	69	76	74
	Specificity	100	100	100	99	99
	Accuracy	76	82	85	88	87
3	Sensitivity	94	78	80	61	64
	Specificity	67	91	100	100	100
	Accuracy	81	85	90	81	82
5	Sensitivity	100	94	75	82	82
	Specificity	17	67	67	67	56
	Accuracy	59	81	71	75	69
7	Sensitivity	0	100	86	80	74
	Specificity	100	46	46	67	67
	Accuracy	50	73	66	74	71

Conclusions

In this paper, we discuss the potential of applying a cluster-based distance algorithm on how to predict the susceptibility for a complex disease, one of important problems in health care association studies. The proposed classification method based on cluster and distance is

shown to have a high prediction rate without finding SNPs associated with the disease which may reduce the running time. The genetic factors associated with the disease have to be identified first. In our future work we are going to continue validation of the proposed method on various type of data.

References

[1] O'donoghue, John; Herbert, John (2012). "Data management within mHealth environments: Patient sensors, mobile devices, and databases". Journal of Data and Information Quality (JDIQ). 4 (1): 5.

[2] Mettler T, Raptis DA (2012). "What constitutes the field of health information systems? Fostering a systematic framework and research agenda". Health Informatics Journal. 18 (2): 147–56. doi:10.1177/1460458212452496. PMID 22733682.

[3] Popularity of usage of software in homeopathy is shown in example video of homeopathic repertorisation: Shilpa Bhouraskar, Working quick acute cases on Homeopathic Software (YouTube).

[4] Cardon, L.R., Bell, J.I.: Association Study Designs for Complex Diseases, Vol.2. Nature Reviews: Genetics (2001), 91-98.

[5] Hirschhorn, J.N.,Daly, M.J.: Genome-wide Association Studies for Common Diseases and Complex Diseases, Vol.6. Nature Reviews: Genetics (2005), 95-108.

[6] He, J. and Zelikovsky, A.: Tag SNP Selection Based on Multivariate Linear Regression, Proc. of International Conference on Computational Science (2006), LNCS 3992, 750-757.

[7] Margaret H.D., Data Mining - Introduction and advanced topics, prentice Hall, ISBN 0-13-088892-3.

[8] Mao, W., Brinza, D., Hundewale, N., Gremalschi, S. and Zelikovsky, A.: Genotype Susceptibility and Integrated Risk Factors for Complex Diseases, Proc. IEEE International Conference on Granular Computing (GRC 2006), pp. 754-757.

[9] Kimmel, G. and Shamir R.: A Block-Free Hidden Markov Model for Genotypes and Its Application to Disease Association. J. of Computational Biology (2005), Vol. 12, No. 10: 1243-1260.

[10] Listgarten, J., Damaraju, S., Poulin B., Cook, L., Dufour, J., Driga, A.,Mackey, J., Wishart, D., Greiner,R., and Zanke, B.: Predictive Models for Breast Cancer Susceptibility from Multiple Single Nucleotide Polymorphisms. Clinical Cancer Research (2004), Vol. 10, 2725-2737.

[11] Daly, M., Rioux, J., Schaffner, S., Hudson, T. and Lander, E.: High resolution haplotype structure in the human genome. Nature Genetics (2001) 29, 229-232.

A new SPH iterative method for solving nonlinear equations

Rahmatjan Imin*, Ahmatjan Iminjan
College of Mathematics and Systems Science, Xinjiang University Urumqi 830046, China
*Corresponding author: rahmatjanim@xju.edu.cn

Abstract: In this paper, based on the basic principle of the SPH method's kernel approximation, a new kernel approximation was constructed to compute first order derivative through Taylor series expansion. Derivative in Newton's method was replaced to propose a new SPH iterative method for solving nonlinear equations. The advantage of this method is that it does not require any evaluation of derivatives, which overcame the shortcoming of Newton's method. Quadratically convergent of new method was proved and a variety of numerical examples were given to illustrate that the method has the same computational efficiency as the Newton's method.
Key words: SPH method; Nonlinear equations; Newton's method; Quadratically convergent; Iterative method.

1. Introduction

A variety of complex problems in different fields of science and engineering require finding the solution of a nonlinear equation (or the system of nonlinear equations) of the form $F(x) = 0$, in order to solve this equation researchers proposed many iterative methods [1–6]. Most of those methods were based on the well-known Newton's method, which is easy to implement and has quadratically convergence under fairly assumptions, however, it requires to compute $F'(x)$ with $F'(x)\neq 0$ in each calculation step, and some times it is difficult to provide the derivatives of function when the function is a complicated function. To overcome this problem, some derivative free iterative methods have been proposed [7–9].

In this paper, concerning derivative calculation in Newton's method, based on the basic principles of the smoothed particle hydrodynamics(SPH) [10, 11] method's kernel approximation, a new kernel approximation method was proposed to compute first order derivatives through Taylor series expansion. By replacing the derivatives in Newton's method, a new SPH iterative method was developed and the convergent of new method was proved. Several examples were calculated using the new method and the results were compared with Newton's method. Results showed that the new method almost has same accuracy and convergence with Newton's method.

2. Conventional SPH method

The first step of SPH method is the kernel approximation. The conventional kernel approximation for a function$f(x)$ and its derivatives at a particular point, whose position vector is x, in a volume Ω is defined as

$$\langle f(x)\rangle = \int_{\Omega} f(x')W(x-x',h)dx' \tag{1}$$

$$\langle \nabla f(x) \rangle = \int_{\Omega} f(x') \nabla W(x - x', h) dx' \tag{2}$$

Where $W(x - x', h)$ is kernel function and depends on two variable: distance of two points$|x - x'|$ and smoothing length h, which determines support domain of $W(x - x', h)$, the kernel function should satisfy some properties [12]. The most frequently used kernel function in SPH method is the cubic B-spline kernel. It has the form as shown in Eq.(3), Fig.1 shows that different smoothed length ($h = 1$ and $h = 0.5$ in one dimensional case) affects influence radius and shape of cubic B-spline kernel.

$$W(x - x', h) = \alpha_d \begin{cases} \frac{2}{3} - R^2 + \frac{1}{2}R^3 & 0 \leq R < 1 \\ \frac{1}{6}(2 - R)^3 & 1 \leq R < 2 < 0 \\ 0 & R \geq 2 \end{cases} \tag{3}$$

Where $\alpha_d = \frac{1}{h}, \frac{15}{7\pi h^2}, \frac{3}{2\pi h^3}$ in one, two and three dimensional cases respectively, R is the relative distance between two points (particles) at points x and x', where $R = \frac{r}{h} = \frac{|x - x'|}{h}$ in which r is the distance between the two points.

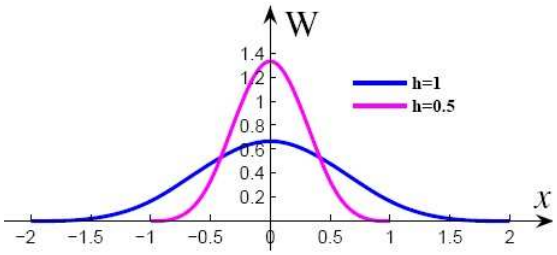


Fig. 1: **Figure of cubic B-spline kernel functionin 1D case**

The second step of SPH method is the particle approximation, in which the problem domain is discretized into a finite number of randomly distributed particles which have mass and volume (see Fig. 2 For illustration in a two-dimensional case). Suppose infinitesimal volume dV' in the above integration at the location of j was replaced by the finite volume of the particle ΔV^j. If the particle mass and density are concerned, the ΔV^j can also be replaced by the corresponding mass to density ratio m^j/ρ^j; the kernel approximation of a function and its derivatives expressed in Eq.(1) and Eq.(2) can be written in the following form of discretized particle approximation.

$$\langle f(x^i) \rangle = \sum_{j=1}^{M} \Delta V^j f^j W^{ij} = \sum_{j=1}^{M} \frac{m^j}{\rho^j} f^j W^{ij} \tag{4}$$

$$\langle \nabla f(x_i) \rangle = \sum_{j=1}^{M} \Delta V^j f^j \nabla W^{ij} = \sum_{j=1}^{M} \frac{m_j}{\rho_j} f^j \nabla W^{ij} \tag{5}$$

where $f^j = f(x^j)$, M is the number of particle in the support domain of particle i, $W^{ij} = W(x_i - x_j, h) = W(r^{ij}, h)$, $\nabla W^{ij} = \nabla W(x - x', h)\big|_{x=x^i, x'=x^j}$, $r^{ij} = |x^i - x^j|$

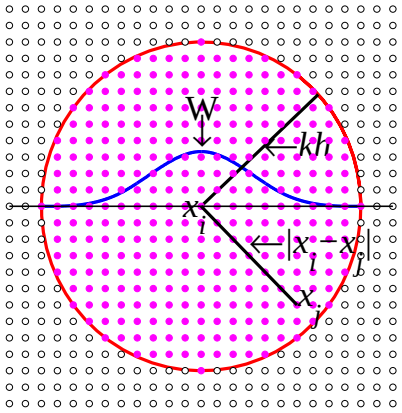


Fig. 2: **SPH particle approximations**

3. New kernel approximation method to compute first order derivatives

Let f be a function of variable x , f'' continuous and f''' exist in neighborhood of $x = x^i$, Taylor series expansion of f about the point $x = x^i$, retaining only three terms, written as

$$f(x) = f(x^i) + \frac{\partial f(x^i)}{\partial x_\alpha}(x_\alpha - x^i_\alpha) + \frac{1}{2}\frac{\partial^2 f(x^i)}{\partial x_\alpha \partial x_\beta}(x - x^i_\alpha)(x - x^i_\beta) + O[(x - x^i)^3] \tag{6}$$

where repeated indices α and β are summed over their ranges, but the repeated index i enclosed in parentheses is not summed.

By multiplying both sides of Eq.(6) with $(x_\gamma - x^i_\gamma)W(x - x^i, h)$ and integration of the resulting equation over the support domain Ω of kernel function $W(x - x^i, h)$ yields.

$$\begin{aligned}\int_\Omega f(x)(x_\gamma - x^i_\gamma)W(x - x^i, h)dx = f(x^i)\int_\Omega (x_\gamma - x^i_\gamma)W(x - x^i, h)dx \\ + \frac{\partial f(x^i)}{\partial x_\alpha}\int_\Omega (x_\alpha - x^i_\alpha)(x_\gamma - x^i_\gamma)W(x - x^i, h)dx \\ + \frac{1}{2}\frac{\partial^2 f(x^i)}{\partial x_\alpha \partial x_\beta}\int_\Omega (x - x^i_\alpha)(x - x^i_\beta)(x_\gamma - x^i_\gamma)W(x - x^i, h)dx + O(h^4)\end{aligned} \tag{7}$$

Due to the symmetry of support domain Ω and the kernel function $W(x - x^i, h)$ is an even function , hence third integral on right hand side of Eq.(7) is equal to zero, therefore the Eq.(7) simplified to

$$\int_\Omega \big[f(x) - f(x^i)\big](x_\gamma - x^i_\gamma)W(x - x^i, h)dx = \frac{\partial f(x^i)}{\partial x_\alpha}\int_\Omega (x_\alpha - x^i_\alpha)(x_\gamma - x^i_\gamma)W(x - x^i, h)dx \tag{8}$$

By applying the particle approximation principle of SPH, from equation (8) leads to the following particle approximation of the derivative at a particle point i

$$\Big\langle \frac{\partial f(x^i)}{\partial x_\alpha} \Big\rangle = \Big(\sum_{j=1}^{M} \frac{m^j}{\rho^j} x^{ji}_\alpha x^{ji}_\gamma W^{ji}\Big)^{-1}\Big(\sum_{j=1}^{M} \frac{m^j}{\rho^j} x^{ji}_\gamma f^{ji} W^{ji}\Big) \tag{9}$$

where $x^{ji}_\alpha = x^j_\alpha - x^i_\alpha$, $x^{ji}_\gamma = x^j_\gamma - x^i_\gamma$, $f^{ji} = f(x^j) - f(x^i)$, $W^{ji} = W(x^j - x^i, h)$.

From the above Eq.(9) it can be seen that the calculation process of derivatives is not need to calculate the derivative of the kernel function. The kernel function could be used directly, reducing differentiability requirement of kernel function, at the same time a larger class of kernel functions could be used and the calculation is made easier.

3. The algorithm of new SPH iterative method

The famous Newton's methods is

$$x_{n+1} = x_n - \left[F'(x_n)\right]^{-1} F(x_n) \tag{10}$$

Applying new kernel approximation method introduced above into Eq.(10), get corresponding new SPH iterative method.

For numerical implementation of the presented method, provide the following algorithm.
Step 1:initial values of x_0, tolerance ε , maximum number of iterations N, $n = 0$ and set the smooth length of kernel function h.
Step 2: while ($n \leq N$) do step(3-7)

Step 3: use Eq.(9) to compute $F'(x_n)$

Step 4: compute $y = -\left[F'(x_n)\right]^{-1} F(x_n)$

Step 5: $x_{n+1} = x_n + y$

Step 6: if $\| F(x_n) \| \leq \varepsilon$ or $\| x_{n+1} - x_n \| \leq \varepsilon$ output approximate solution x^* and stop

Step 7: set $n = n + 1$ and $x_n = x_{n+1}$

Step 8: output "the method failed after N iterations"; Stop

4. The convergent proof of new SPH iterative method

For the sake of simplicity only given and proved one variable case.
Theorem: Let $f''(x)$ continuous and $f'''(x)$ exist in neighborhood of $x = x^*$; and let $f(x^*) = 0$. The iterative method defined by equation (10) is quadratically convergent for smoothed length $h = K\Delta x$ and some Δx is sufficiently small.
Proof:. In equation(9), if we use cubic B-spline kernel and set smoothed length $h = 1.1\Delta x$ in iterative process, the Eq.(10) will become following iterative formula accordingly

$$x_{n+1} = x_n - \frac{878\Delta x f(x_n)}{4[f(x_n + 2\Delta x) - f(x_n - 2\Delta x] + 431[f(x_n + \Delta x) - f(x_n - \Delta x)]}$$

Let $e_n = x_n - x^*$, by using Taylor series expansion,we have

$$f(x_n) = e_n f' + \frac{1}{2} e_n^2 f'' + o(e_n^2)$$

$$f(x_n + \Delta x) = (e_n + \Delta x) f' + \frac{1}{2}(e_n + \Delta x)^2 f'' + o((e_n + \Delta x)^2)$$

$$f(x_n + 2\Delta x) = (e_n + 2\Delta x) f' + \frac{1}{2}(e_n + 2\Delta x)^2 f'' + o((e_n + 2\Delta x)^2)$$

here $f' = f'(x^*), f'' = f''(x^*)$. Therefore,there holds

$f(x_n+\Delta x)-f(x_n-\Delta x)=2\Delta x\big[f'+e_n\Delta x f''+o(e_n)\big]$
so we have
$$e_{n+1}=e_n-\frac{e_nf'+1/2e_n^2f''+o(e_n^2)}{f'+e_n\Delta xf''+o(e_n)}$$
$$=e_n^2\frac{1/2f''+o(1)}{f'+e_n\Delta xf''+o(e_n)}$$
namely,
$$\lim_{n\to\infty}\frac{e_{n+1}}{e_n^2}=\frac{f''}{2f'}$$
which shows that the iterative method is at least a quadratically convergent.

5. Numerical examples

In this section presented some examples to illustrate the efficiency of the new SPH iterative method comparing with Newton's method (NM)(see Table1-3). In each example, the following stopping criteria was used.

$$\parallel F(x_n)\parallel\leq 10^{-15} \tag{11}$$

Computational order of convergence (COC)$\to\rho$ defined as:

$$\rho\approx\frac{\parallel x_{n+1}-x_n\parallel\,/\,\parallel x_n-x_{n-1}\parallel}{\parallel x_n-x_{n-1}\parallel\,/\,\parallel x_{n-1}-x_{n-2}\parallel} \tag{12}$$

Example 1:Application to one variable nonlinear equation

(1.1)$x^2-e^x-3x+2=0$ (1.2)$arctanx+sinx+x-2=0$

(1.3)$sin^2x-x^2+1=0$ (1.4)$x^2-(1-x)^5=0$

Example 2:Application to binary nonlinear equation

(2.1) $\begin{cases}\dfrac{x}{tanx}=-y\\ x^2+y^2=3.5^2\end{cases}$ (2.2) $\begin{cases}e^{x^2}+8xsiny=0\\ x+y=1\end{cases}$ (2.3) $\begin{cases}x^2-2x-y=-0.5\\ x^2+4y^2=4\end{cases}$

Example 3:Application to ternary nonlinear equation

(3.1) $\begin{cases}x^2+y^2+z^2=1\\ 2x^2+y^2-4z=0\\ 3x^2-4y^2+z^2=0\end{cases}$ (3.2) $\begin{cases}x^2+y^2+z^2=9\\ xyz=1\\ x+y-z^2=0\end{cases}$

6. Conclusion

In this paper the SPH method used to finding roots of nonlinear equations and a new SPH iterative method for finding roots of nonlinear equations $F(x)=0$ were proposed which has the following advantages:

(1)This method is available for finding roots multiple dimensions equations.

(2) The method only needs initial approximation of x_0 and it is not only need not to calculate any evaluation of derivatives of $F(x)$ and also need not to calculate of derivatives of kernel function $W(x-x',h)$.

(3)The method is quadratically convergent and keep the same convergent and computational efficiency

Table 1: **Comparison of the presented method and Newton's methods to one variable nonlinear equation**

Examples	Root	Initial Value	Number of Iterations		COC—ρ	
			NM	SPH	NM	SPH
1.1	0.257530285439861	2	5	5	2.0006	2.0006
		-1	5	5	2.0007	2.0007
1.2	0.718586769063582	3	10	10	2.0007	2.0007
		0.1	5	5	1.9997	1.9997
1.3	1.404491648215340	2	5	5	2.0005	2.0005
		1	6	6	1.9999	1.9999
1.4	0.345954815848242	2	7	7	2.0012	2.0012
		-2	11	11	2.0004	2.0004

Table 2: **Comparison of the presented method and Newton's methods to binary nonlinear equation**

Examples	Root	Initial Value	Number of Iterations		COC—ρ	
			NM	SPH	NM	SPH
2.1	(2.389946943809752,	(3, 3)	8	8	1.9957	1.9957
	2.556981346387655)	(2, 2)	6	6	2.0003	2.0003
2.2	(-0.140285010811190,	(0.2, 0.8)	5	5	1.9997	1.9994
	1.140285010811190)	(-0.2, 2)	5	5	1.9621	1.9621
2.3	(-0.222214555059722,	(0.5, 0.5)	7	7	2.0015	2.0016
	0.993808418599834)	(0.5, 1.5)	5	5	1.9996	1.9996

Table 3: **Comparison of the presented method and Newton's methods to ternary nonlinear equation**

Examples	Root	Initial Value	Number of Iterations		COC—ρ	
			NM	SPH	NM	SPH
3.1	(0.560573416264006,	(0.5, 0.5, 0.5)	5	5	2.0000	2.0000
	0.497671169854288,					
	0.219040425836983)	(1, 1, 1)	6	6	2.0065	2.0065
3.2	(2.491375696830689,	(2.5, 0.5, 1.5)	5	5	1.9850	1.9850
	0.242745878757136,					
	1.653517939300274)	(2, 0.1, 2)	5	5	1.9183	1.9181

with the Newton method.

Acknowledgments This work has been supported by the National Natural Science Foundation of China (Grand No:51565054) and the Natural Science Foundation of Xinjiang University(Starting foundation for PhD., Grand No:BS150210).

References

[1] M. Basto, V. Semiao, F. L. Calheiros, A new iterative method to compute nonlinear equations, Applied Mathematics and Computation,2006,17(1):468-483.

[2] V. Kanwar, S. Singh, S. Bakshi, Simple geometric constructions of quadratically and cubically convergent iterative functions to solve nonlinear equations, Numerical Algorithms,2008,47(1):95-107.

[3] C. Chen, B. Neta, Certain improvements of Newton's method with fourth-order convergence, Applied Mathematics and Computation,2009,215(2):821-828.

[4] L. Liu, X. Wang, Eighth-order methods of high efficiency index for solving nonlinear equations, Applied Mathematics and Computation,2012,25(12):3449-3454.

[5] Y. Khan, M. Fardi, K. Sayevand, A new general eighth-order family of iterative methods for solving nonlinear equations, Applied Mathematics Letters,2012,25(12):2262-2266.

[6] Kaysar Rahman, Rahmatjan Yimin, Nurmamat Helil, A Fourth Order Iterative Method for Solving Nonlinear Equations, Computer Science and Information Technology,2010,8:628-630.

[7] YehuiPeng, HeyingFeng, QiyongLi, XiaoqingZhang, A fourth-order derivative-free algorithm for nonlinear equations, Journal of Computational and Applied Mathematics, 2011,235(8):2551-2559.

[8] BeongInYun, Solving nonlinear equations by a new derivative free iterative method, Applied Mathematics and Computation, 2011,217(12):5768-5773.

[9] A. Cordero, J. L. Hueso, E. Martinez, J. R. Torregrosa, Generating optimal derivative free iterative methods for nonlinear equations by using polynomial interpolation, Mathematical and Computer Modelling, 2013,57(7-8):1950-1956.

[10] L. Lucy, A numerical approach to the testing of fission hypothesis, Astrophysical Journal,1977,82(12):1013-1020.

[11] R. Gingold, J. J. Monaghan, Smoothed particle hydrodynamics: Theory and application to non-spherical stars, Monthly Notices of the Royal Astronomical Society,1977,181:375-389.

[12] G. R. Liu, M. B. Liu, Smoothed Particle Hydrodynamics: A Meshless Particle Method, Singapore, World Scientific Publishing, 2003:150-156.

Numerical Simulation of Instantaneous Wave-Free Ratio of Stenosed Coronary Artery

†Wenxin Wang[1], Boyan Mao[1], Bao Li[1], *Youjun Liu[1], Jinshseng Xie[2]

1 College of Life Science and Bioengineering, Beijing University of Technology, Beijing, China
2 Beijing AnZhen Hospital Affiliated to Capital Medical University, Beijing, China

*Presenting author: wangwenxin@emails.bjut.edu.cn
†Corresponding author: lyjlma@bjut.edu.cn

Abstract

Instantaneous wave-free ratio (iFR), an invasive index of coronary artery tree, can evaluate the functional performance of vascular stenosis without pharmacological vasodilators. The non-invasive assessment of diameter stenosis (DS) obtained from coronary computed tomography angiography (CTA) has high false positive rate in contrast to iFR. The aim of this study was to developed a numerical simulation method that predicts the iFR and non-invasive assess the myocardial ischemia. Based on the CTA images, a patient-specific 3 dimensional model of the aorta and coronary arteries were reconstructed. A stenosis was created in the left anterior descending artery (LAD) by reducing the DS of geometric model (40%, 50%, 60%, 75% and 90%). The patient-specific LPM boundary condition were set up to compute iFRct value during the wave-free period at the resting condition. The computed pressure and flow of coronary artery were realistic as compared to literature data. The iFR of the numerical simulation results (iFRct) has very significant positive correlation with the mean flow rate of LAD, which means that iFRct was a functional index of coronary stenosis and was consistent with the clinical significance. In contrast to invasive iFR, the iFRct can make a cost-benefit balance in terms of clinical cost and patient's health.

Keywords: Computational fluid dynamics (CFD), Coronary artery, Instantaneous wave-free ratio (iFR)

Introduction

Coronary artery disease is caused by coronary stenosis, which reduces the myocardial perfusion, results in ischemia, infarction, and even death. Stenosis is only an anatomic description of coronary artery disease. In clinic, in order to treat coronary artery disease, it is important to know whether the physiological or hemodynamic significance of a coronary stenosis could cause the coronary ischemia or not. Coronary stenosis severity has no absolute correlation with myocardial ischemia.

In recent years, instantaneous wave-free ratio (iFR), an important index of coronary stenosis, isolates a specific period in diastole, which called the wave-free period (WFP), and uses the ratio of distal coronary pressure (Pd) to the pressure observed in the aortic (Pa) [1]. In this period, the competing forces (waves) that affect coronary flow are quiescent meaning pressure and flow are linearly related as compared to the rest of the cardiac cycle [2]. During the WFP, the coronary flow rate is higher and the pressure is lower, which leads to coronary microcirculation resistance the most stable and the lowest. At the same time, the pressure and flow of coronary artery are linearly related as compared with the rest of cardiac cycle. The WFP begins 25% into diastole and ends 5ms before the end of diastole [3]. The concept of iFR was originally came from the wave intensity analysis using both coronary pressure and flow rate values [4]. Fractional flow reserve (FFR) is the gold standard for determining whether a coronary artery stenosis causes ischemia [5][6]. Sen S et al. [2] studied that there was a good correlation between iFR and FFR. FFR and iFR can only be measured using invasive coronary artery catheterization, however, which limits its widespread use in clinical

practice. The data have shown that the clinical application rate of FFR is less than 6% in America [7], and that of iFR is even less.

On the other hand, computational fluid dynamics (CFD) has been applied to simulate the blood flow and pressure in patient-specific coronary artery models, which are coupled the lumped parameter model and the 3D model reconstructed from coronary computed tomography angiography (CTA) [8]. The multicenter and prospective clinical trials of DISCOVER-FLOW [9][10], DeFACTO [11] and NXT[12], which predicted the FFRct value through combining patient-specific CTTA images and CFD simulations, have improved diagnostic accuracy than CTA alone to evaluate the ischemic and non-ischemic stenosis. Therefore, iFRct derived from non-invasively through combining CTTA images and CFD can improved diagnostic accuracy as same as FFRct. We just need to simulate at the resting state, without the hyperemia state. In this way, it can avoid the mistake causing by simulating the hyperemia state.

In this study, we developed a numerical simulation method, which could simulate the pressure and flow of coronary artery at the resting (baseline) condition and predict the iFR value (iFRct) from CTA images using 0D/3D coupled method. The 0D/3D coupled method was used in this study to perform a numerical simulation by coupling the lumped parameter model (LPM; 0D sub-model) and 3D vascular sub-models. This method has been used in our previous hemodynamic research and proven useful in studying cardiovascular system [13][14]. The iFRct method can non-invasively assess the level of myocardial ischemia and avoid the deviations when simulate the coronary vascular beds were maximally dilated (hyperemia) condition.

Method

Study Design

The main purpose of this study was to develop a numerical simulation method which could predict the iFRct derived from CTA images and non-invasive assess the myocardial ischemia. The Ethics approval and consent to participate are not applicable, as it is a retrospect study. The patient's records or information have been anonymized prior to analysis. The patient didn't have a cardiac event and/or PCI in CTA examination in this study.

The patient-specific three-dimensional (3D) anatomic geometry was reconstructed, as illustrated in Fig.1. A moderate stenosis (DS≈50%) was applied to left anterior descending (LAD) in patient-specific model. In order to verify the feasibility and correctness of the computational simulation method, the other four 3D stenosis models were built by reducing the local diameter of the geometric model (including 40%, 60%, 75% and 90%) at LAD.

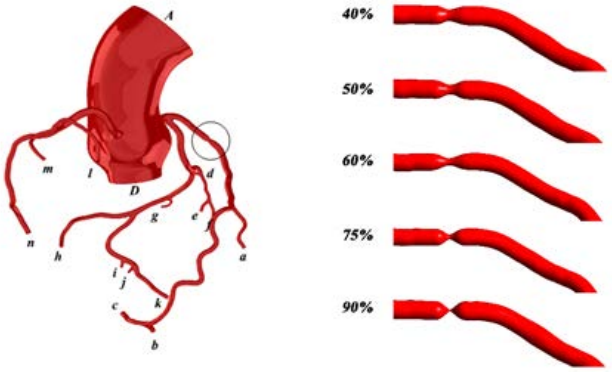


Fig 1. The reconstructed 3D geometry and LAD with different diameter reduction

Imaging acquisition

In this study, CTA images of the patient were provided by Beijing An Zhen Hospital in China. The 3D geometry of coronary arteries was reconstructed from CTA images with a slice thickness of 1mm. The simultaneous acquisition of multi-slice enabled the images of coronary and aorta in a single breath hold. The initial data window was positioned at 70% of the R–R interval. The cardiac output of the patient measured was 5.4 L/min. Four hundred and sixty slices of CTA images with 512×512 pixels were used for 3D reconstruction.

Geometrical models and computational models

Similar to our previous studies, the patient-specific 3D anatomic geometry was reconstructed through both the threshold segmentation and manual segmentation using the software MIMICS [13]-[16]. Based on these data, a moderate stenosis (stenosis degree ≈ 50%) was applied to the LAD artery. A hexahedral mesh was generated mainly by using the ANSYS-CFX (ANSYSTM). A steady state grid sensitivity analysis was conducted to make sure that the relative error in two consecutive mesh refinements < 1% for the maximum velocity.

In order to verify the correctness of the numerical simulation method, other four models with different stenosis in the LAD artery were constructed, reducing the local diameter of the geometric model by 40%, 60%, 75% and 90% respectively. Table 1 provides the geometric characteristics of stenosis LAD. The number of nodes and elements in these five models are within the same order of magnitude.

Similar to our previous study [14][15][17][18], the lumped parameter models were used in this study. On the basis of the 3D vascular models and the LPMs, 0D/3D coupled models were constructed, as illustrated in Fig.2. The boundary conditions of the 3D part were supplied by the 0D calculation and the forcing terms of the 0D part were calculated by the 3D simulation. Five 3D models shared the same 0D part since the patient's peripheral vascular structure was not changed with different stenosis. The algorithm of 0D/3D coupling method used in this study was also applied in our previous study. The difference was that we modified the lumped parameter model of right coronary which was not subjected to the pressure of the left ventricle (Plv). In addition, the coronary venous microcirculation compliance was eliminated from the original model in order to simplify the numerical simulation [8].

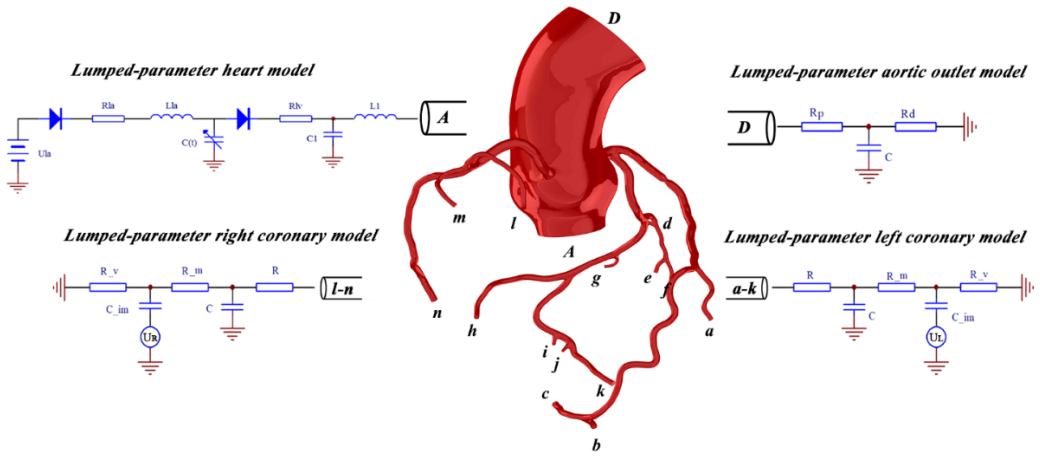


Fig 2. The 0D/3D coupled model

The 0D/3D coupled model

Similar to our previous study [13][14], governing equations were formulated for aorta and coronary arteries. In the 3D model, it was assumed that the vessel walls were rigid. The blood flow was treated as an incompressible viscous Newtonian fluid. The density of the blood flow was 1050 kg/m3 and the dynamic viscosity was 0.0035 Pa s. Navier-Stokes and continuity equations were solved using the commercial software solver ANSYS-CFX. Similar to previous studies [13][14], four cardiac cycles were set to implement convergence for transient analysis. The explicit Euler method was used and the constant time step was set in per cardiac cycle (=0.8s). In the 3D model, the time step was 0.0025s, while that was 0.0001s in 0D model. The boundary conditions of the 3D part were supplied by the 0D calculation and the forcing terms of the 0D part were calculated by the 3D simulation. All the 3D models shared the same 0D part since the patient's peripheral vascular structure did not change with surgery. Therefore, the variation of simulation results can be considered that only caused by the differences of 3D models. In each block of coronary branch, the resistance (R) was used to simulate the flow resistance, the capacitance (C) was used to simulate the compliance of the vessel, and the inductance (L) was used to simulate the inertia of the blood flow.
In the compartment of the ventricle, the function of pressure-volume relationship was applied to demonstrate the cardiac cycle of the left and right ventricle.

$$E(t) = \frac{P(t)}{V(t) - V_0} \tag{1}$$

Where $E(t)$ is the time-varying elastance (mmHg/ml). and are the ventricle volume (ml) and pressure (mmHg) respectively. is the reference volume (ml). Mathematically, the function was used as the approximation.

$$E(t) = (E_{\max} - E_{\min}) \cdot E_n(t_n) + E_{\min} \tag{2}$$

where $E_n(t_n)$ is the normalized time-varying elastance.

$$E_n(t_n) = 1.55 \left[\frac{\left(\frac{t_n}{0.7}\right)^{1.9}}{1 + \left(\frac{t_n}{0.7}\right)^{1.9}} \right] \left[\frac{1}{1 + \left(\frac{t_n}{1.17}\right)^{21.9}} \right] \tag{3}$$

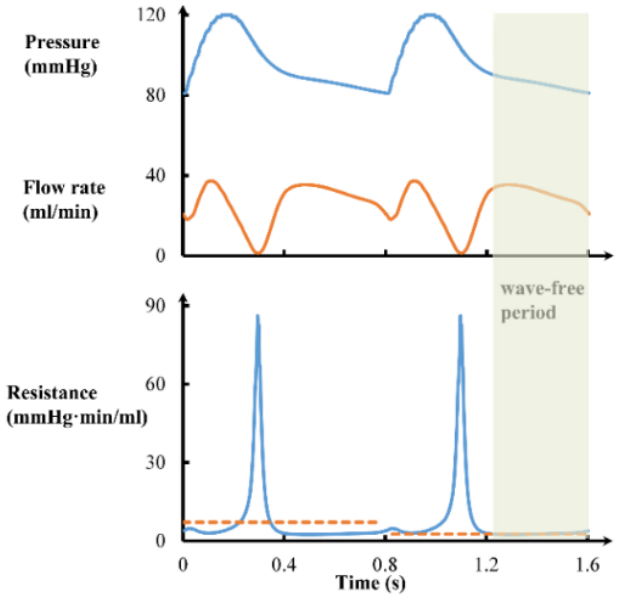
tc is the cardiac cycle interval (s). In this paper, we set $E_{\max} = 2.0$, $E_{\min} = 0.002458$ and $tc = 0.8s$.
The value of the parameters in the 0D network of coupled models were based on data from research into the modeling of coronary arteries [8]. The genetic algorithm was used to tune the parameters of the LPM model. In this way, the systolic pressure, the diastolic pressure and the cardiac output was matched the patient's data.

Results

The iFRct values has been proposed as a new index of stenosis severity that is independent of hyperemia. It used the ratio of distal coronary pressure (Pd) to the pressure observed in the aortic (Pa) over this period.

$$iFR = \frac{Pd}{Pa} \tag{4}$$

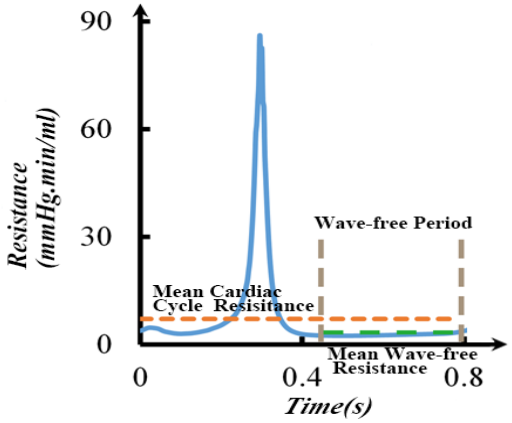
Not only the pressure, flow rate and the resistant of the section, but also the WFP is illustrated



in Fig.3.

Fig 3. Definition of the wave-free period

Moreover, the mean resistant of cardiac cycle and wave-free period are illustrated in Fig.4. The orange line and green line represent the resistance of the whole cardiac cycle and WTP,



respectively. Compared with the resistance of cardiac cycle, the resistance of WTP decreases 51.25%.

Fig 4. Mean resistance in whole cardiac cycle and wave-free period

The aortic pressure is used as Pa in the equation (4), while the coronary pressure distal to the stenosis in the 3D model is used as Pd. For every different stenosis model, the Pd and Pa in cardiac cycle are plotted in Fig.5.

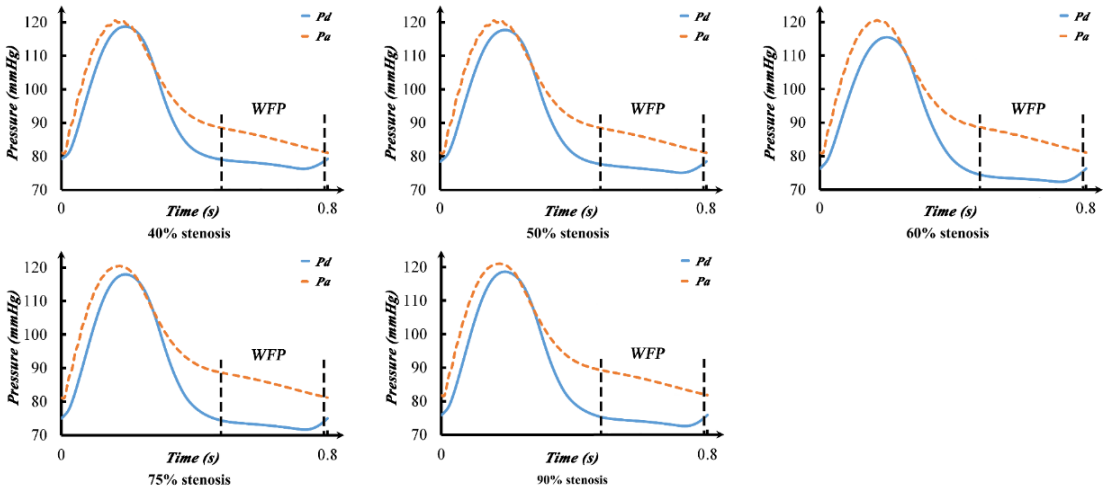


Fig 5. The Pd and Pa of five different stenosis model

During the WFP, the contours of iFRct in different stenosis models are illustrated in Fig.5. With the increase of the degree of the stenosis (DS), the iFRct value gradually reduces.

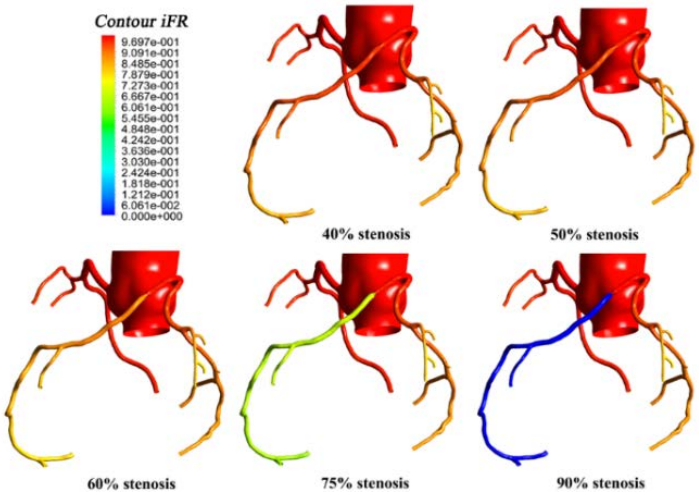


Fig 6. The iFRct contour of different stenosis models

The iFRct value of Model 1 is 0.965. And the iFRct of other four models are listed in Table 1. With the increase of the degree of stenosis, the mean flow rate of LAD is gradually decreased as followed: 81.18 ml/min, 79.64 ml/min, 77.47 ml/min, 64.22 ml/min and 12.79 ml/min. The iFRct contour of different stenosis models are plotted in Fig.7.

Table 1. The iFRct values of different stenosis models

Model	DS (%)	AS (%)	Flow Rate	iFRct
1	40	64.00	81.18	0.965
2	50	75.00	79.64	0.948
3	60	84.00	77.47	0.907
4	75	93.75	64.22	0.680

5	90	99.00	12.79	0.060

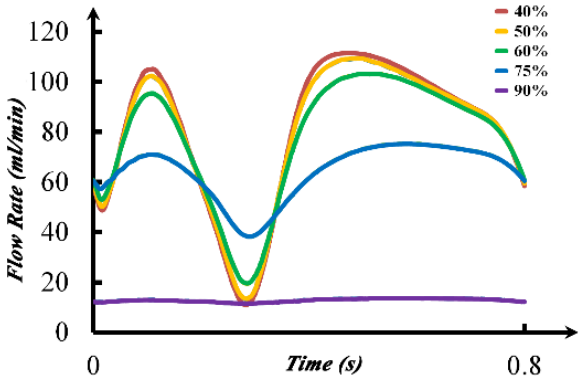


Fig. 7 The flow rate of different stenosis models

Discussion

In order to investigate the numerical simulation method of iFRct of stenosed coronary artery, multi-scale studies were conducted based on patient-specific model with different stenosis.

Similar to some study of Taylor, the inflow boundary condition was used, which coupled the lumped parameter heart model and a lumped parameter coronary model. We have successfully developed a numerical simulation method that coupled the 0D and 3D coronary model to calculate the iFRct. Noninvasive evaluation of the functional significance of coronary stenosis is possible using numerical simulation method based iFR from CTA data.

When the degree of stenosis was more than 75%, iFR was less than 0.86 (the critical value of myocardial ischemia). The DS was negatively correlated with iFR and the mean flow rate of LAD in low level of significance. The iFR had very significant positive correlation with the mean flow rate of LAD. So using iFRct was more accurate, to assess the myocardial blood supply. The iFR was a functional index of coronary stenosis in clinical application, and was consistent with the numerical simulation results. At the same time, it was proved the numerical simulation method could calculate the iFR.

This promising simulation method provided a combined anatomic-physiology-functional evaluation of coronary artery disease with the aim of improving clinical diagnostic accuracy while avoid the vasodilators, invasive pressure wire techniques and the unnecessary costs.

Future Work and Limitation

This paper developed a noninvasive method to assess myocardial ischemia. In the future, the physiologic realism of the boundary conditions can be improved even further in the simulation by the closed-loop lumped parameter network model which may consider the interactions between the heart and arterial system, and the models with different patients (have different afterload) to assess competitive flow is necessary.
Our model just considered the LAD artery with the only one stenosis, since the incidence of atherosclerotic lesions in the LAD artery was the highest and the stenosis was severe. The results and conclusion might be different when the coronary artery with two or more stenoses.

Conclusion

We have successfully developed a numerical simulation method to predict the iFR and non-invasive assess the myocardial ischemia. In anatomical, there is not absolutely correlation between sever stenosis of coronary artery and myocardial ischemia. Therefore, it is necessary to non-invasive assessment of myocardial ischemia by simulating iFRct.

FFR and iFR are both a functional evaluation indicators of coronary artery, however, the numerical calculation of iFRct requires only simulating the resting state without the hyperemia state. Some hypotheses must be set to simulate the hyperemia state of coronary artery, which may cause deviation with real physiological situation and may impair the accuracy of the assessment.

The iFR of the numerical simulation results (iFRct) have very significant positive correlation with the mean flow rate of LAD, which means that iFRct is a functional index of coronary stenosis and is consistent with the clinical significance. In contrast to invasive iFR, the iFRct can make a cost-benefit balance in terms of clinical cost and patient's health.

Competing interests

The authors have declared that no competing interests exist.

Acknowledgement

The study was financially supported by the National Natural Science Foundation of China (11172016, 11472022). The authors thanked all the participants in this study.

References

[1] Sen S, Asrress K, Petraco R, Nijjer S, Broyd C, Foin N, Foale R, Malik I, Mikhail G, Sethi A, Khan M, Khawaja M, Hughes A, Francis D, Baker C, Mayet J, Di Mario C, Escaned J, Redwood S and Davies J (2012) Does Adenosine Administration Improve Diagnostic Classification Of The Instantaneous Wave-Free Ratio (iFR)? Journal of the American College of Cardiology 60:B69-B69.

[2] Sen S, Escaned J, Malik IS, Mikhail GW, Foale RA, Mila R, Tarkin J, Petraco R, Broyd C, Jabbour R, Sethi A, Baker CS, Bellamy M, Al-Bustami M, Hackett D, Khan M, Lefroy D, Parker KH, Hughes AD, Francis DP, Di Mario C, Mayet J and Davies JE (2012) Development and Validation of a New Adenosine-Independent Index of Stenosis Severity From Coronary Wave-Intensity Analysis Results of the ADVISE (ADenosine Vasodilator Independent Stenosis Evaluation) Study. Journal of the American College of Cardiology 59:1392-1402.

[3] Berry C, van 't Veer M, Witt N, Kala P, Bocek O, Pyxaras SA, McClure JD, Fearon WF, Barbato E, Tonino PAL, De Bruyne B, Pijls NHJ and Oldroyd KG (2013) VERIFY (VERification of Instantaneous Wave-Free Ratio and Fractional Flow Reserve for the Assessment of Coronary Artery Stenosis Severity in EverydaY Practice). Journal of the American College of Cardiology 61:1421-1427. doi: 10.1016/j.jacc.2012.09.065

[4] Davies JE, Whinnett ZI, Francis DP, Manisty CH, Aguado-Sierra J, Willson K, Foale RA, Malik IS, Hughes AD, Parker KH and Mayet J (2006) Evidence of a dominant backward-propagating "suction" wave responsible for diastolic coronary filling in humans, attenuated in left ventricular hypertrophy. Circulation 113:1768-78. doi: 10.1161/CIRCULATIONAHA.105.603050

[5] De Bruyne B, Pijls NH, Kalesan B, Barbato E, Tonino PA, Piroth Z, Jagic N, Mobius-Winkler S, Rioufol G, Witt N, Kala P, MacCarthy P, Engstrom T, Oldroyd KG, Mavromatis K, Manoharan G, Verlee P, Frobert O, Curzen N, Johnson JB, Juni P, Fearon WF and Investigators FT (2012) Fractional flow reserve-guided PCI versus medical therapy in stable coronary disease. N Engl J Med 367:991-1001. doi: 10.1056/NEJMoa1205361

[6] Pijls NH and Sels JW (2012) Functional measurement of coronary stenosis. J Am Coll Cardiol 59:1045-57. doi: 10.1016/j.jacc.2011.09.077

[7] Fischer JJ, Samady H, McPherson JA, Sarembock IJ, Powers ER, Gimple LW and Ragosta M (2002) Comparison between visual assessment and quantitative angiography versus fractional flow reserve for native coronary narrowings of moderate severity. American Journal of Cardiology 90:210-215.
[8] Kim HJ, Vignon-Clementel IE, Coogan JS, Figueroa CA, Jansen KE and Taylor CA (2010) Patient-specific modeling of blood flow and pressure in human coronary arteries. Ann Biomed Eng 38:3195-209. doi: 10.1007/s10439-010-0083-6
[9] Koo BK, Erglis A, Doh JH, Daniels DV, Jegere S, Kim HS, Dunning A, DeFrance T, Lansky A, Leipsic J and Min JK (2011) Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. Results from the prospective multicenter DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) study. J Am Coll Cardiol 58:1989-97. doi: 10.1016/j.jacc.2011.06.066
[10] Min JK, Koo BK, Erglis A, Doh JH, Daniels DV, Jegere S, Kim HS, Dunning AM, Defrance T, Lansky A and Leipsic J (2012) Usefulness of noninvasive fractional flow reserve computed from coronary computed tomographic angiograms for intermediate stenoses confirmed by quantitative coronary angiography. Am J Cardiol 110:971-6. doi: 10.1016/j.amjcard.2012.05.033
[11] Min JK, Leipsic J, Pencina MJ, Berman DS, Koo BK, van Mieghem C, Erglis A, Lin FY, Dunning AM, Apruzzese P, Budoff MJ, Cole JH, Jaffer FA, Leon MB, Malpeso J, Mancini GB, Park SJ, Schwartz RS, Shaw LJ and Mauri L (2012) Diagnostic accuracy of fractional flow reserve from anatomic CT angiography. JAMA 308:1237-45. doi: 10.1001/2012.jama.11274
[12] Norgaard BL, Leipsic J, Gaur S, Seneviratne S, Ko BS, Ito H, Jensen JM, Mauri L, De Bruyne B, Bezerra H, Osawa K, Marwan M, Naber C, Erglis A, Park SJ, Christiansen EH, Kaltoft A, Lassen JF, Botker HE, Achenbach S and Group NXTTS (2014) Diagnostic performance of noninvasive fractional flow reserve derived from coronary computed tomography angiography in suspected coronary artery disease: the NXT trial (Analysis of Coronary Blood Flow Using CT Angiography: Next Steps). J Am Coll Cardiol 63:1145-55. doi: 10.1016/j.jacc.2013.11.043
[13] WANG W, LIU Y, ZHAO X, XIE J and QIAO A (2016) HEMODYNAMICS-BASED LONG-TERM PATENCY OF DIFFERENT SEQUENTIAL GRAFTING: A PATIENT-SPECIFIC MULTI-SCALE STUDY. Journal of Mechanics in Medicine and Biology.
[14] Zhao X, Liu YJ, Li LL, Wang WX, Xie JS and Zhao Z (2016) Hemodynamics of the string phenomenon in the internal thoracic artery grafted to the left anterior descending artery with moderate stenosis. Journal of Biomechanics 49:983-991.
[15] Zhao X, Liu YJ, Ding JL, Ren XC, Bai F, Zhang MZ, Ma LC, Wang WX, Xie JS and Qiao AK (2015) Hemodynamic Effects of the Anastomoses in the Modified Blalock-Taussig Shunt: A Numerical Study Using a 0d/3d Coupling Method. Journal of Mechanics in Medicine and Biology 15.
[16] Ma LC, Liu YJ, Zhao X, Ren XC, Bai F, Ding JL, Zhang MZ, Wang WX, Xie JS, Zhang H, Zhao Z and Han H (2014) Hemodynamic Influence of Different Pulmonary Stenosis Degree in Glenn Procedure: A Numerical Study. Advances in Mechanical Engineering.
[17] Zhao X, Liu YJ and Wang WX (2015) Hemodynamic Based Surgical Decision on Sequential Graft and Y-Type Graft in Coronary Artery Bypass Grafting. Molecular & Cellular Biomechanics 12:49-66.
[18] Zhao X, Liu YJ, Ma LC, Wang WX, Xie JS and Qiao AK (2015) Hemodynamic Comparison between Normal Graft and Y-Type Graft in Coronary Artery Bypass Grafting: A Numerical Study Using 0d/3d Coupling Method. Journal of Mechanics in Medicine and Biology 15.

Pressure variation after stent intervention for a giant aneurysm complicated by a stenosis

W.Y. Fu [1], *A.K Qiao [2]

[1] College of Mechanical and Electrical Engineering, Beijing Union University, China.
[2] College of Life Science and Bioengineering, Beijing University of Technology, China.

*Presenting author: jdtwenyu@buu.edu.cn
†Corresponding author: qak@bjut.edu.cn

Abstract

This paper investigates the pressure variation in a stented aneurysm complicated by a stenosis. Computational Fluid Dynamics analyses were performed on a giant aneurysm with and without a stenosis. Three models were constructed to compare the pressure variation. Model M1 is the one with an aneurysm (no stenosis and no stent), Model M2 is the one with a preaneurysm moderate stenosis (50%), and Model M3 is the one with an aneurysm implanted with a stent. Pressure increase in the aneurismal sac caused by the stenosis is about 10.3 mmHg at peak systole (comparison between M2 and M1). The pressure increase in the aneurismal sac is about 7.8 mmHg at peak systole (comparison between M3 and M2). The geometry of the parent vessel and its aneurysmal/stenotic disease do have influence on pressure variation.

Keywords: Aneurysm, Stenosis, Hemodynamics, Computational Fluid Dynamics

1 INTRODUCTION

Four articles published [1-4] in "American Journal of Neuroradiology" raised a drastic debate about that how much is the pressure increase in a segment of aneurysm after stent intervention treatment for an aneurysm complicated by a proximal stenosis. The study [1] demonstrated a 20mm Hg increase of pressure (the first model) in intra-aneurysmal sac after treatment. However, Fiorella thinks that the pressure increase was overestimated by at least a factor of 2 compared with traditional fluid mechanics calculations and experimental measurements[5,6]. Flow resistance will decrease according to the Poiseuille law when a 50% stenosis of a vessel is removed. But the Poiseuille law can only be used for straight tubes. The geometry of the model used by Cebral [1] is tortuous and it cannot be simply solved by the Poiseuille law. So more work needs to be done to explore this problem.

The purpose of this paper is to perform a numerical simulation study by using Computational Fluid Dynamics (CFD) approach and investigate whether a sharp increase of pressure occurs in an intra-aneurysmal sac because of the existence of stenosis after the stent is inserted to the aneurismal region.

2 METHODOLOGY

For the sake of comparison, the internal carotid aneurysm model (M1) without stent in the reference [7] is used in this study. Based on M1, a segmental vessel with moderate stenosis

* This work was supported by the General Program of Science and Technology Development Project of the Beijing Municipal Education Commission of China (Numerical Investigations of the Mechanical Properties of Braided Stents Used for the Treatment of Cerebral Aneurysm) and the National Natural Science Foundation of China (81171107)

about 50% was artificially constructed. Stenosis locates at the proximal neck of the aneurysm in M1. In this way, the second model (M2) was constructed. The third model (M3) is equivalent to the implantation of a stent into M2 and the stenosis in M2 was expanded. These three models are shown in Fig.1.

Figure.1 Aneurismal models (from the left to the right: M1, M2, M3). Lower end is the inlet, and the upper end is the outlet.

The results obtained from the papers [8] and [9] were used to judge the reliability of the present numerical calculation. So numerical simulation was carried out based on the model used in reference [8,9] firstly. The model is two dimensional in the reference [8,9], while the model in the present study is three dimensional. The diameter at the inlet, the diameter at the outlet and the longitudinal length of arterial vessel are 0.38cm, 0.36cm and 5.2cm respectively.

Validation between their experiment and our CFD calculation was carried out. In this indirect way, the reliability of numerical simulation used in this study was tested. Hemodynamic modeling used in this paper is the same as that in the references [8,9]. Comparison of the pressure drop between the measured values and calculated CFD results was shown in Fig.2. The average values of the pressure drop measured from the experiment and obtained from this study are 0.59 mmHg and 0.69 mmHg respectively in a cardiac cycle. The relative error between the experiment in the references [8,9] and this numerical calculation is less than 15%, which means that the method for the CFD calculation used in this study is reliable. The maximum of the pressure drop, about 5.6 mmHg, occurs at times between 0.08s and 0.09s in the present study. The maximum of this pressure drop, 4.5 mm Hg, occurs at 0.11s in the experiment [8,9]. There is a phase difference about 0.02s between the two studies. The main reason is that the vessel wall is assumed to be rigid in the present study and the wall is elastic in the experiment [8,9].

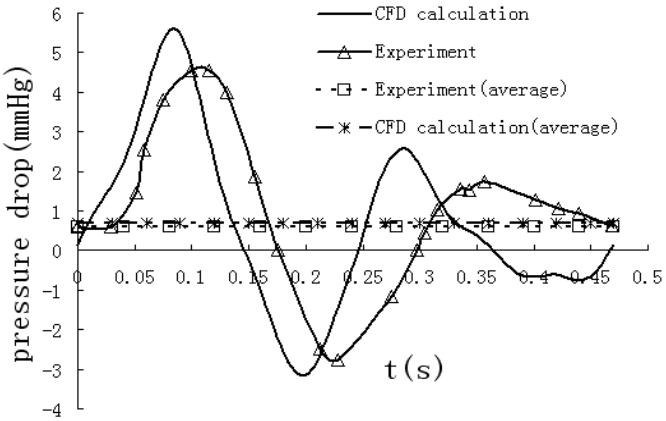


Figure.2 Temporal variation of pressure drop between inlet and outlet in a cardiac cycle

Blood flow is controlled by the three-dimensional incompressible Navier-Stokes equations. The viscosity and density of blood are 0.04 Poise and 1.0 g/cm^3 respectively. Compared with

the large aorta, the elasticity in cerebral arteries is small. Therefore vascular wall was assumed to be rigid. No-slip boundary conditions were applied to the walls. Studies using Newtonian fluid and non-Newtonian fluid on the same cerebral aneurysm models find that the main flow features are not greatly affected by the viscosity model[10], so blood is assumed to be a Newtonian fluid. The maximum of Reynolds number based on the entrance flow velocity and the diameter at the inlet is 704, so the flow is laminar.

As noted in the reference [1,2], the blood flow rate was not measured from the patient from which the model M1 was constructed. So the blood flow rate in the reference [11] was used in the present study. Velocity at the internal carotid artery in different individuals changes greatly. According to the clinical medical statistics data from the literatures [12,13], the velocity range at the internal carotid artery in normal individual is from 20.1 cm/s to 112 cm/s. This range is from 24.1 cm/s to 209.8 cm/s when the aortic stenosis is existed and the degree of stenosis is 10-55%. The maximum velocity is 58.7 cm/s at the inlet in a cardiac cycle in the three models used in the present study and the average velocity is 38 cm/s. Blood pressure was set to 0Pa at the output. The same boundary conditions were used in the 3 models. Based on the consideration of computational accuracy and computational efficiency, the time step of 0.005s was selected for this study. The discrete form of the differential governing equations of blood flow follows the upwind scheme with second order accuracy. ANSYS CFX 12.1 was employed to perform the numerical simulation of Computational Fluid Dynamics (CFD). The solution approach of implicit finite volumes was used. Parallel calculation was used to reduce the running time. The residual convergence criteria of mass and momentum were set to 10^{-5}. The CFD simulations were carried out for 3 cardiac cycles. Results in the third cycle were analyzed.

3 RESULTS AND CONCLUSIONS

As shown in Fig.3, a plane which locates on aneurismal region was assigned to calculate the pressure drop between this plane and the outlet (also between inlet and outlet). Plane positions of inlet, preaneurysm plane and outlet are same for the 3 models. To clearly show the position of this defined plane, streamlines were drawn in Fig.3. The plane of preaneurysm is perpendicular to the centerline of artery model. The average value of pressure on a plane (inlet, plane of preaneurysm, outlet) was calculated at different moments in a cardiac cycle. The average value of pressure on a plane (inlet, plane of preaneurysm, outlet) in a cardiac cycle was calculated too. The results are presented in Table 1.

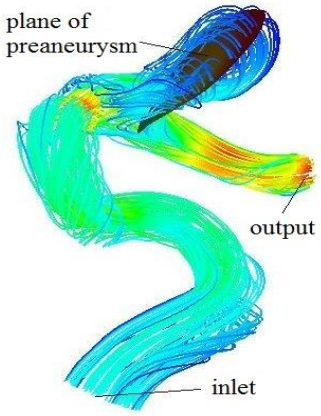


Figure.3 Positions of plane for pressure drop measurement

Table 1: Pressure drop

Model	A[a] (mmHg)	B[b] (mmHg)
M1	5.1	10.9
M2	9.4	21.2
M3	6.4	13.4

a. A indicates "The average of the pressure drop between the preaneurysm and inlet in a cardiac cycle".
b. B indicates "The pressure drop in peak systole between the preaneurysm and inlet".

Comparing M1 with M2, the intra-aneurysmal pressure at peak systole increases about 10.3 mmHg. Pressure drop between the inlet and the preaneurysm plane in a cardiac cycle is calculated for the 3 models and their changes with time are shown in Fig.4. The pressure drop for the M2 is the biggest, this value is moderate for the M3, and it is the minimal for the M1. Because pressure change on aneurismal wall before and after stent implantation is our concern, the difference of the pressure drop (preaneurysm plane and inlet) between M2 and M3 is calculated.

The model used in the references [8,9] is a straight tube with a mild taper. The vessel diameters at the inlet and the outlet are 0.38 cm and 0.36 cm respectively. The axial distance between the inlet and the outlet is 5.2 cm. The maximum velocity at the inlet is 49 cm/s. Pressure drop measured by Banerjee and Back at the cyclic peak is approximately 4.3 mmHg [8,9]. The axial distance along the centerline in the 3 models in this study is also 5.2 cm. It is the same with the model in the references [8,9]. The nominal mean inlet diameter in the models M1 M2 M3 is 0.56 cm. It is bigger than the model used in the references [8,9]. The maximum velocity at the inlet in this study is 46 cm/s. According to the Poiseuilles law for blood flow, pressure difference is related to the length of vascular, the dynamic viscosity of blood, the volumetric flow rate and the radius of the vascular vessel. If the Poiseuilles law could be applied to a patient-specific model, the pressure difference calculated from the model M2 in the present study will not be 21 mmHg. It should be about 3.52 mmHg. This shows that Poiseuilles law cannot be simply applied to a patient-specific model of vascular vessels with geometric characteristics of bent, taper and aneurysm. It also shows that flow resistance increases more than six folds in the model M2 used in the present study compared with the model with a straight tube used in the references [8,9].

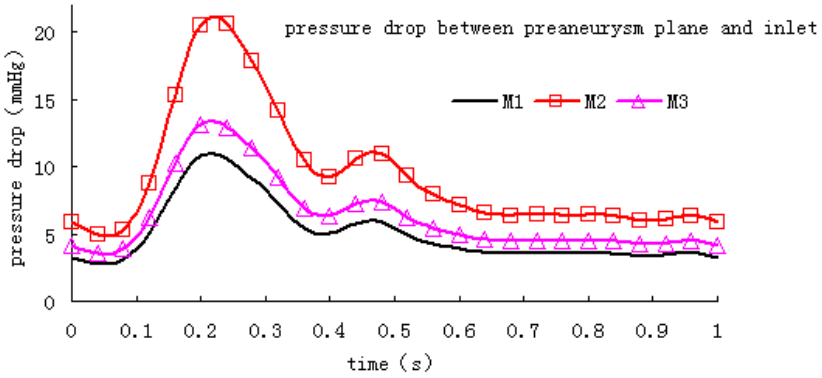


Figure.4 Temporal variation of pressure drop between inlet and preaneurysm plane in a cardiac cycle

The present study demonstrates that a moderate stenosis cannot result in a sharp pressure increase when stent intervention is applied to the treatment of an aneurysm complicated by a 50% stenosis harbored on a tortuous intracranial artery. The geometry of the parent vessel and its aneurysmal/stenotic disease do have influence on pressure variation.

We would like to thank A.G. Radaelli of Universitat Pompeu Fabra, Masahiro Watanabe of Japan Advanced Institute of Science and Technology, Makoto Ohta of Institute of Fluid Science in Tohoku University and organization of Virtual Intracranial Stenting Challenge

2007 for providing us with the patient-specific STL file data. We would also like to thank Mr. Damien Yourrell of University College Dublin for rendering the English of the manuscript.

REFERENCES

[1] Cebral, J.R., Mut, F., Raschi, M., et al. (2011) Aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment, AJNR Am J Neuroradiol 32,27-33.

[2] Fiorella, D., Sadasivan, C., Woo, H.H., et al. (2011) Regarding "aneurysm rupture following treatment with flow-diverting stents: computational hemodynamics analysis of treatment", AJNR Am J Neuroradiol 32, E95-97.

[3] Putman, C.P., Lylyk, P., Cebral, J.R. Reply. (2011) AJNR Am J Neuroradiol 32, E98-100.

[4] Steinman, D.A. (2011) Computational modeling and flow diverters: a teaching moment, AJNR Am J Neuroradiol 32, 977-983.

[5] Young, D.F. Fluid mechanics of arterial stenoses. (1979) J Biomech Eng 101,157-75.

[6] Young, D.F., Cholvin, N.R. (1975) Roth AC. Pressure drop across artificially induced stenoses in the femoral arteries of dogs, Circ Res 36,735-43.

[7] Radaelli, A.G., Augsburger, L., Cebral, J.R., et al. (2008) Reproducibility of haemodynamical simulations in a subject-specific stented aneurysm mode-A report on the Virtual Intracranial Stenting Challenge, Journal of Biomechanics 41, 2069-2081.

[8] Roy, A.S., Back, L.H., Banerjee, R.K. (2008) Evaluation of compliance of arterial vessel using coupled fluid structure interaction analysis, MCB Molecular and Cellular Biomechanics 5, 229-46.

[9] Banerjee, R.K., Back, L.H. (2003) Computed and measured hemodynamics in a compliant tapered femoral artery, In: Proceedings of the 2003 Summer Bioengineering Conference, Key Biscayne, Florida, 261-62.

[10] Kundu, P.K. (2004) Cohen IM. Fluid Mechanics, New York: Academic Press (Elsevier).

[11] Oertel, H. (2008) Prandtl's Essentials of Fluid Mechanics, Beijing: Science Press.

[12] Blackshear, W.M., Phillips, D.J., Chikos, P.M., et al. (1980) Carotid artery velocity patterns in normal and stenotic vessels, Stroke 11, 67-71.

[13] Edward, G.G., Carol, B.B., Gregory, L.M., et al. (2003) Carotid artery stenosis: gray-scale and doppler US diagnosis-society of radiologists in ultrasound consensus conference, Radiology 229, 340-346.

Study of RFD Model Spectrum and the Characteristic Conversion Methods in Industry

J.Y. Li[1] H. Wang[2]
[1,2]Institute of Civil Engineering, Wuhan University, China
*Presenting author: lijy@whu.edu.cn
Corresponding author: 2249633165@qq.com

Abstract: Reversed Flow Diverter (RFD) is a type of maintenance-free device which is applied in pumping and delivering hazardous fluid by using fluidic power. Because of including no moving parts, the operating efficiency and characteristics of RFD are especially sensitive to the structure and operating parameters of the device. This paper presents an optimization design method of RFD based on Genetic Algorithm (GA). System scale and efficiency have been both considered by designing specific fitness function. Based on serialized optimum designs, an example of RFD model spectrum is promoted. Our customers can easily choose a proper type of RFD according to certain flow and head as well as predict the operating of RFD. In the end, the paper discusses the influence of physical characteristics of the delivered fluid. The conversion method of operating flow and efficiency of RFD system is studied.

Key words: RFD; optimization design; operating prediction; model spectrum; genetic algorithm

1 Overview

RFD is a type of fluidic power device designed by the theory of jet. Without moving part, RFD can avoid wear and operate with low maintenance, thus it can be used in occasions where human is inaccessible. RFD is composed of a jet pair, an air-liquid charge vessel and a fluid delivery unit (Fig. 1.). The switching operating jet pair generates compressed air and by the air turns the energy into the pulse liquid with the charge vessel, and then by reversing flow of RFD, the liquid is delivered to outlet. Periodic conveying process is generally divided into two phases, i.e., refill phase and driving phase. Compared with other maintenance-free devices, RFD shows advantages in low waste gas emission, strong adaptability to operating temperature, medium and pressure, et.al.. So it is believed to have a great potential market in the nuclear post-processing industry. [1]-[4][12][13]

Although the structure of RFD is relatively simple, the power of fluidic driven system makes the operating conditions complex and the operating parameters sensitive to the fluidic components. Generally, the industrialization is too expensive for this technology. In China, the serialized design is the key point that influences the application of the technology in nuclear reprocessing facilities and other involved fields. In the late 1980s, AEA almost monopolized this technology and successfully applied it in nuclear post-processing. But due to the complexity and sensitivity of the technology, the key point is still unpublished. What's

more, the industry application depends too much on field tests. Serialized design for industry has not completed yet. As the investment in nuclear industry has been cut down nearly all over the word, the relevant study has been suspended. However, the independent study in our country has been going on constantly since 1990s. It has not been universally applicated for now. The main obstacle comes from the restriction of complicated theory. Therefore, considering the related design theory and equipment test have been done already, serial model spectrum should be provided, besides, a more developed operating prediction program should be served instead of industrial tests. [7]-[10][11][14] In order to achieve industrialization as soon as possible, this paper, based on the study of the GA optimization design, makes serialized design of the RFD system as shown in Fig. 2..

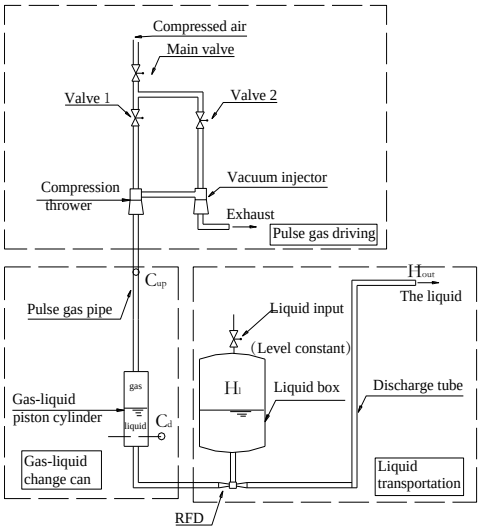


Fig. 1. Schematic construction diagram of RFD set

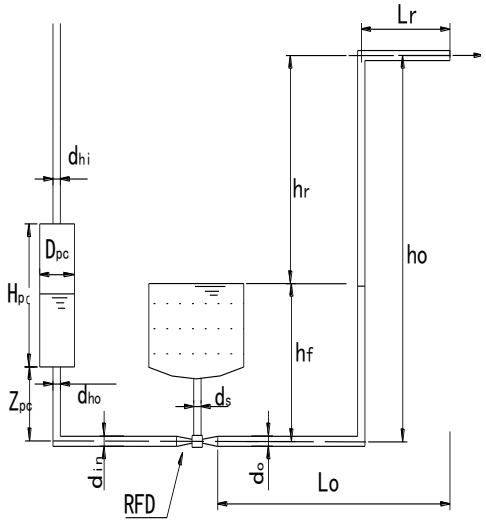


Fig. 2. Calculation sketch of RFD set

2 Theory

2.1 Optimization design model

The optimization of RFD is to achieve the best delivery performance with the smallest set. In a multi-objective problem, customers need to balance their needs. Generally, industrial applications focus on operating and consider the system scale at the same time. Specifically in this paper, to find a valid solution which meets the customer for flow, head, and the limiting conditions, a compound optimization criterion is proposed as:

$$K = N / T^{n} \qquad (1)$$

where N stands for the system efficiency, T for the operating period, a measure of the system scale, n for the scale index ranging from 0 to 1. In the criterion, the smaller the scale index is, the smaller the change rate of the denominator is when T changes, which means the impact of scale is smaller. A larger K indicates that the design can not only satisfy the application goals

(flow, head, size), but also achieve high efficiency. In this way, this multi-objective problem is simplified to a single one with this compound criterion.[5][6]

According to the RFD hydraulic model and the optimization goal, W and M respectively stand for the system operating and scale, numerical simulations and tests experience show that parameters required optimized are as follows: diameter D_{pc} and height H_{pc} of the charge vessel, RFD immersion depth H_f, driving pressure P_i, refill pressure P_r, nozzle diameter D_t and discharge pipe diameter D_o. These parameters are set in domains based on practical industry application, with limits of the parameters matching each other as well. To achieve the highest efficiency and the smallest system while meeting the flow rate, the model is set as:

$$\max W = \max f(D_{pc}, h_{pc}, h_f, P_i, P_r, d_t, d_o)$$

$$\min M = \min f(D_{pc}, h_{pc}, h_f, P_i, P_r, d_t, d_o)$$

$$\text{s. t.} \begin{cases} 0 < D_{pc} <= 2000\text{mm} \\ 0 < h_{pc} <= 5\text{m} \\ 0.2 < h_f <= 5\text{m} \\ 1 < P_i <= 10\text{atm} \\ -1 =< P_r < 0\text{atm} \\ 0 < d_t <= 100\text{mm} \\ d_o >= d_t \end{cases} \quad (2)$$

2.2 Process of genetic algorithm optimization design

In order to use genetic algorithm to solve this multi-factor problem, a fitness function based on RFD optimization model is needed. To optimize those parameters with the given Q, H and n, the GA fitness function is written as follows: (1) set the viscosity N_u, density ρ, coefficients C_p, C_{drf}, C_d, g, k, lengths of inlet and outlet pipelines L_r and L_o, head H, scale index n, average flow rate Q and other known parameters; (2) read a current individual produced by the GA box, i.e., a group of optimized variables: H_f, D_o, D_t, H_{pc}, D_{pc}, P_i, P_r; (3) judge if the variables match the parameters limits, if not, assign 1000 to fitness and then end the program; (4) calculate the operating parameters and judge if they satisfy the pulse period, head, et.al., if not, assign 900 to fitness and then end the program; (5) calculate flow Q_i and Q_o of the inlet and outlet nozzles in the driving phase, the average flow Q_{ra}, and hydraulic efficiency N_h over a period, assign 800 to fitness; (6) judge if Q_{ra}, Q_o/Q_{ra}, H_{pc}/D_{pc}, period T, delivery line loss, and backflow ratio are kept in the predefined limits, and set fitness according to the following principles: 1) if the average flow Q_{ra} is between $Q \pm 0.1\text{m}^3/\text{h}$, set fitness value as 700; 2) if term 1) is met and Q_o/Q_{ra} is between 0.9~1.1, set fitness as 600; 3) if term 2) is met and delivery line loss is lower than 15% of H, set fitness as 500; 4) if term 3) is met and H_{pc}/D_{pc} is

between 2.5 to 3.5, set fitness as 400, 5) if term 4) is met and *T* is less than 1000 seconds, set fitness as 300; 6) if term5) is met and backflow ratio is greater than 8, the solution is considered to be valid, then calculate the compound optimization criterion *K* and assign it to the fitness.

When fitness values of the current generation are calculated, selection, crossover and mutation operations are used to produce the next generation. Then fitness of the new generation is calculated. That cycle repeats and finally a proper group of optimized variables is obtained.

Some information should be input before running the GA program, which includes crossover probability, mutation probability, population size, length of chromosome, the generation gap, the maximum generation, et.al.. Those parameters affect the operating and even influence the convergence of the algorithm. In this paper, gauss mutation is adopt; population size, which directly affects the calculation efficiency and convergence of GA is generally between 10 to 200, in this paper, it is set 200 based on tests; and the maximum generation, acting as a termination condition of algorithm, is set as 2000.Because of the way of random search, GA shows great advantage in CPU comsumption. It takes only about 100 seconds to calculate one case on the current mainstream computer, which makes the serialized design for industrial production possible. However, if one certain case is calculated several times, there will be some small deviation between the results due to the randomness. But compared with basic GA, by debugging control parameters, stability and convergence can be satisfied in industry, also, it seldom results in an local extremum. As shown in Fig. 3., the point near global extreme point appeared at 600 steps, and points got very close to the true value after 600 steps when the algorithm performance was greatly improved.

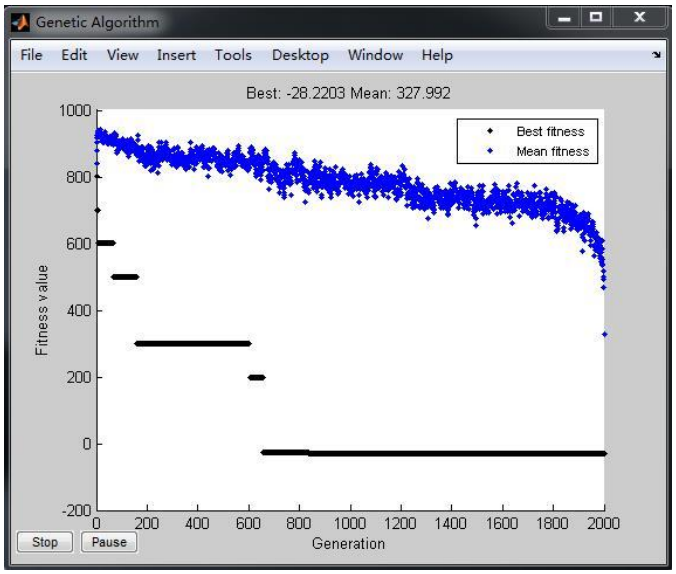


Fig. 3. The best and average fitness values of each generation in GA

By this program we can design for the various parameters and find out the matching of

structure and operating parameters, as well as the effects of parameters on the operating. The model spectrum based on serialized design can also be figured out to provide basic parameters for RFD model design.

3 Results

3.1 The optimization design results

In the calculation the resistance coefficients involved include pressure recovery coefficient of diffusion tube, nozzle suction-flow coefficient, total resistance coefficient of delivery pipe, and nozzle flow coefficient. The scale index n is set 0.1, for a relatively large system scale. In the paper, 68 cases are calculated with the head ranging from 3 to 30 meters and average flow from 1 to 20 m^3/h. Through single factor experiments about nozzle diameter D_t, delivery pipe diameter D_o, charge vessel diameter, and height and driving pressure P_i, the following qualitative conclusions are reached.

Firstly, when the head is fixed and the required flow increases, P_i hardly change, but RFD nozzle diameter D_t, delivery pipe diameter D_o, charge vessel diameter D_{pc} and height H_{pc}, and other parameters related with scale will increase. Otherwise, it is bound to result in an increase in hydraulic loss and a decrease in efficiency of RFD, even lead to the cavitation and a system failure. When the flow is fixed and the head increases, driving pressure P_i increases; while RFD nozzle diameter D_t and delivery pipe diameter D_o are crucial parameters of the delivery capacity, i.e.,flow rate, nearly remain unchanged. That means, the way RFD structure parameters impact the operating is that, cross-section diameters (such as: D_t, D_o, D_{pc}) decide the flow rate, driving pressure *Pi* decides the head. Since the charge vessel is similar to a gas-liquid piston pump, it is reasonable that it is close to conventional piston pump. In addition, within the limits of parameters and all the cases calculated, the refill pressure holds steady at - 5 mH_2O.

Secondly, when the flow and head are fixed and the scale index n increases, scale parameters decrease and system efficiency N also decreases. Customers can choose an appropriate scale index to fit their equipment room.

Summarizing the optimization results of 68 cases we get regression equations of the above relationships, shown in table 1 as empirical formulas, which can be referred when choosing an optimal model.

Table 1 Empirical formulas for RFD optimization design

The known average flow rate Q（m3/h），H（m）	**Applicable condition**
delivery pipe diameter: Do=37.932Q0.4151 (mm)	Pr=-5m
nozzle diameter: Dt=9.668Q0.4966 (mm)	Q:1~20m3/h
charge vessel height Hpc: Hpc=0.5584H0.3Q0.27 (m)	H:3~30m（n=0.1）
charge vessel diameter Dpc: Dpc=Hpc/2.5 (m)	
driving pressure: Pi=0.1463H+0.0175(atm)	

3.2 Model Spectrum design

High efficiency range is important for all kinds of pumps including RFD. To satisfy application, it is necessary to study the high efficiency range and the performance curve of RFD. For RFD pump, head *H* is both operating parameter and structure parameter, so if *H* is set as a GA optimized variable and the other parameters exclude *Q* are all fixed, the algorithm will select out a *H* - *Q* match of high efficiency and then a high efficiency range can be worked out soon. In this paper, 25 RFD models are set corresponding to a certain *H*-*Q* range, and their high efficiency ranges are calculated out and gathered up to draw the RFD model spectrum as shown in Fig. 4..

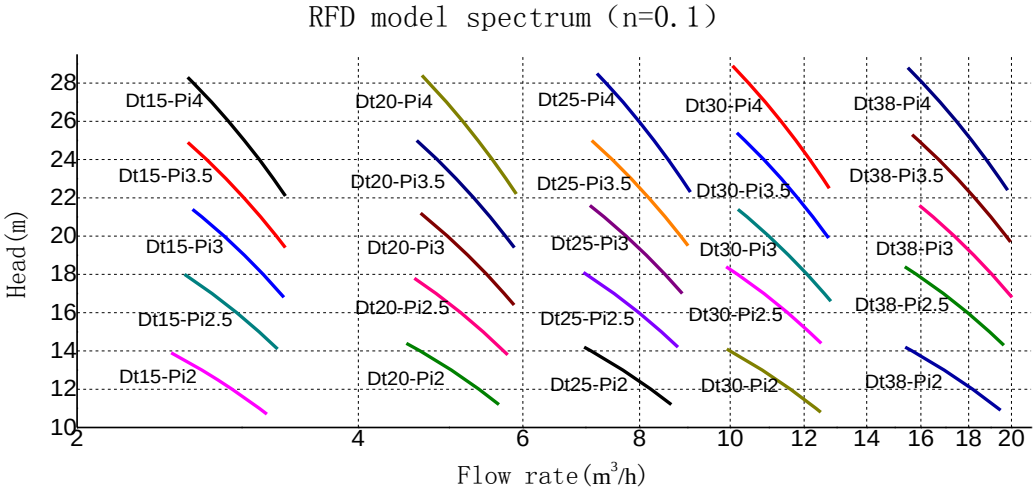


Fig. 4. RFD model spectrum

In application, customers can choose a suitable model shown in Fig. 4. according to flow and head, with supporting structure and operation parameters.

3.3 Performance prediction curve

For a certain RFD model, the paper also studied flow and efficiency change with head in normal operating condition. According to calculation results, the curve *Q* vs *H*, *N* vs *H* are shown as below: (model D_t20 - P_i3 for example)

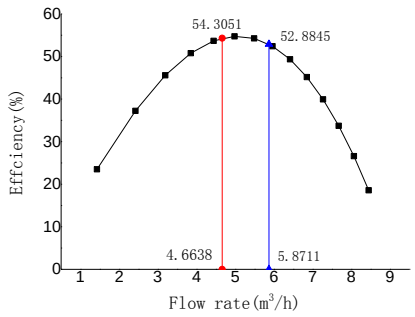


Fig. 5. RFD head H vs flow Q

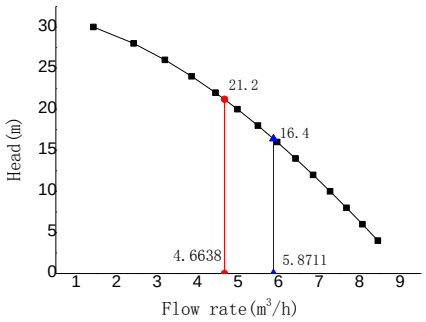


Fig. 6. RFD efficiency N vs flow Q

The *H-Q* curve shows that RFD performance curve is a quadratic parabolic, similar to the centrifugal pump. In Fig. 6., *N-Q* curve between the two vertical is the high efficiency range shown in RFD model spectrum, where not only the efficiency reaches the highest, but also the optimal control conditions in fitness function are satisfied. When *Q* is too low and *H* is too high, injection will inevitably occur, when *Q* is too high and *H* is too low, line loss will increase and lead to the decrease of the system efficiency.

3.4 Performance conversion methods for medium characteristics

When studying prediction curve of performance of a certain RFD model, different kinds of liquid should influence the operating (flow and efficiency), because they have different viscosity and density compared with water. Set model D_t20-P_i3 as an example, comprehensive diagrams of density and viscosity vs flow and efficiency are shown as Fig. 7. and Fig. 8.. The figures show that when changing the kinematic viscosity within the scope of a relatively low level (1e-6~1e-4 m^2/s), as the density and viscosity increases, the efficiency firstly increases and then decreases, which means there is an optimal combination range for density and viscosity where the device reaches the highest efficiency. However, the efficiency change with viscosity is slight overall. When the efficiency reaches the maximum, the viscosity change nearly makes no difference. The flow decrease with viscosity increasing is also small, but with the increase of density, the average flow with a fixed head shows an obvious trend of decline.

1) As the density and viscosity increase at the same time, the hydraulic efficiency shows a decrease after an increase, which means there is a maximum efficiency when flow changes. In the efficiency decreasing phase, delivering resistance (such as energy dissipation in the form of friction) increases as the liquid viscosity increases.

2) There are economic velocities at all density levels, where hydraulic efficiency reaches highest and the viscosity affects little.

3) As the density increases, the flow decreases, and viscosity increases, the flow shows a decreasing trend which is not significant.

4) In industry, attention need to be paid to the decreasing of flow and efficiency when delivering fluid of high viscosity.

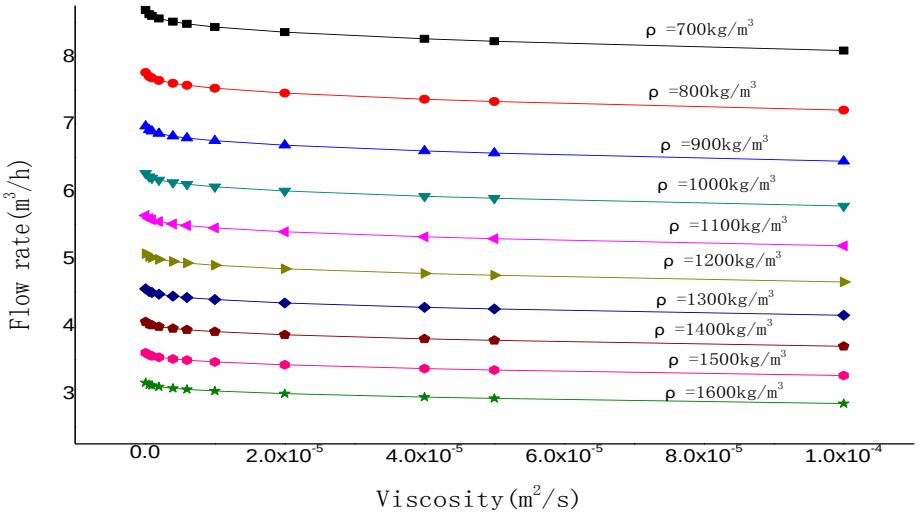


Fig. 7. Effects of Density and viscosity on flow rate

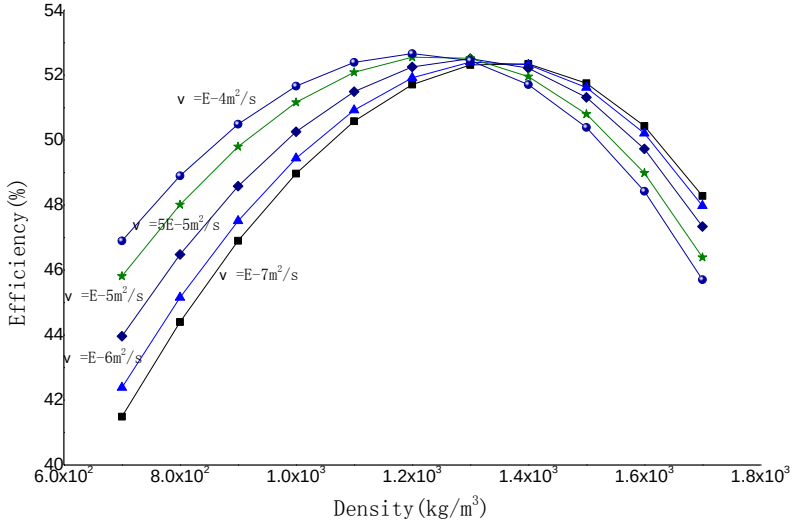


Fig. 8. Effects of Density and Viscosity on efficiency

3.5 Optimization design with medium characteristics considered

In industry, particular designs for RFD devices applied in special delivered mediums are needed. The paper studied the change of structure and operation parameters of the optimized RFD model with different liquid and the same flow and head. For example, setting the flow as 5 m^3/h, head as 15 m, scale index *n* as 0.5, and with changing the density and viscosity based on water, the optimized parameters of RFD device will change.

Analyze the results and conclusions are as follows:

1) While setting the flow constant and increasing liquid density, driving pressure P_i need to increase to maintain the head, otherwise the needed pressure or head can not be reached and it will be unable to meet the flow rate.

2) As to RFD model design, scale parameters show an increasing trend as the density and viscosity increase. When flow keeps constant, nozzle diameter D_t basically remains unchanged.

3) When the delivered water is 30~60℃ in the experiments, it is found that the viscosity changes slightly along with temperature changes, and flow rates remains the same, which has been confirmed in RFD device test.

4 Conclusions

In this paper, an optimization design program of RFD based on GA have been developed, based on single factor tests for structure and operation parameters, the influence laws of parameters on the operating have been found out, and according to a typical range of flow and head in industry, 25 RFD models have been designed to establish a model spectrum with a form of the device parameters. The prediction of performance has also been given out. And the influence of viscosity and density of the delivered liquid, the conversion rules for performance, and the law of liquid characteristics on the optimization design have also been studied.

References:

［1］ J. G. Morgan, W. D. Holland, Pulsatile Fluidic Pump Demonstration and Predictive Model Application[J], ORNL/TM-9913, 1989.4

［2］ D.W.Crass.An Evaluation of Power Fluidics TM mixing and pumping for application in the singel shell tank retrieval program[J].2001.3

［3］ G.V.Smith,B.E.Lewis. Design of a Pulsed-Mode Fluidic Pump Using a Venturi-Like Reverse Flow Diverter[J].ORNL/TM-10281,1987.2

［4］ G. H. Priestman, J. R. Tippetts. Low pressure high head fluidic pump[J]. Source: Flucome, 1991,91:169-174.

［5］ LI Jiangyun, CHEN Youlin, ZHOU Qiyuan, etal. Research On Design Theory and Method of RFD set: Engineering Fluid Mechanical Thermal Physical Society conference in 2009, Dalian, 2009[C].

［6］ LI Jiangyun, CHEN Youlin, LI Yangyu. Research On Marginal utility optimization design theory and Application: Engineering Fluid Mechanical Thermal Physical Society conference in 2010, Nanjing, 2010[C].

［7］ Tippetts, J.R. Priestman, G.H. Power fluidics in the chemical process industry, Institution of Chemical Engineers Symposium Series, n 94, p 1-10, 1985

［8］ P. Fallows, M. Williams, P. Murray, Applications of power fluidics technology in nuclear waste processing plants, WM’5 Conference, February 27-March 3, 2005, Tucson, AZ

［9］ R. Rajeev, C.V. Joyakin and S.B. Koganti，Fluidic pump development studies，Indira Gandhi Centre for Atomic Research，2002

［10］ Priestman, G.H. Tippetts, J.R. Characteristics of a Double-Acting Fluidic pump with hot and cold water, J. Fluid Control 1986, 16(4):19-39

［11］ M. R. Sharon, R. M. Counce, G. V. Smith, Performance Characteristics of a Continuous Flow Fluidic Pump, Ind. Eng. Chem. Res.,1987,26(1698-1704)

［12］ J. R. Tippetts, Definition and properties of a Eulerian 3-terminal gyrator, proceedings of the royal society. A (2005) 461,957-974

［13］ Smith, G.V. Counce, R.M. Performance characteristics of axisymmetric venturi-like reverse-flow diverters, (Univ of Tennessee, USA, Univ of Tennessee, USA); Source: Journal of fluid control, v 16, n 4, p 19-39, 1986

［14］ Saito, Susumu, Sugiyama, Takashi; Sakai, Toshimichi,Development of a fluidic pump driven by single bistable element(Science Univ of Tokyo, Tokyo, Japan); Source: Flucome '91, p 181-186, 1991

Boundary and current elements for simulation of electromagnetic fields of complicated spatial configuration

***†A.A. Trubitsyn, E. Yu. Grachev and V.S. Gurov**

Department of Electronics, Ryazan State Radio Engineering University, Russia

*Presenting and †corresponding author: assur@bk.ru

Abstract

Processes of interaction of charged particle fluxes with substance are a base of operation of a wide range of devices for scientific researches and vacuum technological installations for various purposes. Besides, an important independent scientific and technical task is a task to control parameters of such fluxes by influencing of electrical and/or magnetic fields of a specified configuration on them.

The paper shows a mathematical instrument and algorithms for simulation of the electrical field in electron-optical systems with complicated configuration of electrodes by the boundary element method (BEM). The boundary element method solves an exterior Dirichlet's problem under digitization of a multiplied connected boundary of the area by straight boundary elements. Integral equation being a basic for the method is found from the second Green's formula.

Biot-Savart-Laplace law is used for numerical simulation of the magnetic field for a set of arbitrarily oriented round solenoids. Besides, each turn of the solenoid is divided into current elements having a physically short extent. Calculation of the magnetic field of a two-dimensional shape (turn) in space is reduced to calculation of coordinates of elementary current middles and their projections on the axes 0*x* and 0*y*. Magnetic field of a solenoid is calculated as a superposition of fields of all its turns. Magnetic field of the arbitrarily oriented solenoid can be found by means of direct and reverse rotations of the Cartesian coordinate system for angles being equal to angles setting an orientation of solenoids in space. Magnetic field of a set of arbitrarily oriented solenoids is calculated according to the superposition principle.

Developed methods have been integrated into the computer application and used under design of the microfocus x-ray tube of high power. Acceleration and focusing of the electron beam in the tube is executed by the electrostatic field and its positioning to the required area on the anode – by the magnetic field of two external solenoids.

Keywords: boundary elements, boundary elements method, electron optics system, Biot-Savart-Laplace law, current element, solenoid, numerical modelling, microfocus x-ray tube

Introduction

Tasks of focusing and transportation of charged particle fluxes are main problems under creation of highly qualitative devices for electron and ion optics. Electric and magnetic fields and their combinations are used for solution of such tasks. Mathematical simulation is one of important stages for development of technical means forming electromagnetic fields with required properties. Numerical methods are the most general techniques for simulation of electromagnetic fields as distinct from analytical ones.

Electrostatic field simulation

The problem of numerical analysis of electric fields in systems with a complex configuration of electrodes is brought to the forefront in electron optics which is the basis for analytical instrument making [1] [2]. At present Boundary Elements Method is one of the most advanced numerical techniques to solve problems of the potential theory [3]. The solution of the exterior Dirichlet problem in electron optics, in contrast to the interior one, allows us to predict the parameters of the schemes as close to the real device.

In this case boundary Γ of the researched area G is represented by a combination of closed contours (electrodes of real thickness and configuration) at each of which potential is fixed and integral equation connecting potential $u(\xi)$ in the researched area and its boundary with a normal potential derivative q at the boundary is recorded. The obtained integral relations are based on the second Green's identity [4] that allows simulating fields in areas which boundaries have corners and fractures.

A wide class of electron-optical systems (EOS) can be described within the framework of axi-symmetrical models at the design stage with a high degree of confidence.

For numerical solution of the problem the integral equation has a discrete form. For this purpose boundary Γ is divided into N boundary elements Γ_j. Taking into consideration potential constancy at each contour (electrode) and under assumption of the normal potential derivative constancy at each boundary element, the equation is recorded in the form

$$\gamma u(\xi)+\sum_{j=1}^{N} u_j H_j(\xi)=\sum_{j=1}^{N} q_j F_j(\xi), \qquad (1)$$

where $\xi \in G \cup \Gamma$; $\gamma=\pi$ under $\xi\in\Gamma$ and $\gamma=2\pi$ under $\xi\in$G for two-dimensional tasks, $\gamma=2\pi$ under $\xi\in\Gamma$ and $\gamma=4\pi$ under $\xi\in$G for three-dimensional tasks, and functions $F_j(\xi)$ and $H_j(\xi)$ are integrals from the fundamental solution and from a normal derivative of the fundamental solution of the Laplace equation correspondingly [5] and in a regular case they can be calculated according to the standard Gaussian quadrature.

Calculation of the electrostatic field is executed by two stages. Firstly, by means of equation (1) an unknown vector q_j is calculated according to a known boundary distribution of the potential u ($\xi\in\Gamma$), i.e. "inverse" problem is solved. Then found values q_j and specified u_j for determination of the function $u(\xi)$, $\xi\in\Omega$ from equation (1) are used, i.e. "direct" problem is solved.

Copyright technique [6] is briefly described below for elimination and weakening of peculiarities in sub-integral functions (integrands) under solution of inverse and direct problems providing high accuracy of the task solution in general.

Inverse problem

Collocation method is used for solution of the inverse problem according to which points ξ_i are determined in the middle of each *straight* (then $\gamma(\xi)=2\pi$) element Γ_i and for the whole N- aggregate of points ξ_I a system of N equations is recorded

$$2\pi u(\xi_i)+\sum_{j=1}^{N} u_j H_{ij}=\sum_{j=1}^{N} q_j F_{ij}\ ,\ i=1, 2, \ldots, N, \qquad (2)$$

where $H_{ij}=H_j(\xi_i)$,$F_{ij}=F_j(\xi_i)$.

Integrals with singular integrands exist under $i=j$, i.e. in the case when integration is executed by element Γ_j (let's call it as singular) containing a current collocation point ξ_i.

Calculation of H_{ii}. In consequence of potential jump of a double layer under crossing of the area boundary inside-out [7] we will have the following value of integral H_{ii} for the exterior task:

$$H_{ii}= 4\pi .$$

Calculation of F_{ii}. In this case for elimination of peculiarities, polynomial representation of the complete elliptical integral of I-type is used

$$K(m)=B(m_1)-A(m_1)\ lnm_1, \text{ where } m_1=1-m,\ B(m_1)= \sum_{n=0}^{N_K} p_n m_1^n ,\ A(m_1)= \sum_{n=0}^{N_K} s_n m_1^n ,$$

p_n, s_n - tabulated coefficients of the polynomial representation [8] which ensures enough high speed of convergence to an exact value and allows obtaining the following estimation of integral F_{ii}

$$F_{ii} = 4\Delta\chi_i \left\{ \int_0^1 \left[G_0(\xi_i, x) + \ln 2\left(G_i(-x) + G_i(x) \right) \right] dx + \int_0^1 \left(G_i(-x) + G_i(x) \right) \ln(1/x)\, dx \right\}, \quad (3)$$

where $G_0\left(\xi_i, x\right) = \left[\left(B(m_1) + A(m_1) \ln \frac{a^* + b^*}{\Delta\chi_i^{\,2}} \right) R(x) / \sqrt{a^* + b^*} \right]$,

$a^*+b^*=[R(\xi)+R(x)]^2+[Z(\xi)-Z(x)]^2$, $R(x)=c_ix+d_i$, $Z(x)=a_ix+b_i$,

$$B(m_1) = \sum_{n=0}^{N_k} p_n m_1^n = \sum_{n=0}^{N_k} p_n \left[\frac{\left(\xi_i - x\right)^2 \Delta\chi_i^2}{a^* + b^*} \right]^n ,\ A(m_1) = \sum_{n=0}^{N_k} s_n m_1^n = \sum_{n=0}^{N_k} s_n \left[\frac{\left(\xi_i - x\right)^2 \Delta\chi_i^2}{a^* + b^*} \right]^n ,$$

$$G_i(x) = \left\{ \left((1+x)c_i + 2d_i \right) \frac{\sum_{n=0}^{N_k} s_n \left(\frac{(x\Delta\chi_i)^2}{\left[c_i(2+x) + 4d_i \right]^2 + (xa_i)^2} \right)^n}{\sqrt{\left[c_i(2+x) + 4d_i \right]^2 + (xa_i)^2}} \right\},$$

$\Delta\chi_j$ - a length of straight boundary element Γ_j, $c_i=\sin(\phi)\Delta\chi_i$, $a_i=\cos(\phi)\Delta\chi_i$, b_i and d_i are z- and r-coordinates of the element Γ_i beginning in the cylindrical coordinate system $Z0R$, ϕ - an angle of its slope to the axis z.

The first integral in formula (3) can be calculated by means of the usual Gaussian quadrature and there are special quadratures [8] allowing making calculations with required accuracy for numerical integration of functions of type $f(x)ln(1/x)$. Multiplicative approach has methodologically been realized here for separation of peculiarities.

So, solution q_j of the inverse problem according to (2) can be obtained from a system of linear equations

$$b_i=F_{ij}q_j,$$

where $b_i= \sum_{j=1}^{N} H_{ij}^* u_j$, $i, j = 1, 2, \ldots, N$; $H_{ij}^* = H_{ij}$ for $i \neq j$; $H_{ii}^* = 2\pi + H_{ii} = 6\pi$; F_{ij} is calculated according to formula (3) for $i=j$.

Direct problem

Obvious formula is used for solution of the direct problem (ref. equation (1))

$$u(\xi)=\frac{1}{4\pi}\left[\sum_{j=1}^{N}u_j H_j(\xi)-\sum_{j=1}^{N}q_j F_j(\xi)\right],\ \xi\in\Omega.$$

It is clear that the closer point ξ is to the boundary element Γ_j,, the stronger extremum expression is in behavior of integrands (discontinuity occurs at the limit) that complicates direct application of the standard Gaussian quadrature.

Formulas improving accuracy of estimation of (quasi-) singular integrals $H_j(\xi)$ and $F_j(\xi)$ are mentioned below.

For accurate estimation of integral $H_j(\xi)$ the following formula is used

$$H_j(\xi)=H_j^{*}(\xi)-P(\xi)/(r_j)^k,$$

where an unknown parameter $P(\xi)$ is determined by expression

$$P(\xi)=\frac{\sum_{j=1}^{N}H_j^{*}}{\sum_{j=1}^{N}(r_j)^{-k}}.$$

Here $H_j^{*}(\xi)$ are integrals calculated according to the standard Gaussian quadrature, r_j is a distance between ξ and a point being the closest to it in the segment Γ_j, k is an index of the degree which empirically found optimal value is equal to 4.

Formula for calculation of $F_j(\xi)$:

$$F_j(\xi)=4\Delta\chi_j\left\{\begin{array}{l}\int_0^1\left[\frac{K(m)}{\sqrt{a^*+b^*}}R(x)+g(D/T)\ln(Tx^2-2Dx+P)\right]dx-\\ -g(D/T)\left[\begin{array}{l}(1-D/T)\ln(T-2D+P)+D\ln(P)/T-2+\\ +2\frac{\sqrt{TP-D^2}}{T}\left(arctg\frac{D}{\sqrt{TP-D^2}}+arctg\frac{T-D}{\sqrt{TP-D^2}}\right)\end{array}\right]\end{array}\right\}.$$

Here $g(x)=A(m_1)R(x)/\sqrt{a^*+b^*}$, $T=\Delta\chi_j$, $D=[R(\xi)-d_j]c_j+[Z(\xi)-b_j]a_j$, $P=[R(\xi)-d_j]^2+[Z(\xi)-b_j]^2$.

Additive approach for separation of peculiarities has methodologically been realized here.

Calculation error estimate

Testing of suggested methods for solution of the exterior Dirichlet's problem on model axially symmetrical tasks has allowed making the following conclusions:
- guaranteed accuracy of the potential calculation is ~ 10^{-4}-10^{-3} %;
- for electron-optical systems (EOS) with straight sections of electrodes, calculation errors in the range are limited only by round-off errors.
- for EOS with curvilinear electrodes, calculation errors are determined by an accuracy of approximation of the boundary by linear segments and do not exceed ~ 10^{-3} % under calculation in real time.

Figure 1, as an example, shows a result of simulation of the electrostatic field in the microfocus x-ray tube containing a cathode assembly which elements are under potential -80

kV, focusing electrode with potential -78 kV and grounded thick anode with a narrow channel of the special form intended for increasing of power of x-ray radiation. Distribution of the potential is encoded by shades of grey color.

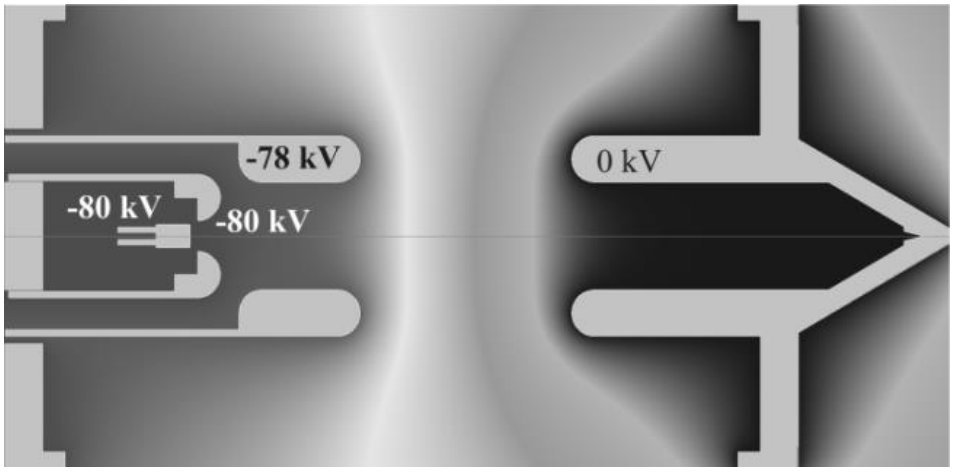


Figure 1. Distribution of an electrical potential in the x-ray tube meridional section

Magnetic field simulation

Known method for the magnetic field excitation in space is based on transmission of the electrical current through conductors. Biot-Savart-Laplace law [9] is one of main laws of electromagnetism and sets a value of the magnetic field induction $d\vec{B}$ created in space by a current element $Id\vec{l}$ according to formula

$$d\vec{B} = \frac{Id\vec{l}\times\vec{r}}{r^3(x_0,y_0,z_0,x,y,z)},$$

where I is a value of current in the element $d\vec{l}(x_0, y_0, z_0)$ of the contour (l), $\vec{r}$ is a distance between element $d\vec{l}$ and point of observation $P(x,y,z)$ (Fig. 2). Here $I=I_{SI}\cdot 10^{-7}$, where I_{SI} – current measured in amperes. The superposition principle allows calculating a magnetic field at any point of space $P(x,y,z)$ by integrating according to contour (l):

$$\vec{B} = \int_{(l)} d\vec{B} = \int_{(l)} \frac{Id\vec{l}\times\vec{r}}{r^3(x_0,y_0,z_0,x,y,z)}. \quad (4)$$

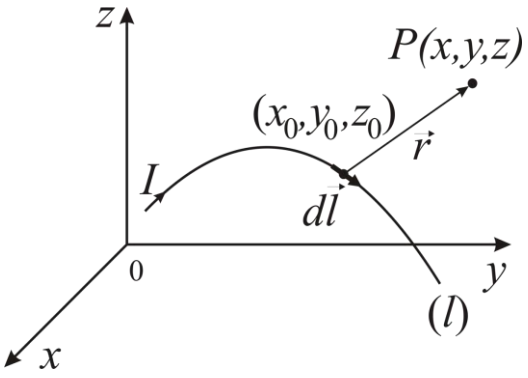


Figure 2. Calculation of the magnetic field at point *P*

One of main sources of the magnetic field in the absence of ferromagnetic is a solenoid as a combination of simple circular currents. However, direct application of the Biot-Savart-Laplace law even for calculation of the field of a circular turn (at any point of space) is complicated by problems with analytical integration in formula (4).

Then a simple method of numerical solution of the task to calculate a field of the circular turn is mentioned and obtained results are generalized for technique of determination of an arbitrary solenoid field.

Field of circular current

Idea of the suggested method is the following. Flat conductor located in the plane $x0y$ is divided into N similar segments of length Δl (Fig. 3). Distance r is considered as constant under integration by one element Δl_i and equal to a distance between a midpoint of the segment with coordinates (x_i,y_i,z_i) and a point P.

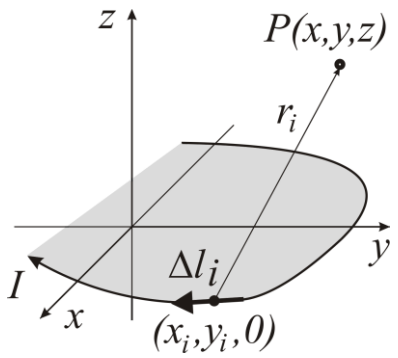


Figure 3. Numerical determination of the flat current magnetic field

In such approximation integral (4) is recorded in the following way:

$$\vec{B} \approx I\sum_{i=1}^{N}\int_{(\Delta l_i)}\frac{d\vec{l}\times\vec{r}_i}{r_i^3} \approx I\sum_{i=1}^{N}\frac{1}{r_i^3}\int_{(\Delta l_i)} d\vec{l}\times\vec{r}_i. \qquad (5)$$

Then we take into account that $d\vec{l} = d\vec{l}(dx, dy, 0)$ and $\vec{r}_i = \vec{r}_i(x - x_i, y - y_i, z)$, and also a vector product is recorded though determinant

$$d\vec{l}\times\vec{r}_i = \begin{vmatrix} \vec{\imath} & \vec{\jmath} & \vec{k} \\ dx & dy & 0 \\ x-x_i & y-y_i & z \end{vmatrix} =$$

$$= zdy\vec{\imath} - zdx\vec{\jmath} + [(y - y_i)dx - (x - x_i)dy]\vec{k}.$$

Taking this record into consideration integral over the segment Δl_i is easily expressed by

$$\int_{(\Delta l_i)} d\vec{l}\times\vec{r}_i = z\Delta y_i\vec{\imath} - z\Delta x_i\vec{\jmath} + [(y - y_i)\Delta x_i - (x - x_i)\Delta y_i]\vec{k}. \qquad 6)$$

Since in the coordinate form a vector of the magnetic induction is recorded in the following way $\vec{B} = B_x\vec{\imath} + B_y\vec{\jmath} + B_z\vec{k}$, then on the basis of formulas (5) and (6) we obtain an expression for components of the vector $\vec{B}$ along coordinate axes

$$B_x = I\sum_{i=1}^{N}\frac{z\Delta y_i}{r_i^3},\ B_y = -I\sum_{i=1}^{N}\frac{z\Delta x_i}{r_i^3},\ B_z = I\sum_{i=1}^{N}\frac{(y-y_i)\Delta x_i-(x-x_i)\Delta y_i}{r_i^3}, \qquad (7)$$

where distance r_i has an obvious expression

$$r_i = \sqrt{(x - x_i)^2 + (y - y_i)^2 + z^2}.$$

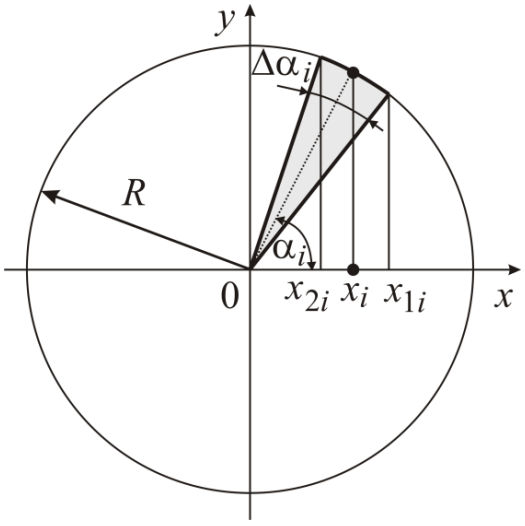


Figure 4. Division of the circular contour into current elements

So, calculation of the magnetic field created by the flat contour with current is reduced to calculation of coordinates x_i, y_i of element middles Δl_i, $i=1\ldots N$, and projections Δx_i, Δy_i of these elements along axes $0x$ and $0y$.

For a circular turn with current located perpendicularly to the axis $0z$ with a center at the beginning of coordinates, enough simple algorithm for calculation of these parameters consists in division of the circumference (Fig. 4) into the same arcs with a small angular size $\Delta\alpha$ and determination of an angular coordinate of the element center

$$\alpha_i = (i - 1)\Delta\alpha + \frac{\Delta\alpha}{2},\ i=1, 2, \ldots, N,$$

determination of coordinates of its center

$$x_i = R\cos\alpha_i,\ y_i = R\sin\alpha_i \tag{8}$$

and values of projections of the element along axes 0*x* and 0*y*

$$\Delta x_i = x_{2i} - x_{1i},\ \Delta y_i = y_{2i} - y_{1i} \tag{9}$$

where $x_{1i} = R\cos[(i-1)\Delta\alpha]$, $x_{2i} = R\cos(i\Delta\alpha)$, $y_{1i} = R\sin[(i-1)\Delta\alpha]$, $y_{2i} = R\sin(i\Delta\alpha)$.

Solenoid field

Suggested technique for calculation of the turn field easily spreads to a problem of calculation of the solenoid magnetic field containing *K* layers and *J* turns in each layer (Fig. 5).

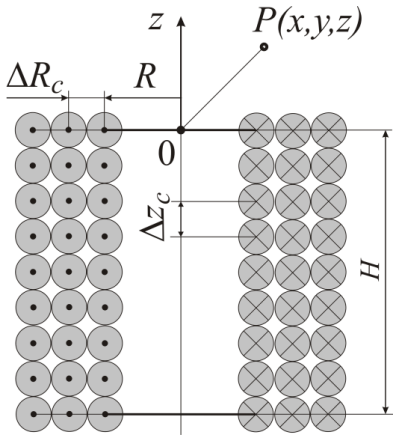


Figure 5. Solenoid cross-section

Solenoid field which upper base is located in the plane *x*0*y* and axis coincides with coordinate axis 0*z* can be calculated as a double sum by a number of layers and turns in each layer

$$\vec{B}_S(x,y,z) = \sum_{k=1}^{K}\sum_{j=1}^{J}\vec{B}_{kj}\left(x,y,z_j(z)\right), \tag{10}$$

where components of the magnetic field $\vec{B}_{kj}$ of each turn along axes are determined according to formula (7) taking into account a value of the turn radius of κ-layer $R_k=R+(k-1)\Delta R_c$ for usage in formulas (8, 9), $k=1\ldots K$, and $z_j=z+(j-1)\Delta z_c$, $j=1\ldots J$. Here ΔR_c is a difference of radiuses of turns and Δz_c is a distance between adjacent turns along axis 0*z*.

Calculation error estimate

Analytical expression for the field of the circular current on the symmetry axis

$$B_0 = \frac{I}{2}\frac{R^2}{(R^2+z^2)^{3/2}},$$

where *z* – a distance from the turn center to a calculated point, allows testing the suggested technique for numerical determination of the magnetic induction.

Fig. 6 shows a dependence of a value of the relative error of induction calculation $B=B_z$ (7) in the turn center (z=0) from a number *N* of elements of division of the rolling circumference.

Fig. 7 shows a dependence of the absolute error of the magnetic field induction calculation $\Delta B = |B_0 - B_z|$ on the turn axis with current I=1 arb. u. and radius R=1 arb. u. for N=180, so, we can make a conclusion on enough fast attenuation of the error with distance from the turn plane.

Errors of the magnetic field calculations in solenoids have been estimated by the same way. For a solenoid with internal radius R=1 and length H=10 containing 5 layers of turns and 1000 turns in each layer under division of the turn circumference into N=360 elements, relative error of the field calculation was 0.1 % in the solenoid center.

Analysis of above mentioned data allows making a conclusion that the suggested method ensures high accuracy of calculations of the magnetic field being enough for correct solution of tasks of contemporary electron optics.

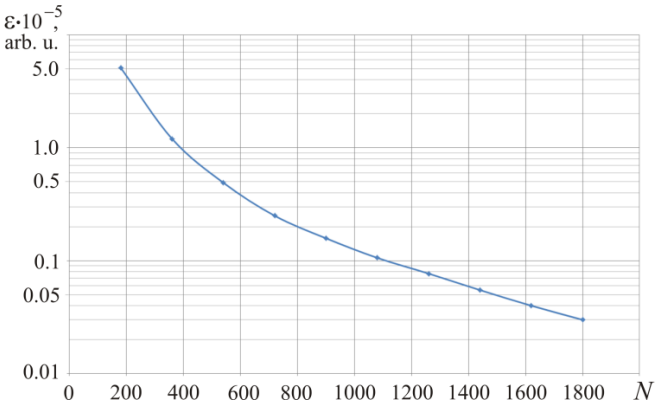


Figure 6. Relative error of the magnetic field calculation at the circular current center

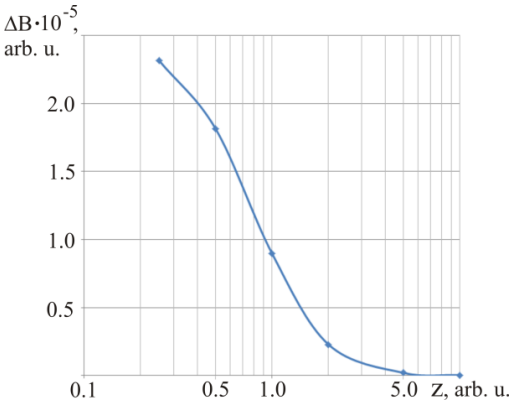


Figure 7. Absolute error of the magnetic field calculation on the circular current axis

Field of the arbitrarily oriented solenoid

Location of the solenoid in space can be definitely set by coordinates x_S, y_S and z_S of the upper base center and angles α and β fixing direction of its axis. Direction of the axial field vector $\vec{B}$ determined according to the right-hand screw rule (Fig. 8) in relation to the current direction is accepted as a positive direction of the solenoid axis $\vec{S}$. Angles α and β are angles of solenoid axis rotation around axes $0x'$ and $0y'$ in the coordinate system $0x'y'z'$ connected with the upper base center which all axes are codirected with axes of the laboratory coordinate system $0xyz$.

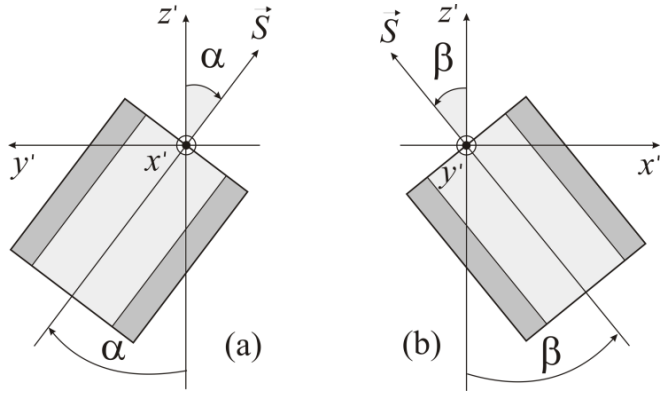


Figure 8. Location of the solenoid in space determined by angles of rotation α and β of its axis $\vec{S}$ around coordinate axes $0x'$ and $0y'$

Under specified center coordinates x_S, y_S, z_S in the coordinate system $0xyz$ and angles of orientation α and β of the solenoid, at the first stage the algorithm for calculation of the magnetic field induction $\vec{B}_S$ at the point (x,y,z) consists in a sequential usage of standard formulas for transformation of coordinates under rotations and application of formula (10) in the following way

$$\vec{B}_{SR}(x,y,z) = \vec{B}_S(x_\beta, y_\beta, z_\beta),$$

where $x_\beta = x_\alpha \cos\beta + z_\alpha \sin\beta$, $y_\beta = y_\alpha$, $z_\beta = z_\alpha \cos\beta - x_\alpha \sin\beta$.

Here $x_\alpha = x - x_S$, $y_\alpha = (y-y_S)\cos\alpha + (z-z_S)\sin\alpha$, $z_\alpha = (z-z_S)\cos\alpha - (y-y_S)\sin\alpha$.

At the second stage for final estimation of magnetic induction components $\vec{B}_S(x,y,z) = B_{Sx}\vec{\imath} + B_{Sy}\vec{\jmath} + B_{Sz}\vec{k}$ inverse rotations of the vector $\vec{B}_{SR} = B_{SRx}\vec{\imath} + B_{SRy}\vec{\jmath} + B_{SRz}\vec{k}$ should be executed according to expressions:

$$B_{x\alpha} = B_{SRx},\ B_{y\alpha} = B_{SRy}\cos\alpha - B_{SRz}\sin\alpha,\ B_{z\alpha} = B_{SRz}\cos\alpha + B_{SRy}\sin\alpha,$$

$$B_{Sx} = B_{x\alpha}\cos\beta - B_{z\alpha}\sin\beta,\ B_{Sy} = B_{y\alpha},\ B_{Sz} = B_{z\alpha}\cos\beta + B_{x\alpha}\sin\beta.$$

Field of a set of arbitrarily oriented solenoids

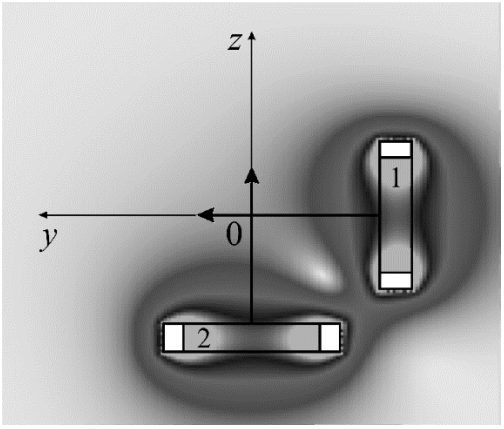


Figure 9. Map of a magnitude of the magnetic induction of two solenoids

It is obvious that in the case of several solenoids calculation of the magnetic field $\vec{B}$ is executed by summation of components of induction $\vec{B}_S$ or in a short form

$$\vec{B}(x,y,z) = \sum_S \vec{B}_S(x,y,z).$$

Fig. 9 shows results of simulation of the magnetic field of two solenoids (1 and 2) with orthogonally related axes. Relation of the exterior diameter to the interior one of each solenoid is 4:3. Distribution of the field is represented in the plane passing through axes of solenoids and encoded by shades of grey color.

Design of the microfocus x-ray tube

At present there is enough rapid expansion of microfocus x-ray tube applications. It caused by the fact that microfocus sources of x-ray radiation have a range of advantages in comparison with macrofocus ones:

• microfocus sources are essentially able to ensure high locality of researches;
• microfocus instruments ensure a higher quality of images under equal doses in the receiver plane;
• microfocus sources allow obtaining increased (in 5-10 times) images.

From the point of view of electron optics, microfocus x-ray source (tube) is an axially-symmetrical electron-beam generator. Traditionally electron generators are made in the form of a sequence of the cathode modulator assembly, several lens systems for acceleration and focusing and, if necessary, an electron-optical circuit for electron beam sweeps in raster on the target surface.

Microfocus tubes of the transmission type ensure the best quality of images. However, power of such tubes with a planar anode cannot exceed 10-20 W due to strong local heating of the anode surface. One of methods to increase power of tubes of transmission type consists in execution of a narrow channel with special form in the thick anode [10]. Under bombardment by accelerated electrons wall of the channel become sources of x-ray quanta.

Above mentioned methods for calculation of electrical and magnetic fields have been integrated in the copyright software FOCUS [11] intended for simulation of electron-optical systems of the wide range. Example of the trajectory analysis of an x-ray microfocus tube with axially-symmetrical construction is shown in Fig. 10. A narrow channel is executed in the anode. Symmetry axis is indicated as 0*x*. Distribution of the electrical field in a tube is shown in Fig. 1.

Exact positioning of the electron beam into the channel can be ensured by the system of magnetic deviation consisting, for example, of two solenoids (ref. Fig. 9) which axes are perpendicularly to the symmetry axis 0*x*. Simulation of the tube with two solenoids 1 and 2 installed outside has shown a possibility of electron beam positioning at any point of the anode surface. So, under the current value 100 ampere-turns in any of solenoids, a value of the trajectory deviation Δr on the anode from the axis 0*x* is approximately 10% of the solenoid exterior radius (Fig. 11).

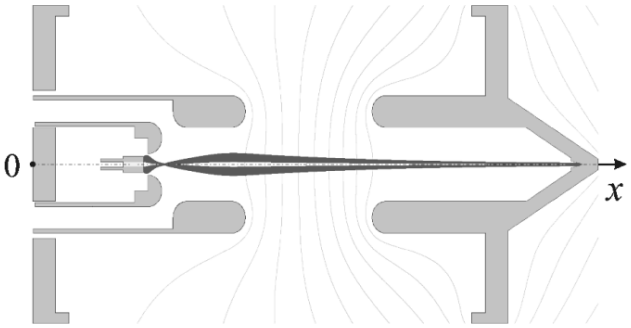


Figure 10. Electron-optical scheme of the microfocus x-ray tube

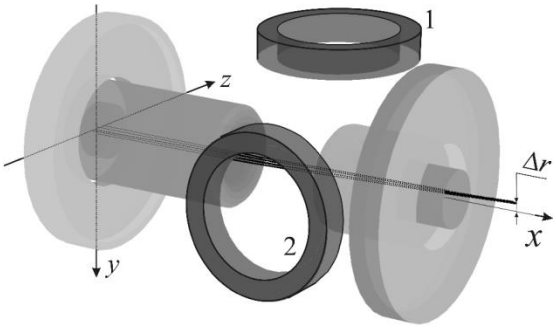


Figure 11. Deviation of the electron beam by the solenoid magnetic field

Conclusions

Boundary Elements Method for simulation of electrostatic fields in axially-symmetrical electron-optical systems with practically arbitrary configuration of electrodes has been developed.

Current Elements Method for simulation of magnetic fields of a set of solenoids arbitrarily oriented in space has been suggested and researched.

Methods for simulation of fields have been integrated into the copyright software FOCUS intended for design of a system with electromagnetic fields of complicated special configuration.

Results of simulation of a microfocus x-ray tube with high power have been represented.

Acknowledgments

The research has been carried out at expenses of the Russian Scientific Fund grant (project No.15-19-00132).

References

[1] Golikov, Yu. K. and Krasnova, N. K. (1999) Electrostatic mass spectrometer for concurrent mass-, energy- and angle-resolved measurements, *Nucl Instrum Meth Phys Res A* **427** (1-2), 208-212.

[2] Cubric, D., De Fanis, A., Konishi and I., Kumashiro, S. (2011) Parallel acquisition electrostatic electron energy analyzers for high throughput nano-analysis, *Nucl Instrum Meth Phys Res A* **645** (1), 227-233.

[3] Cheng, Alexander H.-D. and Cheng, Daisy T. (2005) Heritage and early history of the boundary element method, *Engineering Analysis with Boundary Elements* **29** (3), 268–302.

[4] Greenberg, M. D. (1971*). Application of Green's functions in science and engineering*. Englewood Cliffs, Prentice Hall, NJ, USA.

[5] Brebbia, C. A. , Telles, J. C. F. and Wrobel, L. C. (2012) *Boundary Element Techniques: Theory and Applications in Engineering*, Springer, London, UK.

[6] Gurov, V.S., Saulebekov, A.O. and Trubitsyn A.A. (2015) *Analytical, Approximate-Analytical and Numerical Methods in the Design of Energy Analyzers*, In P.W. Hawkes (Ed.), Advances in Imaging and Electron Physics **192**, Academic Press, London, UK.

[7] Courant, R. and Hilbert, D. (1989) *Methods of mathematical physics*, Interscience Publishers, New York, USA.

[8] Abramovitz, M. and Stegun, I. A. (Eds.), (1964). *Handbook of mathematical functions*, National Bureau of Standards, Gaithersburg, MD, USA.

[9] Griffiths, D. (1998) *Introduction to Electrodynamics*, Addison-Wesley Professional, Prentice Hall, NJ, USA.

[10] Trubitsyn, A.A., Grachev, E.Yu. and Dubkov, M.V. (2016) *A microfocus X-ray Tube with High Power*, Proceedings of the 15th International Seminar "Recent Trends in Charged Particle Optics and Surface Physics Instrumentation". Skalský dvur near Brno, Czech Republic.
[11] http://www.focuspro-soft.com.

Computational Approach to Analyzing 3D Strain Distribution in Opaque Materials via Micro Computer Tomography

† Lingtao Mao[1], Haizhou Liu[1], and *Fu-pen Chiang[2]

[1] State Key Laboratory of Coal Resources and Safe Mining, China University of Mining & Technology, Beijing 100083, China

[2] Dept. of Mechanical Engineering, Stony Brook University, Stony Brook, NY 11794-2300, USA

* Presenting author: mlt@cumtb.edu.cn

† Corresponding author: mlt@cumtb.edu.cn

Abstract

In this paper we introduce a newly developed 3D stress/strain analysis technique, called digital volumetric speckle photography (DVSP), that has the capability of probing internal strain distribution inside opaque solids under load. We take advantage of X-ray computed tomography's ability to record 3D volumetric image of solids with internal markers such as impunities, voids, etc and treat them as 3D volumetric speckles. Under load these markers will move accordingly. We track their displacements via a two-stage 3D FFT process which is an extension of the 2D process used in 2D digital speckle photography technique developed in 1993 by F.P. Chiang and his coworkers. We have successfully applied DVSP to strain analysis of rocks, concrete, and composites. The resolution of the technique is a function of the hardware used. It varies from macro scale with a medical X-ray CT, to micro scale with an industrial X-ray CT, and to nano-scale with a synchrotron radiation CT.

Keywords: 3D strain analysis, digital volumetric speckle photography, computed tomography, FFT.

Introduction

Digital Speckle Photography (DSP) technique evolves from the speckle photography technique originally proposed by J Burch in 1968 years before the invention of lasers. Over the years, it has evolved into techniques such as laser speckle photography, electron speckle photography, white light speckle photography, and one-beam laser speckle interferometry. Up until the advent and ubiquitous usage of digital camera, the process of generating useful information from a specklegram is always done by using a laser beam. In the pointwise approach, a narrow laser beam is directed at a point of a double exposed specklegram. The resulting diffraction pattern consists of a circular halo modulated by an array of parallel fringes, which can be related to the magnitude and direction of the displacement vector experienced by the cluster of speckles within the diameter of the laser beam. For the full field approach, an optical spatial filtering process is employed to display the displacement contours resolved along a particular direction with a sensitivity corresponding to the particular spatial frequency. In the early 1990s the process was digitalized by Chiang and his coworkers[1]. Only the surface deformation of a plane object can be obtained using this 2D-DSP technique.

With the aid of advanced imaging devices, such as high-resolution X-ray computer tomography (Micro CT), micro magnetic resonance imaging (micro-MRI) or laser scanning confocal microscope (LSCM), high-spatial-resolution volumetric images of opaque or semi-transparent materials can be acquired. Combined these advanced imaging devices with 2D Digital Image Correlation (DIC), a novel and useful technique for the quantification of 3D internal deformation, a new technique called DVC(Digital Volume Correlation) has emerged. The DVC method is a 3D extension of 2D-DIC, first proposed by Bay et al for strain analysis in bone [2]. By extending the 2D-DSP, Chiang and Mao recently developed a new 3D strain

analysis technique called Digital Volumetric Speckle Photography (DVSP) [3], which offers a higher computational efficiency than DVC by using the FFT (fast Fourier transform) algorithm. In this paper we present the theory of DVSP and its application to 3D strain analysis of rock, concrete and woven composite.

Theory of Digital Volumetric Speckle Photography(DVSP)

Digital volume images of a 3D solid before and after deformation are reconstructed with advanced imaging techniques using a CT or a MRI (either micro or macro). These two volume images are defined as reference volume image and deformed volume image, respectively. Both of them are divided into volumetric subsets with voxel arrays of 32×32×32 voxels, for example, and 'compared'. The principle is schematically shown in Fig. 1. Based on the theory of 2D digital speckle photography[1, 4], the DVSP principle is as follows[5]:

Let $h_1(x,y,z)$ and $h_2(x,y,z)$ be the gray distribution functions of a pair of generic volumetric speckle subsets, before and after deformation, respectively, and that

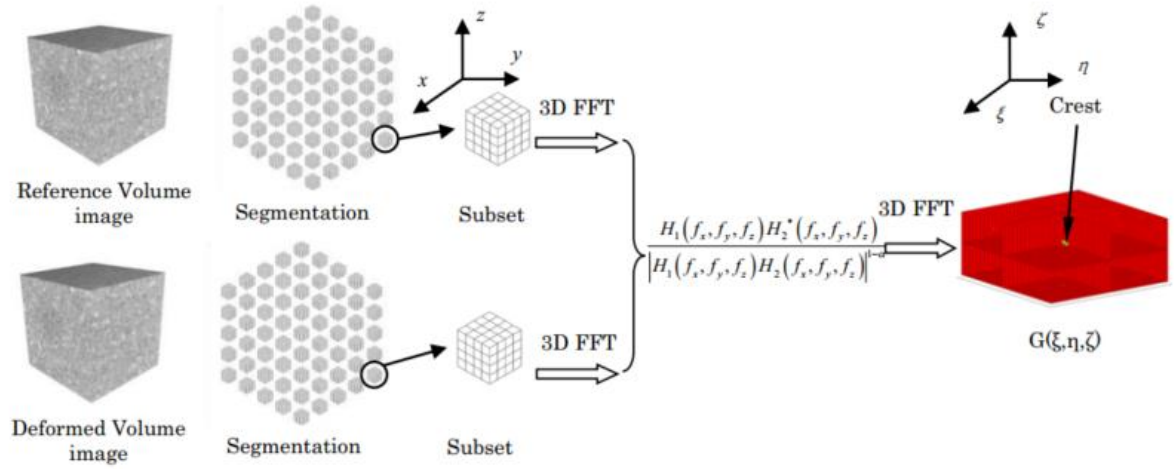


Fig.1 Schematics demonstrating the processing algorithm of DVSP

$$h_1(x,y,z) = h(x,y,z)$$

$$h_2(x,y,z) = h[x-u(x,y,z), y-v(x,y,z), z-w(x,y,z)] \tag{1}$$

where *u*, *v* and *w* are the displacement components experienced by the speckles along the *x*, *y*, and *z* directions, respectively. A first-step 3D FFT (Fast Fourier Transform) is applied to both h_1 and h_2 yielding

$$\begin{aligned} H_1(f_x,f_y,f_z) &= \Im\{h_1(x,y,z)\} = |H(f_x,f_y,f_z)|\exp[j\phi(f_x,f_y,f_z)] \\ H_2(f_x,f_y,f_z) &= \Im\{h_2(x,y,z)\} = |H(f_x,f_y,f_z)|\exp\{j[\phi(f_x,f_y,f_z)-2\pi(uf_x+vf_y+wf_z)]\} \end{aligned} \tag{2}$$

where $H_1(f_x,f_y,f_z)$ is the Fourier transform of $h_1(x,y,z)$, $H_2(f_x,f_y,f_z)$ is the Fourier transform of $h_2(x,y,z)$, and $\Im$ stands for Fourier Transform. $|H(f_x,f_y,f_z)|$ and $\phi(f_x,f_y,f_z)$ are spectral amplitude and phase fields, respectively.

Then, a numerical interference between the two 3D speckle patterns is performed at the spectral domain, i.e.

$$F(f_x, f_y, f_z) = \frac{H_1(f_x, f_y, f_z) H_2^*(f_x, f_y, f_z)}{\left|H_1(f_x, f_y, f_z) H_2(f_x, f_y, f_z)\right|^{1-\alpha}} \tag{3}$$

where * stands for the complex conjugate, and α is an appropriate constant $(0 \le \alpha \le 1)$.

When $\alpha = 0$, Eq.(3) can be expressed as

$$F(f_x, f_y, f_z) = H_1(f_x, f_y, f_z) \frac{\exp\left\{-j\left[\phi(f_x, f_y, f_z) - 2\pi(uf_x + vf_y + wf_z)\right]\right\}}{\left|H(f_x, f_y, f_z)\right|} \tag{4}$$

where $\dfrac{\exp\left\{-j\left[\phi(f_x, f_y, f_z) - 2\pi(uf_x + vf_y + wf_z)\right]\right\}}{\left|H(f_x, f_y, f_z)\right|}$ is essentially an inverse filter (IF).

When $\alpha = 0.5$, Eq.(3) can be expressed as

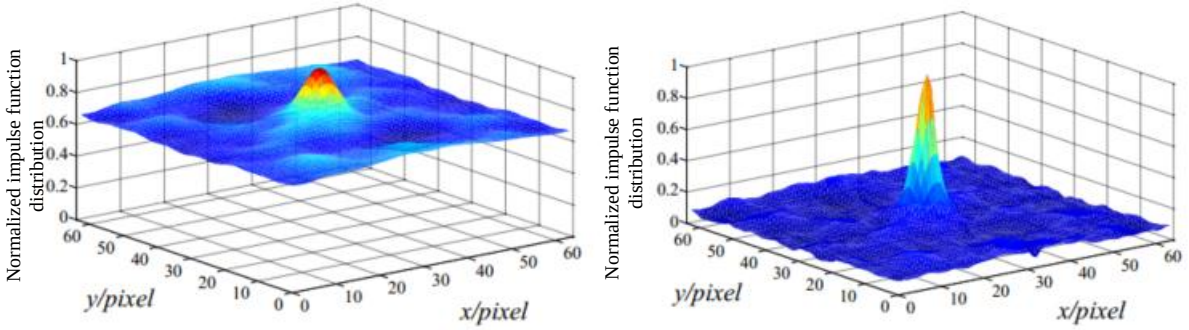
$$F(f_x, f_y, f_z) = H_1(f_x, f_y, f_z) \exp\left\{-j\left[\phi(f_x, f_y, f_z) - 2\pi(uf_x + vf_y + wf_z)\right]\right\} \tag{5}$$

where $\exp\left\{-j\left[\phi(f_x, f_y, f_z) - 2\pi(uf_x + vf_y + wf_z)\right]\right\}$ is a so-called phase-only filter (POF).

When $\alpha = 1$, Eq.(3) can be expressed as

$$F(f_x, f_y, f_z) = H_1(f_x, f_y, f_z) H_2^*(f_x, f_y, f_z) \tag{6}$$

where $H_2^*(f_x, f_y, f_z)$ can be viewed as a classical matched filter(CMF). When a correlation filter is chosen, Peak sharpness and noise tolerance are the criteria to be considered. In the 2D digital speckle photography technique [1,4], α is 0.5, and the algorithm is essentially a POF. In Ref. [6] the influence of CMF, POF and IF filters on the accuracy of 2D electronic speckle photography were analyzed and the results indicated that IF is extremely sensitive to noise, thus cannot be used as a reliable filter. There is no significant difference between CMF and POF filters. But while the POF filter provides somewhat more accurate estimates of the peak position, the reliability of the CMF filter is better. In Fig.2 , normalized impulse function distribution are shown. It is noted that the peak impulse with POF is sharper.



(a) (b)

Fig.2 Normalized impulse function distribution with different filter (a) CMF filter; (b) POF filter

In this paper, $\alpha = 0.5$ is adopted. As a result Eq.(3) can then be written as

$$F(f_x, f_y, f_z) = \frac{H_1(f_x, f_y, f_z) H_2^*(f_x, f_y, f_z)}{\sqrt{\left|H_1(f_x, f_y, f_z) H_2(f_x, f_y, f_z)\right|}} = \left|H_1(f_x, f_y, f_z)\right| \exp\left\{j\left[\phi_1(f_x, f_y, f_z) - \phi_2(f_x, f_y, f_z)\right]\right\} \tag{7}$$

where $\phi_1(f_x, f_y, f_z)$ and $\phi_2(f_x, f_y, f_z)$, are the phases of $H_1(f_x, f_y, f_z)$ and $H_2(f_x, f_y, f_z)$, respectively. It is seen that

$$\phi_1(f_x, f_y, f_z) - \phi_2(f_x, f_y, f_z) = 2\pi(uf_x + vf_y + wf_z) \tag{8}$$

Finally, a function is obtained by performing another 3D FFT resulting

$$G(\xi, \eta, \zeta) = \Im\left\{F(f_x, f_y, f_z)\right\} = \bar{G}(\xi - u, \eta - v, \zeta - w) \tag{9}$$

which is an expanded impulse function located at (*u*, *v*, *w*). This process is carried out for every corresponding pair of the subsets. By detecting the crest of all these impulse functions, an array of displacement vectors at each and every subset is obtained, from which strain tensors can be calculated using an appropriate strain-displacement relation.

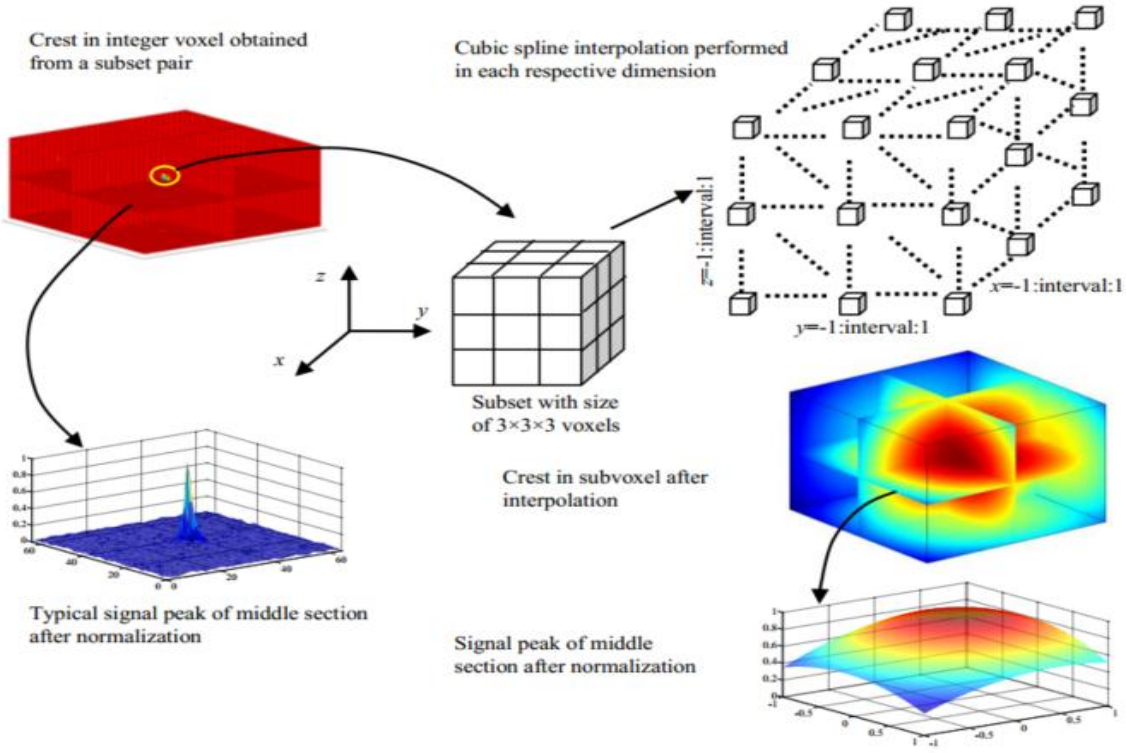


Fig.3 Schematics showing the interpolation procedure

In the above analysis the deformation of the subset itself is neglected. Because of the discrete nature of digital volume images, the displacement vectors evaluated from equation (9)

are integral multiples of one voxel. In order to obtain more accurate and sensitive characterization, a sub-voxel investigation of the crest position is needed. To achieve this, we select a cubic subset with 3×3×3 voxels surrounding an integral voxel of the crest and a cubic spline interpolation is employed to obtain the interpolated values among the integral voxels in each respective dimension. After interpolation, the cubic subset is enlarged and a new three dimensional array is generated with size depending on the interpolation interval. The smaller the interval and the bigger the array size give rise to higher interpolation accuracy. The price to pay, however, is the need for more computational time and more memory space. In practical applications there would be a tradeoff between the two competing needs. By detecting the positions of peak values of the new array, displacements of subvoxel accuracy can be obtained. The interpolation procedure is illustrated schematically in Fig.3.

Strain estimation

The internal strain tensor ε can be derived from the displacement fields. Due to the influence of unavoidable noise contained in the CT images, the displacements determined above contains discontinuities or noise that are not a feature of the material but a consequence of the discrete nature of the analysis performed. The errors in the local displacements may be amplified during the strain computation process. By using PLS (Point Least-Squares) approach, the errors can be largely reduced during the process of local fitting, and the strains estimated will be more accurate [7].

The element of PLS approach is shown as followings. To compute the local strains of each considered point, a regular cubic box with size of $(2N+1) \times (2N+1) \times (2N+1)$ discrete points surrounding the point is selected. If the strain calculation window is small enough, the displacements in each direction can be reasonably assumed to be linearly distributed, and therefore can be mathematically expressed as

$$\begin{aligned} u(x,y,\mathrm{z}) &= a_0 + a_1 x + a_2 y + a_3 z \\ v(x,y,\mathrm{z}) &= b_0 + b_1 x + b_2 y + b_3 z \\ w(x,y,\mathrm{z}) &= c_0 + c_1 x + c_2 y + c_3 z \end{aligned} \tag{10}$$

where $x, y, z = [-N,N]$are the local coordinates within the strain calculation box, $u(x, y, z)$, $v(x, y, z)$ and $w(x, y, z)$ are the displacements directly obtained by DVSP method, and $a_{\mathrm{i}=0,1,2,3}$, $b_{\mathrm{i}=0,1,2,3}$ and $c_{\mathrm{i}=0,1,2,3}$ are the unknown polynomial coefficients to be determined. With the Least-squares or Multiple Regression Analysis, the unknown coefficients can be estimated. Then, the six Cauchy strain components ε_x, ε_y, ε_z, ε_{xy},ε_{xz} and ε_{yz} at the interrogated point can thus be calculated as

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} = a_1 \quad \varepsilon_{xy} = \frac{1}{2}\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) = \frac{1}{2}(b_1 + a_2) \\ \varepsilon_y &= \frac{\partial v}{\partial y} = b_2 \quad \varepsilon_{yz} = \frac{1}{2}\left(\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z}\right) = \frac{1}{2}(c_2 + b_3) \\ \varepsilon_z &= \frac{\partial w}{\partial z} = c_3 \quad \varepsilon_{xz} = \frac{1}{2}\left(\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x}\right) = \frac{1}{2}(a_3 + c_1) \end{aligned} \tag{11}$$

With these Cauchy strain components, the principal strains can be calculated, and then the deviatotic strain ε_{s} and the volumetric strain ε_{v} are written as

$$\varepsilon_s = \frac{\sqrt{2}}{3}\sqrt{(\varepsilon_1 - \varepsilon_2)^2 + (\varepsilon_2 - \varepsilon_3)^2 + (\varepsilon_3 - \varepsilon_1)^2} \tag{12}$$

$$\varepsilon_v = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \tag{13}$$

where ε_1, ε_2 and ε_3 are the major, intermediate and minor principal strains, respectively.

X-ray Micro-CT

Over the years, medical CT scanners have been drastically improved in terms of image quality, imaging speed and deposited radiation dosage. But the spatial resolution remains limited to several hundreds of micrometres due to the dimension of the investigated object, i.e. a human patient. The micro-CT with high-resolution has emerged in the 1980s. Based on the type of X-ray generation, X-ray micro-CT can be divided into synchrotron-based and lab-based micro-CT (using X-ray tubes). Since synchrotron sources have a high X-ray flux, X-ray optics can be used to achieve very high spatial resolution up to tens of nanometers. Combined with different CT scanners, the resolution of DVSP varies from macro scale with a medical x-ray CT, to micro scale with an industrial x-ray CT, and to nano scale with a synchrotron radiation CT.

In this study, the main components of the industrial X-Ray computer tomography system are a microfocus X-ray source from YXLON (Feinfocus 225kV), an X-ray detector unit (1024×1024 pixels) from PerkinElme (XRD 0822AP 14), and a motorized rotation stage from Newport. The X-ray has a focus with size of 3 × 6 μm, a voltage range of 50-225kV, and the tube current ranging from 0 to 1440μA. A uniaxial compression setup has been designed and built to allow performing micro-tomography of a specimen under load (in situ).

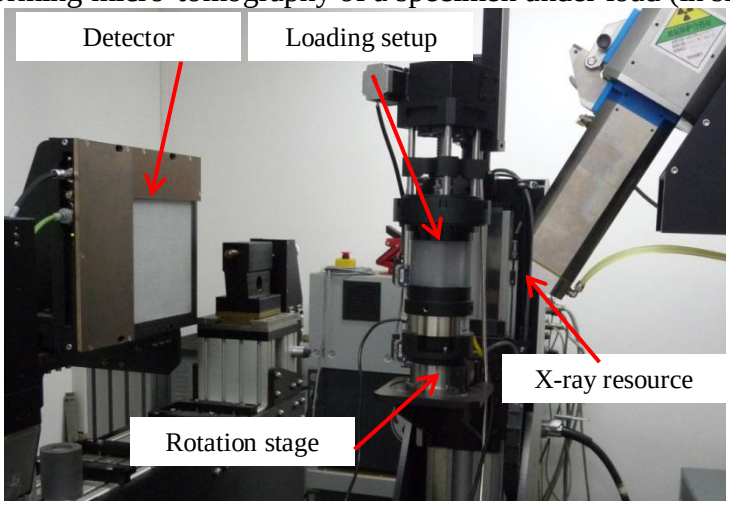


Fig.4 The Micro-CT system and the loading setup used to record the volumetric image of the specimen at each loading step

Applications of DVSP

Internal Strain Analysis of Red Sandstone with a Pre-existing Crack under Uniaxial Compression

A cuboid sample of red sandstone with a pre-existing crack under uniaxial compression was scanned in situ using the X-ray CT system. The sample has the size of 23 mm (L)× 10mm(W)× 40 mmc(H). A partial circular surface crack with an inclination of 45° with respect to the loading axis is carved into the specimen as shown in Fig5(a). The whole compression process was divided into 4 steps. The load-displacement history, reconstructed 3D images of Step 3 and Step 4, and a meridian slice of the specimen are shown in Fig.5 (a) , (b), (c), and (d), respectively. The volume image of step 1 was used as the reference image.

The subsequent deformed images were "compared" to the reference image via the DVSP method and resulted in displacement contours. The sectional image along section AA' shown in Fig.5(d) is depicted in Fig.5(e), and the *u*, *v*, *w* displacement fields and ε_{yz} strain distributions are plotted in Fig.5 (f),(g), (h), and (i), respectively.

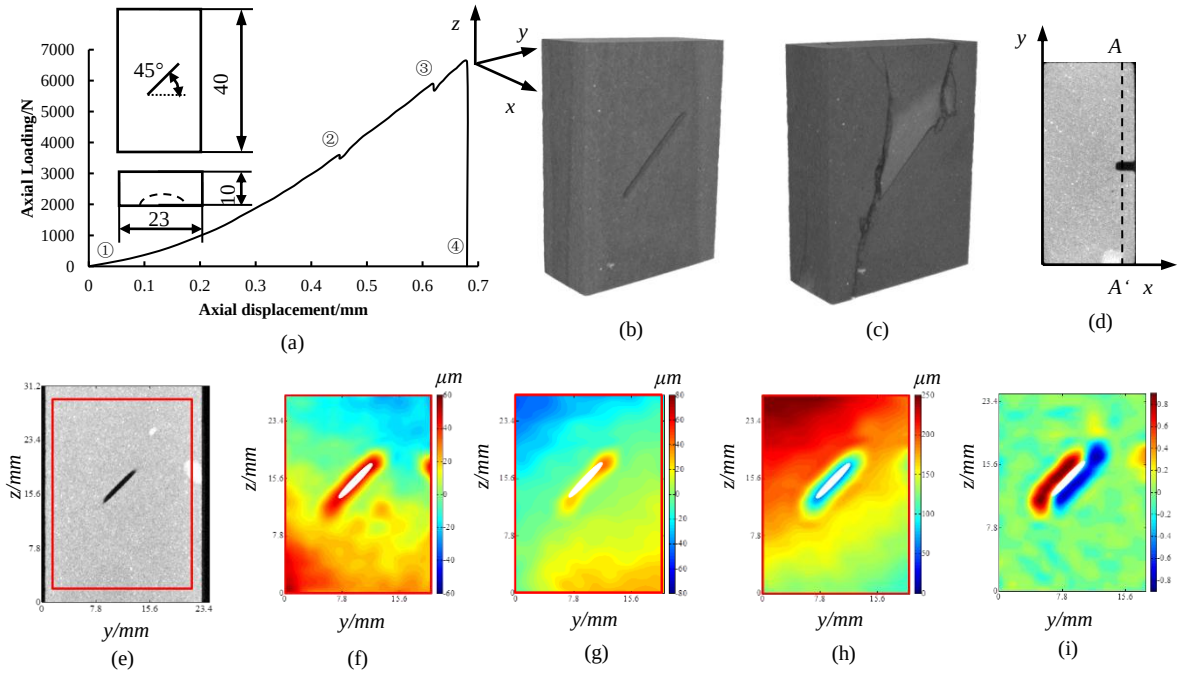


Fig.5 Application in red sandstone with pre-existing crack under Uniaxial Compression; (a) Load-displacement curve; (b) and (c) are reconstructed 3D images of Step 3 and Step 4; (d) Meridian slice; (e) Section along AA'; (f) u-field; (g) v-field; (h) w-field; (i) εyz strain

Internal Strain Analysis of Concrete under Uniaxial Compression

The concrete specimen is made from the subgrade of one highway. The size of the specimen is Φ25×48 mm. The compaction of the specimen was achieved by applying a compressive load in the axial (z) direction. The whole compression process was divided into 8 loading steps. In each step, the loading was kept constant while the specimen was scanned. By using DVSP we calculated the displacement and strain distributions in different sections of the specimen.

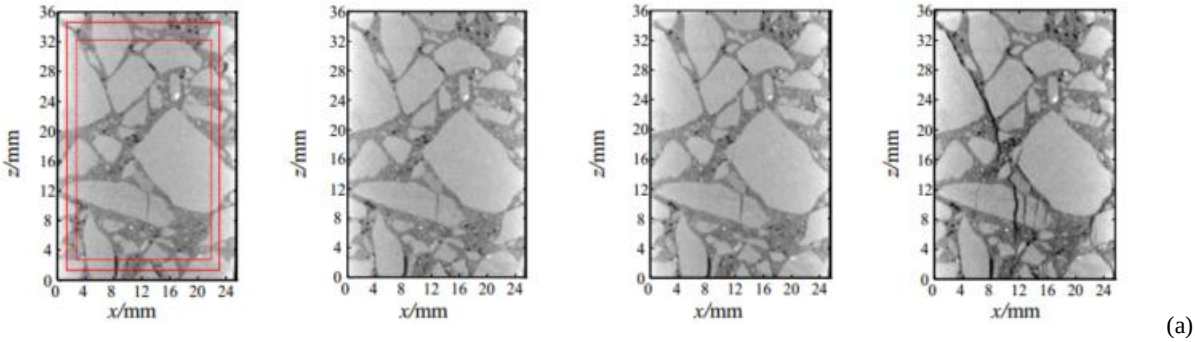


Fig.6.Reconstructed meridian sectional images under different loading; (a)5.3MPa ;(b)14.80MPa ; (c) 18.50MPa;(d)24.70MPa

Fig.6 shows the reconstructed meridian section images under different loading conditions. It is difficult to detect cracks from these gray images until the loading is at 24.70 MPa. Fig.7 shows the distributions of the major principal strain of the meridian section corresponding to the pictures shown in Fig.6. The light yellow mainly occurs at interface zones between aggregates and mortar which indicates the higher strain value appearing in these zones. The specimen tends to break at interface zones as demonstrated in Fig.6 (d).

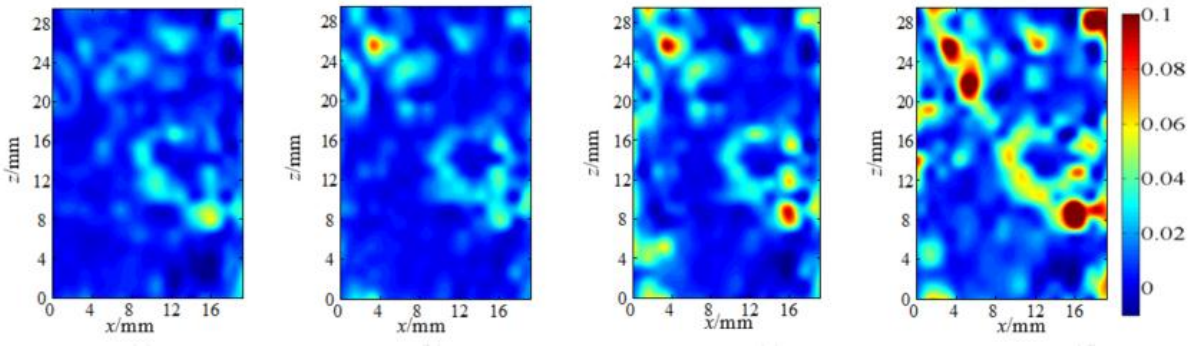


Fig. 7. Distribution contours of the first principal strain; (a)5.3MPa ;(b)14.80MPa ; (c) 18.50MPa;(d)24.70MPa

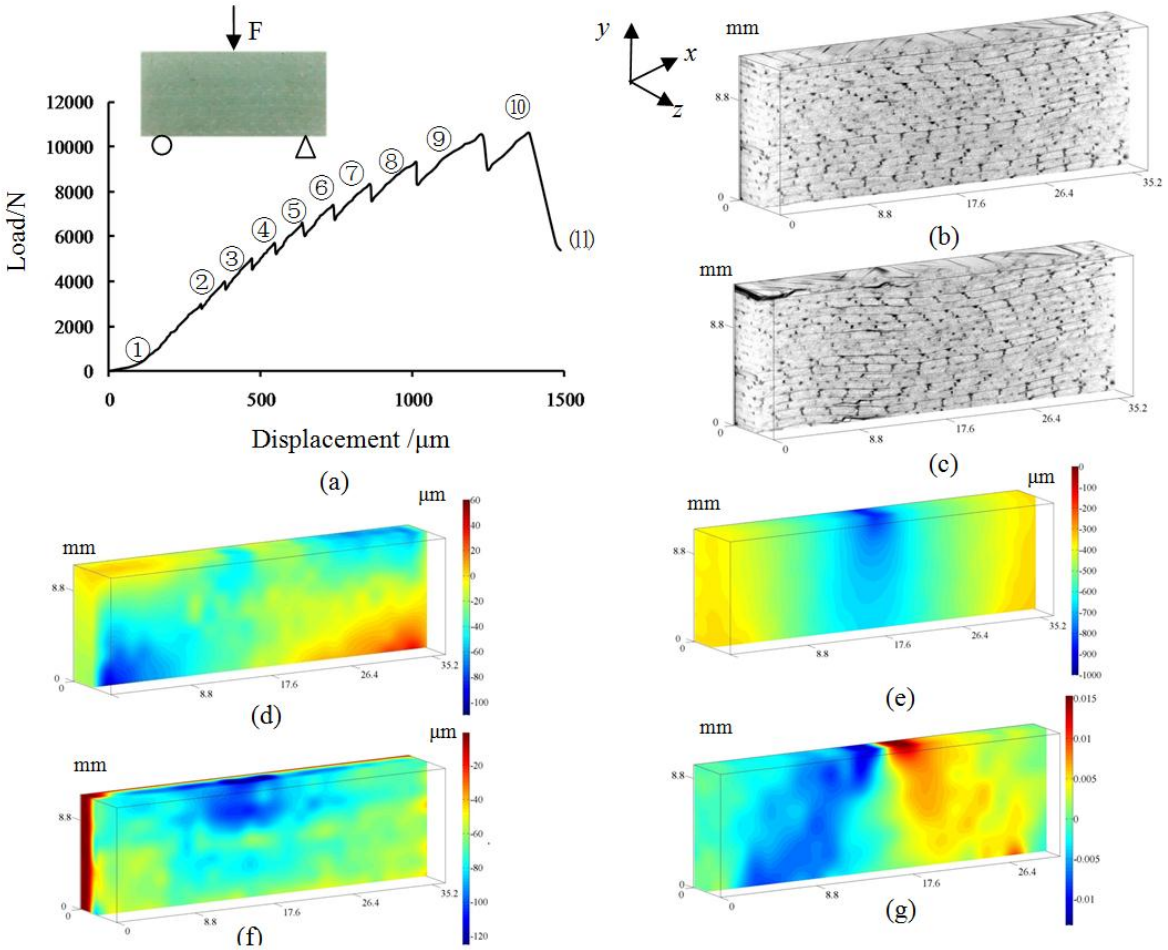


Fig. 8. Load-displacement curve , section images and deformation contours(a) Load-displacement curve (b) section image of Step 10; (c) Section image of Step 11; (d) u-field of section image of Step 10; (e) v-field of section image of Step 10; (f) w-field the section of Step 10, (g) ε_{xy} contours of the section;

Internal Strain Analysis of Composite under 3-point Bending

A tri-direction woven fabric composite beam with the size of 40mm(L)×19mm(H)×9mm(T) under 3-point bending was scanned in situ[8]. The matrix of the composite is epoxy resin, occupying 55% by volume. The filament diameter is 17μm. The experimental process is divided into 11 Step. The load-displacement curve, section images of Step 10 and Step 11 are shown in Fig 8(a), (b), and (c), respectively. The displacement and strain fields are presented in Fig.8 (d), (e), (f), and (g), respectively. As can be seen from the pictures that it is easy to detect the process of deformation localization.

Summary

We have demonstrated that the DVSP method can be effectively applied to analyzing internal strain distribution of red sandstone, concrete and composite, and we believe the DVSP technique has the potential of advancing the art of all 3D stress/strain analysis.

Acknowledgements

This work was financially supported by National Natural Science Foundation of China (51374211), State Key Research Development Program of China (2016YFC0600705), National Key Foundation for Exploring Scientific Instrument of China (2013YQ240803),and ONR Solid Mechanics Program grant N0014-14-1-0419 directed by Dr.Yapa Rajapakse.

References

[1] Chen D.J., Chiang F.P., Tan Y.S., et al. (1993) Digital speckle-displacement measurement using a complex spectrum method, App. Opt. 32, 1839-1849
[2] Bay B. K., Smith T. S., Fyhie D. P., et al, (1999) Digital volume correlation: Three-dimensional strain mapping using X-ray tomography, Experimental Mechanics 39, 217-226
[3] Chiang F.P. ,Mao L.T., (2015) Development of Interior Strain Measurement Techniques Using Random Speckle Patterns, Meccanica 50, 401–410
[4] F. P. Chiang, Super-resolution digital speckle photography for micro/nano measurements, Optics and lasers in Engineering, 47(2009):274-279
[5] Mao Lingtao, Fu-pen Chiang, 3D Strain Mapping in Rocks Using Digital Volumetric Speckle Photography Technique, Acta Mechanica，2016,227(11):3069-3085
[6] M. Sjodahl, Accuracy in electronic speckle photography, App. Opt., 36(13) (1997): 2875-2885
[7] B. Pan, D. F. Wu, Z. Y. Wang, Internal displacement and strain measurement using digital volume correlation : a least-squares framework, Measurement Science and Technology, 23(2012): 45002-45014
[8] Mao L.T., Chiang F.P. , (2015) Interior Strain Analysis of a Woven Composite Beam Using X-ray Computed Tomography and Digital Volumetric Speckle Photography, Composite Structures 134, 782-788

CFD analysis of the heat transfer of fire doors under the standard time-temperature curve

†H.Y. Leung[1]*, H. K. Tam[1], L. M. Tam[1,2] and S. K. Lao[1]

[1]Department of Electromechanical Engineering, Faculty of Science and Technology, University of Macau, Macau SAR, China.
[2]Institute for the Development and Quality, Macau, Macau SAR, China.

*Presenting author: hyleung@idq.org.mo
†Corresponding author: hyleung@idq.org.mo

Abstract

The fire resistance of the construction materials is traditionally measured by the fire resistance test method in accordance with the international standards, i.e. ISO 834-1, BS476-20, and ASTM E119, etc. In fire resistance door test, a full-scale test door is fitted into a standardized furnace for testing and then the test report is issued to ensure how long the door can resist the fire. A single leaf wooden door was tested in this study. During the 1-hour test, the temperatures inside the furnace and six measured points on the unexposed surface were recorded with the thermocouples. COMSOL Multiphysics®, was used to develop a three-dimensional heat transfer model for the fire door under the fire resistance test. In results, it showed that the curve trends of the simulated unexposed temperature of the six measured points (three thermocouples place on the door frame and three thermocouples place on the door) generally agreed with the experimental data. For the door leaf, the maximum temperature error between the experimental data and the simulated data was within ± 25%. For the door frame, the maximum temperature error between the experimental data and the simulated data was within ± 35%. The larger difference on the door frame was calculated because the practical smoke passing through the door crevice between the leaf and frame, due to the door sealant failure, was not simulated in this model.

Keywords: Heat Transfer, Fire Resistance Test, CFD Method

Introduction

In the fully developed fire, the temperature can be up to 1100^{o}C [1]. To avoid the fire passing through an opening of a compartment to another, the separating elements such as fire door should have the sufficient “fire resistance” performance. The fire resistance of the elements is traditionally measured by the fire resistance test method based on the international standards, such as ISO 834-1, BS476-20, and ASTM E119. In fire resistance door test, a full-scale test door is fitted into a standardized furnace for testing and then the test report is issued to ensure how long the door can resist the fire. The furnace is heated up to 1100^{o}C according to the standardized time-temperature curve. Test report relates only to what has been tested and allows very little in the way of variations. Changes to a construction require either another fire test or an assessment. Since the additional fire test is time and cost consuming, an accurate and scientific assessment tool is an alternation to address this issue. The assessment tool is capable to analyze the heat transfer accurately through a fire door under the standard test.

Some researchers have used the numerical method to analyze the thermal performance of the specimens of the fire resistance tests. Welch and Rubini [2] applied the CFD method to simulate a full-size 14-burners fire-resistance furnace with a steel specimen following the ISO-834 time-temperature curve. The results showed that radiation heat transfer was dominant, especially for the steel specimen. Moreover, that study stated that CFD method had potential for investigating the thermal behavior of fire resistance furnace and might be able to assist the harmonization of fire resistance test procedures. Chow and Chan [3] predicted the

fire resistance of building materials based on finite element analysis program. The results showed that the numerical prediction of aluminum sample agreed with the experimental results done by an electric furnace with the temperature following the standard fire curve of BS 476. However, the prediction of the other two samples made of hardwood and cement mortar was not good due to the burning of the material of hardwood and the changes of moisture evaporation from the cement mortar. Ferreira et al. [4] studied the fire resistance of the tabique wall experimentally and numerically. The experimental tests were tested in a fire resistance furnace according to ISO 834 standard fire curve. The experimental results showed the tabique wall panels fulfilled the requirements of the European fire resistance test standard. Also, the numerical model was developed by ANSYS software and the numerical temperatures showed good agreement with the experimental results.

Based on the literature survey, it can be summarized that the numerical method is suitable to simulate the fire resistance tests done with the practical fire test furnace. Therefore, the objectives of this study are to (1) conduct the fire resistance experiments conducted in the gas-fired furnace with the temperature following the standard firer curve of BS 476 : Part 20 and (2) apply the CFD method (COMSOL Multiphysics) to simulate fire resistance experiments for the analysis of heat transfer.

Fire Resistance Experiment and CFD Modeling

In this study, the fire resistance test was conducted in a gas-fired furnace. During the test, the furnace temperature rose on a time basis according to the standard curve of BS 476: Part 20 [5] (see Figure 1). A fire resistant door was used as a test specimen. For testing of the insulation of the fire door, the six Type-K thermocouples were placed on the unexposed side of the door.

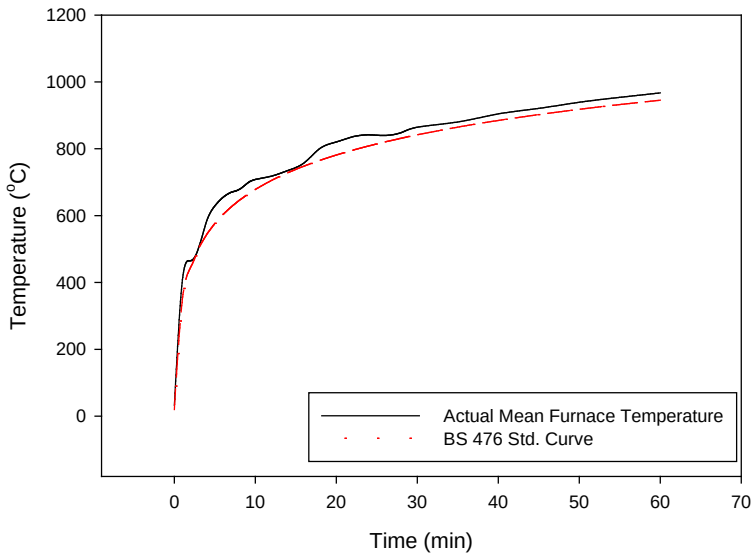


Figure 1. Comparison of the standard time- temperature curve and the mean furnace temperature during the test

Figure 2 shows the wooden door embedded in a concrete wall mounted in a steel frame. In the figure, three thermocouples (labeled as “1”, “3”, “5’) were adhered by the adhesive aluminum foil on the door frame and the other three thermocouples (labeled as “2”, “4”, “6”) were adhered by the adhesive aluminum foil on the door leaf. Since the structure of the wooden door was symmetric and the pressure at the same horizontal level was basically the same, heat transfer through the two horizontal positions (point 2’ and point 2 ; point 6’ and point 6)

should be the same. This assumption was also verified by the COMSOL simulation. After the simulation, the temperature difference between the symmetric positions (point 2' and point 2 ; point 6' and point 6) was calculated of less the 0.04%.

Figure 2. Single leaf - door installed on the furnace

Figure 3 depicts the dimension of the tested wooden door. The outer dimension for the door frame is 2398 mm (H) × 970 mm (W) × 101 mm (D) and that for the wooden door is 2360 mm (H) × 900 mm (W) × 59 mm (D), respectively. The wooden door is composed by sandwiching an acoustic layer into the two pieces of laminated wood boards. Because this study was performed for research purpose, the door components such as door lock, handle, and closer were not installed on the tested door.

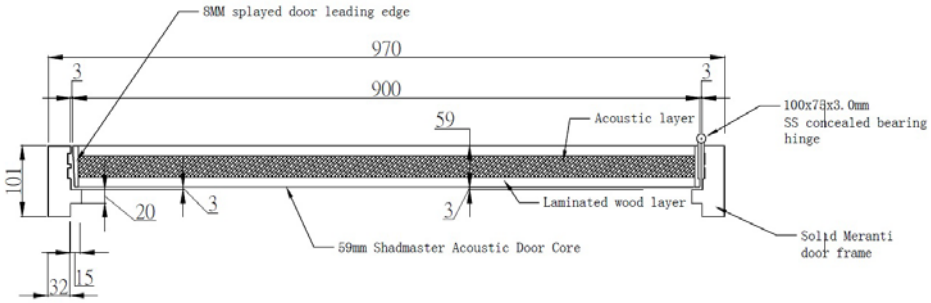


Figure 3. Cross-section of the tested door

COMSOL Multiphysics®, which is an engineering computation software in finite element analysis, has also been used to analyze the thermal behavior of the construction materials [6-9]. Therefore, the COMSOL Multiphysics® was used in this study to design and develop a three-dimensional simulation model to analyze the thermal behavior of the fire door under the fire resistance experiments. Figure 4 represents the three-dimensional model simulated in by Multiphysics. As shown in the figure, the model mainly includes a steel frame, a concrete, and a wooden door. Below lists the setting of those components:

1. Steel frame and concrete – The stainless steel frame is used in the model. The thickness of the concrete is 101 mm, which is the same as the door frame.
2. Door frame – The door frame is an inverted U-shape frame (Figure 5a). The material used for a frame is Meranti wood. The thickness of the door frame is 101 mm.
3. Door leaf – A 53-mm hardboard is sandwiched in two pieces of 3-mm thickness plywood boards. The thickness of the total door leaf is 59 mm.
4. Steel Hinge – Four steel plates are used as the hinges and installed between the door frame and the door leaf (Figure 5b).

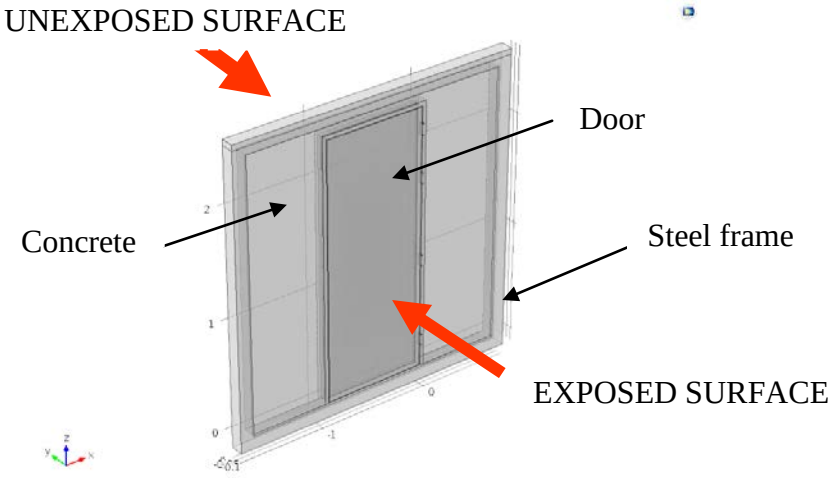
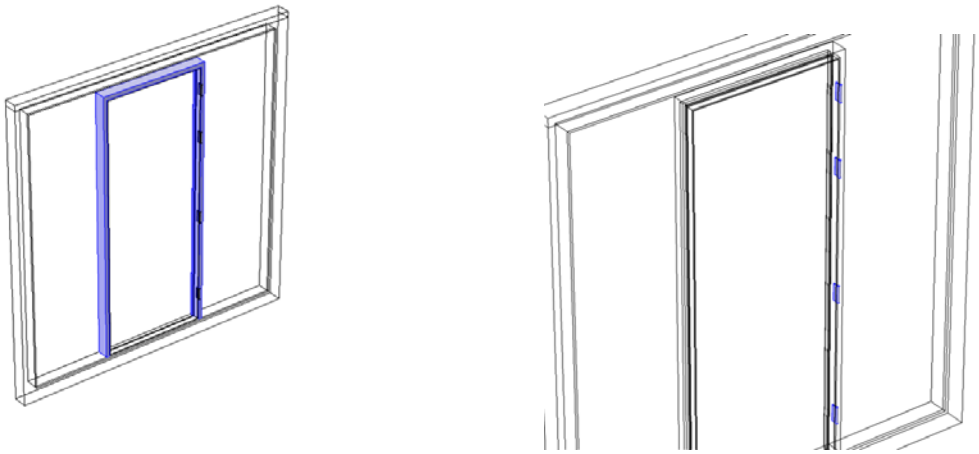


Figure 4. Overview of model



(a) (b)

Figure 5. (a) Wooden Door frame, (b) Steel hinges

In this study, the governing equation used in the simulation is expressed as:

$$\rho C_p \frac{\partial T}{\partial t} + \rho C_p \vec{u} \cdot \nabla T + \nabla \cdot (-k \nabla T) = Q + Q_{ted} \tag{1}$$

where Q_{ted} is the thermoelastic damping heat source, since it is not a loadbearing system, so Q_{ted} is equal to zero.

In the current simulation, four boundary conditions are assigned in the heat transfer model:

(1) A uniform initial temperature was applied to all domains of the model at the initial time $t = 0$;

(2) As shown in Figure 6, the temperature boundary condition of the exposed surface and the door crevice is expressed as [5]

$$T = 345 \cdot \log(8t + 1) + 20 \tag{2}$$

(3) As shown in Figure 7, the radiation boundary condition is applied to unexposed surface of the door, the concrete and the whole steel frame. The radiation equation used is written as

$$-\vec{n} \cdot (-k\nabla T) = \varepsilon\sigma(T_{amb}^4 - T^4) \tag{3}$$

where the surface emissivity of wood with concrete and steel are 0.25 and 0.83, respectively [7].

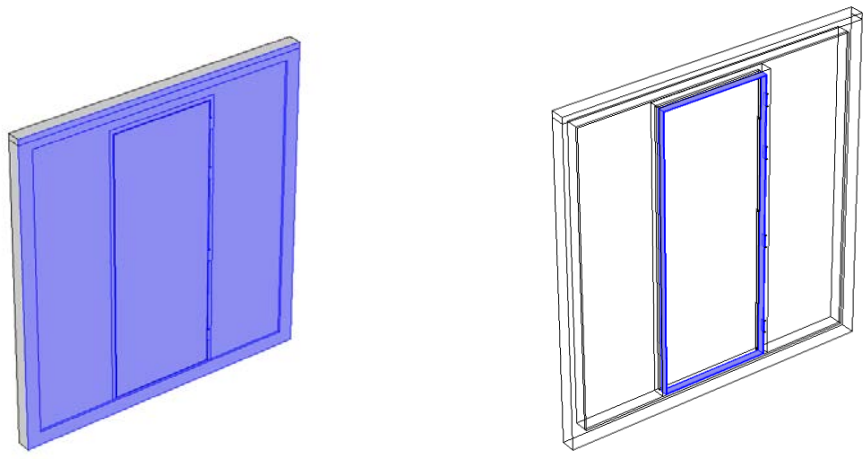
(4) As shown in Figure 8, the natural convection boundary condition is also applied to unexposed surface of the door, the concrete, and the steel frame. The natural convection equation used is written as

$$-\vec{n} \cdot (-k\nabla T) = h(T_{ext} - T) \tag{4}$$

where is restricted by the following condition:

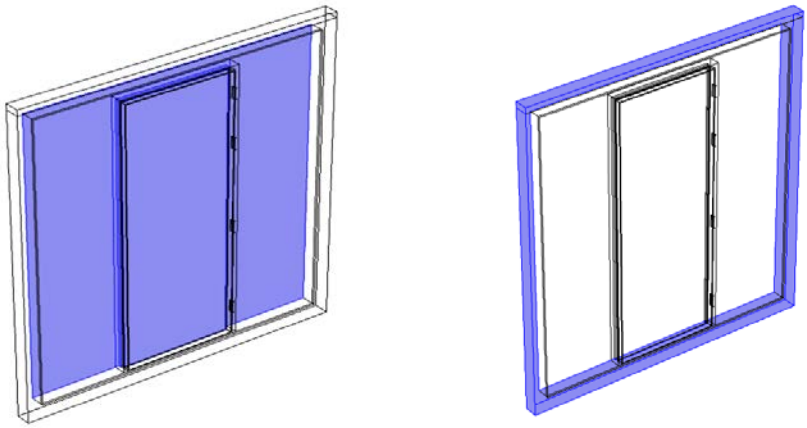
$$\begin{aligned} T &= T_0, \text{ if } \vec{n} \cdot \vec{u} < 0 \\ -\nabla T \cdot \vec{n} &= 0, \text{ if } \vec{n} \cdot \vec{u} \geq 0 \end{aligned} \tag{5}$$

Figure 9 shows the unstructured tetrahedral meshes used in COMSOL. In the figure, there are 897,469 domain elements, 207,354 boundary elements, and 6,977 edge elements used in the model. The finer mesh is suitable for simulating the model, especially, the door hinge.



(a) (b)

Figure 6. Temperature boundary of (a) exposed surface and (b) door crevice



(a) (b)

Figure 7. Radiation boundary of (a) the door and the concrete wall and (b) steel frame

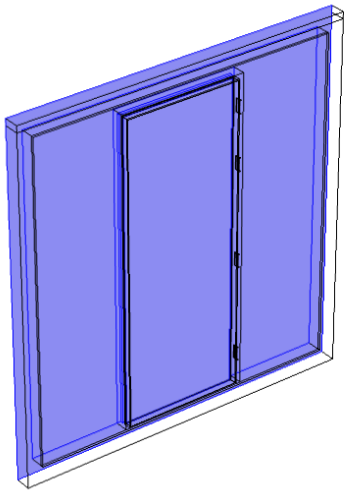


Figure 8. Natural convection boundary

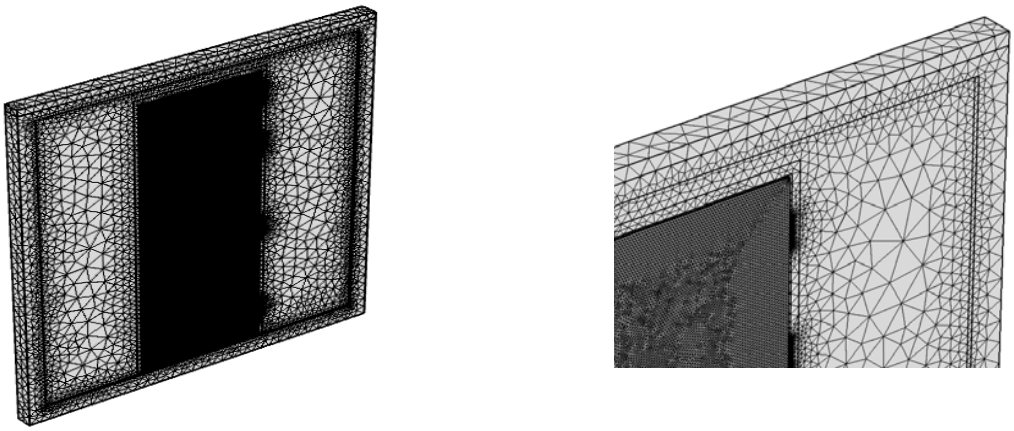


Figure 9. Tetrahedral mesh for the COMSOL model.

Table 1 shows the major properties of the materials used in the simulation. The properties of stainless steel and concrete are obtained from COMSOL library and the others are obtained from the book [10, 11].

Table 1. Material Properties used in this study

Material	Density (kg/m^3)	Thermal conductivity (W/m·k)	Specific heat capacity (J/kg·k)
Hardboard	330	0.15	1500 to 2000
Laminated wood (or plywood)	545	0.12	1210
Meranti wood (or softwood)	513	0.1 to 0.2	950 to 1600
Stainless steel	7194 to 7861	14.7 to 40.7	332 to 476
Concrete	2300	1.8	800

Numerical results and discussion

Figure 10 shows the temperature distribution of the unexposed surface at 1 hour. It is obvious that the temperature of the door is lower than the temperature of the concrete wall and steel frame. Moreover, the door crevice suffers a higher temperature than the wooden door and frame. This observation was also found in the experiment. Figure 11 shows the picture taken at the 1 hour of the experiment. It shows the location near the crevice was scorched and the smoke was released from the crevice so that the temperature of the location near the crevice was higher than the door core.

Figure 12 shows the comparison of the temperature results of the simulation with the experimental results measured at different locations of the tested door. In the figure, the experimental data points are shown in circular symbols and the simulation results based on the standard time-temperature curve and the actual mean furnace temperature measured in the experiment (see Figure 12) are shown in triangular symbols and square symbols, respectively. The difference between the simulations based on different furnace temperatures is nearly the same (within ± 2% error). For the curve trend, all the the simulated results show an agreement with the experiment. However, there is still a deviation between the experiment and the simulation. For the door leaf, the maximum temperature error between the experimental data and the simulated data is calculated within ± 25%. For the door frame, the maximum temperature error between the experimental data and the simulated data is calculated within ± 35%. The larger difference on the door frame is calculated because the practical smoke passing through the door crevice between the leaf and frame, due to the door sealant failure, was not simulated in this model.

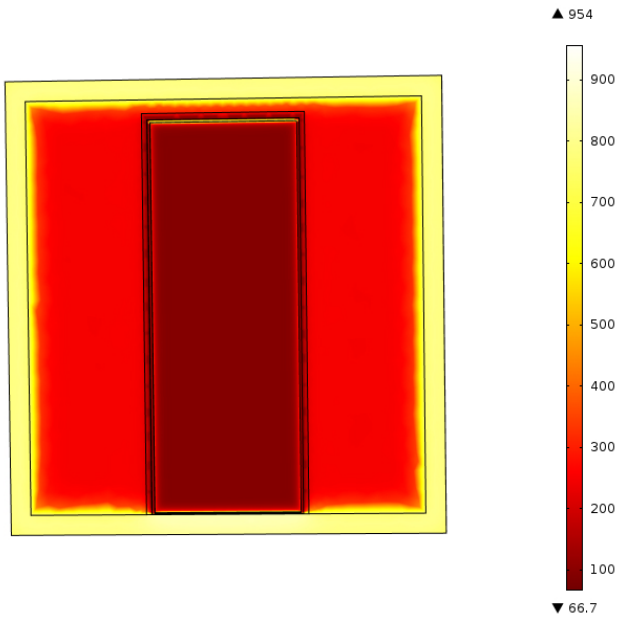


Figure 10. Unexposed surface temperature of model at the time of 60 minutes

Figure 11. Unexposed surface condition of experiment at the time of 60 minutes

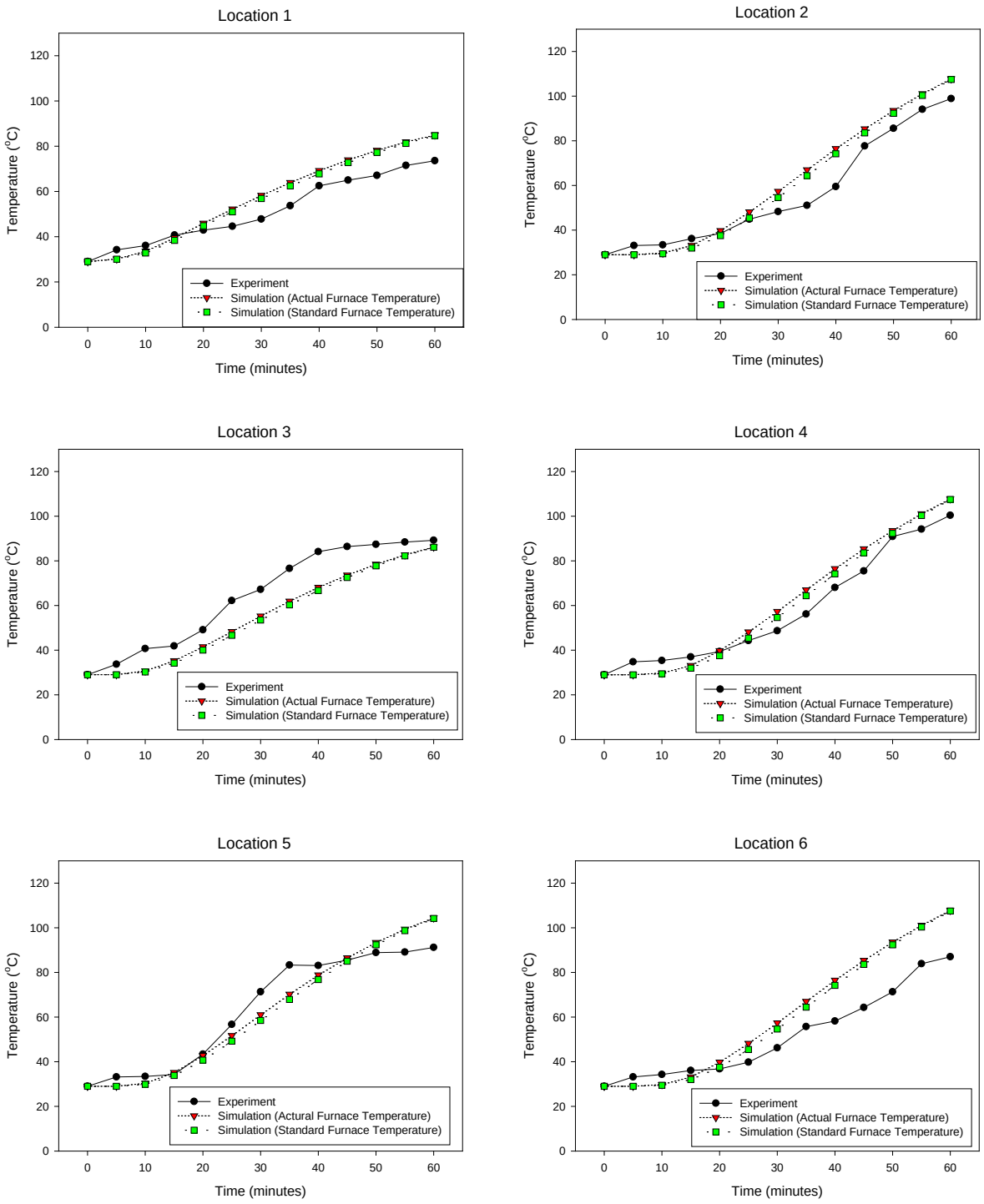


Figure 12. Comparison of the simulation temperatures with the experiments at different measuring locations.

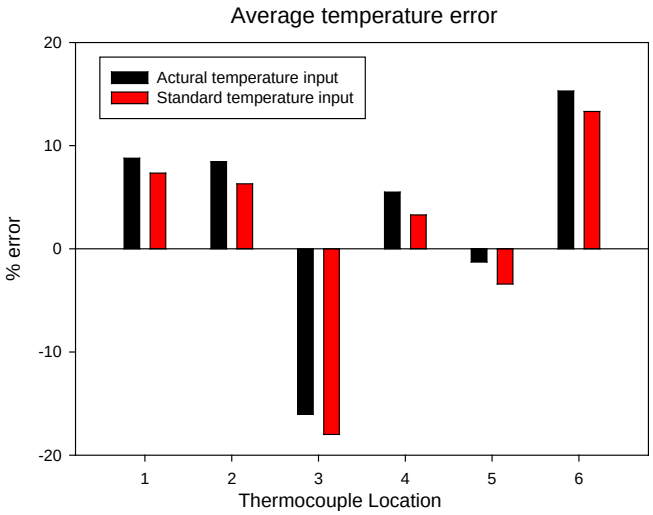


Figure 13. Average temperature error with the experiments at different measuring locations.

Conclusions

In this study, the CFD method is used to simulate the heat transfer through the tested door installed on a fire testing furnace with the standard furnace temperature rise. After comparing the simulation results with the experimental data, it can be concluded that the heat transfer model developed in this study can reasonably predict the unexposed temperature of the door leaf. The insolation of the simulation and experiment result at the unexposed size are consistent. Even the maximum temperature error of door core and door frame were within ± 25% and ± 35%, and the average temperature was within ± 18%. The simulation result of this model, include all the computation error, are capable to adjudge the insolation of this wooden door. In the future, the current heat transfer model can be improved by taking the smoke passing through the crevice, the furnace pressure condition, and the chemical reaction on the wood in the simulation.

Nomenclature

c	specific heat	J/kg·K
T	temperature	ºC or K
k	thermal conductivity	W/m·K
h	convection coefficient	W/ m^2·K
Q	total heat transfer	kJ
t	time	minute

Greek Symbols Letters

ε	surface emissivity (dimensionless)	
ρ	density of material	(kg/m^3)

Acknowledgement

This research is supported by the Institute for the Development and Quality, Macau.

References

[1] D. Drysdale, An Introduction to Fire Dynamics, 2nd edition, John Wiley & Sons, 2007.
[2] S. Welch, P. Rubini; "Three dimensional simulation of a fire-resistance furnace", Proceedings of 5th International Symposium on Fire Safety Science, Melbourne, Australia, March 1997, International Association for Fire Safety Science, pp. 1009-1020.
[3] D. M. Ferreira, A. Araujo, E. M. M. Fonseca, P. A. G. Piloto, J. Pinto, "Behaviour of non-loadbearing tabique wall subjected to fire – Experimental and numerical analysis," Journal of Building Engineering, Vol. 9, pp. 164-176, 2017.
[4] W. K. Chow, Y. Y. Chan, "Computer simulation of the thermal fire resistance of building materials and structural elements," Construction and Building Materials, Vol. 10, No. 2, pp. 131-140, 1996.
[5] --: "Fire Tests on Building Materials and Structures, Part 20, Method of determination of fire resistance of elements of constructions," British Standard Institution, 1987.
[6] J. L. Dréau, R. L. Jensen, K. Kolding, "Thermal behaviour of a gypsum fibre board associated with rigid polyurethane foam under standard fire conditions," Energy Procedia, Vol. 78, pp. 2736-2741, 2015.
[7] B. Weber, "Heat transfer mechanisms and models for a gypsum board exposed to fire," International Journal of Heat and Mass Transfer, Vol. 55, No. 5-6, pp.1661-1678, 2012.
[8] I. Venanzi, M. Breccolotti, A. D'Alessandro and A. L. Materazzi, "Fire performance assessment of HPLWC hollow core slabs through full-scale furnace testing" Fire Safety Journal, Vol. 69, pp. 12-22, 2014.
[9] J. A. Shaw, Thermal Analysis of a 1 Hour Fire Resistive Joint Design for Architectural Expansion Joints, Master Thesis, Rensselaer Polytechnic Institute, Troy, New York, 2013.
[10] Y. Cengel, A. Ghajar, "Heat and Mass Transfer : A Practical Approach," 5th Edition (SI Units), McGraw-Hill, 2015.
[11] ASHRAE Handbook 2013: Fundamentals, SI Editions, Amer Society of Heating, 2013.

Association of simultaneously measured limbs blood pressure differences with ankle-brachial index

***Xiaorui Song[1] , Gaoyang Li[1] , †Aike Qiao[1], Zongping Jin[2], Zhihui Chen[2]**

[1] College of Life Science and Bioengineering, Beijing University of Technology, CHINA

[2] School Hospital, Beijing University of Technology, CHINA

*Presenting author: sonxiaorui@163.com

†Corresponding author: qak@bjut.edu.cn

Abstract

Objects: To investigate the association of simultaneously measured limbs blood pressures with Ankle-Brachial Index as the current non-invasive diagnosis method of peripheral artery disease in clinical primary care.

Methods: 228 subjects (61 males, mean age, 63.92±10.72 years; 167 females, mean age, 59.47±7.33 years) were enrolled. Limbs blood pressure measurements were simultaneously performed using a blood pressure and pulse monitor device in the supine position. Data were statistically analyzed with SPSS 15.0.

Results: The mean age of the 229 subjects was 60.66±8.58 years. Variance analysis presented that RABI and LABI have significant differences with inter-arm difference in SBP (≥10 mmHg VS < 10 mmHg,≥10 mmHg VS ≥15 mmHg and≥15 mmHg VS < 10 mmHg). RABI have significant differences with inter-ankle difference in DBP (≥15 mmHg VS < 10 and≥10 mmHg VS ≥15 mmHg). Multinomial logistic regression analysis presented that LABI (<0.9; OR, 10.028; CI, (1.109-90.682); P=0.040) was independently associated with inter-arm SBP difference ≥ 10mmHg; LABI (<0.9; OR, 15.469; CI, (1.776-134.773); P=0.013) and RABI (0.90-1.00; OR, 4.231; CI, (1.205-14.860); P=0.024) were independently associated with inter-arm SBP difference ≥ 15mmHg. RABI (<0.9; OR, 7.189; CI, (1.010-51.179); P=0.049) and RABI (0.90-1.00; OR, 6.273; CI, (1.783-22.077); P=0.004) were independently associated with inter-ankle SBP difference≥ 15mmHg. LABI (0.90-1.00; OR, 4.331; CI, (1.039-14.330); P=0.016) was independently associated with inter-ankle DBP difference of ≥ 10mmHg. After excluding 99 hypertension patients, LABI (<0.9; OR, 246.330; CI, (5.442-11191.384); P=0.005) was still independently associated with inter-arm SBP difference ≥ 15mmHg.

Conclusion: LABI <0.9 was independently associated with inter-arm SBP difference ≥ 15mmHg, while these differences still existed after excluding 99 hypertensive patients. In addition, the cut off (0.90-1.00) of ABI was independently associated with inter-arm SBP difference ≥ 15mmHg and inter-ankle DBP difference ≥ 10mmHg or ≥ 15mmHg. Hence, detection of limbs blood pressure difference with simultaneous measurement may provide an aid for the non-invasive diagnostic method of peripheral artery disease in clinical primary care.

Key Words: ankle-brachial index; non-invasive diagnosis; blood pressure; statistical analysis; peripheral artery disease

Introduction

The Ankle-Brachial Index (ABI), i.e. the ratio of the blood pressure at the ankle to the blood pressure in the upper arm (brachium), is a simple, non-invasive, and well-documented diagnosis tool for Peripheral Artery Disease (PAD) in lower extremities. ABI<0.9 has been used to identify this condition in both clinical practice and epidemiologic studies [1-2]. Furthermore, the reduction of ABI is not only the independent risk factors of cardiovascular disease occurrence but also the potent predictors of cardiovascular disease mortality. Blanes et al.'s study reported that PAD (ABI<0.9) was associated with risk factors for inducing atherosclerosis, such as aging, male, diabetes, smoking, higher Systolic Blood Pressure (SBP) and coronary artery disease [3]. Thus, ABI is of important application value in clinical intervention treatment.

Recent study reported that a blood pressure difference between bilateral arms has been associated with subclavian stenosis, PAD, cardiovascular mortality and all-cause mortality [4-8]. Meanwhile, recent studies on inter-leg systolic blood pressure difference have added new evidence to this concept [9-11]. The meta-analysis reported by Singh [11] showed that inter-arm systolic blood pressure of 10 mmHg or more was associated with PAD (Risk Ratios(RR), 2.22; Confidence Interval (CI), 1.41–3.5; P =0.0006; sensitivity 16.6%; 6.7–35.4; specificity 91.9%; 83.1–96.3; 8 cohorts; 4774 subjects)，and inter-leg blood pressure difference of 15 mmHg or more was strong predictor of PAD (P =0 .0001).

Current technology has allowed to measure limbs blood pressure simultaneously [12], which could generate true blood pressure differences between four limbs, provide a better comprehensive evaluation of blood pressure [12, 14-15]. Inter-arm or inter-ankle blood pressure differences are known to be associated with low ABI, but little is known about the relationship between simultaneous limbs blood pressure differences and ABI.

Accordingly, this study aims to investigate the association of simultaneously measured limbs blood pressures with ABI as the current non-invasive diagnosis method of PAD in clinical primary care.

Subjects and Methods

Subjects

This study was approved by the Ethics Committee of School Hospital in Beijing University of Technology, and College of Life Science and Bioengineering in Beijing University of Technology. All subjects gave written informed consent. From September 2015 to January 2016, staffs of Beijing University of technology took part in comprehensive examinations of cardiovascular disease and its risk evaluation. Subjects with limb disability, hemiplegia, congenital heart disease, heart failure, and the history of artery intervention were excluded. Finally, 228 subjects (61 males, mean age, 63.92±10.72 years; 167 females, mean age, 59.47±7.33 years) were enrolled in this study.

ABI Measurements

ABI was measured in an air-conditioned room at a temperature of 22–23℃ by using the VS-1500 blood pressure and pulse monitor device (Fukuda Company, Beijing, China), which simultaneously and automatically measured the supine blood pressure of four limbs. Trained technicians placed the blood pressure cuffs on both arms and both ankles, and performed the measurements, after each subject had bared four limbs and taken 10-minute rest in supine position. ABI was calculated by the ratio of the ankle SBP divided by the arm SBP and higher value of the arm systolic blood pressure was used for the calculation.

Limbs Blood Pressure Measurements

Limbs blood pressure was measured in the same condition in supine position. In our study, limbs blood pressures were simultaneously and automatically measured by using the ABI device.

Based on the systolic and diastolic blood pressure, we calculated the inter-arm and inter-ankle blood pressure differences as the absolute value of the blood pressure difference between the right and left arm and blood pressure between the right and left ankle, respectively. Pulse pressure (PP) was the absolute value of difference between systolic and diastolic blood pressure. Pulse pressure index (PPI) was calculated as the ratio of PP to systolic blood pressure. Mean arterial pressure (MAP) was two-thirds diastolic pressure plus one-third systolic pressure. Above mentioned data were used for later statistical analysis.

In addition, the observer also administered a standardized questionnaire to collect information of subjects on age, sex, height, weight, medical history, lifestyle, use of medications, drinking and smoking history. The Body Mass Index (BMI) was calculated as the ratio of weight in kilograms to the square of height in meters.

Statistical Analysis

Data were stored in Excel 2013 and statistical analysis was performed with SPSS 15.0. Data values were expressed as percentages and mean±SD. The differences of inter-arm and inter-ankle were divided into 3 groups (≤10, 10-14 and≥15), and ABI were divided into 3 groups (<0.9, 0.90-1.00, >1.0). The differences between groups were checked by the analysis of variance for continuous variables and by Chi-square test for categorical variable. The differences of ABI on gender, hypertension, smoking and drinking were performed by independent-samples T test. Multinomial logistic regression analysis was used to determine the relationship between ABI and limbs blood pressure. A difference was considered significant if the P value was <0.05.

Results

The mean age of the 228 subjects was 60.66±8.58 years, 8 subjects were younger than 45 years, 31 subjects between 45 and 54 years, 136 subjects between 55 and 64 years, 39 subjects between 65 and 74 years, and 14 subjects were aged 75 years or older.

Table1. Baseline Characteristics of the Study Participants

Characteristics		Male(n=61)	Female(n=167)	P
Age,y		63.92±10.72	59.47±7.33	0.002
Body mass index, kg/m^2		24.54±3.15	25.75±3.70	0.023
Hpertension		32(52.5)	67(40.1)	0.041
Drinking		19(31.1)	8(4.8)	0.000
Smoking		26(42.6)	7(4.2)	0.000
Simultaneous limbs BP measurement, mmHg				
Left arm	Systolic pressure (LArSBP)	136.69±19.01	135.78±18.22	0.740
	Diastolic pressure (LArDBP)	82.50±10.96	81.02±10.28	0.342
Right arm	Systolic pressure (RArSBP)	136.89±18.47	135.78±18.58	0.688
	Diastolic pressure (RArDBP)	83.21±11.87	81.32±10.06	0.230
Left ankle	Systolic pressure (LAnSBP)	151.48±26.19	150.38±22.77	0.754
	Diastolic pressure (LAnDBP)	79.82±10.87	77.25±8.30	0.057
Right ankle	Systolic pressure (RAnSBP)	150.45±25.93	150.47±22.78	0.997
	Diastolic pressure (RAnDBP)	77.05±11.50	75.16±7.58	0.150
BP on the higher arm/ankle side of systolic pressure, mmHg				
Arm	Systolic pressure (HArSBP)	140.27±18.79	139.03±18.44	0.652
	Diastolic pressure (HArDBP)	83.94±11.56	81.72±10.38	0.166
	Pulse pressure (HArPP)	56.34±14.19	57.31±10.38	0.647
	Pulse pressure index	0.40±0.06	0.41±0.06	0.299
	Mean arterial pressure	102.72±12.73	100.03±11.85	0.295
Ankle	Systolic pressure (HAnSBP)	154.50±25.99	153.79±22.51	0.839
	Diastolic pressure (HAnDBP)	80.02±11.81	77.11±8.50	0.041
	Pulse pressure (HAnPP)	74.48±18.27	76.68±17.93	0.414
	Pulse pressure index	0.48±0.07	0.49±0.06	0.046
Inter-arm BP difference, mmHg				
Systolic	mean±SD	6.97±7.66	6.63±5.98	0.725
Diastolic	mean±SD	4.35±3.30	3.95±3.84	0.466
Systolic	≥10mmHg, n(%)	9（14.8）	23（13.7）	0.013
Systolic	≥15mmHg, n(%)	9（14.8）	17（10.2）	0.117
Diastolic	≥10mmHg, n(%)	3（4.9）	7（4.1）	0.206
Diastolic	≥15mmHg, n(%)	1（1.6）	4（2.4）	0.180
Inter-ankle BP difference, mmHg				
Systolic	mean±SD	7.06±7.87	6.74±5.20	0.715
Diastolic	mean±SD	4.87±4.07	4.11±3.48	0.165
Systolic	≥10mmHg, n(%)	5（8.2）	26（15.6）	0.000
Systolic	≥15mmHg, n(%)	8（13.1）	16（9.6）	0.102
Diastolic	≥10mmHg, n(%)	3（6.6）	18（10.8）	0.001
Diastolic	≥15mmHg, n(%)	3（4.9）	0（0）	——
Arm-ankle BP difference, mmHg				
L-ABI	mean±SD	1.08±0.15	1.08±0.10	0.979
R-ABI	mean±SD	1.08±0.16	1.08±0.09	0.734
L-ABI	<0.9, n(%)	4(6.6)	6(3.6)	0.527
	0.9-1.00, n(%)	7(11.5)	32(19.2)	0.000
	>1.00, n(%)	50(82.0)	129(77.2)	0.000
R-ABI	<0.9, n(%)	4(6.6)	4(2.4)	0.100
	0.9-1.00, n(%)	9(14.8)	28(16.8)	0.002
	>1.00, n(%)	48(94.1)	135(80.8)	0.000

Table 1 presents the clinical characteristics of the subjects by gender. The mean of limbs blood pressure and ABI in male was higher than female, but there was no significant difference between them. The distribution of limbs blood pressure differences was shown Figure1. There were 11.35%, 4.37%, 10.92%, and 9.61% of subjects with inter-arm difference in systolic blood pressure of >15 mmHg, inter-arm difference in diastolic blood pressure of >10 mmHg, inter-ankle difference in systolic blood pressure of >15 mmHg and inter-ankle difference in diastolic blood pressure of >10 mmHg, respectively.

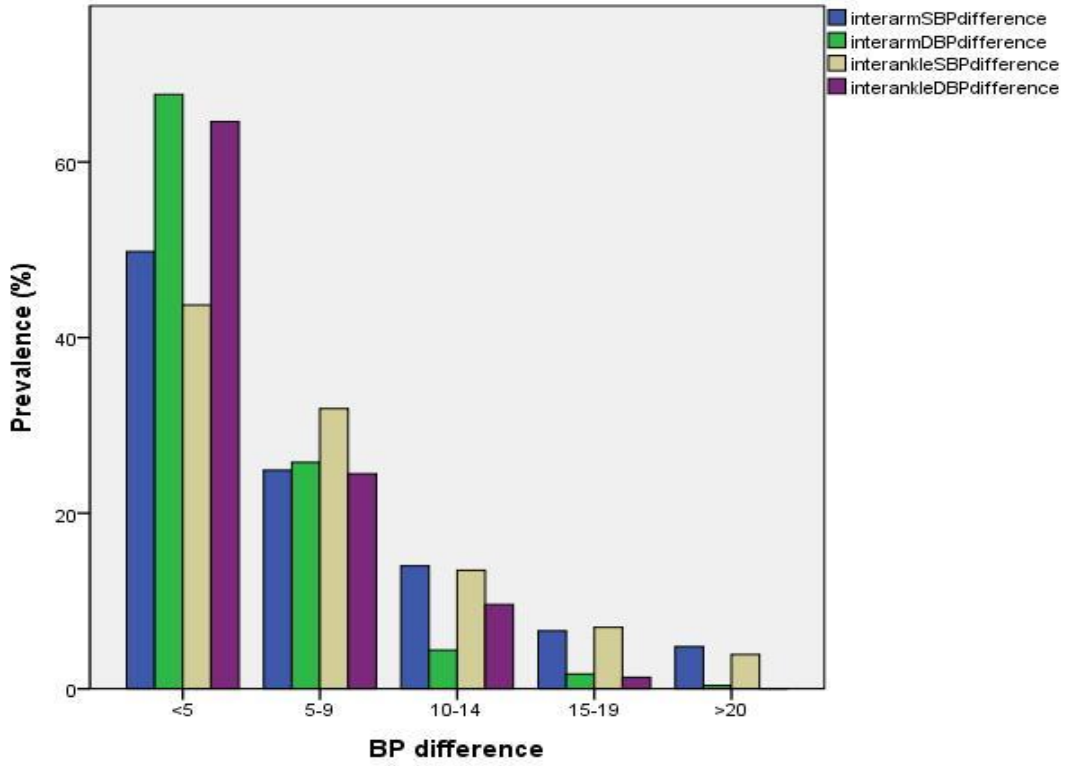


Figure1. The distribution of limbs blood pressure difference in this study subject

Variance analysis was performed to check the difference between limbs blood pressure differences(≤10, 10-14 and≥15) and ABI (<0.9, 0.90-1.00，>1.0), as shown in Figure 2. RABI and LABI have significant differences (P<0.05) with inter-arm difference in systolic blood pressure (≥10 mmHg VS < 10 mmHg, ≥10 mmHg VS ≥15 mmHg and ≥15 mmHg VS < 10 mmHg). RABI have significant differences (P<0.05) with inter-ankle difference in diastolic blood pressure (≥15 mmHg VS < 10 mmHg, ≥10 mmHg VS ≥15 mmHg). RABI and LABI have no significant differences (P>0.05) with inter-arm difference in diastolic blood pressure and inter-ankle difference in systolic blood pressure.

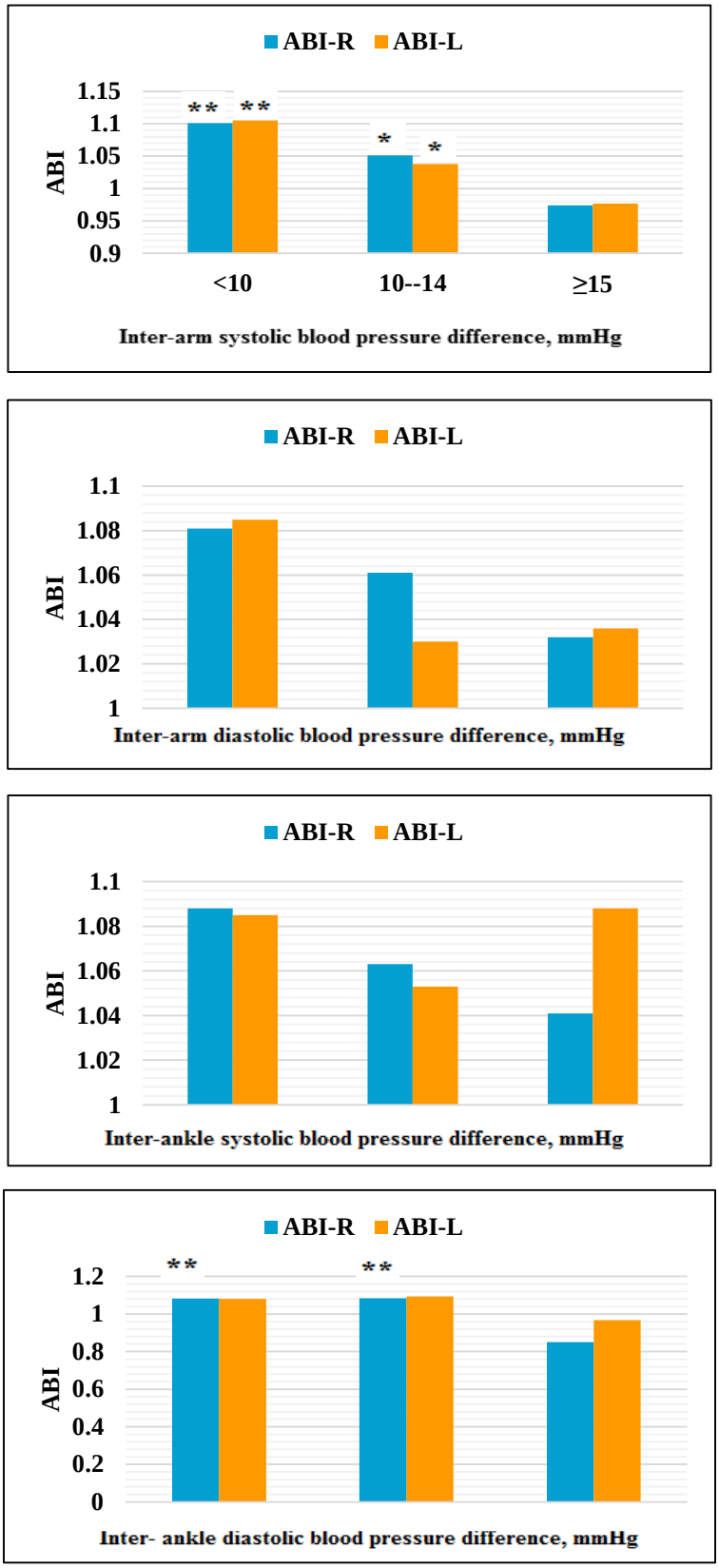


Figure2. Variance analysis between limbs blood pressure differences and ABI

*. The mean difference is significant at the 0.05 level.

**. The mean difference is significant at the 0.001 level.

The reference category is blood pressure difference ≥15 mmHg.

Multinomial logistic regression analysis presented the determinants of limbs blood pressure differences in all subjects, as shown in Table 2. BMI (Odds Ratio (OR), 1.151; 95% Confidence Interval (CI), (1.032-1.282); P=0.011), hypertension (OR, 0.350; CI, (0.152-0.807); P=0.014), and LABI (<0.9; OR, 10.028; CI, (1.109-90.682); P=0.040) were independently associated with inter-arm systolic blood pressure difference ≥ 10mmHg. Hypertension (OR, 0.236; CI, (0.080-0.695); P=0.009), LABI (<0.9; OR, 15.469; CI, (1.776-134.773); P=0.013) and RABI (0.90-1.00; OR, 4.231; CI, (1.205-14.860); P=0.024) were independently associated with inter-arm systolic blood pressure difference ≥ 15mmHg. RABI (<0.9; OR, 7.189; CI, (1.010-51.179); P=0.049) and (0.90-1.00; OR, 6.273; CI, (1.783-22.077); P=0.004) were independently associated with inter-ankle systolic blood pressure difference ≥ 15mmHg. LABI (0.90-1.00; OR, 4.331; CI, (1.039-14.330); P=0.016) was independently associated with inter-ankle diastolic blood pressure difference ≥ 10mmHg.
Since PAD has been associated with hypertension, we also performed a subgroup analysis after excluding 99 hypertension patients. After Multinomial logistic regression analysis in subgroup, we found that LABI (<0.9; OR, 246.330; CI, (5.442-11191.384); P=0.005) was still independently associated with inter-arm systolic blood pressure difference ≥ 15mmHg. In addition, ABI was not independently associated with inter-arm systolic blood pressure difference ≥ 10mmHg and inter-ankle systolic blood pressure difference ≥ 10mmHg or ≥ 15mmHg.

Table2. Multinomial logistic regression analysis on limbs blood pressure differences

≥10 mmHg			**≥15 mmh**		
Parameter	**OR(95% CI)**	**P**	**Parameter**	**OR(95% CI)**	**P**
Inter-arm systolic blood pressure, mmHg					
BMI	1.151(1.032-1.282)	0.011	Hypertension	0.236(0.080-0.69	0.009
Hypertensio	0.350(0.152-0.807)	0.014	LABI<0.9	15.469(1.776-134.7	0.013
LABI<0.9	10.028(1.109-90.682)	0.040	RABI(0.90-1.00)	4.231(1.205-14.860)	0.024
Inter-arm diastolic blood pressure, mmHg					
	——			——	
Inter-ankle systolic blood pressure, mmHg					
	——		RABI<0.9	7.189(1.010-51.179)	0.049
			RABI(0.90-1.00)	6.273(1.783-22.077)	0.004
Inter- ankle diastolic blood pressure, mmHg					
LABI(0.90-1.00)	4.331(1.039-14.330)	0.016		——	

The reference category is blood pressure difference <10 mmHg.
Covariates in the Multinomial logistic regression model include age, BMI, hypertension, smoking, drinking, LABI (<0.9, 0.90-1.00 and >1) and RABI (<0.9, 0.90-1.00 and >1).

In addition, variance analysis was performed to check the ABI difference on age and BMI, the independent-samples T test was performed to check the ABI difference on hypertension, drinking and smoking, and the Pearson correlation analysis was performed to check the correlation degree between limbs blood pressures and ABI. The mean of RABI and LABI

with hypertension was higher than without hypertension, and there was significant difference between them in RABI. The mean of RABI and LABI in drinking was higher than not drinking, but there was no significant difference between them. The mean of RABI and LABI in smoking was higher than not smoking, but there was no significant difference between them. RABI decreased with age (45-75 years), while LABI presented the undulated change, but there was no significant difference between groups. RABI and LABI increased with BMI (18.5-32kg/m^2), but there was no significant difference between groups. RABI and LABI have significant differences ($P<0.05$) and a position correlation with systolic blood pressure, diastolic blood pressure, pulse pressure and pulse pressure index in the ankle.

Discussion

In this cross-sectional study, by using a simultaneous measurement technique, the association between limbs blood pressure differences and ABI was evaluated. Multinomial logistic regression analysis presented that LABI <0.9 was independently associated with inter-arm systolic blood pressure difference ≥ 10mmHg or ≥ 15mmHg; LABI (0.90-1.00) was independently associated with inter-ankle diastolic blood pressure difference ≥ 10mmHg; RABI (0.90-1.00) was independently associated with inter-arm systolic blood pressure difference ≥ 15mmHg and inter-ankle systolic blood pressure difference≥15mmHg. RABI<0.9 was independently associated with inter-ankle systolic blood pressure difference ≥ 15mmHg. However, inter-arm diastolic blood pressure difference ≥ 10mmHg or ≥ 15mmHg, inter-ankle systolic blood pressure difference ≥10mmHg and inter-ankle diastolic blood pressure difference≥15mmHg have no obvious correlation with ABI. LABI (<0.9; OR, 246.330; CI, (5.442-11191.384); P=0.005) was still independently associated with inter-arm systolic blood pressure difference ≥ 15mmHg after excluding 99 hypertension patients.

The values of ABI were good markers for PAD, and lower ABI was reported to be associated with generalized atherosclerosis [16]. Previous studies found that ABI<0.9 had a significant correlation with inter-arm systolic blood pressure difference≥ 15 mmHg or diastolic blood pressure ≥ 10 mmHg [6, 17-18]. Moreover, ABI<0.9 had a significant correlation with inter-ankle systolic blood pressure difference ≥ 15mmHg in hemodialysis patients [9]. This study also illustrated that ABI<0.9 was independently associated with inter-arm systolic blood pressure difference≥ 10mmHg or ≥ 15 mmHg and inter-ankle systolic blood pressure difference≥ 15mmHg. Hence, generalized atherosclerosis might be indicated by inter-arm/inter-ankle systolic blood pressure difference of ≥ 10mmHg or ≥ 15 mmHg.

Chen et al. reported that ABI<0.9 had a significant correlation with inter-ankle diastolic blood pressure difference ≥ 10 mmHg in hemodialysis patients [9]. Su et al. reported that ABI<0.9 had a significant correlation with inter-ankle difference in systolic blood pressure ≥ 15 mmHg [10]. However, this study found that inter-ankle systolic blood pressure difference ≥10 mmHg and inter-ankle diastolic blood pressure difference ≥ 15 mmHg have no obviously independent correlation with ABI. These results were different form previous studies. Verberk et al. reported that the prevalence of inter-arm difference in systolic blood pressure of 10 mmHg or more was roughly doubled when diagnosis measurements method used a sequential approach, or used manual approach rather than automated measurements approach [19].

In this study, by using a simultaneous measurement technique, the association between limbs blood pressure differences and ABI was evaluated. Therefore, the measurement technique of blood pressure might explain the difference or low prevalence between inter-ankle blood pressure difference and ABI in this study. In addition, inter-arm diastolic blood pressure difference ≥ 10mmHg or ≥ 15mmHg was not independent correlation with ABI in this study. Peninsula medical of University of Exeter in England followed 230 patients with hypertension for 10 years and found that inter-arm systolic blood pressure difference might lead to diseases such as heart disease, stroke, cerebral vascular disease and arterial stenosis or hardening, while inter-arm diastolic blood pressure difference might increase risk of death from heart disease, stroke and other diseases [20]. Therefore, the investigation of inter-arm systolic blood pressure difference is more important than inter-arm diastolic blood pressure difference.

The value of ABI from 0.90 to 1.00 is the critical value to estimate PAD. In this study, multinomial logistic regression analysis presents that LABI (0.90-1.00) was independently associated with inter-ankle diastolic blood pressure difference ≥ 10mmHg, and RABI (0.90-1.00) was independently associated with inter-arm systolic blood pressure difference ≥ 15mmHg and inter-ankle systolic blood pressure difference≥15mmHg. Ovbiagele reported that PAD (0.90≤ABI≤1.00) was independently associated with stroke [21]. Thus, this study suggests that further studies are necessary to clarify between ABI (0.90-1.00) and blood pressure differences of four limbs.

Previous studies found that inter-arm difference in systolic blood pressure was associated with risk factors for atherosclerosis, such as aging, hypertension, hypercholesterolemia, obesity, and low ABI [18]. In this study, hypertension, BMI and ABI<0.9 were correlated with inter-arm systolic blood pressure difference of 10 mmHg or 15 mmHg. In addition, gender was not a significant factor associated with inter-arm SBP difference [18]. Other studies showed that female gender was independent risk factor for inter-arm systolic blood pressure difference of 10 mmHg or more [10]. In this study, there was a difference between female and male in inter-arm systolic blood pressure difference ≥10 mmHg and inter-arm systolic/diastolic blood pressure difference ≥ 10 mmHg by Chi-square test, but the gender was not an independent risk factor for inter-arm systolic blood pressure difference. The large gap in sample size between male and female might be a cause of formation of the results in this study. Further studies are necessary to clarify the association between gender and inter-arm systolic blood pressure difference.

In this study, by using a simultaneous and noninvasive measurement technique, we measured four limbs blood pressures. But this technique was not a popular one to measure blood pressure in daily clinical practice. Hence, although four limbs blood pressure with simultaneous measurement could improve the predictive value for PAD, our results might be changed if daily clinical measurement was used. In addition, the subjects were mainly from retired people, whose health care consciousness is better than the serving officer. The factors of smoking, drinking, salting and movement have no significant difference with ABI of subjects in this study. Hence, the clinical utility of this study may be limited in community

people. Furthermore, follow-up data of subjects and enlarged sample size of subjects is necessary for future study.

Conclusions

LABI <0.9 was independently associated with inter-arm SBP difference ≥ 15mmHg, while these differences still existed after excluding hypertensive patients. In addition, the cut off (0.90-1.00) of ABI was found to be independently associated with inter-arm SBP difference ≥ 15mmHg and inter-ankle DBP difference ≥ 10mmHg or ≥ 15mmHg. Hence, detection of limbs blood pressure difference with simultaneous measurement may provide an aid for the non-invasive diagnostic method of PAD in clinical primary care.

Acknowledgements

We would like to acknowledge the support of School Hospital in Beijing University of Technology (Beijing, China), where the fieldwork and measurement took place. This work was supported by National Natural Science Foundation of China **(81171107, 11472023).**

References:

[1] Fowkes F G，Murray G D，Butcher I, et al. (2008) Ankle brachial index combined with Framingham Risk Score to predict cardiovascular events and mortality: a meta-analysis, *JAMA: the journal of the American Medical Association*, **300**, 197.

[2] Su H M, Lin T H, Hsu P C, et al. (2012) Association of interarm systolic blood pressure difference with atherosclerosis and left ventricular hypertrophy, *PloS one*, **7**, e41173.

[3] Blanes J I, Cairols M A, Marrugat J. (2009) Prevalence of peripheral artery disease and its associated risk factors in Spain: The ESTIME Study, *International Angiology*, **28**, 20-25.

[4] Clark C E, Taylor R S, Shore A C, et al. (2012) Interarm blood pressure difference and vascular disease–Authors' reply, *Lancet*, **380**, 24-25.

[5] Clark C E, Campbell J L, Powell R J. (2007) The interarm blood pressure difference as predictor of cardiovascular events in patients with hypertension in primary care: cohort study, *Journal of Human Hypertension*, **21**, 633-638.

[6] Clark C E, Campbell J L, Powell R J, et al. (2007) The interarm blood pressure difference and peripheral vascular disease: cross-sectional study, *Family Practice*, **24**, 420-426.

[7] Cao K, Xu J, Sun H, et al. (2014) The variability of ankle–arm blood pressure difference and ankle–brachial index in treated hypertensive patients, *Journal of the American Society of Hypertension*, **8**, 693-698.

[8] Sheng C S, Liu M, Zeng W F, et al. (2013) Four-limb blood pressure as predictors of mortality in elderly Chinese, *Hypertension*, **61**, 1155-1160.

[9] Chen S C, Chang J M, Tsai Y C, et al. (2012) Association of interleg BP difference with overall and cardiovascular mortality in hemodialysis, *Clinical Journal of the American Society of Nephrology*, **7**, 1646-1653.

[10] Su H M, Lin T H, Hsu P C, et al. (2014) Association of interankle systolic blood pressure difference with peripheral vascular disease and left ventricular mass index, *American Journal of Hypertension*, **27**, 32-37.

[11] Cao K, Xu J, Shangguan Q, et al. (2015) Association of an inter-arm systolic blood pressure difference with all-cause and cardiovascular mortality: An updated meta-analysis of cohort studies, *International Journal of Cardiology*, **189**: 211-219.

[12] Singh S, Sethi A, Singh M, et al. (2015) Simultaneously measured inter-arm and inter-leg systolic blood pressure differences and cardiovascular risk stratification: a systemic review and meta-analysis, *Journal of the American Society of Hypertension*, **9**, 640-650.

[13] Cortez-Cooper M Y, Supak J A, Tanaka H. (2003) A new device for automatic measurements of arterial stiffness and ankle-brachial index, *American Journal of Cardiology*, **91**, 1519-1522.

[14] Hoeven N V, Lodestijn S, Nanninga S, et al. (2013) Simultaneous compared with sequential blood pressure measurement results in smaller inter‐arm blood pressure differences, *Journal of Clinical Hypertension*, 15, 839-844.

[15] Lohmann F W, Eckert S, Verberk W J. (2011) Interarm differences in blood pressure should be determined by measuring both arms simultaneously with an automatic oscillometric device, *Blood pressure monitoring*, **16**, 37-42.

[16] Sodhi H S, Shrestha S K, Rauniyar R, et al. (2006) Prevalence of peripheral arterial disease by ankle-brachial index and its correlation with carotid intimal thickness and coronary risk factors in Nepalese population over the age of forty years, *Kathmandu University Medical Journal*, **5**, 12-15.

[17] Clark C E, Campbell J L, Evans P H, et al. (2006) Prevalence and clinical implications of the inter-arm blood pressure difference: a systematic review, *Journal of Human Hypertension*, **20**, 923-931.

[18] Kimura A, Hashimoto J, Watabe D, et al. (2004) Patient characteristics and factors associated with inter-arm difference of blood pressure measurements in a general population in Ohasama, Japan, *Journal of Hypertension*, **22**, 2277-2283.

[19] Verberk W J, Kessels A G H, Thien T. (2011) Blood pressure measurement method and inter-arm differences: a meta-analysis, *American Journal of Hypertension*, **24**, 1201-1208.

[20] Study shows: too large interarm pressure difference may increase the early death risk. http://tech.qq.com/a/20120322/000292.htm

[21] Ovbiagele B. (2009) Association of ankle–brachial index level with stroke, *Journal of the Neurological Sciences*, **276**, 14-17.

Fatigue life prediction of stents in a realistic coronary stenosis model

***Xinyang Cui[1], Qingshuai Ren[1,2], †Aike Qiao[1], Gaoyang Li[1], Zihao Li[1]**

[1]College of Life Science and Bioengineering, Beijing University of Technology, Beijing, CHINA

[2] Institute of Mechanics, Chinese Academy of Science, Beijing, CHINA

*Presenting author: xinyangcui@qq.com

†Corresponding author: qak@bjut.edu.cn

Abstract

Background The clinical outcome and fatigue life of stent are closely related to the biomechanical environment of coronary stenosis. In order to predict fatigue life of coronary stents, the stents' adaptability and safety in a realistic stenosis model combined with the clinical data was studied in this paper.

Methods Coronary artery stenosis models were built with Mimics based on CT angiography (CTA) data. Finite element method (FEM) was used to simulate the stent expansion in a realistic and an idealized coronary stenosis model. The stress/strain of the two different models was compared. Based on the mechanical analysis, the influence of cyclic loading effect on the fatigue life of the stents was studied. Stents' fatigue rupture was calculated with Goodman diagram, and the fatigue performance parameters such as the cycle to failure, the fatigue life, fatigue safety factor (FSF), cumulative fatigue damage rate were also analyzed.

Results The maximum stresses in the stent, plaque, and the vessel wall in the realistic stenosis model were 413.8 MPa, 6.06 MPa, and 3.39 MPa, respectively. While in the idealized model, they were 418.3 MPa, 4.46 MPa, and 1.13 MPa, respectively. Although the maximum stress was always located at the bending area of crowns, the stress distributions were different largely in the two models. The relative error ratios of the maximum stress in the plaque and vessel wall between the two models were 26.4%, 66.7%. In the fatigue analysis, the stent would not fail for fatigue rupture calculated by Goodman diagram. The closest point to the fatigue limit was also located at the inner bend of crowns. The predicted number of cycles to failure was $5.32*10^8$, the fatigue life was 14 years, the FSF was 2.8, and the maximum cumulative fatigue damage rate was about 71.5%.

Conclusion The mechanical parameters (stress/strain) were associated with the realistic coronary stenosis model. It is feasible to use a realistic model to calculate the accurate stress/strain for a coronary stent to predict the dangerous point and to evaluate the accurate fatigue performance parameters. This method can serve as a useful tool to support interventional planning for stent implantation in order to minimize the risk of fatigue fracture.

Keywords: Endovascular stent, Stenosis, Biomechanics, Fatigue life prediction.

1. Introduction

In the last decades, stent intervention has become common in clinical to treat cardiovascular disease [1]. Currently, 85% of coronary interventions involve stents [2], adding up to more than 1 million stents per year [3]. But stent failure and fracture (SF) is an inherent risk of stenting [4], and coronary stent fracture (CSF) rate occurred in up to 30% [5,6]. Fractured stents lose their ability to scaffold occluded arteries and may cause restenosis [7], thrombosis [8], or artery perforation [9].

Stent failure can be caused by the mechanical loading, either monotonic or cyclic loading during deployment or service, respectively. Cyclic loading is due to the pulsatile blood pressure or due to vessel movements imposing bending, torsion, or tension/compression on the stent [10]. In the following, we will refer to the failure due to cyclic loading as fatigue. The computer-based design modeling represents an assessment tool for the prediction of stent performance and fatigue life [11, 12].

Computational models have been widely used for their ability to replicate the biomechanical response of medical devices under physiological conditions [14]. From the viewpoint of mechanical, the complex biomechanical environment caused by plaques may lead to CSF [13].

Researchers usually simulate the stent expansion process using idealized models [14,15,16], while many researchers have also used realistic models with the angiography and vascular ultrasound technology in recent years [17]. But the contrast of mechanical parameters (stress/strain) in the two models needs further research.

A study used a patient's arterial geometry data to evaluate the fatigue performance of peripheral stents [18]. But there was few research on the fatigue parameter based on the expansion of stents in a realistic coronary model. [19, 20], which provides new insights into the coronary stent design and failure mechanism.

In this study, structural finite element models were implemented to simulate the stent expansion in a realistic stenosis coronary artery model with the angiography and vascular ultrasound technology. The stresses and strains in the stent, plaque and vessel wall in an idealized model and a realistic model were compared to quantify the response with the presence of a realistic geometrical plaque. To investigate the effects of cyclic blood pressure on the fatigue resistance of the stent, not only Goodman diagram but also the fatigue performance parameters such as the cycle to failure, the fatigue life, fatigue safety factor (FSF), cumulative fatigue damage rate were analyzed.

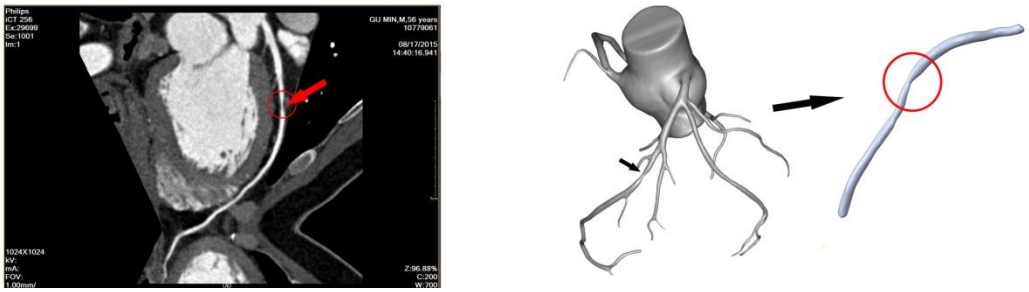
The information may be helpful in predicting the stent strut fracture and assessing the fatigue life for a coronary stent. This study may eventually help designers to optimize the stent geometry structure, for physicians in their diagnosis and intervention surgical decision

making process.

2. Methods

2.1 3D model reconstructed

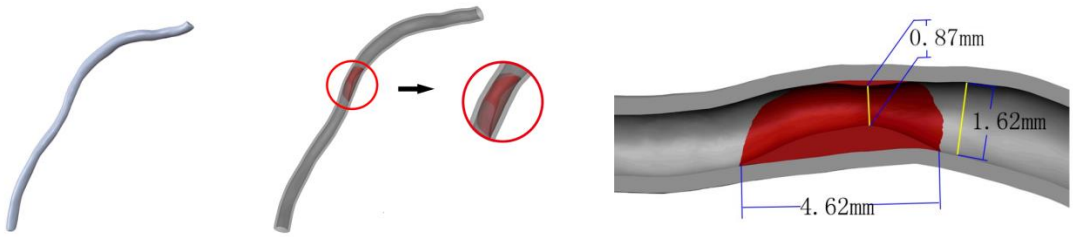
Coronary angiography was used to detect coronary atherosclerosis for a 56-year-old male patient. Left coronary artery (LCA) lesion (Fig. 1a) was detected. The realistic stenosis vascular model was reconstructed based on the clinical data of blood vessel and plaque. Firstly, the CT angiography (CTA) data at end-diastole was selected and digitized by using the medical image processing software Mimics 15.0 to eliminate the respiratory and cardiac motion artifacts. Then the coronary arterial tree was reconstructed, as shown in Fig. 1b. Figure 1c shows the narrow segment of blood vessel, which was obtained from the coronary arterial tree with red marker.



a b c

Figure 1. 3D model reconstruction of coronary stenosis

The stenosis segment was repaired by using the software "Freeform Modeling Plus" based on the data of coronary physiological parameters. A realistic healthy blood flow model was established, as shown in Fig. 2a. The realistic plaque geometry was obtained from the logic subtraction between the healthy and stenotic vessel. Besides, the realistic plaque model was reconstructed combined with its geometry data by "Solid Works 15.0". The healthy vessel model was obtained by offsetting the healthy blood flow from the physical parameters of the coronary artery wall. Eventually, the realistic plaque model was implanted into the healthy blood vessel to establish the realistic coronary stenosis model, which was shown in Fig. 2b. Figure 2c shows the geometry size of the realistic atherosclerotic lesions of blood vessels. The internal diameter of the blood vessel was 1.62 mm, the length of the stenosis (i.e., the length of the plaque) was 4.62 mm, and its diameter was 0.87 mm.



a b c

Figure 2. Reconstruction of vessel wall and plaque

2.2 Parametrized stent design

In order to design a new stent, the size of the narrow blood vessels was taken into consideration to ensure the implementation of the percutaneous coronary artery. The 3D geometry of the repeatable units of the stent was generated by using Pro/Engineer. The single unit cell consisted of 12 crowns with square shaped struts formed a circular stent (diameter=0.6 mm) as shown in Fig.3. The 6 unit cells were linked by I – link bars, and the parameters of the designed stent were shown in Fig. 3, too.

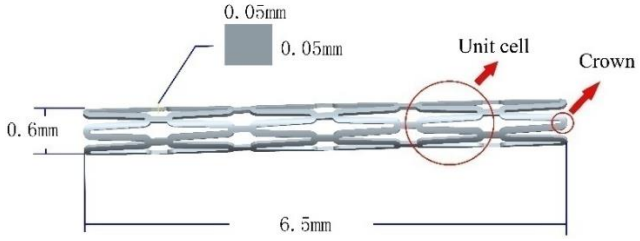
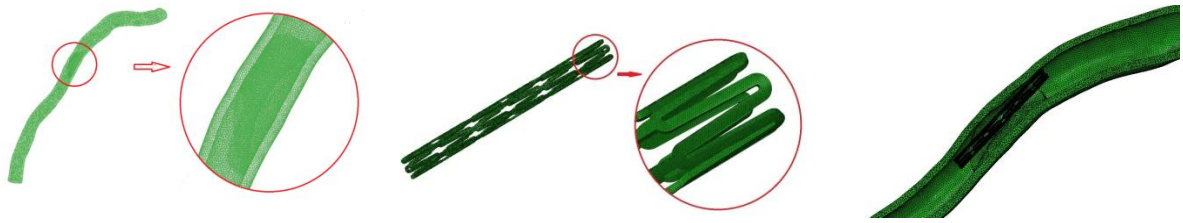


Figure 3. Geometrical model of the stent with I-link

2.3 Meshed models and material models

In this part, the vessel wall, plaque and the stent were meshed with the aid of the software Hypermesh11.0 respectively. Both the realistic vessel wall and corresponding plaque model (as shown in Fig. 4a) were meshed with tetrahedron element, which was relatively denser and would avoid the disturbance of the irregular structures in a realistic stenosis model. Since the realistic model and the stent were deformed largely in the expansion process, the stent was meshed with the linear tetrahedron element to prevent deformity interference deformation, as shown in Fig. 4b. The stent was positioned inside the stenosis coronary artery, and the combined model was meshed shown in Fig. 4c. The detail information of the meshes was shown in Tab. 1.



a b c

Figure 4. Mesh of the realistic model.

Table 1. Mesh information and material properties of the stent and vessel model

Models	Element types	Number of elements	Number of nodes	Material properties	Young's modulus (GPa)	Poisson ratio	Yield strength (MPa)
Stent	C3D4	674 335	175 809	304 Stainless steel(304SS)	193.00000	0.270	207
Plaque	C3D4	37286	9366	Calcified Coronary Atherosclerotic Plaque	0.00219	0.499	—
Vessel wall	C3D4	543 974	124930	Calcified vascular cell	0.00175	0.499	—

The material of the vessel wall and the corresponding calcified plaque were ideally linear

elastic, isotropic and incompressible, while the material of the stent was bilinear elastic-plastic. The detailed information of the combination was shown in Tab. 1, too.

2.4 Boundary conditions

Finite Element Method (FEM) was applied to simulate the deployment of a 304 Stainless steel (304SS) stent expanding in the coronary artery and the subsequent deformation under diastolic-systolic cyclic loading conditions. The commercial FEM tool ABAQUS/Standard 6.13 was used to run the simulations.

The stent was positioned inside the coronary artery (Fig. 4c). Then uniform radial deformation load was imposed on the stent's inner surface along the vertical axis to simulate the expansion process. Furthermore, nodes belonging to the stent surface were constrained in the tangential direction to avoid any rotation inside the vessel. The artery was constrained in the section near the plaque to prevent any displacements and rotations. To mimic possible interactions between the stent and the plaque, the sliding friction contact surfaces were introduced with the friction factor being 0.2 [21], with the same value for the interaction ratio between the stent and the vessel wall. Plaque and the vessel inner surfaces were assumed to be bonded, in order to ensure the models' displacement.

The simulated loading steps were as follows:

The first step: stent expansion. After percutaneous transluminal angioplasty (PTA), the balloon inflation was modeled through a displacement-driven analysis, where a uniform radial deformation was imposed to inner stent surface, till the diameter is equal to 1.1 times of the vessel inner diameter. This value was chosen to ensure the achievement of a residual stenosis lower than 30% that is required by clinicians for continuing with the stenting procedure [22]. At the end of this step, the deformation was deflated, and the plaque and artery vessel wall were elastically recoiled.

The second step: diastolic-systolic cyclic loading. After expanding in the stenosis coronary artery, the stent undergoes pulsatile loading that arisen from oscillation of the internal blood pressure. This phenomenon was modeled by applying the loading boundary conditions illustrated in Fig. 5. In particular, the internal blood pressure oscillated between 80 mmHg and 120 mmHg from diastole to systole and was applied to the inner surfaces of the stent. Diastolic heart filling and systolic contraction steps were simulated for three cardiac cycles (Fig. 5). It was believed that the constitutive model response get stabilized after three loading cycles [22]. The pressures of the three loading cycles were performed to simulate the effect on the material model by using ABAQUS. These cardiac cycles simulation were conducted to determine the alternating maximum strain induced in the stent as a result of these cyclic loading patterns, and the pressure loading was based on Food and Drug Administration (FDA) test requirements [23]. It was also recommended by the FDA that Goodman fatigue life analysis be used to determine the FSF subjected to physiologic loading up to $4*10^8$ fatigue cycles.

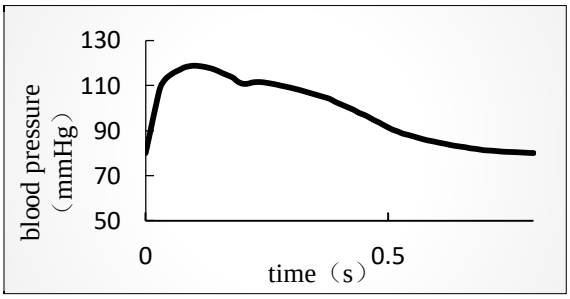


Figure 5. The blood pressure wave form in a cardiac cycle

2.5 Establishment of an idealized model as the control group

A 3D model for the stenosis coronary artery with the idealized structure was established, and the same plaque and vessel material properties were defined in the realistic model, as shown in Fig. 6. The geometry of idealized model is obviously uniform and symmetric, and other information and boundary conditions are the same in the realistic coronary model. The mechanical properties in the different models were compared, and the effect of realistic geometry of a plaque on the stent mechanical behavior was explored.

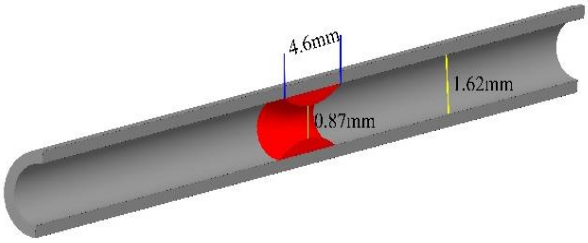
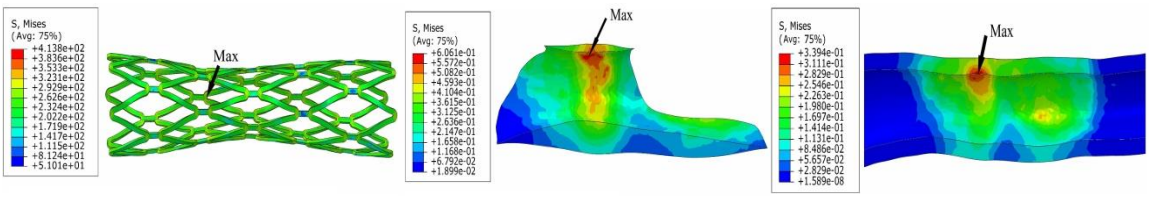


Figure 6. The geometry of the idealized model

3. Results

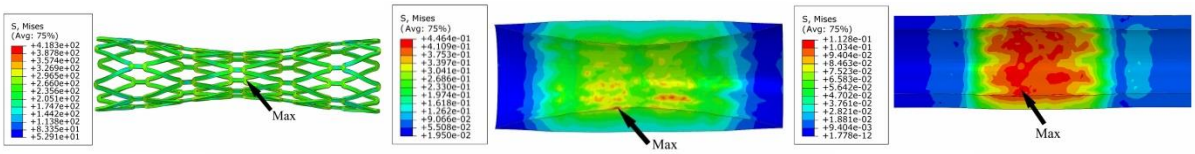
3.1 Analyses of mechanical properties (stress/strain)

At the end of stent expansion, the deformation loading was deflated. The plaque and artery vessel wall were elastically recoiled, and consequently the stent protrusion moved towards the center of the lumen. The presence of plaque generated a greater resistance to the vessel. In order to compare the effect of two different models on the mechanical properties, the stress distributions of the stent, plaque and inner vessel wall were shown in Fig. 7 and Fig. 8, respectively. The stress distribution of idealized model was symmetric (Fig. 8). However, the results of the realistic model were completely different (Fig. 7).



a b c

Figure 7. The stress distribution of complete expansion stage in a realistic model



a b c

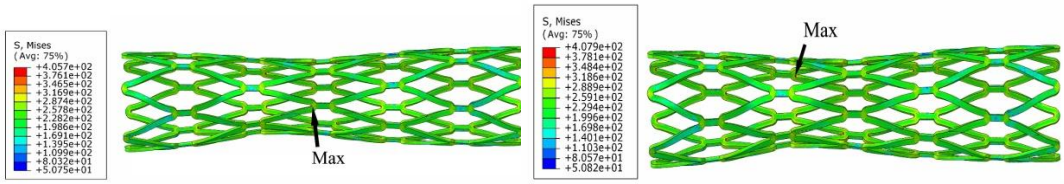
Figure 8. The stress distribution of complete expansion stage in an idealized model

In Fig.7a, the higher stress region was observed in the stent struts close to the calcified plaque, and the maximum stress was 413.8 MPa localized in the bend of the crown, where the dangerous position in the stent structure must be located. The minimal stressed regions were located in the middle position of the bridging I-links. While in the idealized model, the maximum stress of the stent was 418.3 MPa (Fig. 8a). Comparing Fig. 7a with Fig. 8a, the position of the maximum stress changed, but was still located in the bend of the crown.

The higher stress region of plaque was located in the stenosis, and there was a correlation between the stress distribution and the shape of the plaque. The stress distributed uniformly along the axis in the idealized model (Fig. 8b), and was different from the stress distribution in the realistic model (Fig. 7b). Quantitatively, the maximum stress in the plaque was 6.06 MPa in the realistic model (Fig. 7b), while that in the idealized model was 4.46 MPa (Fig. 8b). The relative error ratio was about 35.9%. This revealed that the realistic geometric parameters greatly influenced the result of the mechanics.

From the Fig. 7c and the Fig.8c, the difference of stress distribution in the inner vessel wall between the realistic model and idealized model was obvious. In the higher stress area, the stress was mainly 1.13 MPa~3.39 MPa (Fig. 7c), while in the realistic model it was mainly 0.56 MPa~1.13 MPa (Fig. 8c). The maximum stress on the inner vessel wall of the realistic model was 3.39 MPa, while in the idealized model it was 1.13 MPa. The relative error ratio was about 200.9%.

We supposed that a realistic coronary model can provide accurate stress/strain calculations. It is evident that the highest stresses are observed near the connectors between the stent struts (in agreement with the Ref. [17]). In addition, higher stresses were observed in the stent struts close to the calcified plaque region (in agreement with the Ref. [20]). The coronary stenosis model including a realistic plaque can serve as a useful tool for physicians in their diagnosis and intervention surgical decision making process. Therefore, the data we used in the following fatigue analysis were calculated by the realistic model.



a b

Figure 9. The stress distribution in the stent under different pressures

Figure 9 shows the stress distribution under the systolic and diastolic blood pressure using a realistic model. When the systolic pressure was 120 mmHg, the maximum stress was 405.7 MPa, as shown in Fig. 9a. When the diastolic pressure was 80 mmHg, the position of the maximum stress changed, and the maximum stress changed. The maximum stress was 407.9 MPa, as shown in Fig. 9b. The numerical fluctuation of the stress was not obviously compared with the stress during the stent expansion process. It can be seen from the Fig. 9a and Fig. 9b that stress distribution has the same trend under different pressures, and the maximum stress located in the inner bend of crowns, while the minimal stress located in the I-link. In order to analyze the influence of cyclic loading effect, we calculated the fatigue parameters of the stents.

3.2 Fatigue life analysis

Stress values obtained after the second cycle were equal to those obtained after the third cycle; hence, a Goodman life analysis was performed using the oscillating multi-axial stress state obtained during the second cardiac cycle [13].

3.2.1 Goodman diagram for failure prediction

The FDA recommends that the fatigue resistance of the stent to physiologic loading should be determined using Goodman analysis. Due to the lack of exact information on the mechanical behavior of the material used to fabricate the stent, the literature data were used to draw the stent material fatigue limit on the Goodman diagram [24]. In particular, the endurance limit for zero mean stress was assumed to be equal to 186 MPa while ultimate stress was equal to 520 MPa.

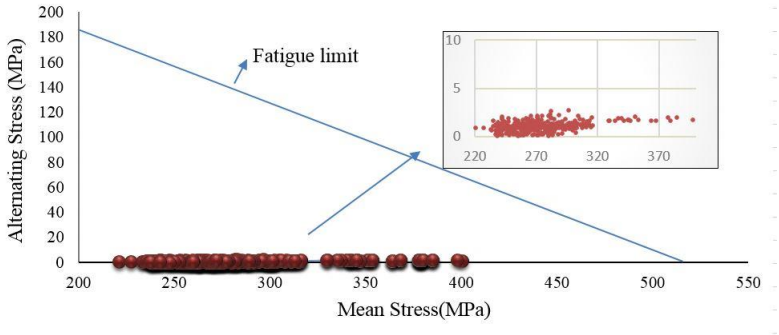


Figure 10. Goodman diagram for the fatigue analysis of stent

The mean stress and alternating stress could be calculated with the stress state of dilation recoil and external loading phase, then all the calculating results on the coordinate figure could be plotted to obtain the Goodman diagram (Fig. 10). As shown in Fig. 10, the x-axis represented the mean stress and the y-axis represents the alternating stress. The slanting line in Fig. 10 was the fatigue limit line (Goodman line), and the intersection point between the fatigue limit and the x-axis was the ultra-tensile strength of 304SS material. The intersection point between the fatigue limit and the y-axis indicated the fatigue endurance strength of the 304SS material. In theory, the closer it was to the fatigue limit, the more dangerous to the structure. As shown in Fig. 10, all the points fallen below the limit line of the material and the stent was able to pass the fatigue life of $4*10^8$ cycles under pulsatile fatigue loading. The stent would not fail with fatigue rupture.

3.2.2 Stent Service Life Prediction

Goodman fatigue life analysis has been widely used for assessing the device fatigue resistance and providing an indication of device chronic durability. The results showing the most dangerous points (Fig. 11) were located in the bend of the crowns, and the data of the most dangerous points were shown in the Tab. 2. This finding is in agreement with the Ref. [16], which means that the maximum strain occurred at the inner surface of the crowns.

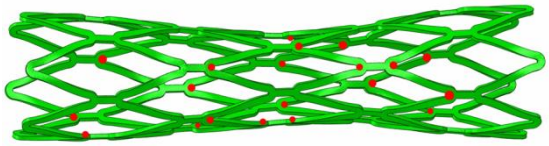


Figure 11. The most dangerous points (red marks) in the stent fatigue prediction

According to the cumulative fatigue damage rate descending order, the predicted cycles to failure of the most dangerous point was $5.32*10^8$ (14 years equivalent), FSF>1, and the cumulative fatigue damage rate was 71.45%. The FDA requires the fatigue testing should include in vitro testing at least $4*10^8$ cycles (10 years equivalent), which indicates the stent is safe in theory.

In this study, the FSF of the most dangerous point was 2.8>1. Compared with the results in Ref. [16] (where FSF=2.06), the stent structure used in this paper was relatively safer. The FSF is defined as the ratio of the strain amplitude against the modified endurance limit, and FSF of less than 1.0 indicates the stent fatigue failure. As shown in the Tab.2, the trend of FSF is similar to that of the fatigue cycle, while the fatigue cycle quantified the proximity of the mean strain and strain amplitude at the given integration points to the Goodman failure line.

In Tab.2, the cumulative fatigue damage rate of the most dangerous point was 71.45%. The referenced methods for calculating cumulative fatigue damage rate (Ref. [24] and Ref. [25]) were quantitative standard for evaluating the fracture rupture. The calculating cumulative fatigue increases with applied load cycles in a cumulative manner, and the stent structure would fracture when the cumulative fatigue damage rate reached 100%. As shown in the Tab. 2, the cumulative fatigue damage rate was all below the 71.45%, and hence the stent structure is safe.

Table 2. The fatigue performance parameters of the most dangerous points

Number	Predicted number of cycles to failure	Fatigue life (a)	FSF	Cumulative fatigue damage rate (%)
1	5.32E+08	14	2.8	71.45
2	5.41E+08	14.25	2.81	70.18
3	5.56E+08	14.63	2.81	68.33

4	7.00E+08	18.42	2.84	54.29
5	7.43E+08	19.55	2.86	51.16
6	7.78E+08	20.47	2.86	48.86
7	8.20E+08	21.58	2.85	46.34
8	8.77E+08	23.07	2.86	43.34
9	8.87E+08	23.34	2.86	42.85
10	8.91E+08	23.44	2.89	42.66
11	9.38E+08	24.7	2.86	40.49
12	9.71E+08	25.56	2.88	39.12
13	1.19E+09	31.27	2.92	31.98
14	1.22E+09	32.15	2.9	31.11
15	1.24E+09	32.64	2.92	30.64
16	1.25E+09	32.97	2.93	30.33
17	1.26E+09	33.17	2.95	30.15
18	1.27E+09	33.4	2.95	29.94
19	1.33E+09	35.04	2.94	28.54
20	1.34E+09	35.39	2.99	28.25

According to the stress analysis and fatigue analysis, we can deduce that maximum stress was always located at the bend of crowns in all simulation procedures, the closest point to the fatigue limit was also located at the inner bend of crowns of stent, and the fatigue analysis results confirmed the mechanical analysis again.

4. Discussion

Numerical models are nowadays a widely recognized useful tool to study the interaction of the stent with tissues and organs [26]. Stent fracture caused by cyclic loading is becoming a key issue in interventional treatment [1]. Clinical studies showed that factors such as stent design, vascular district properties and loading conditions are possible causes of stent fracture. In this study, the interaction of stents, the plaque and the vessel wall in a realistic model and an idealized model were simulated, and the importance of the plaque geometry in the mechanical analysis was confirmed. The finite element method to simulate the effects of cyclic loading experienced by a stented artery during a cardiac cycle was presented. The evaluation was based on the numerical results to build the Goodman Diagram. In the fatigue life prediction, the FSF quantified the proximity of the mean stress and stress amplitude at the given integration points to the Goodman failure line. Cumulative fatigue damage rate analysis played a key role in life prediction of components and structures subjected to field load histories. This method provided an accurate way to predict the stent fatigue life.

This study can surely be improved as from the presence of some limitations. From a modeling point of view, the absence of the balloon has an influence during the inflating process, and the influence of residual stress in the process of crimping was ignored. Future research will be devoted to removing this assumption to have a more realistic description of the stent deployment process. In the fatigue analysis, the stent in the stenosis coronary artery undergoes pulsatile loading conditions that arise from two different phenomena: oscillation of the

internal blood pressure and cardiac wall movement [13], but in this paper we only take the blood pressure into consideration, the next step should focus on the bending fatigue caused by cardiac wall movement. In this paper, material data cited as the strain-based Goodman fatigue life analysis were from several references [24], we should do the material test to get the material data, in order to get more accurate results.

5. Conclusions

The CSF was affected by the complex biomechanical environment caused by the plaques. The mechanical stress/strain was calculated more accurately using a realistic coronary stenosis model. Our results indicated that the maximum stress at all stages was always located at the inner bend of the crowns and the maximum alternating stress was also distributed at the same position, which indicated that the dangerous points were located at the curvature area of the crowns. The effects of cyclic loading on the identification of mechanisms governing stent fatigue rupture were investigated. A better understanding of the failure cycle, the fatigue life, FSF, and cumulative fatigue damage rate may lead to better assessment for the stent fatigue life. This study provided an approach to select and position a stent in order to minimize the risk of fatigue fracture. In fact, this is a general approach and is not limited to 304 stainless steel, and further applications could focus on other metallic materials.

Acknowledgement

The research was supported by National Natural Science Foundation of China (81171107).

References

[1]. Dordoni E, Meoli A, Wu W, et al. (2014) Fatigue behaviour of Nitinol peripheral stents: The role of plaque shape studied with computational structural analyses. *Medical Engineering & Physics* **36**, 842-9.

[2]. Garg S, Serruys PW. (2010) Coronary Stents: Current Status. *Journal of the American College of Cardiology* **56**, S1-S42.

[3]. Li J, Luo Q, Xie Z, et al. (2010) Fatigue life analysis and experimental verification of coronary stent. *Heart & Vessels* **25,** 333-337.

[4]. Satjit A, Mujeeb S, Jason WU, et al. (2010) Stent Fracture in the Coronary and Peripheral Arteries. *Journal of Interventional Cardiology* **23**, 411-419.

[5]. Lim HB, Hur G, Kim SY, et al. (2008) Coronary stent fracture: detection with 64-section multidetector CT angiography in patients and in vitro. *Radiology* **249**, 810-9.

[6]. Nakazawa G, Finn AV, Vorpahl M, et al. (2009) Incidence and Predictors of Drug-Eluting Stent Fracture in Human Coronary Artery. *A Pathologic Analysis. Journal of the American College of Cardiology* **54**:1924-31.

[7]. Scheinert D, Scheinert S, Sax J, et al. (2005) Prevalence and clinical impact of stent fractures after femoropopliteal stenting. J Am CollCardiol. *Journal of the American College of Cardiology* **45**.

[8]. Bessias N, Sfyroeras G, Moulakakis KG, et al. (2005) Renal artery thrombosis caused by stent fracture in a single kidney patient. *Journal of Endovascular Therapy* **12,**516-520.

[9]. Lewitton S, Babaev A. (2008) Superficial Femoral Artery Stent Fracture that Led to Perforation, Hematoma and Deep Venous Thrombosis. *Journal of Invasive Cardiology* **20,** 479-81.

[10]. Auricchio F, Constantinescu A, Conti M, et al. (2015) Fatigue of Metallic Stents: From Clinical Evidence to Computational Analysis. *Annals of Biomedical Engineering*, 1-15.

[11]. Dordoni E, Petrini L, Wu W, et al. (2015) Computational Modeling to Predict Fatigue Behavior of NiTi Stents: What Do We Need? *Journal of Functional Biomaterials* **6**, 299-317.

[12]. James BA, Sire RA. (2010) Fatigue-life assessment and validation techniques for metallic vascular implants. *Biomaterials* **31**, 181-186.

[13]. Morlacchi S, Pennati G, Petrini L, et al. (2014) Influence of plaque calcifications on coronary stent fracture: A numerical fatigue life analysis including cardiac wall movement. *Geological Magazine* **47**, 899-907.

[14]. Ragkousis GE, Curzen N, Bressloff NW. (2014) Simulation of longitudinal stent deformation in a patient-specific coronary artery. *Medical Engineering & Physics* **36**, 467-476.

[15]. Hsiao HM, Chiu YH, Lee KH, et al. (2012) Computational modeling of effects of intravascular stent design on key mechanical and hemodynamic behavior. *Computer-Aided Design* **44**, 757-765.
[16]. Hsiao HM, Wu LW, Yin MT, et al. (2014) Quintupling fatigue resistance of intravascular stents via a simple design concept. *Computational Materials Science* **86**, 57-63.
[17]. Gijsen FJ, Migliavacca F, Schievano S, et al. (2008) Simulation of stent deployment in a realistic human coronary artery. *Biomedical Engineering Online* **7**, 23-23.
[18]. Petrini L, Trotta A, Dordoni E, et al. (2015) A Computational Approach for the Prediction of Fatigue Behaviour in Peripheral Stents: Application to a Clinical Case. *Annals of Biomedical Engineering*, 1-12.
[19]. Marrey RV, Burgermeister R, Grishaber RB, et al. (2006) Fatigue and life prediction for cobalt-chromium stents: A fracture mechanics analysis. *Biomaterials* **27**, 1988-2000.
[20]. Schievano S, Taylor A M, Capelli C, et al. (2010) Patient specific finite element analysis results in more accurate prediction of stent fractures: Application to percutaneous pulmonary valve implantation. *Journal of Biomechanics* **43**, 687-693.
[21]. Yang Z, Zhang HP, Marder M. (2008) Dynamics of static friction between steel and silicon. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 13264-8.
[22]. Dordoni E, Meoli A, Wu W, et al. (2014) Fatigue behaviour of Nitinol peripheral stents: The role of plaque shape studied with computational structural analyses. *Medical Engineering & Physics* **36**, 842-9.
[23]. The guidance for industry and FDA staff: Non-clinical tests and recommended labeling for intravascular stents and associated delivery systems[EB/OL]. Http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm071863.htm,2010-04-18.
[24]. Li HX, Gao YH, Wang XC. (2013) Optimization of stent-balloon system based on surrogate modeling technique. *Journal of Medical Biomechanics*
[25]. Pelton AR, Schroeder V, Mitchell MR, et al. (2008) Fatigue and durability of Nitinol stents. *Journal of the Mechanical Behavior of Biomedical Materials* **1**, 153-64.
[26]. Morlacchi S, Migliavacca F. (2013) Modeling Stented Coronary Arteries: Where We are, Where to Go. *Annals of Biomedical Engineering* **41**, 1428-1444.

A finite element method used for contact analysis of rolling bearings

†*S. Li[1] and M. Motooka[1]

[1]Interdisciplinary Graduate School of Science and Engineering, Shimane University, Japan

*Presenting author: shutingli@ecs.shimane-u.ac.jp
†Corresponding author: shutingli@ecs.shimane-u.ac.jp

Abstract

This paper deals with contact analysis method of rolling bearings. A mathematical model is presented for the purpose of contact analysis of rolling bearings based on the principle of mathematical programming method at the first. Then, three-dimensional (3D), finite element method (FEM) is introduced in the mathematical model to calculate deformation influence coefficients and gaps of assumed pairs of contact points between contact surfaces. Special software is developed to realize the procedures of contact analysis. With the help of the developed software, contact analyses are conducted for a deep groove ball bearing and a cylindrical roller bearing. In the case of the ball bearing, it is found that the calculated contact pressures on ball surface are more reasonable and accurate than the ones obtained by commercial CAE software. In the case of the roller bearing, it is found that edge-loads (non-Hertz contact that cannot be analyzed by Hertz theory) on the two ends of the roller surface are analyzed successfully by the method presented in this paper when the rollers are not crowned longitudinally. It is also found that the edge-loads are disappeared and the contact pressure becomes uniform distribution on the roller surface when the rollers are crowned on the two ends using Johson-Gohar [1] curve. Since the calculated results given in this paper cannot be obtained by using commercial CAE software and other numeric methods, the mathematical model and numeric method presented in this paper have a great practical meaning in engineering design and calculations of the rolling bearings.

Keywords : Finite element method, Contact analysis, Mathematical programing method, Rolling bearing

1. Introduction

It is a very important thing for machine designers to evaluate lifetime and radial rigidity of the rolling bearings when they decide to use. Unfortunately, it is still a difficult thing to evaluate contact strength and lifetime of the rolling bearings accurately in theory. Also, it is still a difficult thing to calculate contact pressure and radial rigidity of the roller bearings accurately in theory. This is because there have been still some unsolved problems remained in strength and performance analyses of the rolling bearings, though it is a very long history to use the rolling bearings in various kinds of machines.

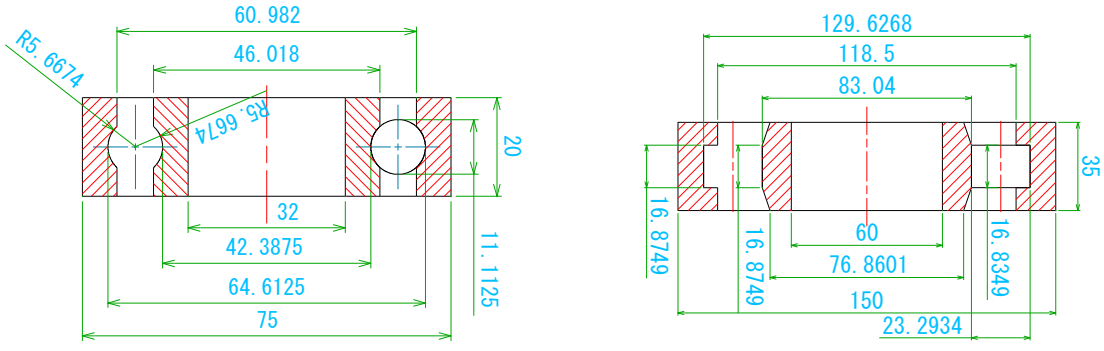
In the case of the ball bearings, usually, Hertz theory [2] is used to calculate the contact pressure and radial rigidity of the ball bearings. Since Hertz theory can only consider local deformation of contact areas of the ball bearings and the total structural deformation of the ball, outer ring and inner ring as well as housings cannot be included, Hertz theory has a limit in engineering calculations when the total structural deformation mentioned above is considered. In the case of the roller bearings, since edge-loads exist between the two contact surfaces, contact problem

of the roller bearings belongs to a non-Hertz contact problem and Hertz theory cannot be used for contact analysis of the roller bearings. In order to solve the contact problem of the roller bearings, an approximate contact model of using a roller contacting a surface of infinite length was used [3]-[5]. But since the surface of infinite length also cannot consider the effects of structural sizes and shapes of the inner and outer rings, this contact model is also not so accurate for contact analysis of the roller bearings. Finally, FEM was suggested to do contact analysis of the roller bearings [3][6]. Indeed, FEM is a very practical method for structural analysis and very successful in many kinds of engineering calculations. But, unfortunately, this method is not so successful in contact analysis of machines and machine elements. The problem of using some commercial CAE software in contact analysis of the bearings shall be introduced in Section 3 of this paper.

This paper tries to present a new FEM that can conduct contact analysis of the rolling bearings accurately. Firstly, a new mathematical model is presented in this paper for contact analysis of the rolling bearings based on the principle of the mathematical programming method. Then, 3D, FEM is introduced in the mathematical model to calculate deformation influence coefficients and gaps of the assumed pairs of contact points on the contact surfaces. Special software is developed through efforts of many years. With the help of the developed software, contact analyses of a deep groove ball bearing and a cylindrical roller bearing are conducted successfully. Calculation results shows that the special software can calculate more reasonable and accurate contact pressure distribution of the rolling bearings than the commercial software SolidWorks and some other finite element method [10] stated in this paper. The maximum contact pressure and radial contact rigidity of the ball bearing are also analyzed with Hertz theory. It is found that the results obtained by the special software are similar to the results obtained by Hertz theory, but they are not exactly equal. The total structural deformation of the ball, outer ring and inner ring can be thought to be the main reason to result in the difference between the method presented in this paper and Hertz theory. This assumption shall be confirmed experimentally in the near future.

2. Structural dimensions of the bearings used as research objects

Structures and dimensions of the rolling bearings used as research objects are illustrated in Fig.1. In Fig.1, (a) and (b) are a deep groove ball bearing (type number 6332) and a cylindrical roller bearing (type number NU412) respectively. They are made by NTN, a Japanese bearing company [8]. Contact analyses are conducted for them with commercial software SolidWorks and special FEM software developed in this paper respectively. Hertz theory is also used to calculate contact pressure and radial rigidity of the ball bearing in this paper in order to make a comparison with the results obtained by the special FEM software.



(a) Ball bearing (b) Roller bearing

Figure 1. Structures of the ball and roller bearings used as research objects

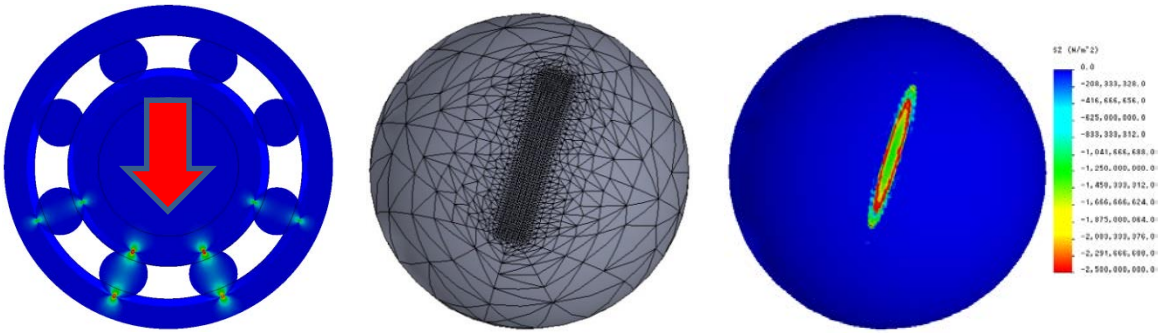
3. Problems of some commercial software used for contact analysis

As stated above, some commercial software is very successful in many kinds of engineering calculations, but it is not so successful in contact analysis of machines and machine elements. In order to make this problem clear in this paper, some results of using the commercial software SolidWorks for contact analysis of the ball and roller bearings are introduced in the following.

In SolidWorks software, there is a function called SolidWorks Simulation that can be used to do CAE simulations. SolidWorks Simulation is originated from the famous CAE software COSMOSWorks [9]. In the COSMOSWorks, contact analysis function is also included. So, this paper uses this function to conduct contact analyses of the ball and roller bearings given in Fig. 1. When the contact analyses are conducted, outside surfaces of the bearing outer rings are fixed as boundary conditions and a radial load P is applied on the inside surfaces of the bearing inner rings as shown in Fig. 2(a) and 3(a) through bearing shafts that are inserted into the central holes of the inner rings. The bearing shaft and the inner ring are unified as one elastic body in the analyses.

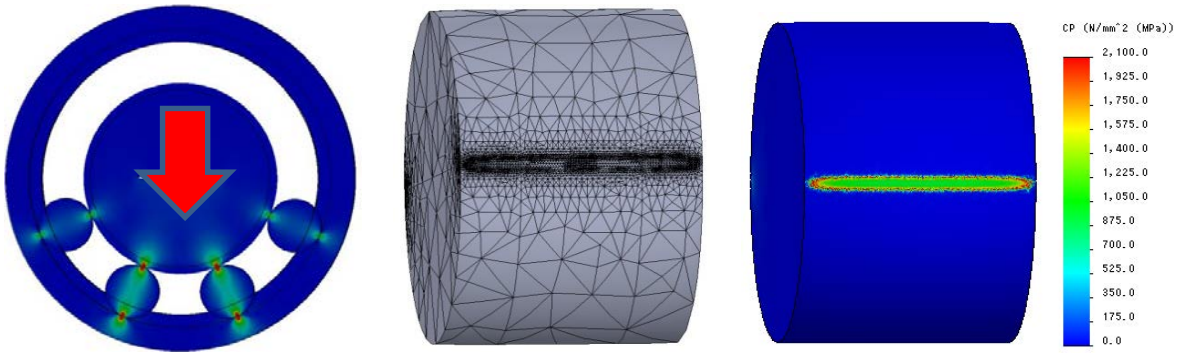
Calculation results of the ball bearing are given in Fig. 2. In Fig. 2, (a) is a contour map of calculated Von Mises stresses distributed on the section that goes through the ball center and is perpendicular to the bearing axis. Fig. 2(a) indicates that only four balls at the lower part of the bearing are in contact with the raceways of the inner and outer rings. Fig. 2(b) is FEM mesh-diving pattern of the ball. As shown in Fig. 2(b), contact areas of the ball surface are fine mesh-divided in order to ensure high calculation accuracy. Of course, the contact areas of the raceways of the outer and inner rings are also fine mesh-divided responsively. Fig. 2(c) is a contour map of calculated contact pressure distributed on the ball surface. From Fig. 2(c), it is found that though contact pattern of the ball takes the shape of an elliptical contact, the maximum contact pressure is not located at the center of the elliptical area. It distributes along a closed elliptical curve as shown in Fig .2(c) illustrated in the red line. Also the maximum contact pressure is calculated to be about twice the value calculated by Hertz theory. This means that SolidWorks cannot calculate contact pressure distribution and the maximum contact pressure correctly if it is used to conduct contact analysis of the ball bearings.

Calculation results of the roller bearing are given in Fig. 3. In Fig. 3, (a) is a contour map of calculated Von Mises stresses distributed on the section going through the center point of the roller width and being perpendicular to the bearing axis. Fig. 3(a) also indicates that only four rollers at the lower part of the bearing contact the raceways of the inner and outer rings. Fig. 3(b) is FEM mesh-diving pattern of the roller. As shown in Fig. 3(b), contact areas of the roller surface are fine mesh-divided. Of course, contact areas of the raceways are also fine mesh-divided responsively. Fig. 3(c) is a contour map of calculated contact pressure distributed on the roller surface. Fig. 3(c) indicates that contact pattern of the roller bearing takes the shape of an elliptical contact (the roller is crowned longitudinally using Johson-Gohar curve [1]), but the maximum contact pressure is also not located at the center of the contact area. It distributes along a closed elliptical curve as shown in Fig .3(c) illustrated in the red line. It means that there is also a problem existing for the roller bearing that SolidWorks cannot calculate contact pressure distribution of the roller bearing correctly. Per a long-time experience of the author on CAE analysis using commercial CAE software, it is found that not only SolidWorks, but also some other commercial software, such as ANSYS and ADINA, have the similar problem like SolidWorks that they cannot calculate contact pressure distribution accurately when they are used to do contact analyses of machines or machine elements.



(a) Von Mises stresses (b) FEM mesh-dividing (c) Contact pressure

Figure 2. Calculation results of the ball bearing using SOLIDWORKS

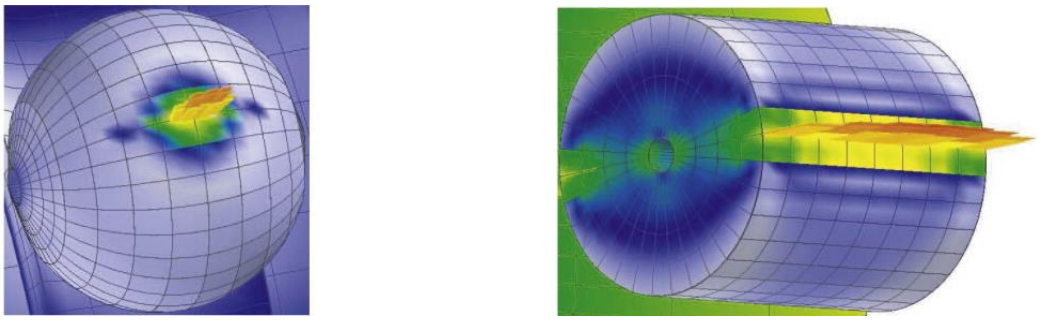


(a) Von Mises stresses (b) FEM Mesh-dividing (c) Contact pressure

Figure 3. Calculation results of the roller bearing using SOLIDWORKS

Guo and Parker [10] also conducted contact analyses for a deep groove ball bearing and a cylindrical roller bearing using FEM and specially developed software. Calculation results obtained by Guo and Parker are given in Fig. 4. Fig. 4(a) and (b) are calculated contact loads distributed on the ball and the roller surfaces respectively. From Fig. 4, it is found that quite rough results were obtained in Guo and Parker's research.

Based on the results mentioned above, it can be understood well that it is a quite difficult thing to conduct loaded bearing contact analysis and get correct contact pressure distribution of the bearings using available commercial CAE software and finite element techniques at the present situation. So, it is necessary to develop a new method and technology that can conduct contact analysis of the bearings correctly. This paper tries to present a new mathematical model and numeric method for contact analysis of the rolling bearings.

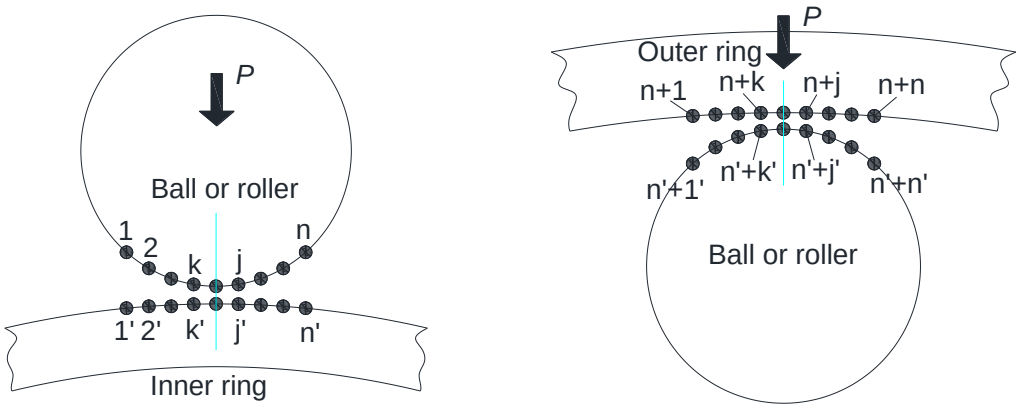


(a) Load distribution on ball surface (b) Load distribution on roller surface

Figure 4. Guo and Parker's results on bearing contact analysis [10]

4. A new mathematical model and numeric method for bearing contact analysis

4.1 Principle used for contact analysis of the rolling bearings



(a) Contact of ball with inner ring (b) Contact of ball with outer ring

Figure 5. Mathematical model used for contact analysis of rolling bearings

Models used for contact analysis of the ball and roller bearings are given in Fig. 5. In Fig. 5, (a) is used to stand for the contact of a ball (or roller) with the inner ring raceway and (b) is used to stand for the contact of the ball (or roller) with the outer ring raceway. It is assumed that an external load *P* (usually, equals to radial load of the bearings) is applied on the bearings in vertical direction as shown in Fig. 5. It is assumed that only elastic deformation occurred in the contact problem of the rolling bearings.

In Fig. 5(a), firstly, a lot of pairs of contact points, such as (1-1'), (2-2'), …, (j-j'), …, (k-k') and (n-n'), are assumed on the contact surfaces of the ball (roller) and the inner ring raceway along the vertical direction. In Fig. 5(a), 1,2, …, j, …, k and n are the assumed points on the contact area of the ball (roller) surface and 1',2', …, j', …, k' and n' are the responsive points on the contact area of the inner ring raceway. The common normal lines of these assumed pairs of contact points are parallel to the vertical direction and pass through the pairs of contact points. It is assumed that these pairs of contact points have possibility to come into contact when the external load *P* is applied.

4.1.1 Deformation compatibility relationship of the pairs of contact points

As shown in Fig. 5(a), for an optional pair of contact points (k-k'), F_k is used to denote the contact force between the pair of contact points (k-k'). Of course, direction of F_k is along the direction of its common normal line. Also, F_j is the contact force between the pair of contact points (j-j') along its common normal line. Gaps between the pairs (j-j'), (k-k') and (n-n') are denoted as ε_j, ε_k and ε_n respectively. Relative deformation of the ball (roller) relative to the inner ring along the vertical direction is denoted as δ_1. Elastic deformation of the pair of contact points (k-k') along its common normal line direction are denoted as ω_k and $\omega_{k'}$ respectively. If (k-k') comes into contact after *P* is applied, $(\omega_k + \omega_{k'} + \varepsilon_k)$, the amount of the deformation and the gap of the pair of points (k-k'), shall be equal to the relative deformation δ_1. But, if (k-k') doesn't come into contact, $(\omega_k + \omega_{k'} + \varepsilon_k)$ shall be greater than δ_1. These relationships are called deformation compatibility relationships and they can be expressed with Eq. (1) and (2) in the following. Eq. (1) and (2) can be summarized into Eq. (3).

$$\omega_k + \omega_{k'} + \varepsilon_k - \delta_1 > 0 \quad \text{(Not contact)} \tag{1}$$

$$\omega_k + \omega_{k'} + \varepsilon_k - \delta_1 = 0 \quad \text{(Contact)} \tag{2}$$

$$\omega_k + \omega_{k'} + \varepsilon_k - \delta_1 \geq 0 \qquad (k = 1, 2, \dots, n) \tag{3}$$

Eq. (3) is not only suitable for an optional pair of contact points (k-k'), but also suitable for all the pairs of contact points assumed on the contact surfaces of the ball (roller) with the inner ring raceway. In Eq. (3), n is the total number of the assumed pairs of contact points.
Since the elastic deformation ω_k and $\omega_{k'}$ can be expressed with deformation influence coefficients a_{kj} and $a_{k'j'}$, then Eq. (4) and (5) can be obtained. If Eq. (4) and (5) are substituted into Eq. (3), then Eq. (6) can be obtained.

$$\omega_k = \sum_{j=1}^{n} a_{kj} F_j \tag{4}$$

$$\omega_{k'} = \sum_{j=1}^{n} a_{k'j'} F_j \tag{5}$$

$$\sum_{j=1}^{n} \left[a_{kj} + a_{k'j'}\right] \times F_j + \varepsilon_k - \delta_1 \geq 0 \qquad (k = 1, 2, \dots, n) \tag{6}$$

Where, a_{kj} and $a_{k'j'}$ are deformation influence coefficients of the pairs of contact points along their common normal lines. a_{kj} and $a_{k'j'}$ can be calculated through 3D, finite element analysis.

4.1.2 Load equilibrium relationship of the pairs of contact points

Except for the deformation compatibility relationship as shown in Eq. (6), a load equilibrium relationship of the pairs of contact points can also be built as given in Eq. (7). Where, P is the external load applied on the bearing.

$$\sum_{k=1}^{n} F_k = P \qquad (k = 1, 2, \dots, n) \tag{7}$$

In the case of the ball (roller) contacting the outer ring raceway as illustrated in Fig. 5(b), the deformation compatibility relationship and load equilibrium relationship can also be built for the assumed pairs of contact points on the contact surfaces of the ball (roller) with the outer ring raceway in the same way. Eq. (8) and (9) are the two relationships for the pairs of contact points on the contact surfaces of the ball (roller) with the outer ring raceway.

$$\sum_{j=n+1}^{n+n} \left[a_{kj} + a_{k'j'}\right] \times F_j + \varepsilon_k - \delta_2 \geq 0 \qquad (k = n+1, \dots, n+n) \tag{8}$$

$$\sum_{k=n+1}^{n+n} F_k = P \qquad (k = n+1, \dots, n+n) \tag{9}$$

Where, δ_2 is the relative deformation of the ball (roller) relative to the outer ring along the vertical direction. By adding Eq. (6) and Eq. (8) together, then Eq. (10) can be obtained. Also, by adding Eq. (7) and Eq. (9) together, then Eq. (11) can be obtained. Where, $\delta = \delta_1 + \delta_2$ is the total relative deformation among the outer ring, the ball (roller) and the inner ring along the vertical direction.

$$\sum_{j=1}^{2n} \left[a_{kj} + a_{k'j'}\right] \times F_j + \varepsilon_k - \delta \geq 0 \qquad (k = 1,2, \dots, 2n) \tag{10}$$

$$\sum_{k=1}^{2n} F_k = 2P \qquad (k = 1,2, \dots, 2n) \tag{11}$$

If Eq. (10) is written into a matrix expression, then Eq. (12) can be obtained.

$$[S]\{F\} + \{\varepsilon\} - \delta\{e\} \geq \{0\} \tag{12}$$

Where,

$$[\mathrm{S}] = \begin{bmatrix} [S1] & [0] \\ [0] & [S2] \end{bmatrix}$$

$$[S1] = \left[S_{kj}\right] = \left[a_{kj} + a_{k'j'}\right] \ , k = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, n$$

$$[S2] = \left[S_{kj}\right] = \left[a_{kj} + a_{k'j'}\right] \ , k = n+1, n+2, \dots, n+n; j = n+1, n+2, \dots, n+n$$

$$[0] = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

$$\{F\} = \{F_1, F_2, \dots, F_k, \dots, F_{n+n}\}^T$$

$$\{\varepsilon\} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k, \dots, \varepsilon_{n+n}\}^T$$

$$\{e\} = \{1, 1, \dots, 1\}^T$$

$$\{0\} = \{0, 0, \dots, 0\}^T$$

Also, If Eq. (11) is written into a matrix form, Eq. (13) can be obtained.

$$\{e\}^T\{F\} = 2P \tag{13}$$

Eq. (12) and (13) can be used as constrain conditions in contact analysis of the bearings to identify which pair of contact points is in contact and which pair is not in contact when the external load P is applied. Contact problem of bearings can be explained as looking for the contact force F_k $(k = 1,2,3, \dots, 2n)$ of the pairs of contact points that must satisfy Eq. (12) and (13) under the conditions of knowing the deformation influence coefficients a_{kj}, $a_{k'j'}$, the gaps ε_k and the external load P in advance.

4.2. A new mathematical model used for contact Analysis of the rolling bearings

A new mathematical model is built to solve Eq. (12) and Eq. (13) based on the principle of the mathematical programming method [11]-[12] as follows. Since Eq. (12) is an inequality constraint equation that may be strictly positive or identically zero, it can be transformed into

an equality constraint equation by introducing a so-called slack variable $\{Y\}$ (consists of positive variables) based on the principle of the modified simplex method [11]-[12]. Then Eq. (14) and (15) can be obtained.

$$[S]\{F\} + \{\varepsilon\} - \delta\{e\} - [I]\{Y\} = \{0\} \tag{14}$$

or

$$-[S]\{F\} + \delta\{e\} + [I]\{Y\} = \{\varepsilon\} \tag{15}$$

Where

$\{Y\} = \{Y_1, Y_2, \dots, Y_k, \dots, Y_{2n}\}^T$ (Slack variables)

$[I]$= a unit matrix of 2n×2n

Then the two equality constraint equations of Eq. (13) and (15) are obtained. The next task is to make an objective function Z that is necessary to build the mathematical programming model. The objective function Z can be made artificially through introducing some positive variables X_{2n+1}, X_{2n+2}, ..., X_{2n+2n}, $X_{2n+2n+1}$ (usually called artificial variables) to every constrain equation based on the principle of the modified simplex method [11]-[12]. Then the mathematical programming model used for contact analysis of the bearings can be made as follows.

Mathematical programming model used for bearing contact analysis

Objective Function:

$$Z = X_{2n+1} + X_{2n+2} + \cdots + X_{2n+2n} + X_{2n+2n+1} \tag{16}$$

Constraint Conditions:

$$-[S]\{F\} + \delta\{e\} + [I]\{Y\} + [I]\{Z'\} = \{\varepsilon\} \tag{17}$$

$$\{e\}^T\{F\} + X_{2n+2n+1} = 2P \tag{18}$$

Where,

$\{Z'\} = \{X_{2n+1}, X_{2n+2}, \dots, X_{2n+2n}\}^T$ (Artificial variables)

$$[\mathrm{S}] = \begin{bmatrix} [S1] & [0] \\ [0] & [S2] \end{bmatrix}$$

$[S1] = [S_{kj}]$, $k = 1, 2, 3, \dots, n; j = 1, 2, 3, \dots, n$

$[S2] = [S_{kj}]$, $k = n+1, n+2, \dots, n+n; j = n+1, n+2, \dots, n+n$

$$[0] = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \qquad (n \times n)$$

$\{F\} = \{F_1, F_2, \dots, F_k, \dots, F_{2n}\}^T$

$\{Y\} = \{Y_1, Y_2, \dots, Y_k, \dots, Y_{2n}\}^T$ (Slack variable)

$\{\varepsilon\} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k, \dots, \varepsilon_{2n}\}^T$

$\{e\} = \{1, 1, \dots, 1\}^T$

$F_k \geq 0, Y_k \geq 0, \theta_k \geq 0, \theta \geq 0, k = 1, 2, \dots, 2n$

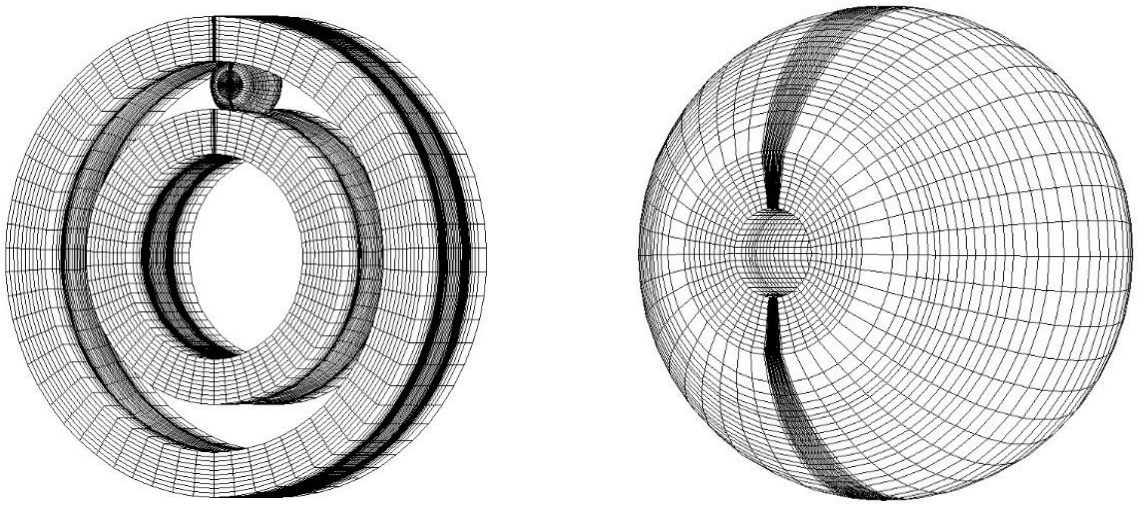
$X_{2n+m} \geq 0, m = 1, 2, \dots, 2n+1$

The contact force F_k and the total radial deformation δ can be calculated by minimizing the objective function Z in Eq. (16) under the constrain conditions of Eq. (17) and (18) using the modified simplex method [11]-[12].

4.3 Software development

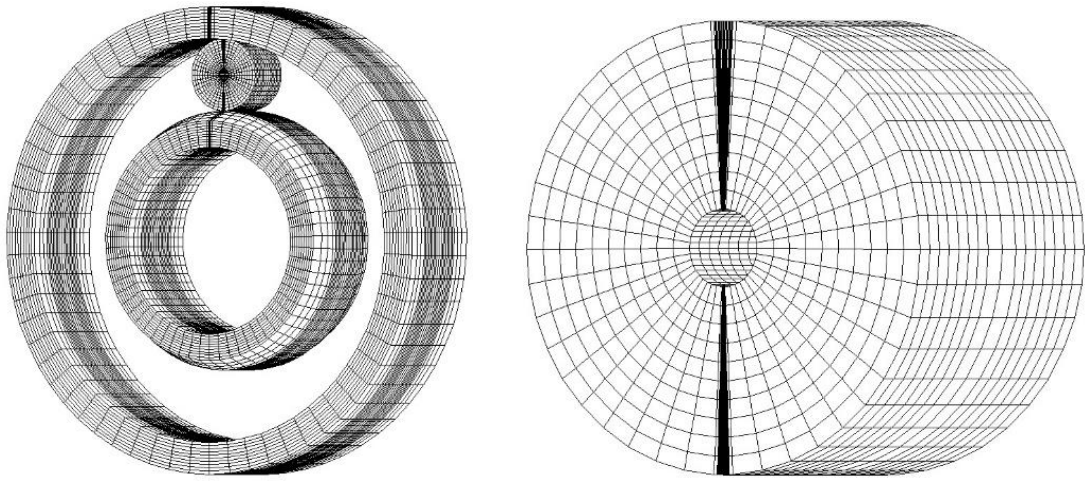
Software development is conducted to realize procedures of the bearing contact analysis. Firstly, 3D, FEM method is used to calculate the deformation influence coefficients a_{kj},

$a_{k'j'}$ that are necessary to form the [S1] and [S2] in the matrix [S]. Special FEM software is developed using Super-parametric hexahedron solid element, which has 8 nodes at the corner and 3 nodes inside the element [13]. FEM models and mesh-dividing patterns of the ball and roller bearings are given in Fig. 6 and 7 respectively. Fig. 6(a) and Fig. 7(a) are FEM models and mesh-dividing patterns of the whole ball and roller bearings respectively. Fig. 6(b) and Fig. 7(b) are enlarged views of the mesh-dividing patterns of the ball and roller only respectively. As shown in Fig. 6 and Fig. 7, meshes on the contact areas of the outer rings, ball (roller) and inner rings are fine divided in order to ensure high calculation accuracy of FEA.



(a) Ball bearing (b) Ball

Figure 6. Mesh-dividing patterns of the ball bearing



(c) Roller bearing (d) Roller

Figure 7. Mesh-dividing patterns of the roller bearing

Special software development is also conducted to realize the procedures of the mathematical programming for the bearing contact analysis after the deformation influence coefficients are available by FEA. Then, contact load $\{F\}$ and the total radial deformation δ can be available after the mathematical programming is conducted with the help of the developed software. Contact pressure distribution can be calculated after the contact load $\{F\}$ is available through calculating the contact load distributed on unit contact area. Also, radial rigidity K of the bearings can be calculated through this expression $K = P/\delta$. Calculation results are introduced in the following.

5. Calculation results and discussions

5.1 Contact pressure distribution

Firstly, loaded bearing contact analysis is conducted for the deep groove ball bearing as shown in Fig. 1(a) with the developed FEM software. Fig. 8(a) and 9(b) are calculated contact pressure distributed on the ball surfaces. Fig. 8(a) is a contour map of the contact pressure between the ball and the outer ring raceway. Fig. 8(b) is a contour map of the contact pressure between the ball and the inner ring raceway. The external load P is equal to 40kN when the contact analysis is conducted. From Fig. 8, it is found that the contact pressure on the ball surfaces is calculated to be beautiful elliptical distribution and the maximum contact pressure point is located at the center of the contact areas. These results are more reasonable than the results obtained by SolidWorks and the reference [10] as given in Fig. 2 and Fig. 4(a). It is also found that the maximum contact pressure on the upper part of the contact surfaces (the ball with the outer ring raceway) is a little smaller than the one on the lower part of the contact surfaces (the ball with the inner ring raceway). This is because the radius of curvature of the inner ring raceway is smaller than that of the outer ring raceway. The smaller radius of curvature of the contact surface shall bring greater contact stress based on Hertz theory.

The maximum contact pressure of the ball bearing is also calculated with Hertz theory. Fig. 9 is a comparison of the maximum contact pressure between the developed FEM software and Hertz theory. In Fig. 9, abscissas are radial load P applied on the bearing and the ordinates are the maximum contact pressure on the ball surface. Fig. 9(a) is the maximum contact pressure between the ball and the inner ring raceway and Fig. 9(b) is the one between the ball and the outer ring raceway. Fig. 9(a) indicates that the results obtained by the FEM software are smaller than the ones obtained by Hertz theory. Fig. 9(b) indicates that the results obtained by the FEM software are greater than the ones obtained by Hertz theory. The difference between the two methods can be thought to be the effect of the total structural deformation of the bearing. As it has been stated above, Hertz theory cannot consider of the total structural deformation of the ball, inner and outer ring while FEM can consider the total structural deformation of the bearing. Secondly, loaded bearing contact analysis is conducted for the cylindrical roller bearing as shown in Fig. 1(b) with the developed FEM software. Fig. 10(a) and 10(b) are calculated contact pressure distributed on the roller surfaces when the roller is not crowned. Fig. 10(a) is the contact pressure between the roller and the outer ring raceway. Fig. 10(b) is the contact pressure between the roller and the inner ring raceway. The external load P is equal to 4kN when the contact analysis is conducted. From Fig. 10, it is found that contact pressure on the roller surface is calculated to be uniform distribution along axial direction of the roller except for the two end areas of the roller. It is also found that the edge-loads are calculated on the two end areas of the roller beautifully. It is a big success or progress that the developed FEM software can analyse edge-loads of an uncrowned roller bearing successfully. By comparing Fig. 10 with Fig. 4(b), it is found that the mathematical model and numeric method presented in this paper can calculate more reasonable results than method given in reference [10]. The fact is that it is still a difficult thing for some commercial CAE software to analyse the edge-loads correctly at the present situation.

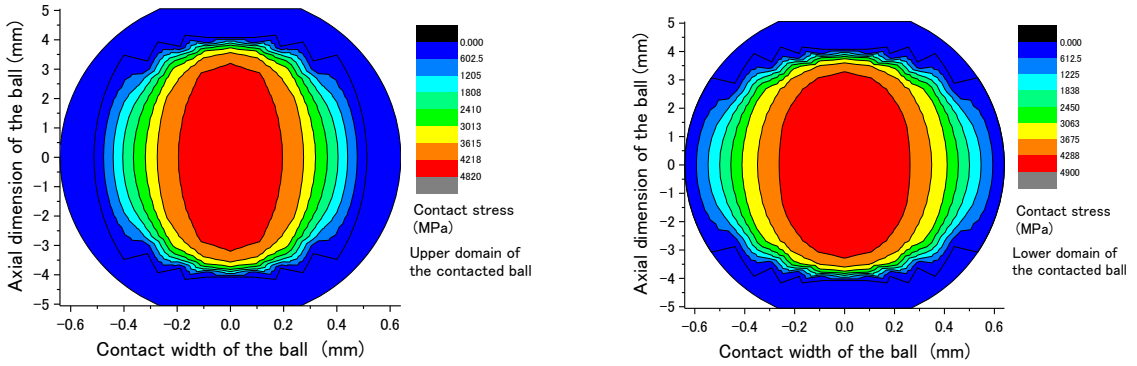
Loaded bearing contact analysis is conducted also for the roller bearing when the roller is crowned on the two end areas using Johson-Gohar curve [1] as given in Eq. (19). Fig. 11(a) and 11(b) are imagines of the roller before and after crowned. Calculation results for the crowned roller bearing are given in Fig. 12. Fig. 12(a) is the contact pressure distribution between the crowned roller and the outer ring raceway. Fig. 12(b) is the contact pressure distribution

between the crowned roller and the inner ring raceway. From Fig. 12, it is found that the edge-loads disappeared on the two end areas of the roller and contact pressure becomes uniform distribution longitudinally in comparison with the results given in Fig. 10. It is also found that the maximum contact pressures are reduced about 17% and 21% when the roller is crowned by comparing Fig. 12(a) with Fig. 10(a) and Fig. 12(b) with Fig. 10(b). The results in Fig. 12 indicate that Johson-Gohar curve is a very nice curve to be used as crowning curve for the roller bearings. It can reduce edge-loads greatly and bring the roller bearing a uniform contact pressure distribution.

$$q(x) = \frac{2P}{\pi l E'} ln \frac{1}{1-(1-0.3033b/a)(2x/l)^2} \tag{19}$$

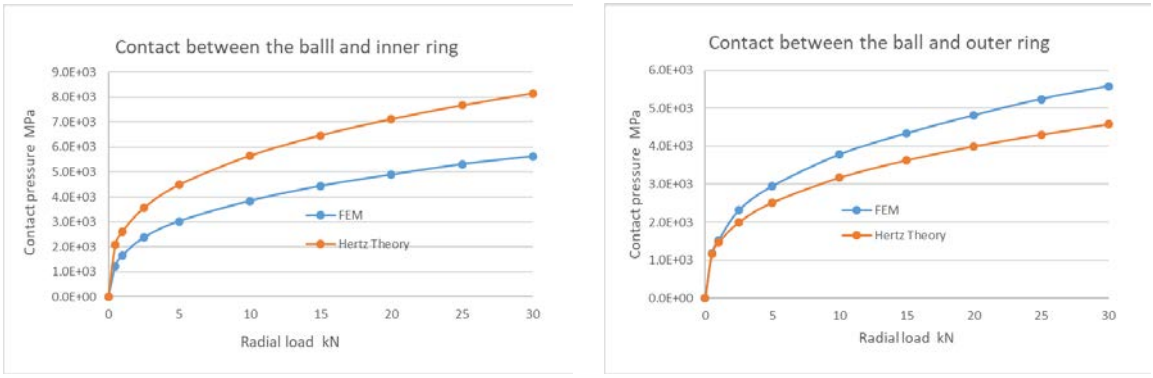
Where, l is an effective contact length of the roller and a is a half of the effective contact length l. b is a half width of the contact. E is Young's modulus and ν is Poisson's ratio. E' is equivalent Young's modulus that can be obtained by following Eq. (20). P is a load applied on the roller. x is used to stand for longitudinal position of a point along the axial direction. $q(x)$ is used to denote the drop (quantity of crowning) at the position x in the axial direction.

$$E' = \frac{E}{1-\nu^2} \tag{20}$$



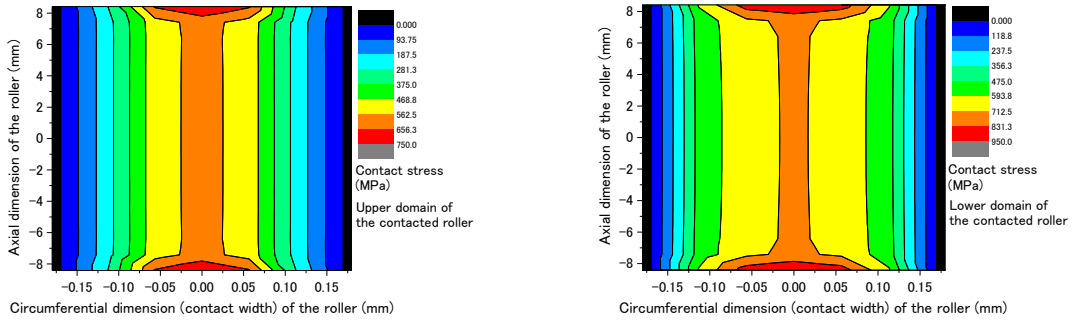
(a) The upper part of the contact domain (b) The lower part of the contact domain

Figure 8. Contour maps of contact stresses distributed on the ball surface



(a) Ball and inner ring (b) Ball and outer ring

Figure 9. Contact pressure comparison between FEM software and Hertz theory

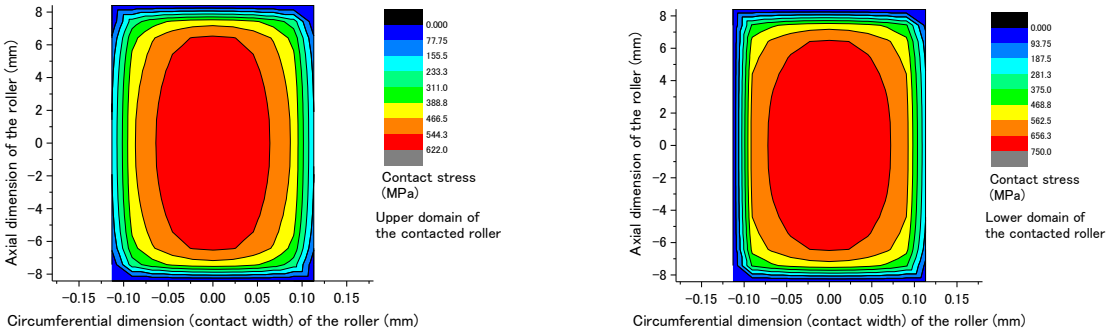


(a) The upper part of the contact domain (b) The lower part of the contact domain

Figure 10. Contour maps of contact stresses distributed on the roller surface

(a) The roller before crowned (b) The roller after crowned

Figure 11. Crowning on the two ends of the roller with Johson-Gohar curve



(a) The upper part of the contact domain (b) The lower part of the contact domain

Figure 12. Contour maps of contact stresses distributed on the roller surface

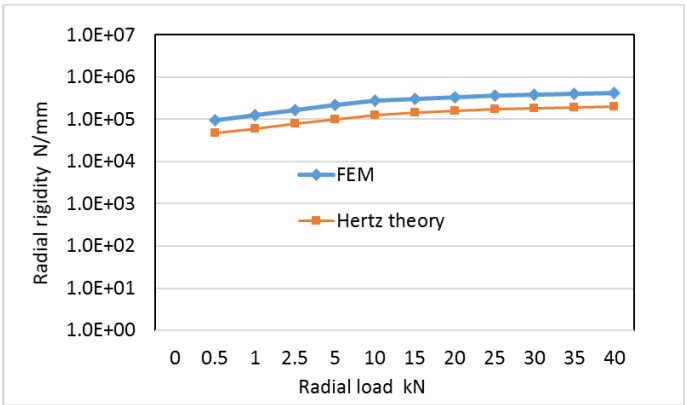


Figure 13. Comparison of radial rigidity of the ball bearing

5.2 Radial rigidity of the ball bearing

In the case of the ball bearings, since Hertz theory can be used to calculate radial rigidity of the

bearings, a comparison of the radial rigidity is made for the ball bearing between the developed FEM software and Hertz theory in Fig. 13. From Fig. 13, it is found that the FEM results are a little greater than the Hertz theory results. An experimental research is scheduled to identify which method is more reasonable and accurately in the near future.

6. Conclusions

A mathematical model and numeric method are presented in this paper in order to conduct contact analysis of rolling bearings based on the principle of the mathematical programming method. Three-dimensional, finite element method is introduced to calculate deformation influence coefficients and gaps of the assumed pairs of contact points between contact surfaces. Special software is developed to realize the procedures of the contact analysis. With the help of the special software, loaded bearing contact analyses are conducted for a deep groove ball bearing and a cylindrical roller bearing. Calculation results shows that the special software can calculate more reasonable and accurate contact pressure distribution of the rolling bearings than the commercial software SolidWorks and some other methods. The maximum contact pressure and radial contact rigidity of the ball bearing are also analyzed with Hertz theory. It is found that the results obtained by Hertz theory are similar to the results obtained by the special software, but they are not equal exactly. An experimental research is scheduled to identify which method is more reasonable and accurately in the near future.

Acknowledgment

NSK Foundation for the Advancement of Mechatronics is appreciated for funding this research. Mr. Ryuichi SONEZAKI, a student who graduated from Machine Design Lab in Shimane University, is also appreciated for conducting contact analysis of the ball and roller bearings with SolidWorks software under the direction of the first author.

References

[1] John, P. M. and Gohar, R. (1981) Roller bearings under radial and eccentric loads, Tribology International, 14, 131-136.
[2] Tedric, A. Harris, Michael, N. Kotzalas (2007) Essential Concept of Bearing Technology: Rolling Bearing Analysis, Fifth Edition, CRC Press.
[3] Tedric, A. Harris, Michael, N. Kotzalas, (2007) Advanced Concepts of Bearing Technology: Rolling Bearing Analysis, Fifth Edition, CRC Press.
[4] Kamamoto S., Fujimoto K. and Yamamoto T. (2001) Research on crowning profile to obtain maximum load carrying capacity for roller bearings, KOYO Engineering Journal English Edition, 159, 47-52.
[5] Nagatani, H. (2016) A new resolution to contact problem of roller bearings (in Japanese), The Tribology, 6(346), 44-46.
[6] Zhao, H. (1998) Analysis of load distributions within solid and hollow roller bearings, Trans. ASME, J. Tribology 120(1), 134-139.
[7] Ryuichi SONEZAKI, R. and LI, S. (2014) Contact analysis and stiffness testing device design of rolling bearings, JSME Mechanical Engineering Congress, 1-4.
[8] Product catalogue of NTN Corporation
[9] Dassault Systemes SolidWorks Corporation, SolidWords Essentials
[10] Guo Y. and Parker R. G. (2012) Stiffness matrix calculation of rolling element bearings using a finite element/contact mechanics model, Mechanism and Machine Theory, 51, 32–45.
[11] Wolfe, P. (1959) The Simplex method for quadratic programming, Econometrica, 27, 382-398.
[12] Hiramoto, I. and Hase, A. (1973) Linear Programming Method, Baifukan Co., LTD. Press.
[13] Liu G. (1994), Structural Dynamics of the Finite Element Method (In Chinese), National Defense Industry Press.

Two-dimensional SPH simulation of liquid sloshing in a rotating tank

J.R. Shao[1], †M.B. Liu[2]

[1]College of Mechanical Engineering, Chongqing University of Technology, Chongqing, China
[2]College of Engineering, Institute of Ocean Research, State Key Laboratory for Turbulence and Complex systems, Peking University, Beijing, China

† Corresponding author: mbliu@pku.edu.cn

Abstract

This paper presents the application of an improved smoothed particle hydrodynamics (SPH) method to simulate violent liquid sloshing in a rotating liquid tank. The dynamic response of the sloshing system and the change of the pressure profiles are investigated in detail while the SPH simulations are conducted under different external excitations and different water waves. It is revealed that for small amplitude liquid sloshing, the linear wave theory can apply and free surface does not break up. For liquid sloshing with strong nonlinear effects, free surface evolves violently with breakup and impinging onto bulky water and solid walls, and therefore the linear wave theory is no longer valid. A circular frequency deviated from the natural frequency from linear wave theory may produce bigger maximal pressure load than the case with equivalent natural frequency.

Keywords: SPH, liquid sloshing, dynamic response,circular frequency

1. Introduction

Sloshing refers to the movement of liquid inside a partially-filled container due to external excitations. As a complex hydrodynamic phenomenon, liquid sloshing involves the change of free surfaces and the strong coupling between liquid and containers, which are usually thin-shelled structures. Liquid sloshing can be frequently observed in daily life, engineering and sciences. When the amplitude of an external excitation is very large or its frequency is close to the natural frequency of the liquid sloshing system, the liquid inside the container can exhibit violent oscillations, and exert strong impact load on the container[1]. For example, large liquid sloshing in an oil or liquefied natural gas ship may result in local breakages and global instability to the ship, and further lead to leakage of oil, and capsizing of ship. The sloshing of liquefied fuel inside the fuel tank in an aeronautic or astronautic craft can disturb or even breakdown normal navigation of the craft.

Liquid sloshing has been a hot research subject attracting much attention over the last decades. The majority of theoretical publications are restricted to linear theory associated with non-resonant excitation and small forcing magnitude[2-5]. There are also some works based on weakly nonlinear semi-analytical theories, which used the assumptions that an inviscid incompressible liquid is in irrotational flow, the liquid surface does not overturn or break, and the surface tension does not matter [6-8]. All these theories in general, may result in big errors in the time-history response when the external excitation is large or near the natural frequency. Recently more and more researches on liquid sloshing are focused on numerical simulations with the advancement of the computer techniques [9-11]. However, most of the numerical simulations are focused on grid-based methods, such as FDM [12, 13], FEM [14, 15] and BEM [16-19]. Liquid sloshing is a complex fluid motion, which usually involves changing and breakup of free surfaces, strong turbulence and vortex, and violent fluid-solid interaction. Traditional grid-based numerical methods have difficulties in tracking changing free surfaces or moving interfaces, and usually need mesh adjustment or rezoning [20-22].

The meshfree and particle methods provide new alternates for investigating liquid sloshing problems [23, 24]. Among these methods, smoothed particle hydrodynamics (SPH) [25-27] combines the advantages of meshfree, Lagrangian and particle methods. In SPH, particles are used to represent the state of a system and these particles can freely move according to internal particle interactions and external forces. Therefore it can naturally obtain history of fluid motion, and can easily track material interfaces, free surfaces and moving boundaries. There are a few literatures addressing the application of SPH method to liquid sloshing dynamics. For example, Delorme et al. simulated the sloshing loads in LNG tankers with SPH [28]. Iglesias et al. simulated the anti-roll tanks and sloshing type problems[29]. Rhee and Engineer studied liquid tank sloshing with Reynolds-averaged Navier-Stokes[30]. Souto-Iglesias et al. assessed the liquid moment amplitude in sloshing type problems with smooth particle hydrodynamics[31]. Anghileri investigated the fluid-structure interaction of water filled tanks during the impact with the ground[32]. These works have demonstrated the feasibility of SPH method in modeling liquid sloshing dynamics. However, previous works are generally based on traditional SPH method, which have poor computational accuracy, and it is hard to track the variations of pressure. In this paper, an improved SPH model is used to simulate the liquid sloshing in a rotating tank.

2. SPH methodology

The governing motion of liquid sloshing in isothermal condition can be described by the following continuity and momentum equations

$$\frac{d\rho}{dt} = -\rho\nabla\cdot\boldsymbol{v}\,, \tag{1}$$

$$\frac{d\boldsymbol{v}}{dt} = -\frac{1}{\rho}\nabla P + \frac{\mu}{\rho}\nabla^2\boldsymbol{v} + \boldsymbol{g} + \frac{1}{\rho}\vec{\nabla}(\rho\boldsymbol{R})\,, \tag{2}$$

where ρ is fluid density, $\boldsymbol{v}$ is the velocity vector, P is pressure, μ is the dynamic viscosity, $\boldsymbol{g}$ is the gravitational acceleration, and $\boldsymbol{R}$ is the Reynolds stress tensor. The eddy viscosity assumption is used to model the Reynolds stress tensor as:

$$\boldsymbol{R} = 2\nu_t\boldsymbol{S} - \frac{2}{3}k\boldsymbol{I}\,, \tag{3}$$

where ν_t is the eddy viscosity, $\boldsymbol{S}$ is the mean rate-of-strain tensor, k is turbulence kinetic energy and $\boldsymbol{I}$ is a unit tensor. The Smagorinsky model of eddy viscosity $\nu_t = (C_s\Delta l)^2\sqrt{2S_{ij}S_{ij}}$ is widely used, where C_s is the Smagorinsky constant usually taken as 0.12 and Δl is a mixing length which is assumed to be the initial particle spacing in SPH. S_{ij} is the elements of $\boldsymbol{S}$ given by:

$$S_{ij} = \frac{1}{2}\left(\frac{\partial v_j}{\partial x_i} + \frac{\partial v_i}{\partial x_j}\right). \tag{4}$$

Considering $k = \overline{v_i'v_j'}\big/2$ and $R_{ij} = -\overline{v_i'v_j'}$, the relation between k and the Reynolds stress tensor can be written as $k = -R_{ii}/2$.

In SPH, the state of a system is represented by a set of particles, and flow field variables (such as density, velocity, acceleration) can be obtained through approximating the governing

equations which are discretized on the set of particles. A field function and its derivative can then be written in the following forms

$$< f(\boldsymbol{x}_i) >= \sum_{j=1}^{N} \frac{m_j}{\rho_j} f(\boldsymbol{x}_j) W(\boldsymbol{x}_i - \boldsymbol{x}_j, h), \tag{5}$$

$$< \nabla \cdot f(\boldsymbol{x}_i) >= \sum_{j=1}^{N} \frac{m_j}{\rho_j} f(\boldsymbol{x}_j) \nabla_i W_{ij}, \tag{6}$$

where $< f(\boldsymbol{x}_i) >$ is the approximated value of particle *i*; $f(\boldsymbol{x}_j)$ is the value of $f(\boldsymbol{x})$ associated with particle *j*; $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are the positions of corresponding particles; *m* denotes mass; *h* is the smooth length; *N* is the number of the particles in the support domain; *W* is the smoothing function representing a weighted contribution of particle *j* to particle *i*. The cubic spline function has continuous second-order derivative, and it is not sensitive to particle disorder. It is given by

$$W(R,h) = a_d \times \begin{cases} 1 - \frac{3}{2}R^2 + \frac{3}{4}R^3, & 0 \le R < 1 \\ \frac{1}{4}(2-R)^3, & 1 \le R < 2 \\ 0, & R \ge 2 \end{cases}. \tag{7}$$

where a_d is $3/2h$, $10/7\pi h^2$ and $1/\pi h^3$ in one-, two- and three-dimensional space, respectively, so that the condition of unity can be satisfied for all the three dimensions. $R = |\boldsymbol{x} - \boldsymbol{x}'| / h$.

Through particle approximation, the governing motion of fluid flow in isothermal condition can be described by the following continuity and momentum equations

$$\frac{d\rho_i}{dt} = \sum_{j=1}^{N} m_j \boldsymbol{v}_{ij} \nabla_i W_{ij}, \tag{8}$$

$$\begin{aligned} \frac{d\boldsymbol{v}_i}{dt} = &-\sum_{j=1}^{N} m_j \left(\frac{p_i}{\rho_i^2} + \frac{p_j}{\rho_j^2}\right) \cdot \nabla_i W_{ij} + \sum_{j=1}^{N} \frac{4 m_j \left(\mu_i + \mu_j\right) \boldsymbol{x}_{ij} \cdot \nabla_i W_{ij}}{\left(\rho_i + \rho_j\right)^2 \left(x_{ij}^2 + 0.01h^2\right)} \boldsymbol{v}_{ij} \\ &+ \sum_b m_b \left(\frac{\boldsymbol{R}_a}{\rho_a} + \frac{\boldsymbol{R}_b}{\rho_b}\right) \cdot \nabla_a W_{ab} + \boldsymbol{g} \end{aligned}. \tag{9}$$

To solve the equations of motion, an artificial compressibility technique is usually used to model the incompressible flow as a slightly compressible flow. Therefore, it is feasible to use a quasi-incompressible equation of state to model the incompressible flow. A commonly used artificial equation of state is

$$P = C_0^2 (\rho - \rho_0), \tag{10}$$

where ρ_0 is the reference density , and it is taken as the initial density of the water. C_0 is the sound speed.

It is known that the conventional SPH method has been hindered with low accuracy as it cannot exactly reproduce quadratic and linear functions, and even cannot exactly reproduce a

constant. The accuracy is also closely related to the distribution of particles, selection of smoothing function and the support domain (described by the smoothing length *h* multiplied by a scalar factor). In liquid sloshing problem, the changing and breakup of free surfaces as well as splashing and fall of water particles lead to highly disordered particle distribution, which can seriously influence computational accuracy of SPH approximations. Therefore improving the particle inconsistency and hence the approximation accuracy is necessary. In this work, the KGC model[33] for approximating kernel gradient is used. It is obtained based on Taylor series expansion on the SPH approximation of a function, and can guarantee the second order accuracy in the whole computational domain. It is given by

$$\nabla_i^C W_{ij} = L(\boldsymbol{r}_i)\nabla_i W_{ij}\,, \tag{11}$$

$$L(\boldsymbol{r}_i) = \left(\sum_j \begin{pmatrix} x_{ji}\dfrac{\partial W_{ij}}{\partial x_i} & y_{ji}\dfrac{\partial W_{ij}}{\partial x_i} \\ x_{ji}\dfrac{\partial W_{ij}}{\partial y_i} & y_{ji}\dfrac{\partial W_{ij}}{\partial y_i} \end{pmatrix} V_j \right)^{-1}. \tag{12}$$

It is noted that, the new kernel gradients are obtain after searching particles, and then they can be used for the approximation of and field variables, therefore it can save a lot of computing time. And there is no need to significantly change the structure of SPH computer programs and procedure of SPH simulations.

In SPH, solid boundary conditions are not able to be directly and rigorously implemented as in the grid-based numerical models. In this work, a coupled dynamic solid boundary treatment algorithm(SBT) [34] (Fig.1) is used to construct the solid boundary, which can predict the pressure of the solid boundary accurately.

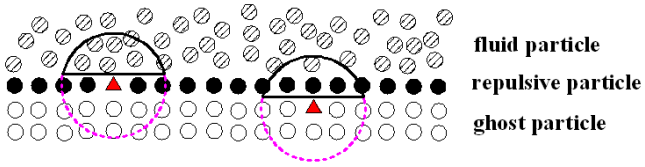


Figure 1. Illustration of the coupled dynamic SBT algorithm.

In this coupled dynamic SBT algorithm, the information of repulsive particles only comes from fluid particles, and the information of ghost particles comes from both fluid and repulsive particles. For non-slip boundary condition, variables of the boundary particles (both repulsive and ghost particles) can be obtained as follows,

$$\rho_i^{new} = \sum_{j=1}^{N} \rho_j W_{ij}^{new} \frac{m_j}{\rho_j} = \sum_{j=1}^{N} m_j W_{ij}^{new}\,, \tag{13}$$

$$v_i = -\sum_{j=1}^{N} v_j W_{ij}^{new} \frac{m_j}{\rho_j}\,, \tag{14}$$

$$W_{ij}^{new} = \frac{W_{ij}}{\sum_{j=1}^{N} W_{ij} \dfrac{m_j}{\rho_j}}. \tag{15}$$

For particles near the free surface, the support domains are usually cut off, and the kernel function do not keep the normalized nature. It has great influence to particle approximation. To resolve the problem, a simple but very effective treatment is presented, which is obtained based on interpolation of kernel function,

$$k_i = \sum_{j=1}^{N} W_{ij} \frac{m_j}{\rho_j} \le 0.9 \ . \tag{16}$$

Here, k_i is a coefficient, it is calculated each time step. If the value is less than 0.9, it is considered to be the free surface particle, and the pressure will be equal to atmospheric pressure, its density is equal to the initial value.

3. SPH modelling of sloshing in a rotating liquid tank

In this section, the liquid sloshing in a rotating tank is investigated under different external excitations. Fig. 2 shows the geometry of the liquid sloshing system, which is similar to what Iglesias provided [31]. i.e., L=0.64 m, H=1.15 m, the water depth is h_w=0.03 m, and the centre of rotation is 0.1 m below the baseline.

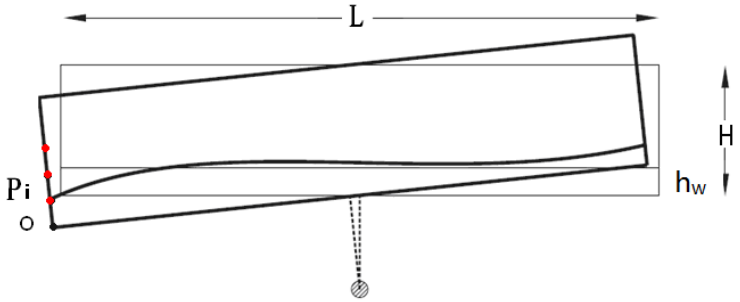


Figure 2. Illustration of liquid sloshing in a rotating tank.

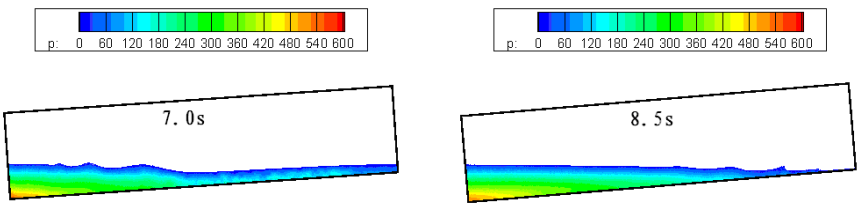
The external excitation can be described as $\theta = \theta_0 \sin(\omega t)$, where θ_0 is the angular displacement, ω is the circular frequency of the rotating motion. In the left wall, eight probe points P1-P8 are employed, OP1= P1P2=…= P7P8= 0.01 m. In this model, the water depth is very small comparing with the size of the tank, therefore the corresponding first resonance frequency can be obtained from the linear wave theory[31] ,

$$\omega = \sqrt{\frac{g\pi}{L} \tanh\left(\frac{\pi h_w}{L}\right)} . \tag{17}$$

Therefore, the value of the first resonance frequency is 2.65 rad/s.

3.1 Liquid sloshing under different circular frequencies

To investigate the free surface evolution and the pressure load on the tank under different circular frequencies, the sloshing models are set as θ_0=6°, and ω=1.0, 2.65, 4.34, 8.21rad/s respectively.



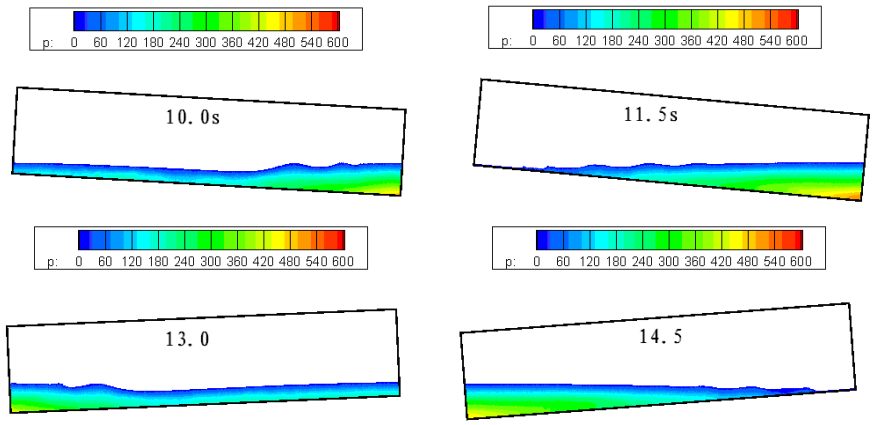


Figure 3. Pressure field with ω=1.0 rad/s.

When ω =1.0 rad/s, the period of the external excitation is 6.28s. Fig.3 shows the corresponding liquid sloshing mode and free surface evolution of the liquid sloshing system in a period. In this case, the frequency of the external excitation (1.0 rad/s) is less than the natural frequency (2.65 rad/s) and the free surface does not break up. Though there are nonlinear phenomena, the linear wave theory still applies. In the whole process of the sloshing, the water waves are travelling waves. Fig.4 is the pressure history at P3. It shows that the water impact the solid wall periodically, with an amplitude about 380 Pa.

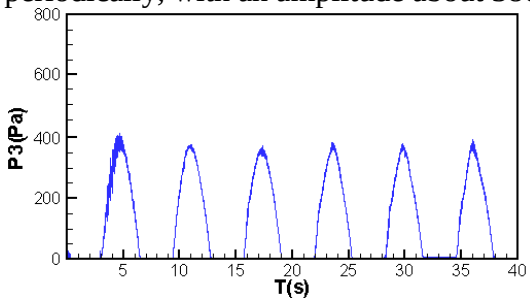


Figure 4. Pressure values at P3, ω=1.0 rad/s.

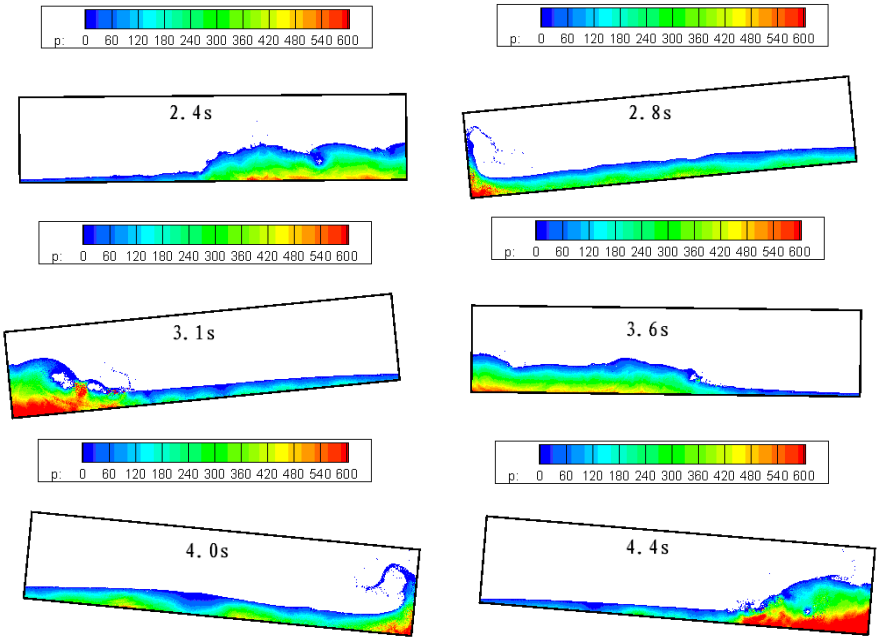


Figure 5. Pressure field with ω=2.65 rad/s.

Next, we increase the value of ω to 2.65 rad/s, which equals to the natural frequency of sloshing system. The free surface evolution as well as the pressure field distribution is shown on Fig.5. Comparing Fig.5 with Fig.3, it is found that, liquid sloshing in this case is more violent with free surface break up. It can also be observed that with the development of the liquid sloshing, the water column impacts brutally against the solid wall, generates a bounce-back flow pattern after a short period of interaction with the vertical wall, and finally forms some cavities. The cavities change their shapes and disappear finally. The water wave is breaking wave and the linear wave theory may not be valid. Fig.6 shows that water exerts strong impact load on the solid wall (left and right wall) periodically, and the pressure amplitude is about 800 Pa at P3, which is much bigger than that in the above case.

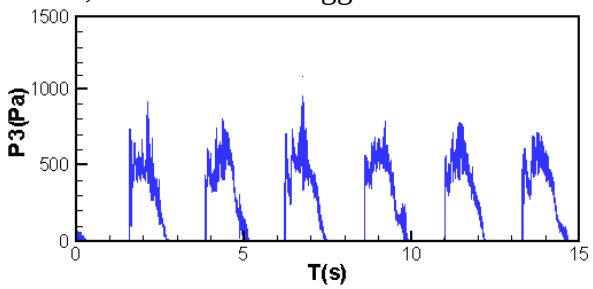
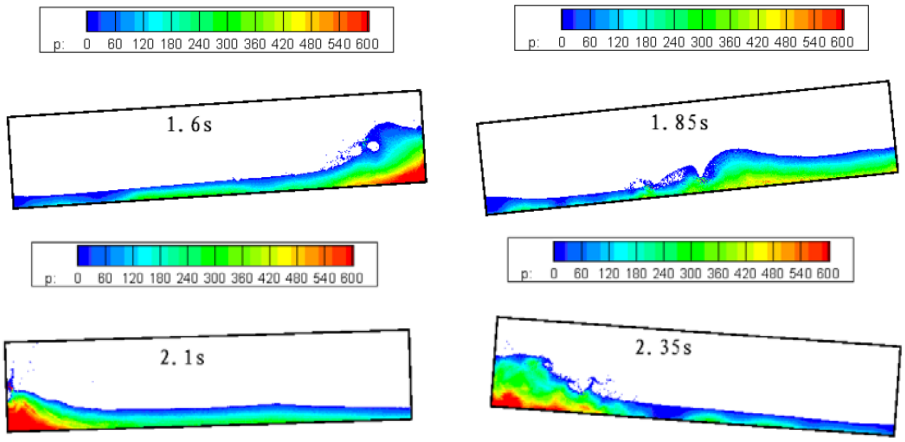


Figure 6. Pressure values at P3, ω=2.65rad/s.

Increasing the value of ω to 4.34 rad/s, the moving condition in Fig.7 can be obtained. The sloshing is still violent and the water wave has two main wave peaks. The front one is higher than the later one. These two surge fronts with two water wave peaks then impact the solid wall (left and right wall) twice, as shown in Fig. 8. The first impact is more violent with an instantaneous pressure peak (about 1200-1400 Pa). The pressure peak of the second impact is about 700-1000Pa. One notable observation is that in this case, the strongest wave impact can happen on the left-top and right-top corners. Also for this case, though the circular frequency is away from the first natural frequency (2.65 rad/), the maximal impact pressure values both for the first and second water impacts are bigger than the above case in which the circular frequency equals to the natural frequency. This is because the liquid sloshing is associated with strong nonlinear effects (with breaking free surface and violent water slamming on the solid wall & on the bulky water). For such highly nonlinear liquid sloshing, the linear wave theory does not apply.



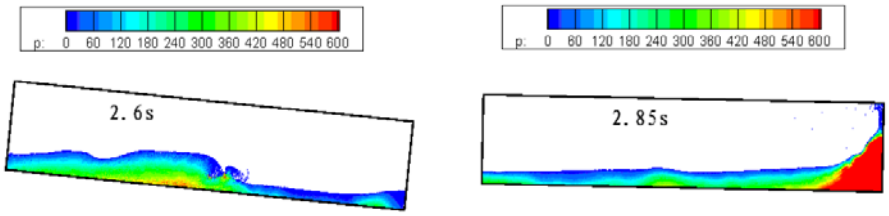


Figure 7. Pressure field with ω =4.34 rad/s.

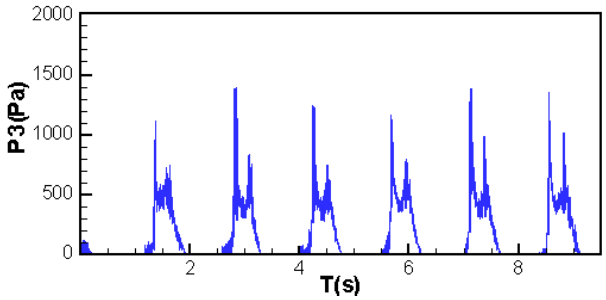


Figure 8. Pressure values at P3 with ω =4.34 rad/s.

Increasing the frequency of the external excitation to 8.21 rad/s, the amplitude of the liquid sloshing is much smaller than that in Fig.5. It is noted that as the frequency of the external excitation is so big, the movement of water is not able to follow the movement of the container. Free surface does not break up and the water wave becomes travelling wave again. Two and even more wave peaks occurs during the sloshing process (see Fig. 9) At the same time, the pressure peak will also decrease, as shown in Fig. 10, its value is about 250- 600Pa, and is different in different periods due to the interaction of different wave surge fronts.

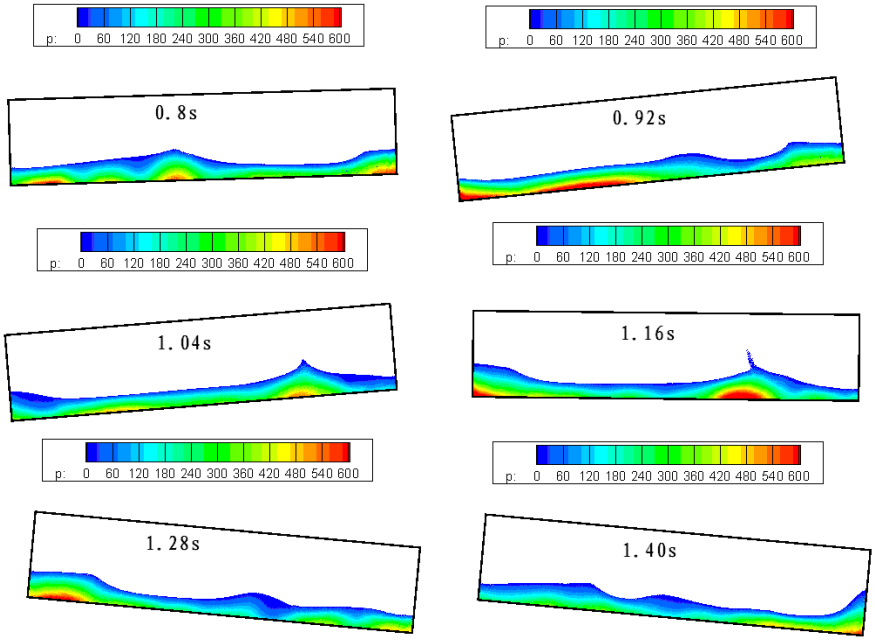


Figure 9. Pressure field with ω =8.21 rad/s.

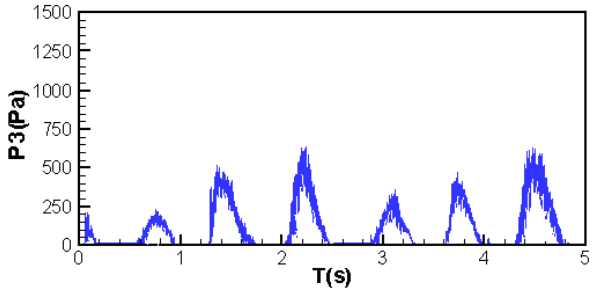


Figure 10. Pressure values at P3, ω=8.21 rad/s.

3.2 Liquid sloshing under different angular displacements

To investigate the change of the free surface evolution, and the pressure load on the tank under different angular displacement, the sloshing models are set ω = 2.65 rad/s (corresponding to the natural frequency), θ_0 =4°, 6°, 8°, 10°, 12° separately. The liquid sloshing with θ_0 =4° can be observed in Fig.10. Comparing the pressure field with Fig.5 (θ_0 =6°) and Fig.11 (θ_0 =10°), it is found that the amplitude of the sloshing increase with the increasing of the angular displacement.

Fig.12 shows the pressure curves in the eight probes with θ_0 = 4°. The peak values of the pressure change from 180 Pa to 950 Pa. When the angular displacement adds up to 10°, free surface breaks up and the water wave becomes rougher and the whole flow domain involves strong turbulence and vortex. To describe the turbulence more accurately, the RANS turbulence model is necessary. Fig.13 shows the evolution of the cavities in the flow domain. The water column impacts brutally against the solid wall. When *t*= 6.3 s, the water front impacted on the top of the container, and generated a bounce-back flow pattern. Fig. 14 shows that the peak values of the pressure change from 400Pa to 1400 Pa, which is bigger than that when θ_0 = 4°. For a specific angular displacement at natural frequency, the pressure load on the right or left solid wall will reduce with the increase of the height of the observation points.

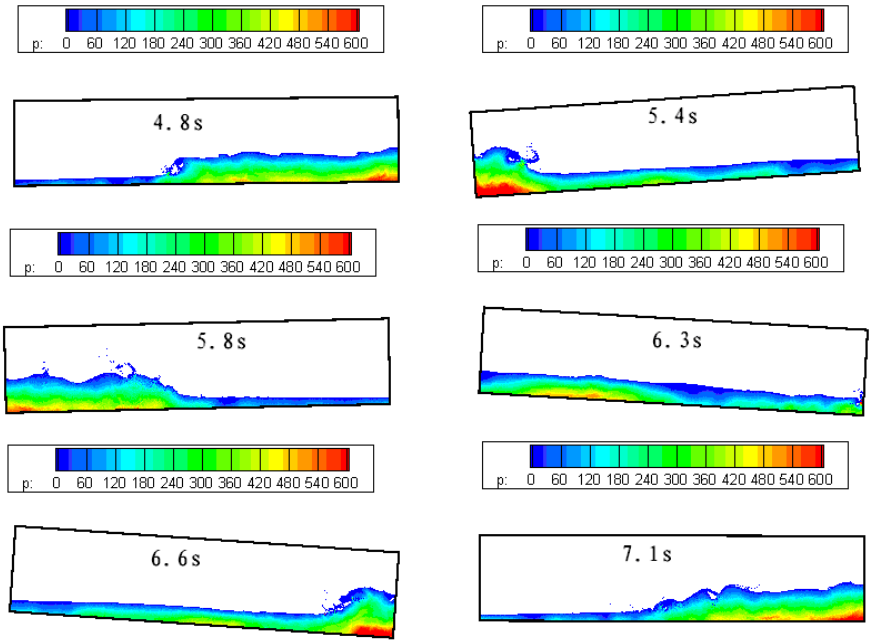


Figure 11. Pressure field with θ_0 = 4°, ω = 2.65rad/s

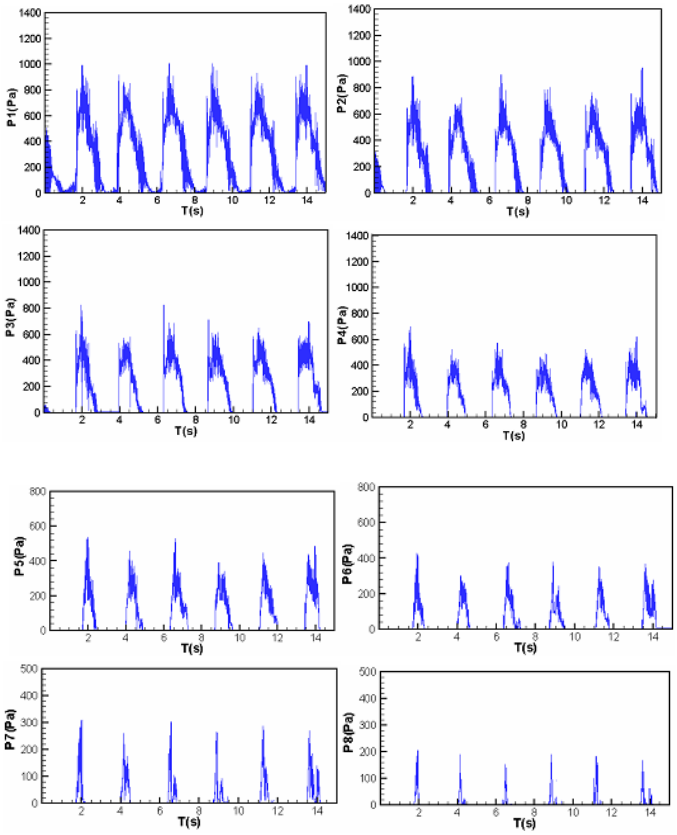


Figure 12. Pressure values at P1-P8, θ_0 = 4°

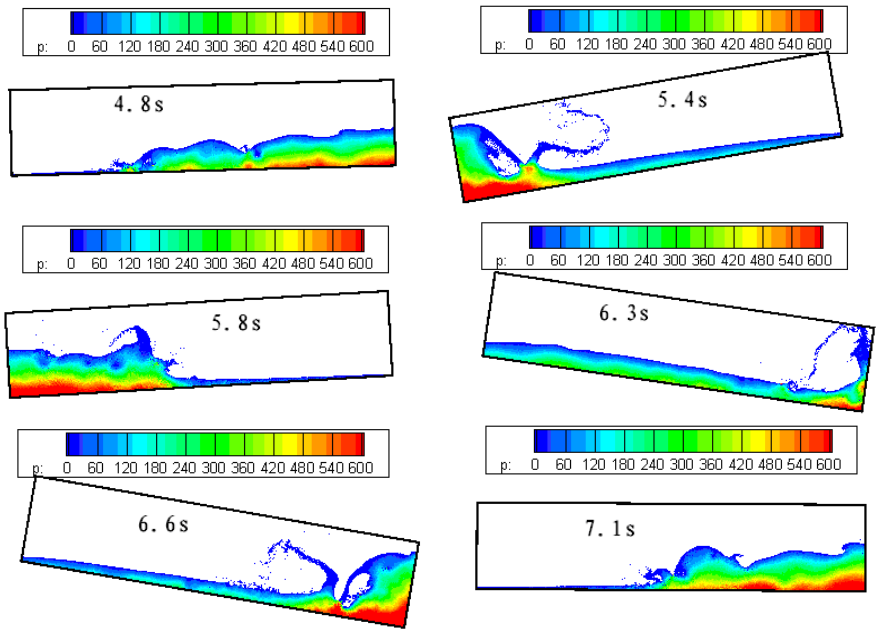


Figure 13. Pressure field with θ_0 = 10°, ω = 2.65 rad/s.

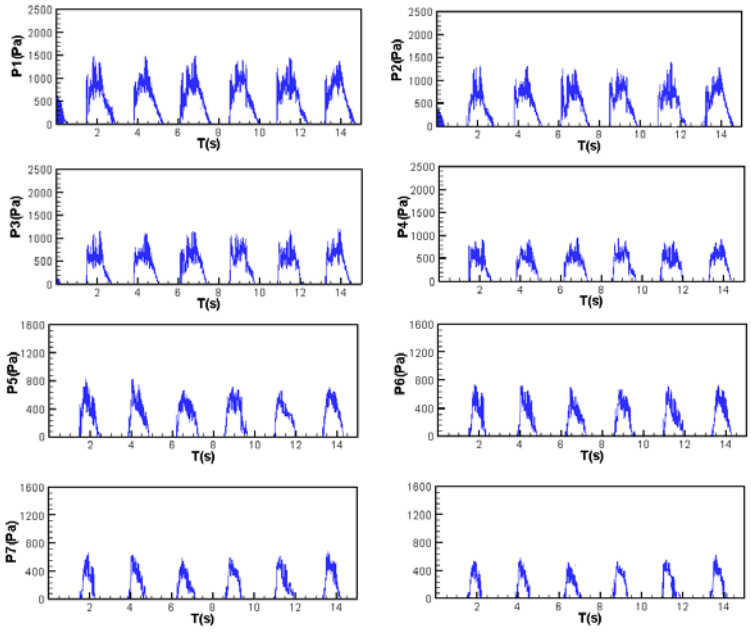


Figure 14. Pressure values at P1-P8, θ_0 = 10^o.

Conclusions

In this paper, an improved SPH model is used for the simulation of liquid sloshing in a rotating tank which moves under different external excitations. The KGC and density correction can improve the computational accuracy and obtain smoothed pressure field; the coupled dynamic solid boundary treatment can remove the numerical oscillation near the solid boundary and ensure no penetration condition.

From the numerical simulations, the following conclusions can be drawn:

1) The liquid inside the container can demonstrate different sloshing behaviours with different circular frequency. For small circular frequencies, liquid sloshing is associated with smoothly evolving free surface and the linear wave theory is valid. The resultant water wave behaves as traveling waves and the maximal impact pressure happens on the left and right walls of the container.
2) For cases with circular frequencies close to or slightly bigger than the natural frequency, the sloshing liquid is highly nonlinear, and the linear wave theory is no longer valid. Therefore, the maximal pressure load obtained with the circular frequency equivalent to the natural frequency from linear wave theory may not necessary bigger than those from circular frequencies deviated from the natural frequency from linear wave theory.
3) For cases with very large circular frequencies, the movement of water is not able to follow the movement of the container. Free surface does not break up and the water wave becomes travelling waves again. Two and even more wave peaks may occur during the sloshing process.
4) Changing the angular displacement can also lead to different sloshing patterns. Increasing the angular displacement can lead to more violent movement. For a specific angular displacement at the natural frequency, the pressure load on the right or left solid wall will reduce with the increase of the height of the observation points.

Acknowledgement This work has been supported by the National Natural Science Foundation of China (11602045), NSAF (Grant No.U1530110), Natural Science Foundation of Chongqing (cstc2016jcyjA0373), STRPCMEC (Grant No.KJ1600918).

References

[1] Ibrahim,R.A. (2005) Liquid sloshing dynamics: theory and applications, Cambridge University Press.
[2] Karamanos, S.A., Patkas, L.A. and Platyrrachos, M.A. (2006) Sloshing effects on the seismic design of horizontal-cylindrical and spherical industrial vessels. *Journal of pressure vessel technology*, **128(3)**, 328-340.
[3] Patkas, L.A. and Karamanos,S.A. (2007) Variational solutions for externally induced sloshing in horizontal-cylindrical and spherical vessels. *Journal of engineering mechanics*, **133(6)**, 641-655.
[4] Drodos, G.C., Dimas, A.A, and Karabalis, D.L. (2008) Discrete modes for seismic analysis of liquid storage tanks of arbitrary shape and height. *J. Pressure Vessel Technology*, **130**, 0418011.
[5] Hasheminejad, S.M. and Aghabeigi, M. (2009) Liquid sloshing in half-full horizontal elliptical tanks. *Journal of Sound and Vibration*, **324 (1)**, 332-349.
[6] Gavrilyuk, I.P., Lukovsky, I.A. and Timokha, A.N. (2000) A multimodal approach to nonlinear sloshing in a circular cylindrical tank. 2000. *Hybrid Method Engng*, **2(4)**, 463-183.
[7] Faltinsen, O.M., Rognebakke, O.F. and Timokha,A.N. (2003) Resonant three-dimensional nonlinear sloshing in a square-base basin. J*ournal of Fluid Mechanics*, **487**, 1-42.
[8] Gavrilyuk, I., Lukovsky, I., Trotsenko, Y. and Timokha, A. (2006) Sloshing in a vertical circular cylindrical tank with an annular baffle. Part 2. NonLinearresonant waves, *Journal of Engineering Mathematics*, **54(1)**, 71-88.
[9] Ibrahim, R.A., Pilipchuk, V.N. and Ikeda, T. (2001) Recent advances in liquid sloshing dynamics. *Applied Mechanics Reviews*, **54**,133-199.
[10]Faltinsen, O.M. and Timokha, A.N. (2009) Sloshing, Cambridge University Press.
[11]Cariou, A. and Casella, G. (1999) Liquid sloshing in ship tanks: a comparative study of numerical simulation. *Marine structures*, **12(3)**, 183-198.
[12]Chen, B.F. and Nokes, R. (2005) Time-independent finite difference analysis of fully non-linear and viscous fluid sloshing in a rectangular tank. *Journal of Computational Physics*, **209(1)**: 47-81.
[13]Chen, B.F. (2005) Viscous fluid in tank under coupled surge, heave, and pitch motions. J*ournal of waterway, port, coastal, and ocean engineering*, **131(5)**, 239-256.
[14]Wu, G.X., Ma, Q.W. and Eatock Taylor R. (1998) Numerical simulation of sloshing waves in a 3D tank based on a finite element method. *Applied Ocean Research*, **20(6)**, 337-355.
[15]Mitra, S., Upadhyay, P.P. and Sinhamahapatra, K.P. (2008) Slosh dynamics of inviscid fluids in two-dimensional tanks of various geometry using finite element method. *International Journal for Numerical Methods in Fluids*, **56(9)**, 1625-1651.
[16]Faltinsen, O.M. (1978) A numerical nonlinear method of sloshing in tanks with two-dimensional flow. *Journal of Ship Research*, **22(3)**, 193-202.
[17]Faltinsen, O.M., Rognebakke, O.F., Lukovsky, I.A. and Timokha, A.N. (2000) Multidimensional modal analysis of nonlinear sloshing in a rectangular tank with finite water depth. *Journal of Fluid Mechanics*, **407**, 201-234.
[18]Faltinsen, O.M. and Timokha, A.N. (2001) An adaptive multimodal approach to nonlinear sloshing in a rectangular tank. *Journal of Fluid Mechanics*, **432**, 167-200.
[19]Faltinsen, O.M, and Timokha, A.N. (2010) A multimodal method for liquid sloshing in a two-dimensional circular tank. *Journal of Fluid Mechanics*. **665**, 457-479.
[20]Liu, D.M., and Lin, P.Z. (2008) A numerical study of three-dimensional liquid sloshing in tanks. *Journal of Computational Physics*, **227(8)**, 3921-3939.
[21]Veldman, A.E.P., Gerrits, J., Luppes, R., Helder, J.A. and Vreeburg J.P.B. (2007) The numerical simulation of liquid sloshing on board spacecraft. *Journal of Computational Physics*, **224(1)**, 82-99.
[22]Fang, Z.Y., Duan, M.Y.and Zhu, R.Q. (2007) Numerical simulation of liquid sloshing in a liquid tank based on Level-set method. *Journal of Ship Mechanics*, **11(1)**, 62-67.
[23]Pan, X.J., Zhang, H.X. and Lu, Y.T. (2008) Numerical simulation of viscous liquid sloshing by moving-particle semi-implicit method. *Journal of Marine Science and Application*, **7(3),** 184-189.
[24]Koshizuka, S., Nobe, A. and Oka, Y. (1998) Numerical analysis of breaking waves using the moving particle semi-implicit method. *International Journal For Numerical Methods In Fluids*, **26(7)**, 751-769.
[25]Gingold, R.A. and Monaghan J.J. (1977) Smoothed particle hydrodynamics-theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, **181**, 375-389.
[26]Lucy, L.B. (1977) A numerical approach to the testing of the fission hypothesis. *The Astronomical Journal*,. **82**, 1013-1024.

[27]Liu, G.R. and Liu, M.B. (2003) Smoothed particle hydrodynamics: a meshfree particle method, World Scientific, Singapore.
[28]Delorme, L., Iglesias, A.S. and Perez, S.A. (2005) Sloshing loads simulation in LNG tankers with SPH. *International Conference on Computational Methods in Marine Engineering*. Barcelona.
[29]Iglesias, A.S., Rojas, L.P. and Rodriguez, R.Z. (2004) Simulation of anti-roll tanks and sloshing type problems with smoothed particle hydrodynamics. *Ocean Engineering*, **31(8-9)**, 1169-1192.
[30]Rhee, S.H. and Engineer, L. (2005) Unstructured grid based Reynolds-averaged Navier-Stokes method for liquid tank sloshing. *Journal of Fluids Engineering*, **127**, 572-582.
[31]Souto-Iglesias, A., Delorme, L., Perez-Rojas, L. and Abril-Perez, S. (2006) Liquid moment amplitude assessment in sloshing type problems with smooth particle hydrodynamics. *Ocean Engineering*, **33(11-12)**, pp. 1462-1484.
[32]Anghileri, M., Castelletti, L.M.L. and Tirelli, M. (2005) Fluid structure interaction of water filled tanks during the impact with the ground. *International Journal of Impact Engineering*, **31(3)**,235-254.
[33]Shao, J.R., Li, H.Q., Liu, G.R. and Liu, M.B. (2012) An improved SPH method for modeling liquid sloshing dynamics. *Computers & Structures*, **100-101**, 18-26.
[34]Liu, M.B., Shao, J.R. and Chang, J.Z. (2012) On the treatment of solid boundary in smoothed particle hydrodynamics. *Science China*, **55**, 244-254.

Periodic DFT study of structural transformations of crystalline dihydroxylammonium 5,5′-bistetrazole-1,1′-diolate under high pressures

G.Z. Zhao[1]*, H.R. Sun[2], J.F. Jia[1], and H.S. Wu[1]

[1] School of Chemistry and Material Science, Shanxi Normal University, PR China

[2] School of Foreign Languages, Shanxi Normal University, PR China

*Presenting author: zhaoguozheng99@126.com

†Corresponding author: zhaoguozheng99@126.com

Abstract

Density functional theory (DFT) periodic calculations were performed to study the crystal and electronic structures of energetic compound dihydroxylammonium 5,5′-bistetrazole-1,1′-diolate (TKX-50) under the pressure ranging from 0 to 400 GPa. The optimized crystal structure by the local density approximation (LDA) with CA-PZ functional matches well with the experimental values under the ambient pressure. When the structural transformations occur under the pressure of 126, 288, and 334 GPa, with the pressure increasing, the lattice constants and unit cell volume of TKX-50 change gradually. First of all, TKX-50 is rearranged in the crystal and the improvement of the molecular planarity occurs. Next, structural transformation appears with the distortion of the tetrazole rings. Finally, the rotation of molecular conformation occurs. The results of density of states show that TKX-50 crystal, with the increase of pressure, undergoes an electronic transition from the semiconductor to the metallic system. These results provide basic information for the high pressure behavior of crystalline TKX-50.

Keywords: density functional theory; TKX-50; high pressure; crystalline structures

Introduction

High energetic compounds have been widely used in military and civilian applications [1–4]. To meet the continuing need for novel high energetic compounds with high explosive performance and insensitivity, many scientists have paid considerable attention to the compounds' designing and synthesizing over several decades [5–7]. Among various types of high energetic compounds, energetic ionic salts come to be a unique class of energetic compound and have received a substantial amount of interests due to their merits, such as low vapor pressures, favorable insensitivity, excellent explosive performance, and environmental acceptability [8,9].

The energetic ionic salts of azoles, which are composed of high-nitrogen cations and anions, possess a large number of energetic N=N, N–N, C=N and C–N bonds. Therefore, they exhibit high energy and positive heats of formation [10–12]. Compared with the others, the tetrazole cations or anions are more energetic owing to their higher nitrogen content [13]. Klapötke *et al.* [14], detailed the preparation of a new explosive dihydroxylammonium 5,5′-bistetrazole-1,1′-diolate (TKX-50), which is not only easily prepared and exceedingly powerful (detonation velocity 9.70 $km \cdot s^{-1}$, detonation pressure 42.4GPa), but also possessing the required thermal insensitivity (decomposition onset 222℃), low toxicity, and handling safety. Luo *et al.* [15], investigated the structures, mechanical properties, and mechanical responses of TKX-50 and TKX-50 based PBX with molecular dynamics. Goddard *et al.* [16], developed a flexible classical force field for TKX-50, which reproduces the cell parameters, densities, lattice energy and mechanical properties derived from quantum mechanics (QM) and experiments.

Detailed information on the behavior of TKX-50 under high pressures is of great significance for the understanding of its chemical reactivity, detonation process, structural stabilities, and sensitivity. In the present work, the periodic DFT calculations were performed

to study the crystal structure and properties of TKX-50 under hydrostatic pressure of 0–400 GPa. Figure 1 displays crystal and molecular structures of TKX-50.

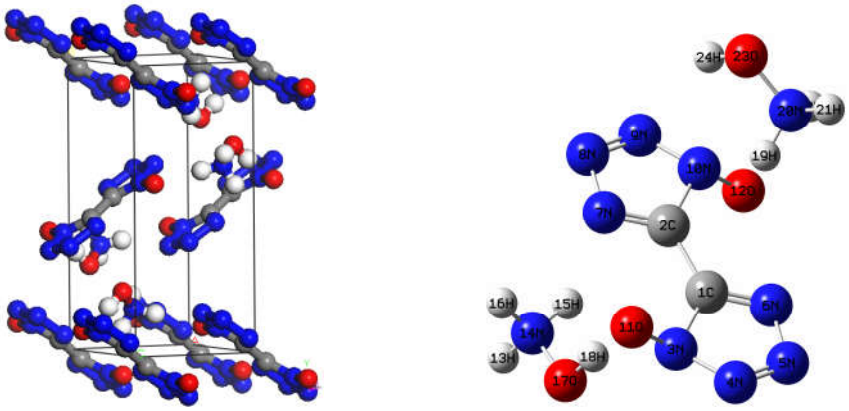


Figure 1. Crystal and molecular structures of TKX-50

Computational Methods

The periodic DFT calculations were performed in combination with the Vanderbilt-type ultrasoft pseudopotential and a plane-wave expansion of the wave functions [17]. The local density approximation (LDA) with the Ceperley−Alder exchange-correlation potential parametrized by Perdew and Zunger (CA-PZ) [18,19] implemented in the CASTEP [20] module of Materials Studio 6.0 [21] were used to perform the test calculations on crystalline TKX-50. The cutoff energy of plane waves was set to 340 eV. Brillouin zones sampling was performed by using the Monkhost−Pack scheme with a *k*-point grid of 3 × 1 × 2. The initial crystal was taken from Klapötke *et al*. (CCDC 872232) [14] and used for the computations.

Results and Discussion

Table 1. Comparison of the lattice constants of TKX-50 with experimental data at ambient conditions

Method	*a* (Å)	*b* (Å)	*c* (Å)	Cell volume (Å^3)
LDA/CA-PZ	5.522 (1.50)	11.047 (-5.99)	6.235 (-4.97)	377.83 (-9.58)
GGA/PW91	5.061 (-6.97)	13.171 (12.08)	7.086 (8.00)	468.31 (12.07)
GGA/PBE	5.066 (-6.86)	13.139 (11.81)	7.083 (7.96)	467.51 (11.88)
Exp.	5.441	11.751	6.561	417.86

The values in parentheses correspond to the percentage differences relative to the experimental data.

Two different functionals, local density approximation (LDA) and generalized gradient approximation (GGA) were applied to the computation of crystalline TKX-50 as a test. To benchmark the performance of the theoretical approach, LDA/CA-PZ, GGA/PW91(Perdew−Wang-91) [22], and GGA/PBE (Perdew−Burke−Ernzerhof) [23] were selected to fully relax the TKX-50 at ambient pressure without any constraint. Table 1 lists the experimental and relaxed cell parameters of TKX-50 crystal. The relative errors of the calculated values to the experimental ones show that the calculated values of LDA/CA-PZ agree better with the experimental ones than those of GGA. Thus, LDA/CA-PZ method has been employed in the present study.

Crystal Structure

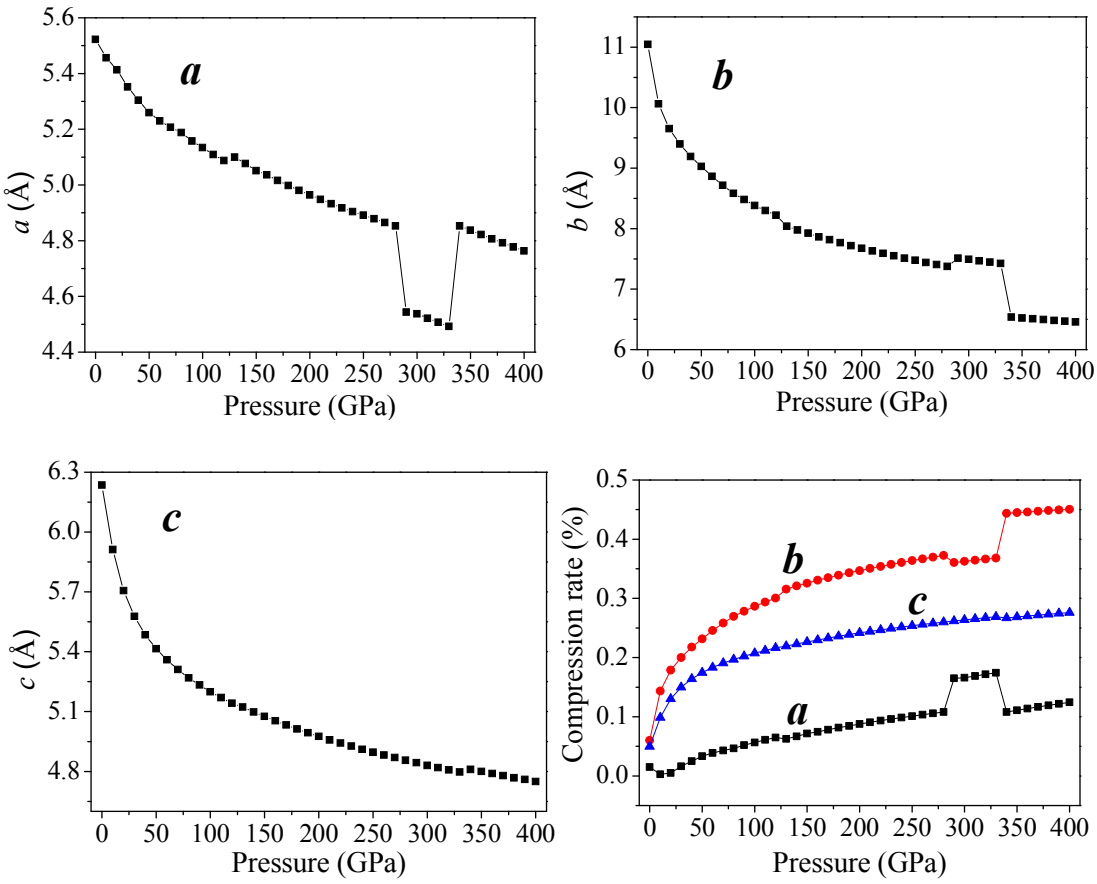


Figure 2. Lattice constants (*a*, *b*, *c*) and compression rates of TKX-50 as a function of pressure

The relaxed lattice constants (*a*, *b*, c) and compression rates in the pressure range of 0−400 GPa are depicted in Figure 2. With the pressure increasing, the lattice constants (*a*, *b*, *c*) and unit cell volume decrease gradually. It is because the external pressure is large enough to overcome the intermolecular repulsion along the crystallographic directions and makes the crystal structure shrink. The curves *a*, *b*, and *c* have sudden changes in magnitudes at about 126, 288, and 334 GPa, which suggest large changes have taken place. The value of *a* at 126 GPa is anomalously larger than that at 125 GPa, and increases abnormally again at 334 GPa. At 126 GPa, *b* value is much smaller than that at 125 GPa, and decreases sharply at 334 GPa. During 150−250 GPa, the lattice constants (*a*, *b*, *c*) decrease steadily, while at 334 GPa *c* increases dramatically.

The largest compression of the unit cell takes place in the pressure region below 50 GPa. With the pressure increasing from 50 to 125 GPa, and from 130 to 280 GPa, the lattice parameters decrease slowly. In the pressure range of 0–280 GPa, the compression ratios along three directions are not tantamount. At 80 GPa, the total compression ratios along the directions of *a*, *b*, and *c* are 6.9%, 28.8%, and 26.4%, respectively, which indicats that the compressibility of TKX-50 crystal is anisotropic and the structure is much stiffer in *a* direction than those in *b* and *c* directions.

Electronic Structure

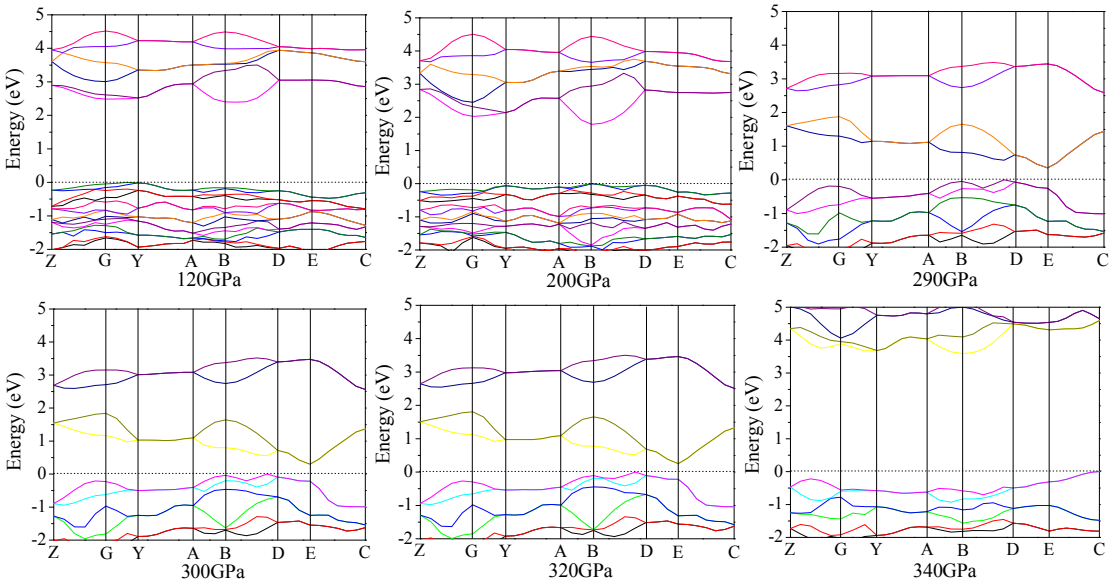


Figure 3. Self-consistent band structures of crystalline TKX-50 under different pressures. The Fermi energy is shown as a dashed horizontal line

On the basis of the equilibrium crystal structures obtained under different pressures, the self-consistent band structures along different symmetry directions of the Brillouin zone have been calculated and are shown in Figure 3. In order to have a visualized view, there is merely a presentation of range from –2.0 to 5.0 eV. It is clear that as the pressure is raised, the energy bands of valence and conduction bands change in an obvious way. For example, if the pressure is less than 200 GPa, energy bands will be in a stable state and the intermolecular interactions are not obvious. Some important changes in the electronic properties can be got on the basis of the transformation in configuration of TKX-50 at 288 GPa. As is shown, the energy bands get to a higher energy region. That is to say, the conduction bands become wider, while the valence bands become narrower in comparision with the ones in lower pressures.

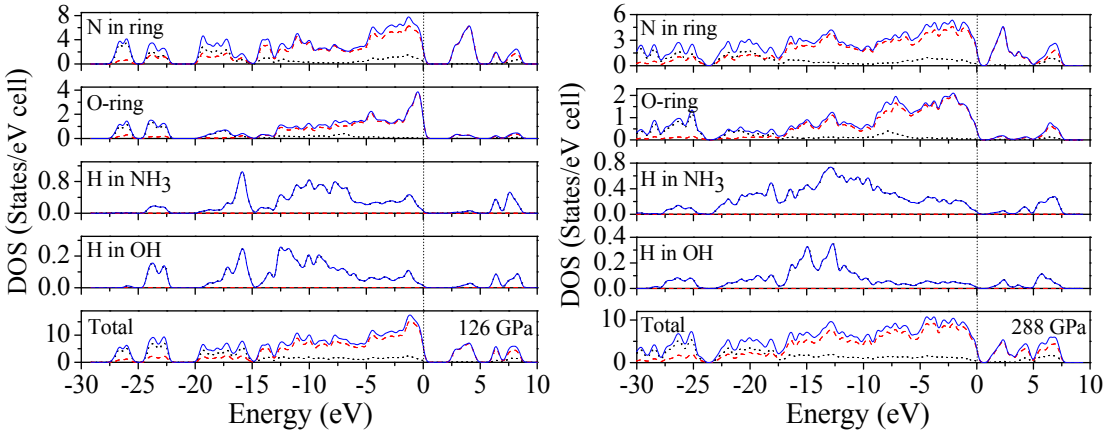


Figure 4. The atom-resolved DOS and PDOS of TKX-50 at 126 GPa, and 288 GPa

The atom-resolved DOS and PDOS of TKX-50 under pressure relate closely to the structural transformation. The DOS and PDOS at 126 and 288 GPa are shown in Figure 4. At 126 GPa, the upper valence bands and the lower conduction bands are predominated by the *p* states of N atoms in tetrazole rings. While at 288 GPa, the upper valence bands and the lower conduction bands are mainly contributed to the N states in tetrazole rings. Furthermore, the N states in tetrazole rings move toward the Fermi level. The structural transformations occur due to the deformation of tetrazole rings.

Conclusions

In this work, periodic DFT calculations have been performed to study the effect of high pressure on the crystal and electronic structures of the energetic salt TKX-50 under hydrostatic pressure of 0−400 GPa. Pressure-induced molecular structure transformations occur at 126, 288, and 334 GPa. At 126 GPa, TKX-50 molecules in crystal are rearranged, and the molecular planarity are improved by the high pressure. With the pressure increasing, the molecular structure and unit cell parameters change gradually. The second structural transformation occurs at 288 GPa with the distortion of the tetrazole rings. When the pressure reaches as high as 334 GPa, the tetrazole ring is severely distorted, and a new structure appeared with the rotation of the molecular conformation.

Acknowledgements

This work was supported by Scientific and Technologial Innovation Programs of Higher Education Institutions in Shanxi (No. 2016157)

References

[1] Forquet, V., C. M. Sabate, H. Chermette, G. Jacob, E. Labarthe, H. Delalu, and C. Darwich. 2016. *Chemistry-An Asian Journal*, 11: 730–744.
[2] Tang, Y. X., C. L. He, L. A. Mitchell, D. A. Parrish, and J. M. Shreeve. 2016. *Journal of Materials Chemistry A*, 4: 3879–3885.
[3] Zhao, G. Z., and Lu M. 2017. *Journal of Energetic Materials*, 35: 63–76.
[4] Keshavarz, M. H., K. Esmaeilpour, M. Oftadeh, and Y. H. Abadi. 2015. *RSC Advances*, 5: 87392–87399.
[5] Liu, Q. Q., B. Jin, R. F. Peng, Z. C. Guo, J. Zhao, Q. C. Zhang, and Y. Shang. 2016. *Journal of Materials Chemistry A*, 4: 4971–4981.
[6] He, P. A., J. G. Zhang, K. Wang, X. Yin, and T. L. Zhang. 2015. *Journal of Organic Chemistry*, 80: 5643–5651.
[7] Myers, T. W., J. A. Bjorgaard, K. E. Brown, D. E. Chavez, S. K. Hanson, R. J. Scharff, S. Tretiak, and J. M. Veauthier. 2016. *Journal of The American Chemical Society*, 138: 4685–4692.
[8] Yin, P., C. L. He, and J. M. Shreeve. 2016. *Chemistry: A European Journal*, 22: 2108–2113.
[9] Ghosh, J., and A. Bhattacharya. 2016. *Chemical Physics*, 464: 26–39.
[10] Yin, P., D. A. Parrish, and J. M. Shreeve. 2015. *Journal of The American Chemical Society*, 137: 4778–4786.
[11] Abe, T., G. H. Tao, Y. H. Joo, R. W. Winter, G. L. Gard, and J. M. Shreeve. 2009. *Chemistry: A European Journal*, 15: 9897–9904.
[12] Klapötke, T. M., and C. M. Sabate. 2008. *European Journal of Inorganic Chemistry*, 34: 5350–5366.
[13] Fischer, D., T. M. Klapötke, and J. Stierstorfer. 2015. *Angewandte Chemie International Edition*, 54: 10299–10302.
[14] Fischer, N., D. Fischer, T. M. Klapötke, D. G. Piercey, and J. Stierstorfer. 2012. *Journal of Materials Chemistry*, 22: 20418–20422.
[15] Ma, S., Y. J. Li, Y. Li, and Y. J. Luo. 2016. *Journal of Molecular Modeling*, 22: 43.
[16] An, Q., T. Cheng, W. A. Goddard, and S. V. Zybin. 2015. *Journal of Physical Chemistry C*, 119: 2196–2207.
[17] Vanderbilt, D. 1990. *Physical Review B*, 41: 7892–7895.
[18] Ceperley, D. M., and B. J. Alder. 1980. *Physical Review Letters*, 45: 566–569.
[19] Perdew, J. P., and A. Zunger. 1981. *Physical Review B*, 23: 5048–5079.
[20] Clark, S. J., M. D. Segall, C. J. Pickard, P. J. Hasnip, M. J. Probert, K. Refson, and M. C. Payne. 2005. *Zeitschrift für Kristallographie*, 220: 567–570.
[21] Materials Studio 6.0 2012. Accelrys.
[22] Perdew, J. P., and Y. Wang. 1992. *Physical Review B*, 45: 13244–13249.
[23] Perdew, J. P., K. Burke, and M. Ernzerhof. 1996. *Physical Review Letters*, 77: 3865–3868.

3D Nonlinear dynamical analysis of cable-stayed offshore structures

†K. Wang, G. K. Er, and V. P. Iu

Department of Civil and Environmental Engineering, University of Macau, Macau SAR, P. R. China.

†Presenting and corresponding author: yb37402@umac.mo

Abstract

The nonlinear vibrations of the moored floating structures under horizontal sinusoidal excitations are studied in three dimensions. Four mooring lines are connected to the floating structure and fixed to the sea bed. The nonlinear equations of motions of the mooring lines are formulated using the cable elements formulated based on the extended Hamilton principle. The floating structure is considered as a rigid body with six degrees of freedom. Then the equations of motion of the floating structure and mooring lines are formulated through their connection conditions. In the last, the equations of motion of the whole structure are analyzed numerically. The influences of different sag-to-span ratio and inclined angle of the mooring cables on the responses of the floating structure are studied.

Keywords: Floating structure, mooring lines, cable elements, connection conditions

Introduction

The moored floating structures can find their applications in ocean engineering to exploit marine resources such as oil, gas and minerals. It consists of the floating platform and mooring cables. If the floating body is subjected to external excitations, the movements of floating body can induce the geometry change of mooring lines. The geometric nonlinearity of the mooring lines plays an important role in the dynamical analysis due to their flexibility. Therefore accurate modeling of mooring cables is necessary for the vibration analysis of the whole structure. Some researches simplified the mooring lines as a linear spring [1, 2, 3] to support the floating body for convenient and efficient analysis. The constant stiffness of the spring is derived and added to the linear stiffness matrix of the floating body. The mooring lines were also modeled as nonlinear spring [4, 5, 6, 7]. The restoring forces from mooring lines are determined on the static analysis of catenary cables with the assumption that the floating body moves slowly. However, the above two methods cannot reflect the real behavior of the cable. Therefore, fully modeling of cable is required. The lumped mass model [8, 9, 10, 11] or the bar element [12, 13] was used for the vibration analysis of mooring lines. The cables were modeled using the finite element method based on the principle of minimum energy including strain energy due to bending and torsion [14, 15], in which the equations of motions of the mooring lines and those of floating body were solved separately and iteratively.

In this paper, the nonlinear vibrations of three-dimensional floating structure and mooring system under the horizontal sinusoidal excitations are studied. The nonlinear equations of motions of the mooring lines are formulated using the 3D cable elements formulated based on the extended Hamilton principle [16, 17]. The cable element is simplified as a flexible tension member without considering its bending and torsion stiffness because of the extremely large ratio of length over cross-section dimension. The floating body is considered as a rigid body with six degrees of freedom, i.e., three translational displacements and three rotational displacements. The equations of motions of both the floating body and mooring system are formulated through their connection conditions and they are solved numerically as a whole.

Problem Statement

Consider the floating structure and mooring system as shown in Fig. 1. It consists of the floating body and four catenary mooring lines C_1, C_2, C_3 and C_4. The floating body and mooring lines are connected through four nodes *A*, *B*, *C* and *D*. *O* is the centroid of the floating body. The other ends of the mooring lines are fixed on the sea bed. w_a, w_b and w_c are the length, height and width of the floating body, respectively; h is the depth of sea; h_s is the submerged height of the floating body in the sea in static state. The top view and side view of the three-dimensional floating system are shown in Fig. 2. The mooring lines C_1, C_2 and C_3, C_4 are symmetric about the y-axis in the plane x_1Oy and x_2Oy, respectively. θ, l and d are the inclined angle, inclined length and initial sag of the mooring line, respectively. w_l is the distance between the nodes *A* and *B*.

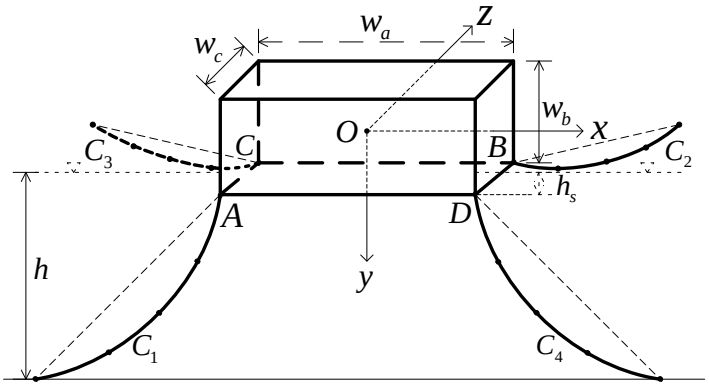
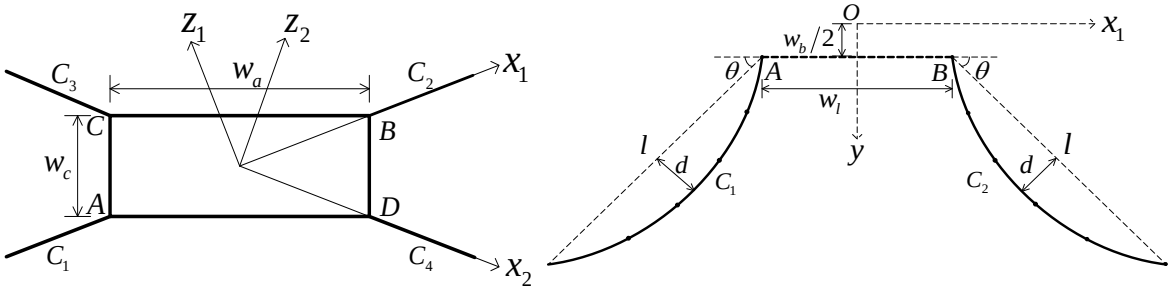


Figure 1. Configuration of the three-dimensional floating system



(a) (b)

Figure 2. (a) Top view (b) Side view of the three-dimensional floating system

The exact catenary profile of the mooring cables in static state is governed by the initial pretension and self-weight of cable. The catenary profile of cable is needed for given sag-to-span ratio d/l.

The submerged height h_s of the floating body in static state is obtained as follows referring to Fig. 3. F_{A1}, F_{A2}, F_{A3}, F_{B1}, F_{B2}, F_{B3}, F_{C1}, F_{C2}, F_{C3}, F_{D1}, F_{D2}, F_{D3} are the components of cable pretensions at nodes *A*, *B*, *C*, and *D* in *x*, *y*, *z* axes, respectively, and

$$F_f + F_{A2} + F_{B2} + F_{C2} + F_{D2} = Mg \tag{1}$$

where M is the mass of the floating body; F_f is the buoyancy of the floating body in the sea and expressed by $F_f = \rho_s g h_s w_a w_c$, in which ρ_s is the density of sea water.

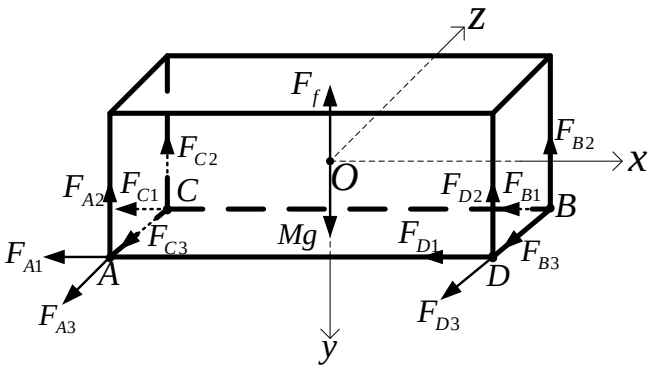


Figure 3. Static equilibrium of the floating body

Nonlinear Dynamical Analysis of the Moored Floating System

The equations of motions of the mooring lines and the floating body are derived individually first and then they are assembled together through their connection conditions.

Finite Element Formulation for the Dynamics of Cable

The equations of motion of the mooring cables are formulated with finite element. The cable element is formulated based on the extended Hamilton principle in the following.

Consider the differential cable element in dynamical state as shown in Fig.4. Let ds and ds' denote the length of cable element in static state and dynamical state, respectively. u, v and w are the dynamical displacements in x, y and z directions, respectively. Then

$$\begin{aligned} (ds)^2 &= (dx)^2 + (dy)^2 \\ (ds')^2 &= (dx+du)^2 + (dy+dv)^2 + (dw)^2 \end{aligned} \tag{2}$$

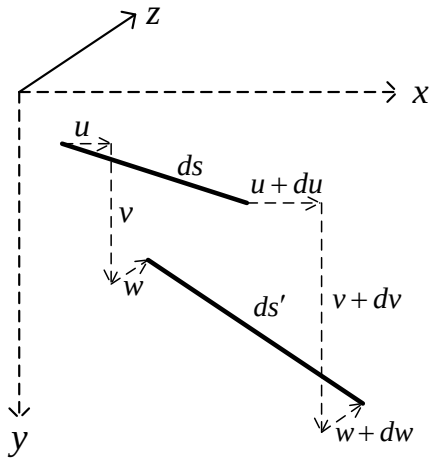


Figure 4. Differential cable element in dynamical state

Retaining the terms up to second order, the axial strain of cable is given by

$$\varepsilon = \frac{ds'-ds}{ds} = x'u' + y'v' + \frac{1}{2}(u')^2 + \frac{1}{2}(v')^2 + \frac{1}{2}(w')^2 \tag{3}$$

where $(\ )' \equiv \partial/\partial s$. Taking the derivatives of ε with respect to u', v' and w', respectively, it gives

$$\frac{\partial \varepsilon}{\partial u'} = x' + u', \quad \frac{\partial \varepsilon}{\partial v'} = y' + v', \quad \frac{\partial \varepsilon}{\partial w'} = w' \tag{4}$$

From which we have the following variation of ε,

$$\delta\varepsilon = \left(x' + u'\right)\delta u' + \left(y' + v'\right)\delta v' + w'\delta w' \tag{5}$$

The variation of potential energy relative to the unstressed state is given by

$$\delta\Pi = \int_0^L \left(T + EA\varepsilon\right)\delta\varepsilon ds \tag{6}$$

The variation of kinetic energy is given by

$$\delta K = -\int_0^L \rho A\left(\ddot{u}\delta u + \ddot{v}\delta v + \ddot{w}\delta w\right)ds = 0 \tag{7}$$

The variation of virtual work associated with gravity and damping force is given by

$$\delta W = \int_0^L \left(\rho Ag\delta v + f_1\delta u + f_2\delta v + f_3\delta w - c_1\dot{u}\delta u - c_2\dot{v}\delta v - c_3\dot{w}\delta w\right)ds \tag{8}$$

where f_1, f_2 and f_3 are the external distributed loads per unit length along x, y and z directions, respectively; c_1, c_2 and c_3 are the damping coefficient per unit length along x, y and z directions, respectively.

Substituting Eq. (5) into Eq. (6), the variation of potential energy can be expressed as

$$\delta\Pi = \int_0^L \left\{EA\varepsilon\left[\delta u'\left(x' + u'\right) + \delta v'\left(y' + v'\right) + \delta w'w'\right] + T\left[\delta u'\left(x' + u'\right) + \delta v'\left(y' + v'\right) + \delta w'w'\right]\right\}ds \tag{9}$$

With Eqs. (7)-(9) and applying the static equilibrium equations of cable element into $\int_{t_1}^{t_2}\left(\delta K - \delta\Pi + \delta W\right)dt = 0$, we have

$$\int_{t_1}^{t_2}\left(\delta K - \delta\Pi' + \delta W'\right)dt = 0 \tag{10}$$

where Π' is the potential energy relative to static equilibrium state and W' is the virtual work done by the external forces from static equilibrium state and done by the damping forces.

The variation of potential energy relative to the static state is expressed as

$$\delta\Pi' = \int_0^L \left\{EA\varepsilon\left[\delta u'\left(x' + u'\right) + \delta v'\left(y' + v'\right) + \delta w'w'\right] + T\left(\delta u'u' + \delta v'v' + \delta w'w'\right)\right\}ds \tag{11}$$

and the variation of the virtual work done from the state of initial profile is expressed as

$$\delta W' = \int_0^L \left(f_1\delta u + f_2\delta v + f_3\delta w - c_1\dot{u}\delta u - c_2\dot{v}\delta v - c_3\dot{w}\delta w\right)ds \tag{12}$$

Let $\mathbf{u}^e = \left\{u^e, v^e, w^e\right\}^T = \mathbf{N}^e\mathbf{d}^e$ and $\mathbf{d}^e = \left\{u^i, v^i, w^i, u^j, v^j, w^j\right\}^T$ for element e, where i and j are two node numbers of element e and $\mathbf{d}^e$ is the displacement vector of element e in local coordinate system O-x_1yz_1 or O-x_2yz_2 of the cable.

The linear shape function of element e is given as follows,

$$\mathbf{N}^e=\begin{bmatrix} N_1 & 0 & 0 & N_2 & 0 & 0 \\ 0 & N_1 & 0 & 0 & N_2 & 0 \\ 0 & 0 & N_1 & 0 & 0 & N_2 \end{bmatrix} \tag{13}$$

where $N_1=1-s/l^e$, $N_2=s/l^e$ and l^e is the length of element e. Denoting $\mathbf{D}^e=d\mathbf{N}^e/ds$ and $\mathbf{x}^e=\{x',y',0\}^T$ for element e, the strain of element e or ε^e can be expressed as

$$\begin{aligned}\varepsilon^e&=\mathbf{x}^{eT}\mathbf{u}'^{eT}+\frac{1}{2}\mathbf{u}'^{eT}\mathbf{u}'^e=\mathbf{x}^{eT}\mathbf{D}^e\mathbf{d}^e+\frac{1}{2}\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e\mathbf{d}^e\\&=\left(\mathbf{x}^{eT}\mathbf{D}^e+\frac{1}{2}\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e\right)\mathbf{d}^e=\mathbf{B}_1^e\mathbf{d}^e\end{aligned} \tag{14}$$

where $\mathbf{B}_1^e=\mathbf{x}^{eT}\mathbf{D}^e+\frac{1}{2}\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e$. The variation of strain $\delta\varepsilon^e$ is obtained to be

$$\delta\varepsilon^e=\left(\mathbf{x}^{eT}\mathbf{D}^e+\frac{1}{2}\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e\right)\delta\mathbf{d}^e=\mathbf{B}_2^e\delta\mathbf{d}^e \tag{15}$$

where $\mathbf{B}_2^e=\mathbf{x}^{eT}\mathbf{D}^e+\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e$. Therefore, the variation of the potential energy is given by

$$\begin{aligned}\delta\Pi'&=\sum_{e=1}^{N^e}\int_0^{l^e}\left[EA\left(\delta\varepsilon^e\right)^T\varepsilon^e+T^e\delta\mathbf{u}'^{eT}\mathbf{u}'^e\right]ds\\&=\sum_{e=1}^{N^e}\int_o^{l^e}\left[EA\delta\mathbf{d}^{eT}\mathbf{B}_2^T\mathbf{B}_1\mathbf{d}^e+T^e\delta\mathbf{d}^{eT}\mathbf{D}^{eT}\mathbf{D}^e\mathbf{d}^e\right]ds\\&=\sum_{e=1}^{N^e}\delta\boldsymbol{d}^{eT}\int_0^{l^e}\left[EA\mathbf{B}_2^T\mathbf{B}_1+T^e\mathbf{D}^{eT}\mathbf{D}^e\right]\mathbf{d}^e ds\end{aligned} \tag{16}$$

where N^e is the total number of elements. The stiffness matrix of element e is then obtained to be

$$\mathbf{k}^e=\int_0^{l_e}\left[EA\mathbf{B}_2^T\mathbf{B}_1+T^e\mathbf{D}^{eT}\mathbf{D}^e\right]ds \tag{17}$$

The variation of kinetic energy of the whole system is given by

$$\delta K=-\sum_{e=1}^{N_e}\int_0^{l^e}\rho A\delta\mathbf{u}^{eT}\ddot{\mathbf{u}}^e ds=-\sum_{e=1}^{N_e}\delta d^{eT}\int_0^{l^e}\rho A\mathbf{N}^{eT}\mathbf{N}^e\ddot{\mathbf{d}}^e ds \tag{18}$$

Thus the mass matrix of element e is obtained to be

$$\mathbf{m}^e=\int_0^{l^e}\rho A\mathbf{N}^{eT}\mathbf{N}^e ds=\frac{\rho Al^e}{6}\begin{bmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix} \tag{19}$$

The lumped mass matrix $\mathbf{m}_l^e$ of element e is

$$\mathbf{m}_l^e = \frac{\rho A l^e}{2}\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{20}$$

In the following numerical analysis, the lumped mass matrix of element *e* is used for simplicity.

The equations of motion for the element *e* in local coordinate systems *O*-x_1yz_1 and *O*-x_2yz_2 are given as follows.

$$m_l^e \ddot{\mathbf{d}}^e + \mathbf{c}^e \dot{\mathbf{d}}^e + \mathbf{k}^e(\mathbf{d}^e)\mathbf{d}^e = \mathbf{f}^e \tag{21}$$

where $\mathbf{c}^e$ and $\mathbf{f}^e$ are the damping matrix and force vector of element e.

Denote $\mathbf{d}_g^e = \left\{u_g^i, v_g^i, w_g^i, u_g^j, v_g^j, w_g^j\right\}^T$ for element e, which is the displacement vector of element e in global coordinate system *O-xyz* of the cable. Using the transformation matrix $\mathbf{T}$ between the local and global coordinate systems and the relationship $\mathbf{d}^e = \mathbf{T}\mathbf{d}_g^e$, Eq. (21) becomes

$$\mathbf{M}_g^e \ddot{\mathbf{d}}_g^e + \mathbf{C}_g^e \dot{\mathbf{d}}_g^e + \mathbf{K}_g^e(\mathbf{d}_g^e)\mathbf{d}_g^e = \mathbf{F}_g^e \tag{22}$$

where $\mathbf{M}_g^e = \mathbf{T}^T\mathbf{M}_l^e\mathbf{T}$, $\mathbf{C}_g^e = \mathbf{T}^T\mathbf{C}_l^e\mathbf{T}$, $\mathbf{K}_g^e = \mathbf{T}^T\mathbf{K}_l^e\mathbf{T}$, and $\mathbf{F}_g^e = \mathbf{T}^T\mathbf{F}_l^e$, which are the stiffness matrix, damping matrix, stiffness matrix and force vector of element *e* in global coordinate system *O-xyz* of the cable.

The equations of motion of the mooring cables are then formulated as follows using the above cable element when ρ is replaced by $\rho - \rho_s$.

$$\mathbf{M}_m \ddot{\mathbf{U}}_m + \mathbf{C}_m \dot{\mathbf{U}}_m + \mathbf{K}_m(\mathbf{U}_m)\mathbf{U}_m = 0 \tag{23}$$

where the subscript *m* denotes the number of mooring lines and $m = 1, 2, 3$ or 4 for this structure. $\mathbf{U}_m$ denotes the global displacement vector of the *m*th mooring cable; $\mathbf{K}_m(\mathbf{U}_m)$ denotes the global stiffness matrix of the *m*th mooring cable; $\mathbf{M}_m$ denotes the global mass matrix of the *m*th mooring cable; $\mathbf{C}_m$ denotes the global damping matrix of the *m*th mooring cable; and $\mathbf{F}_m$ denotes the force vector of the *m*th mooring cable.

Dynamics of the Floating Body

The floating body is considered as a rigid body with six degrees of freedom, which are displacements u_f, v_f, w_f along *x*, *y*, *z* axes and rotations α, β, γ in *xOy*, *xOz*, *yOz* planes, respectively. The equations of motion of the floating body are given as follows referring to Fig. 5.

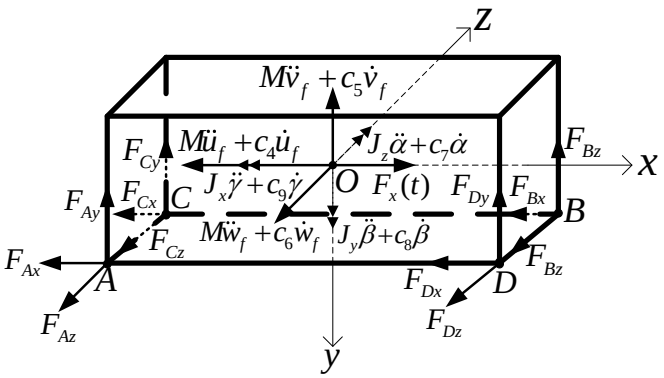


Figure 5. Forces applied on floating body

$$\sum F_x = 0: \ M\ddot{u}_f + c_4\dot{u}_f + F_{Ax} + F_{Bx} + F_{Cx} + F_{Dx} = F_x(t) \tag{24}$$

$$\sum F_y = 0: \ M\ddot{v}_f + c_5\dot{v}_f + F_{Ay} + F_{By} + F_{Cy} + F_{Dy} + F_b = 0 \tag{25}$$

$$\sum F_z = 0: \ M\ddot{w}_f + c_6\dot{w}_f + F_{Az} + F_{Bz} + F_{Cz} + F_{Dz} = 0 \tag{26}$$

$$\begin{aligned}\sum M_z = 0: J_z\ddot{\alpha} + c_7\dot{\alpha} + F_{Ax}\frac{w_b}{2} + F_{Bx}\frac{w_b}{2} + F_{Ay}\frac{w_a}{2} - F_{By}\frac{w_a}{2} \\ + F_{Cx}\frac{w_b}{2} + F_{Dx}\frac{w_b}{2} + F_{Cy}\frac{w_a}{2} - F_{Dy}\frac{w_a}{2} + F_{b1}\frac{2w_a}{3} = 0\end{aligned} \tag{27}$$

$$\begin{aligned}\sum M_y = 0: J_y\ddot{\beta} + c_8\dot{\beta} + F_{Ax}\frac{w_c}{2} - F_{Bx}\frac{w_c}{2} - F_{Az}\frac{w_a}{2} + F_{Bz}\frac{w_a}{2} \\ - F_{Cx}\frac{w_c}{2} + F_{Dx}\frac{w_c}{2} - F_{Cz}\frac{w_a}{2} + F_{Dz}\frac{w_a}{2} = 0\end{aligned} \tag{28}$$

$$\begin{aligned}\sum M_x = 0: J_x\ddot{\gamma} + c_9\dot{\gamma} + F_{Ay}\frac{w_c}{2} - F_{By}\frac{w_c}{2} + F_{Az}\frac{w_b}{2} + F_{Bz}\frac{w_b}{2} \\ - F_{Cy}\frac{w_c}{2} + F_{Dy}\frac{w_c}{2} + F_{Cz}\frac{w_b}{2} + F_{Dz}\frac{w_b}{2} + F_{b2}\frac{2w_c}{3} = 0\end{aligned} \tag{29}$$

where $F_x(t)$ is the external force applied on floating body in x axis; J_z, J_y and J_x are the moment of inertia of the floating body in xOy, xOz and yOz planes, respectively; F_{Ax}, F_{Ay}, F_{Az}, F_{Bx}, F_{By}, F_{Bz}, F_{Cx}, F_{Cy}, F_{Cz}, F_{Dx}, F_{Dy}, F_{Dz} are the dynamical tensions of the cable at node A, B, C, and D in x, y, z axes, respectively, which are induced by the displacements of the elements of mooring lines connected to floating body; F_b, F_{b1} and F_{b2} are the dynamical buoyancy of the floating body due to the change of submerged volume of the floating body, which are expressed as

$$\begin{aligned}F_b &= \rho_s g w_a w_c v_f \\ F_{b1} &= \frac{1}{8}\rho_s g w_a^2 w_c \alpha \\ F_{b2} &= \frac{1}{8}\rho_s g w_a w_c^2 \gamma\end{aligned} \tag{30}$$

Connection Conditions

In order to formulate the equations of motion of the mooring cables and the floating body as a whole, the connection conditions between the mooring lines and floating body are needed. Their relationships are

$$u_A = u + \frac{w_b}{2}\alpha + \frac{w_c}{2}\beta,\ v_A = v + \frac{w_a}{2}\alpha + \frac{w_c}{2}\gamma,\ w_A = w - \frac{w_a}{2}\beta + \frac{w_b}{2}\gamma \tag{31}$$

$$u_B = u + \frac{w_b}{2}\alpha - \frac{w_c}{2}\beta,\ v_B = v - \frac{w_a}{2}\alpha - \frac{w_c}{2}\gamma,\ w_B = w + \frac{w_a}{2}\beta + \frac{w_b}{2}\gamma \tag{32}$$

$$u_C = u + \frac{w_b}{2}\alpha - \frac{w_c}{2}\beta,\ v_C = v + \frac{w_a}{2}\alpha - \frac{w_c}{2}\gamma,\ w_C = w - \frac{w_a}{2}\beta + \frac{w_b}{2}\gamma \tag{33}$$

$$u_D = u + \frac{w_b}{2}\alpha + \frac{w_c}{2}\beta,\ v_D = v - \frac{w_a}{2}\alpha + \frac{w_c}{2}\gamma,\ w_D = w + \frac{w_a}{2}\beta + \frac{w_b}{2}\gamma \tag{34}$$

where u_A, v_A, w_A, u_B, v_B, w_B, u_C, v_C, w_C, u_D, v_D, w_D are the displacements of the nodes *A*, *B*, *C*, and *D* in *x*, *y*, *z* axes, respectively. Then the equations of motion about the nodes A, B, C and D in Eq. (23) are removed and replaced by Eqs. (24)-(29) using the connections conditions given by Eqs. (31)-(34). The variables of displacements related to nodes A, B, C and D in other equations of motion in Eq. (23) are also expressed by Eqs. (31)-(34). Thus the final equations of motion of the whole system are obtained to be

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{C}\dot{\mathbf{U}} + \mathbf{K}(\mathbf{U})\mathbf{U} = \mathbf{F} \tag{35}$$

where **U** is the global displacement vector of the whole system; **K**(**U**) is the global stiffness matrix of the whole system; **C** is the Rayleigh damping matrix.

Numerical example

Consider the mooring cables and floating body with their parameter values shown in Table 1 and Table 2, respectively. The density of sea water is $\rho_s = 1.025\times10^3$ kg/m^3.

Table 1. Properties of mooring cables

Parameter	Value
Young's modulus E (N/m^2)	2×10^{11}
Diameter D (m)	0.1
Mass density ρ (kg/m^3)	8×10^3
Damping ratio ξ	0.03
Sea depth h (m)	100
Inclined angle θ (degree)	50
Sag-to-span ratio d/l	1/60

Table 2. Properties of floating body

Parameter	Value
Length w_a (m)	26
Height w_b (m)	5
Width w_c (m)	10
Mass M (kg)	1.2×10^5

The submerged height of the floating body in static state is calculated with Eq. (1) to be $h_s = 0.896$ m, Consider that the floating body is subjected to the sinusoidal force $F_x(t) = \mathrm{A}\sin(\omega t)$ with amplitude A being 10^5 N and $\omega = 1.6$ rad/s. Each mooring line is divided into 11 elements. The system starts to move from static state. The time history of the responses of the floating body and the maximum cable tensile force is shown in Fig. 6.

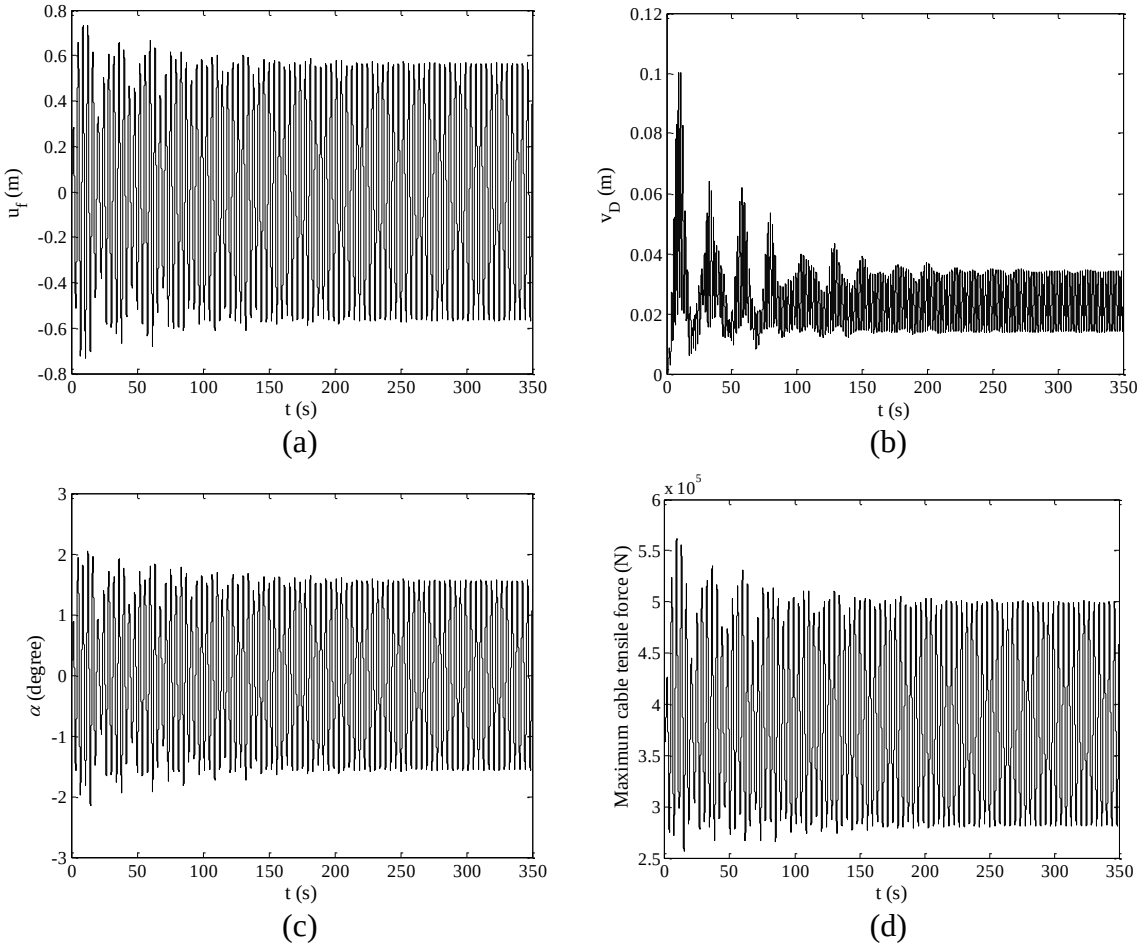


Figure 6. Time history of (a) Displacement along *x*-axis (b) Displacement along *y*-axis (c) Rotational angle in *xOy* plane (d) Maximum tensile force in cable with $d/l = 1/60$ and $\theta = 50^\circ$

When the initial inclined angle equals 50 º, the response amplitudes of the floating body and the maximum tensile force in the cable at steady state are presented in Fig. 7 for different sag-to-span ratios of cable. It is seen that the displacements along *x*-axis, y-axis and the rotational angle in *xOy* plane of floating body decreases obviously as the sag-to-span ratio decreases

from 1/45 to 1/80. This means that the displacements along *x*-axis, y-axis and the rotational angle in *xOy* plane of floating body are much influenced by the sag-to-span ratio of the cable. The maximum cable tensile force decreases as the sag-to-span ratio decreases from 1/45 to 1/60. After that, it increases as the sag-to-span ratio decreases from 1/60 to 1/80. It means that the maximum cable tensile force is much influenced by the sag-to-span ratio of the cable.

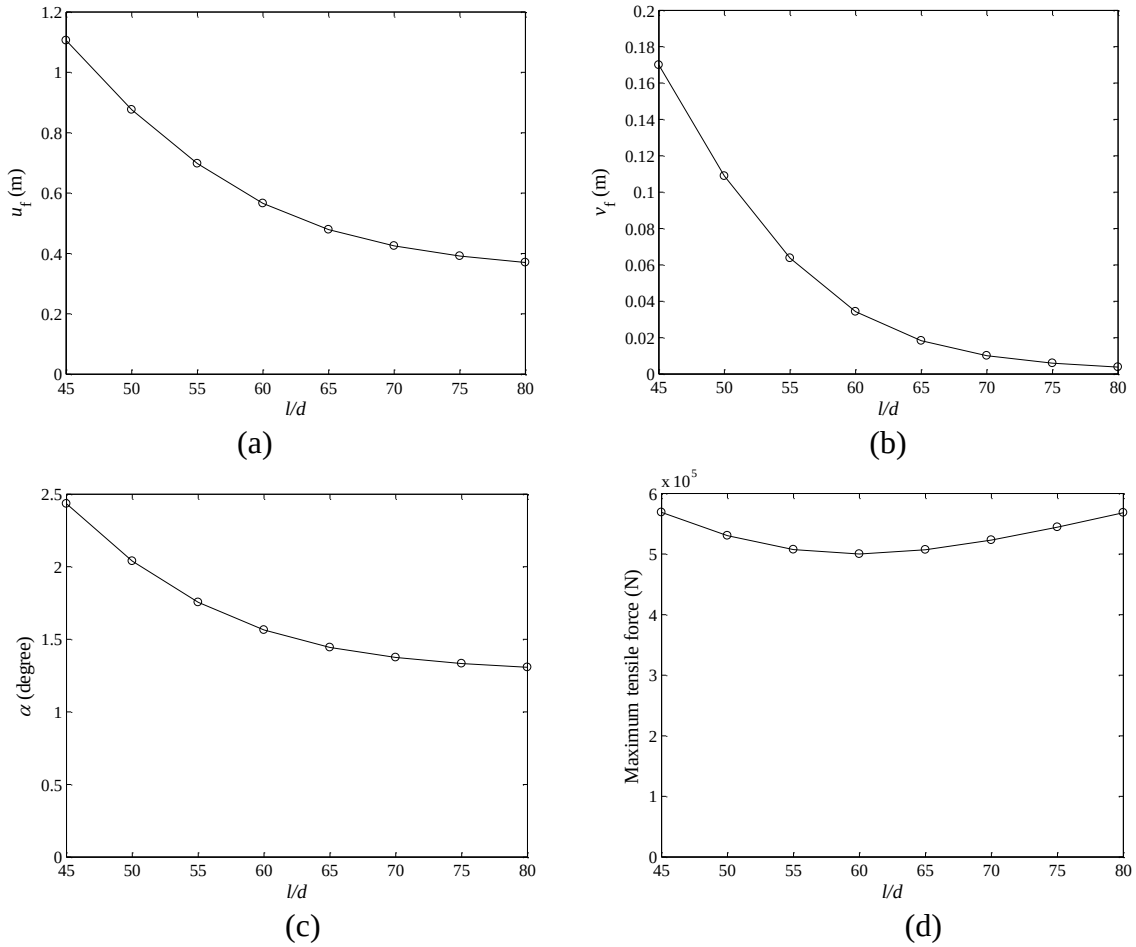


Figure 7. Amplitude of (a) Displacement along *x*-axis (b) Displacement along *y*-axis (c) Rotational angle in *xOy* plane (d) Maximum tensile force in cable at steady state for different sag-to-span ratios of cable with $\theta = 50^{\circ}$

When the sag-to-span ratio equals 1/60, the response amplitudes of the floating body and the maximum tensile force in the cable at steady state are presented in Fig. 8 for different initial inclined angles of cable. It is seen that the displacements along *x*-axis, y-axis and the rotational angle in *xOy* plane of floating body decreases obviously as the inclined angle increases from 40 º to 54 º. This means that the displacements in *x*-axis, y-axis and the rotational angle in *xOy* plane of floating body are much influenced by the initial inclined angle of the cable. The maximum cable tensile force decreases as the inclined angle increases from 40 º to 50 º. After that, it increases as the inclined angle increases from 50 º to 54 º. It means that the maximum cable tensile force is much influenced by the inclined angle of the cable.

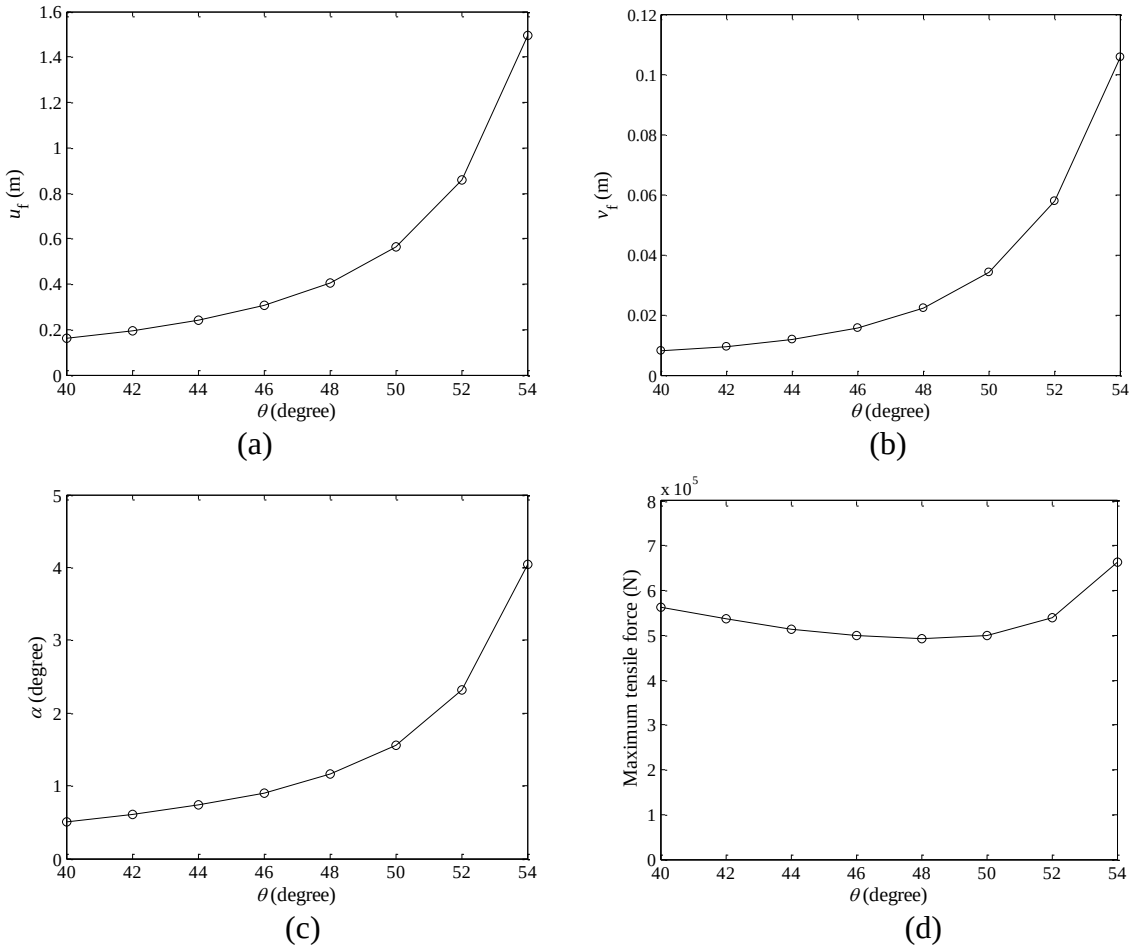


Figure 8. Amplitude of (a) Displacement along *x*-axis (b) Displacement along *y*-axis (c) Rotational angle in *xOy* plane (d) Maximum tensile force in cable at steady state for different initial inclined angle of cable with $d/l = 1/60$

Conclusions

The nonlinear vibrations of the three-dimensional floating structure moored by cables are analyzed. The floating body is modeled as a rigid body with six degrees of freedom. The mooring cables are modeled with the 3D nonlinear cable elements which are formulated with the extended Hamilton principle. The connection conditions between the mooring cables and the floating platform are introduced and hence the nonlinear equations of motions of both the mooring cables and floating platform are formulated as a whole through these connection conditions. Then the nonlinear equations of motion of the system under horizontal sinusoidal excitations are solved numerically as a whole. The influence of the mooring cables on the responses of the floating body and the maximum cable tensile force are discussed for different values of initial sag-to-span ratio or initial inclined angle of the mooring cables. It is seen from the numerical results that the initial sag-to-span ratio and the initial inclined angle of the mooring cables at static state have much influence on the dynamical displacements and rotations of the floating platform and the cable tensile forces.

Acknowledgement

The research work presented in this paper were obtained under the supports of the Research Committee of University of Macau (Grant No. MYRG2014-00084-FST) and the Science and Technology Development Fund of Macau (Grant No. 043/2013/A).

References

[1] Yamamoto, T., Yoshida, A. and Ijima, T. (1980) Dynamics of elastically moored floating objects, *Applied Ocean Research* **2**, 85–92.
[2] Sannasiraj, S. A., Sundar, V. and Sundarvadivelu, R. (1998) Mooring forces and motion response of pontoon-type floating breakwaters, *Ocean Engineering* **26**, 27-48.
[3] Tang, H. J., Chen, C. C. and Chen, W. M. (2011) Dynamics of dual pontoon floating structure for cage aquaculture in a two-dimensional numerical wave tank, *Journal of Fluids and Structures*, **27**, 918-936.
[4] Agarwal, A. K. and Jain, A. K. (2003) Dynamic behavior of offshore spar platforms under regular sea waves, *Ocean Engineering* **30**, 487-516.
[5] Esmailzadeh, E. and Goodarzi, A. (2001) Stability ananlysis of a CALM floating offshore structure, *International Journal of Non-Linear Mechanics* **36**, 917–926.
[6] Umar, A and Datta, T. K. (2003) Nonlinear response of a moored buoy, *Ocean Engineering* **30**, 1625-1646.
[7] Rosales, M. B. and Filipich, C. P. (2006) Full modeling of the mooring non-linearity in a two-dimensional floating structure, *International Journal of Non-Linear Mechanics* **41**, 1–17.
[8] Huang, S. (1994) Dynamic analysis of three-dimensional marine cable, *Ocean Engineering* **6**, 587–605.
[9] Driscoll, F., Lueck, R. and Nahon, M. (2000) Development and validation of a lumped-mass dynamics model of a deep-sea ROV system, *Applied Ocean Research* **22**, 169–182.
[10] Buckham, B., Nahon, M., Seto, M., Zhao, X. and Lambert, C. (2003) Dynamics and control of a towed underwater vehicle system, part I: model development, *Ocean Engineering* **30**, 453-470.
[11] Zhu, X. Q. and Yoo, W. S. (2016) Dynamic analysis of a floating spherical buoy fastened by mooring cables, *Ocean Engineering* **121**, 462–471.
[12] Garrett, D. L. (1982) Dynamic analysis of slender rods, *Journal of Energy Resources Technology* **104**, 302-306.
[13] Garrett, D. L. (2005) Coupled analysis of floating production system, *Ocean Engineering* **32**, 802–816.
[14] Kim, B. W., Sung, H. G., Hong, S. Y. and Jung, H. J. (2010) Finite element nonlinear analysis for catenary structure considering elastic deformation, *Computer Modeling in Engineering and Sciences* **63**, 29-45.
[15] Kim, B. W., Sung, H. G., Kim, J. H. and Hong, S. Y. (2013) Comparision of linear spring and nonlinear FEM methods in dynamic coupled analysis of floating structure and mooring system, *Journal of Fluid and Structures* **42**, 205–227.
[16] Pai, P. F. (2007) *Highly Flexible Structure: Modeling, Computation, and Experimentation*. American Institute of Aeronautics and Astronautics, Inc., Reston.
[17] Wang, K., Er, G. K. and Iu, V. P. (2016) Finite element method for nonlinear dynamics of inclined cable with support motion, in: *The 8th Asian Conference on Multibody Dynamics*, Kanazawa, Japan.

Flows in out-phase slip-patterned micro-channels using boundary element methods

***Chandra Shekhar Nishad[1], Anirban Chandra[2], and G.P. Raja Sekhar[1]**

[1]Department of Mathematics, Indian Institute of Technology Kharagpur, India.
[2]Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, India

*Presenting & Corresponding author: csnishad@maths.iitkgp.ernet.in

Abstract

Flow through two-dimensional rectangular micro-channel with patterned slip on the walls under the low-Reynolds number limit ($Re << 1$) is studied using boundary element method (BEM). We assume that the pattern of the slip on the upper and the lower walls maintain a phase difference (i.e., out-phase configuration). We considered two subcases of out-phase patterned slip, namely large and fine depending on the characteristic length of the patterning. In order to obtain a deep insight of flow mechanics, we investigated the streamlines, velocity profiles, shear stress, and pressure gradients with varying slip-length (l_s).

Keywords: Stokes equation, micro-channel, slip patterning, boundary element method.

Introduction

In order to model the flow of a viscous incompressible fluid over solid boundaries, proper understanding about the boundary conditions at the interface of solid and liquid are required. The most applicable boundary condition is the no-slip condition which generally state that the fluid velocity at the wall is equal to the wall velocity. The no-slip boundary condition does not work in certain occasions like flows over hydrophobic surfaces, problems involving multiple interfaces, and flow of rarefied fluids. In the past, several researchers have proposed various models to correctly defines the interaction between a solid wall and a liquid. A detailed discussion about the slip boundary condition can be found in [1].

The range of slip-length is affected by several factors. A perfect slip (large slip) can be observed in situations where the nanobubbles are trapped over hydrophobic surfaces. In most of the micro-fluidic applications, there is a high possibility of encountering the irregular boundaries. In order to handle such complex boundary conditions, attempting analytical solution might be impossible. In that case one has to rely on numerical solution.

There are several domain discretization techniques (such as finite difference method (FDM), finite element method (FEM), and finite volume method (FVM)) available in literature to solve such fluid flow problems involving complex boundaries. In case of linear operators another technique available in literature which is based on the integral equation method known as boundary element method (BEM). The major advantage of this technique over the others is that the dimension of the solution domain is reduced by order one, which results in saving computational resources.

Understanding flow through rectangular micro-channels is useful in designing and optimization of lab on chips systems. At the micro, and nano length scales, the ability to fabricate patterns and structures has produced a board area of scientific research. The interaction between solid wall and liquid defines slip or no-slip conditions on a wall. The hydrophilic and hydrophobic surfaces are characterized by no-slip and partial slip boundary conditions. In order to design hydrophobic/ hydrophilic surfaces, various investigations on chemical modification of surfaces have been done ([2] [3]). In this connection, surfaces coated with alternate slip and no-slip are important in reducing drag [4]. The ability to manipulate the flow to enhance the mixing in the low-Reynolds number limit is another very interesting phenomena of the surfaces coated with alternate slip and no-slip boundary conditions. In [1], the idea of in-phase patterned slip has been used, while studying Stokes flow in a two-dimensional rectangular micro-channels. In this study, we extend the work of [1] by considering the out-phase configuration of patterned slip in rectangular micro-channels.

Mathematical Formulation

Let us consider a steady, two-dimensional,viscous incompressible Newtonian flow at low Reynolds number that is governed by Stokes equation together with the continuity equation given by the respective non-dimensional form,

$$\nabla P = \nabla^2 \boldsymbol{U},\ \nabla \cdot \boldsymbol{U} = 0, \tag{1}$$

where the non dimesionalization is done using the characteristic variables,

$$\boldsymbol{U} = \frac{\boldsymbol{u}}{\tilde{U}},\ \boldsymbol{X} = \frac{\boldsymbol{x}}{L},\ P = \frac{pL}{\mu \tilde{U}}\ .$$

Eliminating the pressure term from Eq. (1) while introducing stream function, Stokes equation can be reduced to a biharmonic equation which is given by,

$$\nabla^4 \psi = 0. \tag{2}$$

In order to solve Eq. (2) using boundary element method, the biharmonic equation is rewritten as a coupled system of Poisson and Laplace equation in terms of stream function-vorticity variables,

$$\nabla^2 \psi = -\omega,\ \nabla^2 \omega = 0. \tag{3}$$

In order to solve Eq. (3) we use direct biharmonic boundary integral equation methods (BBIE). Of course one may recall boundary integral formulation based on velocity-traction variables [15]. However, here we restrict to stream function-vorticity formulation without debating much on the relative advantages and disadvantages. In the literature, several authors have reviewed the BBIE methods ([5]-[13]) in significant detail. Therefore, in this study we present this method in brief. Let us denote a general field point by $p(X,Y)$ and an integration point on the boundary by $q(X_0,Y_0)$, so that $p \in \Omega \cup \partial\Omega$ and $q \in \partial\Omega$. Let G^L and G^B be the fundamental solutions

corresponding to Laplacian and biharmonic operators which satisfies $\nabla^2 G^L = \delta(|q-p|)$, and $\nabla^4 G^B = \delta(|q-p|)$, where δ is the dirac delta function and $G^L = log|p-q|$, and $G^B = |p-q|^2 (log|p-q|-1)$ [14].

Application of Green's second identity to Eq. (3) gives rise to the following pair of integral equations at a general field point p,

$$\lambda(p)\psi(p) = \int_{\partial\Omega}[\psi(q)\frac{\partial G^L(p,q)}{\partial n_q} - G^L(p,q)\frac{\partial \psi(q)}{\partial n_q}]ds(q) - \frac{1}{4}\int_{\partial\Omega}[\omega(q)\frac{\partial G^B(p,q)}{\partial n_q} - G^B(p,q)\frac{\partial \omega(q)}{\partial n_q}]ds(q), \tag{4}$$

$$\lambda(p)\omega(p) = \int_{\partial\Omega}[\omega(q)\frac{\partial G^L(p,q)}{\partial n_q} - G^L(p,q)\frac{\partial \omega(q)}{\partial n_q}]ds(q), \tag{5}$$

where, $\lambda(p)$ is defined by,

$$\lambda(p) = \begin{cases} 2\pi, & if \quad p \in \Omega, \\ \pi, & if \quad p \in \partial\Omega, \\ 0, & if \quad p \notin \Omega \cup \partial\Omega. \end{cases} \tag{6}$$

We discretized the boundary $\partial\Omega$ into N constant elements $\partial\Omega_j$ containing the mid-element boundary nodes $q_j \quad (j = 1,2,...,N)$. We use piecewise constant functions ψ_j, $\frac{\partial \psi_j}{\partial n}$, ω_j, and $\frac{\partial \omega_j}{\partial n}$, for $j = 1,2,...,N$ to approximate the values of ψ, $\frac{\partial \psi}{\partial n}$, ω, and $\frac{\partial \omega}{\partial n}$ over each element. Applying the discretized form of Eqs. (4) and (5) at the mid-point $p \equiv q_i \in \partial\Omega_i \quad (i = 1,2,...,N)$ of each element, gives a system of linear equations,

$$A_{ij}\psi_j + B_{ij}\frac{\partial \psi_j}{\partial n_q} + C_{ij}\omega_j + D_{ij}\frac{\partial \omega_j}{\partial n_q} = 0, \tag{7}$$

$$A_{ij}\omega_j + B_{ij}\frac{\partial \omega_j}{\partial n_q} = 0, \tag{8}$$

where the coefficients A_{ij}, B_{ij}, C_{ij}, and D_{ij} are given by,

$$A_{ij} = \int_{q\in\partial\Omega_j}\frac{\partial G^L(q_i,q)}{\partial n_q}ds(q) - \lambda_j\delta_{ij}, \quad C_{ij} = -\frac{1}{4}\int_{q\in\partial\Omega_j}\frac{\partial G^B(q_i,q)}{\partial n_q}ds(q),$$

$$B_{ij} = -\int_{q\in\partial\Omega_j} G^L(q_i,q)ds(q), \qquad D_{ij} = \frac{1}{4}\int_{q\in\partial\Omega_j} G^B(q_i,q)ds(q).$$

where δ_{ij} denotes the Kronecker delta function. In this paper, we use the analytical expression of the above integrals derived by [5]. On solving Eqs. (7) and (8), one obtain the unknown set of boundary variables at each node $q_j \in \partial\Omega_j$, $j = 1,2,...,N$. Subsequently, one can obtain the values of ψ and ω at a general field point $p \in \Omega$ by solving the Eqs. (7) and (8) for $\lambda(p) = 2\pi$.

Once the values of stream function (ψ) and vorticity (ω) are known inside the domain, one can compute the velocity profiles and pressure gradients by taking their appropriate spatial derivatives, and also evaluate wall shear stress. please refer [1] for details.

Results and Discussion

In this study, we have considered out-phase patterned slip in rectangular micro-channels (see Figs. 1 and 8). Based on the characteristic length of the patterning, we further considered two subcases namely large and fine patterns. When the characteristic length of the pattering (a) is bigger than the width of the channel, we have large patterning, while fine pattering arises when the characteristic length of pattering (a) is smaller than the width of the channel. Depending on the location of the slip or no-slip surface, the boundary conditions are used as follows,

$$\begin{aligned}
&\text{On the top wall (no slip surface)} &&: \quad \psi = 1, \ \frac{\partial \psi}{\partial Y} = 0 : (NS)\\
&\text{On the top wall (slip surface)} &&: \quad \psi = 1, \ \frac{\partial \psi}{\partial Y} + l_s\, \omega = 0 : (S)\\
&\text{On the bottom wall (no slip surface)} &&: \quad \psi = 0, \ \frac{\partial \psi}{\partial Y} = 0 : (NS)\\
&\text{On the bottom wall (slip surface)} &&: \quad \psi = 0, \ \frac{\partial \psi}{\partial Y} - l_s\, \omega = 0 : (S)\\
&\text{Inlet} &&: \quad \psi = 3Y^2 - 2Y^3, \ \ \omega = 12Y - 6\\
&\text{Outlet} &&: \quad \frac{\partial \psi}{\partial X} = 0, \ \frac{\partial \omega}{\partial X} = 0
\end{aligned}$$

where l_s represents the dimensionless slip length.

In the present study, we use 240 constant elements to discertize the boundary in a counter-clockwise sense.

Out-phase patterned slip

Here, we consider out-phase configuration of slip-patterning in rectangular micro-channels (see Figs. 1 and 8). The length and width of the channels are 10 and 1 units respectively. The no-slip and slip boundary conditions on the channel wall are characterized by black and grey line segments on the wall. The value of periodicity (a) is same for both the no-slip and slip regime. $a = 2.5$ represents the large patterning, while $a = 0.5$ corresponds to fine patterning. To ignore the effects of inlet and outlet in our discussion we present the results for a certain part of the channel. The two subcases of out-phase slip-patterning (large and fine) have been discussed below.

Large Patterning

In this case, the periodicity of patterning (a) is taken as 2.5 (refer Fig. 1). Fig. 2 shows the contour plots of the streamlines for different values of l_s ranging from 0.1 to 10. It is observed that the streamline profiles look qualitatively similar for different values of l_s.

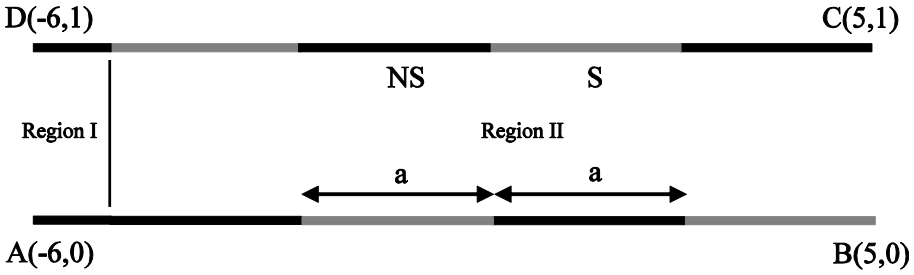


Figure 1. Schematic of the rectangular channel with out-phase large patterned slip. Region I denotes the extension, which is introduced to match the inlet boundary condition. Region II is the actual channel through which the flow is desired.

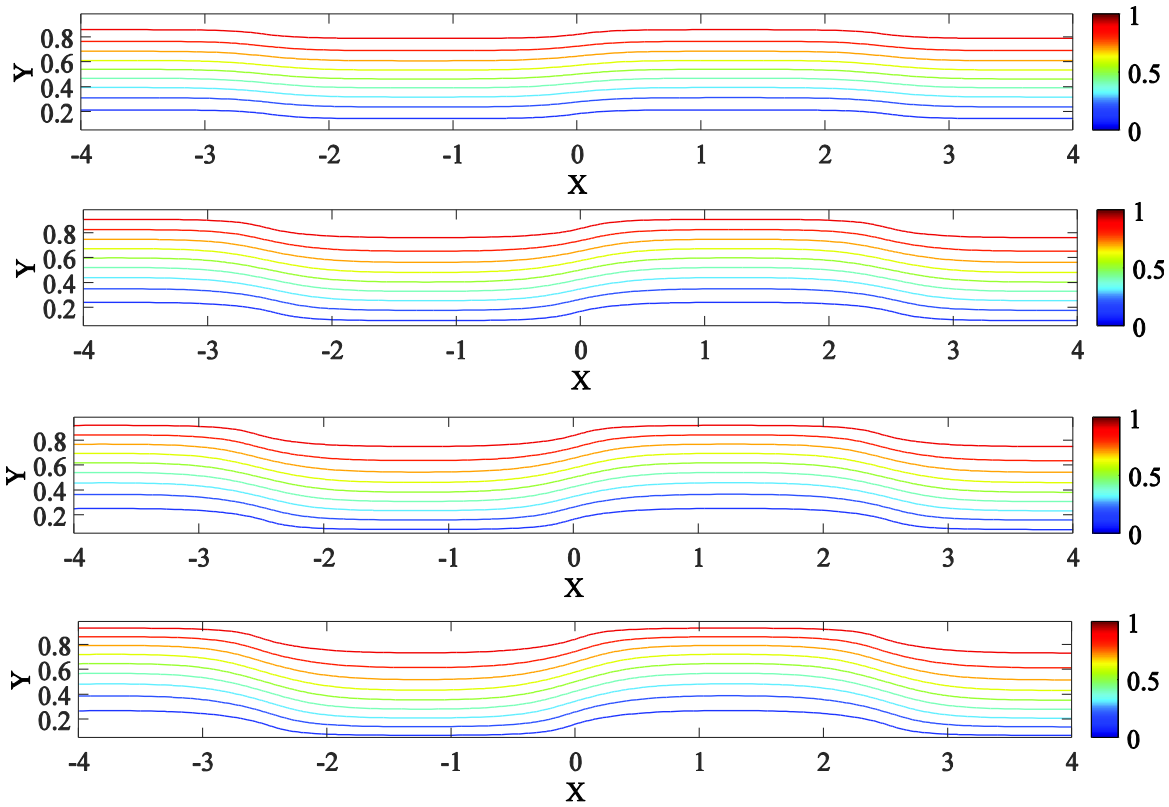
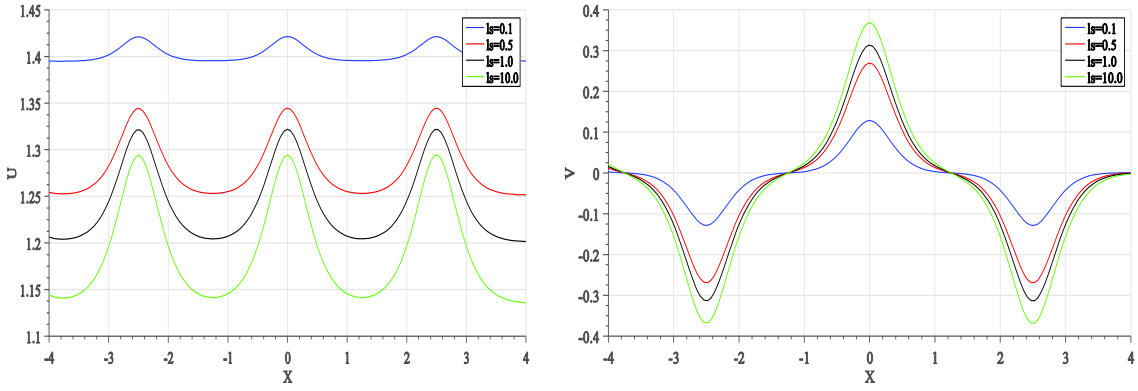


Figure 2. Contour plots of the streamlines for out-phase large patterned slip in rectangular channels with different values of $l_s = 0.1, 0.5, 1$ and 10.

The variation of the horizontal (U) and vertical (V) velocities with X at $Y = 0.5$ are shown in Fig. 3. The U velocity has significant dependency on the slip-length : with decreasing l_s the value of U at $Y = 0.5$ becomes almost constant (similar to the case where complete no-slip boundary condition is applied on both the walls). On the other hand with increasing l_s, the overall center-line velocity decreases. This fact can be understood on the basis of principle of mass conservation. Interestingly, the maximum U for a particular l_s is attained at the points where a change in boundary conditions take place, i.e., at $X = -2.5$, 0 and 2.5. The variation of V with X at $Y = 0.5$ is also a function of l_s and its maximum value increases as l_s increases. Moreover the peaks in V are obtained where there is a change of boundary conditions. Similar variation has been observed in the V velocity with X at other values of Y.

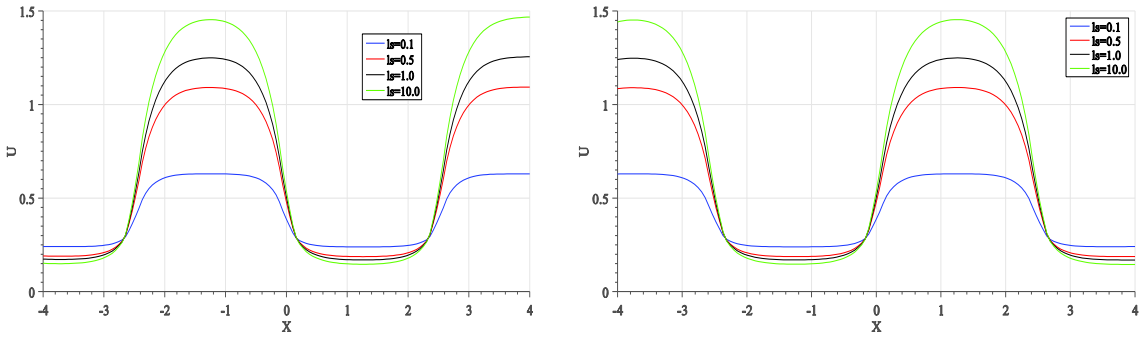


(a) U vs X

(b) V vs X

Figure 3. Variation of the horizontal (U) and vertical (V) velocities with X at $Y = 0.5$ for varying slip-length l_s from 0.1 to 10.

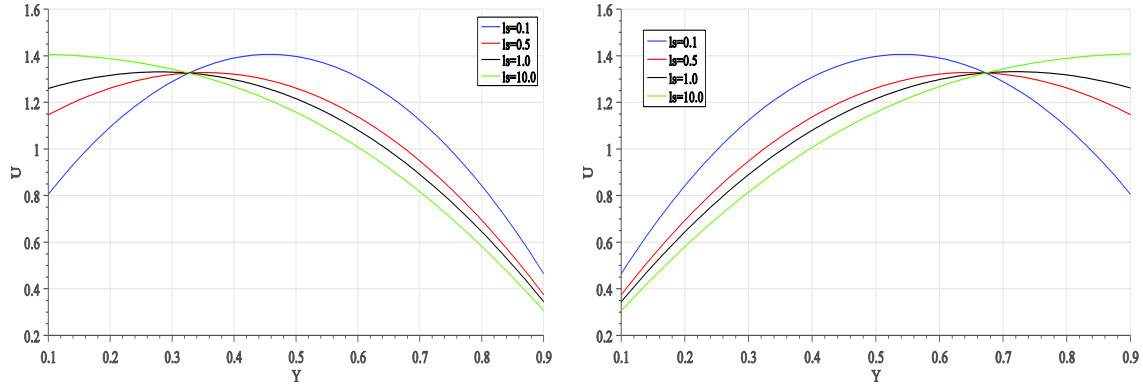
In order to gain a proper understanding of the effect of patterns on the walls, U is plotted against X for $Y = 0.05$ and 0.95, see Fig. 4. It is observed that in the no-slip regime, the effect of the slip-length is not that prominent, and thus U does not seem to vary much with l_s as opposed to the slip-regime where a prominent variation with l_s is noted. In fact when l_s is significantly large then the value of U approaches 1.5, which is theoretically the maximum velocity possible for a channel flow, with the particular inlet conditions that we have utilized. Furthermore, for a particular value of l_s, the U velocity profiles presented in Figs. 4 (a) and 4 (b) are completely out of phase. This is simply because of the fact that the patterning is out of phase, and near the boundaries the flow profiles closely replicate the boundary conditions implemented at the walls.



(a) Y=0.05 (b) Y=0.95

Figure 4. Variation of the horizontal velocity (U) with X at two different values of Y for varying slip length, l_s, from 0.1 to 10.

The variation of U with Y, presented in Fig. 5, shows that at the no-slip end, the U velocity tends to zero and in the slip regime the horizontal velocity varies according to the slip-length. Another interesting fact is that, in Figs. 5 (a) and 5 (b), there exist a particular value of Y for which the U velocity remains the same with varying l_s.



(a) X=-1.7 (b) X=1.7

Figure 5. Variation of the horizontal velocity (U) with Y at two different values of X for varying slip length, l_s, from 0.1 to 10.

In order to understand the effect of pressure drop, the variation of pressure gradients (PG) in X-direction against the horizontal distance X at the channel center-line with varying l_s is plotted in Fig. 6. Dips in the pressure gradient (PG) profile are observed at the places where the boundary condition changes. Also, when the slip-length is very low ($l_s = 0.1$ almost represents the no-slip condition) the value of PG tends to a constant value, i.e., the amplitude of variation of PG is reduced. With increasing slip-length there is a net upward shift (tending towards zero) of the PG profile which is justified because, as the slip-length increases there is less resistance to the flow and thus a smaller pressure gradient would be sufficient to drive the flow.

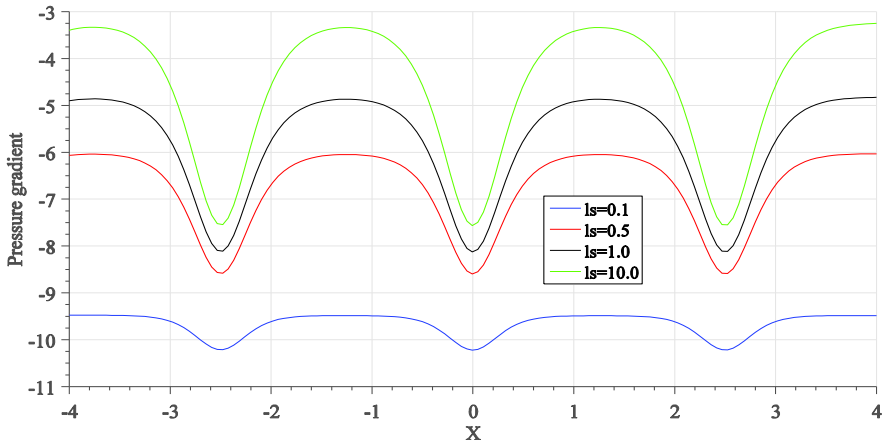


Figure 6. Plot of the pressure gradient (PG) in the X direction with X at $Y = 0.5$ for varying l_s , from 0.1 to 10.

The shear stress on the top wall ($Y = 1$) is plotted with X in Fig. 7. In the slip regime for large value of slip-length (l_s =10) shear stress vanishes on the wall, while in the no-slip regime there is a small variation in the shear stress profiles.

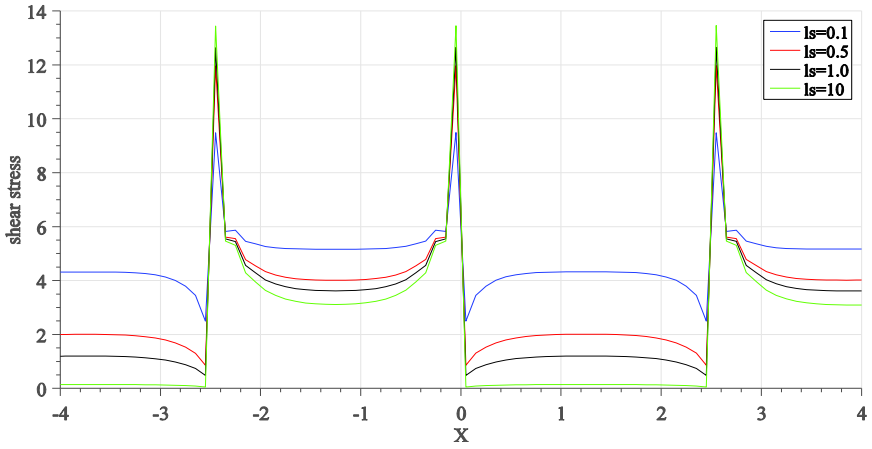


Figure 7. Plot of the shear stress with X at $Y = 1$ for varying l_s , from 0.1 to 10.

Fine Patterning

In this case $a = 0.5$ as shown in Fig. 8. It is observed from Fig. 9 that the streamline profiles remain qualitatively same with varying l_s .

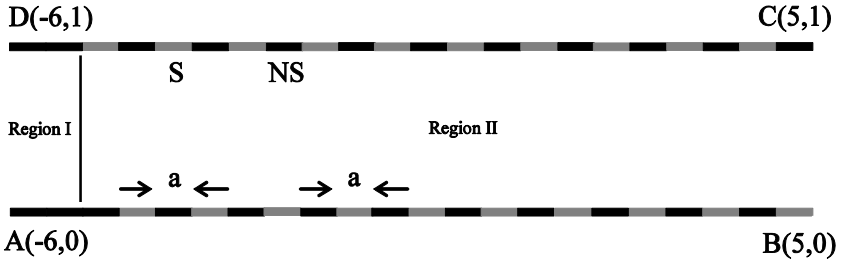


Figure 8. Schematic of the rectangular channel with out-phase fine patterned slip. Region I denotes the extension, which is introduced to match the inlet boundary condition. Region II is the actual channel through which the flow is desired.

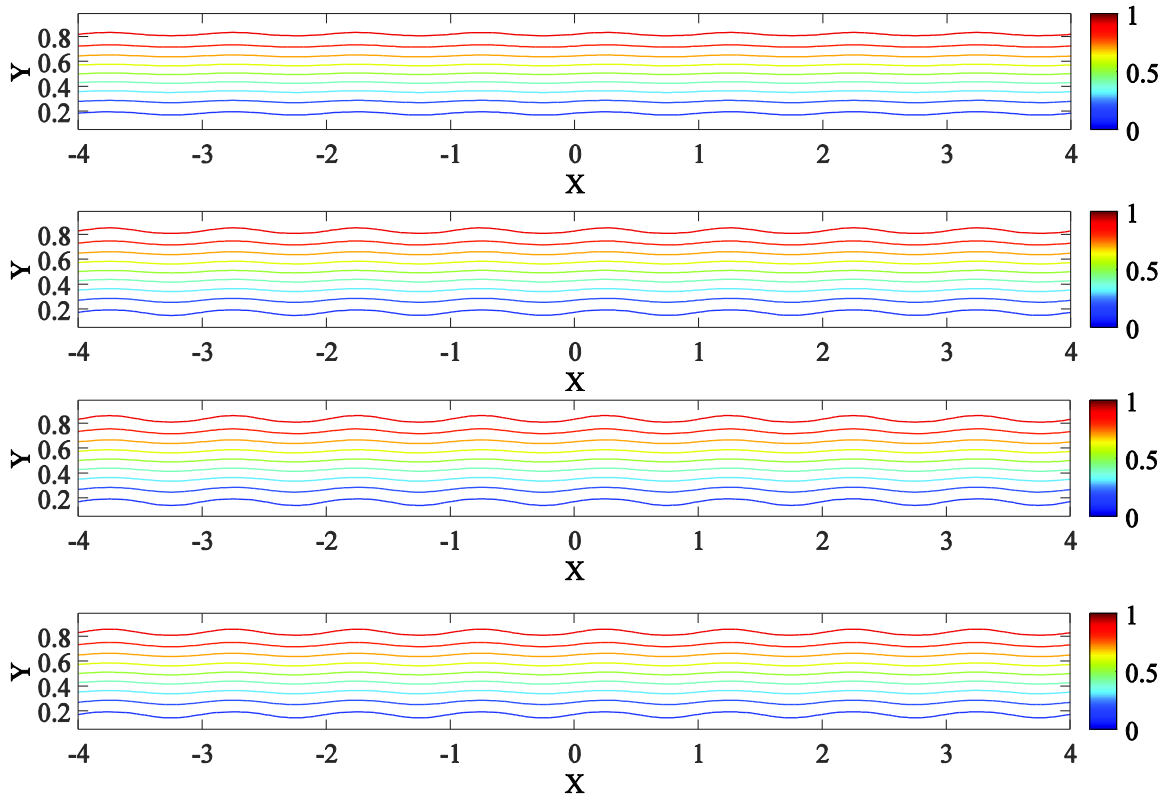
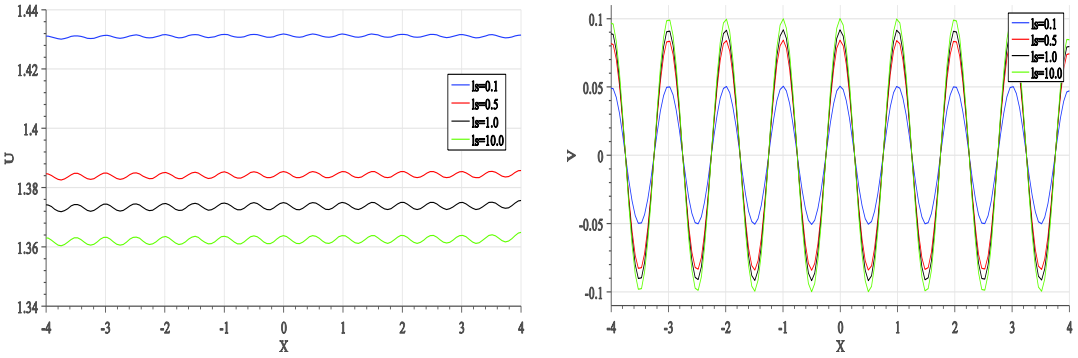


Figure 9. Contour plots of streamline for out-phase fine patterned slip in rectangular micro-channels with different values of $l_s = 0.1, 0.5, 1$ **and** 10**.**

The U and V velocities are plotted with X in Fig. 10. The U velocity has greater dependency over the slip-length l_s. As opposed to the earlier case of out-phase large patterned slip, the amplitude of variation of U is very marginal. This is mainly because of the fact that the characteristic length of patterning is very small which does not let the flow develop as in the previous case. Also, with increase in l_s the center-line U velocity decreases. This fact can be

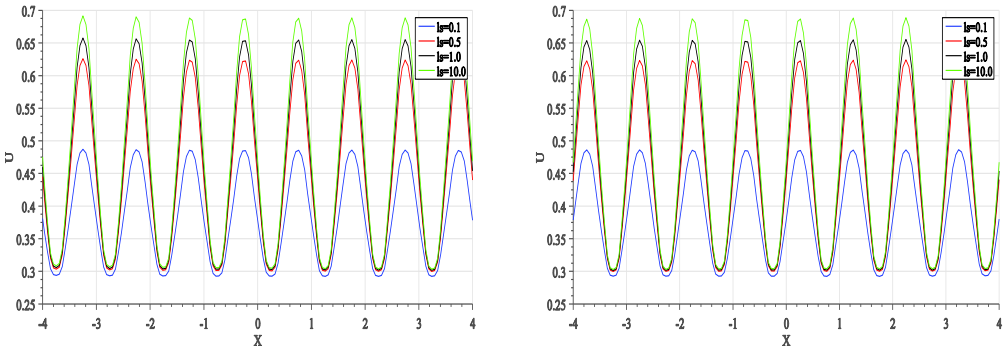
understood on the basis of principle of mass conservation. From Fig. 10 (b) it is recognized that the variation in V velocity with X is quite similar to the previous case of out-phase large patterned slip but there is a remarkable difference in the magnitude of the variation.



(a) U vs X (b) V vs X

Figure 10. Variation of the horizontal (U) and vertical (V) velocities with X at $Y = 0.5$ for different values of l_s, ranging from 0.1 to 10.

To understand the effect of the boundaries, the variation of U with X at $Y = 0.05$ and 0.95 is shown in Fig. 11. In the slip-regime a prominent variation with l_s is observed, more the slip-length more is the U velocity. Interestingly, when l_s is significantly large, in contrary to the case of large out-phase slip, the value of U does not approach 1.5. The reason for this is that because of the small characteristic length a the flow does not develop fully. In Fig. 11, it can be clearly seen that the U velocity profile at $Y = 0.05$ and $Y = 0.95$ are completely out of phase for varying slip length $l_s = 0.1, 0.5, 1,$ and 10.

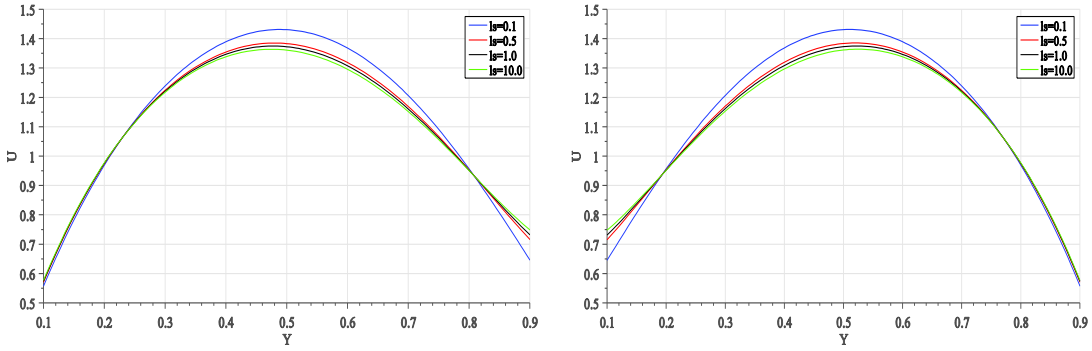


(a) Y=0.05 (b) Y=0.95

Figure 11. Variation of the horizontal velocity (U) with X at two different values of Y for varying slip length, l_s, from 0.1 to 10.

The U velocity with Y at two different values of X is plotted in Fig. 12. Similar variation in U velocity is observed as in the case of out-phase large patterning, but there is marked difference in U velocity close to the boundary as there is no prominent variation in this

component with l_s.



(a) X=-1.7 (b) X=1.7

Figure 12. Variation of the horizontal velocity (U) with Y at two different values of X for varying slip-length l_s, from 0.1 to 10.

To clarify the effect of pressure, the variation of the pressure gradient (PG) in the X direction with X is plotted for $Y = 0.5$ (Fig. 13). Due to the small characteristic length of patterning the pressure gradient profile deviates and fluctuates with a smaller amplitude as compared to the previous case.

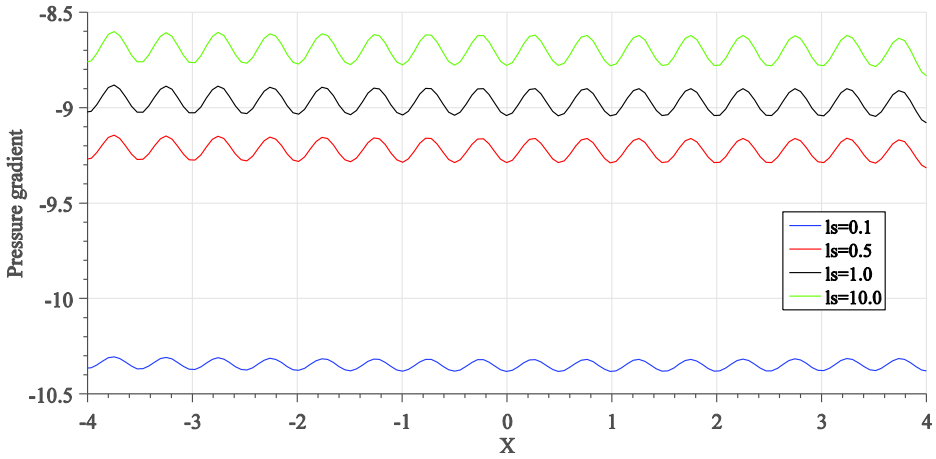


Figure 13. Plot of the pressure gradient (PG) in the X direction with X at $Y = 0.5$ for varying l_s from 0.1 to 10.

The variation of the shear stress on the top wall as shown in Fig. 14 is quite similar to the previous case. Due to the small characteristic length of the patterning, shear stress profiles deviates from the previous case.

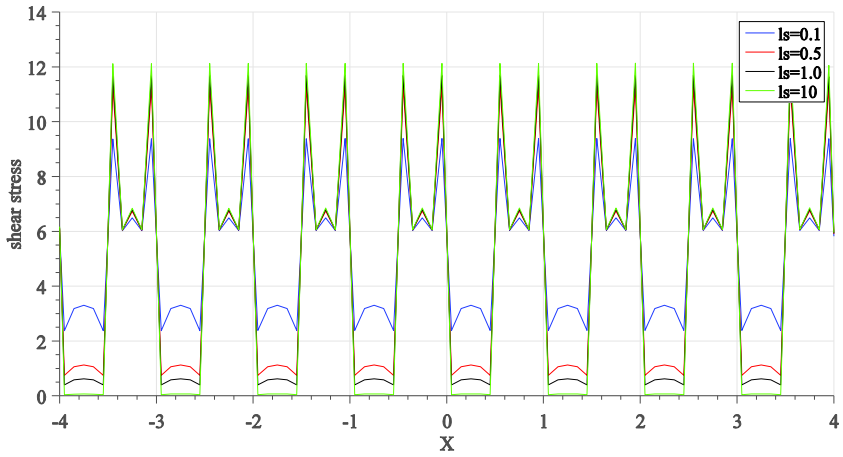


Figure 14. Plot of the shear stress with X at $Y = 1$ for varying l_s from 0.1 to 10.

Conclusions

In this study, flow in out-phase slip-patterned rectangular micro-channels is analyzed using boundary element method (BEM). Two subcases of out-phase slip-patterning namely, large and fine have been discussed. For each of the mentioned cases, the streamline profiles for different values of slip length have been investigated. It was understood that the U and V velocities varies significantly with slip length. It is also observed that the pressure gradient (PG) and wall shear stress have greater dependency over slip-length. When the slip length is very high, it was observed that there is practically no resistance to the flow and therefore shear stress vanishes on the wall.

It is found that the analytical solution is almost not possible for such configuration of boundary conditions. In order to handle such complexities in boundary conditions the presented boundary element solution is very useful.

References

[1] Nishad, C. S., Chandra , A., Raja Sekhar, G.P. (2016) Flows in slip-patterned micro-channels using boundary element methods, *Engineering Analysis with Boundary Elements* **73** 95–102.
[2] Hu, S., Ren, X., Bachman, M., Sims, C. E., Li, G., Allbritton, N. (2002) Surface modification of poly (dimethylsiloxane) microfluidic devices by ultraviolet polymer grafting, *Analytical Chemistry* **74** (16) 4117–4123.
[3] Choi, C.-H., Kim, C.-J. (2006) Large slip of aqueous liquid flow over a nanoengineered superhydrophobic surface, *Physical Review Letters* **96** (6) 066001.
[4] Lauga, E., Stone, H. A. (2003) Effective slip in pressure-driven Stokes flow, *Journal of Fluid Mechanics* **489** 55–77.
[5] Kelmanson, M. (1983) An integral equation method for the solution of singular slow flow problems, *Journal of Computational Physics* **51** (1) 139–158.
[6] Kelmanson, M. (1983) Boundary integral equation solution of viscous flows with free surfaces, *Journal of Engineering Mathematics* **17** (4) 329–343.
[7] Lu, W.-Q., Chang, H.-C. (1988) An extension of the biharmonic boundary integral method to free surface flow in channels, *Journal of Computational Physics* **77** (2) 340–360.
[8] Mazouchi, A., Homsy, G. (2001) Free surface Stokes flow over topography, *Physics of Fluids* (1994-present) **13** (10) 2751–2761.

[9] Mazouchi, A., Gramlich, C., Homsy, G. (2004) Time-dependent free surface Stokes flow with a moving contact line. I. Flow over plane surfaces, *Physics of Fluids* (1994-present) **16** (5) 1647–1659.
[10] Goodwin, R., Homsy, G. (1991) Viscous flow down a slope in the vicinity of a contact line, *Physics of Fluids A: Fluid Dynamics* (1989-1993) **3** (4) 515–528.
[11] Krechetnikov, R., Homsy, G. M. (2006) Surfactant effects in the landau–levich problem, *Journal of Fluid Mechanics* **559** 429–450.
[12] Ingham, D., Yuan, Y. (1997) Multiple solutions of Stokes flow with a free surface, *Engineering Analysis with Boundary Elements* **19** (4) 261–268.
[13] Pozrikidis, C. (1992) Boundary integral and singularity methods for linearized viscous flow, Cambridge University Press.
[14] Cheng, A. (2000) Particular solutions of laplacian, helmholtz-type, and polyharmonic operators involving higher order radial basis functions, *Engineering Analysis with Boundary Elements* **24** (7) 531–538.
[15] Zeb, A., Elliott, L., Ingham, D.B., Lesnic, D. (1998) The boundary element method for the solution of Stokes equations in two-dimensional doamins, *Engineering Analysis with Boundary Elements* **22** (4) 317–326.

Numerical investigations of centrifugal compressor with corrosion pit defect preset on the disk and blade

†S. Y. Chen, Y. P. Qu, X. X. Tang, D. D. Wan and J. Y. Pan

Key Laboratory of High-efficiency and Clean Mechanical Manufacture, School of Mechanical Engineering, Shandong University, Jinan, Shandong, 250062, P. R. China
†Corresponding author: chensy66@sdu.edu.cn

Abstract

Three dimensional numerical investigations of a centrifugal compressor with pitting corrosion defect preset on the disk and blade are presented by FLUENT software. The pitting defect parameters such as the velocity field and pressure distribution in the gas passage are analyzed. With that purpose, the RNG k-ε two−equation turbulence model and simple algorithm are used based on the Navier-Stokes equation. Meanwhile, the DPM model is used to simulate the second phase particles deposit around the pitting corrosion defect. The chosen model provides a reference for the compressor impeller material stress corrosion test. The simulation results showed that the velocity gradient and static pressure decreased sharply while dynamic pressure increased and the chloride concentration was 50 times higher than before in the pitting area. The results could be as guidance for compressor serving in nature gas containing chloride ion to estimate the importance of ion concentration in the stress corrosion originated pitting defect.

Key words: centrifugal compressor, impeller, numerical simulation, pitting corrosion defect

1 Introduction

With respect to a compressor, the service life mainly depends on integrity and safety of the impellers. It is very difficult to do theoretical and experimental studies in fluid flow inside the impeller. Lots of researchers investigate the flow peculiarity inner the compressor without minor damage in the gas passage. The velocity and pressure changes due to minor defect such as pitting and micro cracks in the impeller will influence the generation and extension of cracks which affecting the performance and operation security of a centrifugal compressor. The crack propagation on the blade surface will lead to efficiency droop and abnormal vibrations. Numerical investigation on the compressor performance with defect on the blade is of alternative important to determine the velocity and pressure redistribution [1].

The turbulence model method is widely used for solving viscous flow in recent years. The k-ε two-equation model has reached success in engineering application for its own clear physical background [2]. The standard k-ε model is widely used in the centrifugal compressor impeller. However, the standard k-ε model has great restriction for that in the analysis of internal flow impeller machinery ignoring the anisotropy of turbulence. For that reasons, the standard k-ε model is only suitable for predicting simple isotropic turbulence [3-4]. In 1986, Yakhot and Orzag [5] came up with the dynamic renormalization group (RNG) k-ε model for the first time and the model allows us to evaluate transport coefficients and transport equations for the large-scale transport modes. The k equation and ε equation are as follows:

$$\frac{\partial(\rho k)}{\partial t}+\frac{\partial(\rho k u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\alpha_k \mu_{eff}\frac{\partial k}{\partial x_j}\right]+G_k+\rho\varepsilon \tag{1}$$

$$\frac{\partial(\rho \varepsilon)}{\partial t}+\frac{\partial(\rho \varepsilon u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\alpha_\varepsilon \mu_{eff}\frac{\partial k}{\partial x_j}\right]+\frac{c_{1\varepsilon}^{*}\varepsilon}{k}-G_{2\varepsilon}\rho\varepsilon\frac{\varepsilon^2}{k} \tag{2}$$

Where α_k, α_z are the Prandtl numbers. G_k is turbulence kinetic energy generated by laminar velocity gradient, k is turbulence kinetic energy, μ_{eff} is eddy viscosity, ε is dissipation term.

Compared with the standard k–ε equation, the RNG k-ε equation takes the average flow in rotational flow. Based on the k–ε equation, it adds an item which reflects the strain rate of the mainstream. The RNG k-ε model is not only related with the flow movement, but also a function in spatial coordinates. So, it is easier to manipulate the high strain rate and large gradient of streamlining curvature.

Shengbo[6] discovered that different k-ε models calculation results varied widely in the simulation of the centrifugal compressor impeller. RNG k-ε model can be used to study the separation movement of the blade suction surface while standard k-ε and Realizable k-ε cannot. The RNG k-ε model can be used to not only simulates a wide range of separation vortex, but also distinguish the nuances of small-scale vortex structure. For this reason, the researchers give priority to use the RNG k-ε model in the flow simulation of centrifugal compressor impeller [7-8]. Before 1980s, inviscid flow simulation is adopted because of the limitation of hardware. Zhonghua Wu proposed the theory of S1/ S2 streaming surface, and Li [9] successfully carried out a simulation of impeller inner flow field to display the motion in average S2 stream surface. In the stage of inviscid simulation, quasi-orthogonal calculation also played an important role in numerical simulation [10]. In1980-1990, the researchers took viscosity of the fluid, backflow and eddy influence into consideration. In the period, many computational methods were adopted. Epureanu [11] simulated the impeller inner fluid with inviscid-viscous turbulence model, potential function calculated inviscid part while integral for viscous part. Gang Zhu carried the impellor inner fluid simulation with vorticity-steam function method in viscous fluid condition. Borello [12] put up with the finite element overlap base on the development of computer with high configuration and parallel computing technique.

The paper focuses on the compressor impeller in natural gas transportation project to explore the influences of chloride abundant that could lead stress corrosion. It can be the guidance of preventative maintenance and security operating.

2 Numerical analysis method

The compressor works at a constant angular speed. The control equations for the system are Navier-Stokes equation, continuity equation and energy equation in a relatively cylindrical coordinate system. The calculation solver is Pressure-Based type coupled implicit solver. The turbulence model is RNG k-ε model and the coupled of velocity and pressure is SIMPLE. Computed variables adopted the second order upwind format and residual error is 10^{-3} convergence. The second phase particle Cl^- will be added from the inlet of the impeller with 2

percent of the compressor mass-flow after the calculation completed. The control parameters are that the time step is 1 *ms*, calculation steps are 5000 with unsteady particle tracking. The particle injects into continuous phase each iteration step 30. The simulation is based on steady-state.

Navier-Stokes equation:

$$\frac{Du_i}{Dt} = f_i - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu\nabla^2 u_i + \frac{1}{3}\nu\nabla(\nabla u_i) \tag{3}$$

Continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho u_i) = 0 \tag{4}$$

Gas state equation:

$$p = \rho RT \tag{5}$$

Energy equation:

$$\frac{\partial(\rho T)}{\partial t} + div(\rho u_i t) = div(\frac{k}{c_p} grad T) + S_T \tag{6}$$

Particle trajectory control equation:

$$\frac{du_i^p}{dt} = F_D + F_g + F_B + F_L + F_T \tag{7}$$

Particle stokes law FD expression for resistance equation:

$$F_D = \frac{18\mu}{d_p^2 \rho_p c_c}(\mu_i - \mu_i^p) \tag{8}$$

Thermophoresis force FT expression equation:

$$F_T = -D_T \frac{1}{m_p T}\frac{\partial T}{\partial x} \tag{9}$$

Where μ is the kinematic viscosity, ρ is the density, t is time, u_i is the velocity, R is molar gas constant, Cp is specific heat capacity, T is temperature and k is heat transfer coefficient and the S_T is the internal heat source. Cc is slip coefficient, Fg is gravity, F_B is brown power, Δt is time step, S_0 is spectral intensity function, F_L is Saffman lift force and D_T is the swimming thermal coefficient.

3 Analysis model and the grid

3.1 *Boundary conditions*

The numerical simulation model is the stage 2 impeller containing 11 long blades and 11 short blades. The circumferential velocity of compressor is 296.9 *m/s* and the inlet pressure is 0.17 *MPa*. The outlet diameter is 918 *mm*. The inlet temperature is 350 *K*. The mass-flow of the impeller is 26 *Kg/s* with compression ratio 4:1. The impeller geometric 3-D model and flow channels 3-D model are shown in Figure 1.

In the outlet, pressure-outlet is taken as boundary condition..

Figure 1. Impeller geometric model and flow channels model.

3.2 Grid

As shown in Table 1, the static pressure and total pressure extracted on the pitch-averaged surface would keep steady with the increase of mesh number. The Tet/Hybrid mixed grid type is used to mesh flow channel with 2642992. The mesh is shown in Figure 2.

The pitting is taken as the first step of the stress corrosion where the micro-cracks appear. The flow field pressure changes and distribution of corrosion ion plays a vital role in crack initiation and propagation. In order to examine the influential factors of stress corrosion in crack initiation and propagation, two corrosion pit defects of 2 *mm* in diameter on the disk and pressure surface are preset for simulating the flow field of impeller, respectively.

Table 1. Test for grid independence

Mesh number	Static pressure /MPa	Total pressure /MPa
1506543	0.1366	0.1898
2642992	0.1366	0.1798
3207642	0.1366	0.1775

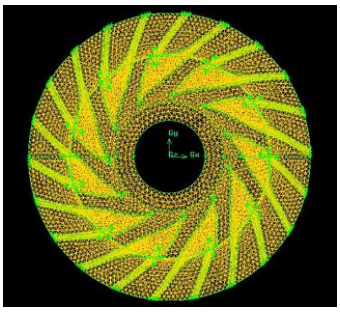


Figure 2. The sketch mesh of flow channel model.

The grid of pitting corrosion defect on disk and blade are refined as 63 and 106, respectively, as shown in Figure 4.

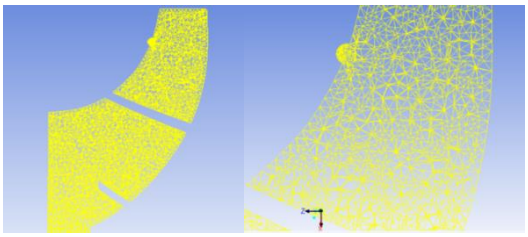


Figure 3. The grid of pitting on disk

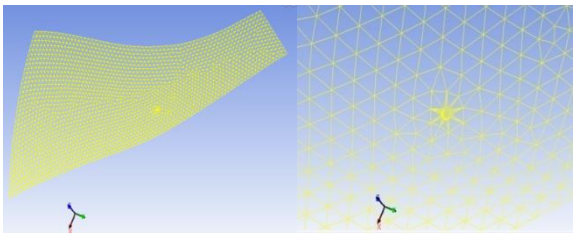
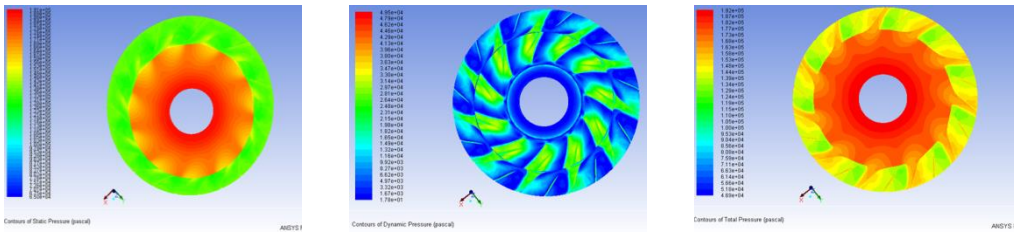


Figure 4. The grid of pitting on blade

4 Simulation results and analysis

4.1 *Pressure and velocity distribution of impeller*

According to the distribution of static pressure and total pressure of the impeller showed in Figure 5. In the inlet of the blade, there is significant low pressure area and pressure gradient at the head of the blade. Velocity distribution is seen in Figure 6.



(a) static pressure (b) dynamic pressure （c）total pressure

Figure 5. The pressure distribution of the impeller

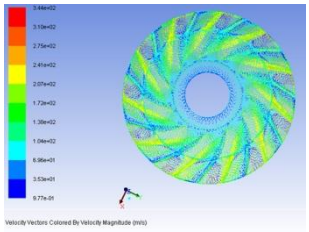
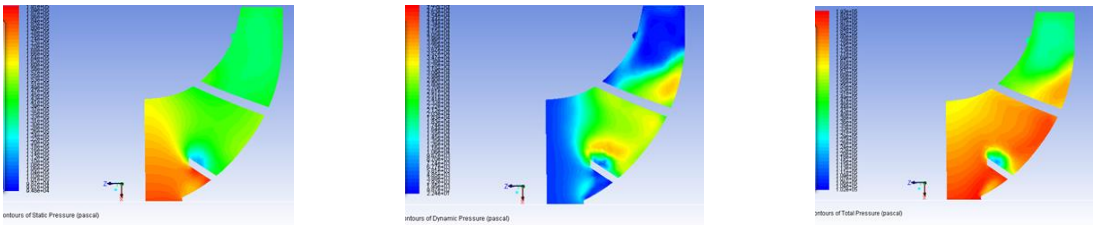


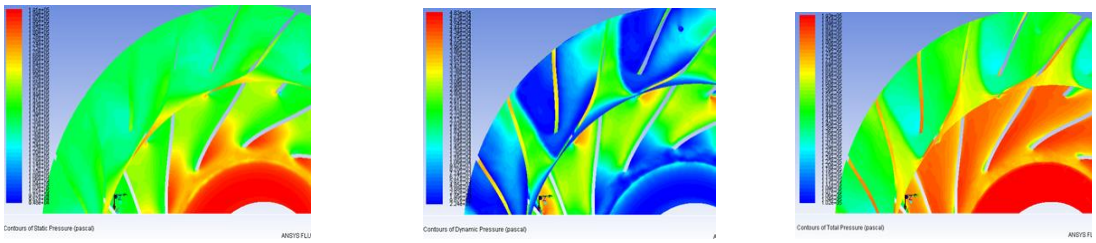
Figure 6. The velocity vector of the impeller

4.2 *Pressure, velocity and corrosion ion distribution at the corrosion pit defect of the disk*

As shown in Figure 7 and 8, static pressure increases and dynamic pressure reduces quickly in the incident surface of pitting defect. The velocity and velocity gradient decrease sharply and backflow vortex appears in pitting in which the static pressure decreased while dynamic pressure increased. There are a large number of backflow eddies especially in the passage of impeller according to Figure 9 and 10.

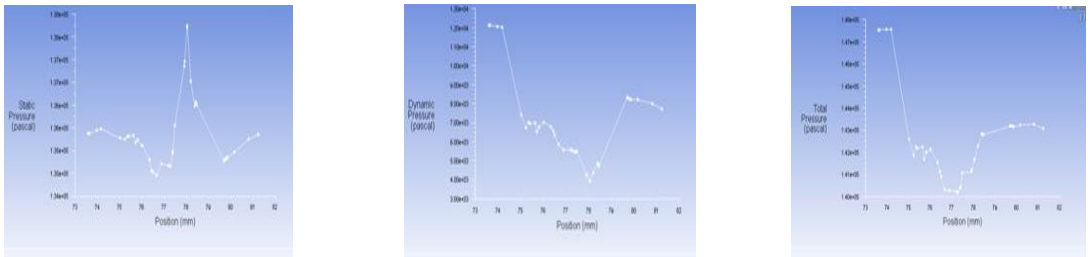


(a) static pressure in passage (b) dynamic pressure in passage (c) total pressure in passage



(d) static pressure on disk (e) dynamic pressure on disk (f) total pressure on disk

Figure 7. The pressure distribution of the passage and disk



（a）static pressure (b) dynamic pressure (c) total pressure

Figure 8. The pressure distribution on the corrosion pit defect of the disk

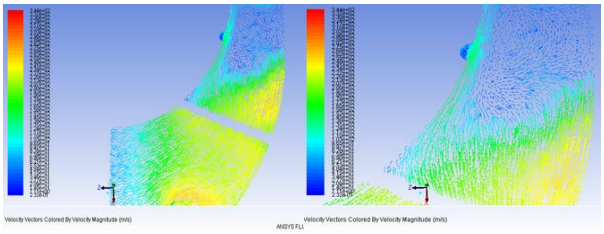


Figure 9. The velocity vector on the pit defect of the disk

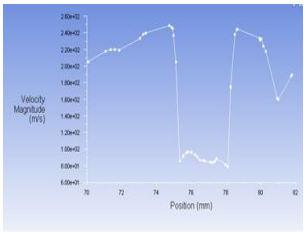


Figure 10. The velocity distribution on the of the disk

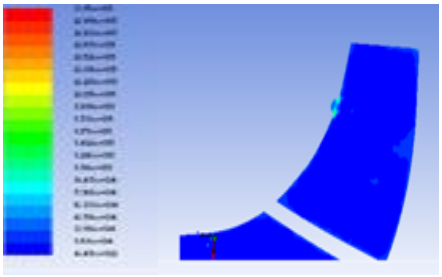


Figure 11. The strain rate on the pit of the disk

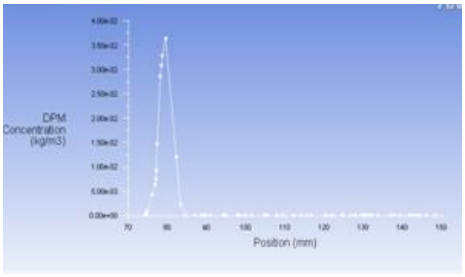


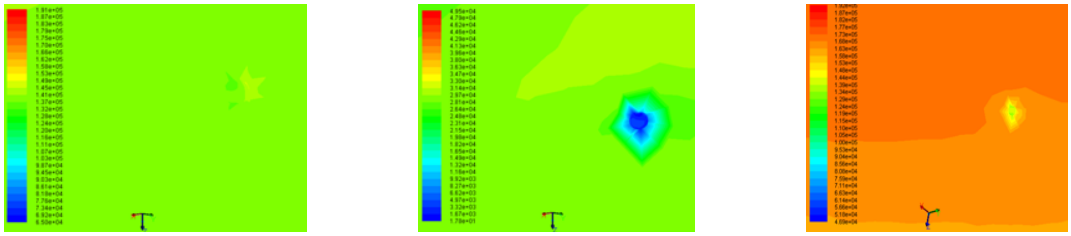
Figure 12. The sedimentary concentration of Cl- at pit of the disk

The flow strain rate at the pitting defect of the disk changes considerably as shown in Figure 11. The sedimentary concentration at the pitting defect of the disk is shown in Figure 12. It is found that the Cl- concentration is 50 times higher than before in the pitting area. The pitting defect area should be easy to be the origin of the crack because the Cl- is reactive ion to the corrosion of high strength stainless steel. The deduction can be further proved by the fact that the results of simulation coinciding with the stress corrosion test.

4.3 Pressure, velocity and corrosion ion distribution at the corrosion pit defect of the blade

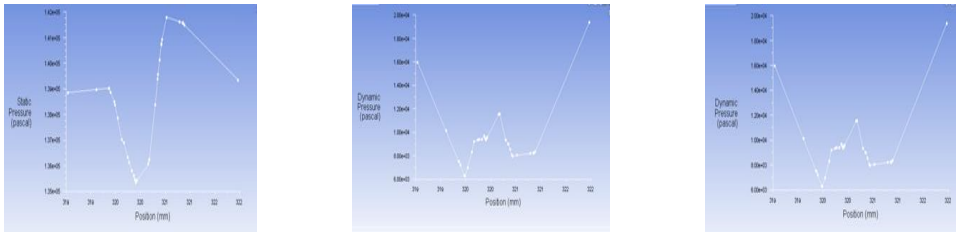
The static pressure increases in the incident surface of pitting defects and dynamic pressure reduces quickly as shown in Figure 13, 14, 15 and Figure 16. The static pressure reduces and

dynamic pressure increase in pitting defect. Fluid strain rate on the pitting defect of the blade changes considerably as shown in Figure 17, 18 and19.



(a) static pressure (b) dynamic pressure (c) total pressure

Figure 13. The pressure at corrosion pit of the blade



(a) static pressure (b) dynamic pressure (c) total pressure

Figure 14. The pressure distribution curve at pit of the blade

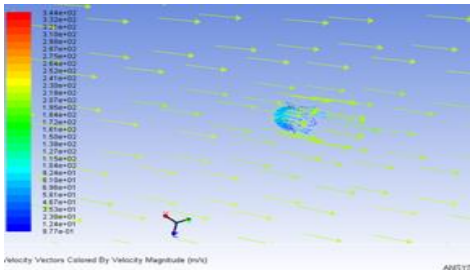


Figure 15. The velocity vector at the blade

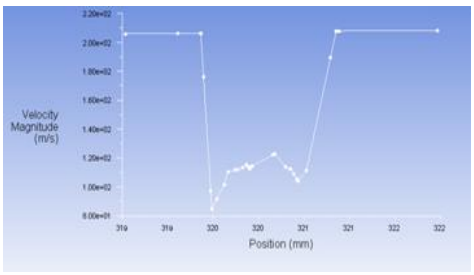


Figure 16. The velocity curve at the blade pit

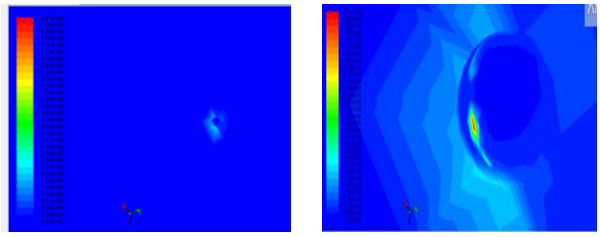


Figure 17. The Fluid strain rate in corrosion pit of the blade

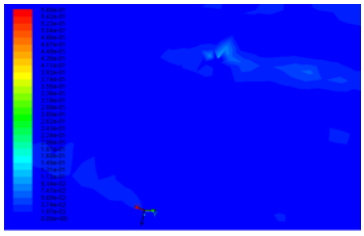


Figure 18. The sedimentary concentration of Cl^- in the corrosion pit of the blade

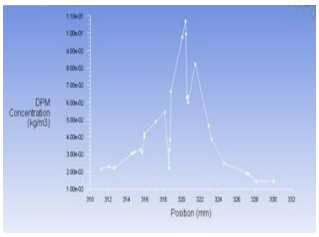


Figure 19. The sedimentary concentration of Cl^- at the corrosion pit of the blade

5. Conclusion

Based on the computational simulation of blades with pitting corrosion as well as the disk, the change of flow field parameters are made out as follows：

(1) The static pressure increased and dynamic pressure decreased in the upstream face of pitting and the fluid velocity decreased sharply. Vortex backflows appear at the pitting defect.

(2) The flow becomes complicated near the pitting and the blade would be easy to crack with fluid shear rate varying widely and the stress distribution deteriorates under the influence of fluid in the pitting area.

(3) The Cl- concentration is 50 times higher than before in the pitting area where is easily to be the origin of the crack and the phenomenon coincides with stress corrosion test results.

(4) The expanded crack would deteriorate the flow field. Due to such alternates, the appearance of micro-crack would be the safety hazard to the blades. In the actual condition piping natural gas containing Cl^-, the results could be used to estimate the expansion of crack originated from the pitting defect that should design with higher safety coefficient.

Acknowledgement

The work was supported by Shandong Provincial Science and Technology Development Program (2014GGX108001 and 2016GGX104018).

References

[1] Hou, S. Q., Wang, C. X., Lin, J. Z. (2005) Summary of Numerical Simulation of the Internal Flow in the Turbomachinery. *FLUID MACHINERY*, **5**, 33.
[2] Chen, X., Miao, Y. M. (1992) The computation of internal flow fields in centrifugal compressor impellers. *Journal of Thermos-physics Engineering*, **3**, 13
[3] Hu, S. B., Su, M. M. (2009) Numerical Study on Loading Rules of Centrifugal Compressor Impeller. *Compressor Blower & Fan Technology*, **4**, 25-27.
[4] Liu, X. Z., Cao, S. Z., Gu, C. G. (2000) Computational and Experimental Investigation of 3-D Turbulent Flow Field in Centrifugal Impeller. *FLUID MACHNERY*, **28**(4):9-12.
[5] Yakhot V., Orzag S. A.（1986） Renormalization group analysis of turbulence：basic theory. *Scient Comput*.**1,** 3-11.
[6] Hu, S. B., Su, M. M.. (2010) Comparison of Different Turbulence Model for Numerical Simulation of Internal Flow of Centrifugal Compressor. *FLUID MACHNERY*, **8**, 38
[7] Balje O. E. Turbo-machine A : Guide to Design, Selection and Theory. America: Wiley Inter-science Publication, 1981.
[8] Krain H. (1999) High Pressure Ratio Centrifugal Compressor with Transonic Flow. *Proceeding of the 3rd ASME/ JSME Joint Fluids Engineering Conference*. July 18 -23.
[9] Zhang, L., Chen, H. P., Xu, Z. (2000) Quasi-three Dimensional Iterative Numerical Analysis for Flow Field in Compressor Impellor. *Fan Technology*, **5**, 3-7.
[10] Chen, Q., Zhang, B. Z., Cui, C. L.（2010） Quasi Orthogonal Surface Calculation for Three dimensional Flow in Centrifugal Fan. *Journal of Hangzhou Institute of Application Engineering*, **12**, 21-23.
[11] Epureanu B. I., Hall K. C., Dowell E. H. (2001) Reduced Order Models of Unsteady Viscous Flows in Turbomachinery Using Viscous Inviscid Coupling. *Journal of Fluids and Structures*, **15**, 255-273.
[12]Borello D., Corsini A., Rispoli F. (2003) A finite element overlapping scheme for turbomachinery flows on parallel platforms. *Computers & Fluids*, **32**, 1017-1047.

Building Language Models with Fuzzy Weights

***Tsan-Jung He[1], †Shie-Jue Lee[2], and Chih-Hung Wu[3]**

[1]Department of Electrical Engineering National Sun Yat-Sen University, Kaohsiung, Taiwan.
[2]Department of Electrical Engineering National Sun Yat-Sen University, Kaohsiung, Taiwan
[3]Department of Electrical Engineering National University of Kaohsiung, Kaohsiung, Taiwan

*Presenting author: zlhe@water.ee.nsysu.edu.tw
†Corresponding author: leesj@mail.ee.nsysu.edu.tw

Abstract

Word2Vec is a recently developed tool for building neural network language models. The purpose of this work is to propose an improvement to Word2Vec by adding fuzzy weights related to the distances in the context to use more information than the way adopted in the original linear bag-of-word structure. In Word2Vec, the same weights are given regardless of different distances between words. We consider that word distances in the context bear certain semantic sense which can be exploited to reinforce connections more effectively for the network model. In order to formalize the influence of different distances in the context, we adopt Gaussian functions to represent fuzzy weights which take part in the training of the connections of network. Various experiments show that our proposed improvement can result in better language models than Word2Vec.

Keywords: word2vec, skip-gram, fuzzy weights, neural network language model, word embedding, natural language processing (NLP).

I. Introduction

In order to deal with many tasks like text processing and near-duplicate detecting easily, an appropriate method to select features is extremely important. In these tasks, one of the most employed method is Term Frequency-Inverse Document Frequency (TF-IDF) [1]-[3] which represents documents as a vector by using word counts in a corpus. However, it would be less efficient when a corpus becomes larger. Worst of all, different words in a corpus are considered as independent objects. The limit of this type of representation is that similar words or synonyms are not taken into account. In the real world, people may use different words to represent the same or similar concepts. When two similar documents contain some synonyms, calculating the similarity between these documents would encounter difficulties. Therefore, if we want to do the calculation in a semantic way, word representation is certainly an issue.

The simplest way of word representation is the one-hot representation [4][5] which is a $1\times N$ vector to map different words in a document. The vector is as large in size as the vocabulary which is the collection of all different words in a corpus of documents. All the word vectors form a matrix. In the matrix, an entry is 1 if the word appears in the underlying document, and is 0 otherwise. However, this type of representation has the problem of the curse of dimensionality, and it is also difficult to express the relationship among different words. Recent researches based on Mikolov et al. [6][7] proposed the Word2Vec model, employed in Neural-network Language Models [8], which transfers the one-hot representation into small dense vectors. Words are expressed as low-dimensional distributed vectors. Word2Vec is also an efficient method to capture complex semantics and word relationships, and can be used to find synonymous words by considering positions in the vector space. Therefore, Word2Vec

has been widely applied in many Natural Language Processing (NLP) tasks such as Part-Of-Speech (POS) Tagging and text classification [9]-[11].

To train a network language model, Word2Vec adopts a linear bag-of-words training. During the training phase, documents are divided into many word sets, usually sentences, and then words in the context windows are selected as targets and surroundings to reinforce the connections. If two words occur simultaneously more frequently, the relationship between these words is more increasingly enhanced. However, all the involved words are treated of equal importance and the information of their positions is not properly exploited in the training phase. In this paper, we overcome this disadvantage by integrating the weights induced from the positions of contextual words in the context window into the Word2Vec training architecture. Intuitively, words nearer to the target in the context window should have a closer relationship to the target and should be more important to the target. So we take advantage of the distance information to improve the semantic accuracy of the resulting language model. We express the importance of words in terms of fuzzy weights. To formalize the influence of different distances in the context, we adopt Gaussian functions to represent fuzzy weights which take part in the training of the connections of network. Various experiments show that our proposed improvement can result in better language models than Word2Vec.

The rest of this paper is organized as follows. Section II briefly introduce the Word2Vec neural network architecture. Section III describes our proposed method to integrate fuzzy weights into the Word2Vec algorithm. In Section IV, we present some experimental results. Finally, a conclusion is given in Section V.

II. Related Work

Word2Vec is a tool released by Google in 2013 [6][7]. It provides a simple and high-quality method to train the representation of words and is used in many natural language processing (NLP) applications. In Word2Vec, it has two main training architectures: continuous bag-of-word (CBOW) and skip-gram (SG). The difference between the two is that CBOW predicts the current word based on the context while SG predicts the surrounding words by the current word.

Training data can rely on the unlabeled text data as the input [12], and the model produces the vectors of the words as the output. The vocabulary is first constructed from the input and the vector representation of words is learned. SG and CBOW work the same in the early steps of the training phase, picking a word as a center and selecting contextual words according to the window size. Then CBOW takes contexts as input and SG takes a center word to infer contexts, as presented in Fig 1. After training, each word is represented as a vector in the vector space. These word vectors can then be used as features in the succeeding classification applications.

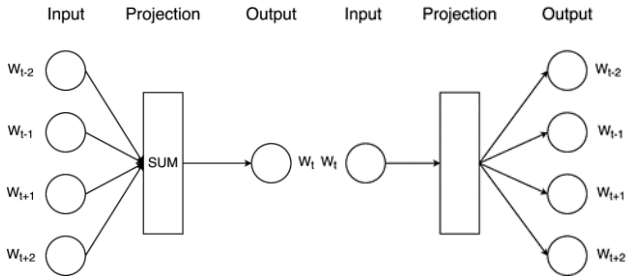


Figure 1: Word2Vec training architectures.

In this work, we focus on the skip-gram architecture. The SG proposed in the original Word2Vec is re-drawn in Fig 2. The input is a token of word w_t, which is mapped to its word vector $v(w)$, and is then used to predict the word vectors of its neighboring words. Given a word sequence w_1, w_2, ..., w_N, the training process maximizes the following objective function.

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{-C\le j\le C, j\neq 0}\log P(w_{i+j} \mid w_i) \tag{1}$$

where is the softmax function defined as

$$P(w_O \mid w_I) = \frac{\exp({v'_{w_O}}^T v_{w_I})}{\sum_{w=1}^{W}\exp({v'_w}^T v_{w_i})}. \tag{2}$$

Note that v_w and v'_w are the input and output vector representations, respectively, of w, and W is the number of words in the vocabulary. As shown in Fig 2, SG models do not take into account the word order to predict the output. Therefore, these methods allow us to find out close vectorial representation of a given word or context, but are not optimal in predicting word sequences based on grammatical and syntactic properties of words.

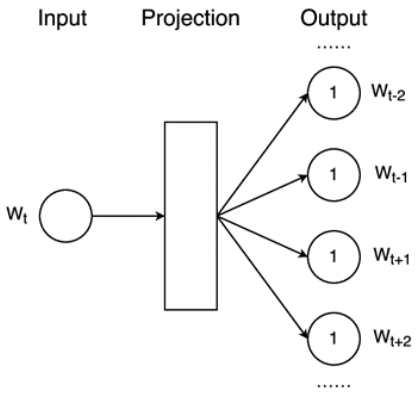


Figure 2: Original Skip-gram model.

III. Proposed Method

In the bag-of-words structure, distances from target words in the context window are not taken into account. In our proposed method, we try to make use of the distances in terms of fuzzy weights in the training phase. Our method can basically be divided into two main processes. The first process is to define adaptive fuzzy weights which denote sensibly the importance of different words according to their distances to the target. The second process is to integrate the fuzzy weights into the SG model of Word2Vec.

Two algorithms are proposed and implemented in our method, skip-gram and negative sampling. Fuzzy weights are integrated in both algorithms. Negative sampling selects a

positive sample (target word w_t) and some negative samples, trying to make these negative samples differ from the target word in the training phase. Although negative sampling may lower the accuracy a bit, but it improves the efficiency in model training.

Word2Vec works as follows. Before beginning to train a network model, as shown in Fig. 3, a vocabulary is built depending on the corpus under consideration. Then random vectors are generated to initialize all the word vectors in the vocabulary, named $syn0$, and set zero to the other synonyms, named $syn1$, which connect word vectors to the output.

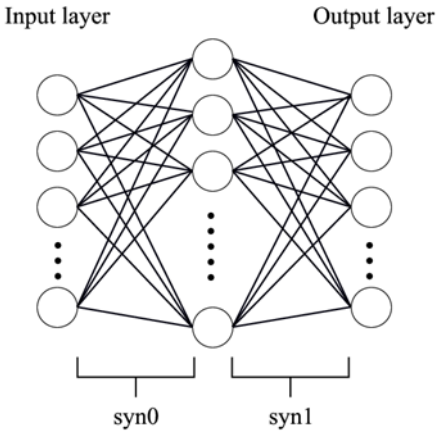


Figure 3: Connections in the network model.

To train the model, the first step is to select the target word, or positive word, *pword*, and contextual words, or last words, *lword*, according to the window size. The skip-gram holds a target word to train a last word in the window one at a time. The model adopts sigmoid function as activate function and expects the output of similar words to be 1. In negative sampling mode, the label of the target word would be set as 1 and the other samples would be 0. The accumulated error err_h is calculated as

$$output = \frac{1}{1+e^{-syn0^{(l)} \cdot syn1^{(p)}_{(l)}}} \tag{3}$$

$$err_{output} = (label - output) \times \alpha \tag{4}$$

$$err_h = err_h + err_{output} \cdot syn1^{(p)}_{(l)} \tag{5}$$

where α is a learning rate, $syn0^{(l)}$ is the vector of the *lword* in the context window and $syn1^{(p)}_{(l)}$ is the connections from vector *lword* to vector *pword*, and $syn1$ is updated by

$$syn1^{(p)}_{(l)} = syn1^{(p)}_{(l)} + err_{output} \cdot syn0^{(l)} \tag{6}$$

After training the positive sample and negative samples, the accumulated errors err_h is for back-propagation to fix $syn0$ by

$$syn0^{(l)} = syn0^{(l)} + err_h. \tag{7}$$

Our proposed method allows fuzzy weights into play with the updating of $syn1$ and $syn0$.

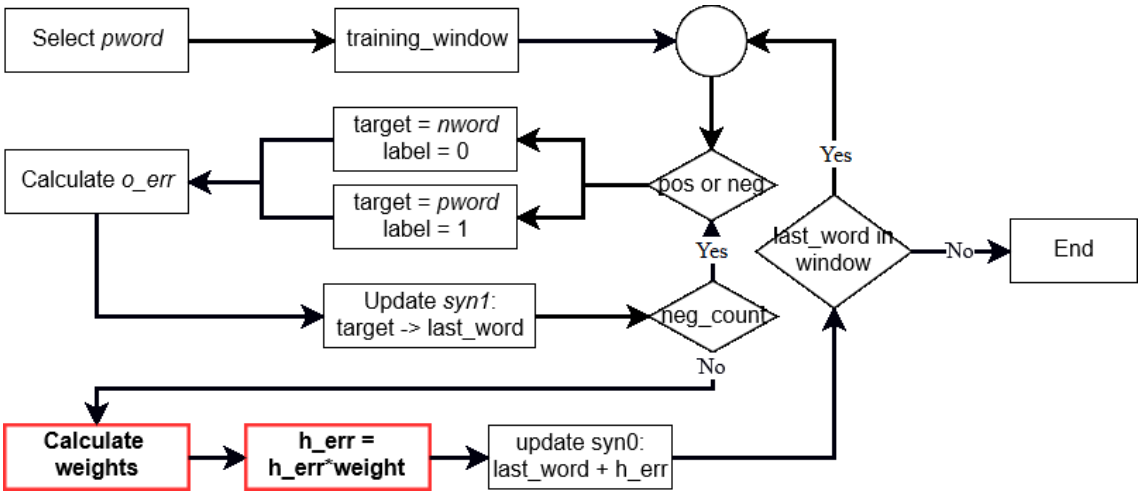


Figure 4: Improved skip-gram algorithm.

As shown in Fig. 4, we calculate the fuzzy weights which are transformed from distances by applying Gaussian functions. For the first process, to take advantage of the relationship of distances between the contextual words and the target word, we replace the original binary weighting of output vectors with Gaussian function values. Words in context are weighted with $G(w_i)$ for word w_i:

$$G(w_i) = \frac{-d(w_i)^2}{2 \cdot \sigma^2} \tag{8}$$

where $d(w_i)$ is the distance between a training word and the target word, i is the order of words in context, and the σ is the radius of Gaussian function. Fig 5 shows the value of each position as a function of the distance from the center of the context window. This function is then used to evaluate the distance relationship of the contextual words. If the position of a contextual word is closer to the target word in the center, the target word would be more influential in the training phase. Consequently, we adopt the following modifications:

$$new\_err_h = G(w_i) * err_h \tag{9}$$

$$syn0^{(l)} = syn0^{(l)} + new\_err_h. \tag{10}$$

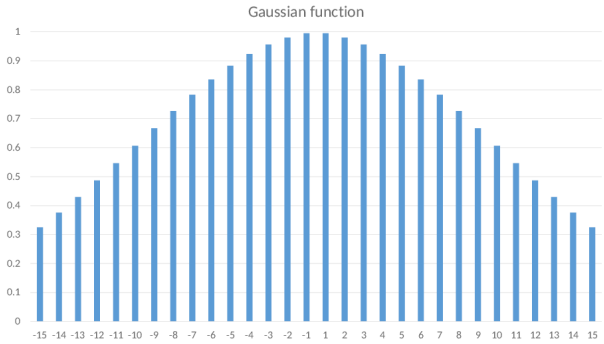


Figure 5: A Gaussian function with $\sigma^2 = 100$.

Fig 6 presents the skip-gram model which incorporates the context weighting approach. The skip-gram model refers to a word to predict its surrounding words. The training process selects a center target word and one of contextual words to form a pair, and then does the training pair by pair. If the weights were applied on word vector directly, the new vector would be regarded as a different word, so we propose to apply the weights in the phase of updating errors. Updating the accumulated error affects the most important learning steps of skip-gram, so we adjust the weights to make similar words closer easily.

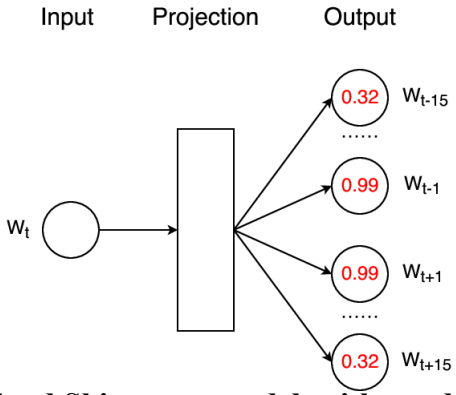


Figure 6: Fuzzy Weighted Skip-gram model, with and context window size .

Example
Assume in an article we have the following sentence:

Lions and jaguars hunt zebras, giraffes and gazelles in the savanna.

Let the context window size be 5, $\sigma^2 = 100$, and the target word be *giraffes*. Table 1 presents the involved fuzzy values. In this table, the first row lists the contextual words, the second row shows the relationship according to the position of the target word, and fuzzy weights are listed in the third row.

Table 1. Fuzzy weights of words

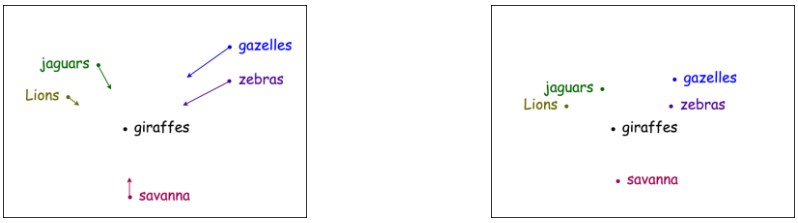
contextual	**Lions**	**and**	**jaguars**	**hunt**	**zebras**
position	-5	-4	-3	-2	-1
weights	0.29	0.45	0.64	0.82	0.95
contextual	**and**	**gazelles**	**in**	**the**	**savanna**
position	+1	+2	+3	+4	+5
weights	0.95	0.82	0.64	0.45	0.29

The following are training pairs:

(Lions, giraffes), (and, giraffes), (jaguars, giraffes), (hunt, giraffes), (zebras, giraffes), (and, giraffes), (gazelles, giraffes), (in, giraffes), (the, giraffes), (savanna, giraffes).

When a weight is greater, the error feedback caused is more effective. In this example, the vectors of 'zebras' and 'gazelles' may have more chance to get closer to 'giraffes' than to 'Lions' and 'jaguars'.

Fig. 7(a) shows the distribution of the words in the vector space. In this figure, the length of arrows stand for the strength of training weights. Note that the strengths are different due to their positions. After updating the vectors, 'zebras' and 'gazelles' take a bigger step to approach 'giraffes', but 'Lions', 'jaguars', and 'savanna' are less effective, as shown in Fig. 7(b).



(a) Before. (b) After

Figure 7: Words in the vector space: (a) Before updating; (b) after updating.

IV. Experimental Results

In this section, we evaluate the effectiveness of our proposed method. We also compare our method with the original Word2Vec method. We conduct our experiments with the Google-One-Billion-Word-Language-Modeling Benchmark containing around 1-billion words for training. After training, there are 553,402 words in the vocabulary. Testing words are also taken from Google, a file named question-word, containing 14 topics: capital-common-countries, capital-word, currency, city-in-state, family, and 9 grammars. The Semantic-syntactic Word relationship test is made of 19,558 questions. Its main objective is to verify if a distributed representation of words captures complex syntactic and semantic relations between words. A question is made of two pairs of words sharing the same relation. For example, for the following question

'woman' 'man' = 'queen' 'king',

The model has to infer from 'queen' to 'king' through the relationship between 'woman' and 'man'. A question in the capital-common-countries topic is something like

Athens Greece Baghdad Iraq.

All the test questions contain the rules of inference that people do in daily life. In other words, if a model gets a large score in this test, the model is accurate in the semantic sense.
We perform experiments with different settings of the parameters for the window size, dimension of vector space, and radius of Gaussian functions. The results with different window sizes are shown in Table 2.

Table 2. Accuracy results with different window sizes

dimension = 100	**w5**	**w10**	**w20**	**nMAX**
capital-common-countries:	363	369	364	462
capital-world:	941	984	976	1295
currency:	7	5	7	40
city-in-state:	418	466	476	1700
family:	260	249	218	342
gram1-adjective-to-adverb:	191	215	235	930
gram2-opposite:	109	112	104	462
gram3-comparative:	1060	986	915	1332
gram4-superlative:	546	515	416	756
gram5-present-participle:	677	668	617	870
gram6-nationality-adjective:	1073	1088	1104	1160
gram7-past-tense:	876	868	844	1560
gram8-plural:	696	689	722	992
gram9-plural-verbs:	411	375	327	650
Total	**7628**	7589	7325	12551

Note that dimension = 100 for this table. Three window sizes are set, namely window size = 5, 10, and 15, represented by w5, w10, and w15, respectively. From this table, we can see that a larger window may not improve the semantic accuracy, since the words far away the target may act as noise for training. In many situations, further words may be less related to the central word syntactically, but for some sentences in which an object mentioned far away may have more influence for a setting with long context window. In Table 2, the topics like 'capital-world' and 'capital-in-state' contain more terms when the window size increases. As a result, accuracy increases as the window size increases. The experimental results with different dimensions are shown in Table 3.

Table 3. Accuracy results with different dimension sizes

window size = 5	**d100**	**d200**	**d400**	**nMAX**
capital-common-countries:	371	392	420	462
capital-world:	932	1044	1103	1295
currency:	5	8	10	40
city-in-state:	400	724	916	1700
family:	260	277	274	342
gram1-adjective-to-adverb:	191	183	167	930
gram2-opposite:	115	165	189	462
gram3-comparative:	1069	1171	1191	1332
gram4-superlative:	573	606	578	756
gram5-present-participle:	678	711	690	870
gram6-nationality-adjective:	1070	1090	1112	1160
gram7-past-tense:	871	920	948	1560
gram8-plural:	698	744	784	992
gram9-plural-verbs:	416	491	460	650
Total	7649	8526	**8842**	12551

For this table, window size is set to be 5, and three dimensions are set, namely dimension = 100, 200, and 400, represented by d100, d200, and d400, respectively. From this table, we can see that the bigger the dimension of the vector space is, the higher accuracy is obtained. As the dimension increases, more information can be stored in the hidden layer and the semantic relationship among words can be more completely maintained. The experimental results with different radii of Gaussian functions are shown in Table 4.

Table 4. Accuracy results with different radii

	Word2Vec	**r100**	**r125**	**r150**	**r175**	**r200**	**r225**
capital-...:	392	**412**	408	406	403	404	403
capital-...:	1078	1081	**1089**	1087	**1089**	1085	1087
currency:	8	9	9	9	**10**	9	9
city-in-...:	920	**1001**	968	972	971	957	962
family:	246	**269**	267	264	260	257	259
gram1-...:	216	227	229	222	221	221	**237**
gram2-...:	**186**	171	178	178	174	174	172
gram3-...:	1120	1121	1108	1123	1112	1125	**1129**
gram4-...:	518	533	540	543	546	**547**	542
gram5-...:	691	698	690	696	698	693	**703**
gram6-...:	**1129**	1128	1128	1126	1128	1126	1128
gram7-...:	938	966	970	**974**	965	969	955
gram8-...:	784	786	794	795	**800**	791	787
gram9-…:	410	418	418	415	418	**427**	423
Total	8636	**8820**	8796	8810	8795	8785	8795

For this table, six values are set to the radius, namely σ^2 = 100, 125, 150, 175, 200, and 225, denoted by r100, r125, r150, r175, r200, and r225, respectively. As can be seen, all the versions of our proposed method perform better than the original Word2Vec in terms of semantic accuracy. Our method can improve the semantic accuracy. However, more execution time is needed. The computer system we have used for the experiments is shown in Table 5. Table 6 shows the execution times required for Word2Vec and our method.

Table 5. Computing environment

Experimenting equipment	
CPU	Intel(R) Core(TM) Core i7-6700 @ 3.40GHz.
RAM	2133 GHz, 8 GB
OS	Ubuntu 14.04 LTS x64
Develop tools	Python 3.5
Tool Library	gensim-0.13.2

Table 6. Comparison on execution time

	w5	Time:(s)	**w10**	Time:(s)	**w15**	Time:(s)
Word2Vec	8769	11837	8718	18574	8636	23576
r100	8798	11981	8796	18698	**8820**	23849
r125	8757	12345	8784	19134	8796	24481
r150	8782	12407	8801	19327	8810	24579
r175	8764	12504	**8817**	19709	8795	25032
r200	**8808**	12304	8801	19157	8785	24386
r225	8751	11993	8794	18768	8795	23845

As can be seen, our method runs slower than Word2Vec, but only slightly.

V. Conclusion

There are many situations where the words with similar meanings appear at the same time in the same sentence. Word2Vec adopts the bag-of-words model to train a type of word representation with the idea of context window. It makes senses that relevant words may appear simultaneously in a context window. However, Word2Vec ignores the distance information and treats each word equally in the context. In our work, we offer an alternative method to weight the contextual words at different positions. Fuzzy values are exploited to express the degree of importance a context word imposed on the target due to the distance between them. The derived fuzzy values take part in the training of network language models. From the experimental results, we have seen that our method can improve the semantic accuracy for the Google-One-Billion-Word-Language-Modeling Benchmark dataset.

References

[1] Salton, G.and Lesk, M. E. (1968) Computer evaluation of indexing and text processing, *Journal of the ACM (JACM)* **15**, 8–36.

[2] Sparck Jones, K. (1972) A statistical interpretation of term specificity and its application in retrieval, *Journal of documentation* **28**, 11–21.

[3] Salton, G. and Buckley, C. (1988) Term-weighting approaches in automatic text retrieval, *Information processing & management* **24**, 513–523.

[4] Turian, J., Ratinov , L. and Bengio, Y. (2010) Word representations: a simple and general method for semi-supervised learning, *Proceedings of the 48th annual meeting of the association for computational linguistics*, 384–394.

[5] Bengio, Y., Ducharme, R., Vincent, P. and Jauvin, C. (2003) A neural probabilistic language model, *Journal of machine learning research* **3**, 1137–1155.

[6] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S. and Dean, J. (2013) Distributed representations of words and phrases and their compositionality, *Advances in Neural Information Processing Systems* **26**, 3111–3119.

[7] Mikolov, T., Chen, K., Corrado, G. and Dean, J. (2013) Efficient estimation of word representations in vector space, *CoRR*, **abs/1301.3781**.

[8] Kim, Y. (2014) Convolutional neural networks for sentence classification, *arXiv preprint arXiv:1408.5882*.

[9] Socher, R., Bauer, J., Manning, C. D. and Ng, A. Y. (2013) Parsing with compositional vector grammars. *ACL (1)*, 455–465.

[10] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., Potts, C. et al., (2013) Recursive deep models for semantic compositionality over a sentiment treebank, *Proceedings of the conference on empirical methods in natural language processing (EMNLP)* **1631**, 1642

[11] Xue, B., Fu, C. and Shaobin, Z. (2014) A study on sentiment computing and classification of sina weibo with word2vec, *IEEE International Congress on Big Data (Big Data Congress)*, 358–363.

[12] Nigam, K., McCallum, A. K., Thrun, S. and Mitchell, T. (2000) Text classification from labeled and unlabeled documents using EM, *Machine learning* **39**, 103–134.

About control of convection intensity in chemical active equilibrium gas

†I.B. Palymskiy[1], V.I. Palymskiy[2], P.A. Fomin[3], l.V. Frolov[4]

[1]Siberian State University of Telecommunications and Information Sciences, 630102, Novosibirsk, Russia

[2]Kutateladze Institute of Thermophysics, SB RAS, 630090, Novosibirsk, Russia

[3]Lavrentyev Institute of Hydrodynamics, SB RAS, 630090, Novosibirsk, Russia

[4]Novosibirsk State University, Russia

† Presenting and corresponding author: palymsky@hnet.ru

Abstract

The problem of the Rayleigh-Benard convection for a chemical equilibrium gas is investigated using methods of linear analysis. The Boussinesq approximation is used for a plane layer with isothermal horizontal boundaries free from shearing stresses. A previously developed model for chemical equilibrium is applied to calculate the thermodynamic parameters of the gas analyzed (hydrogen-oxygen mixture). Recombination and dissociation processes are shown to introduce an additional factor multiplying the Rayleigh number defined as in inert gas. Formulas for the increment of growth o infinitesimal perturbation and the critical Rayleigh number as a function of absolute temperature are obtained.

Key words: Rayleigh-Benard convection, Rayleigh number, Prandtl number, chemical equilibrium, recombination, dissociation.

Introduction

The Rayleigh-Benard convection is a type of natural convection belonging to the area of classic science in whose frameworks a mathematical model based on the Boussinesq approximation and numeric methods were developed for this convection [1-4,13]. Scientists mainly focused on modelling convection for chemically inert gases and liquids.

Convection processes happen in chemically reacting media as well. Such type of convection can be observed while gas mixture burns in large-scale reactors, technical installations and experimental facilities.

Convection of a gas in chemical equilibrium is a particular, though quite common, case of convection in gases accompanied by chemical reactions, which can occur if the time scale of the reaction is much shorter than that characteristic for the convection process.

It should be expected that quantitative and qualitative characteristics of convection mainly depend on the type of the thermal expansion coefficient dependence on temperature, which determines the buoyancy force casing convection.

The thermal expansion coefficient of a chemically inert gas is monotonic and inversely related to the temperature. In the case of a chemically active equilibrium gas, the dependence of this coefficient on temperature changes considerably and becomes non-monotonic with a clear local maximum, where the value of the coefficient for a chemical equilibrium gas can exceed that of a chemically inert gas by several times [5-9]. Such anomalous dependence of the coefficient on the temperature expansion (other than $1/T$) leads to considerable changes in quantitative and qualitative characteristics of convection regimes in a chemical equilibrium gas compared to those in a chemically inert gas.

Authors in [10-12] were the first to formulate and solve numerically a two-dimensional problem for Rayleigh-Benard convection in a chemically active equilibrium gas, as well as considered linear and non-linear regimes. The molar mass and thermodynamic parameters of

the gas were calculated according to a previously developed model of chemical equilibrium [5-9]. The results of the linear and non-linear analysis [10-12] need additional investigations and improvements. In addition, the physical and mathematical model provided in [10-12] is quite complicated and inconvenient; while representing the dependence of the thermal expansion coefficient on the whole range of the real (not dimensionless!) temperature, a second degree polynomial equation with four constants is used instead.

Further investigations into the problem showed that the realization of the model proposed in [10-12] can be considerably simplified as recombination and dissociation processes in gases under the temperature given can be taken into account through multiplying the Rayleigh number by an additional factor, the function of absolute temperature.

In terms of convection in a laboratory, the gas compressibility can be neglected and the convection media can be considered as incompressible. Moreover, if the domain has a certain geometry and small supercriticality, convection develops as two-dimensional rollers [13], which allows us to consider two-dimensional convective flows of a viscous incompressible fluid even at the first stage of our investigation.

Following [10-12], we consider a convective two-dimensional and non-stationary flow of a viscous incompressible fluid in the horizontal layer heated from below in the Boussinesq approximation. The boundaries of the layer are assumed isothermal and free from shearing stresses. Thermodynamic parameters of a chemical equilibrium gas are calculated according to the model in [5-9].

The present work is aimed at describing a simpler and improved realization of a physical and mathematical model for convection of a chemically active gas in equilibrium and at specifying the results of the linear analysis.

Nomenclature

The following notations will be used:

ρ, P, T *and* μ	-	the density, pressure, absolute temperature and molar mass of the gas;
R	-	the universal gas constant;
μ_{min}, μ_{max}	-	the molar masses of the gas in the completely dissociated and completely recombined states;
A, K_+	-	the rate constants of dissociation and recombination of the generalized reaction products;
θ	-	the effective excitation temperature of the vibrational degrees of freedom of the molecules;
E	-	the mean dissociation energy of the reaction products;
β	-	the thermal expansion coefficient of the gas with chemical reactions;
$\beta_{in} = 1/T$	-	the thermal expansion coefficient of the inert ideal gas;
u, v	-	the velocities in x and y directions;
ν, χ	-	the kinematic and thermal diffusivity coefficients;
T_h, T_c	-	the temperatures of the lower and upper boundaries;
H	-	the thickness of the layer;
ψ	-	the stream function ($\psi_y = u$ and $\psi_x = -v$);
ω	-	the vorticity $\omega = v_x - u_y$;
$Ra = g\beta H^3 \delta T/\chi\nu$	-	Rayleigh number;
Ra_{cr}, α_m	-	the critical values of Rayleigh and wave numbers;
$Pr = \nu/\chi$	-	Prandtl number;
$\lambda, \alpha, \pi m$	-	the eigenvalue and the wavenumbers of linear problem in x and y directions;
$Q = T\text{-}T_e$	-	the temperature deviation from linear equilibrium distribution.

Kinetic model

To describe chemical equilibrium in the gas mixture, we use a highly accurate kinetic model [5-9]. According to this approach, the density of a chemical equilibrium ideal gas can be considered as a function of the absolute temperature and pressure:

$$\rho = \frac{p\mu}{RT},$$

$$\mu = \frac{B\mu_{\min} - 2\mu_{\max} + ((B\mu_{\min} - 2\mu_{\max})^2 + 4(B-1)\mu_{\max}^2)^{0.5}}{2(B-1)}, \quad (1)$$

$$B = \frac{AT^{0.75}}{4K_+}\cdot(1-\exp(-\Theta/T))^{1.5}\exp(-E/RT))\frac{\mu_{\max}^2}{\mu_{\min}\rho}.$$

The molar mass of the gas can vary within the interval $\mu_{min} \le \mu \le \mu_{max}$. We have $\mu \to \mu_{max}$ with decreasing temperature and $\mu \to \mu_{min}$ with increasing temperature.

We assume that the pressure P in the gas mixture (external pressure) is constant. Then, the gas density ρ, as is seen from system (1), depends on the absolute temperature T only. For the purposes of comparing our results with those in [10-12], our linear analysis was performed with the pressure changing in the same range, from 10^{-5}atm to 10^{5}atm.

The model of chemical equilibrium used [5-9] favors simplicity, accuracy and compatibility with the principle of entropy increase. It takes into account considerable changes of the molar mass, the temperature effect, heat capacities and the ratio of specific heats due to the change of chemical equilibrium. All the constants in the model are physically clear.

Such equations are usually used to calculate chemical equilibrium in hydrogen-oxygen mixture and hydrocarbon-oxygen mixtures [6-8], in heterogeneous mixtures of gas-condensed-phase type in the presence of the interphase mass transfer [7] and heterogeneous mixtures of gases with chemically inert particles [9]. Two-stage models of detonation combustion of methane in oxygen and air have also been proposed [14].

The model of chemical equilibrium in question is particularly useful for scientific and applied problems, e.g., while modelling detonation processes in chemically active bubble systems [14,15], designing equipment for fluidized bed coating by means of gaseous detonation [16], using detonation suppression by inert particles [17-19], calculating concentration limits for gas mixture flammability [20].

As it is seen from system (1), the ratio for the molar mass as a function of temperature is as follows:

$$\mu(T) = \mu_{\max} + 0.5\cdot(f - \sqrt{f^2 + 4f(\mu_{\max} - \mu_{\min})}),$$

$$f = \frac{AR\mu_{\max}^2 T^{1.75}}{4K_+ P\mu_{\min}}\cdot(1-\exp(-\Theta/T))^{1.5}\exp(-E/RT)).$$

The present paper refers to a stoichiometric hydrogen-oxygen mixture. The numerical values of the constants were as follow: *E = 459.2·kJ/mole, Θ = 4000K,* $K_+ = 6\cdot10^8 m^6/(kmole^2\cdot sec)$, $A = 5.1\cdot10^{10} m^3/(kmole\cdot sec\cdot K^{3/4})$, μ_{min} *= 6kg/kmole,* μ_{max} *= 18kg/kmole* [5-9].

Considering the ratio for f at comparatively "low" and high temperatures T and taking into account

$$f \sim T^{1.75}\cdot\exp(-\frac{E}{RT}) \sim \exp(-\frac{E}{RT}) \ (T \gg \Theta), \ f \sim T^{0.25} \ (E/R \gg T),$$

we have the following for the "low" temperature

$$T \to 0, \ f \to 0, \ \mu_{as1} = \mu_{\max} - f^{0.5}(\mu_{\max} - \mu_{\min})^{0.5} + f/2 + O(f^{3/2})$$

and for the high temperature

$T \to \infty,\ f \to \infty,\ \mu_{as2} = \mu_{\min} + (\mu_{\max} - \mu_{\min})^2 / f + O(f^{-2}).$

The thermal expansion coefficient β can be found dependent on the density or the molar mass as

$$\beta = -\frac{1}{\rho}\cdot\frac{\partial\rho}{\partial T} = \frac{d}{dT}\ln(\mu/T) = \frac{1}{T} - \frac{1}{\mu}\frac{d\mu}{dT}.$$

Notice that the asymptotic approximation of the molar mass μ to constants as the temperature is reduced or increased allows us to find the asymptotics for the thermal expansion coefficient β as $\beta = 1/T$.

Figures 1-4 and 7 show the asymptotic curves calculated with the full formula for f where all the coefficients in expression for f were used. Figures 1 - 7 were plotted with $P = 1atm$.

The solid curve in Fig. 1 shows the molar mass $\mu = \mu(T)$, asymptotics for the "low" ($\mu_{as}1$, curve 1) and high ($\mu_{as}2$, curve 2) temperatures as well as $\mu = \mu_{max}$ (3) and $\mu = \mu_{min}$ (4). The asymptotic curves $\mu_{as}1$ (curve 1) and $\mu_{as}2$ (curve 2) are correct for both comparatively low ($T \le 3500K$) and high ($T \ge 5000K$) temperatures.

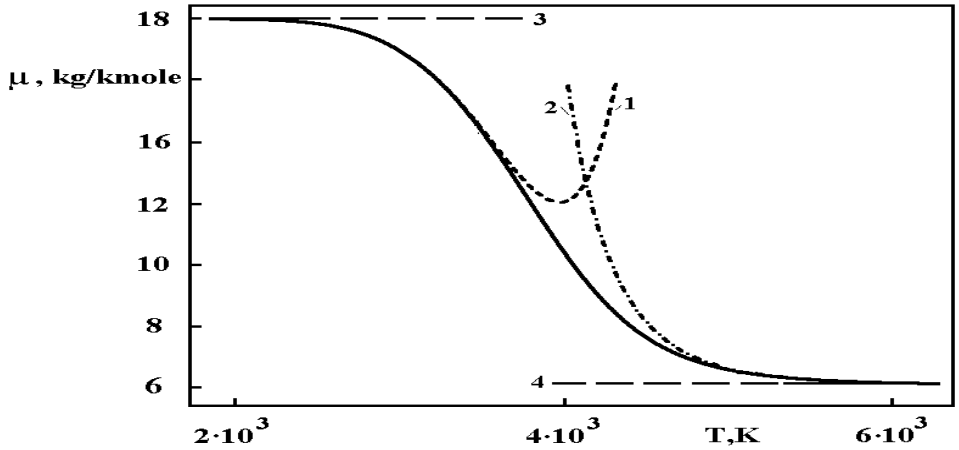


Fig. 1. The molar mass and its asymptotics

The solid curve in Fig. 2 shows the density of the chemical equilibrium gas $\rho = \rho(T)$ at $2\cdot10^3K \le T \le 6\cdot10^3K$. Again, we see two asymptotics $\rho_1 = \mu_{as1}P/(RT)$ (dashed curve 1) and $\rho_2 = \mu_{as2}P/(RT)$ (dashed curve 2) corresponding to comparatively low ($T \le 3500K$) and high ($T \ge 5000K$) temperatures, in comparison to curves $\rho_3 = \mu_{max}P/(RT)$ (dot-and-dash curve 3) and $\rho_4 = \mu_{min}P/(RT)$ (dot-and-dash curve 4) for absolutely recombined and dissociated conditions respectively. The curves are quite accurate for both low and high temperatures. Notice that throughout the range of temperature changes the density depends on the temperature monotonically.

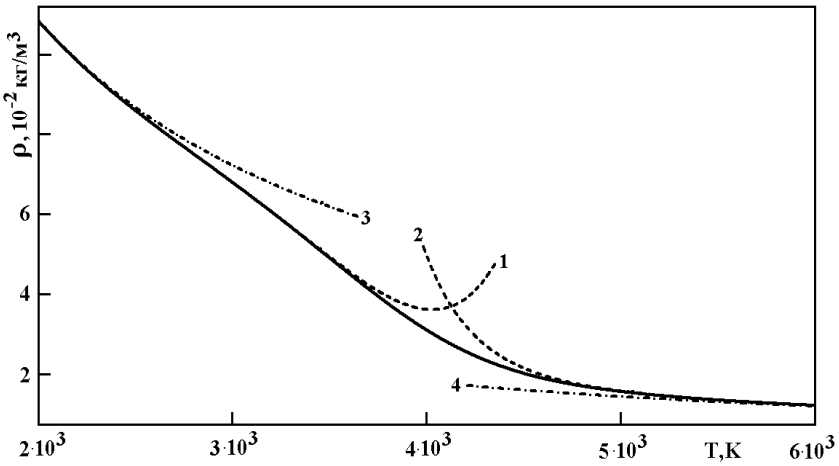


Fig. 2. The gas density

The solid curve in Fig. 3 shows the thermal expansion coefficient $\beta = \beta(T)$ while the dot-and-dash line $\beta = 1/T$ for ideal gas shows the close values at $T \leq 2000K$ and $T \geq 6000K$. Again, we see two asymptotics (dashed curves 1 and 2) with close values at comparatively low ($T \leq 3500K$) and high ($T \geq 5000K$) temperatures. We see the temperature interval where the thermal expansion coefficient β demonstrates an anomalous behavior (different from $1/T$) with noticeable local maximum. As it is shown below, it is this anomalous behavior that causes a possibility of convective instability amplification in comparison with inert gas.

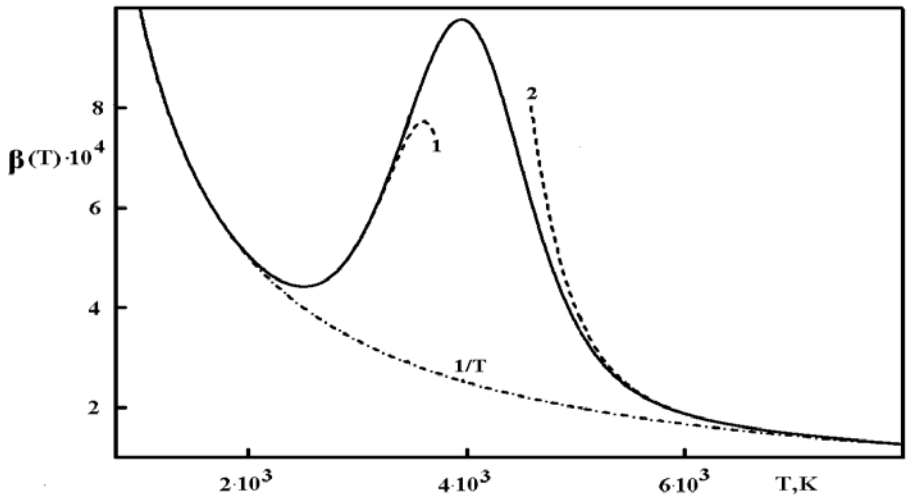


Fig.3. The thermal expansion coefficient

Mathematical model

Convection of incompressible fluids in the Boussinesq approximation is described by the following system of equations [1,2]:

$$\omega_t + \frac{1}{\mathrm{Pr}}(\psi_y\omega_x - \psi_x\omega_y) = \Delta\omega + C\cdot Ra\mathrm{Q}_x,$$
$$\Delta\psi = -\omega,$$
$$\mathrm{Q}_t + \frac{1}{\mathrm{Pr}}(\psi_y\mathrm{Q}_x - \psi_x\mathrm{Q}_y) = \frac{1}{\mathrm{Pr}}\Delta\mathrm{Q} - \frac{1}{\mathrm{Pr}}\psi_x, \quad (2)$$
$$C = \frac{\beta(T)}{\beta_{in}(T)} = T\cdot\beta(T).$$

Here, $Ra = g\beta_{in}H^3\delta T/\chi\nu = gH^3\delta T/(T\chi\nu)$ and $Pr = \nu/\chi$ are Rayleigh and Prandtl numbers, here $\beta_{in} = 1/T$ is the thermal expansion coefficient of the inert ideal gas as function of the absolute temperature T.

We choose the following reference values for the layer: H for length, H^2/ν for time, χ/H for velocity, $\rho_0\nu\chi/H^2$ for pressure, $\rho_0 = \rho(T)$ for density, and $\delta T = T_h - T_c$ for temperature, Q is temperature deviation from equilibrium linear distribution, $\Delta f = f_{xx} + f_{yy}$ is the Laplace operator, φ and ω are stream function and vorticity respectively. The thermal expansion coefficient β is considering as constant in system (2) at chosen absolute temperature T.

According to the buoyancy force in the first equation of system (2), we see that the convection rate is defined by $C\cdot Ra$. The Rayleigh criterion specifies the ratio between the buoyancy force and the viscosity and depends on the temperature gradient or the degree of heating, while the dimensionless factor C depends on the recombination and dissociation processes under the temperature given.

Figure 4 shows function $C = C(T)$ at $P = 1atm$ and two asymptotes (dashed curves 1 and 2). It can be seen that in domain ($2000K \le T \le 8000K$) the chemical reactions in gas increase the convection intensity. When the temperature falls ($T < 2000K$) or rises ($T > 8000K$), the convection intensity decreases up to convection intensity level in an ideal inert gas. The maximum convection intensity is realized at the points of local extremum of the factor $C = 3.907$ at $T = 4051K$.

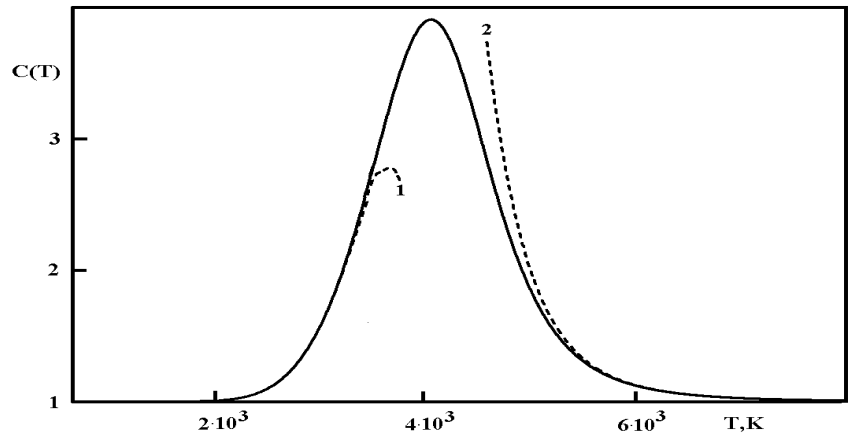


Fig. 4. Factor $C = C(T)$

Considering the relation for f at comparatively "low" and high absolute temperatures T, we have the following asymptotics for the "low" and high temperature

$$fas1 = \frac{AR\mu_{\max}^2 T^{1.75}}{4K_+ P\mu_{\min}}\cdot\exp(-E/RT)) \quad (T \square \Theta),$$
$$fas2 = \frac{AR\mu_{\max}^2 T^{0.25}\Theta^{1.5}}{4K_+ P\mu_{\min}} \quad (E/R \square T).$$

The solid curve in Fig. 5 shows the function $f = f(T)$ and its two asymptotes fas1 and fas2 (dashed curves 1 and 2) with close values at comparatively "low" ($T \le 4000K$) and high ($T \ge 55000K$) temperatures. The two vertical lines correspond $T = \Theta = 4000K$ and $T = E/R = 5.523 \cdot 10^4$.

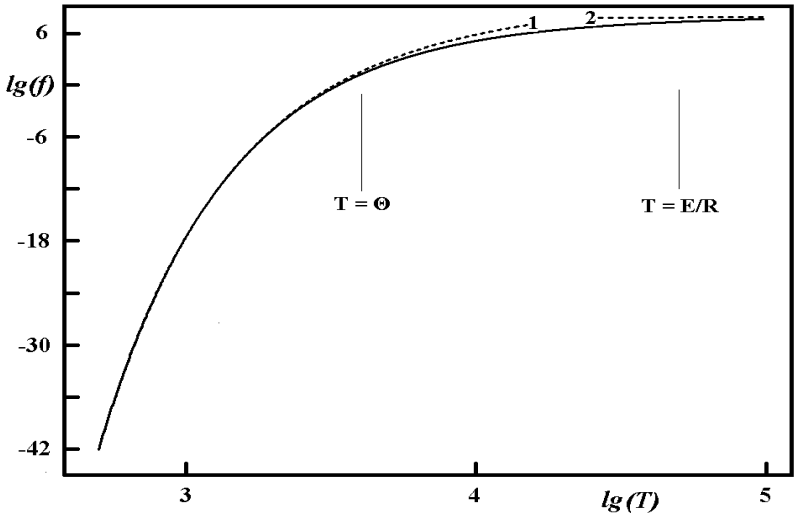


Fig. 5. The function $f(T)$

At a considerably high temperature ($T \sim E/R \sim 5 \cdot 10^4 K$) power asymptotics for f is stated, which allows us to obtain far-field asymptotics for β and C.
Actually, taking into account

$$f \to fas2 = \frac{AR\,\mu_{\max}^2 \Theta^{1.5}}{4K_+ P\,\mu_{\min}} T^{0.25} \ (T \to \infty),$$

we find $\beta = \frac{1}{T}\left(1 + \frac{(\mu_{\max} - \mu_{\min})^2 K_+ P}{AR\,\mu_{\max}^2\,\Theta^{1.5}} \frac{1}{T^{0.25}}\right)$, $C = 1 + \frac{(\mu_{\max} - \mu_{\min})^2 K_+ P}{AR\,\mu_{\max}^2\,\Theta^{1.5}} \frac{1}{T^{0.25}}$.

The dependences given show that β and C approach asymptotics with the rise of the temperature quite slowly, which means that the corresponding boundary is fuzzy.

Linear analysis

We consider the linear analog of system (2):

$$\begin{aligned} \omega_t &= \Delta\omega + C \cdot Ra\,\mathrm{Q}_x, \\ \Delta\psi &= -\omega, \\ \mathrm{Q}_t &= \frac{1}{\mathrm{Pr}}\Delta\mathrm{Q} - \frac{1}{\mathrm{Pr}}\psi_x. \end{aligned} \tag{3}$$

We seek approximate solutions of system (3) in the form:

$$\begin{aligned} \omega(t,x,y) &= \Omega \exp(-\lambda t)\cos(\alpha x)\sin(m\pi y), \\ \psi(t,x,y) &= \Omega \exp(-\lambda t)\cos(\alpha x)\sin(m\pi y)/S, \\ \mathrm{Q}(t,x,y) &= \Theta \exp(-\lambda t)\cos(\alpha x)\sin(m\pi y), \\ S &= \alpha^2 + m^2\pi^2. \end{aligned} \tag{4}$$

Here λ is the eigenvalue, where $\lambda > 0$ is matched by amplitude losses and $\lambda < 0$ by amplitude rises in harmonic solutions, α and $m\pi$ are the wavenumbers in x and y directions, and Ω and Θ are constants.

Plugging (4) into (3) and using standard calculations [1,2] we find:

$$\lambda_{1,2} = \frac{1+\mathrm{Pr}}{2\,\mathrm{Pr}} S \pm \sqrt{(\frac{1-\mathrm{Pr}}{2\,\mathrm{Pr}})^2 S^2 + \frac{C \cdot Ra \cdot \alpha^2}{S\,\mathrm{Pr}}}, \tag{5}$$

where $S = \alpha^2 + m^2\pi^2$, too.

Let us consider the ratio $d\lambda = \lambda_{in}/\lambda_r$ between maximal increments in the inert λ_{in} and reacting λ_r gases. The asymptotical value of $d\lambda$ can be found as

$$\frac{\lambda_{in}}{\lambda_r} \to 1/\sqrt{C\max} = 0.506 \ \ at\ Ra \to \infty.$$

Here the maximum of the factor *Cmax* is equal to 3.907 at *P = 1atm and T = 4051K*.

Figure 6 shows the value $d\lambda = \lambda_{in}/\lambda_r$ as function of the supercriticality $r = Ra/657.511$ where the pressure and absolute temperature are equal to 1atm and 4051K respectively. It can be seen that three distinct domains exist at various supercriticality r. At $0 < r < 1/Cmax$, $1/Cmax = 0.256$ (domain d1) only the quantitative difference is observed since the convection in inert and reacting gases is stable. At $1/Cmax \le r < 1$ (d2) the qualitative difference is seen since the convection in inert gas is stable and convection in reacting gas is unstable or neutral. At $1 \le r$ (d3) the quantitative difference is observed since the convection in the inert and reacting gases is unstable (or neutral for inert gas).

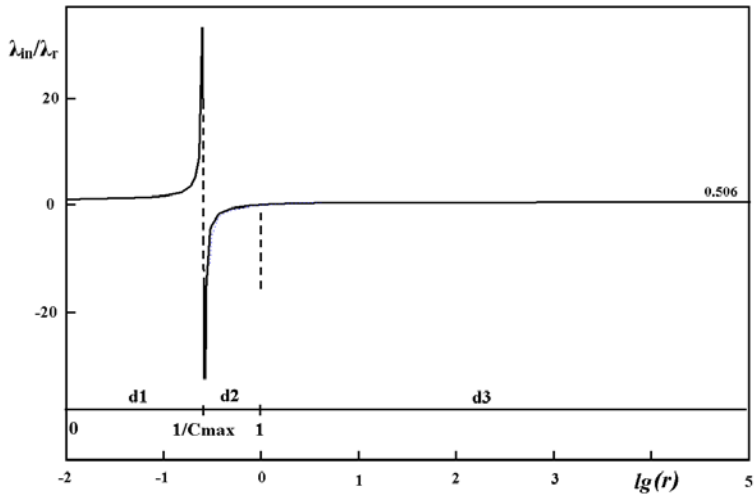


Fig.6. The increment ratio $d\lambda$ versus supercriticality r

Assuming increment λ_2 in (5) equal to 0, one can obtain the Rayleigh number corresponding to the neutral curve:

$$Ra = \frac{(m^2\pi^2 + \alpha^2)^3}{C \cdot \alpha^2}.$$

The critical value of the Rayleigh number (minimum of *Ra* versus α with $m = 1$) can be found by the formula

$$Ra_{cr} = \frac{27}{4C}\pi^4 = \frac{657.511}{C},\ \alpha_{cr} = 2.221.$$

The factor 657.511 is identified here as the critical value of the Rayleigh number for the convection in a chemically inert media [1,2], and the factor $1/C$ as the function of temperature and pressure is due to recombination and dissociation processes.

Providing *C - 1* is small (for $|T - \Theta| >> 1$), one can obtain the critical value of the Rayleigh number as the principal member of the power series expansion:

$$Ra_{cr} = \frac{657.511}{C} = 657.511 \cdot (2 - C).$$

Using the *C* asymptotics found above under the temperature approaching infinity, we can obtain the far-field asymptotics for Ra_{cr} and lengthy asymptotic formula for increment λ.
Taking into account

$$C = 1 + \frac{(\mu_{\max} - \mu_{\min})^2 K_+ P}{4 A R \mu_{\max}^2 \Theta^{1.5}} \frac{1}{T^{0.25}} \ (T \to \infty), \quad \text{we have}$$

$$Ra_{cr} = 657.511 \cdot (1 - \frac{(\mu_{\max} - \mu_{\min})^2 K_+ P}{4 A R \mu_{\max}^2 \Theta^{1.5} T^{0.25}}) \quad \text{and}$$

$$\lambda_r = \lambda_{in} \pm \frac{(\mu_{\max} - \mu_{\min})^2 K_+ P \, Ra \, \alpha^2}{4 A R \mu_{\max}^2 \Theta^{1.5} S^{0.5} \sqrt{(1 - \mathrm{Pr})^2 S^3 + 4 Ra \, \alpha^2 \, \mathrm{Pr}}} T^{-0.25}.$$

Here λ_r and λ_{in} are increments for the reacting and inert gases correspondingly.
Figure 7 shows the critical Rayleigh number as a function of *T* for pressure *P = 1atm* (solid curve), two its asymptotes (dashed curves 1 and 2) and $Ra_{cr} = 657.511 \cdot (2\text{-}C)$ (dashed curve 3). It can be seen that in domain ($3000K \le T \le 5000K$, $T \approx \Theta = 4000K$) the chemical reactions in gas considerably decrease the critical Rayleigh number.

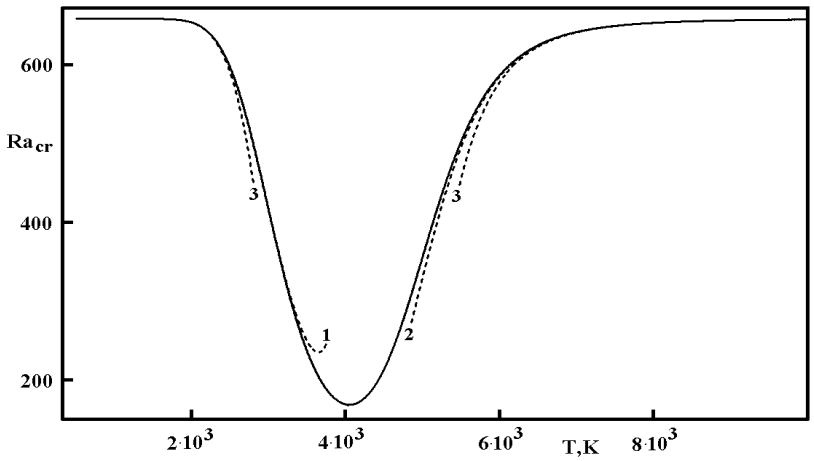


Fig.7. The neutral curve in the plane *T-Ra* for pressure 1atm

Figure 8 shows the critical Rayleigh number as a function of *T* for pressures *P = 1atm* (solid curve 1), *P = 0.1atm* (dashed curve 2), and *P = 10atm* (dashed curve 3). The pressure increasing, the convection intensity amplification domains move while expanding towards higher temperatures. The latter is more obvious for the high-temperature boundary of domain.

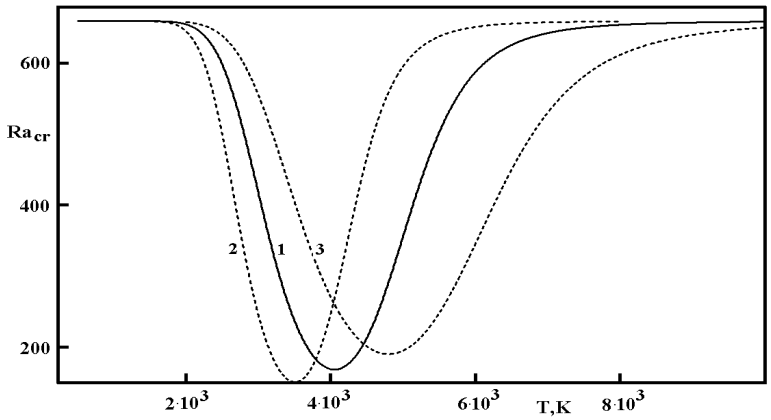


Fig.8. The neutral curve in the plane *T-Ra* for various pressure

As one can see, Figure 8 shows a relatively low increase in the critical Rayleigh number (minimum with respect to *T*) while the pressure increases (about 3.6 times while the pressure increases by 10 orders of magnitude):
$Ra_{cr} = 657.511/C \approx 657.511/(3.9241 - 0.44393 \cdot lg(P/1atm))$ and
$Ra_{cr} \approx 159.33 + 8.2324 \cdot P/1atm$ at the near-normal pressure $P \approx 1atm$.
The convective processes intensiveness can be influenced. The critical Rayleigh number Ra_{cr} is weakly depending on the pressure, the intensiveness decreases as the pressure grows, changing from the maximum values at low pressure to minimal values at high pressure. Another and cardinal way to affect the intensiveness is to change the temperature. The intensiveness of convective processes is equal to zero at high and low temperatures and is different in between. Of course, it is possible to change both the pressure and the temperature.
In order to study the possibility to control convective processes by changing the pressure and the temperature in more details, consider the instability domain on *R-T* plane in Fig. 9.
A curve 1 in Fig. 9 correspond to the temperatures at which the maximum convection intensiveness at given pressure is observed (maximum *C* with respect to *T* corresponds to curve Ra_{cr}). Curves 2 and curve 3 correspond to the top and low boundaries of domain where reacting-inert gas transitions are observed.

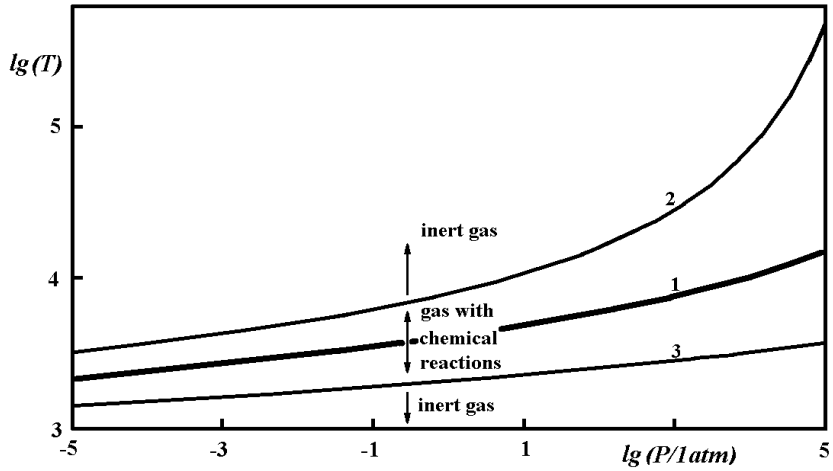


Fig. 9. The instability domain on the plane *R-T*

It is obvious that the highest increase in temperature by 155 times occurs at the upper boundary of the reacting-inert gas transition (curve 2), whereas the other curves demonstrate a substantially smaller change of temperature. For instance, at the lower boundary of the reacting-inert gas transition (curve 3), the change is by one order of magnitude less.

Curve 1 (the maximum convection intensiveness at given pressure) on the plane *R-T* is accurately described throughout the range of pressure change by

$$T(P) = \frac{1212.4}{1.2957 \cdot (P/1atm)^{-0.015848} - 1},$$

and at the near-normal pressure $P \approx 1atm$ by

$T(P) = 3814.9 + 284.67 \cdot lg(P/1atm)$.

Similar approximate expressions for curves 2 and 3 on the plane *R-T*, where the top and low boundaries of the reacting-inert gas transition are realized as follow:

$$T(P) = 10^{\frac{3.8932 - 0.49630 \cdot \lg(P/1atm)}{1 - 0.15697 \cdot \lg(P/1atm) + 0.0013110 \cdot (\lg(P/1atm))^2}}$$ - for top curve 2

and

$$T(P) = \frac{2373.9}{1.4891 \cdot (P/1atm)^{-0.026968} - 1}$$ - for low boundary curve 3.

Considering the expressions above for $P \sim 1atm$, we obtain that:

$T(P) = 6922.0 + 897.84 \cdot P/1atm$ for curve 2 and

$T(P) = 1968.8 + 79.253 \cdot P/1atm$ for curve 3.

At the same time, the value of local maximum *C* that characterize the convection intensiveness vanishes in a practically linear way with respect to *lg(P)*:

$C(P) = 3.9241 - 0.44393 \cdot lg(P/1atm)$ along curve 1 and

$C(P) = 4.1169 - 0.19280 \cdot P/1atm$ at the near-normal pressure $P \approx 1atm$.

Conclusion

A new physical and mathematical model for convective flows of a gas in chemical equilibrium has been suggested. The flow is considered in a horizontal layer with heating from below. The gas is assumed to be viscous and incompressible, and the convective flow two dimensional and non-stationary. The horizontal boundaries of the layer are isothermal and free from shearing stresses. The parameters of the gas in chemical equilibrium (hydrogen-oxygen mixture) are calculated according to the model of chemical equilibrium proposed earlier. Given recombination and dissociation processes, an additional factor $C = C(T)$ appears as factor with the Rayleigh number in the system of equations describing convection. Hence, all the regimes and characteristics of convective flows of a chemical equilibrium gas can be derived from the corresponding characteristics of convection in a chemically inert medium when the Rayleigh number is overdetermined.

The results of linear analysis of the stability of a convective flow are presented. In particular, for a chemically active gas in equilibrium we obtained formula for the infinitesimal perturbation growth rate increments and formula for the critical Rayleigh number depending on temperature at given pressure. The neutral curve on the plane *T-Ra* shows the significant decreasing of the critical Rayleigh number in the reacting gas. The critical Rayleigh number (minimum with respect to *T*) depends on the pressure relatively weakly. The domain of the intensity convection increase due to chemical reactions in reacting gas on the plane *R-T* extends significantly towards higher temperatures when the pressure increases, which is especially true for the top high-temperature inert-reacting gas boundary.

Discussion

If we take into account recombination and dissociation processes, there appears an additional factor $C = C(T)$ with the Rayleigh number in the system of equations describing convection. Hence, all the regimes and characteristics of convective flows of a chemical equilibrium gas can be derived from the corresponding characteristics of convection in a chemically inert

medium when the Rayleigh number is re-determined. In this respect, calculations of non-linear regimes of convection of a chemically active equilibrium gas yield little information.
In the present work the far-field high-temperature asymptotics for factor *C*, critical Rayleigh number and the increment λ is derived. It would be of certain interest to investigate also the low-temperature asymptotics.
The performed linear analysis shows the existence of the anomalous (non-monotonic) behavior of the thermal expansion coefficient in spite of the monotonic dependence of density on temperature. It seems to be possible that such anomalous (non-monotonic) behavior of the thermal expansion coefficient must bring to the secondary instability because of thermal expansion coefficient temperature dependence.
It would be of certain interest to prove experimentally the existence of the secondary instability because of thermal expansion coefficient temperature dependence. Another point of interest is related to the generation of the new physical and mathematical model and numerical investigation of such secondary instability. Of cause, the numerical investigation of this secondary instability must be performed with investigation of all regimes of flow together with defining the boundaries of temperature intervals where such regimes are observed, followed by researching integral characteristics, etc.
We acknowledge financial support of this work by Russian Foundation for Basic Research Gr. №15-08-05166 and №17-58-53100.

Literature

1. G.Z. Gershuni and E.M. Zhuchovitskii. Convective Stability of Incompressible Fluids (Konvektivnaya ustojchivost' neszhimaemoj zhidkosti). Nauka, Moscow, 1972 (in Russian). English translation: G.Z. Gershuni, E.M. Zhukhovitskii. Convective Stability of Incompressible Fluids (Israel Program for Scientific Translations). Jerusalem, 1976.
2. Palymskiy, I.B. [2011], Turbulent Rayleigh-Benard Convection. Numerical Method and Results of Simulation (Turbulentnaya konvektziya Rehleya-Benara. Chislennyi metod i resultaty raschetov), LAP, Germany (in Russian).
3. V.M. Paskonov, V.I. Polezhaev, L.A. Chudov. Numerical Modeling of Heat and Mass Exchange Processes. Moscow, Nauka, 1984 (in Russian).
4. Palymskiy I.B. Vorticity scale and integral values of Rayleigh-Benard convection. Computational Thermal Sciences. 2014. Vol.6. N2, P.113-127.
5. Nikolaev Yu.A. Model of the kinetics of chemical reactions at high temperatures // Combustion, Explosion and Shock Waves.- 1978.- V. 14.- N. 4.- P. 468-471.
6. Yu.A. Nikolaev, P.A. Fomin. Analysis of equilibrium flows of chemically reacting gases. Combustion, Explosion and Shock Waves, 1982, 18, 1, pp.53-58.
7. Yu.A. Nikolaev, P.A. Fomin. Approximate equation of kinetics in heterogeneous systems of gas-condensed-phase type. Combustion, Explosion and Shock Waves, 1983, 19, 6, pp. 737-745.
8. Yu.A. Nikolaev, D.V. Zak. Agreement of models of chemical reactions in gases with the second law of thermodynamics. Combustion, Explosion and Shock Waves, 1988, 24, 4, pp. 461-464.
9. P.A. Fomin, A.V. Trotsyuk. An approximate calculation of the isentrope of a gas in chemical equilibrium. Combustion, Explosion and Shock Waves, 1995, 31, 4, pp. 455-457.
10. Palymskiy I., Fomin P.A., Hieronymus H. Rayleigh-Benard convection in chemical equilibrium gas. Proceedings of the Fourth International Conference on Computational Heat and Mass Transfer (ICCHMT'05), 17-20 May 2005, Paris, France, in book: Progress in Computational Heat and Mass Transfer.- V.1.- P. 116-122.

11. Palymskiy I., Fomin P.A., Hieronymus H. The Rayleigh–Benard convection in gas with chemical reactions // Siberian Journal of Numerical Mathematics. – 2007.- V. 10.- N. 4.- P.371–383.
12. Palymskiy I., Fomin P.A., Hieronymus H. Rayleigh-Benard convection in a chemical equilibrium gas (simulation of surface detonation wave initiation) // Applied Mathematical Modelling.- 2008.- V. 32.- Is. 5.- P. 660–676.
13. A.V. Getling. Rayleigh-Bénard convection: Structures and Dynamics. World Scientific Publishing Co (Singapore), 1998, ISBN 9810226578.
14. Fomin P.A., Mitropetros K.S., Hieronymus H. Modeling of detonation processes in chemically active bubble systems at normal and elevated initial pressures // Journal of Loss Prevention in the Process Industries.- 2003.- V. 16.- N. 4.- P. 323-331.
15. Fomin P.A., Chen J.-R. Shock induced condensation in a fuel-rich oxygen containing bubble in a flammable liquid // Chemical Engineering Science.- 2008.- V. 63.- N. 3.- P. 696-710.
16. Gavrilenko T.P., Grigoriev V.V., Zhdan S.A., Nikolaev Yu.A., Boiko V.M., Papyrin A.N. Acceleration of solid particles by gaseous detonation products // Combustion and Flame.- 1986.- V. 66.- N. 2. - P. 121-128.
17. Fomin Pavel A., Chen Jenq-Renn. Effect of Chemically Inert Particles on Thermodynamic Characteristics and Detonation of a Combustible Gas // Combustion Science and Technology.- 2009.- V. 181.- Is. 8.- P. 1038-1064.
18. Fedorov A.V., Fomin P.A., Fomin V.M., Tropin D.A., Chen J.R. Mathematical Analysis of Detonation Suppression by Inert Particles.- Taiwan: Kao Tech Publishing, Kaohsiung, 2012.
19. Fomin P.A., Fedorov A.V., Chen Jenq-Renn. Control of explosions in silane-air mixtures by means of chemically inert microparticles. Proceedings of the Tenth International Symposium on Hazards, Prevention and Mitigation of Industrial Explosions (X ISHPMIE), Bergen, Norway, 10-14 June 2014, P. 951-958.
20. Fomin P.A., Chen J.-R. New simple method for calculation flammability limits of mixtures of flammable fuels. 12th International Symposium on Loss Prevention and Safety Promotion in the Process Industries, UK, 2007, IChemE Symposium Series No. 153, paper 104, 6 pages.

The Hydrodynamics of the WIG (Wing-In-Ground) Effect Craft

†B.C. Khoo[1], and H.B. Koe[1]

[1]Department of Mechanical Engineering, National University of Singapore, Singapore.

†Presenting and corresponding author: mpekbc@nus.edu.sg

Abstract

Due to its close proximity to the surface of water, Wing-In-Ground (WIG) crafts have much benefits from the increased aerodynamic efficiency which easily translates into fuel saving. This is besides the much higher speed possible of up to O(100) nautical miles per hour for the WIG craft which is at least twice the speed of the fastest water-surface speed boat of less than 50 nautical miles per hour. However, a WIG craft needs to be able to overcome significant hydrodynamic drag to take-off from water. The objective of this work is to investigate the hydrodynamics of a WIG craft through towed-tank test experiment as well as computational fluid dynamics simulation. From the model test, the resistance, sinkage, running trim angle and wetted area are obtained throughout the take-off speed range. Region associated with the highest resistance called hump drag is identified as well as the possibility of secondary hump and slight oscillation are discussed. Despite the complex FSI (Fluid-Structure Interaction) between the hull and water, the good comparison between experiment and simulation shows that the present state-of-the art numerical simulation is a powerful tool for WIG craft designers. An important finding is the critical presence of the stepped hull in overcoming the mentioned humped drag.

Keywords: Wing-in-Ground, Hydrodynamics, Stepped Hulls, Towing Test, CFD

Introduction

In the search of an efficient marine transportation, Wing-In-Ground (WIG) craft provides a promising solution. It is comfortable since it's flying above the sea, thus away from the wavy seas. It has a large Lift to Drag ratio, which means WIG craft has an efficient aerodynamics form due to its close proximity to the surface of the water where the ground effect takes place [1]. These means that WIG is an attractive vehicle for commercial application.

Apart from aerodynamics differences with aircraft due to ground effect [2], the hydrodynamics of the craft is another important aspect of WIG design since it needs to take-off from water. During take-off phase the hydrodynamic drag can be relatively very high compared to the aerodynamics drag, thus this often leads to high take-off thrust requirement. Initially, most of the weight is supported by water through buoyancy, as the craft speed ups the hydrodynamics drag builds up and reaches a maximum point which is called Hump Drag. It is generally occurs between 30 - 50% of the take-off speed [3]. At faster speed, the aerodynamics lift becomes significant enough to lift the craft and the hydrodynamic drag starts to reduce gradually up to the take-off point.

In order to make sure the hydrodynamics drag at high speed is acceptable, a stepped hull design is employed. The shape of the WIG craft hull is similar to those of high speed planing boat with a sudden discontinuity called step located amidships. Flying boat has been employing such design and interested reader can refer to [4]. The sudden discontinuity might induces high hydrodynamics drag, however, at high speed the flow will be separated from it and thus reducing the wetted area significantly which in turn makes the total drag is acceptable. If the step is nonexistent, as shown in [5], a planing hull shape will tend to have small trim angle at high speed which translated to significant viscous drag arising from the wetted area. As

compared to planing hull where semi-empirical method for analysis is available [6], stepped hull flow is more complex and this makes performance prediction more challenging and typically done through towing test which is costly and time consuming.

The objective of this work is to study the hydrodynamics of a WIG craft by employing both experimental method which is done via towing test and numerical method through computational fluid dynamics (CFD). The experimental method also serves as a benchmark in validating the solution that is obtained from numerical simulation, which will help to check the viability of CFD as a design tool which will reduce significantly the design cost. The value of drag, trim, sinkage and the wetted area during take-off are presented. Furthermore, a detail discussion is provided on the important hydrodynamics features such as the hump drag, possibility of secondary hump as well as porpoising instability.

Experimental Method

Model Description

A scaled model of AirFish-8 WIG craft is built and tested in the Davidson Laboratory towing tank at the Stevens Institute of Technology as shown in Fig. 1. The geometry of the model is given in Fig. 2 and the principal dimension in Table 1. For the study, only the parts which is important to the hydrodynamics of the craft are built, notably the main fuselage hull and the two floats called sponsons. The hull and sponsons are of a hard chine type which is primarily used for high speed boat and there is a step amidships of the main hull. The step divides the hull into two regions, forebody which is the part in front and afterbody which is the part behind the step. The model is made primarily from fiberglass and an aluminum structure is built to make sure there is a rigid connection between the hull and sponsons which is 0.8 m away from the symmetry plane of the hull.

Figure 1: Model Tested

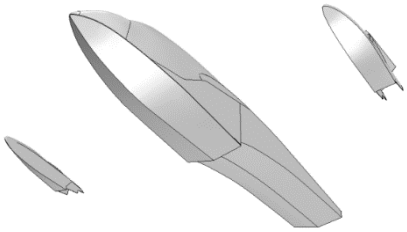


Figure 2: 3D Cad of the WIG Hull

Table 1: Model Principal Properties

Parameter	Symbol	Value
Overall Length	LOA	1.93 m
Beam	b	0.29 m
Deadrise	β	12 degree
Displacement	Δ_0	13.12 kg
Longitudinal Center of Gravity	LCG	0.90 m

Testing Procedure

The towing test is done similar to flying boat towing test as discussed in [7]. Here, the aerodynamics forces contribution is modelled in the tank. The lift is provided through the parabolic unloading method following [7] which is essentially the same as assuming a constant

lift coefficient (C_L). Aerodynamic moment coefficient (C_m) is assumed constant. Pitch damping from the tail is calculated as in [8] and is given by using a damper filled with oil and calibrated accordingly. The towing pivot is located at the LCG location and 0.26 m above the keel to simulate the propeller point of action. The models were free to trim and heave, but is restricted in the yaw, roll, surge and sway. The sinkage was measured using a motion transducer attached, trim was measured using an inclinometer mounted on the model and the drag was measured using a drag balance. Free to trim and free to heave test were carried out in the speed coefficient (C_V) ranging from 0.58 to 5.21 which covers most of the take-off speed range. The speed coefficient is defined as

$$C_V = \frac{V}{\sqrt{gb}}. \tag{1}$$

Here V is the speed and g is the gravitational acceleration. Fixed trim test were also carried out at near take-off speed. The time histories of drag, sinkage and trim are captured as well as underwater photos to determine the wetted area. Here, the hydrodynamics drag is the interest, the aerodynamics of the model is captured by running the model above the water and then deducted from the total drag as in [9].

Numerical Method

Governing Equations

The governing equations solved by the software are described by the Reynold-averaged Navier-Stokes (RANS) equations which in incompressible fluid is given below:

$$\frac{\partial}{\partial x_i}(\overline{u_i}) = 0 \tag{2}$$

$$\frac{\partial}{\partial t}(\rho\overline{u_i}) + \frac{\partial}{\partial x_j}\left(\rho\overline{u_i u_j} + \bar{p}\delta_{ij} - \mu\frac{\partial \overline{u_i}}{\partial x_j} + \rho\overline{u_i' u_j'}\right) = S_i. \tag{3}$$

Here, ρ is the density of the fluid, u_i is the fluid velocity component, p is the pressure, and S is the source term. The SST (Menter) K-ω is chosen as the turbulence model. In the simulation, both air and water are simultaneously simulated using the Volume of Fraction (VOF) approach whereby an additional convection equation is solved in the domain where the scalar solved is the volume fraction. That is,

$$\rho = \alpha\rho_{water} + (1-\alpha)\rho_{air} \tag{4}$$

$$\mu = \alpha\mu_{water} + (1-\alpha)\mu_{air} \tag{5}$$

$$\frac{\partial \alpha}{\partial t} + \frac{\partial}{\partial x_i}(\alpha u_i) = S_\alpha \tag{6}$$

where α is the volume fraction of water and $\alpha = 0.5$ is used when visualizing the free surface. HRIC (High Resolution Interface Capturing) Method is used in discretizing the VOF flux. Implicit unsteady and segregated flow technique is then employed to solve these equations. The details on how the software is employed to solve these equations are given in [10].

Mesh and Computational Domain

The domain as shown in Fig. 3 is only half since flow symmetry is assumed. The reference of $x = 0$ is located at the bow while positive x-axis means downstream, $z = 0$ is assumed to be the calm water surface position while positive z-axis means positive vertical. The size of the domain is given by $-4\ m \leq x \leq 10\ m$, $0 \leq y \leq 2.45\ m$ and $-2\ m \leq z \leq 1.5\ m$. In order to tackle the issue of large hull movement (in heave and pitch), the overset mesh methodology is utilized. In this method, the domain mesh is fixed while an additional domain to encapsulate the model will move together with the model. Additional refinement is used near z = 0 to capture the free surface accurately. Refinement of mesh in the wake region on both the main hull and sponson as well as step area is used. The mesh close to the wall is designed to make sure the value of $y+\leq 80$ is adhered to. The typical mesh used is given in Fig.4.

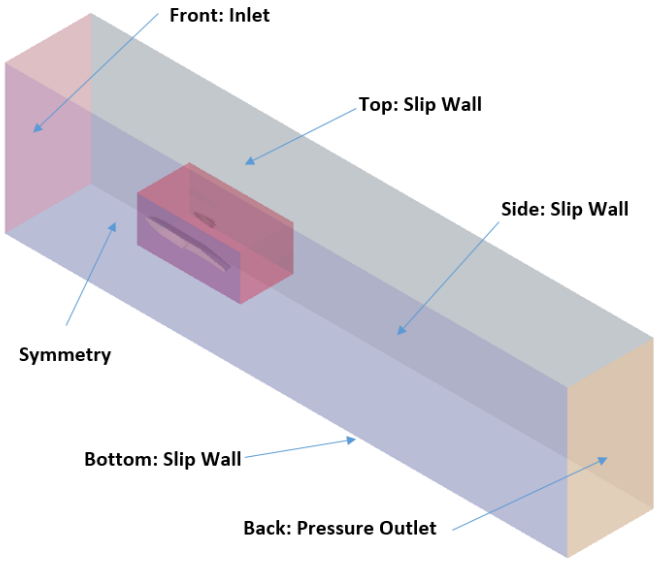


Figure 3: Simulation Domain

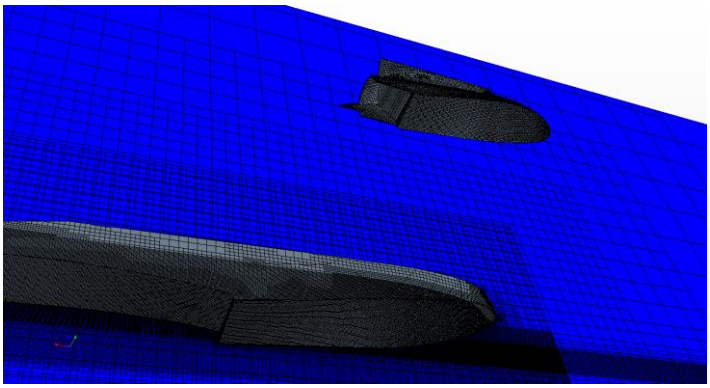


Figure 4: Mesh refinement around the hull.

In order to make sure that the solution has converged well, a mesh sensitivity study is carried to make sure all the relevant complex flow phenomena are captured. Table 2 presented the study with three different mesh configurations. The difference in the computed value of D/W (non-dimensional drag), trim angle and sinkage between them are less than 5%. The medium mesh is then chosen so that the computational cost is still acceptable.

Table 2: Mesh independence study

Mesh	No. of Cells	D/W	Trim (deg)	Sinkage (beam)
Coarse	1.1 Millions	0.172	4.40	0.081
Medium	1.8 Millions	0.171	4.44	0.081
Fine	3.6 Millions	0.175	4.54	0.082

Result and Discussion

Fig. 5 shows the hydrodynamic drag of the WIG craft that is obtained from through the free-to-trim test. The attitude of the craft in water described by the trim and sinkage are shown in Fig. 6 and Fig. 7, respectively. Here, the drag (D) is non-dimensionalized with respect to the weight (W) of the craft. It is observed that the characteristics of the drag are different from a typical high speed boat where it typically increases as the speed increased. In WIG craft the drag increases, but will reach a maximum drag called hump drag at $C_V = 1.7$. Once this speed is passed, the drag starts to reduce towards the take-off speed at $C_V = 5.8$. Here, one observes that the CFD solution is able to produce a good agreement on drag with the tow tank result both qualitatively in term of trend and quantitatively. The absolute error between CFD and tow test result has an average of 8.4%. CFD solution is able to predict the existence of hump drag at $C_V = 1.7$ despite slightly over predicting the drag value. The reduction of drag is primarily caused by the dominance of aerodynamics lift at higher speed which reduces the load on water as well as the existence of the step.

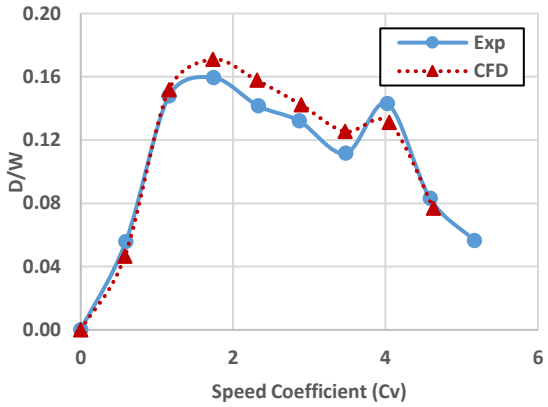


Figure 5: Drag vs. Speed Comparison.

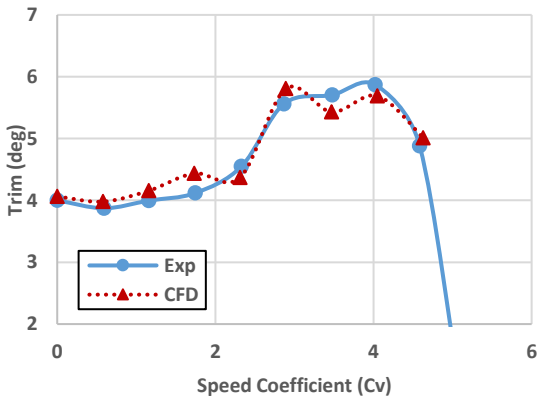


Figure 6: Trim vs. Speed Comparison.

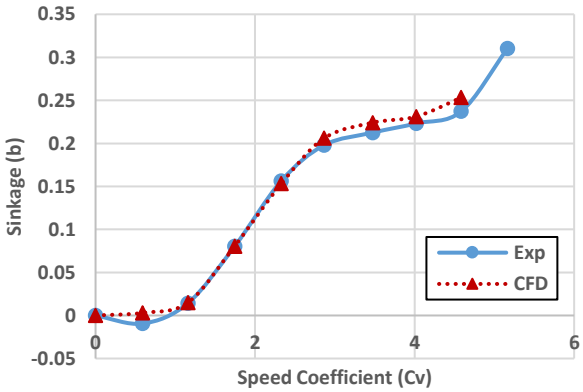


Figure 7: Sinkage vs. Speed Comparison.

Interestingly, there exists another local maximum on the drag curve at $C_V = 4$, this is called the secondary hump. This is in fact not unique to WIG craft, as can be seen from towing test done on a series flying boat done by NACA [11]. If this secondary hump is higher than the available thrust, it will prevent the craft from taking-off. One reason provided in [12] is that as the speed is getting faster, the wetted beam at the step is getting smaller than the actual beam and as a result a heavy spray escapes backward (typically called blister spray). The tangential contact of this spray with the hull (afterbody) will increase the frictional resistance significantly. CFD also shows the secondary hump despite slightly under predicting it. It was found that a proper mesh resolution is needed in order to capture the blister spray emanating from the step since the spray dimension is thin. Table 3 shows three different mesh configuration where for meshes 2 and 3 additional refinements added in the area behind the step to capture the blister spray. On Mesh 1 with the lowest resolution, there is under prediction of D/W of 22.9% and on the highest mesh resolution the error reduced to only 8.5%. Fig. 8 reveals that as the mesh is refined, the simulation seems to capture the area wetted by the blister spray better. Hence it is important for WIG craft designers to make sure that the simulation capture the blister spray correctly to make sure whether the secondary hump exist on certain design or not. Based on Mesh 3, mesh refinement of 0.5% beam around the blister spray is recommended.

Table 3: Secondary hump mesh sensitivity study.

Mesh	No. of Cells	D/W	Error (%)
1	2 Millions	0.110	22.9
2	3.5 Millions	0.127	11.4
3	11 Millions	0.131	8.5
Towing Test		0.143	

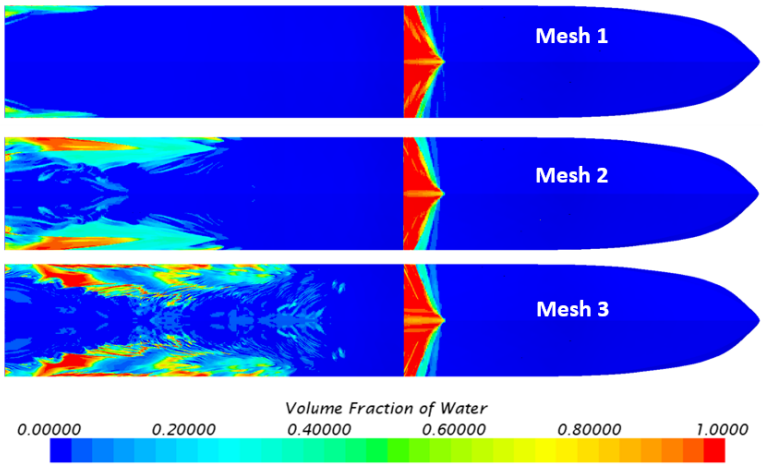


Figure 8: Volume fraction of Water on the fuselage at Cv = 4 on different mesh.

In investigating the secondary hump, the result of fix trim test done at this particular speed ($C_V = 4$) reveals the general behavior of the WIG craft hull. Fig. 9 shows the hydrodynamic drag behavior at different trim angle. The lowest drag of $D/W = 0.118$ is found when $\tau = 4.1°$ while for the free-to-trim test result in $\tau = 5.8°$ and $D/W = 0.143$, a 21% increase in drag. This means that if the WIG is to maintain the trim at the optimum trim angle, the secondary hump will not be seen. Slight oscillation is also found in both towing tank test and simulation near the secondary hump speed. Several researchers Savitsky & Morabito [5]; Garland [13] suggest that the blister spray is able to create such instability on planing hull. However, the amplitude of the oscillation of 0.5° is still acceptable [14]. Moreover, during the take-off

process this phase will pass quickly since the craft is accelerating and thus deemed to be acceptable.

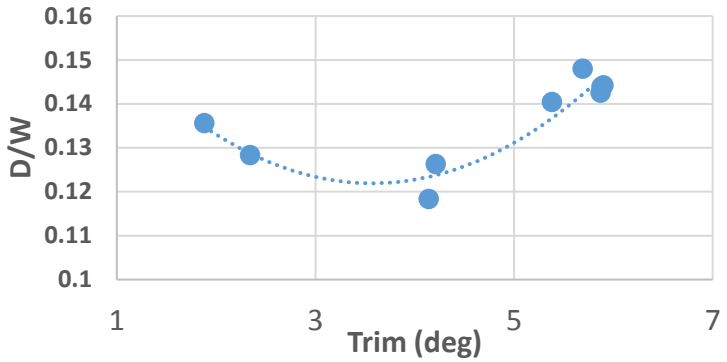


Figure 9: Drag on varying trim at Cv = 4.

The comparison on the equilibrium trim (τ) condition is given in Fig. 6 The tangent to the forebody keel at the step is used as a reference for the trim angle. Generally, there is a good agreement between the tow tank and CFD solution with an average error of only 4.3% which is 0.2° in absolute number. WIG craft trim does not vary significantly during the take-off process, unlike typical high speed boat where the trim will get smaller as the speed gets higher [5]. The trim is slowly increasing from 4° to the maximum of 6° at $C_V = 3.5$. The reduction in trim seen when $C_V \geq 4$ which is near the take-off speed is primarily due to the pitching down moment that the aerodynamics surface provided. At this speed the WIG craft can be trimmed up for taking-off. The comparison of sinkage which measures the center of gravity movement (in vertical axis) w.r.t the static condition is given in Fig. 7. Here, it is non-dimensionalized with respect to the beam (b) of the main hull. During the displacement mode (stationary) where the hull is mainly supported by buoyancy, the sinkage is small. Once the dynamic lift of the planing hull build up there is an appreciable increase in sinkage as seen when $C_V > 1$. There is also a good comparison between the numerical and experimental data with average error of 3.8%. The comparisons on trim and sinkage show that CFD is able to produce the equilibrium state accurately and is capable of simulating the complex fluid structure interaction between the fluid and the WIG hull.

The wetted area at $C_V = 1.7$ and $C_V = 4$ are given by Fig. 10 and Fig. 11, respectively. In each figure, the top half is the underwater photo taken in the tank while the bottom half is the underwater perspective obtained from CFD. At $C_V = 1.7$ which corresponds to hump speed, the step is already aerated and the wake from the forebody hit the afterbody again. As the craft speeds up, there is less weight on the water and higher pressure, thus less wetted area is needed to support the craft. In both comparisons, CFD is able to reproduce the wetted area, this means that the numerical method is able to predict the flow separation at the step, the wake generated by the forebody and the re-attachment point at the afterbody correctly. From the comparison on drag, trim, sinkage and wetted area with the towing tank, it is concluded that the present state of simulation technology is able to reproduce quite accurately the physics of WIG craft hydrodynamics. This means that CFD is a very powerful tool that designer can use when designing WIG craft before embarking on the experimental towing test which is costly and time consuming.

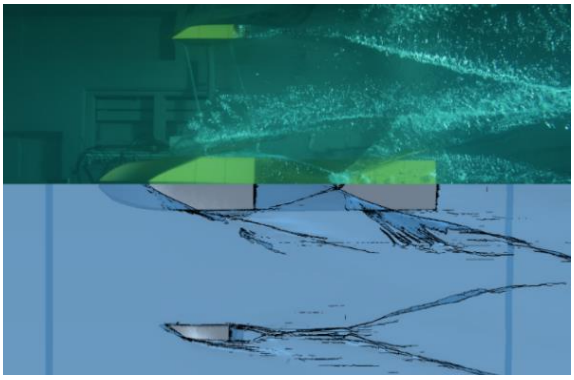


Figure 10: Wetted area comparison at Cv = 1.7. Top is tow test and bottom is CFD.

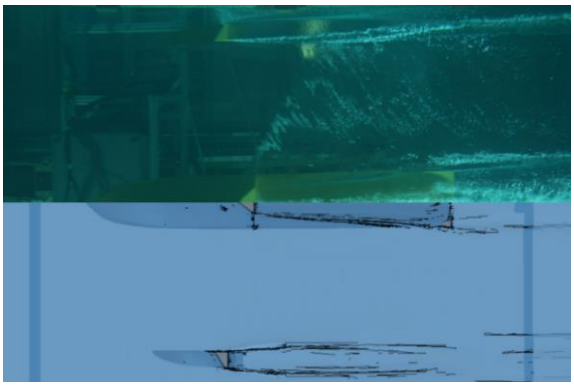


Figure 11: Wetted area comparison at Cv = 4.0. Top is tow test and bottom is CFD.

In order to gain some understanding on the use of step on WIG craft. A comparison against (hypothetical) hull without step and no sponsons as well as purely prismatic is performed here. It is assumed that this hull has the same beam (b), deadrise (β), displacement (Δ_0), LCG and the same parabolic unloading for simulating aerodynamics lift. Semi-empirical procedure based on [6] with the lift and moment coefficient obtained from [15] is used. The non-dimensional drag (D/W) is then calculated and given in Fig. 12. Without step, the drag at lower speed is lower ($C_V = 2$) but then it rises dramatically by 190% higher at $C_V = 4$. The main reason of this is the trim angle turns out to be very low ($\leq 1.5°$) and this makes the surface of the hull to be wetted and resulting in high viscous drag. On the other hand, the trim of the stepped WIG craft hull does not vary significantly as shown in Fig. 6 which results in a more optimum trim angle. Hence, despite the hull without step has lower drag at low speed, it is impractical to be used for WIG craft since high drag at high speed means that a very powerful engine is needed to take-off.

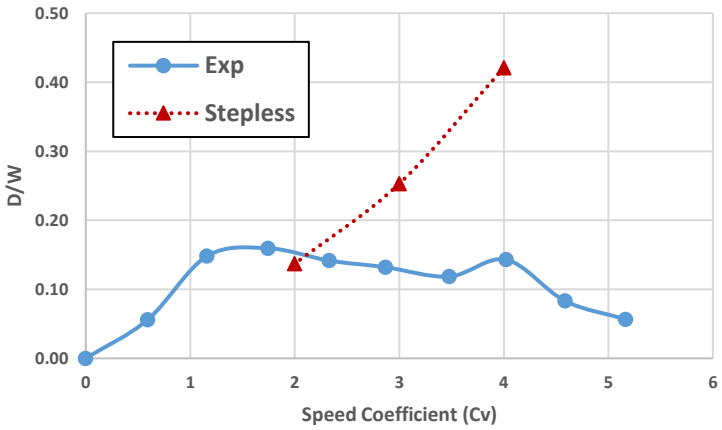


Figure 12: Comparison of Airfish-8 against a stepless hull.

Conclusion

Detailed investigation on the hydrodynamics of Wing-In-Ground craft has been performed by using towing test experiment and CFD. Through the validation of the numerical method, the current state-of-the-art CFD is a promising design tool for WIG craft designers. Several important features such as hump drag, secondary hump as well as slight oscillation are identified both experimentally and numerically. The blister spray emanating from the step is found to be the possible source of increase in drag at secondary hump. It has been shown as well that this secondary hump can be avoided when the hull is running at its optimum trim

angle. Slight oscillation was found as well near the region where secondary hump appears, albeit it is considered acceptable, the study is important during the design process in order to make sure that the WIG craft is able to take-off.

References

[1] K. V. Rozhdestvensky, "Wing-In-Ground Effect Vehicles," *Progress of Aerospace Science,* vol. 42, pp. 211-283, 2006.

[2] R. Irodov, "Criteria of the Longituinal Stability of the Ekranoplane," *Ucheniye Zapiski Tsagi,* vol. 1, no. 4, pp. 63-74, 1970.

[3] D. Ingo and L. M. de, "Hydrodynamic Characteristics of Seaplanes as Affected by Hull Shape Parameters," *AIAA Intersociety Advanced Marine Vehicle Conference,* p. 275, 1989.

[4] E. Stout, "Development of High-Speed Water-Based Aircraft," *Journal of Aeronautical Science,* vol. 17, no. 8, pp. 457-480, 1950.

[5] D. Savitsky and M. Morabito, "Surface Wave Contours Associated With The Forebody Wake of Stepped Planing Hulls," *Marine Technology,* vol. 47, no. 1, pp. 1-16, 2010.

[6] D. Savitsky, "Hydrodynamic Design of Planing hulls," *Marine Technology,* vol. 1, no. 1, pp. 71-95, 1964.

[7] H. J. Wc and W. Axt, "Hydrodynamic investigation of a series of hull models suitable for small flying boats and amphibians," Stevens Institute of Technology, New Jersey, 1951.

[8] L. Chu, "Damping Coefficients Due To Tail Surfaces," *NACA TR 136,* 1923.

[9] W. H. Jr and W. C. Axt, "Hydrodynamic Investigation of a Series of Hull Models Suitable for Small Flying Boats and Amphibians," *NACA TN 2503,* 1961.

[10] CD-Adapco, "Star-CCM+ v10.04 user Guide," 2015.

[11] J. Allison, "The Effect of The Angle of Afterbody Keel On The Water Performance of a Flying-Boat Hull Model," *NACA TN 541,* 1935.

[12] W. Sottorf, "The Design of Floats," *NACA TM 860,* 1938.

[13] R. W. Garland, "Stepped Planing Hull Investigation," *SNAME Transaction,* 2011.

[14] A. Smith and H. White, "A Review of Porpoising Instability of Seaplanes," *Royal Aircraft Establishment,* vol. 2, no. 2, 1954.

[15] P. W. Brown, "An Experimental and Theoretical Study of Planing Surfaces with Trim Flaps," SIT Report SIT-DL-71-1463, Hoboken, New Jersey, 1971.

Study of applied tissue power in microwave ablation

Tong Dong, Qun Nan*, Zhen Tian, Xiaohui Nie, Yanyan Cheng

College of Life Science and Bioengineering, Beijing University of Technology, Beijing, China, 100124

*Corresponding author: nanqun@bjut.edu.cn.

Abstract

Objective: Causing the different ablation results at the same power output or net power output, but the same applied tissue power can create the same ablation results. Calculating this value is the focus of research, which could provide accurate reference basis.

Methods: We took two methods (water specific heat capacity measurement and bio-heat transfer equation solution) to calculate the applied tissue power when power outputs were 40W, 60W and 80W. The first one is through measuring the raising temperature of water to calculate at thermal isolation environment. The other one is to calculate by bio-heat transfer equation according to the transient variation of temperature.

Results: When power outputs were 40W, 60W and 80W, the net power outputs were 39.4W, 57.5W and 74.7W by power instrument. The applied tissue powers were 32.2W, 46.6W and 58.5W by the method of specific heat capacity; by the other method, the applied tissue powers were 29.2W, 41.3W and 47.7W. Thus, we built the numerical relationship between the net power output and the applied tissue power to calculate applied tissue power efficiently. Meanwhile, this fitting equations are true by comparing applied tissue power through fitting equation with that calculated by the two methods at 70W power output.

Conclusion: Compared two fitting equations, the one by the method of bio-heat transfer equation is more suitable and has less error to calculate. And applied tissue power is an important reasonable reference to assist clinical guidance, which avoids the loss power to cause different and error results.

Keywords: The applied tissue power; Net power output; Water specific heat capacity; Bio-heat transfer equation;

Introduction

The tumor is one of the most dangerous diseases to human health and survival. According to previous studies and results, surgical resection is not the only method, and hyperthermia plays an important role in the treatment of tumor.

With the development of medical science, hyperthermia has two kinds of method, the local and whole body. And in the treatment of local body, microwave ablation(MWA) is the most great development in the field of tumor ablation. The mechanism of microwave ablation therapy takes advantage of specific conduction heating device and body's blood circulation to make the tumor internal temperature of 52℃[1] or above 60 ℃[2], and does not damage the normal tissue. This treatment aims to destroy the whole tumor and not to damage the surrounding healthy tissue. Sometimes, to make sure destruction the tumor completely, we need to have a 5mm margin of surrounding healthy tissue along the whole boundary of the tumor[3-4]. Compared with existing ablation technologies, microwave ablation will have better results consistently, larger coagulation, higher intratumoral temperatures, faster ablation times[5].

Due to the first success of treatment of small liver cancer through ultrasonic-guided percutaneous MWA[6], researchers have studied the material or structures[7] of antenna to achieve lager coagulation. In fact, they always consider the core wire material and exposed length of the radiation antenna[8], instead of ignoring the waste power of the antenna itself.

Unlike radiofrequency ablation, microwave energy is not an electrical current but rather a propagating electromagnetic field[9]. Thus, the longer time we spare to ablate the tumor, the more waste power antenna produces. In one word, the applied tissue power is the key to cause the ablation results. In the previous studies, looking for better antenna structures and materials to achieve the purpose of larger ablation zone blindly is our main object, which leads us to neglect the major factor. In addition, ignoring the applied tissue power makes confusion of the experiment data. Under the same initial conditions, the results of ablation zone are different. This could misunderstand the truth of experimental data and even influence clinical ablation effect. Therefore, grasping the applied tissue power is essential, which can consummate standardization and unification data of microwave ablation, as well as be benefit for clinical guidance and applications.

In this study, we adopt two methods to achieve our aim. Fist one is relate to water specific heat capacity, and another one is about bio-heat transfer equation. Moreover, several assumptions have been offered to use water specific heat capacity from energy transfer. So we can do experiment to measure the applied tissue power. The reason using two methods measure the applied tissue power on the different power output and heating-up time conditions is to demonstrate the accuracy of experimental results.

Methods

In previous experiments and studies, many works have ignored the crucial factor of power loss, including transmission loss and other loss, to lead some faulty studies. Thus, considering the lose energy is benefit for studies to improve the feasibility and accuracy of experimental data. What is more, transmission loss could be calculated by dynamometer. Using standard symbols, the governing equations describing the imposed energy phenomenon are given as follows:

$$p_1 = p - p' \qquad (1)$$

where p_1 is the net power output(W), p is the imposed power(W), p' is the loss power of cable(W). And the net power output is measuring by power meter without the power of cable(in Fig.1). However, the applied tissue power excludes not only the loss power of cable, but also the loss power of antenna itself, which is the whole power to be absorbed by tissue. Fig.1 shows the distribution of different powers.

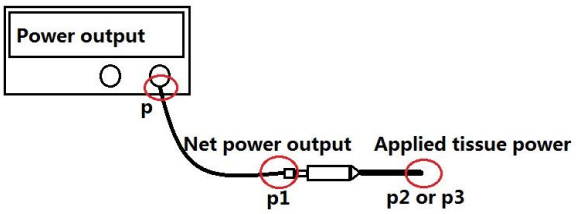


Figure1. The distribution of different powers

Water specific heat capacity measurement

The specific heat capacity is the absorbed energy of unit mass when the temperature is increasing by 1℃. This word is connected with energy and temperature which is relate to our study without other influencing factors. The power of microwave is becoming the absorbed energy per second of water to increasing the water temperature. And due to the large value about water specific heat capacity, changes of the water temperature, compared with other substances, will be small at the same conditions. So it has a great feasibility to study the applied tissue power through using water specific heat capacity. The equation describing the heat transfer phenomenon in water is given by:

$$Q_{r1} = cm\Delta T \qquad (2)$$

where Q_{r1} is the absorbed energy of water, which is the applied tissue energy(J); c is the water specific heat capacity($4.2\times10^3 J/(kg\cdot℃)$); m is the water quality(kg); ΔT is temperature difference(℃).

Among that, we should do experiments to get the value of water quality and temperature difference. We use the antenna insert water, and release energy to cause heating in the special organic glass with heat preservation. Then we applied the thermometer to observe and record temperatures at different times(before the heating starts and after the heating ends). In order to have accurate data, we apply 10 minutes[10] in the heating up time to ensure that the water temperature is diffusing equably on the different powers. Finally, though formula of specific heat capacity, we can get the results of water absorbed energy.

According to the physics knowledge, the relationship between energy and power is as follows.

$$p_2 = Q_{r1}/t \qquad (3)$$

Where Q_{r1} is the absorbed energy of water(J),t is the heating up time(s); p_2 is the applied tissue power(W).

To make sure the importance and significance of p_2 , we choose the ratio η_1 to explain the relation of applied tissue power and net power output.

$$\eta_1 = \frac{p_2}{p_1} \times 100\% \qquad (4)$$

Bio-heat transfer equation

The treatment of microwave ablation is based on the Pennes' bio-heat transfer equation[11], which has been widely used in the biological heat transfer field. It is followed by:

$$\rho C \frac{\partial T}{\partial \tau} = \nabla \cdot K \nabla T + W_b C_b (T_a - T) + Q_m + Q_r \qquad (5)$$

where ρ is the tissue density(kg / m^3); Cis tissue specific heat capacity($J / (kg \cdot ℃)$); K is the thermal conductivity of tissue($J / (m \cdot s \cdot ℃)$); T is the temperature of tissue(℃); τ is the time (s); C_b is the specific heat capacity of blood($J / (kg \cdot ℃)$); W_b is the blood perfusion rate(kg/ (m^3·s)); T_a is the blood temperature in the heating zone; Q_m is the heat energy of biological tissue($kg / (m^3 \cdot s)$); Q_r is the energy of the external heat quantity.

In the ex vivo experiment, there is no metabolism and blood flow perfusion. Besides, Sherar MD[12] considered that if the duration time of the energy emission could keep less than 10s, the influence of thermal diffusion in tissue could reduce to minimize. Because the influence of temperature changes only depends on microwave energy, and the factors of heat conduction and convection is not considered. Moreover, the thermal conductivity is ignored due to 10s heating-up time. On this condition, the equation(5) could be simplified as:

$$\rho C \frac{\partial T}{\partial \tau} = Q_r \qquad (6)$$

According to our previous studies and experiments, Q_r is equal to specific absorption rate[13]. Thus, $\partial T / \partial \tau$, which describes temperature rise curve, represents slope between temperature and time, as well as it is the crucial numerical value to achieve the purpose of calculating applied tissue power in our study. Consequently, in order to compute Q_r , we

should get the slope between temperature and time at the first ten seconds in the experiment.

In the experiment, we use some temperature needles of thermocouple to capture changes of temperatures at different points. On account of slope which indicates per-point of measuring temperature in temperature rise curve, Q_r indicates the power per unit volume. Thus we should apply integration algorithm to calculate the whole power that is the applied tissue power. Based on the volume, calculate the whole power is following as equation(7).

$$p_3 = \iiint Q_r dV \qquad (7)$$

At the same principle, we also measure the net power output from the formula(1). In statistics, the ratio between applied tissue power and net power output could predict influence of applied tissue power in experiments and emphasizes the importance as well as significance of this study.

$$\eta_2 = \frac{p_3}{p_1} \times 100\% \qquad (8)$$

According to the method of bio-heat transfer equation, we chose a special material to make a tissue phantom, which is similar to livers in thermophysical parameter[14] (Table 1). Instead of ex vivo liver, this tissue phantom could avoid the other factors to affect the ablation results

Table 1 The parameters of the tissue phantom and liver

	$\rho\,(kg/m^3)$	$C(J/(kg\cdot℃))$	$\sigma\,(s/m)$
tissue phantom	1.070	3567	0.116
liver	1.060	3600	0.148

Results

In this study, we have calculated the results of applied tissue power and net power output by above-mentioned two methods at the same initial conditions including the same power 40W, 60W and 80W respectively.

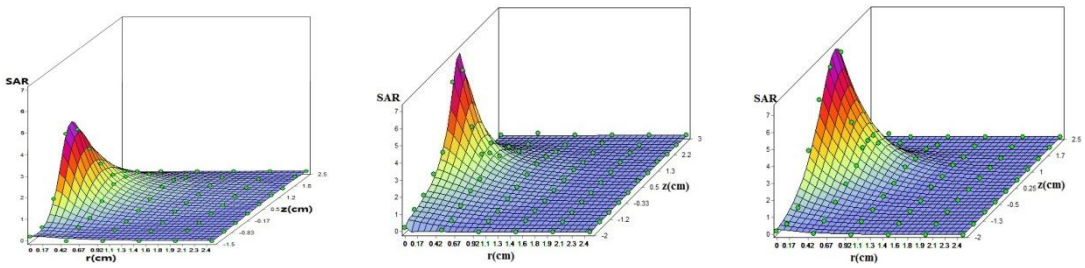
On the basis of the specific heat capacity method, the experiments were divided into three groups, including 40W, 60W and 80W. Meanwhile, the heating-up time was all set as 600s[10], which caused the homogenization of the temperature distribution to get accurate data. Besides, we have taken experiments of every group to repeat five times so as to reduce the artificial errors. When power output was 40W, the net power output was about 39.4W. Moreover, we calculated that the applied tissue power by the method of specific heat capacity was 32.2W, and the percentage η_1 was 81.73%. The detail result is as table2 below.

Table2 Net powers of different power outputs

	net power output(p_1)	applied tissue power(p_2)	the ratio(η_1)
40W	39.4W	32.2W	81.73%

60W	57.5W	46.6W	81.04%
80W	74.7W	58.5W	78.31%

According to the other method, bio-heat transfer equation, we have also calculated the net power output and applied tissue power. Based on the fitted equation and multiple integral to get the results, different distributions about power per unit volume were showed as Fig.2.



(a)The power output is 40W (b)The power output is 60W (c)The power output is 80W

Figure2. The different distributions of SAR

In Table3, when the power output was 40W, the net power output was about 39.4W, and the applied tissue power was 29.2W, the percentage η_2 is 74.11%. Besides that, we found that when we set power output as 60W, the net power output was about 57.5W, the applied tissue power was 41.3W. So the percentage to illustrate the power utilization was 71.82%. Furthermore, in case of power output 80W, the net power output and applied tissue power were about 74.7W and 47.7W, respectively. What's more, the percentage was 63.86% shown in Table 3. As we known, with the power output increasing, the net power output and applied tissue power were rising in sequence. However, it showed that the rising range of the applied tissue power was smaller and smaller with increasing the power output, this meant that the power output was higher, the power loss was more. Moreover, the change of power loss was nonlinear.

Table3 Net powers of different power outputs

	net power output(p_1)	applied tissue power(p_3)	the percentage(η_2)
40W	39.4W	29.2W	74.11%
60W	57.5W	41.3W	71.82%
80W	74.7W	47.7W	63.86%

In experiment, the net power output could be measured by power instrument. But calculating the applied tissue power is very difficult. To get the applied tissue power efficiently, building the numerical relationship between the net power output and the applied tissue power is the best way. Owing to the limitation of valid data, we have fitted the curve when the net power output was from 0W to 80W. Fig.3 is presented two curves to illustrate tendency of applied tissue power with increasing the net power output between the two methods. Based on the fitting curves, we got the fitting equation, and the fitting equation is as Equation (9) by the method of water specific heat capacity. And the other formula is showed as Equation (10) by the second method of bio-heat transfer equation.

$$y_1 = -3.428663\times10^{-5}x^3 + 2.945353\times10^{-3}x^2 + 7.544372\times10^{-1}x + 5.173639\times10^{-15} \quad (9)$$

$$y_2 = -9.550572\times10^{-5}x^3 + 7.991748\times10^{-3}x^2 + 5.745011\times10^{-1}x + 3.352874\times10^{-15} \quad (10)$$

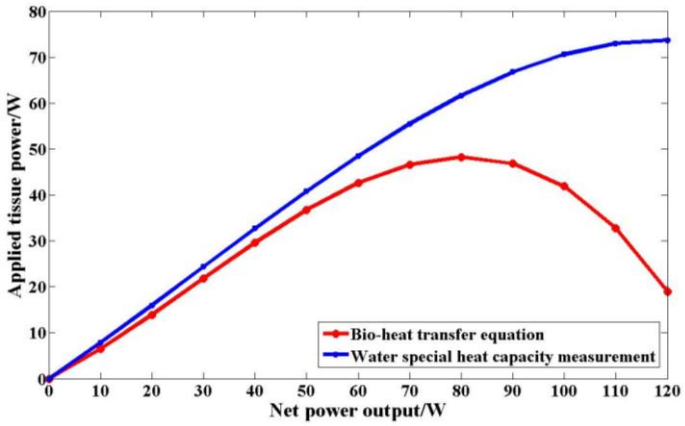


Figure3. The power curve by fitting equation

To validate the rationality of this equation, we made an experiment of 70W power output by the two methods. When the power output was 70W, the net power output was 69.1W, and the applied tissue power was 45.4W by the method of bio-heat transfer equation. Additionally, the applied tissue power was 46.3W by fitting equation of power curve. In a similar way, the net power output was 69.1W, and the applied tissue power was 53.2W by the water specific heat capacity and the data 54.88W by fitting equation of power curve.

Discussion and Conclusion

In this study, we presented the results of net power output and applied tissue power by two different methods. According to the results, the range between the applied tissue power and the net power output is changing larger and larger as the power output increasing. However, the power loss by the method of water specific heat capacity is smaller than that by the method of bio-heat transfer equation. Among that, the rates of applying the tissue power are 81.73%, 81.04% and 78.31% respectively by the method of water specific heat capacity. And the rates are 74.11%, 71.82% and 63.86% in the situation of bio-heat transfer equation method. It's observed that the gap of power conversion between the two methods is from 7.62% to 14.45%. The percentage implies that with the net power output increasing, percentage is reducing gradually, which illustrates that net power output is higher and the heat loss is more. Consequently, it’s essential to calculate the applied tissue power to help doctor to achieve the best effect of microwave ablation in clinic.

It was reported that studying the applied tissue power was calculated in the ex-vivo and in vivo experiments[10]. When the power output was 80W, the applied tissue power was 47.3W in ex-vivo experiments. In our study, the applied tissue power is 47.7W by the method of bio-heat transfer equation while the power output is 80W. But the applied tissue power is 58.5W by the other method. Compared between them, the error(0.85%) by the method of bio-heat transfer equation is much less than that(22.64%) by the method of water specific heat

capacity. Owing to thermal loss, it's inevitable to cause inaccuracy in the experiment by the method of water specific heat capacity, which leads to measure aberration of temperature. In the experiment, as the water heats up during the ablation procedure, the temperature becomes higher and higher, but the effect of heat preservation is not as good as assumption. Because it will take long time to release into the water completely and spread evenly to lead water to have a balanced temperature in microwave ablation. As a result, thermal loss and unbalanced temperature engender inaccurate data to get deviation of applied tissue power. Thus, compared with the method of water specific heat capacity, the other method is very precise to calculate the applied tissue method.

For the sake of verifying the best method to calculate the applied tissue power, we have compared the whole applied tissue powers through the two fitting curves with other study[15]. In Table4, it's presented the applied tissue powers of calculating two equations and other's study results, as well as the percentage errors between them. It can be seen that percentage errors about Eq.(10) are all smaller than those about Eq.(9). Among that, the whole percentage errors are all less than 8.76%. What is more, when the net power outputs are 45W, 57W, 64W and 71W respectively, the percentage errors are 0.91%, 2.50%, 1.11% and 2.90% in sequence. So the equation(10) could be applied to calculate the applied tissue power according to the net power output. Nevertheless, the percentage errors about Eq.(10) are 8.76% and 6.47% perceptively when the net power outputs are 37W and 75W. So the limitation is not to be neglected, and the range of application about this equation is between 37W and 75W of net power output.

Table4 The comparison of applied tissue powers

net power output in previous study	37.0W	45.0W	57.0W	64.0W	71.0W	75.0W
applied tissue power in previous study[15]	25.1W	33.0W	40.0W	45.0W	48.3W	51.0W
power by equation(9)	30.2W	36.7W	46.2W	51.3W	56.1W	58.7W
power by equation(10)	27.3W	33.3W	41.0W	44.5W	46.9W	47.7W
percentage error about Eq.(9)	20.32%	11.21%	15.50%	14.00%	13.90%	15.10%
percentage error about Eq.(10)	8.76%	0.91%	2.50%	1.11%	2.90%	6.47%

Depending on the power curve, we could achieve the aim to get the applied tissue power according to the net power output, not using complicated algorithmic method any more. Likewise, this research of the applied tissue power is important and it can reduce unsuccessful operations caused by many uncertainties in clinic, especially uncertain applied tissue power which is the biggest effect factor in microwave ablation. Owing to our data validation, the power curve by the method of bio-heat transfer equation as shown in Figure2, is suitable for calculate the applied tissue power. This fitting equation of power curve is scientific and reasonable to apply to compute applied tissue power, which is also attributed to our clinical

treatment.

Acknowledgment

This research is supported by National Science Foundation of China (No.31070754), Beijing Municipal Commission of Education Project Scientific and Technological Program (KM201410005028), the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions for three years(2013-2015)-Nan Qun, and Basic Research Foundation of Beijing University of Technology(X4015999201401).

References

[1] Wang T, Zhao G, Qiu B. Theoretical evaluation of the treatment effectiveness of a novel coaxial multi-slot antenna for conformal microwave ablation of tumors[J]. International Journal of Heat & Mass Transfer, 2015, 90:81-91.

[2] Stone H B, Dewey W C. Biologic basis and clinical potential of local-regional hyperthermia[M]//Radiation oncology. Volume 2. 1987.

[3] Tito L, Franca M, Luigi S, et al. Complications of Microwave Ablation for Liver Tumors: Results of a Multicenter Study[J]. Cardiovascular & Interventional Radiology, 2012, 35(4):868-874.

[4] Dodd GD 3rd1, Frank MS, Aribandi M, et al. Radiofrequency thermal ablation: Computer analysis of the size of the thermal injury created by overlapping ablations[J]. AJR Am J Roentgenol,2001, 177(4):777-782.

[5] Violari E G, Petre E N, Feldman D R, et al. Microwave ablation (MWA) for the treatment of a solitary, chemorefractory testicular cancer liver metastasis.[J]. Cardiovascular & Interventional Radiology, 2015, 38(2):488-493.

[6] T. Seki, M. Wakabayashi, T. Nakagawa, et al. Ultrasonically guided percutaneous microwave coagulation therapy for small hepatocellular carcinoma[J]. Cancer.1994, 74(3):817-825.

[7] Chris Hancock, Christopher I. Duff, Malcolm White. New microwave antenna structures for treating gastro-esophageal reflux disease (GERD)[J], Transactions On Microwave Theory And Techniques,2013, 61(5):2242-2243.

[8] Acikgoz H, Turer İ. A Novel Microwave Coaxial Slot Antenna for Liver Tumor Ablation[J]. Advanced Electromagnetics, 2014, 3(1):19.

[9] Erica M. Knavel, ChristopherL.Brace. Tumor Ablation: Common Modalities and General Practices[J]. Tech Vasc Interventional Rad.2013,16:192-200.

[10] Jiang Hua, Fan Weijun, Zhang Liang, et al. Net power and energy of cooled antenna microwave ablation: ex vivo versus in vivo results in porcine liver[J]. Natl Med J China,2012,92(35):2513-2517.

[11] H. H. Pennes. Analysis of tissue and arterial blood temperatures in the resting human forearm.Applied Physiology. 1948, 1(2):93-122.

[12] Sherar M D, Gladman A S, Davidson S R H, et al. Infrared thermographic SAR measurements of interstitial hyperthermia applicators: errors due to thermal conduction and convection.[J]. International Journal of Hyperthermia the Official Journal of European Society for Hyperthermic Oncology North American Hyperthermia Group, 2004, 20(5):539-55.

[13] Zafar T, Zafar J, Zafar H. Development and microwave analysis of slot antennas for localized hyperthermia treatment of hepatocellular liver tumor.[J]. Australasian Physical & Engineering Sciences in Medicine, 2014, 37(4):1-7.

[14] Zhang Bulin, Hu Bing, Kuang Shengli. A tissue-mimicking phantom for visualizing coagulation temperature distribution at the same range of biological tissue during thermal ablation[J]. C hin J M ed Imagin g T echnol, 2007,23(9):1405-1408.
[15] Tang Tian. The experimental study of microwave ablation liver tissue for the effective damaged power and the thermal deposition dose of tissue[D]. Zhongshan University,2010.

A computational study of the movement of an object driven by the centrifugal pump in the pipeline based on overset meshes

†Jie Chen[1], *JingYin Li[1], Le Wang[2], and JunWei Su[3]

[1]School of Energy And Power Engineering, Xi'an Jiaotong University, China.
[2]School of Civil Engineering, Southwest Jiaotong University, China.
[3]School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, China.

*Presenting author: vrnetgod@qq.com
†Corresponding author: jyli@mail.xjtu.edu.cn

Abstract

The overset mesh method was used for numerical calculation of the movement process of a moving object in a pipe driven by a pump, the appropriate set solving parameters, such as mesh scale, iteration, time step, etc., were determined by reducing the fluctuation of the force applied to the object, the movement process of the object in the pipe from stillness to linearly increased revolution speed was explored, the movement process and movement law of the moving objects with different diameters were analyzed, the change in flow field of each moving object at the gap between it and the pipe wall was compared, and the effect of the gap on the fluid force on the object was analyzed in this paper. The simulation results showed that the forces on and movement fluctuations of moving objects at different gaps were different, and the gap sizes also significantly affected the fluctuations of fluid forces on the moving objects.

Keywords: transient flow; CFD; overset; centrifugal pump

Introduction

The transient operation of a centrifugal pump exists in many occasions. For example, underwater weapon launching is a typical engineering application of the start process of a centrifugal pump. Such process requires the lift and flow of the pump to reach the given indices within a short time. The pump always is starting until it stably works. This start process is significantly different from the stable operation process of this pump. When this centrifugal pump is sued as an underwater launcher, this launching process is equivalent to the movement process of an object in a pipe driven by it. The problem about the transient performance of this centrifugal pump which drives the object in this pipe is different from the previous studies on the start characteristics of a single pump and the flow in the pipe. This problem leads to correlation between start process of this pump and the movement process of the object. Thus, it is necessary to further study the numerical calculation of this transient performance. The proposal of an appropriate numerical calculation method is very important to predict this dynamic process by the start characteristics of this centrifugal pump and numerical calculation of the object movement process. The study result will provide a reference for design and application of the start process of a centrifugal pump.

The study on the transient characteristics during transient start of a centrifugal pump goes all the way back to relevant researchers' theoretical analysis and numerical calculation of water hammer phenomenon in the early 20th century. Allievi[1] proposed a method of solving the water hammer equation. Schnyder and Bergeron[2] proposed a graphic method for analysis of the water hammer phenomenon, respectively. Wylie and Streeter[3] systematically discussed the transient flow mechanism and proposed the theoretical model of solving several transient processes of hydraulic machinery. Akimoto. Narumi[4] defined the physical concepts of water hammer and pressure fluctuation and systematically described the calculation method and the method for prevention of water hammer. In addition, foreign researchers did a lot of experimental studies. For example, Tsukamoto and Tanaka[5]-[7] experimentally studied various transient processes of a centrifugal pump, and used high speed photography to capture

the cavitation damage in the transient processes. Domestic researchers did a lot of studies on the transient start characteristics of a pump. Wu Dazhuan[8] used MRF method for numerical simulation of transient start process of one centrifugal pump to obtain the development process of flow field in the transient start process. Wang Leqin et al [9][10] used power bond graph method to build a mathematical model for start process of a vane pump, and proved the applicability of the power bond graph method by comparing with the experimental results. Li Zhifeng et al[11] did a lot of experimental studies on the various transient processes of a vane pump and used the moving mesh method and large eddy simulation technique for unsteady numerical simulation of the start process of a centrifugal pump; their studies showed that one transient pressure shock phenomenon existed in the pump during quick start, and the lift increased with increasing revolution speed; when the revolution speed reached a maximum value and was steady, the transient performance became steady state performance, and the flow increment showed a lag effect in the transient start process. Hu Fangfang et al. [12] used the moving mesh method for numerical simulation in the transient start process of a centrifugal pump, and their simulation results showed that a large eddy zone, which existed in the impeller channel during start, gradually shifted from impeller outlet to the middle of the impeller channel in the start process.

Relevant studies showed that the study on the transient characteristics of a pump by sliding mesh method is relatively mature. The movement of an object in a pipe is usually studied by sliding mesh method, which requires call by value through an interface between pipe wall zone and object zone, and use of dynamic layering method to split and merge the mesh layer at boundary between moving zone and stationary zone according to the change in height of the moving mesh at the boundary. To ensure the calculation accuracy, sliding mesh method requires mesh densification in the pipe wall zone and object zone. A very small gap between the pipe wall and object will greatly increase the number of meshes. In addition, the mesh splitting and merging at the boundary also increase the additional calculation load. The overset mesh method is an important mesh method to process the relative movement of an object [13][14]. In this method, computational mesh is split into many overset or overset submesh. When an object is moving, the body-fitted mesh moves. The numerical calculation is performed on each submesh, and information transfer of the flow field is achieved by interpolation on the overset submesh. It should be noted that the sliding mesh method [15][16] is a special overset mesh method with minimum overset area. The main advantage of the overset mesh method is to reduce the difficulty in mesh generation, enhance the mesh generation flexibility, guarantee initial mesh quality, and inherit the initial solver better. A mesh zoning strategy, the overset mesh method is very favorable for parallel calculation. When the overset zone is moving, the interface for the sliding mesh method need not be artificially defined and the mesh at the interface need be split and merged by dynamic model. Thus, the number of meshes can be reduced and the mesh splitting and merging time can be saved when the overset mesh is used for the object movement in a pipe. However, the overset mesh quality greatly influences the value solving accuracy and convergence rate. The author of the literature [17] proposed the advice on meshblock and hole boundary position, overset zone size, and mesh quality at the overset zone, An overset mesh method suitable for the object movement in a pipe was studied in consideration of relevant advice in this paper.

Calculation Assumption and Physical Model

In the simulation calculation, it is assumed that the revolution speed of a pump is linearly increased to 960 rpm within 5 s, i.e. transient speed, N=960/5*t, and the effect of the gravity on water and moving object is negligible.

Continuity equation:

$$\nabla \cdot \vec{u} = 0 \tag{1.}$$

Momentum equation:

$$\rho \frac{\partial u}{\partial t} + \rho div(u\vec{u}) = \mu \nabla^2 u - \nabla p \tag{2.}$$

$$\rho \frac{\partial v}{\partial t} + \rho div(v\vec{u}) = \mu \nabla^2 v - \nabla p \tag{3.}$$

$$\rho\frac{\partial w}{\partial t}+\rho div(w\vec{u})=\mu\nabla^2 w-\nabla p \tag{4.}$$

Turbulence model: The turbulence model used in this paper is $k-\omega$ SST model. $k-\omega$ model was initially proposed by Wilcdx[19], and its advantage is the processing of near wall in calculation at low Reynolds number. It does not the complicated non linear damping function required in the $k-\varepsilon$ model, and therefore is more accurate and steadier than $k-\varepsilon$ model.
k and ω meet the following transport equations:

$$\frac{D(\rho k)}{Dt}=P_k-\beta^*\rho\omega k+\frac{\partial}{\partial x_j}[(\mu+\sigma^*\mu_T)\frac{\partial k}{\partial x_j}] \tag{5.}$$

$$\frac{D(\rho\omega)}{Dt}=\frac{\gamma\omega P_k}{k}-\beta\rho\omega^2+\frac{\partial}{\partial x_j}[(\mu+\sigma\mu_T)\frac{\partial\omega}{\partial x_j}] \tag{6.}$$

$$\mu_T=\gamma^*\frac{\rho k}{\omega} \tag{7.}$$

Menter believed that the initial $k-\varepsilon$ model was very sensitive to free flow conditions [20]. The calculation results from the model are very different at different ω values at the inlet. To solve this problem, Menter developed a method: $k-\omega$ model was used for the near wall; $k-\varepsilon$ model was used for the outside of the boundary layer, and both the turbulence models are used for the inside of the boundary layer. The $k-\omega$ SST model equation is as follows:

$$\frac{D(\rho k)}{Dt}=P_k-\beta^*\rho\omega k+\frac{\partial}{\partial x_j}[(\mu+\sigma_k\mu_T)\frac{\partial k}{\partial x_j}] \tag{8.}$$

$$\frac{D(\rho\omega)}{Dt}=\frac{\gamma\rho P_k}{\mu_T}-\beta\rho\omega^2+\frac{\partial}{\partial x_j}[(\mu+\sigma\mu_T)\frac{\partial\omega}{\partial x_j}]+2(1-F_1)\frac{\rho\sigma_{\omega 2}}{\omega}\frac{\partial k}{\partial x_j}\frac{\partial\omega}{\partial x_j} \tag{9.}$$

$$\mu_T=\frac{a_1\rho k}{\max(a_1\omega,F_2\Omega)} \tag{10.}$$

Where $P_k=\tau_{ij}\frac{\partial u_i}{\partial x_j}$ is turbulence kinetic energy production term. For all the parameters in this equation, see the above literature.

Calculation Model

Geometry Model and Boundary Conditions

IS 65-40-250 centrifugal pump in chemical industry was used in calculation. The appearance dimension of the moving object was shown in Figure 1.The objects had the diameter, D, of 80 mm, 90 mm and 100 mm and were expressed as Object-A, Object-B, and Object-C, respectively. One object was installed in transparent pipe with inner diameter of 120 mm. The pump had two pipes with the inner diameter of 32 mm at its ends. One end is connected to a water tank (During calculation, the change in water level was negligible, and the boundary conditions of a stagnation inlet were used), and the other end was transitionally connected to the transparent pipe with inner diameter of 120 mm. The arrangement and boundary conditions were shown in Figure 2.

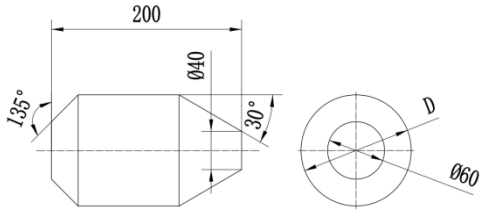


Figure 1. Appearance Dimension of Moving Object

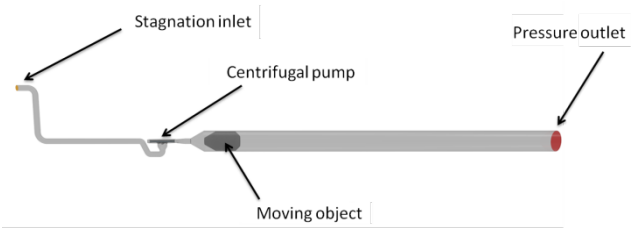
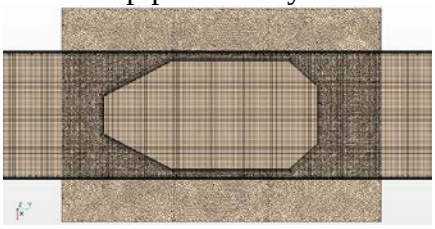


Figure 2. Piping Figure 1. A plate not for meal

Meshing

A computed field mainly is divided into four zones, pump inlet pipe zone, rotary zone, pump outlet pipe zone and moving object zone in the sequence of water flow. In this paper, Trimmer mesh method was used to give pipe zone, rotary zone, and pump outlet pipe zone during numerical calculation. The polyhedral mesh method was used to give a motion zone. Different mesh sizes are used. The calculated forces on the objects were compared. It was found that the forces on the objects fluctuated very sharply, and the force fluctuation was not significantly improved after mesh densification and reduction of time step. Finally, the mesh size was set to 0.002 m considering that a lot of time was spent on calculation. The meshes at each zone generated the boundary layer meshes (Figure 3). Two pipe zones were connected to the rotary zone including pump vane by the interface, while the independent zone including the moving object was connected to the pipe zones by Overset Mesh Interface.



A) Uninitialized overset mesh

B) Initialized overset mesh

Figure 3. Mesh in Computed Field (Object-C)

In the overset mesh method, the calculation was performed by mesh overlapping. This method has such as significant advantage that it can solve the problems of very complicated relative movement at some zones, solid contact in space, etc. However, these problems always are impossibly solved by discontinuous mesh method. The overset mesh mechanism is calculation information transfer between two different zones by overset meshes. Its principle is as follows: In the case of calculation by the overset mesh method, the meshes between two zones (master zone and slave zone) are overset. Based on the interpolation information of the slave zone, the overset zone will re-generate combined overset mesh according to the connection information of the point of the adjacent mesh outside the overset zone as well as the interpolation of the point of the mesh at the overset zone edge and master zone after the calculation begins. In the calculation of the

combined overset mesh, the point of the adjacent point outside the overset zone and the point at the overset zone edge are regarded as the boundary points. The values are transferred and exchanged between the mesh at the slave zone and the mesh at the master zone by interpolation. [18] The operation procedure of the overset mesh method mainly consists of five steps: (1) two or more computed fields are created: the background zone covers the flow zone of the far field, and the independent zone surrounds the moving object of concern (overset bodies); (2) all the zones are independently meshed; here would be an overset mesh zone, where the meshes from different zones are overset at one place; (3) the outer boundary of the overset body (overset body) is set to "Overset Mesh", and "Overset Mesh Interface" is built in the two computed fields; and (4) with moving overset body in the background zone, the overset zone will change; and (5) The information is transferred in two zones by the overset mesh.

Result and Analysis

During calculation, the time step is 0.005s, the six degrees of freedom solver has the number of solving steps of 15 and maximum number of internal circulation steps of 40, and the residual error is set to less than 0.001 in the condition of convergence. A double core (Intel Xeon E5-2680 v4 14C 2.40GHz) workstation is used for calculation, 18 threads parallel calculations of 5s movement takes about 20 days.

The forces on Object-A, Object-B, and Object-C applied by the fluid in a pipe are shown in Figure 4. With decreasing gap between the object and pipe wall, the forces on the moving objects fluctuated very sharply. With increasing revolution speed of the pump, the force fluctuation increases. The fluid forces on three moving objects are in the reverse direction of water flow (positive direction of y-axis) and continuously increase at the time between 2.2s and 2.7s and between 2.7s and 3.2s. The fluid force on Object-A and Object-B sharply reduces at the time of 3.5s and 3.7s, respectively, but that on Object-C does not sharply reduce.

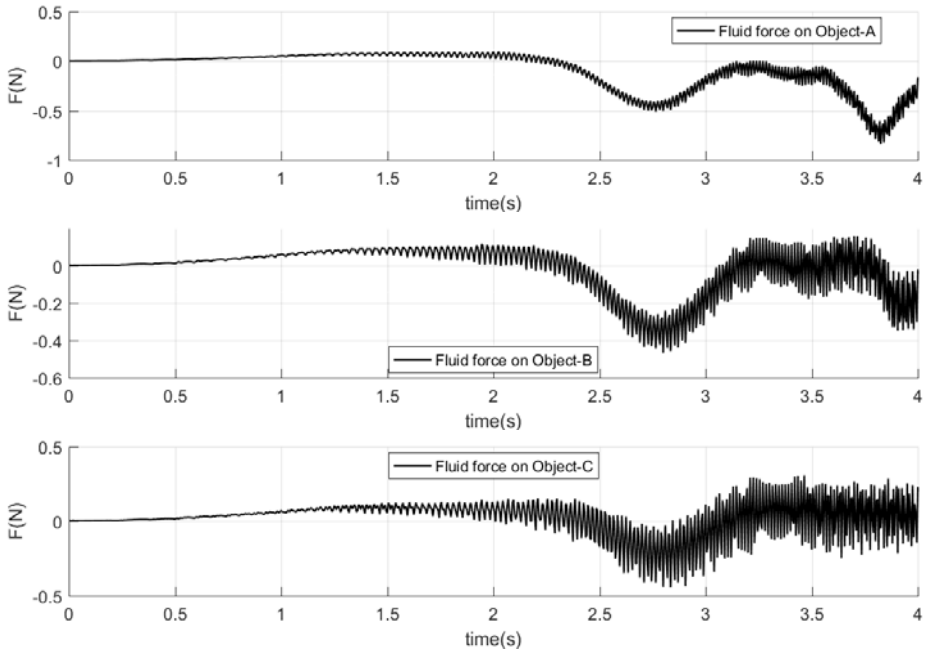


Figure 4. Calculated Force

For different moving objects in the pipe, the fluid force on the pump vane (Figure 5) and torque (Figure 6) change similarly over time, indicating that the loads on the pump vane are not significant different when the pump vane drives the moving objects with different diameters to move mainly because the objects do not significantly affect the flow resistance in the pipe. The change in mass flow (Figure 7) also accounts for this phenomenon.

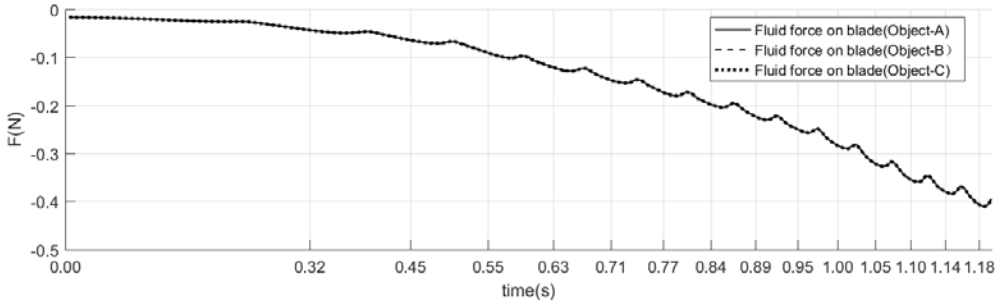


Figure 5. Fluid force on vane

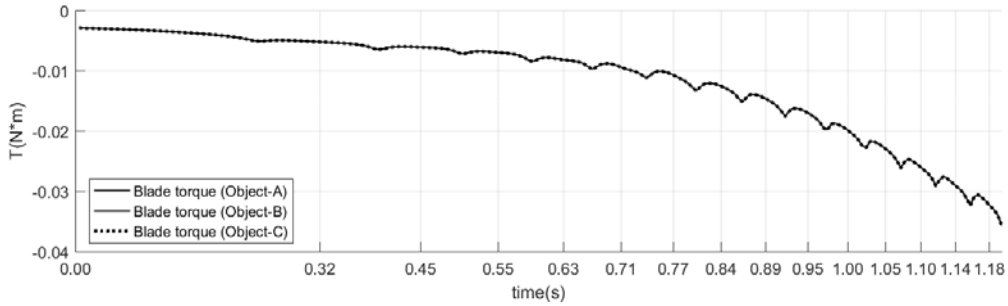


Figure 6. Vane torque

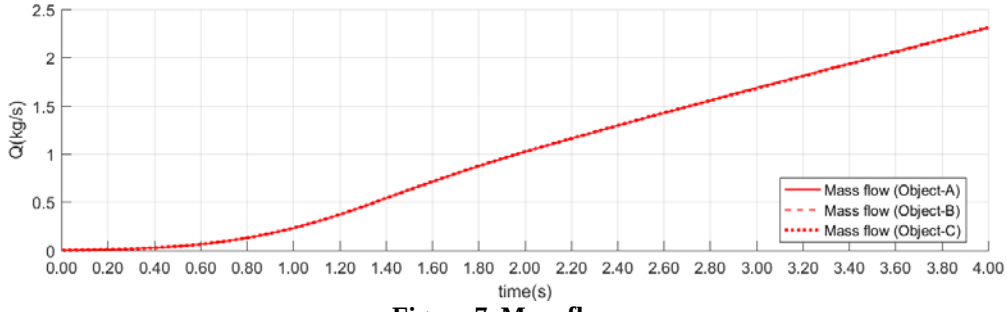


Figure 7. Mass flow

It was found by comparing the fluctuations of the fluid forces on Object-A, Object-B and Object-C (Figure 8) that the force fluctuation closely related to pump revolution; the force fluctuation had the same main frequency as the vane; with decreasing gap between the moving object and pipe wall, the force fluctuated more sharply because the differential pressure on the moving object dominated and the pressure fluctuation caused by pump revolution could be transferred by the fluid to the downstream.

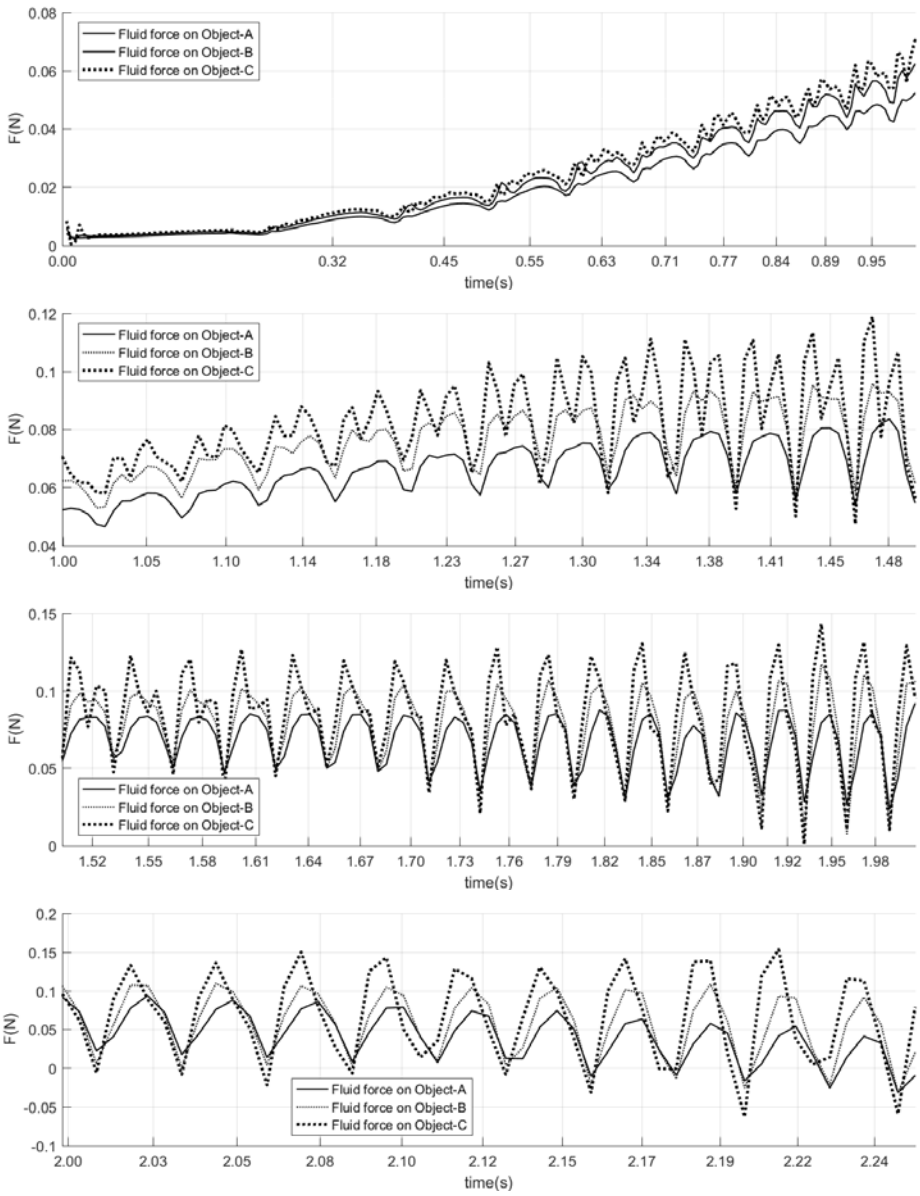


Figure 8. Fluid Force Fluctuation (Time scale as the inverse number of vane frequency)

It could be seen from comparison of fluid forces on all the objects that a law of force fluctuation existed, but the high pressure fluid from the pump flowed away from a big gap without force on the object (Figure 9), so the forces on Object-A, Object-B and Object-C were significant different and finally affected their movement processes as shown in Figure 10 and Figure 11. it could be seen that Object-C always moved to the outlet, whereas Object-A and Object-B moved in the reverse direction because the high pressure fluid flowed through the gap before a high pressure first and then a low pressure, which stopped the forward movement of the object, occurred in the front of the object (Figure 12).

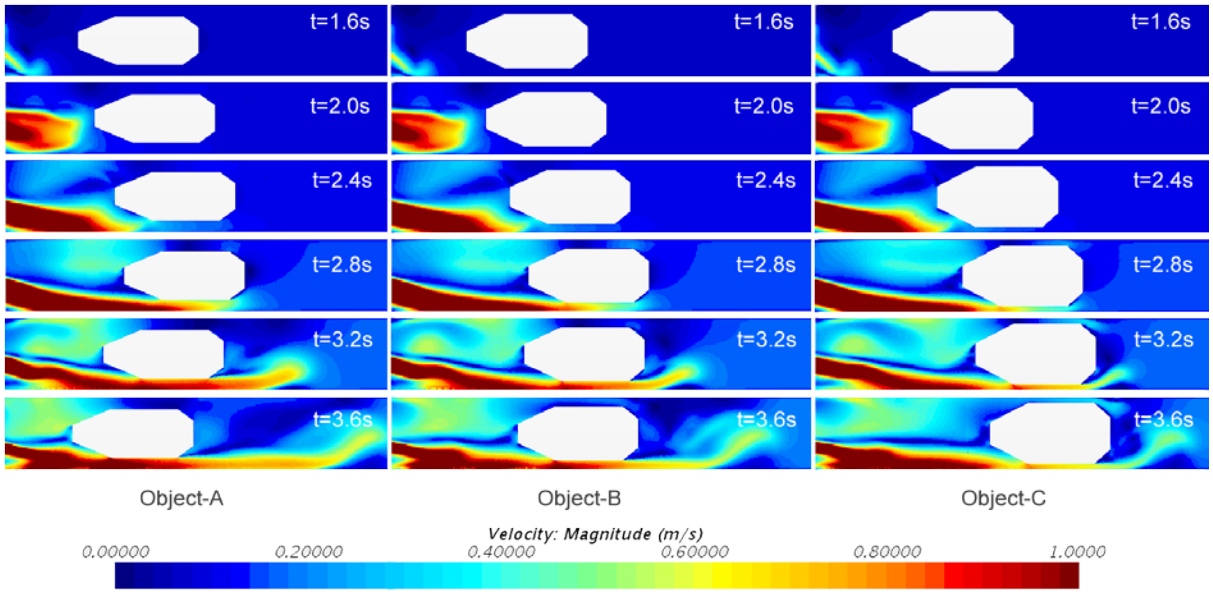


Figure 9. Velocity Distribution Cloud (z=0 m)

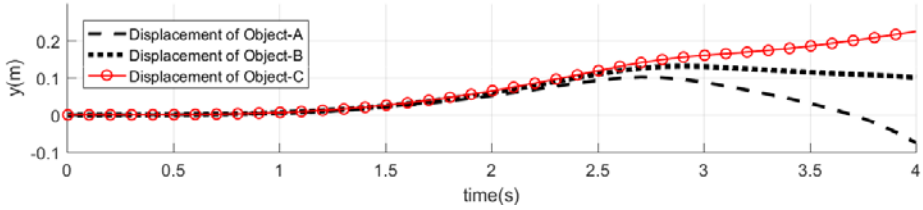


Figure 10 Moving object displacement

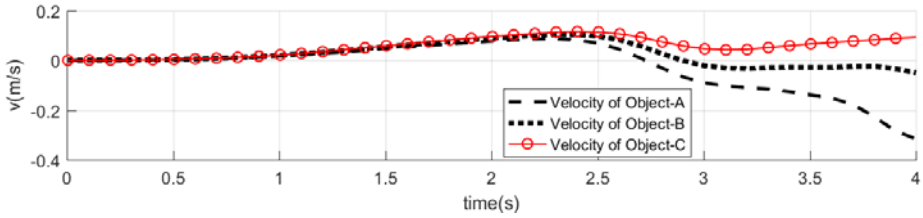


Figure 11 Velocity

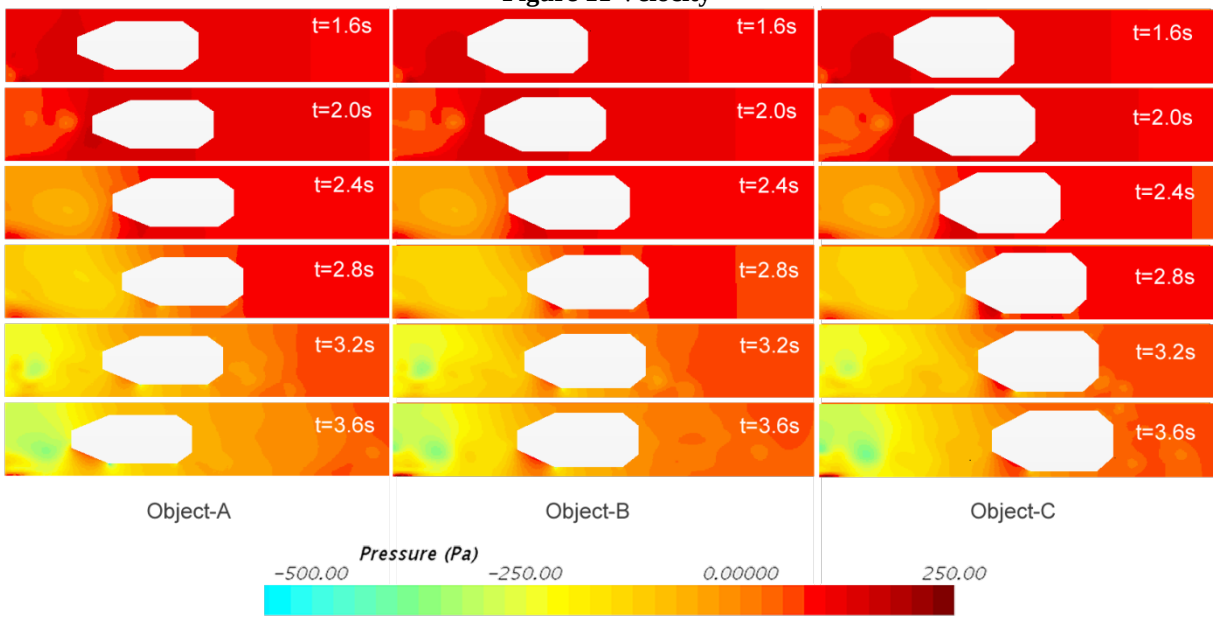


Figure 12. Pressure Distribution Cloud (z=0 m)

The pump revolution would lead to pressure fluctuation of the water flow at the pump outlet; with increasing revolution speed of the pump, both fluctuation frequency and magnitude increased,

resulting in increasing force fluctuation on the object. However, the force on the object fluctuated more greatly when the gap was reduced, especially when Object-B and Object-C moved more slowly than Object-A (Fig. 7), maybe because the limited calculation capacity of the overset mesh method, which required mesh densification at the gap according to the gap dimension or improved the fluid force by further reducing the time step, leaded to impossible numerical calculation on a smaller mesh scale in this paper.

Conclusions

The overset mesh method was used for numerical calculation of the movement process of a moving object in a pipe driven by a pump, the fluctuation of the fluid force on the moving object is minimized by relevant settings, and the movement processes and laws of the moving objects with different diameters were analyzed in this paper. The following conclusions are given:

1. The fluid forces on the moving object have the same fluctuation frequency as the vane;
2. The vane had the same fluid forces and torques when the pump was driving different moving objects;
3. Object-A, Object-B, and Object-C moved in a law, but differently. In particular, Object-B and Object-C moved in the reverse direction mainly because of a pressure difference between before and after the object. This maybe did not match with the actual situation maybe because the time step used for calculation in this paper did not meet the calculation of the flow field at high vane frequency. Thus, the calculation will be improved by further mesh densification or time step reduction if the calculation condition permits.

References

[1] Allievi.The theory of water hammer.English Translation by Halmos.ASME, New York, 1925.
[2] Bergeron.Study on the steady Fluctuation in Water-filler conduits.General Graphical Solution.Revue General de L Hydraulique, 1935, 1(1).
[3] Wylie, V.L.Streeter.Fluid transients[M]. New York: McGraw-Hill, 1978
[4] Akimoto.Narumi. *Water Hammer and Pressure Fluctuation*[M]. Beijing: China Power Press, 1981.
[5] Tanaka T, Tsukamoto H.Transient Behavior of a Cavitating Centrifugal Pump at Rapid Change in Operating Conditions—Part 1: Transient Phenomena at Opening/Closure of Discharge Valve[J].Nihon Kikai Gakkai Ronbunshu B Hen/transactions of the Japan Society of Mechanical Engineers Part B, 1999, 121(4).
[6] Tanaka T, Tsukamoto H, Tanaka T, et al.Transient Behavior of a Cavitating Centrifugal Pump at Rapid Change in Operating Conditions—Part 2: Transient Phenomena at Pump Startup/Shutdown[J].Journal of Fluids Engineering, 1999, 121(4): 850-856.
[7] Tanaka T, Tsukamoto H.Transient Behavior of a Cavitating Centrifugal Pump at Rapid Change in Operating Conditions—Part 3: Classifications of Transient Phenomena[J].Nihon Kikai Gakkai Ronbunshu B Hen/transactions of the Japan Society of Mechanical Engineers Part B, 1999, 121(4): 857-865.
[8] Wu Dazhuan. THE STUDY ON TRANSIENT CHARACTERISTICS OF SPECIAL CENTRIFUGAL PUMP DURING QUICK START PROCESS, PhD Thesis, Hangzhou: *ZHEJIANG UNIVERSITY*, 2004.
[9] Wang Leqin, Wu Dazhuan, Hu Zhengyu et al. NUMERICAL SIMULATION ON STARTING CHARACTERISTIC OF VANE PUMP BY BOND GRAPH[J]. *JOURNAL OF ENGINEERING THERMOPHYSICS*, 2004, 25(3): 417-420.
[10] Wu Dazhuan, Wang Leqin, Hu Zhengyu. Numerical simulation of centrifugal pump's transient performance during rapid starting period[J]. *JOURNAL OF ZHEJIANG UNIVERSITY* (ENGINEERING SCIENCE), 2005, 39(9): 1427-1430.
[11] Li Zhihui, Wu Dazhuan, Dau Weiping et al. *2-D numerical simulation on transient flow in centrifugal pump during starting period*[J] *JOURNAL OF ENGINEERING THERMOPHYSICS*.2007.
[12] Hu F F, Ma X D, Wu D Z, et al. Transient internal characteristic study of a centrifugal pump during startup process [C]. //26th IAHR Symposium on Hydraulic Machinery and Systems, Beijing, China, August 19-23, 2012: IAHRXXVI-103.
[13] Steger JL, Dougherty FC, Benek JA. A chimera mesh scheme. Mini Symposium on Advances in Mesh, Generation, ASME, 1982
[14] Juvigny X, Canonne E, Benoit C. Multimesh algorithms for the chimera method. AIAA 2004-758, 2004
[15] Henshaw WD. Automatic mesh generation. Acta Nummerica, 1996, 5. 121 148

[16] Zhang Laiping, Deng Xiaogang, Zhang Haixin. Reviews of Moving Mesh Generation Techniques and Numerical Methods for Unsteady Flow[J]. *Advances in Mechanics*, 2010, 40(4): 424-447
[17] Shih T I-P. Overset meshes: fundamentals and practical issues. AIAA 2002-3259, 2002
[18] Koomullil R, Soni B, Singh R. A comprehensive generalized mesh system for CFD applications[J].Mathe and Computers in Simulation, 2008, 78: 605-617.
[19] Wilcdx D. Turbulence modeling for CFD[J].La Canada, CA:DCW industries,Inc,1993
[20] Menter F. Two-equation eddy-viscosity turbulence models for engineering applications[J].AIAA journal,1994,32(8):1598-1605.

Transient thermal response of a functionally graded piezoelectric laminate with a crack normal to the bimaterial interface

Y.Nakaue[1] and S. Ueda[2]

[1, 2] Department of Mechanical Engineering, Osaka Institute of Technology,
5-16-1 Omiya, Asahi-ku, Osaka 535-8585, Japan

[1] Presenting author: n.yoshihisa.oit@gmail.com
[2] Corresponding author: sei.ueda@oit.ac.jp

Abstract

In this paper, the fracture problem of a functionally graded piezoelectric material strip (FGPM strip) containing a crack perpendicular to the interface between an FGPM strip and a homogeneous layer is considered. The problem is solved for the laminate that is suddenly heated from the surface of the FGPM strip. The surface of the homogeneous layer is maintained at the initial temperature. The crack faces are supposed to be completely insulated. Material properties are assumed to be exponentially dependent on the distance from the interface. By using the Laplace and Fourier transforms, the thermo-electro-mechanical fracture problem is reduced to a singular integral equation, which is solved numerically. The stress intensity factors are computed and presented as a function of the normalized time, the nonhomogeneous and geometric parameters.

Keywords: functionally graded piezoelectric material, fracture mechanics, stress intensity factor, elasticity

Introduction

The concept of the well-known functionally graded materials (FGMs) can be extended to the piezoelectric material to improve its reliability [1]. As a result, it is also important to investigate the fracture behavior of functionally graded piezoelectric materials (FGPMs) under thermal load, and some interesting results have been reported. For example, Wang and Noda [2] treated the thermally induced fracture of a smart functionally graded composite structure. The present author has investigated the singular fields around cracks in FGPMs under the static thermal loading condition [3-5] and under the thermal shock loading condition [6-8]. It was found that by selecting the material constants appropriately, the steady stress and electric displacement intensity factors can be lowered substantially. Moreover, the overshooting phenomenon of the stress and electric displacement intensity factors was observed in a homogeneous piezoelectric strip [9, 10] and in an FGPM strip [6-8].

On the other hand, piezoelectric composites have been used in a wide variety of applications including vibration control and actuators. These systems can be achieved by incorporating a thin piezoelectric layer into a structural system. Several kinds of piezoelectric actuators have been designed. Uchino et al. [11] fabricated a monomorph actuator made from semi-conductive piezoelectric ceramics.

In this paper, we focused on the transient thermal fracture problem of monomorph actuators using an FGPM strip. The analytical model of the monomorph actuator consists of an FGPM strip and a homogeneous elastic layer. The problem of the normal crack in the FGPM strip is analyzed under transient thermal loading conditions. Material properties are exponentially dependent on the distance from the interface between the FGPM strip and the homogeneous

elastic layer. The superposition technique is used to solve the governing equations. The transient temperature and transient thermal stress in an un-cracked FGPM strip are calculated by the Laplace transform, and a numerical method is employed to obtain time-dependent solutions by way of a Laplace inversion technique [12]. The obtained thermal stress is used as the crack surface tractions with opposite sign to formulate the mixed boundary value problem. By using the Fourier transform techniques [13, 14], the electromechanical problem is reduced to a singular integral equation, which are solved numerically [15]. The stress intensity factors are computed and presented as functions of the normalized time, the nonhomogeneous and geometric parameters.

Formulation of the problem

Consider a strip of FGPM of thickness h_1 containing a finite crack bonded to an elastic layer of thickness h_2 with the rectangular Cartesian coordinate system (x, y, z) as shown in Figure 1. The crack of length $2c$ is located along z-axis from a to b $(b-a=2c, \quad 0<a<b<h_1)$. The piezoelectric strip is poled in the z-direction and is in the plane strain conditions perpendicular to the y-axis. It is assumed that initially the medium is at the uniform temperature T_1 and is suddenly subjected to a uniform temperature rise $T_0H(t)$ along the boundary $z=h_1$, where $H(t)$ is the Heaviside step function and t denotes time. The temperature along the boundary $z=-h_2$ is maintained at T_1. The crack faces remain thermally and electrically insulated.

The material property parameters are taken to vary continuously along the z-direction inside the FGPM strip. The material properties of FGPM, such as the elastic stiffness constants $c_{kl}(z)$, the piezoelectric constants $e_{kl}(z)$, the dielectric constants $\varepsilon_{kk}(z)$, the stress-temperature coefficients $\lambda_{kk}(z)$, the coefficient of heat conduction $\kappa_x(z), \kappa_z(z)$ and the pyroelectric constant $p_z(z)$, are one-dimensionally dependent as

$$\left.\begin{aligned}
(c_{kl}, e_{kl}, \varepsilon_{kk}) &= (c_{kl0}, e_{kl0}, \varepsilon_{kk0})\exp(\beta z)\\
(\lambda_{kk}, p_z) &= (\lambda_{kk0}, p_{z0})\exp[(\beta+\omega)z]\\
(\kappa_x, \kappa_z) &= (\kappa_{x0}, \kappa_{z0})\exp(\delta z)
\end{aligned}\right\} \tag{1}$$

where β, ω and δ are positive or negative constants, and the subscript 0 indicates the properties at the plane $z=0$. For some materials, the thermal diffusivity τ_0 indeed doesn't vary dramatically, then τ_0 is assumed to be a constant. The material properties of the homogeneous elastic layer are the elastic stiffness constants c_{kl}^E, the stress-temperature coefficients λ_{kk}^E, the coefficient of heat conduction κ^E and the thermal diffusivity τ_0^E. The superscript E denotes the physical quantities of the homogeneous elastic layer.

The crack problem may be solved by superposition. In the problem considered here, since the heat conduction is one-dimensional and straight cracks do not obstruct the heat flow in this arrangement, determination of the temperature distribution and the resulting thermal stress would be quite straightforward and the related crack problem would be one of model I. We suppose that each crack is opened under the action of the same distribution of the internal

pressure $\sigma_0^T(z,t)$, where $\sigma_0^T(z,t)$ is the thermal stress induced by the time-dependent temperature change. In the following, the subscripts x, y, z will be used to refer to the direction of coordinates.

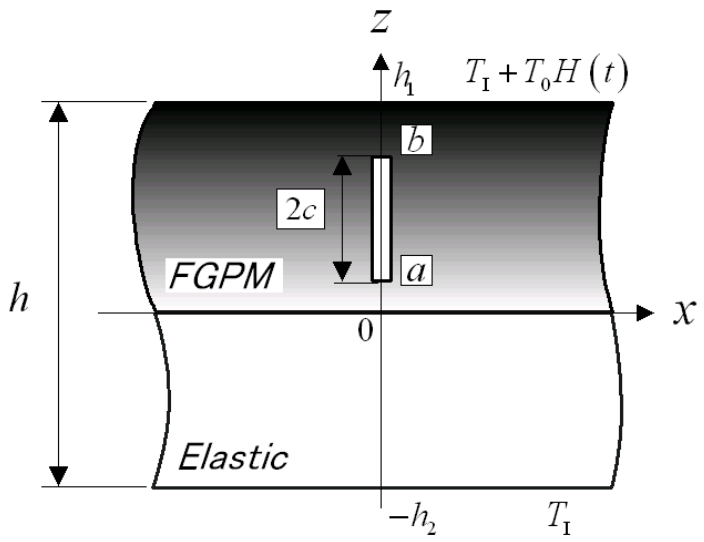


Figure 1 : Geometry of the crack problem in a functionally graded piezoelectric laminate

Temperature distribution and thermal stress in the un-cracked strip

By using the Laplace transform method, the temperatures $T^*(z,p)\quad(0\le z\le h_1)$ and $T^{E*}(z,p)\quad(-h_2\le z\le 0)$ in the Laplace transform plane can be easily obtained as follows:

$$\left.\begin{aligned} T^*(z,p) &= \sum_{j=1}^{2} D_{1j}\exp(\mu_{1j}z) \quad (0\le z\le h_1) \\ T^{E*}(z,p) &= \sum_{j=1}^{2} D_{2j}\exp(\mu_{2j}z) \quad (-h_2\le z\le 0) \end{aligned}\right\} \tag{2}$$

where the superscript $*$ denotes the physical quantities in the Laplace domain and p is the Laplace parameter. The functions μ_{ij} and D_{ij} $(i,j=1,2)$ are given in Appendix A. Thus the temperature fields $T(z,t)\ (0\le z\le h_1)$ and $T^E(z,t)\quad(-h_2\le z\le 0)$ in the time domain may be evaluated as:

$$\left.\begin{aligned} T(z,t) &= \frac{1}{2\pi i}\int_{Br} T^*(z,p)\exp(pt)dp \quad (0\le z\le h_1) \\ T^E(z,t) &= \frac{1}{2\pi i}\int_{Br} T^{E*}(z,p)\exp(pt)dp \quad (-h_2\le z\le 0) \end{aligned}\right\} \tag{3}$$

The temperature fields $T(z,t)$ $(0 \le z \le h_1)$ and $T^E(z,t)$ $(-h_2 \le z \le 0)$ can be found from Eq. (3) by using the numerical Laplace inversion scheme [12].

Once $T(z,t)\,(0 \le z \le h_1)$ and $T^E(z,t)\,(-h_2 \le z \le 0)$ is known, the thermal stress component $\sigma_{xx}^T(z,t)\,(0 \le z \le h_1)$ can be also obtained by the following equation:

$$\sigma_{xx}^T = \bar{c}_{110} \exp(\beta z) \frac{\partial u_x^T}{\partial x} - \bar{\lambda}_{110} \exp[(\beta + \omega)z]T \tag{4}$$

where the superscript T denotes the thermally induced quantities, $u_x^T(z,t)$ is the displacement component and

$$\left.\begin{aligned} \bar{c}_{110} &= c_{110} - \frac{c_{130}(c_{130}\varepsilon_{330} + e_{310}e_{330}) + e_{310}(c_{130}e_{330} - c_{330}e_{310})}{c_{330}\varepsilon_{330} + e_{330}^2} \\ \bar{\lambda}_{110} &= \lambda_{110} - \frac{\lambda_{330}(c_{130}\varepsilon_{330} + e_{310}e_{330}) - p_{z0}(c_{130}e_{330} - c_{330}e_{310})}{c_{330}\varepsilon_{330} + e_{330}^2} \end{aligned}\right\} \tag{5}$$

Similarly, $\sigma_{xx}^{ET}(z,t)\,(-h_2 \le z \le 0)$ is given by

$$\sigma_{xx}^{ET} = \bar{c}_{11}^E \frac{\partial u_x^{ET}}{\partial x} - \bar{\lambda}_{11}^E T^E \tag{6}$$

where $u_x^{ET}(z,t)$ is the displacement component and

$$\bar{c}_{11}^E = c_{11}^E - \frac{\left(c_{13}^E\right)^2}{c_{33}^E}, \quad \bar{\lambda}_{11}^E = \lambda_{11}^E - \frac{c_{13}^E \lambda_{33}^E}{c_{33}^E} \tag{7}$$

The compatibility conditions that need to be satisfied become

$$\frac{\partial^2}{\partial z^2}\left[\frac{\partial u_x^T}{\partial x}\right] = 0, \quad \frac{\partial^2}{\partial z^2}\left[\frac{\partial u_x^{ET}}{\partial x}\right] = 0 \tag{8}$$

giving

$$\left.\begin{aligned} &\frac{\partial u_x^T}{\partial x} = A^T z + B^T \\ &\sigma_{xx}^T = \overline{c}_{110}\exp(\beta z)(A^T z + B^T) - \overline{\lambda}_{110}\exp[(\beta+\omega)z]T \end{aligned}\right\} \tag{9}$$

$$\left.\begin{aligned} &\frac{\partial u_x^{ET}}{\partial x} = A^T z + B^T \\ &\sigma_{xx}^{ET} = \overline{c}_{11}^E(A^T z + B^T) - \overline{\lambda}_{11}^E T^E \end{aligned}\right\} \tag{10}$$

where $A^T(t)$ and $B^T(t)$ are unknown functions to be obtained from boundary conditions for the laminate. If the laminate is unconstrained along its boundaries, we have

$$\left.\begin{aligned} &\int_0^{h_1}\sigma_{xx}^T(z,t)dz + \int_{-h_2}^0 \sigma_{xx}^{ET}(z,t)dz = 0 \\ &\int_0^{h_1}\sigma_{xx}^T(z,t)zdz + \int_{-h_2}^0 \sigma_{xx}^{ET}(z,t)zdz = 0 \end{aligned}\right\} \tag{11}$$

In the crack problem under considered, the equal and opposite of the stress $\sigma_0^T(z,t) = \sigma_{xx}^T(z,t) \quad (a < z < b)$ given by Eq. (4) will be used as the crack surface traction and the laminate will be assumed to be under plane-strain conditions.

The crack problem

Referring to Figure 1, it is assumed that $x = 0$ is a plane of symmetry regarding to geometry and loading conditions. Thus, in analyzing the problem it is sufficient to consider one-half ($0 \le x < \infty$) of the FGPM strip and the homogeneous elastic layer only. Also, through a proper superposition, the problem is assumed to have been reduced to a perturbation problem in which the crack surface tractions are the only nonzero external loads and the stresses in the layered strip vanish for $x \to \infty$.

Taking Eqs. (1) into consideration, the governing equations for the electromechanical fields of the FGPM strip and the homogeneous elastic layer may then be expressed as follows:

$$\left.\begin{aligned}
&c_{110}\frac{\partial^2 u_x}{\partial x^2}+c_{440}\frac{\partial^2 u_x}{\partial z^2}+\left(c_{130}+c_{440}\right)\frac{\partial^2 u_z}{\partial x\partial z}+\left(e_{310}+e_{150}\right)\frac{\partial^2 \phi}{\partial x\partial z}\\
&\quad+\beta\left\{c_{440}\left(\frac{\partial u_x}{\partial z}+\frac{\partial u_z}{\partial x}\right)+e_{150}\frac{\partial \phi}{\partial x}\right\}=0\\
&c_{440}\frac{\partial^2 u_z}{\partial x^2}+c_{330}\frac{\partial^2 u_z}{\partial z^2}+\left(c_{130}+c_{440}\right)\frac{\partial^2 u_x}{\partial x\partial z}+e_{150}\frac{\partial^2 \phi}{\partial x^2}+e_{330}\frac{\partial^2 \phi}{\partial z^2}\\
&\quad+\beta\left(c_{130}\frac{\partial u_x}{\partial x}+c_{330}\frac{\partial u_z}{\partial z}+e_{330}\frac{\partial \phi}{\partial z}\right)=0\\
&e_{150}\frac{\partial^2 u_z}{\partial x^2}+e_{330}\frac{\partial^2 u_z}{\partial z^2}+\left(e_{150}+e_{310}\right)\frac{\partial^2 u_x}{\partial x\partial z}-\varepsilon_{110}\frac{\partial^2 \phi}{\partial x^2}-\varepsilon_{330}\frac{\partial^2 \phi}{\partial z^2}\\
&\quad+\beta\left(e_{310}\frac{\partial u_x}{\partial x}+e_{330}\frac{\partial u_z}{\partial z}-\varepsilon_{330}\frac{\partial \phi}{\partial z}\right)=0
\end{aligned}\right\} \tag{12}$$

$$\left.\begin{aligned}
&c_{11}^E\frac{\partial^2 u_x^E}{\partial x^2}+c_{44}^E\frac{\partial^2 u_x^E}{\partial z^2}+\left(c_{13}^E+c_{44}^E\right)\frac{\partial^2 u_z^E}{\partial x\partial z}=0\\
&c_{44}^E\frac{\partial^2 u_z^E}{\partial x^2}+c_{33}^E\frac{\partial^2 u_z^E}{\partial z^2}+\left(c_{13}^E+c_{44}^E\right)\frac{\partial^2 u_x^E}{\partial x\partial z}=0
\end{aligned}\right\} \tag{13}$$

The boundary conditions can be written as

$$\left.\begin{aligned}
&\sigma_{xx}(0,z,t)=-\sigma_0^T(z,t) \quad (a<z<b)\\
&u_x(0,z,t)=0 \quad (0\le z\le a, b\le z\le h_1)
\end{aligned}\right\} \tag{14}$$

$$\left.\begin{aligned}
&\sigma_{zx}(0,z,t)=0\\
&D_x(0,z,t)=0
\end{aligned}\right\} \quad (0\le z\le h_1) \tag{15}$$

$$\left.\begin{aligned}
&\sigma_{zx}^E(0,z,t)=0\\
&u_x^E(0,z,t)=0
\end{aligned}\right\} \quad (-h_2\le z\le 0) \tag{16}$$

$$\left.\begin{aligned}
&\sigma_{zx}(x,h_1,t)=0\\
&\sigma_{zz}(x,h_1,t)=0\\
&D_z(x,h_1,t)=0
\end{aligned}\right\} \quad (0\le x<\infty) \tag{17}$$

$$\left.\begin{aligned}
&\sigma_{zx}(x,0,t)=\sigma_{zx}^{E}(x,0,t)\\
&\sigma_{zz}(x,0,t)=\sigma_{zz}^{E}(x,0,t)\\
&D_{z}(x,0,t)=0\\
&u_{x}(x,0,t)=u_{x}^{E}(x,0,t)\\
&u_{z}(x,0,t)=u_{z}^{E}(x,0,t)
\end{aligned}\right\}\quad (0\le x<\infty) \tag{18}$$

$$\left.\begin{aligned}
&\sigma_{zx}^{E}(x,-h_2,t)=0\\
&\sigma_{zz}^{E}(x,-h_2,t)=0
\end{aligned}\right\}\quad (0\le x<\infty) \tag{19}$$

By using the Fourier integral transform technique, the stress intensity factors $K_{\mathrm{IA}}(t)$ at $z=a$ and $K_{\mathrm{IB}}(t)$ at $z=b$ may be evaluated as

$$K_{\mathrm{IA}}(t)=\left\{\begin{array}{ll}
-Z^{\infty}\exp(\beta a)(\pi c)^{1/2}\Phi(a,t) & (b<h_1)\\
 & \\
-Z^{\infty}\exp(\beta a)(4\pi c)^{1/2}\Phi(a,t) & (b=h_1)
\end{array}\right\} \tag{20}$$

$$K_{\mathrm{IB}}(t)=\left\{\begin{array}{ll}
Z^{\infty}\exp(\beta b)(\pi c)^{1/2}\Phi(b,t) & (b<h_1)\\
 & \\
0 & (b=h_1)
\end{array}\right\} \tag{21}$$

In Eqs.(20) and (21), the constant Z^{∞} is given in Appendix B and the function $\Phi(u,t)$ is given by

$$G(\xi,t)=\frac{\Phi(u,t)}{(1+u)^{1/2}(1-u)^{\alpha}} \tag{22}$$

where $\alpha=1/2$ for $0<b<h_1$ (embedded crack) and $\alpha=-1/2$ for $b=h_1$ (edge crack), and $\xi=(b-a)u/2+(b+a)/2$. The function $G(\xi,t)$ is the solution of the following singular integral equation obtained from the mixed boundary conditions (14) with the boundary conditions (15)-(19).

$$\int_a^b G(\xi,t)\left[\frac{1}{\xi-z}+\sum_{i=1}^{4}M_i(\xi,z)\right]d\xi=\exp(-\beta z)\frac{\pi}{\Re[Z^{\infty}]}\sigma_0^{T}(z,t)\quad (a<z<b) \tag{23}$$

In the integral equation, the kernel functions $M_i(\xi,t)\ (i=1,2,3,4)$ are also given in Appendix B. The singular integral equation (23) for $0<b<h_1$ (embedded crack) is to be solved with the following subsidiary conditions obtained from the second mixed boundary condition of Eqs.(14).

$$\int_a^b G(\xi,t)d\xi = 0 \tag{24}$$

Numerical result and discussion

For the numerical calculations, the properties of cadmium selenide [16] are used as the properties of the FGPM strip at the plane $z=0$.

$$\left.\begin{aligned}
&c_{110} = 7.41\times10^{10}[\mathrm{N/m^2}], && c_{130} = 3.93\times10^{10}[\mathrm{N/m^2}],\\
&c_{330} = 8.36\times10^{10}[\mathrm{N/m^2}], && c_{440} = 1.32\times10^{10}[\mathrm{N/m^2}],\\
&e_{310} = -0.16[\mathrm{C/m^2}], && e_{330} = 0.347[\mathrm{C/m^2}],\\
&e_{150} = -0.138[\mathrm{C/m^2}], &&\\
&\varepsilon_{110} = 0.825\times10^{-10}[\mathrm{C/Vm}], && \varepsilon_{330} = 0.903\times10^{-10}[\mathrm{C/Vm}],\\
&\lambda_{110} = 0.621\times10^{6}[\mathrm{N/Km^2}], && \lambda_{330} = 0.551\times10^{6}[\mathrm{N/Km^2}],\\
&p_{z0} = -2.94\times10^{-6}[\mathrm{C/Km^2}]. &&
\end{aligned}\right\} \tag{25}$$

Since the values of the coefficients of heat conduction for cadmium selenide could not be found in the literature, the value $\kappa^2 = \kappa_x/\kappa_z = 1/1.5$ is used. The normalized nonhomogeneous parameters βh_1, ωh_1, δh_1 and the thermal diffusivities τ_0 and ${\tau_0}^E$ are assumed to be $\beta h_1 = \omega h_1 = \delta h_1$ and $\tau_0 = {\tau_0}^E$. The properties of titanium (Ti) and brass with following properties are also used as the properties of the elastic layer.

$$\left.\begin{aligned}
&c_{11}^E = c_{33}^E = \frac{2(1-\nu)}{1-2\nu}\mu, && c_{13}^E = \frac{2\nu}{1-2\nu}\mu,\\
&c_{44}^E = \mu, && \lambda_{11}^E = \lambda_{33}^E = \frac{2(1+\nu)}{1-2\nu}\mu\alpha^E.
\end{aligned}\right\} \tag{26}$$

where α^E is the coefficient of linear thermal expansion. The values of μ, ν and α^E of titanium (Ti) and brass are

$$\left.\begin{aligned}
\mu &= \begin{cases} 42.6\times10^9\ [\mathrm{N/m^2}] & \text{(Ti)}\\ 41.4\times10^9\ [\mathrm{N/m^2}] & \text{(Brass)}\end{cases}\\
\nu &= \begin{cases} 0.28 & \text{(Ti)}\\ 0.33 & \text{(Brass)}\end{cases}\\
\alpha^E &= \begin{cases} 8.4\times10^{-6}\ [1/\mathrm{K}] & \text{(Ti)}\\ 19.9\times10^{-6}\ [1/\mathrm{K}] & \text{(Brass)}\end{cases}
\end{aligned}\right\} \tag{27}$$

First of all, we consider the effect of the thickness ratio h_2 / h_1 on the stress intensity factors K_{IA} and K_{IB} for the Ti elastic layer. Figures 2(a) and 2(b) show the normalized stress intensity factors $(K_{\mathrm{IA}}, K_{\mathrm{IB}}) / \lambda_{110} T_0 (\pi c)^{1/2}$ versus time t for $h_2 / h_1 = 1.0$, 0.5 and $\rightarrow 0.0$. It is supposed that the crack length parameter is $c / h_1 = 0.1$, the crack location parameter is $(a+b) / 2h_1 = 0.5$ and the nonhomogeneous parameter is $\beta h_1 = 0.0$. In the following figures, the time t is represented through the dimensionless Fourier number defined by

$$F = \frac{\tau_0 t}{h^2} \tag{28}$$

Note that the values of those intensity factors rise sharply at first, reach maximum values and then decrease and approach the static values with increasing F. The results for the case of $h_2 / h_1 \rightarrow 0.0$ are coincident with the results of the previous paper [3]. The magnitudes of $(K_{\mathrm{IA}}, K_{\mathrm{IB}}) / \lambda_{110} T_0 (\pi c)^{1/2}$ depend remarkably on h_2 / h_1. The normalized maximum values of the stress intensity factors for $h_2 / h_1 = 0.5$ are larger than those for h_2 / h_1 =1.0 and $h_2 / h_1 \rightarrow 0.0$.

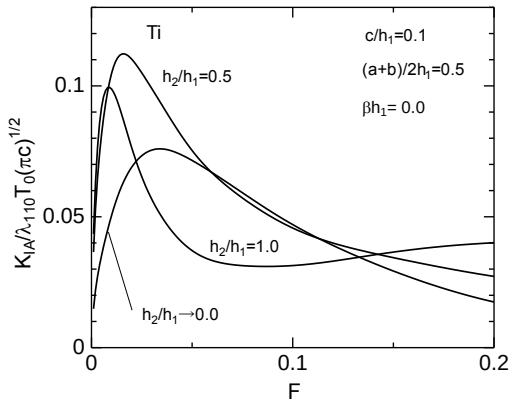


Figure 2(a). The effect of the thickness ratio h_2 / h_1 on the stress intensity factor K_{IA} for the Ti layer (embedded crack).

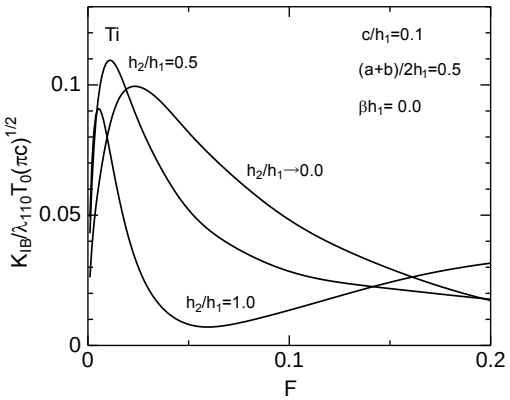


Figure 2(b). The effect of the thickness ratio h_2 / h_1 on the stress intensity factor K_{IB} for the Ti layer (embedded crack).

Next, we examine the effect of the crack length parameter c/h_1 on the time dependences of the stress intensity factors K_{IA} and K_{IB} for the Ti elastic layer. The normalized stress intensity factors $(K_{IA}, K_{IB})/\lambda_{110}T_0(\pi c)^{1/2}$ are plotted versus F for $c/h_1 = 0.1$, 0.2 and 0.3 with $h_2/h_1 = 1.0$, $(a+b)/2h_1 = 0.5$ and $\beta h_1 = 2.0$ in Figures 3(a) and 3(b). The same tendencies as the cases shown in Figures 2(a) and 2(b) are observed. For the case of $c/h_1 = 0.3$, the value of $K_{IB}/\lambda_{110}T_0(\pi c)^{1/2}$ at some time becomes negative, and the crack would be closed, because the stress $\sigma_0^T(z,t)$ $(a < z < b)$ on the surfaces of the crack becomes compressive.

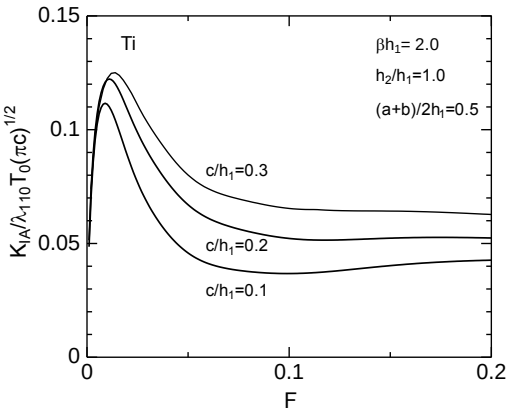


Figure 3(a). The effect of the crack length c/h_1 on the stress intensity factor K_{IA} for the Ti layer (embedded crack).

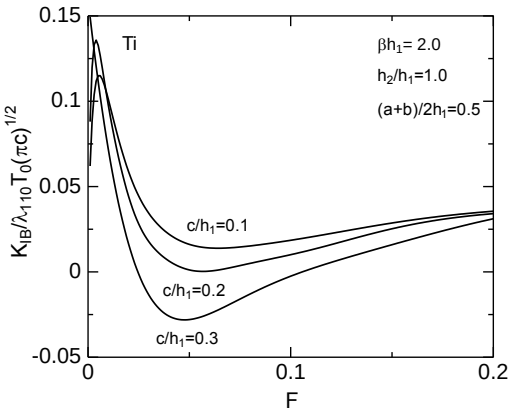


Figure 3(b). The effect of the crack length c/h_1 on the stress intensity factor K_{IB} for the Ti layer (embedded crack).

Thirdly, we consider the effect of the crack location parameter $(a+b)/2h_1$, the nonhomogeneous parameter βh_1 and the material properties of the elastic layer. Figures 4(a) and (b) indicate the time dependences of the stress intensity factors K_{IA} and K_{IB} for the Ti elastic layer. It is supposed that the geometric parameters are $c/h_1 = 0.1$, $h_2/h_1 = 1.0$ and $(a+b)/2h_1$=0.2, 0.5 and 0.8. In these figures, the solid, dashed and dotted lines indicate the results for the $\beta h_1 = 2.0$, -2.0 and 0.0, respectively. Figures 5(a) and (b) are the same as Figures 4(a) and (b) for the Brass elastic layer. For the case of $(a+b)/2h_1$=0.8, we can also see the crack contact phenomenon. The values of the stress intensity factors for the Brass layer are much larger than those for the Ti layer. The most remarkable difference between the

results for the Ti elastic layer and the Brass layer is whether the time dependences of the stress intensity factors have the peak value or not. For the case of the Brass layer, the time dependences of $(K_{\mathrm{IA}},\ K_{\mathrm{IB}})/\lambda_{110}T_0(\pi c)^{1/2}$ do not have the peak value.

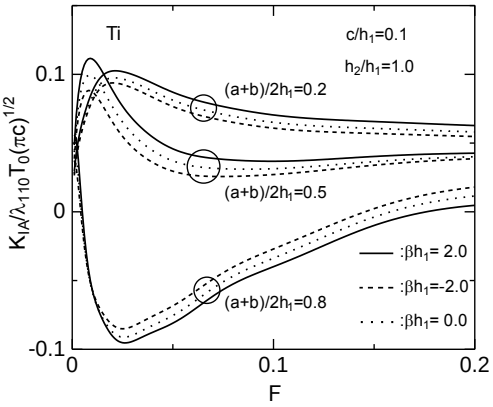


Figure 4(a). The effect of the crack location $(a+b)/2h_1$ and the nonhomogeneous parameter βh_1 on the stress intensity factor K_{IA} for the Ti layer (embedded crack).

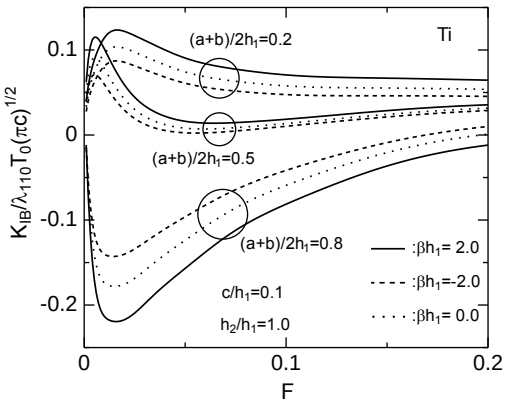


Figure 4(b). The effect of the crack location $(a+b)/2h_1$ and the nonhomogeneous parameter βh_1 on the stress intensity factor K_{IB} for the Ti layer (embedded crack).

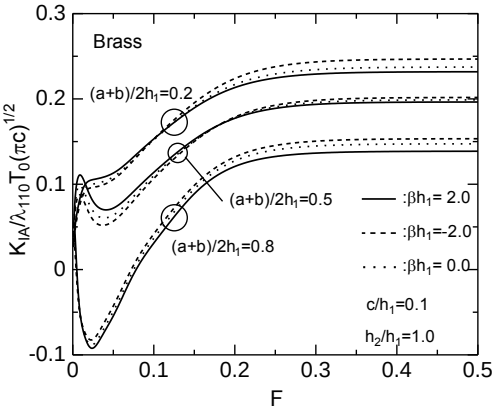


Figure 5(a). The effect of the crack location $(a+b)/2h_1$ and the nonhomogeneous parameter βh_1 on the stress intensity factor K_{IA} for the Brass layer (embedded crack).

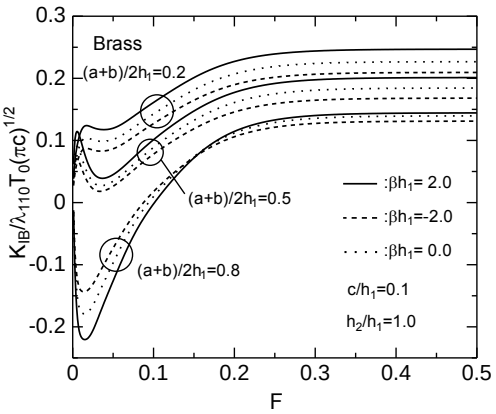


Figure 5(b). The effect of the crack location $(a+b)/2h_1$ and the nonhomogeneous parameter βh_1 on the stress intensity factor K_{IB} for the Brass layer (embedded crack).

Finally, we consider the case of $b = h_1$ (edge crack). Assume the top surface of the strip is cooled from initial temperature T_I to $T_I + T_0$ $(T_0 < 0)$ suddenly, the normalized stress intensity factor $K_{IA}/\lambda_{110}|T_0|(2\pi c)^{1/2}$ is plotted versus F for the $\beta h_1 = 2.0$, -2.0 and 0.0 with $2c/h_1 = 0.2$ in Figure 6 and for $2c/h_1 = 0.2, 0.4$ with $\beta h_1 = 2.0$ in Figure 7, respectively. The normalized value of stress intensity factor decreases with decreasing βh_1 and $2c/h_1$. The influence of the material nonhomogeneity on the stress intensity factor is the same as the results for the embedded crack shown in Figure 4(a). For the case of large F, the stress intensity factor may be negative and the crack contact occurs.

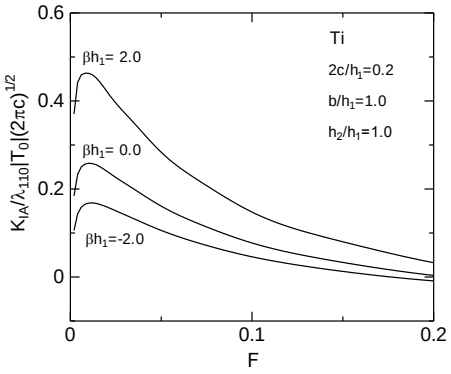


Figure 6. The effect of the nonhomogeneous parameter βh_1 on the stress intensity factor K_{IA} for the Ti layer (edge crack).

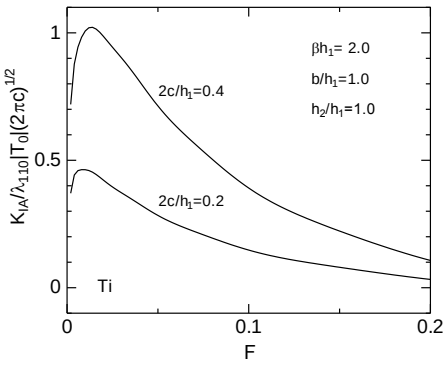


Figure 7. The effect of the crack length $2c/h_1$ on the stress intensity factor K_{IA} for the Ti layer (edge crack).

CONCLUSION

The transient fracture problem of the cracked functionally graded piezoelectric strip bonded to the homogeneous elastic layer is studied. The effects of the thickness of the elastic layer, the crack length, the crack location and the material nonhomogenity on the fracture behavior are considered. The following facts can be found from the numerical results.

(1) The distinct overshooting phenomenon for the case of the Ti elastic layer can be observed and this fact may suggest the importance of these transient analyses. The effect of the thickness of the elastic layer on the time dependence of the stress intensity factors is large (Figs. 2(a) and 2(b)).

(2) The maximum values of the stress intensity factors and the static values of them indicating the inertial effect increase with increasing c/h_1. For the case of $c/h_1 = 0.3$, the stress intensity factor K_{IB} becomes negative (Figs. 3(a) and 3(b)).

(3) While the time dependences of the stress intensity factors for the Ti elastic layer have the peak values, those for the Brass layer do not have the peak values. Generally, the decrease of βh_1 is beneficial for reducing the stress intensity factors. However, the static values of the stress intensity factors for the Brass elastic layer decreases with increasing βh_1. For the case of the crack near the heating surface, the crack contact phenomenon can be found (Figs. 4(a), 4(b) and 5(a), 5(b)).

(4) In some cases, the stress intensity factors under the thermal load become negative and the results have no physical meaning. However, when the thermal load is combined with the mechanical load which induces the positive stress intensity factor, those results can be used effectively.

Appendix A

The functions μ_{ij} $(i, j = 1,2)$ are

$$\left.\begin{aligned} \mu_{11} &= -\frac{\delta}{2} - \mu_0, \quad \mu_{21} = \left(\frac{p}{\tau_0^E}\right)^{1/2} \\ \mu_{12} &= -\frac{\delta}{2} + \mu_0, \quad \mu_{22} = -\left(\frac{p}{\tau_0^E}\right)^{1/2} \end{aligned}\right\} \tag{A.1}$$

where

$$\mu_0 = \left(\frac{\delta^2}{4} + \frac{p}{\tau_0}\right)^{1/2} \tag{A.2}$$

The functions D_{ij} $(i, j = 1, 2)$ are

$$\left.\begin{aligned} D_{11} &= \frac{\rho_{01}\rho_{22}}{\rho_{11}\rho_{22} - \rho_{12}\rho_{21}} \frac{T_0}{p} \\ D_{21} &= -\frac{\rho_{01}\rho_{12}}{\rho_{11}\rho_{22} - \rho_{12}\rho_{21}} \frac{T_0}{p} \\ D_{12} &= \frac{(\rho_{11}\rho_{22} - \rho_{12}\rho_{21})\exp(-\mu_{12}h_1) - \rho_{01}\rho_{22}\exp(-2\mu_0 h_1)}{\rho_{11}\rho_{22} - \rho_{12}\rho_{21}} \frac{T_0}{p} \\ D_{22} &= \frac{\rho_{01}\rho_{12}\exp(-2\mu_{21}h_2)}{\rho_{11}\rho_{22} - \rho_{12}\rho_{21}} \frac{T_0}{p} \end{aligned}\right\} \tag{A.3}$$

where

$$\left.\begin{aligned} \rho_{01} &= (\mu_{12} - \mu_{22})\exp(-\mu_{12}h_1) \\ \rho_{02} &= (\mu_{12} - \mu_{22})\exp(\mu_{22}h_2) \\ \rho_{11} &= \mu_{22} - \mu_{11} + (\mu_{12} - \mu_{22})\exp(-2\mu_0 h_1) \\ \rho_{21} &= \mu_{21} - \mu_{22} \\ \rho_{12} &= \mu_{12} - \mu_{11} \\ \rho_{22} &= \mu_{21} - \mu_{12} + (\mu_{12} - \mu_{22})\exp(-2\mu_{21}h_2) \end{aligned}\right\} \tag{A.4}$$

Appendix B

The constant Z^{∞} is given by

$$\left.\begin{aligned} Z^{\infty} &= \lim_{s\to\infty} Z(s) \\ Z(s) &= \sum_{j=1}^{3} p_{11j}(s)\delta_j(s) \end{aligned}\right\} \tag{B.1}$$

where

$$\left.\begin{aligned} \delta_1(s) &= \frac{\rho_{22}(s)\rho_{33}(s)}{a_{11}(s)\rho_{22}(s)\rho_{33}(s) - a_{12}(s)\rho_{21}(s)\rho_{33}(s) - a_{13}(s)\rho_{22}(s)\rho_{31}(s)} \\ \delta_j(s) &= -\frac{\rho_{j1}(s)}{\rho_{jj}(s)}\delta_1(s) \quad (j = 2, 3) \end{aligned}\right\} \tag{B.2}$$

with

$$\left.\begin{aligned}\rho_{2j}(s) &= p_{313}(s)p_{21j}(s) - p_{213}(s)p_{31j}(s)\\ \rho_{3j}(s) &= p_{312}(s)p_{21j}(s) - p_{212}(s)p_{31j}(s)\end{aligned}\right\} \quad (j = 2,3) \tag{B.3}$$

The kernel functions $M_i(\xi, z)$ $(i = 1,2,3,4)$ are

$$\left.\begin{aligned}M_1(\xi,z) &= \int_0^\infty \left[\frac{\Re[Z(s)]}{\Re[Z^\infty]} - 1\right]\sin[s(\xi - z)]ds\\ M_2(\xi,z) &= \int_0^\infty \frac{\Im[Z(s)]}{\Re[Z^\infty]}\cos[s(\xi - z)]ds\\ M_3(\xi,z) &= -\frac{2}{\Re[Z^\infty]}\int_0^\infty m^\infty(s,\xi,z)ds\\ M_4(\xi,z) &= -\frac{2}{\Re[Z^\infty]}\int_0^\infty \left[m(s,\xi,z) - m^\infty(s,\xi,z)\right]ds\end{aligned}\right\} \tag{B.4}$$

where

$$\left.\begin{aligned}m(s,\xi,z) &= \sum_{i=1}^{8}\sum_{j=1}^{3} p_{12j}(s)r_{ji}(s)Q_i(s,\xi)\exp(s\gamma_{2j}z)\\ &\quad + \sum_{i=1}^{8}\sum_{j=4}^{6} p_{12j}(s)r_{ji}(s)Q_i(s,\xi)\exp[-s\gamma_{2j}(h_1 - z)]\\ m^\infty(s,\xi,z) &= \sum_{k=1}^{3}\left\{\sum_{i=1}^{3} p_{i+2,1k}^\infty\left[\sum_{j=1}^{3} F_{ijk}\exp(-s\theta_{1jk}) + \sum_{j=4}^{6} F_{ijk}\exp(-s\theta_{3jk})\right]\right.\\ &\quad + \sum_{i=4}^{6} p_{i-1,1k}^\infty\left[\sum_{j=1}^{3} F_{ijk}\exp(-s\theta_{2jk}) + \sum_{j=4}^{6} F_{ijk}\exp(-s\theta_{4jk})\right]\\ &\quad \left. + \sum_{i=7}^{8} p_{i-1,1k}^\infty\left[\sum_{j=1}^{3} \bar{F}_{ijk}\exp(-s\theta_{1jk}) + \sum_{j=4}^{6} \bar{F}_{ijk}\exp(-s\theta_{3jk})\right]\right\}\end{aligned}\right\} \tag{B.5}$$

$$\left.\begin{aligned}
F_{ijk} &= \frac{1}{2\gamma_{1k}^{\infty}} p_{12j}^{\infty} r_{ji}^{\infty} \delta_k^{\infty} && (i, j = 1,2,...,6)\\
\bar{F}_{ijk} &= \frac{1}{2} p_{12j}^{\infty} r_{ji}^{\infty} \delta_k^{\infty} && (i = 7,8,\ j = 1,2,...,6)\\
\theta_{1jk} &= -\frac{\xi}{\gamma_{1k}^{\infty}} - \gamma_{2j}^{\infty} z && (j = 1,2,3)\\
\theta_{2jk} &= -\frac{h_1 - \xi}{\gamma_{1k}^{\infty}} - \gamma_{2j}^{\infty} z && (j = 1,2,3)\\
\theta_{3jk} &= -\frac{\xi}{\gamma_{1k}^{\infty}} - \gamma_{2j}^{\infty} (h_1 - z) && (j = 4,5,6)\\
\theta_{4jk} &= -\frac{h_1 - \xi}{\gamma_{1k}^{\infty}} - \gamma_{2j}^{\infty} (h_1 - z) && (j = 4,5,6)
\end{aligned}\right\} (k = 1,2,3) \qquad \text{(B.6)}$$

$$\left.\begin{aligned}
p_{i1j}^{\infty} &= \lim_{s\to\infty} p_{i1j}(s) && (i = 1,2,...,8,\ \ j = 1,2,3)\\
p_{i2j}^{\infty} &= \lim_{s\to\infty} p_{i2j}(s) && (i = 1,2,...,8,\ \ j = 1,2,...,6)\\
r_{ji}^{\infty} &= \lim_{s\to\infty} r_{ji}(s) && (i, j = 1,2,...,10)\\
\gamma_{1j}^{\infty} &= \lim_{s\to\infty} \gamma_{1j}(s) && (j = 1,2,3)\\
\gamma_{2j}^{\infty} &= \lim_{s\to\infty} \gamma_{2j}(s) && (j = 1,2,...,6)\\
\delta_j^{\infty} &= \lim_{s\to\infty} \delta_j(s) && (j = 1,2,3)
\end{aligned}\right\} \qquad \text{(B.7)}$$

where

$$\left.\begin{aligned}
p_{11j}(s) &= c_{110}\gamma_{1j}a_{1j}(s) - c_{130} - e_{310}b_{1j}(s)\\
p_{21j}(s) &= e_{150}[\gamma_{1j} + a_{1j}(s)] - \varepsilon_{110}\gamma_{1j}b_{1j}(s)\\
p_{31j}(s) &= c_{440}[\gamma_{1j} + a_{1j}(s)] + e_{150}\gamma_{1j}b_{1j}(s)\\
p_{41j}(s) &= c_{130}\gamma_{1j}a_{1j}(s) - c_{330} - e_{330}b_{1j}(s)\\
p_{51j}(s) &= e_{310}\gamma_{1j}a_{1j}(s) - e_{330} + \varepsilon_{330}b_{1j}(s)\\
p_{61j}(s) &= a_{1j}(s)\\
p_{71j}(s) &= 1\\
p_{81j}(s) &= b_{1j}(s)
\end{aligned}\right\} \ (j = 1,2,3) \qquad \text{(B.8)}$$

$$\left.\begin{aligned}
p_{12j}(s) &= c_{110}a_{2j}(s) - \gamma_{2j}[c_{130} - e_{310}b_{2j}(s)] \\
p_{22j}(s) &= e_{150}[\gamma_{2j}a_{2j}(s) - 1] - \varepsilon_{110}b_{2j}(s) \\
p_{32j}(s) &= c_{440}[\gamma_{2j}a_{2j}(s) - 1] + e_{150}b_{2j}(s) \\
p_{42j}(s) &= c_{130}a_{2j}(s) + \gamma_{2j}[c_{330} - e_{330}b_{2j}(s)] \\
p_{52j}(s) &= e_{310}a_{2j}(s) + \gamma_{2j}[e_{330} + \varepsilon_{330}b_{2j}(s)] \\
p_{62j}(s) &= a_{2j}(s) \\
p_{72j}(s) &= 1 \\
p_{82j}(s) &= -b_{2j}(s)
\end{aligned}\right\} \quad (j = 1,2,...,6) \tag{B.9}$$

In Eqs. (B.2), (B.7), (B.8) and (B.9), the functions $\gamma_{1j} = \gamma_{1j}(s)$, $a_{1j}(s)$ and $b_{1j}(s)$ $(j = 1,2,3)$ are given in Appendix A of the previous paper [7], and the functions $\gamma_{2j} = \gamma_{2j}(s)$, $a_{2j}(s)$ and $b_{2j}(s)$ $(j = 1,2,...,6)$ are given in Appendix B of the previous paper [4].The functions $r_{ij}(s)$ $(i, j = 1,2,...,10)$ are the elements of a square matrix $R = \Omega^{-1}$ of order 10. The elements $\omega_{i,j}(s)$ $(i, j = 1,2,...,10)$ of the square matrix Ω are given by

$$\left.\begin{aligned}
\omega_{i-2,j}(s) &= p_{i2j}(s) && (i = 3,4,5) \\
\omega_{i+1,j}(s) &= p_{i2j}(s)\exp(s\gamma_{2j}h_1) && (i = 3,4,5) \\
\omega_{i+1,j}(s) &= p_{i2j}(s) && (i = 6,7)
\end{aligned}\right\} (j = 1,2,...,6) \tag{B.10}$$

$$\left.\begin{aligned}
&\omega_{1,7}(s) = 2\mu, && \omega_{1,8}(s) = 2\mu(1-2\nu) \\
&\omega_{1,9}(s) = -2\mu, && \omega_{1,10}(s) = 2\mu(1-2\nu) \\
&\omega_{2,7}(s) = 2\mu, && \omega_{2,8}(s) = 2\mu(1-\nu) \\
&\omega_{2,9}(s) = 2\mu, && \omega_{2,10}(s) = -2\mu(1-\nu) \\
&\omega_{7,7}(s) = -1, && \omega_{7,9}(s) = -1 \\
&\omega_{8,7}(s) = -1, && \omega_{8,8}(s) = -(3-4\nu) \\
&\omega_{8,9}(s) = 1, && \omega_{8,10}(s) = -(3-4\nu) \\
&\omega_{9,7}(s) = \exp(sh_2), && \omega_{9,8}(s) = -(sh_2 - 1 + 2\nu)\exp(sh_2) \\
&\omega_{9,9}(s) = -\exp(-sh_2), && \omega_{9,10}(s) = (sh_2 + 1 - 2\nu)\exp(-sh_2) \\
&\omega_{10,7}(s) = \exp(sh_2), && \omega_{10,8}(s) = -(sh_2 - 2 + 2\nu)\exp(sh_2) \\
&\omega_{10,9}(s) = \exp(-sh_2), && \omega_{10,10}(s) = -(sh_2 + 2 - 2\nu)\exp(-sh_2)
\end{aligned}\right\} \tag{B.11}$$

The functions $Q_i(s,\xi)(i=1,2,...,6)$ are

$$
\begin{aligned}
Q_1(s,\xi) = & \frac{1}{2}\sum_{j=1}^{3}\Re\left[\frac{p_{31j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\exp\left(\frac{s\xi}{\gamma_{1j}^{\infty}}\right)\right] \\
& -\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\left\{\begin{array}{l}\Re\left[\dfrac{p_{31j}\delta_j}{\gamma_{1j}^2}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right] \\ -\Re\left[\dfrac{p_{31j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty 2}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\cos(\eta\xi)d\eta \\
& +\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\Im\left[\frac{p_{31j}\delta_j}{\gamma_{1j}^2}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\sin(\eta\xi)d\eta
\end{aligned}
\tag{B.12}
$$

$$
\begin{aligned}
Q_{k-2}(s,\xi) = & \frac{1}{2}\sum_{j=1}^{3}\Re\left[\frac{p_{k1j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\exp\left(\frac{s\xi}{\gamma_{1j}^{\infty}}\right)\right] \\
& +\frac{1}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\eta\left\{\begin{array}{l}\Re\left[\dfrac{p_{k1j}\delta_j}{\gamma_{1j}^2}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right] \\ -\Re\left[\dfrac{p_{k1j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty 2}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\sin(\eta\xi)d\eta \\
& +\frac{1}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\eta\Im\left[\frac{p_{k1j}\delta_j}{\gamma_{1j}^2}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\cos(\eta\xi)d\eta \quad (k=4,5)
\end{aligned}
\tag{B.13}
$$

$$
\begin{aligned}
Q_4(s,\xi) = & \frac{1}{2}\sum_{j=1}^{3}\Re\left[\frac{p_{31j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\exp\left(\frac{s(h_1-\xi)}{\gamma_{1j}^{\infty}}\right)\right] \\
& -\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\left\{\begin{array}{l}\Re\left[\dfrac{p_{31j}\delta_j}{\gamma_{1j}^2}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right] \\ -\Re\left[\dfrac{p_{31j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty 2}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\cos[\eta(h_1-\xi)]d\eta \\
& -\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\Im\left[\frac{p_{31j}\delta_j}{\gamma_{1j}^2}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\sin[\eta(h_1-\xi)]d\eta
\end{aligned}
\tag{B.14}
$$

$$Q_{k+1}(s,\xi)=\frac{1}{2}\sum_{j=1}^{3}\Re\left[\frac{p_{k1j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\exp\left(\frac{s(h_1-\xi)}{\gamma_{1j}^{\infty}}\right)\right]$$
$$+\frac{1}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\eta\left\{\begin{array}{l}\Re\left[\dfrac{p_{k1j}\delta_j}{\gamma_{1j}^{2}}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\\ -\Re\left[\dfrac{p_{k1j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty 2}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\sin[\eta(h_1-\xi)]d\eta \tag{B.15}$$
$$-\frac{1}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\eta\Im\left[\frac{p_{k1j}\delta_j}{\gamma_{1j}^{2}}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\cos[\eta(h_1-\xi)]d\eta\quad(k=4,5)$$

$$Q_7(s,\xi)=\frac{1}{2}\sum_{j=1}^{3}\Re\left[p_{61j}^{\infty}\delta_j^{\infty}\exp\left(\frac{s\xi}{\gamma_{1j}^{\infty}}\right)\right]$$
$$-\frac{s^2}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\frac{1}{\eta}\left\{\begin{array}{l}\Re\left[\dfrac{p_{61j}\delta_j}{\gamma_{1j}^{2}}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\\ -\Re\left[\dfrac{p_{61j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty 2}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\sin(\eta\xi)d\eta \tag{B.16}$$
$$-\frac{s^2}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\frac{1}{\eta}\Im\left[\frac{p_{61j}\delta_j}{\gamma_{1j}^{2}}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\cos(\eta\xi)d\eta$$

$$Q_8(s,\xi)=\frac{1}{2}\sum_{j=1}^{3}\Re\left[\frac{p_{71j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\exp\left(\frac{s\xi}{\gamma_{1j}^{\infty}}\right)\right]$$
$$-\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\left\{\begin{array}{l}\Re\left[\dfrac{p_{71j}\delta_j}{\gamma_{1j}}\dfrac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\\ -\Re\left[\dfrac{p_{71j}^{\infty}\delta_j^{\infty}}{\gamma_{1j}^{\infty}}\dfrac{1}{\eta^2+(s/\gamma_{1j}^{\infty})^2}\right]\end{array}\right\}\cos(\eta\xi)d\eta \tag{B.17}$$
$$+\frac{s}{\pi}\sum_{j=1}^{3}\int_0^{\infty}\Im\left[\frac{p_{71j}\delta_j}{\gamma_{1j}}\frac{1}{\eta^2+(s/\gamma_{1j})^2}\right]\sin(\eta\xi)d\eta$$

References

[1] Wu, C. M., Kahn, M. and Moy, W. (1996) Piezoelectric Ceramics with Functionally Gradients : A New Application in Material Design, *Journal of American Ceramics Society*, Vol.79, pp.809-812.

[2] Wang, B. L. and Noda, N. (2001) Thermally Induced Fracture of a Smart Functionally Graded Composite Structure, *Theoretical and Applied Fracture Mechanics*, Vol.35, pp.93-109.

[3] Ueda, S. (2004) Thermally Induced Fracture of a Functionally Graded Piezoelectric Layer, *Journal of Thermal Stresses*, Vol.27, pp.291-309.

[4] Ueda, S. (2007) Thermal Intensity Factors for a Parallel Crack in a Functionally Graded Piezoelectric Strip, *Journal of Thermal Stresses*, Vol.30, pp.321-342.

[5] Ueda, S. (2007) A Penny-Shaped Crack in a Functionally Graded Piezoelectric Strip under Thermal Loading, *Engineering Fracture Mechanics*, Vol.74, pp.1255-1273.

[6] Ueda, S. and Kondo, H. (2008) Transient Intensity Factors for a Parallel Crack in a Plate of a Functionally Graded Piezoelectric Material under Thermal Shock Loading Conditions, *Journal of Thermal Stresses*, Vol.31, pp.211-232.

[7] Ueda, S. (2008) A Cracked Functionally Graded Piezoelectric Material Strip under Transient Thermal Loading, *Acta Mechanica*, Vol.199, pp.53-70.

[8] Ueda, S. (2008) Transient Thermoelectroelastic Response of a Functionally Graded Piezoelectric Strip with a Penny-Shaped Crack, *Engineering Fracture Mechanics*, Vol.75, pp.1204-1222.

[9] Wang, B. L. and Mai, Y.-W. (2002) A Cracked Piezoelectric Material Strip under Transient Thermal Loading, *Transactions of the ASME, Journal of Applied Mechanics*, Vol.69, pp.539-546.

[10] Ueda, S. (2006) Transient Response of a Cracked Piezoelectric Strip under Thermoelectric Loading, *Journal of Thermal Stresses*, Vol.29, pp.973-994.

[11] Uchino, K., Yoshizaki, M., Kasai, K., Yamamura, H., Sakai, N. and Asakura, H. (1987) Monomorph Actuators using Semiconductive Ferroelectrics, *Japan Journal of Applied Physics*, Vol.26, pp.1046-1049.

[12] Miller, M. K. and Guy, W. T. (1966) Numerical Inversion of the Laplace Transform by Use of Jacobi Polynomials, SIAM *Journal of Numerical Analysis*, vol. 3, pp. 624-635.

[13] Sneddon, I. N. and Lowengrub, M. (1969) *Crack Problems in the Classical Theory of Elasticity*, John Wiley & Sons, Inc., New York.

[14] Erdogan, F. and Wu, B. H. (1996) Crack Problems in FGM Layers Under Thermal Stresses, *Journal of Thermal Stresses*, vol. 19, pp. 237-265.

[15] Erdogan, F., Gupta, G. D. and Cook, T. S. (1972) *Methods of Analysis and Solution of Crack Problems*, in G. C. Sih (ed), Noordhoff, Leyden.

[16] Ashida, F. and Tauchert, T. R. (1998) Transient Response of a Piezothermoelastic Circular Disk Under Axisymmetric Heating, *Acta Mechanica*, vol. 128, pp. 1-14.

Solving nonlinear singular boundary value problems using a newly constructed scaling function

Yanan Liu*

China Special Equipment Inspection and Research Institute，Beijing 100029, China.

*Presenting author: liuyanan@csei.org.cn
*Corresponding author: liuyanan@csei.org.cn

Abstract

In this paper, a scaling function constructed using special filter coefficients is used for solution of nonlinear singular boundary value problems. The basis functions in interval originated from the newly constructed scaling function are used in function approximation, Galerkin method and iteration approach are used for solution. Some numerical examples are given to demonstrate the validity of the technique. Numerical results prove that the new basis functions have good approximation ability and the present method is very efficient and highly accurate in solving nonlinear singular boundary value problems.

Keywords: filter coefficients, scaling functions, Galerkin method, iteration

Introduction

Many problems in applied mathematics leads to singular boundary value problems which arise in a variety of differential applied mathematics and physics such as gas dynamics, nuclear physics, chemical reaction, studies of atomic structures and atomic calculations. These problems also occur very frequently in the study of electrohydrodynamics and the theory of thermal explosions. There is a vast amount of literature on numerical solutions on singular boundary value problems. Some of the well-known techniques used in solving these problems are finite differences method[1][2], B spline method[3][4], sinc method[5], and reproducing kernel space method[6,7].

Wavelet is a powerful mathematics tool in solving many problems in science and engineering. In recent years, there has been an increasing interest in wavelet-based methods due to their successes in some applications. Wavelet-based numerical method has been developed in recent years. At present, there are mainly three kinds of wavelet-based numerical methods: wavelet finite element method[8][9], wavelet collocation method[10][11]and wavelet-Galerkin method[12][13]. In these methods, wavelet scaling functions and wavelet functions are used as basis functions in functions approximation.

The main aim of this paper is to introduce a new scaling function constructed using special filter coefficients to solve nonlinear singular boundary value problems. The basis functions in interval originated from the new scaling function are directly used to approximate the unknown functions. Using the Galerkin discretization method and iteration approach, the problem will be reduced to a set of algebraic equations. Some numerical examples are given to illustrate the stability and the effectiveness of the present method.

2. The nonlinear singular boundary value problems

In this paper, we consider following nonlinear singular boundary value problems The m degree B-spline is defined as

$$a(x)y'' + \frac{p}{x}y' + b(x)M(y) = f(x) \tag{1}$$

Subject to the boundary conditions

$$y'(0) = 0 \quad y(1) = \beta \tag{2}$$

where $0 < x < 1$, M are nonlinear functions of y, $a(x)$, $b(x)$ and $f(x)$ are given continuous functions, and p, β are finite constants.

3. Functions approximation by new scaling functions

According to the traditional theory of wavelets, the so-called scaling function $\phi(x)$ and wavelet function $w(x)$ both satisfy two-scaling relation

$$\begin{aligned}\phi(x) &= \sum_k h_k \phi(2x-k) \\ w(x) &= \sum_k g_k \phi(2x-k)\end{aligned} \tag{3}$$

The h_k and g_k are called filter coefficients. The dual scaling function $\tilde{\phi}(x)$ and wavelet $\tilde{w}(x)$ also generate a multiresolution analysis. They satisfy refinement relations like (3) with coefficients $\tilde{h}_k$ and $\tilde{g}_k$, respectively. From the theory of wavelets and filter banks, the conditions for perfect reconstruction of dual filters h_k, g_k, $\tilde{h}_k$ and $\tilde{g}_k$ can be stated as

$$\begin{aligned}h(z)\tilde{h}(z^{-1}) + g(z)\tilde{g}(z^{-1}) &= 2 \\ h(z)\tilde{h}(-z^{-1}) + g(z)\tilde{g}(-z^{-1}) &= 0\end{aligned} \tag{4}$$

where, $h(z)$ denotes the z-transform of h_k

$$h(z) = \sum_k h_k z^{-k} \tag{5}$$

The lifting scheme [14][15]demonstrates that the new filters can be constructed as follows

$$\begin{aligned}g^{\text{new}}(z) &= g(z) + h(z)s(z^2) \\ \tilde{h}^{\text{new}}(z) &= \tilde{h}(z) - g(z)s(z^{-2})\end{aligned} \tag{6}$$

and the dual lifting scheme can be expressed as

$$\begin{aligned}h^{\text{new}}(z) &= h(z) + g(z)t(z^2) \\ \tilde{g}^{\text{new}}(z) &= \tilde{g}(z) - \tilde{h}(z)t(z^2)\end{aligned} \tag{7}$$

where, $s(z)$ and $t(z)$ are the arbitary Laurent polynomials. The lifting scheme tell us that one can start with the lazy wavelet and use lifting to build filters with particular properties. We can obtain scaling functions and wavelet functions which are suitable for numerical simulation from special filters. In this paper, the filter

$$h(z) = \sum_{k=0}^{5} h_k z^k = 0.05 + 0.3z^{-1} + 0.65z^{-2} + 0.65z^{-3} + 0.3z^{-4} + 0.05z^{-5} \tag{8}$$

is used to construct the scaling functions $\phi(x)$. It is obvious that the support of $\phi(x)$ is

$$\text{supp}\phi(x) = [0,5] \tag{9}$$

From the two-scaling relation (1), the following equation can be obtained

$$\boldsymbol{\phi} = \mathbf{M}\boldsymbol{\phi} \tag{10}$$

where, $\boldsymbol{\phi}$ is a vector

$$\boldsymbol{\phi} = \left[\phi(1), \phi(2), \phi(3), \phi(4)\right]^{\mathrm{T}} \tag{11}$$

$\mathbf{M}$ is a 4 x 4 matrix

$$\mathbf{M} = \begin{bmatrix} 0.3 & 0.05 & 0 & 0 \\ 0.65 & 0.65 & 0.3 & 0.05 \\ 0.05 & 0.3 & 0.65 & 0.65 \\ 0 & 0 & 0.05 & 0.3 \end{bmatrix} \tag{12}$$

From Eq.(10) and additional condition $\phi(1)+\phi(2)+\phi(3)+\phi(4)=1$, the values of $\phi(i), i=1,2,3,4$ can be obtained. Then, $\phi(\frac{k}{2^j}),\ j,k\in\mathbb{Z}$ can be easily evaluated using the two-scaling relation (3) and $\phi(i)$. Furthermore, the $\phi'(\frac{k}{2^j}),\phi''(\frac{k}{2^j}),\ j,k\in\mathbb{Z}$ can be obtained by the similar method. Figure 1 shows the scaling function $\phi(x)$ and its first derivative and second derivative. The scaling functions constructed above can be used as basis functions to approximate the function u defined on interval $[0,1]$.

$$u(x)=\sum_{k=-4}^{i-1}c_k\phi_{i,k}(x) \tag{13}$$

where, $\phi_{i,k}(x)=\phi(ix-k)$ and i denotes the scale in approximation. The support of $\phi_{i,k}(x)$ is

$$\mathrm{supp}\,\phi_{i,k}=[\frac{k}{i},\frac{5+k}{i}] \tag{14}$$

In order to apply boundary conditions effectively, we use the boundary scaling functions in this paper. For $i\geq 5$, the left boundary scaling functions are defined as

$$\phi_{i,-4}^{\mathrm{L}}(x)=\begin{cases}\alpha_{-4}\phi_{i,-4}(x) & 0\leq x<1/i\\ 0 & \text{else}\end{cases} \tag{15}$$

$$\phi_{i,-3}^{\mathrm{L}}(x)=\begin{cases}\phi_{i,-3}(x)-\alpha_{-3}\phi_{i,-4}(x) & 0\leq x<1/i\\ \phi_{i,-3}(x) & 1/i\leq x<2/i\\ 0 & \text{else}\end{cases} \tag{16}$$

$$\phi_{i,-2}^{\mathrm{L}}(x)=\begin{cases}\phi_{i,-2}(x)-\alpha_{-2}\phi_{i,-4}(x) & 0\leq x<1/i\\ \phi_{i,-2}(x) & 1/i\leq x<3/i\\ 0 & \text{else}\end{cases} \tag{17}$$

$$\phi_{i,-1}^{\mathrm{L}}(x)=\begin{cases}\phi_{i,-1}(x)-\alpha_{-1}\phi_{i,-4}(x) & 0\leq x<1/i\\ \phi_{i,-1}(x) & 1/i\leq x<4/i\\ 0 & \text{else}\end{cases} \tag{18}$$

$$\alpha_{-4}=1+\alpha_{-3}+\alpha_{-2}+\alpha_{-1},\quad \alpha_{-3}=\frac{\phi(3)}{\phi(4)},\quad \alpha_{-2}=\frac{\phi(2)}{\phi(4)},\quad \alpha_{-1}=\frac{\phi(1)}{\phi(4)} \tag{19}$$

The right boundary scaling functions $\phi_{i,i+k}^{\mathrm{R}}(x),k=-4,-3,-2,-1$ can be constructed by similar method. Then in approximation Eq.(15), the ordinary scaling functions $\phi_{i,k}(x)$ and $\phi_{i,i+k}(x),k=-4,-3,-2,-1$ are respectively replaced by boundary scaling functions $\phi_{i,k}^{\mathrm{L}}(x)$ and $\phi_{i,i+k}^{\mathrm{R}}(x),k=-4,-3,-2,-1$. For the sake of uniform expressions, the subscript L and R in boundary scaling functions will omit in the following part. The four left boundary scaling functions with $i=5$ are shown in Figure 2.

4. Numerical implementation

In order to solve the Eq.(1), we can approximate the functions $y(x)$ as follows

$$y(x)=\sum_{k=1}^{i+4}c_k\phi_k(x) \tag{20}$$

where $\phi_k(x)$ are corresponding to $\phi_{i,k-5}(x)$ discussed in section 3. Introducing (20) into Eq.(1) and using Galerkin discretization method, we have

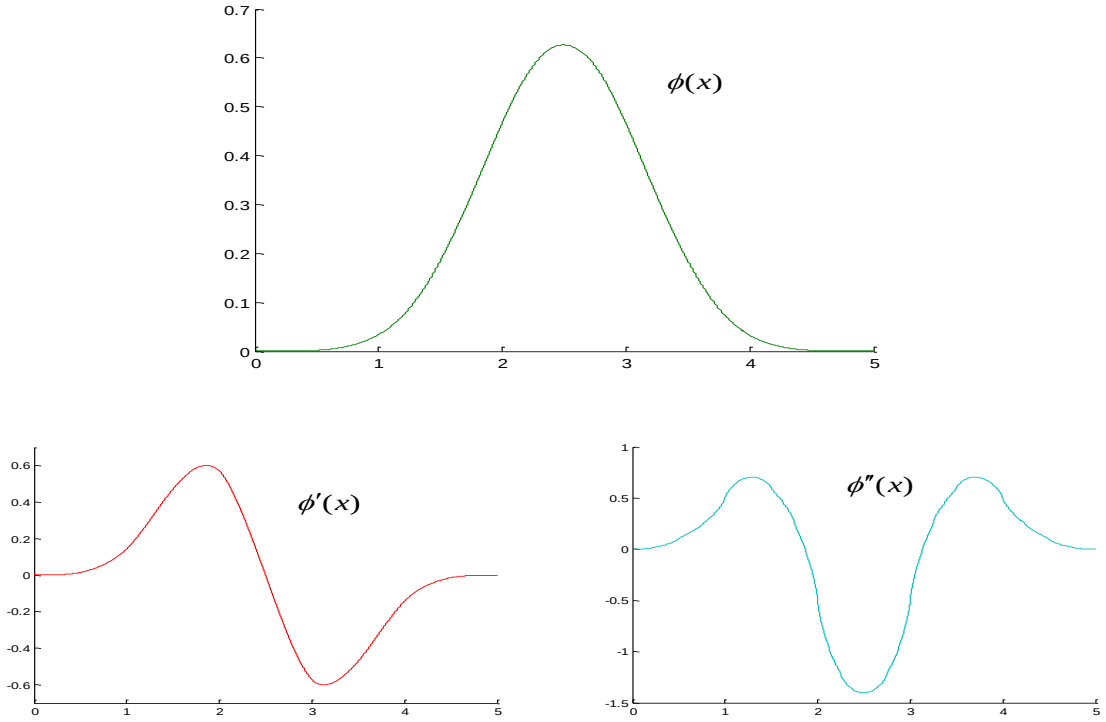


Figure 1. The basic scaling function $\phi(x)$ and its first and second derivative

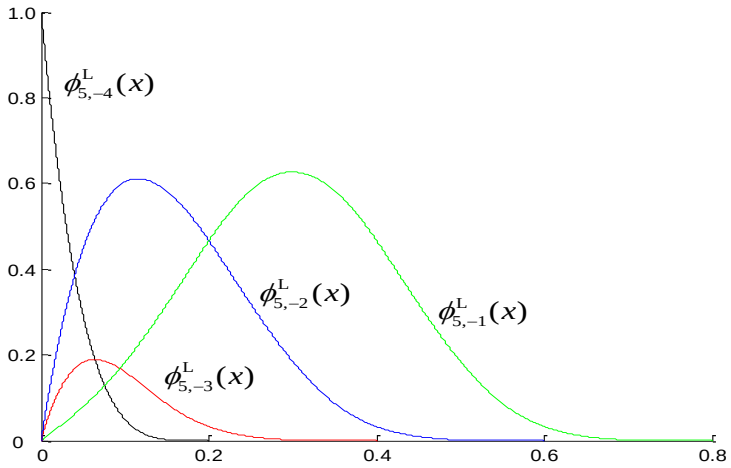


Figure 2. The four left boundary scaling functions with $i=5$

$$\int_0^1 (a(x)\sum_{k=1}^{i+4} c_k\phi_k''(x) + \frac{p}{x}\sum_{k=1}^{i+4} c_k\phi_k'(x))\phi_j(x)dx = \int_0^1 (f(x) - b(x)M(y))\phi_j(x)dx \tag{21}$$

where $j = 1, 2, \cdots i+4$. From (21), we can obtain a set of algebraic equations

$$\mathbf{Kc} = \mathbf{f} \tag{22}$$

where

$$\mathbf{K}(i,j) = \int_0^1 (a(x)\phi_i''(x) + \frac{p}{x}\phi_i'(x))\phi_j(x)dx \quad i, j = 1, 2, \cdots i+4 \tag{23}$$

$$\mathbf{c}=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{i+2} \end{bmatrix} \tag{24}$$

$$\mathbf{f}(j)=\int_0^1 (f(x)-b(x)M(y))\phi_j(x)dx, \quad j=1,2,\cdots i+4 \tag{25}$$

Because there is nonlinear part in Eq. (22), we should use iteration approach for solution. In this case, we have

$$\mathbf{K}\mathbf{c}^{(n+1)}=\mathbf{f}^{(n)} \quad n=0,1,\cdots \tag{26}$$

where, n is the iteration number and

$$\mathbf{c}^{(0)}=[\mathbf{0}] \tag{27}$$

The computation of $\mathbf{f}^{(n)}$ is as follows

$$\mathbf{f}^{(n)}(j)=\int_0^1 (f(x)-b(x)M(y^{(n)}))\phi_j(x)dx \tag{28}$$

and

$$y^{(n)}(x)=\sum_{k=1}^{i+4} c_k^{(n)}\phi_k(x) \tag{29}$$

5 Numerical examples

In this section, we will apply the present new method to solve some nonlinear singular boundary value problems. The computed results are compared with the exact solutions.

Example1

$$-y''-\frac{2}{x}y'+y^2=x^4-2x^2+7 \qquad 0<x<1 \tag{30}$$

Subject to the boundary conditions

$$y'(0)=0 \quad y(1)=0 \tag{31}$$

The exact solution of (30) is $y(x)=1-x^2$. Table 1 shows the comparison of exact solution and numerical results of $y(x)$. The scale used in approximation is $i=30$, and the iteration number is $n=8$. It can be found that the results evaluated by present method are highly accurate.

Example2

$$y''+\frac{2}{x}y'+y^5=0 \qquad 0<x<1 \tag{32}$$

Subject to the boundary conditions

$$y'(0)=0 \quad y(1)=\frac{\sqrt{3}}{2} \tag{33}$$

The exact solution of (32) is $y(x)=1/\sqrt{1+\frac{x^2}{3}}$. Table 2 shows the comparison of exact solution and numerical results of $y(x)$. The scale used in approximation is still $i=30$. The iteration number is $n=10$, and the numerical results are highly accurate.

Table 1. The comparison of exact and numerical results of $y(x)$ for example 1

x	Exact results	Numerical results	Absolute error
0	1.0	0.99993999595962	6.0E-005
0.1	0.99	0.98995337910760	4.7E-005
0.2	0.96	0.95996029190128	4.0E-005
0.3	0.91	0.90996593278944	3.4E-005
0.4	0.84	0.83997127834294	2.9E-005
0.5	0.75	0.74997635685368	2.4E-005
0.6	0.64	0.63998121355224	1.9E-005
0.7	0.51	0.50998591179173	1.4E-005
0.8	0.36	0.35999053476071	9.5E-006
0.9	0.19	0.18999518777879	4.8E-006
1.0	0.0	0.0	0.0

Table 2. The comparison of exact and numerical results of $y(x)$ for example 2

x	Exact results	Numerical results	Absolute error
0	1.0	0.99995346837571	4.7E-005
0.1	0.9983375	0.99829788110237	4.0E-005
0.2	0.9933993	0.99336315548509	3.6E-005
0.3	0.9853293	0.98529691402213	3.2E-005
0.4	0.9743547	0.97432681393413	2.8E-005
0.5	0.9607689	0.96074595606071	2.3E-005
0.6	0.9449112	0.94489331093333	1.8E-005
0.7	0.9271455	0.92713268820308	1.3E-005
0.8	0.9078413	0.90783318696177	8.1E-006
0.9	0.8873565	0.88735271237105	3.8E-006
1.0	0.8660254	0.86602540378444	0.0

Conclusions

In this paper, a scaling function constructed using special filter coefficients is used for solution of nonlinear singular boundary value problems. The basis functions in interval originated from the new scaling function are directly used in function approximation, and the Galerkin discretization method and iteration approach are used for solution. Numerical results demonstrate that the new basis functions are suitable for numerical simulation and the present solution method is very efficient and highly accurate in solving nonlinear singular boundary value problems.

Acknowledgements

This paper was supported by the National Natural Science Foundation of China (11272358).

References

[1] Kanth, A.S.V.R. and Reddy, Y.N. (2004) Higher order finite difference method for a class of singular boundary value problems, *Appl. Math. Comput.* 155, 249–258.
[2] Chawla, M.M.and Katti, C.P. (1984) A finite difference method for a class of singular two point boundary value problems, *IMA. J. Numer.Anal.* 4, 457–466.
[3] Kanth, A.S.V.R. and Reddy, Y.N. (2005) Cubic spline for a class of singular two-point boundary value problems, *Appl. Math. Comput.* 170, 733–740.
[4] Kadalbajoo, M.K. and Aggarwal, V.K. (2005) Numerical solution of singular boundary value problems via Chebyshev polynomial and B-spline, *Appl. Math. Comput.* 160, 851–863.
[5] Mohsen, A. and El-Gamel, M. (2008) On the Galerkin and collocation methods for two-point boundary value problems using sinc bases, *Comput. Math. Appl.* 56(4), 930-941
[6] Cui, M. and Geng, F. (2007) Solving singular two-point boundary value problem in reproducing kernel space, *J. Comput. Appl. Math.* 205, 6–15.
[7] Wang, W., Cui, M. and Han, B. (2008) A new method for solving a class of singular two-point boundary value problems, *Appl.Math. Comput.* 206(2) , 721-727.
[8] Ko, J., Kurdila, A.J. and Pilant, M.S. (1995) A class of finite element methods based on orthonormal compactly supported wavelets, *Comput. Mech.* 16, 235–244.
[9] Han, J.G., Ren, W.X. and Huang, Y. (2006) A spline wavelet finite-element method in structural mechanics, *Int. J. Numer. Methods Engrg.* 66, 166–190
[10] Bertoluzza, S. and Naldi, G. (1996) A wavelet collocation method for the numerical solution of partial differential equations, *Appl. Comput. Harmonic Anal.* 3, 1–9.
[11] Reckinger, S.M., Vasilyev, O.V. and Fox-Kemper, B. (2014) Adaptive wavelet collocation method on the shallow water model, *J. Comput. Phys.* 271(9), 342-359.
[12] Amaratunga, K. and Williams, J.R. (1994) Wavelet-Galerkin solutions for one dimensional partial differential equation, *Int. J. Numer. Methods Engrg.* 37(16), 2703－2716.
[13] Yousefi, M.R., Jafari, R. and Moghaddam, H.A. (2014) Imposing boundary and interface conditions in multi-resolution wavelet Galerkin method for numerical solution of Helmholtz problems, *Comput. Methods Appl. Mech. Engrg*. 276(7), 67-94
[14] Sweldens, W. (1996) The Lifting Scheme: A Custom-Design Construction of Biorthogonal Wavelets. *Appl. Comput. Harmonic Anal.* 3(2), 186-200.
[15] Daubechies, I. and Sweldens, W. (1998) Factoring wavelet transforms into lifting steps, *J. Fourier Anal. Appl.* 4(3), 247-269.

Biot's Consolidation analysis using the node-based smoothed point interpolation method (NS-PIM)

Shiyang Pan[1], Tongchun Li[1], †Jing Cheng[1, 2], Ping Yuan[3], and Xinyang Ning[1]

[1] College of Water Conservancy and Hydropower Engineering, Hohai University, Nanjing 210098, Peoples Republic of China

[2] 210 KAP Hall, University of Southern California, Los Angeles, CA 90089, U.S.A

[3] Zhongye Changtian International Engineering Co., Ltd , Changsha 410007, Peoples Republic of China

†Corresponding author: mscj042@hotmail.com.

ABSTRACT

The node-based smoothed point interpolation method (NS-PIM) is developed for soil consolidation analysis based on the Biot's theory. Both the shape functions for displacements and pore pressures are constructed using the point interpolation method which is easier to be programmed than the other meshless methods. Then, a 2D consolidation problem under ramp load is solved. The results show good agreement with those certified results. Meanwhile, convergence features of different solutions are studied and useful convergence features are found. Thus a simple method is introduced to estimate the errors of the model with rough grids. Obviously, it is promising to apply NS-PIM to the analysis of consolidation problems.

Keywords: Meshless Methods, NS-PIM, Consolidation, Biot's theory, FEM

1 INTRODUCTION

Consolidation analysis is necessary in many situations such as foundation settlement analysis, earth and rock fill dam analysis and so on.[1][2] Generally, the consolidation analysis in one dimensional field is based on the Terzaghi theory[3] in soil mechanics, while the Biot's consolidation theory[4][5] is usually applied in two dimensional and three dimensional fields.

The finite element method[6] has been developed to solve lots of problems and great achievements has been made[7]. With the development of finite element method and the improvement of computer performance, numerous programs using the Biot's theory based on the finite element method have been developed to solve practical problems in consolidation analysis. Those programs using the finite element method are relatively easy to be programmed and the results can be checked in many commercial soft wares and monographs[8]. However, the building and mesh of models takes place of most of the time for the analysis of FEM solution. Many mistakes may also take place during the building of the models. And it is also hard to get accurate results, unless the dense of the grid is sufficient.

Thus, the mesh-free(or meshless) method[9] was put forward to deal with the problems mentioned above. And there are many kinds of element-free or mesh-free methods, such as reproducing kernel particle method (RKPM), HP-cloud method, point interpolation method (PIM) and so on. The node-based smoothed point interpolation method (NS-PIM or LC-PIM) [10][11], as one of the mesh free methods, has been developed using the node-based strain smoothing operation[12]. Formulated by polynomial PIM [13] or radial PIM (RPIM) [14] shape functions, it possesses the Kronecker delta property and the boundary can be enforced like that of the FEM solution[15], so that the programming of the NS-PIM can be easier than the programming of many other mesh free methods.

It has been certified that the NS-PIM and FEM solutions possess different convergence features[16][17]. And it is significant to apply the NS-PIM solution to consolidation problem in order to get more accurate results with the comparison to those by the FEM solution[18]. Also, the adaptability of mesh free methods to large deformation as the shape functions are based on nodes is meaningful for consolidation problems like foundation settlement.

Recent years, some PIM programs for elastic problem and consolidation problem[19] have been developed, but NS-PIM solution is rarely applied to consolidation problems. With all the features mentioned above, it is meaningful and necessary to apply NS-PIM to the analysis of consolidation problems and do some further research.

In this work, the node-based smoothed point interpolation method(NS-PIM) is built and coded for Biot's consolidation analysis. Then, the NS-PIM solution is applied to a classic example to check the correctness and validation together with the FEM solution. It is confirmed that the results of the NS-PIM solution fit well with the certified FEM results. So that it is feasible to apply the NS-PIM solution to the Biot's consolidation analysis. Further more, different mesh schemes are applied to investigate the convergence features of different solutions. The investigation shows that the NS-PIM and FEM solutions still possess the certified convergence features which can be used as a method to get more accurate results. So that a simple method is introduced to estimate the errors of the results with rough grids.

2 BRIEFING ON FORMULATIONS AND SHAPE FUNCTIONS

Some basic formulations and shape functions for NS-PIM are introduced in this part. The programming of the solution is also introduced.

2.1 Basic equations for Biot's theory

The basic theory used to write the NS-PIM program is the Biot's consolidation theory .

On the one hand, the sets of partial differential equations for two dimensional equilibrium in consolidation problems are listed as follows :

$$\begin{aligned} \frac{\partial \sigma_x'}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial u_w}{\partial x} &= f_x \\ \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_y'}{\partial y} + \frac{\partial u_w}{\partial y} &= f_y \end{aligned} \tag{1}$$

where u_w is the excess pore pressure ,τ_{xy} is the shearing stress , f_x f_y are the forces in x- , y- directions and σ_x', σ_y' are the effective stresses.

$$\{\sigma'\} = [D]\{\varepsilon\} \tag{2}$$

where $[D]$ is the constitutive matrix and $\{\varepsilon\}$ is the strain matrix.

In linear elastic problem, the constitutive matrix $[D]$ is an elastic matrix. Then,

$$\begin{aligned} \sigma_x' &= \frac{E'}{(1+\upsilon')}\left(\frac{\upsilon'}{(1-2\upsilon')}\varepsilon_v + \varepsilon_x\right) \\ \sigma_y' &= \frac{E'}{(1+\upsilon')}\left(\frac{\upsilon'}{(1-2\upsilon')}\varepsilon_v + \varepsilon_y\right) \\ \tau_{xy} &= \frac{E'}{2(1+\upsilon')}\gamma_{xy} \end{aligned} \tag{3}$$

where ε_v is the volumetric strain, ε_x ε_y are the x- y- strain, γ_{xy} is the shearing strain, E' is the effective Young's modulus and v' is the poisson's ratio.

Further more, assuming small strains , the geometric equation is

$$\{\varepsilon\} = [\partial]\{u\} \tag{4}$$

where $\{\varepsilon\}$ is the strain matrix , $\{u\}$ is the displacement matrix and $[\partial]$ is the partial differential operator matrix .

Thus, equation (1) can be rewritten as follows:

$$\begin{aligned}
&\frac{\mathrm{E}'(1-\upsilon')}{(1+\upsilon')(1-2\upsilon')}\left[\frac{\partial^2 u_x}{\partial x^2}+\frac{(1-2\upsilon')}{2(1-\upsilon')}\frac{\partial^2 u_x}{\partial y^2}+\frac{1}{2(1-\upsilon')}\frac{\partial^2 u_y}{\partial x\,\partial y}\right]+\frac{\partial \mathrm{u}_w}{\partial x}=f_x \\
&\frac{\mathrm{E}'(1-\upsilon')}{(1+\upsilon')(1-2\upsilon')}\left[\frac{1}{2(1-\upsilon')}\frac{\partial^2 u_x}{\partial x\,\partial y}+\frac{\partial^2 u_y}{\partial y^2}+\frac{(1-2\upsilon')}{2(1-\upsilon')}\frac{\partial^2 u_y}{\partial x^2}\right]+\frac{\partial \mathrm{u}_w}{\partial y}=f_y
\end{aligned} \tag{5}$$

where u_x is the x-displacement and u_y is the y-displacement .

On the other hand, assuming fluid incompressibility and 2D continuity ,the continuity equation takes the following form :

$$\frac{\partial \varepsilon_v}{\partial t}+\frac{k}{\gamma_w}\nabla^2 u_w=0 \tag{6}$$

where $\varepsilon_v=\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}$ is the volumetric strain , t is the time, k is the matrix of permeabilities in the x- and y-directions , γ_w is the unit weight of water and $\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$ is the Laplace operator .

Now ,the coupled 'Biot' equations for a 2D homogeneous poroelastic material can be listed as follows by combining equations (5) - (6) :

$$\begin{aligned}
&\frac{\mathrm{E}'(1-\upsilon')}{(1+\upsilon')(1-2\upsilon')}\left[\frac{\partial^2 u_x}{\partial x^2}+\frac{(1-2\upsilon')}{2(1-\upsilon')}\frac{\partial^2 u_x}{\partial y^2}+\frac{1}{2(1-\upsilon')}\frac{\partial^2 u_y}{\partial x\,\partial y}\right]+\frac{\partial \mathrm{u}_w}{\partial x}=f_x \\
&\frac{\mathrm{E}'(1-\upsilon')}{(1+\upsilon')(1-2\upsilon')}\left[\frac{1}{2(1-\upsilon')}\frac{\partial^2 u_x}{\partial x\,\partial y}+\frac{\partial^2 u_y}{\partial y^2}+\frac{(1-2\upsilon')}{2(1-\upsilon')}\frac{\partial^2 u_y}{\partial x^2}\right]+\frac{\partial \mathrm{u}_w}{\partial y}=f_y \\
&\frac{\partial}{\partial t}\left(\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}\right)+\frac{k_x}{\gamma_w}\frac{\partial^2 u_w}{\partial x^2}+\frac{k_y}{\gamma_w}\frac{\partial^2 u_w}{\partial y^2}=0
\end{aligned} \tag{7}$$

2.2 Discretization of Biot's theory

To solve the set of simultaneous equations in equation (7), the displacement variables u_x,u_y and excess pore pressure u_w need to be discretized in the equations showed as follows ,

$$\begin{aligned}
\bar{u}_x &= [N]\{u_x\} \\
\bar{u}_y &= [N]\{u_y\} \\
\bar{u}_w &= [N]\{u_w\}
\end{aligned} \tag{8}$$

where $\bar{u}_x, \bar{u}_y$ are the element or nodal displacements in x- , y- directions , $\bar{u}_w$ is the element or nodal excess pore pressure and $[N]$ is the corresponding shape function matrix.

In practice, it may be more appropriate to use higher order shape functions for u_x and u_y to avoid oscillation[20]. But, in this work, all three variables are described by the same order of shape functions which means the solution in this work may have the problem of oscillation under some extreme conditions.

After all the discretization and Galerkin process , equation (7) can be rewritten as follows :

$$\begin{aligned} [k_m]\{u\} + [c]\{u_w\} &= \{f\} \\ [c]^T \left\{\frac{du}{dt}\right\} - [k_p]\{u_w\} &= \{0\} \end{aligned} \tag{9}$$

where $[k_m]$ is the elastic stiffness matrix , $[k_p]$ is the fluid conductivity matrix and $[c] = \iint \frac{\partial N_i}{\partial x} N_j \, dxdy$ is the coupling matrix where N_i is the shape function of the displacement filed and N_j is the shape function of the excess pore pressure filed .

The following equations for writing the program can be obtained by integrating Equation (9) in time and doing interpolation using θ, and then Equation (10) can be obtained :

$$\begin{aligned} &\begin{bmatrix} \theta[k_m] & \theta[c] \\ \theta[c]^T & -\theta^2 \Delta t[k_p] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{u_w\} \end{Bmatrix}_1 \\ &= \begin{bmatrix} -(1-\theta)[k_m] & -(1-\theta)[c] \\ \theta[c]^T & \theta(1-\theta)\Delta t[k_p] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{u_w\} \end{Bmatrix}_0 + \begin{Bmatrix} (1-\theta)f \\ \{0\} \end{Bmatrix}_0 + \begin{Bmatrix} \theta f \\ \{0\} \end{Bmatrix}_1 \end{aligned} \tag{10}$$

where θ is the time-stepping parameter ($0.5 \leq \theta \leq 1$). To avoid oscillatory results ,the fully implicit method with θ = 1 is used in the NS-PIM solution. In this work, the left side matrix $\begin{bmatrix} \theta[k_m] & \theta[c] \\ \theta[c]^T & -\theta^2 \Delta t[k_p] \end{bmatrix}$ in Equation (10) is called 'Kt' and the right side matrix $\begin{bmatrix} -(1-\theta)[k_m] & -(1-\theta)[c] \\ \theta[c]^T & \theta(1-\theta)\Delta t[k_p] \end{bmatrix}$ is called 'Kd'. Also, in equation (10)

$$\begin{aligned} k_m &= \iint [B]^T [D][B] \, dxdy \\ k_p &= \iint [T]^T [K][T] \, dxdy \end{aligned} \tag{11}$$

where $[B]$ is the strain–displacement matrix , $[D]$ is the stress–strain matrix, $[T]$ is the '$[B]$ matrix' for fluid and $[K]$ is the '$[D]$ marix' for fluid.

2.3 Simulate load in construction period

The load is usually applied step by step in a construction period. To simulate this process, the total load should be divided into several part according to the construction period. The previous load {Rtp}, the current load {Rtc} and the effective stress {Rp} generated by the previous load is needed to calculate the load {Rc} for the current time. And

$$\{Rc\} = \{Rtc\} + (\{Rtp\} - \{Rp\}) \tag{12}$$

This load {Rc} would be the external force applied in current time step. Thus, the simulation is realized.

2.4 Shape functions for NS-PIM

Triangular cells are used for the mesh of models, since the triangular cells are more adaptive to complex geometry. Polynomials which are usually built utilizing Pascal triangle are used in the interpolation to create shape functions for NS-PIM. The Figure 1 is the Pascal triangle of which the first two lines are used for complete polynomial basis of first order and the first three lines are used for complete polynomial basis of second order in two-dimensional domain, as is listed below.

$$p^T(x) = \{1 \ \mathrm{x} \ \mathrm{y}\} \tag{13}$$

$$p^T(x) = \{1 \ \mathrm{x} \ \mathrm{y} \ \mathrm{x}^2 \ \mathrm{xy} \ \mathrm{y}^2\} \tag{14}$$

where $p^T(x)$ is the complete polynomial basis mentioned above , x y are the coordinates of each node.

$$1$$

$$\mathrm{x} \quad \mathrm{y}$$

$$\mathrm{x}^2 \ \mathrm{xy} \ \mathrm{y}^2$$

$$\mathrm{x}^3 \ \mathrm{x}^2\mathrm{y} \ \mathrm{xy}^2 \ \mathrm{y}^3$$

$$\cdots\cdots\cdots\cdots$$

Figure 1. Pascal triangle for two-dimensional domains

The vector of shape functions is

$$\Phi^T(\mathrm{x}) = p^T(x)P_n^{-1}(x) = \{\varphi_1 \ \varphi_2 \ \cdots \ \varphi_n\} \tag{15}$$

where n is the number of nodes, φ is the shape function of each node, $P_n^{-1}(x)$ is the moment matrix and

$$P_n(x) = \begin{bmatrix} 1 & x_1 & y_1 & x_1y_1 & \cdots & p_n(x_1) \\ 1 & x_2 & y_2 & x_2y_2 & \cdots & p_n(x_2) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & y_n & x_ny_n & \cdots & p_n(x_n) \end{bmatrix} \tag{16}$$

2.4.1 Supporting nodes for interpolation

For NS-PIM, there are two schemes available to create the needed shape functions, T3-scheme and T6/3-scheme [21]. The T3-scheme is a scheme utilizing the vertexes of each cell to create the shape functions. In the T6/3-scheme, the vertexes of an interior cell that has no edge on the boundary of the problem domain and its neighboring cells sharing one edge with this interior cell are needed to create the shape functions, while the vertexes of a boundary cell that has at least one edge on the boundary are enough to create the corresponding shape functions .

As is showed in Figure 2 , cell 1 (ABD) is a boundary cell and cell 2 (BPD) is an interior cell. Node A, B, D would be enough to create shape functions of cell 1 and cell 2 for the T3-scheme of NS-PIM. However, six nodes (node A, B, D, E, F, P) are needed to create the shape functions of cell 2 for the T6/3-scheme of NS-PIM.

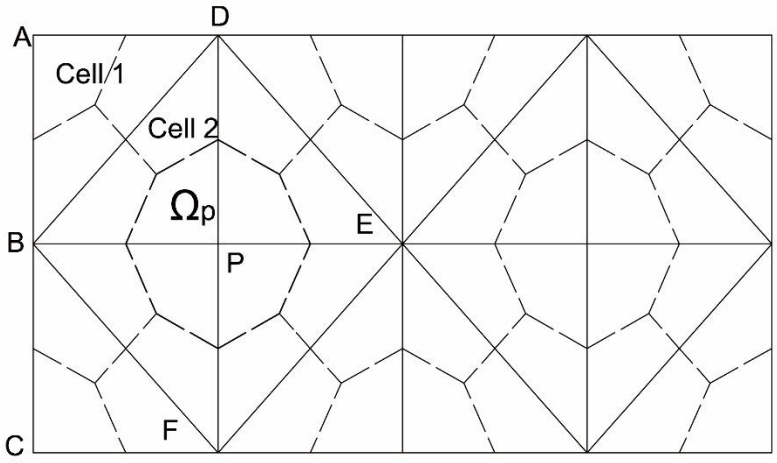


Figure 2. Background cells and the construction of smoothing cells

2.4.2 Integral domain

Shape functions for NS-PIM can be constructed by the nodes attached to the cells. In practice, to ensure the continuity of the shape functions ,the problem domain of NS-PIM needs to be divided into smoothing domains based on background cells [22].

As shown in Figure 2, the solid lines is the lines for the background cells and the smoothing domains are constructed by the dotted lines that is connecting the mid-edge-point to the centroids of the cells. For example, in Figure 2 ,the sub-domain Ω_p is the domain for the point P.

2.5 Programming of the NS-PIM solution

The consolidation analysis program for NS-PIM solution is based on the Fortran90 language and applicable for linear elastic Biot's consolidation analysis.

The programming of the NS-PIM solution is best explained by the flowchart of Figure 3. There are two main loops in this program: node-looping and time-looping. Many global matrices of equation (10) are formed in the node-looping and the results of each loading step can be obtained in the time-looping. **Table** I is the list of some main subroutines and their functions.

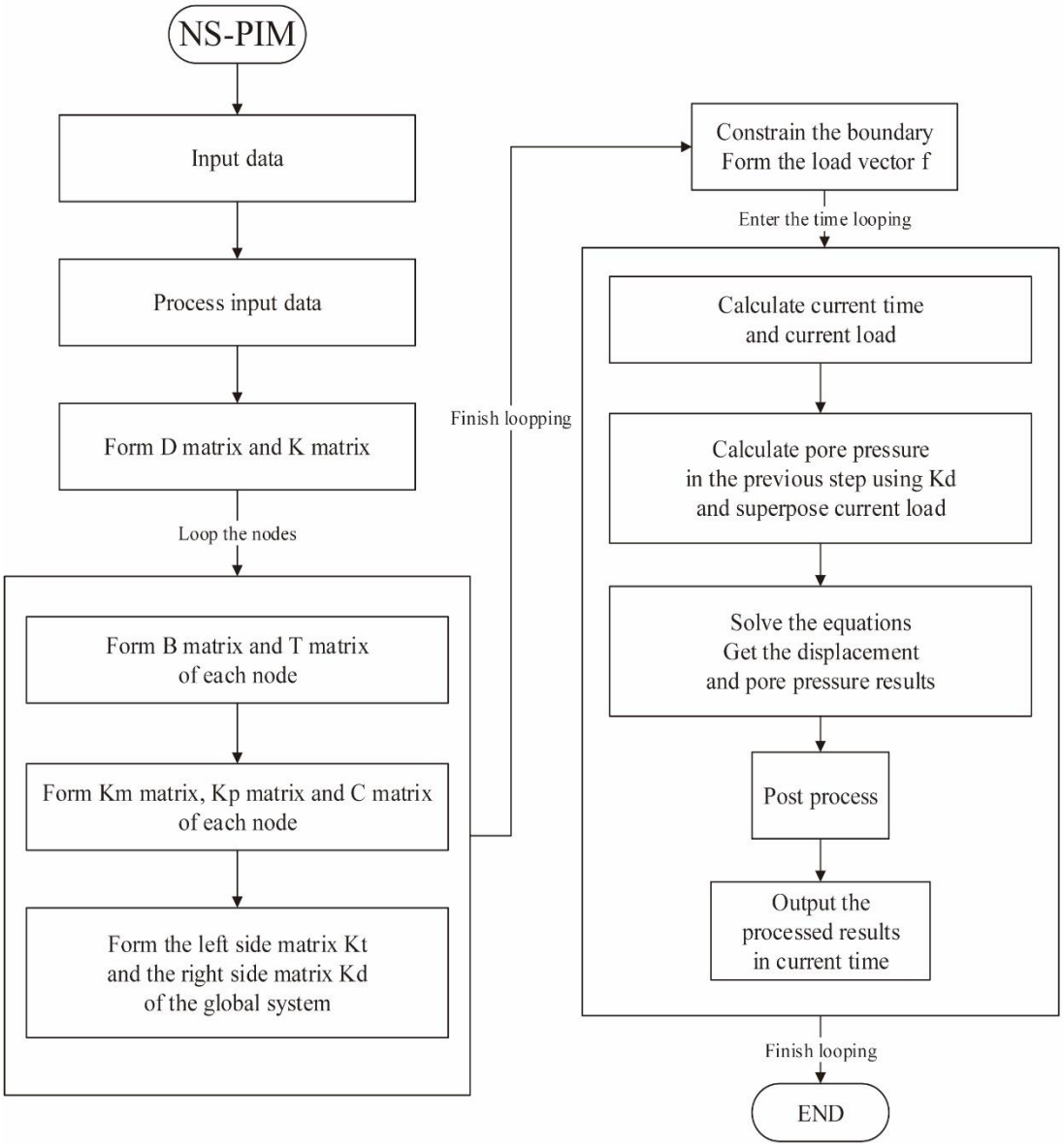


Figure 3. Flowchart of NS-PIM program

Table I. Main subroutines and functions

Subroutine Name	Function
D_Form, K_Form	Form the D matrix and K matrix
FormBCT	Form the B T and C matrices of each node
Stiff_point_Km, Stiff_point_Kp	Form the Km and Kp matrices of each node
Stiff_point_KN	Form the left side matrix of each node
Stiff_point_KD	Form the right side matrix of each node
Form_BG	Form the global left side matrix
Form_Pb	Form the global right side matrix
Traction	Form the total load vector
EBCs	Constrain the boundary
EquaSolvBand	Solve the equations
Tecout_mult	Output the results

3 NUMERICAL EXAMPLES

Firstly, a classic 2D Biot elastic solid example is applied to certify the correctness and validation of the PIM solution and T3-scheme is applied to create the shape functions for NS-PIM. Then, the mesh of the same model is refined to investigate the convergence feature of the NS-PIM and FEM solutions.

3.1 Classic 2D consolidation problem

3.1.1 Model and Material parameters

A classic homogeneous 2D Biot elastic solid example is simulated in this part[8]. For convenience, the permeabilities in the x- and y-directions are considered numerically equal to the unit weight of water which means the effect of Mandal would not be observed easily. Also, the effective Young's modulus and the poisson's ratio is taken respectively as 1 and 0.

A plane strain consolidation model is built in this problem, as shown by the mesh and constraints given in Figure 4. The width and depth of the model is 1m and 10m. The base and sides of the mesh are impermeable boundaries and both restrained in the normal direction of the boundaries. The top of the model is drained, and subjected to the ramp loading shown in Figure 5 ,which indicates a linearly increasing load reaching a maximum of 1.0 at time t_0. In this case, there are ten load steps, twenty time steps, and the duration of each step is one second.

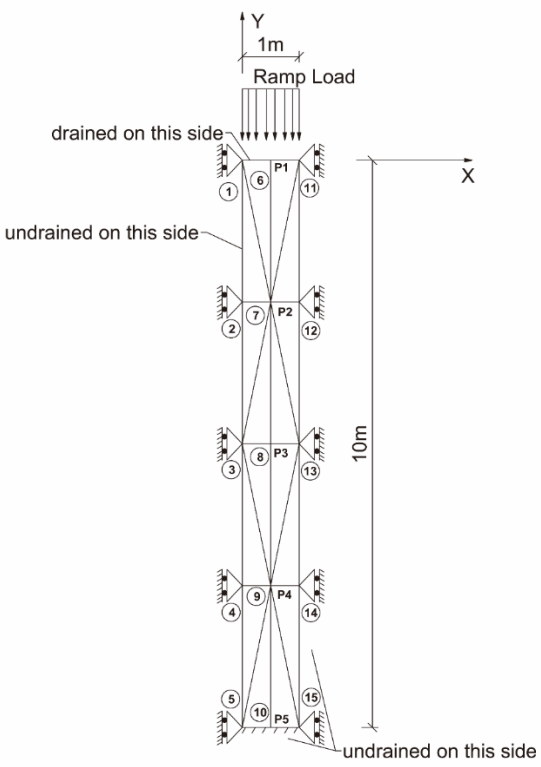


Figure 4. Mesh and constraints of the 2D Biot elastic solid model

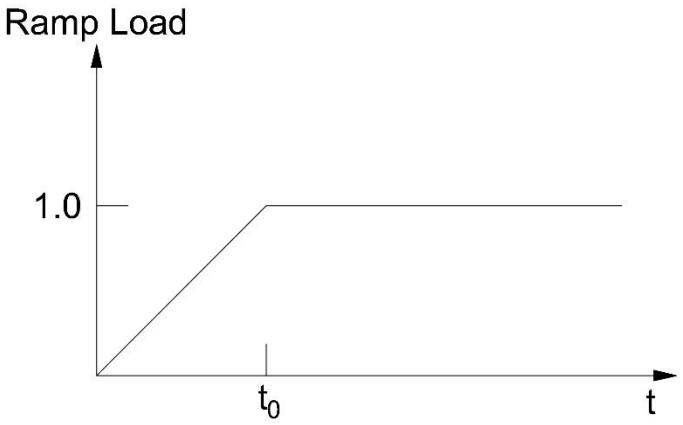


Figure 5. Ramp loading

3.1.2 Distributions of the results

The distributions of Y-displacement and pore pressure results using NS-PIM solution are calculated and the results of step 1, step 5, step 10, step 11, step 15 and step 20 are shown in Figure 6 and Figure 7.

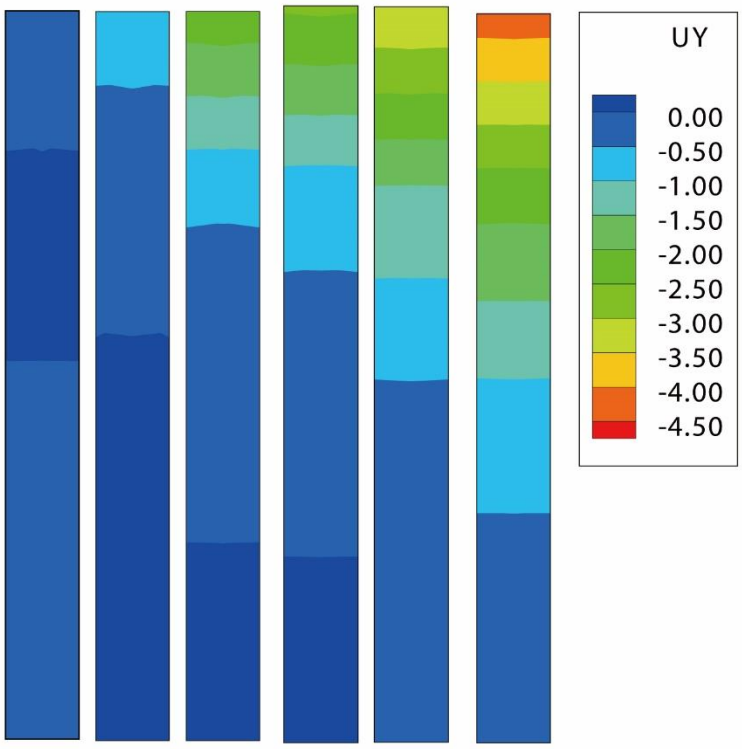


Figure 6 Distributions of Y-Displacement by NS-PIM(T3) in different steps (mm)

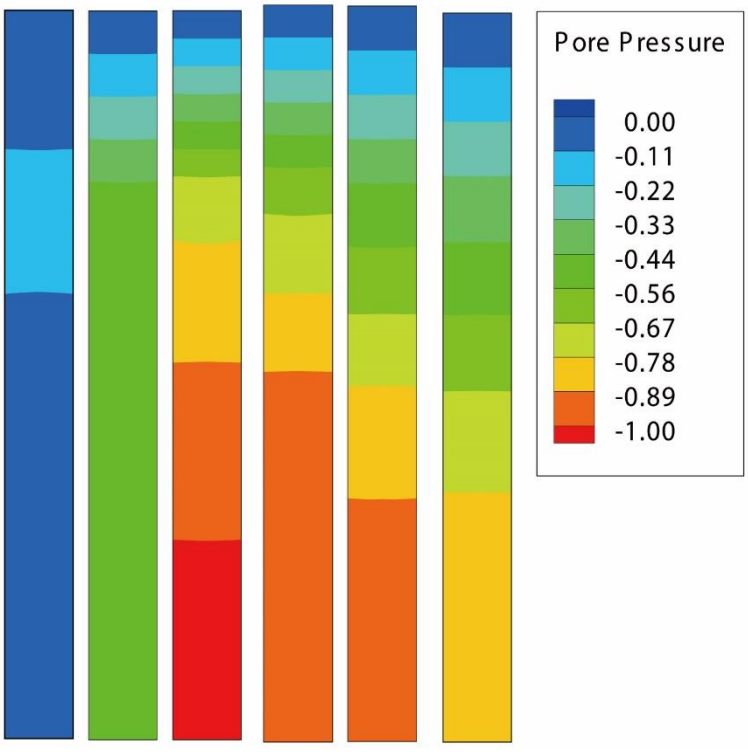


Figure 7 Distributions of pore pressure by NS-PIM(T3) in different steps (Pa)

Extracting the Y-displacement of different steps from top to bottom(‘P1’-‘P5’) to plot a graph of the distribution of Y-displacement by different solutions as is showed in Figure 8 and Figure 9.

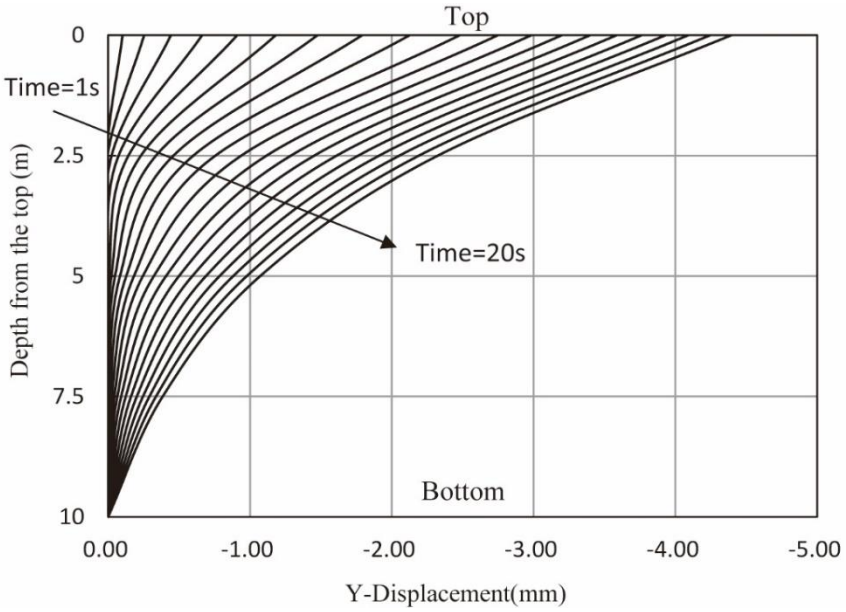


Figure 8 Distributions of Y-Displacement at different times by NS-PIM(T3)

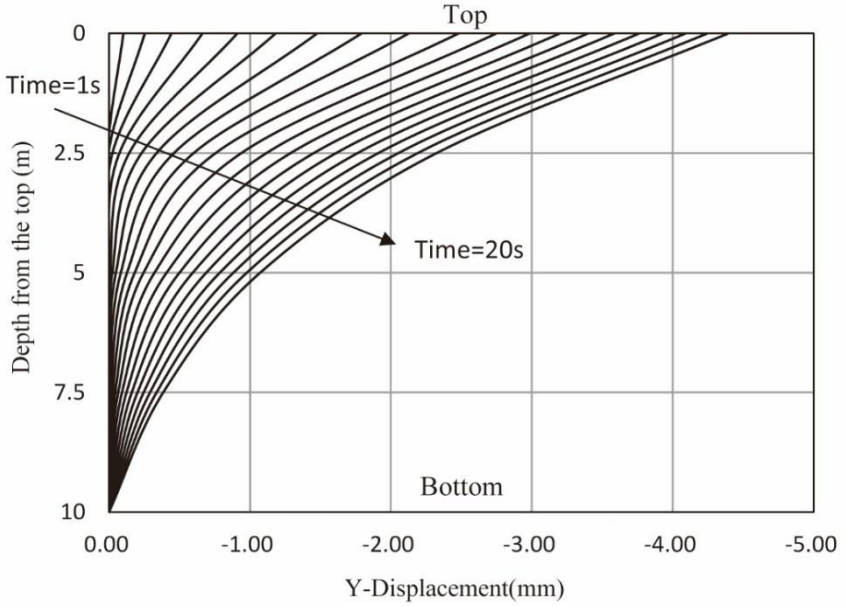


Figure 9 Distributions of Y-Displacement at different time by FEM

As is described in the figures above, the Y-displacement values by both solutions increase gradually from bottom to top, while the pore pressure values increase gradually from top to bottom. Also, the Y-displacement values at each point keep increasing, while the pore pressure values first increase then decrease with the increase of time. Comparing with the general rules, the distributions of the results obtained by both the FEM solution and the NS-PIM solution described above are reasonable.

3.1.3 Time history of the results

In order to study the time-dependent change of the results obtained by NS-PIM, five points from top to bottom ('P1'-'P5') in the middle of the model are selected to plot the time series graphs of the Y-Displacement results and pore pressure results. The graphs are shown in Figure 10 and Figure 11.

To take a close investigation of different solutions, combining the displacement values of 'P1' and the pore pressure values of 'P5' by NS-PIM and FEM. The graphs are shown in Figure 12 and Figure 13.

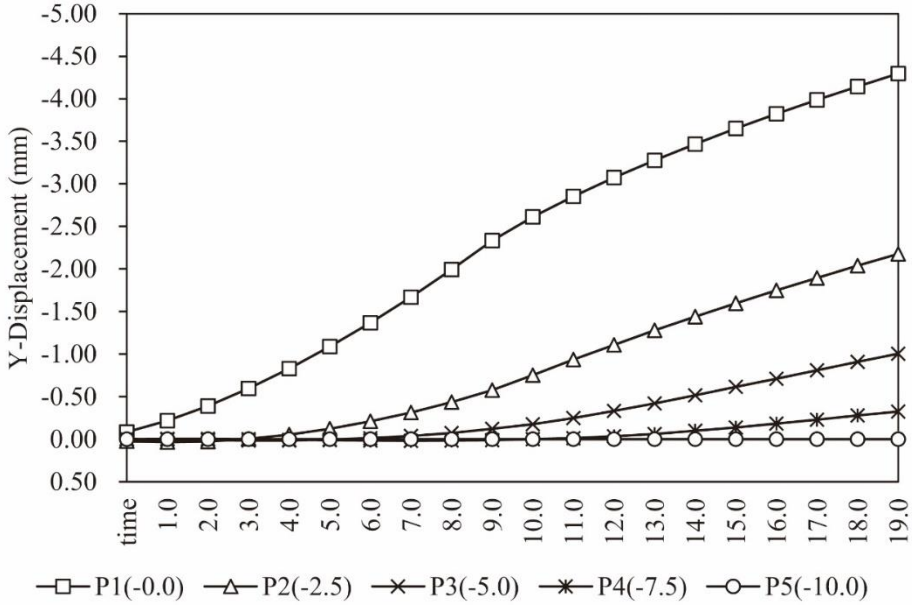


Figure 10 Time series graph of the Y-Displacement results by NS-PIM(T3) (mm)

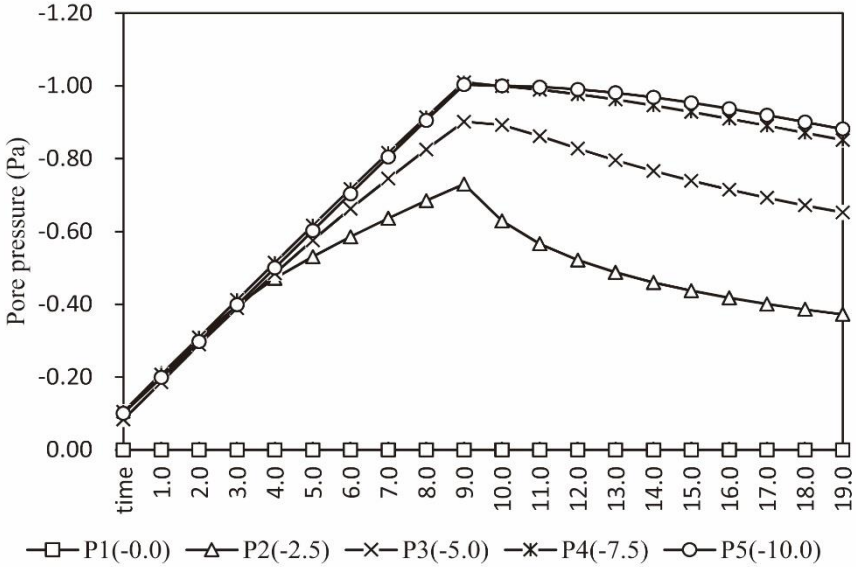


Figure 11 Time series graph of the pore pressure results by NS-PIM(T3) (Pa)

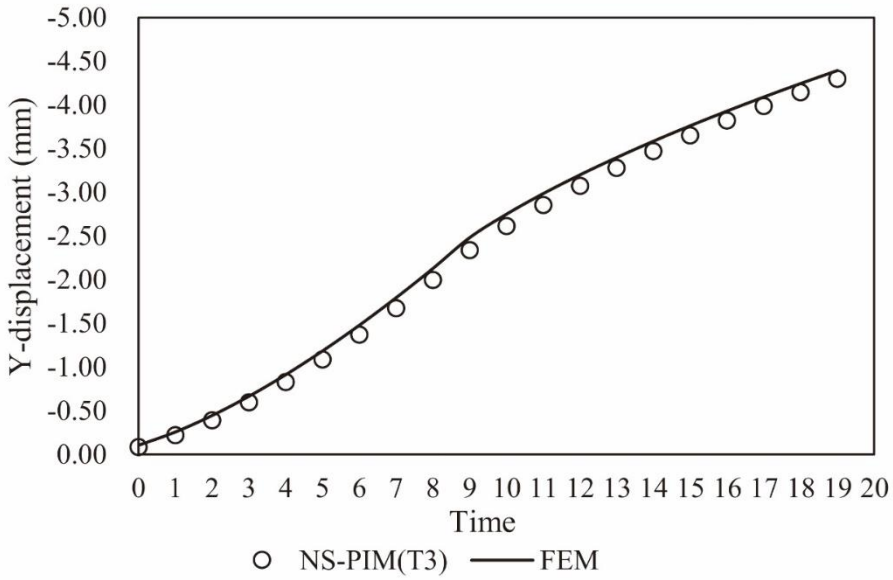


Figure 12 The Y-displacement values of 'P1' by different solutions(mm)

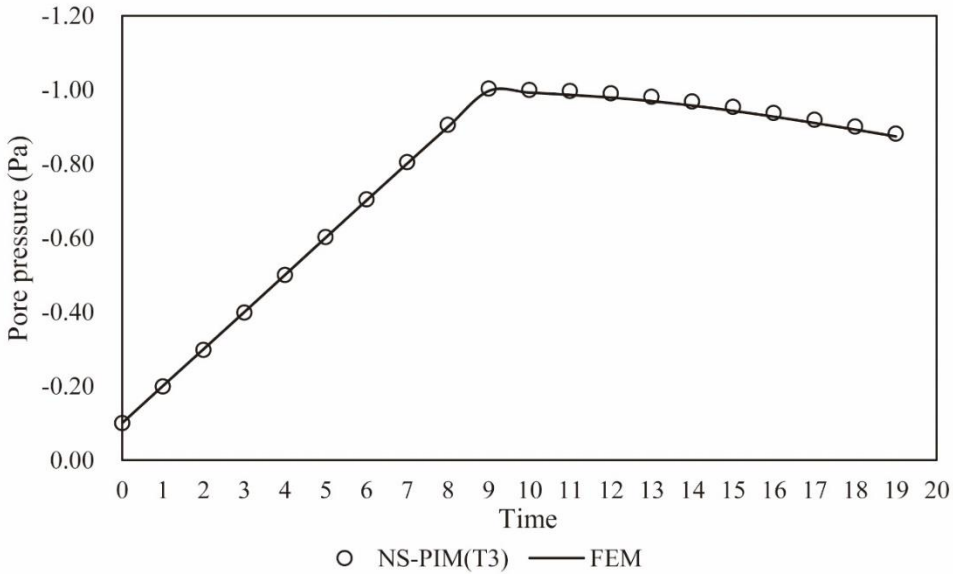


Figure 13 The Pore pressure values of 'P5' by different solutions(Pa)

The relations between the results and time is well simulated by NS-PIM, that is the Y-Displacement values increase gradually over time and the pore pressure values first increase then decrease with the increase of time. Also, the results are equivalent to those of the FEM solution. However, the time series curve in Figure 11 is a little abnormal during the time from 0s to 4s, and this is considered to be the computational error of the NS-PIM program.

Although the distributions and time series graphs of NS-PIM and FEM solutions share the similar rules as is described above, there are still some differences between the results of two solutions which can be found from careful comparison. Thus, it is necessary to compare the results with those of the certified results in Schiffman RL's report(1960) [8,23] to confirm the correctness and validation of the NS-PIM solution.

3.1.4 Certification of the results

In order to compare the calculated results in this example with the certified results in Schiffman RL's report(1960), the time of the simulation is extended to 1000 seconds. Also, the dimensionless 'time factor' T is defined as

$$\mathrm{T} = \frac{c_v t}{D^2} \tag{17}$$

Where D is the 'maximum drainage path'(depth) of the model, t is the time of the simulation and c_v is the coefficient of consolidation. c_v is defined as

$$c_v = \frac{k_y}{m_v \gamma_w} \tag{18}$$

where k_y is the permeability in the y-direction, γ_w is the unit weight of water and m_v is defined as

$$m_v = \frac{(1-v')(1-2v')}{E'(1-v')} \tag{19}$$

where E' is the effective Young's modulus and v' is the poisson's ratio.

At the beginning of this part ,it is mentioned that D is 10m, k_y equals to γ_w, E' is 1 and v' is 0. Thus, the dimensionless 'time factor' T equals to 0.01t, and the pore pressure of different solutions at the base of the mesh (P5) is plotted against time in Figure 14 .

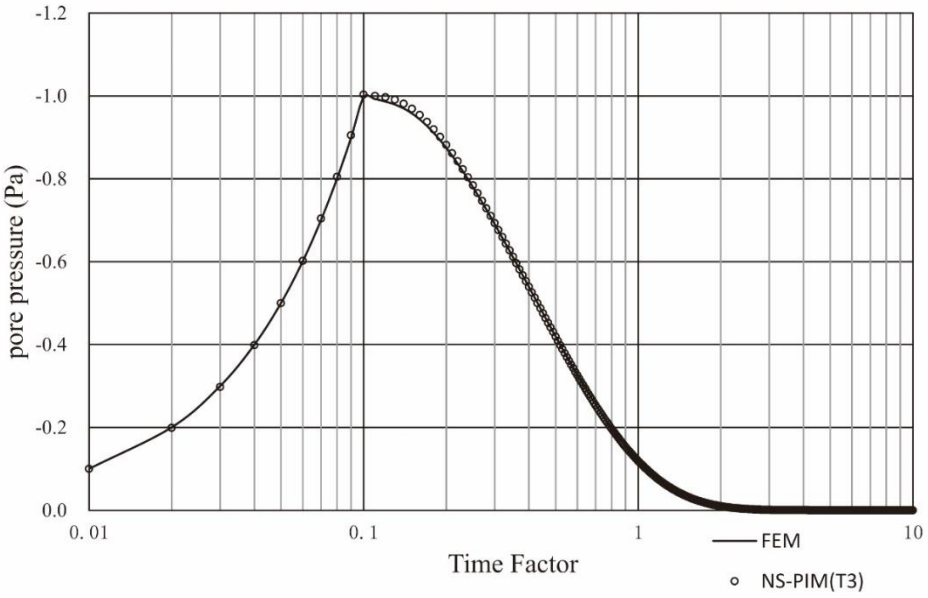


Figure 14 Pore pressure response to ramp loading at 'P5' by different solutions

As is described in the figures above, the pore pressures of both solutions rises to peak(1.0Pa) till the time factor is 0.1, then the pore pressure decreases slowly to zero. Comparing the figures above with the corresponding figures in Schiffman RL's report (1960)[23], conclusions can be drawn that the results calculated by both solutions are correct and valid.

3.1.5 Mesh refinement

Further research is carried out in order to study the convergence features of the NS-PIM solution. The same model as Figure 4 is created in this part. Main parameters are also listed in part 3.1.1 .

The Figure 15 is the same model using different mesh schemes. Details of the models and the results calculated by the NS-PIM and FEM solutions in the 20th time step(last time step of the analysis) are listed in Table II . Figure 16 and Figure 17 are the logarithmic graphs of the maximum Y-Displacement and maximum pore pressure in the 20th time step which indicate that the results of both solutions possess good convergency .

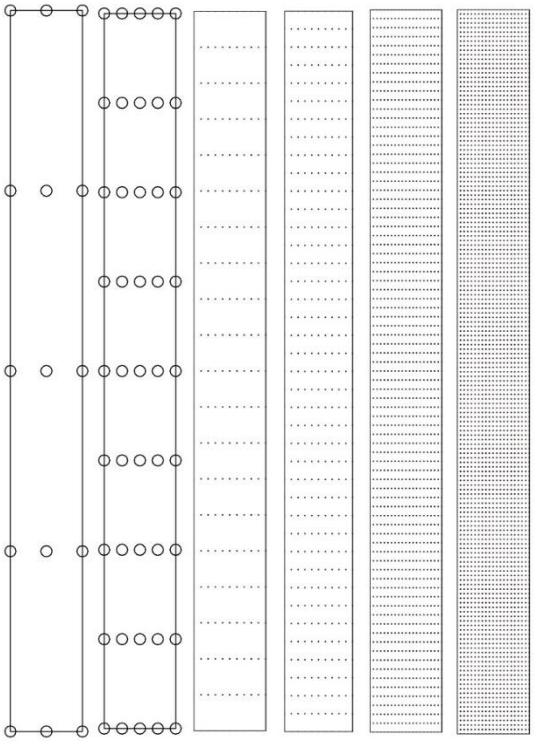


Figure 15 Six different mesh schemes

Table II. Results of different mesh schemes

Mesh Scheme	1	2	3	4	5	6
Number of nodes	15	45	231	451	1701	3381
Number of cells	16	64	400	800	3200	6400
Maximum Y-Displacement in the 20th time step (NS-PIM(T3))	-4.29540	-4.36070	-4.37970	-4.38250	-4.38310	-4.38330
Maximum Pore pressure in the 20th time step (NS-PIM(T3))	-0.88144	-0.85975	-0.85398	-0.85312	-0.85292	-0.85286
Maximum Y-Displacement in the 20th time step (FEM)	-4.39280	-4.38450	-4.38350	-4.38340	-4.38340	-4.38340
Maximum Pore pressure in the 20th time step (FEM)	-0.87418	-0.85822	-0.85369	-0.85305	-0.85289	-0.85285

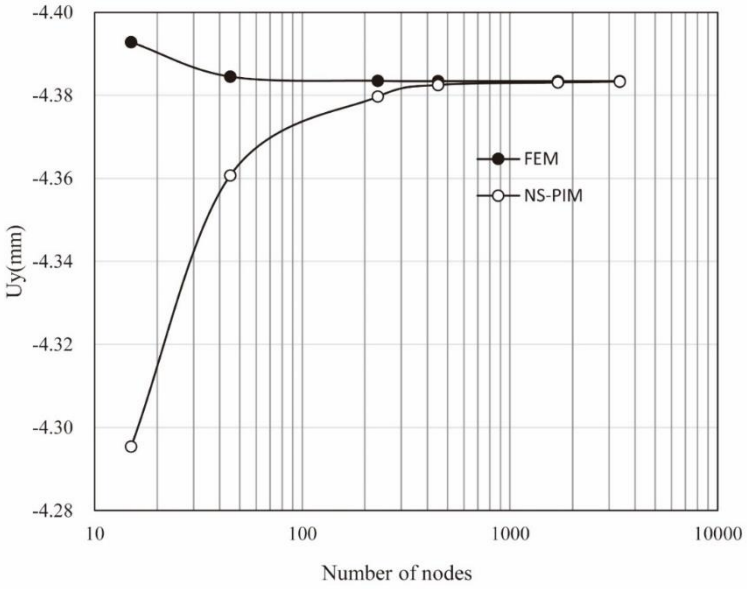


Figure 16 Maximum Y-Displacement in the 20th time step by FEM and NS-PIM

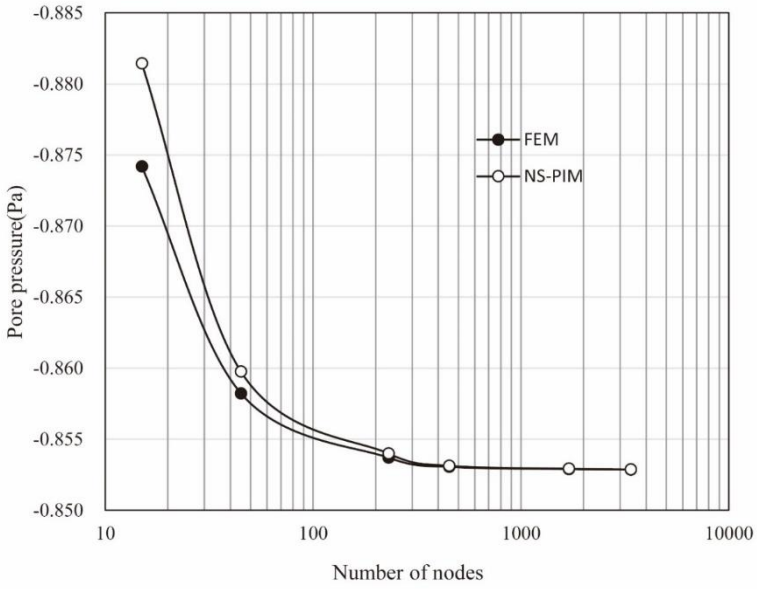


Figure 17 Maximum pore pressure in the 20th time step by FEM and NS-PIM

The graphs of Figure 16 and Figure 17 indicate that the calculated displacement values of NS-PIM and FEM solutions converge towards the exact values from different sides , while the calculated pore pressure values of both solutions converge towards the exact values from the same side. Thus, more accurate results and computational error can be obtained by comparing the results calculated by the NS-PIM and FEM solutions.

3.1.6 Error Estimation

Based on the convergence features of the NS-PIM and FEM, equation (20) and equation (21) is defined to estimate the errors of the current model.

$$ERR = \frac{|D_{exact} - D_{numer}|}{|D_{exact}|} \tag{20}$$

$$ERR1 = \frac{|D_{NS-PIM} - D_{FEM}|}{\min\{|D_{NS-PIM}|, |D_{FEM}|\}} \tag{21}$$

where ERR is the exact error, $ERR1$ is the calculated error, and D_{exact}, D_{numer}, D_{NS-PIM}, D_{FEM} mean the exact maximum value, maximum value from numerical method, NS-PIM and FEM, respectively.

As regard to the displacement values, it can be seen from Figure 16 that $|D_{NS-PIM} - D_{FEM}| > |D_{NS-PIM} - D_{exact}|$ and $|D_{NS-PIM} - D_{FEM}| > |D_{exact} - D_{FEM}|$, hence $ERR1 > ERR$ which means the $ERR1$ for displacement is an overestimate of ERR.

As for the pore pressure values, the exact value is less than the calculated values by both solutions. But, it still can be found from Figure 17 that the values by both solutions getting closer to each other, while they are approaching the exact value. Thus, Equation (21) can still be used as an estimation of the error.

Take the results of mesh scheme 1 for example, the $ERR1$ for displacement is equal to 2.27% and the $ERR1$ for pore pressure is equal to 0.83%. If the accuracy is enough, mesh scheme 1 would be enough to get the results. Otherwise, finer mesh schemes should be used to get the satisfactory results.

This approach for estimating the calculation error is quite useful ,as the estimated error can be obtained with just one set of model and cells which means more time can saved from the boring work of modeling and remeshing.

4 CONCLUSIONS AND DISCUSSIONS

Node-based smoothed point interpolation method is applied to the Biot's consolidation analysis in this work. Results are compared to those of the certified FEM solution to check the correctness and validation of the NS-PIM solution. Also, a simple method is introduced to estimate the errors of the results with rough grids.

The conclusions of the comparison are listed as follows:

1) It is feasible to apply the NS-PIM solution to the Biot's consolidation analysis. And the results of the solution fit well with the certified results.

2) The NS-PIM solution applied to Biot's consolidation still possesses different convergence in displacement results and similar convergence in pore pressure results compared to the FEM solution. This property can be used to determine the computational errors of the solutions and get more accurate results. According to the results of this work, when the ERR1 for displacement is less than 0.5% and the ERR1 for pore pressure is less than 0.15%, the accuracy would be enough for general analysis.

3) The proposed method for estimating the computational error makes it possible to obtain accurate results with fewer 'remeshing works'. NS-PIM solution as one of the mesh free methods definitely makes the meshing and calculation for consolidation analysis easier and more efficient.

Thus, it is promising to apply the NS-PIM to the consolidation analysis. Also, more accurate and convincing results can be got based on the different convergence features between the FEM and NS-PIM solutions.

Meanwhile, it cannot be denied that the calculation of NS-PIM solution takes more time than that of the FEM solution, as the point interpolation for NS-PIM shape functions involves more nodes and procedures than that of the FEM solution. Although optimizations like node renumerating have been done in the solutions, more works are still necessary to be carried out to optimize the NS-PIM solution. Also, in further study, the feasibility of nonlinear, elastoplastic and 3D model can still be probed in the NS-PIM solution.

5 Acknowledgements

The support of the Fundamental Research Funds for the Central Universities (No. 2015B15114) and the Natural Science Foundation of China (No. 51009056) is gratefully acknowledged.

Reference

[1] SM Gourvenec, C Vulpe, TG Murthy. A method for predicting the consolidated undrained bearing capacity of shallow foundations. Geotechnique 2014; 64: 215–225.
[2] ZY Ai, CC Yi. 3-d consolidation analysis of layered soil with anisotropic permeability using analytical layer-element method. Acta Mechanica Solida Sinica 2013; 26(1): 62-70.
[3] DV Griffiths, J Huang. One-dimensional consolidation theories for layered soil and coupled and uncoupled solutions by the finite-element method. Géotechnique 2010; 60: 709–713.
[4] ZY Ai, YD Hu. Multi-dimensional consolidation of layered poroelastic materials with anisotropic permeability and compressible fluid and solid constituents. Acta Geotech 2015; 10: 263-273.
[5] MA Biot. General theory of three-dimensional consolidation. J. Appl. Phys 1941; 12 : 155–164.
[6] Zienkiewicz OC, Taylor RL. The Finite Element Method (5th edn). (V1: The basis), Butterworth-Heinemann: Oxford, 2000.
[7] S Cen, C Li, S Rajendran, Z Hu. Advances in finite element method, Math. Probl. Eng. 2014; 2014: 0–3.
[8] I M Smith, D V Griffiths, L Margetts. Programming the finite element method. John Wiley & Sons Ltd: United Kindom, 2014 .

[9] GR Liu, D Karamanlidis. Mesh Free Methods: Moving Beyond the Finite Element Method, Appl. Mech. Rev 2003; 56(2): B17-B18.
[10] GR Liu, Y Li, KY Dai, MT Luan, W Xue. a Linearly Conforming Radial Point Interpolation Method for Solid Mechanics Problems. Int. J. Comput. Methods 2011; 3(4): 401–428.
[11] GY Zhang, GR Liu, YY Wang, HT Huang, ZH Zhong, GY Li, X Han. A linearly conforming point interpolation method (LC-PIM) for three-dimensional elasticity problems. Int. J. Numer. Methods Eng 2007; 72 (13): 1524–1543.
[12] J Cheng, XL Chang, SC Wu, GY Zhang. Certified solutions for hydraulic structures using the node-based smoothed point interpolation method (NS-PIM). Int. J. Numer. Anal. Methods Geomech 2009; 34 (15): 1560–1585.
[13] GR Liu, YT Gu. A point interpolation method for two-dimensional solids. Int. J. Numer. Methods Eng 2001; 50(4): 937–951.
[14] JG Wang, GR Liu. A point interpolation meshless method based on radial basis functions. Int. J. Numer. Methods Eng 2002; 54(11): 1623–1648.
[15] Y Krongauz, T Belytschko. Enforcement of essential boundary conditions in meshless approximations using finite elements. Comput. Methods Appl. Mech. Eng 1996; 131(1-2): 133–145.
[16] GR Liu, GY Zhang. Upper bound solution to elasticity problems: A unique property of the linearly conforming point interpolation method (LC-PIM). Int. J. Numer. Methods Eng 2008; 74(7): 1128–1161.
[17] GY Zhang, GR Liu, TT Nguyen, CX Song, X Han, ZH Zhong, GY Li. The upper bound property for solid mechanics of the linearly conforming radial point interpolation method (LC-RPIM). Int. J. Comput. Methods 2011; 04(3): 521–541.
[18] GY Zhang, GR Liu, Y Li. An efficient adaptive analysis procedure for certified solutions with exact bounds of strain energy for elasticity problems. Finite Elem. Anal. Des 2008; 44(14): 831–841.
[19] JG Wang, GR Liu, YG Wu. A point interpolation method for simulating dissipation process of consolidation. Comput. Methods Appl. Mech. Eng 2001; 190(45): 5907–5922.
[20] P Mira, M Pastor, T Li, X Liu. Failure problems in soils An enhanced strain coupled formulation with application. Revue Française De Génie Civil 2004; 8(5-6):735-759.
[21] GR Liu, GY Zhang. Edge-Based Smoothed Point Interpolation Methods. Int. J. Comput. Methods 2011; 5(4): 621–646.
[22] J Chen, C Wu, S Yoon. A stabilized conforming nodal integration for Galerkin mesh-free methods. Int. J. Numer. Meth. Eng 2001; 50(2): 435–466.
[23] R Schiffman. Consolidation of soil under time-dependent loading and varying permeability. Highway Research Board Proceedings 1958; 37.

Temperature influences on the performance of a dielectric elastomer generator with consideration of dissipation processes and failures

***Shoue Chen[1], †Zhicheng He[1], Eric Li[2]**

[1]State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, China
[2]Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, China

*Presenting author: chenshoue725@163.com
†Corresponding author: hezhicheng815@163.com

Abstract

As a category of promising material, the membrane of dielectric elastomer (DE) sandwiched between two compliant electrodes has the capacity for converting the electrical energy into mechanical energy, vice versa. Owing to the large deformation produced by relatively small stimulations, this elastomer is also component to function as a generator. Investigations on the dielectric elastomer generator (DEG) have achieved wide attention recently. The previous studies have indicated that the performance of DE depends on the major dissipation processes including viscoelasticity and current leakage and also varies with temperature. However, very few works take these factors together into consideration when investigating the performance of DEGs. Therefore, a model that involves the temperature-dependent permittivity and shear modulus of the DE membrane, viscoelastic relaxation and current leakage is established in this study. Then, based on a specific Carnot-shape conversion cycle, the performances of the dissipative generator made of very-high-bond (VHB) elastomer can be discussed at different sampling temperatures. The parameters characterizing the performances of the DEG include the energy densities of different kinds and conversion efficiency. Moreover, the mechanisms of typical failure modes including material rupture, loss of tension (LT), electrical breakdown (EB) and electromechanical instability (EMI) are studied with the influences of temperature to ensure that the generator is operated in an allowable area. It can be concluded from the numerical results that the temperature plays an important role in the performance of the DEG, which could possibly improve its conversion efficiency.

Keywords: Dielectric elastomer generators (DEGs); Temperature; Failure modes; Dissipation processes;

1. Introduction

The dielectric elastomer (DE) which appears as a thin membrane is often coated with softly conductive electrodes on both sides in thickness [1]-[4]. When subject to a voltage through its thickness, the membrane of DE will shrink in thickness and expand in area due to Maxwell stress, converting the electrical energy into mechanical energy, which can be exploited as an actuator [5]-[7]. Owing to the large deformation caused by small mechanical forces and an appropriate voltage, the DE membrane is also competent to function as a generator [8]-[10]. When a membrane of DE operates under cyclic loadings, the reduction in tensile force will lead to the enhancement in both the thickness and voltage through the electrodes, converting the mechanical energy into electrical energy. Nowadays, there is an increasing attention on the field of dielectric elastomer generator (DEG). For instance, a heel-strike generator made of DE has been implanted into shoes [11], a membrane device placed behind the knee is able

to recycle the energy from human motion [12], and even a wave power generator made of DE has been developed for harvesting renewable energy [13].

Being a generator, except the unavoidable dissipation processes including viscoelasticity and current leakage [14], the DE membrane may lose efficacy due to multiple failure modes, such as material rupture, loss of tension (LT), electrical breakdown (EB), and electromechanical instability (EMI) [15]-[22], which may restrict the area of allowable states during the energy harvesting. As a result, it is of great importance to study the mechanism of failure modes for advanced development of DEGs.

Most of the previous studies on the electromechanical performance of DE were implemented in an isothermal environment but ignored the influences of temperature [23][24]. However, according to the experiments on very-high-bond (VHB) elastomers [25][26], there is a strong dependency between the relatively permittivity and the planar stretches of the DE membrane at different operating temperatures. The shear modulus of the membrane of DE is also proved to be related with temperature and the relationship will not be described by the T/T_0 factor anymore. To the authors' best knowledge, there is few research on the effects of temperature on both electromechanical performance and energy conversion of DEGs.

In this study, the most promising DE material, VHB 4910, which has high electromechanical conversion efficiency and can produce a large deformation within a wide temperature range (233~363 K) [26] is selected to study the effect of temperature on the performance of DEGs. For the first time, an integrated model that combines the temperature-dependent permittivity and shear modulus, current leakage and viscoelasticity is established to simulate the energy harvesting. Afterwards, the mechanism and processes of a four-stroke conversion cycle for the DEG is introduced. Based on the cyclic loads, different energies as well as the conversion efficiency can be figured out. By comparing the numerical results at different temperatures, one can summarize the temperature influence on the DEG. At the same time, common failure modes of the DE membrane are taken into consideration to ensure the allowable states.

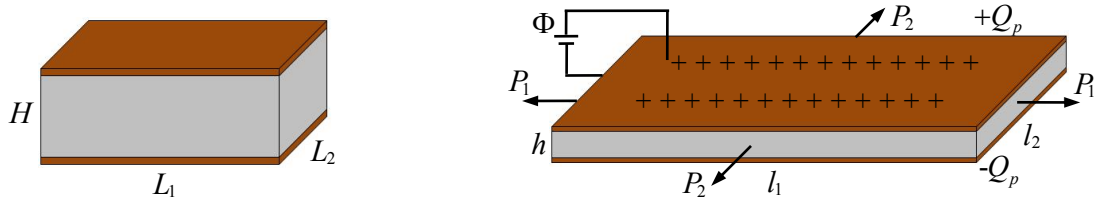
2. Governing equations for the dissipative DEG

To focus on the influences of temperature on the electromechanical performance of the DEG, the governing equations for the membrane of DE should be first derived and can be separated into the following parts: (1) the free energy function of DE with consideration of viscoelastic relaxation and temperature-related factors; (2) equations of the plane-parallel capacitor; (3) equations of the current leakage. These segments constitute the integrated equations of state for the DEG.

2.1. Free energy function of the DE membrane

A fundamental configuration is utilized to conduct the simulation, in which a thin membrane of VHB 4910 is sandwiched between two electrodes with negligible electrical resistance and mechanical stiffness. Subject to in-plane biaxial forces P_1 and P_2 and a voltage Φ in thickness, the membrane deforms from its initial dimensions $L_1\times L_2\times H$ to the current dimensions $l_1\times l_2\times h$ with charges of opposite sign $\pm Q_p$ generated on the electrodes at a fixed temperature T, as demonstrated in Fig. 1. As a homogeneous and isotropic elastomer, the stretches in plane can be defined as $\lambda_i=l_i/L_i$ (i=1,2), while the stretch in thickness direction can be substituted by $\lambda_1^{-1}\lambda_2^{-1}$ due to the nearly-incompressibility of the membrane [2][27]. Similarly, the nominal stresses in plane are defined by $s_1=P_1/(L_2H)$, $s_2=P_2/(L_1H)$. In electrical category, $\tilde{E}=\Phi/H$

and $\tilde{D}=Q_p/(L_1L_2)$ represent the nominal electric field and the nominal electric displacement, respectively.



(a) (b)

Figure 1. Schematic of the DE membrane in (a) undeformed state; (b) deformed state, subject to forces and a voltage at a fixed temperature *T;*

Due to obvious viscoelasticity, the membrane of DE will no longer show the simple elasticity during the deformation. According to the recent studies [28]-[31], the viscoelastic relaxation can be simulated by a nonlinear rheological model, as illustrated in Fig. 2. The viscoelastic model consists of two parallel networks: spring *A* in parallel with spring *B* and a dashpot with a viscosity η. Factors μ^A and μ^B represent the shear moduli of the springs which vary with temperature, J^A and J^B are the dimension parameters referring to the dependence on chain extension limits of the elastomer. It is assumed that J^A and J^B are independent of temperature.

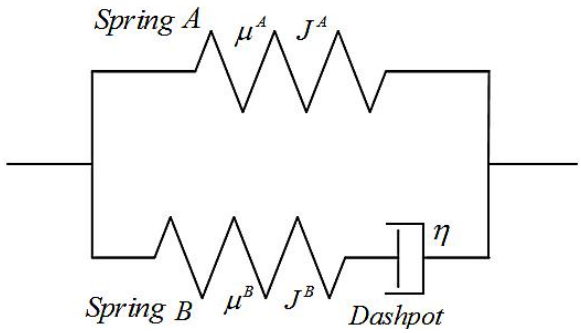


Figure 2. Viscoelasticity of the DE membrane is modeled by a rheological model.

In the rheological model, the stretch of the DE membrane in a certain direction is equal to the net stretches in both networks. For spring *A*, the stretch in it is the same as the stretch of the membrane λ. While the stretch of the spring *B* is described by λ^e, which cooperates with the inelastic stretch ξ due to the dashpot. A well-established rule [14] is chosen to represent the relationship between these stretches, $\lambda=\lambda^e\xi$.

In order to measure the mechanical work done by the dashpot, that is, the energy dissipated by the viscoelasticity during the harvesting, the condition of stress in the rheological model should also be solved. When subject to mechanical forces in plane and a voltage in thickness, the DE membrane will be stretched by λ due to the stress arising form the force σ_{force} and the Maxwell stress $\sigma_{Maxwell}$. The sum of the external stresses should be balanced by the stresses in both networks, $\sigma^A+\sigma^B=\sigma_{Maxwell}+\sigma_{force}$, where the stress in spring *A* equals the stress in the top network σ^A, and the stress acting on the dashpot is the same as the stress in spring *B*, which is also the stress in the bottom network σ^B.

Previously, the permittivity of DE is assumed as constant in the majority of cases, which goes against the molecular physics: molecules in the elastomer will contain more thermodynamic energy at a higher temperature, resulting in a greater amplitude of random thermal motion.

Therefore, the molecules are less closely aligned with each other, and the relative permittivity will decline with the enhancement in temperature. The stretch will change the active range of the molecules too [32]. Indeed, the recent research has proved that the relative permittivity of VHB 4910 is affected by the temperature and stretch simultaneously [26], and the specific expression of ε_r for VHB 4910 can be depicted in Eq. (1).

$$\varepsilon_r(\lambda_1,\lambda_2,T)=a\lambda_1\lambda_2+\frac{b}{T}+c \tag{1}$$

where a is the parameter describing the electrostriction of DE, and b is determined by the expression $M\eta_d^2/(3\kappa\varepsilon_0)$, where M, η_d and κ represent the dipole current density, the dipole moment and the Boltzmann constant, respectively. ε_0=8.85×10^{-12} F/m is the permittivity of vacuum. Parameter c describes the relative permittivity at the reference temperature without deformation. The parameters obtained from the experimental data of Jean-Mistral [26] will be selected here: a=-0.0533 F/m, b=645.4224 F.K/m and c=3.1834 F/m.

The dielectric elastomer itself, the mechanical forces, the external voltage and the thermal force constitute a thermodynamic system, and the free energy density function of this system can be expressed as follows:

$$W=W_s(\lambda_1,\lambda_2,T)+\frac{\tilde{D}^2}{2\varepsilon_0\varepsilon_r(\lambda_1,\lambda_2,T)}\lambda_1^{-2}\lambda_2^{-2}+\rho_0 c_0\left[T-T_0-T\ln(\frac{T}{T_0})\right] \tag{2}$$

where the thermo-elastic energy W_s denotes the variation in strain energy from the reference configuration to the current configuration at a fixed temperature T. The electrostatic energy is represented by the second item in the right hand side of Eq. (2). The last item reflects the absolute thermal contribution [18][19], in which the density and the specific heat capacity of the DE film are described by ρ_0 and c_0, respectively.

In practice, the VHB elastomer exhibits strain-stiffening effect during the deformation due to finite configurations of the polymer chains [2][32]. Considering the extension limits, the Gent model [33] is adopted to characterize the thermo-elastic energy W_s. Combine the rheological model and the relative permittivity of VHB 4910, the specific function of W_s can be described in Eq. (3).

$$\begin{aligned}W_s=&-\frac{\mu^A(T)J^A}{2}\ln\left(1-\frac{\lambda_1^{\ 2}+\lambda_2^{\ 2}+\lambda_1^{-2}\lambda_2^{-2}-3}{J^A}\right)-\frac{\mu^B(T)J^B}{2}\cdot\\&\ln\left(1-\frac{\lambda_1^{\ 2}\xi_1^{-2}+\lambda_2^{\ 2}\xi_2^{-2}+\lambda_1^{-2}\lambda_2^{-2}\xi_1^{\ 2}\xi_2^{\ 2}-3}{J^B}\right)\end{aligned} \tag{3}$$

where the shear moduli of the springs meet the relation, μ^A/μ^B=3/7 [14] and an instantaneous modulus of DE is defined as $\mu(T)=\mu^A+\mu^B=Y(T)/3$, where $Y(T)$ stands for the isothermal elastic modulus in small deformation at temperature T. In this study, a set of material parameters for

VHB 4910 are chosen as: J^A=90, J^B=30 [34] and $Y(T)$=0.2001(1000/T)^2-1.078(1000/T)+1.518 MPa [35].

According to the equilibrium in thermodynamics, when the DE membrane is in mechanical and electrical equilibrium, the equations of state can be achieved from: $s_1=\partial W/\partial\lambda_1$, $s_2=\partial W/\partial\lambda_2$ and $\tilde{E}=\partial W/\partial\tilde{D}$ [2][16]. The specific equations of state are demonstrated in Eqs. (4)-(6).

$$s_1=\frac{\mu_B(T)(\lambda_1\xi_1^{-2}-\lambda_1^{-3}\xi_1^{2}\lambda_2^{-2}\xi_2^{2})}{(1-(\lambda_1^{2}\xi_1^{-2}+\lambda_2^{2}\xi_2^{-2}+\lambda_1^{-2}\lambda_2^{-2}\xi_1^{2}\xi_2^{2}-3)/J_B)}+\frac{\mu_A(T)(\lambda_1-\lambda_1^{-3}\lambda_2^{-2})}{(1-(\lambda_1^{2}+\lambda_2^{2}+\lambda_1^{-2}\lambda_2^{-2}-3)/J_A)}-\frac{\tilde{D}^2}{\varepsilon_0(a\lambda_1\lambda_2+b/T+c)\lambda_1^{3}\lambda_2^{2}}-\frac{a\tilde{D}^2}{2\varepsilon_0(a\lambda_1\lambda_2+b/T+c)^2\lambda_1^{2}\lambda_2} \tag{4}$$

$$s_2=\frac{\mu_B(T)(\lambda_2\xi_2^{-2}-\lambda_2^{-3}\xi_1^{2}\lambda_1^{-2}\xi_2^{2})}{(1-(\lambda_1^{2}\xi_1^{-2}+\lambda_2^{2}\xi_2^{-2}+\lambda_1^{-2}\lambda_2^{-2}\xi_1^{2}\xi_2^{2}-3)/J_B)}+\frac{\mu_A(T)(\lambda_2-\lambda_2^{-3}\lambda_1^{-2})}{(1-(\lambda_1^{2}+\lambda_2^{2}+\lambda_1^{-2}\lambda_2^{-2}-3)/J_A)}-\frac{\tilde{D}^2}{\varepsilon_0(a\lambda_1\lambda_2+b/T+c)\lambda_1^{2}\lambda_2^{3}}-\frac{a\tilde{D}^2}{2\varepsilon_0(a\lambda_1\lambda_2+b/T+c)^2\lambda_1\lambda_2^{2}} \tag{5}$$

$$\tilde{E}=\frac{\tilde{D}}{\varepsilon_0(a\lambda_1\lambda_2+b/T+c)\lambda_1^{2}\lambda_2^{2}} \tag{6}$$

The membrane is subject to a homogeneous, equal-biaxial force P in the following simulation, so setting $s_1=s_2=s$, $\lambda_1=\lambda_2=\lambda$ and $\xi_1=\xi_2=\xi$ in the equations of state above. With elimination of the variable $\tilde{D}$ in either Eq. (4) or (5), the function denoting the nominal stress of the membrane is illustrated in Eq. (7).

$$\frac{P}{LH}=\frac{\mu^A(T)(\lambda-\lambda^{-5})}{1-\frac{2\lambda^2+\lambda^{-4}-3}{J^A}}+\frac{\mu^B(T)(\lambda\xi^{-2}-\lambda^{-5}\xi^{4})}{1-\frac{2\lambda^2\xi^{-2}+\lambda^{-4}\xi^{4}-3}{J^B}}\quad\tilde{E}^2\varepsilon_0(1.5a\lambda^2+\ b/T+\ c)\lambda^3 \tag{7}$$

Based on the expression of the nominal stress, a function $F(\lambda)$ that characterizes the elasticity of the elastomer can be obtained in Eq. (8). In the absence of a voltage (E=0), the function $F(\lambda)$ corresponds to the nominal stress $s=P/LH$. In the absence of mechanical loads (s=0), the membrane of DE will also contract in thickness and expand in area when subject to a voltage. The deformation induced by electric filed equals that caused by the Maxwell stress, which is equivalent to the equal-biaxial nominal stress $(1.5a\lambda^2+b/T+c)\lambda^3\varepsilon_0\tilde{E}^2$. The two items in Eq. (8) also represent the nominal stresses of the networks, respectively.

$$F(\lambda)=\frac{\mu^{A}(T)(\lambda-\lambda^{-5})}{1-\dfrac{2\lambda^{2}+\lambda^{-4}-3}{J^{A}}}+\frac{\mu^{B}(T)(\lambda\xi^{-2}-\lambda^{-5}\xi^{4})}{1-\dfrac{2\lambda^{2}\xi^{-2}+\lambda^{-4}\xi^{4}-3}{J^{B}}} \tag{8}$$

The dashpot in the nonlinear rheological model is regarded as a Newtonian fluid, so the rate of deformation in the dashpot can be denoted by $\xi^{-1}\mathrm{d}\xi/\mathrm{d}t$, and the relationship between the stretch λ and the inelastic stretch ξ can be described in Eq. (9). The dashpot will dissipate the mechanical energy with a viscoelastic relaxation time $\tau(T)=\eta/\mu^{B}(T)$ [14], which obviously depends on temperature. Refer to the experiment data from Sheng et. al [36], the viscoelastic relaxation time corresponding to four sampling temperatures are illustrated: τ(273K)=87.216 s, τ(293K)=72.377 s, τ(313K)=67.802 s and τ(333K)=65.573 s.

$$\frac{\mathrm{d}\xi}{\mathrm{d}t}=\frac{\mu^{B}(T)}{6\eta}\frac{\lambda^{2}\xi^{-1}-\lambda^{-4}\xi^{5}}{1-(2\lambda^{2}\xi^{-2}+\lambda^{-4}\xi^{4}-3)/J^{B}} \tag{9}$$

2.2. Equations of the plane-parallel capacitor

When the two electrodes are connected to an electric power source through a conducting wire, charges will be activated and transferred from one electrode to another, leading to the electric potential difference between the surfaces of the DE membrane. Along with the deformation induced by the equal-biaxial force, the additional charges can be stored and released through the electric circuit. From an electrical point of view, the DEG can be regarded as a stretchable plane-parallel capacitor. The membrane of DE functions as a dielectric medium between two compliant electrodes plates. The governing equation for this capacitor is given by $Q_p=\Phi C$, where C represents the capacitance of the DEG determined by the configuration of the DE membrane. By definition, the electric displacement is $D=Q_p/(\lambda L)^2$, and the electric field is $E=\Phi/h$. These two variables can be related by $D=\varepsilon_0\varepsilon_r E$. A combination of the relations above is shown in Eq. (10).

$$Q_P=C\Phi=\left[\frac{L^{2}}{H}\varepsilon_0(a\lambda^{6}+\frac{b}{T}\lambda^{4}+c\lambda^{4})\right]\Phi \tag{10}$$

2.3. Equations of the current leakage

As an electrical component, there is no doubt that the generator made of DE will suffer from unavoidable leakage problem, as shown in Fig. 3. The current leaks through the membrane i_{leak} can be modeled by a conductor which is in parallel with a capacitor. The amount of the current which flows through the electrodes i_p is obtained by differentiating the magnitude of the positive and negative charges $\pm Q_p$ that polarize the membrane with respect to time. The total amount of the charges which transported through the conducting wire is Q and can be acquired by Eq. (11).

$$\frac{\mathrm{d}Q}{\mathrm{d}t}=\frac{\mathrm{d}Q_P}{\mathrm{d}t}+i_{leak} \tag{11}$$

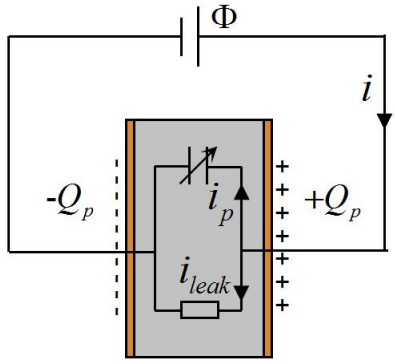


Figure 3. Current leakage model of the DE membrane.

A variable describing the density of the leaked current is defined, $j_{leak}=i_{leak}/(\lambda L)^2$. In addition, the experiment has indicated that the conductivity of the VHB elastomer rises exponentially with the rising electric field, and the following relation has be fitted, $j_{leak}=\sigma_0\exp(E/E_B)E$ [37], where σ_0 is the conductivity under the low electric fields, and E_B is an empirical constant with the same dimension as the electric field. The variables $\sigma_0=3.32\times10^{-14}$ S/m, $E_B=40$ MV/m are chosen for the following analysis [14][38], and i_{leak} can be obtained from Eq.(12).

$$\frac{i_{leak}}{L^2}=\sigma_0\exp(\frac{\tilde{E}\lambda^2}{E_B})\tilde{E}\lambda^4 \tag{12}$$

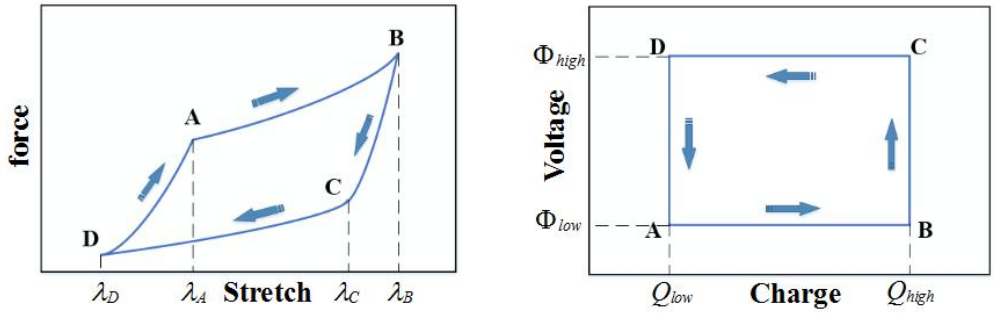
3.Mechanisms of energy conversion and failure modes

Different kinds of processes such as constant voltage, constant charge and constant electric [39]-[41] have been designed as energy conversion cycles. Although each conversion cycle has its own merits and demerits in the suitability and stability, the mechanism of conversion cycles can be summarized: the electrical energy is extracted from a four-stroke cycle by the variability in the capacitance of the DE membrane. To narrow down our research objective, the Carnot-shape cycle [42][43] that consists of two constant voltage processes and two open circuit processes is introduced and employed to study the behaviors of the DEG at different temperatures.

3.1. Basic energy conversion cycle

The Carnot-shape cycle applied to the DEG can be realized via a three-way switch [43][44]. A switch can connect the generator to the input battery that provides the charges at a low voltage, or connect it to the output battery that stores the charges at a high voltage, or keep it in an open circuit. The mechanism of an ideal four-stroke cycle is demonstrated in Fig. 4, where a line describes a process and a contour is a cycle. In the process A→B, as illustrated in Fig. 4(b), the amount of the charges on the electrodes supplied by the low-voltage battery increases from Q_{low} to Q_{high} with the continuous stretch of the membrane from λ_A to λ_B as depicted in Fig. 4(a) (the capacitance improves). In the process B→C, with constant charge (open circuit), the voltage between the electrodes increases when the stretch is partly released from λ_B to λ_C (the capacitance decreases). The release degree is measured by the expression: $\Phi_{low}C(\lambda_B)=\Phi_{high}C(\lambda_C)$. In the process C→D, the charges on the electrodes are pumped to the high-voltage battery due to the further release of the stretch from λ_C to λ_D (the amount of the charges that the DEG can store declines with a smaller capacitance). In the last process, with charges maintained again (open circuit), the voltage between the electrodes decreases when the membrane is stretched from λ_D to the initial stretch λ_A (the capacitance increases) and the

following relation should be satisfied: $\Phi_{low}C(\lambda_A)=\Phi_{high}C(\lambda_D)$. After each cycle, external forces pump a certain amount of charges from the low-voltage battery to the high-voltage one.



(a) force-stretch diagram (b) voltage-charge diagram

Figure 4. The mechanism of an ideal four-stroke cycle.

In the following analysis, the DEG will be operated in the conversion cycle with four courses: starting with the initial stretch λ_0 at state A, the DEG is subject to a low-voltage battery and an equal-biaxial force stretches the membrane at the same time. The first process ends at state B, in which the maximum stretch achieves λ_{max}. In the process B→C, the low-voltage battery disconnects the membrane and the force is partly released to make the membrane thicker. In the process C→D, a high-voltage battery is applied to the membrane and the force is further released. The minimum stretch λ_{min} is attained at state D. In the last process, the equal-biaxial force stretches the membrane to the initial state in the open circuit condition. The complete energy conversion cycle is sketched in Fig. 5, where the initial state of the DEG is depicted by the color image, and the dotted line indicates the final state.

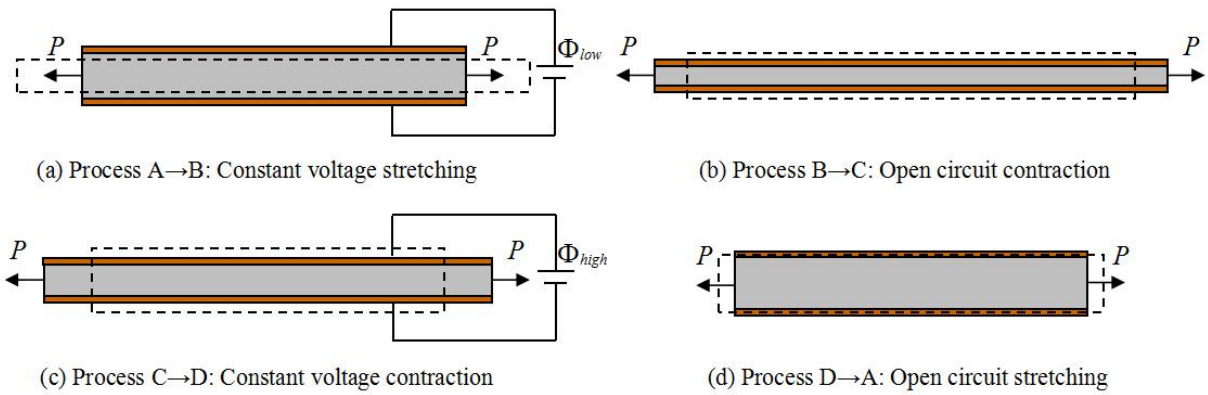


Figure 5. The DEG operates in different processes of the conversion cycle.

The previous study has proved theoretically that the pre-stretch can produce a larger nominal electric field and improve the stability of the elastomer [45]. Thus, the DEG is pre-stretched before activation. Because of viscoelasticity, the mechanical behavior of the DE membrane is affected by the viscoelastic relaxation time. Considering that the relaxation time reaches its maximum 87.216 s at 273 K, so the period of the cycle should be set above this value to give the membrane enough time to fully relax. Here, the period is set as t_{cycle}=90 s. By determining the initial and maximum stretches (λ_0=4 and λ_{max}=6), the stretch at state C (λ_C=4.8), and the low-level voltage Φ_{low}=1.8 kV, the minimum stretch and the high-level voltage are calculated by expressions $\Phi_{low}C(\lambda_{max})=\Phi_{high}C(\lambda_C)$ and $\Phi_{low}C(\lambda_0)=\Phi_{high}C(\lambda_{min})$. However, the capacitance

of the DEG will be altered by temperature, resulting in various results, as can be seen in Tab. 1. For simplicity, the average values are employed and the periods of different courses can be obtained at a constant stretch rate $|d\lambda/dt|=2\times(\lambda_{max}-\lambda_{min})/t_{cycle}$ (t_{AB}=33 s, t_{BC}=20 s, t_{CD}=25 s and t_{DA}=12 s). The stretch applied to the DEG is demonstrated in Fig. 6a, and the nominal electric field solved from Eq. (10) is fitted under different temperature situations and illustrated in Fig. 6b (the curve is assumed to be continuous and independent of temperature).

Table 1. The minimum stretch and high-level voltage calculated at different temperatures.

Temperature	273 K	293 K	313 K	333 K
λ_{min}	3.2954	3.3002	3.3048	3.3091
Φ_{high} (kV)	3.6918	3.6645	3.6389	3.6149

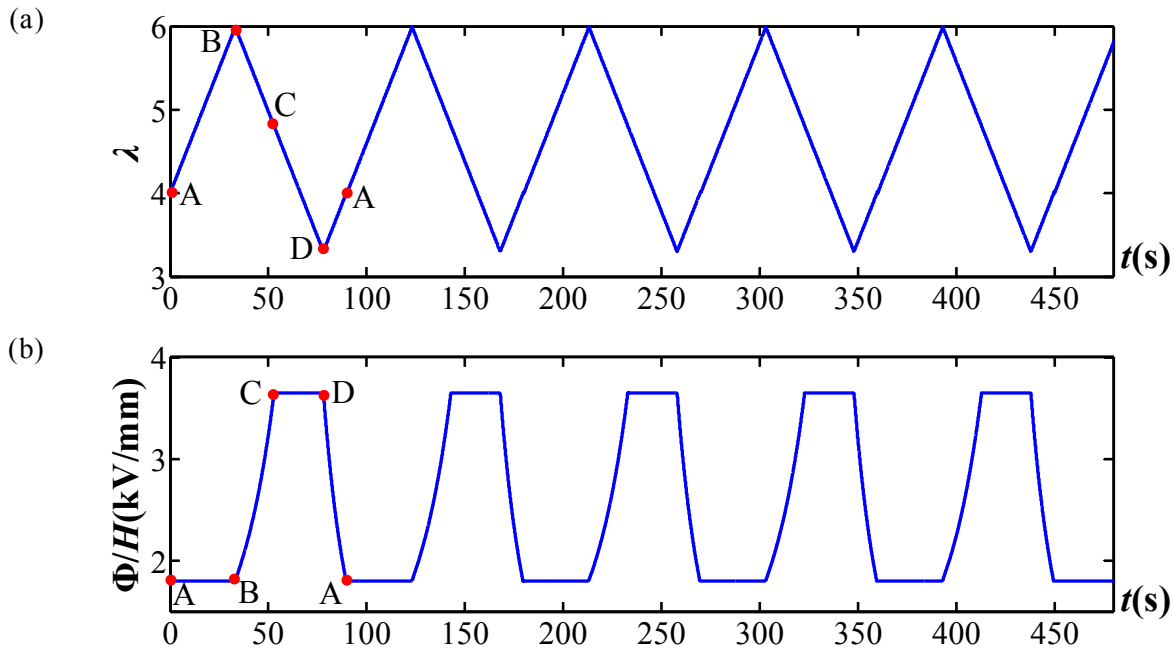


Figure 6. Cyclic loadings for the conversion. (a) The stretch; (b) The nominal electric field;

3.2. Mechanisms of failure modes and verification

As mentioned before, the behavior of the DE membrane is badly restricted by failure modes. Therefore, the cyclic loadings applied in Fig. 6 must locate in the allowable area to ensure the normal function of the DEG [43][44]. In this section, the mechanisms of the common failures are described and the feasibility of the cyclic loadings applied is validated.

3.2.1. Material rupture

The deformation of the DE membrane is finite because of the extension limits of the polymer. When polymer chains are seriously pulled, the chains approach the limit of configuration and the rupture may cause the failure of the generator. The critical stretch of the DEG may also be restricted by the stretch limit of the electrodes. However, more compliant electrodes are used in this work, letting the critical stretch of the generator be the stretch limit of DE. The critical condition for rupture is $\lambda=\lambda_R$, and the previous research has suggested that the critical stretch of VHB 4910 under equal-biaxial stretch is λ_R=6 [46].

3.2.2. Loss of tension

It is vital to maintain the membrane of DE in tension, as wrinkles can be caused easily by any compressive stress in plane, which may affect the normal operation of the DEG. The critical condition for LT is s=0 (P=0), and can be figured out by vanishing the nominal stress in Eq. (7), as illustrated in Eq. (13). Substitute the applied time-dependent stretch into Eq. (13), the boundaries of LT corresponding to different sampling temperatures can be plotted in Fig. 7.

$$\tilde{E}_{LT}^2 = \frac{\dfrac{\mu^A(T)(\lambda-\lambda^{-5})}{1-\dfrac{2\lambda^2+\lambda^{-4}-3}{J^A}} + \dfrac{\mu^B(T)(\lambda\xi^{-2}-\lambda^{-5}\xi^4)}{1-\dfrac{2\lambda^2\xi^{-2}+\lambda^{-4}\xi^4-3}{J^B}}}{\varepsilon_0(1.5a\lambda^2+\dfrac{b}{T}+c)\lambda^3} \tag{13}$$

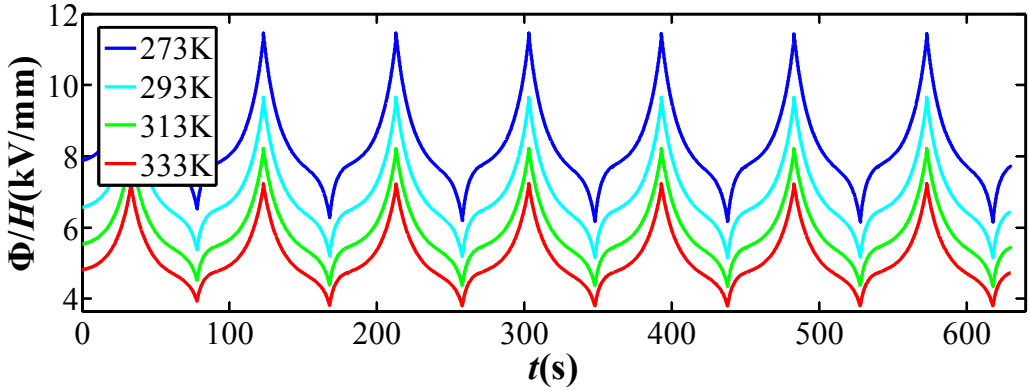


Figure 7. The critical condition for LT at different temperatures.

It is found that the boundary of LT will decline with a higher temperature and gradually reach a steady state after several cycles due to viscoelasticity. In this simulation, the lowest critical value for this failure mode is 3.79×10^6 V/m at T=333K, which is still larger than the highest nominal electric field applied. As a result, no failure of LT will take place.

3.2.3. Electrical breakdown

When the membrane is connected to a voltage source, electrical breakdown may happen with the increasing electric field, making the device fail. The electrical breakdown field is always assumed as constant in the previous studies [42][43], but the latter experiments on the VHB elastomer have indicated that the critical value of electrical breakdown E_{EB} depends on the stretch of DE [47][48], and the experimental data can be fitted in Eq. (14).

$$\tilde{E}_{EB} = E_{EB}(\lambda)\lambda^{-2} = E_{EB}(1)\lambda^{R-2} \tag{14}$$

where $E_{EB}(1)$ is the electrical breakdown field of the membrane of DE in the reference state, and the exponent coefficient R represents the sensitivity of the electrical breakdown field toward the stretch. Parameters for the VHB elastomer with the thickness H=1.0mm are fitted as $E_{EB}(1)$=30.6 MV/m and R=1.13 [20]. Substitute the time-dependent stretch into Eq. (14), it can be easily observed that the smallest critical value for EB 6.44×10^6 V/m is larger than the maximum of the applied nominal electric field, leading to no EB during the harvesting.

3.2.4. Electromechanical instability

Exposure to a voltage, the membrane of DE will compress its thickness, resulting in a higher true electric field. The positive feedback between the true electric field and the thickness may make the membrane thin down sharply, leading to EMI. The boundary of this failure can be defined as: when the mechanical forces are fixed, the voltage becomes a function of stretch. The peak of $\Phi(\lambda)$ corresponds to the critical condition for EMI. Differentiate Eq. (7) with respect to stretch, and the boundaries of EMI can be derived as:

$$\tilde{E}_{EMI}^2 = \frac{\left(\begin{array}{l} \dfrac{\mu^B(\lambda\xi^{-2}-\lambda^{-5}\xi^4)(4\lambda\xi^{-2}-4\lambda^{-5}\xi^4)}{(1-\dfrac{2\lambda^2\xi^{-2}+\lambda^{-4}\xi^4-3}{J^B})^2 J^B} + \dfrac{\mu^A(\lambda-\lambda^{-5})(4\lambda-4\lambda^{-5})}{(1-\dfrac{2\lambda^2+\lambda^{-4}-3}{J^A})^2 J^A} \\ + \dfrac{\mu^B(\xi^{-2}+5\lambda^{-6}\xi^4)}{(1-\dfrac{2\lambda^2\xi^{-2}+\lambda^{-4}\xi^4-3}{J^B})} + \dfrac{\mu^A(1+5\lambda^{-6})}{(1-\dfrac{2\lambda^2+\lambda^{-4}-3}{J^A})} \end{array} \right)}{(7.5a\lambda^4+\frac{3b}{T}\lambda^2+3c\lambda^2)\varepsilon_0} \tag{15}$$

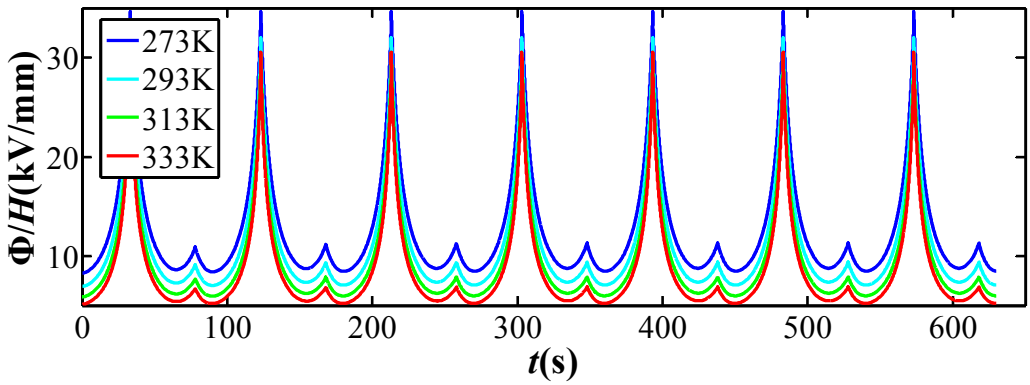


Figure 8. The critical condition for EMI at different temperatures.

Similar to the failure of LT, the boundary of EMI also depends on temperature and a higher temperature lowers the boundary of EMI. According to the numerical results, the minimum critical value for EMI is 5.12×10^6 V/m, which is larger than the highest nominal electric field applied. In summary, for the generator made of VHB 4910, the cyclic loadings applied will not lead to any failures within the temperature range 273K~333K.

4. Temperature effects on the performance of the DEG

Due to obvious viscoelastic relaxation, polymer chains in the membrane of DE will not attain stable configurations instantly. Therefore, it will take the DEG some time to possess steady mechanical parameters such as inelastic stretch and nominal stress. According to Eq. (9), the inelastic stretch can be obtained once the time-dependent stretch is determined. The inelastic stretches at different sampling temperatures are plotted in Fig. 9. It is observed that a higher temperature gives rise to a larger inelastic stretch, and the inelastic stretch will reach the peak value behind stage B in which the peak stretch achieves in every single cycle. Furthermore, the largest inelastic stretch will never exceed λ_{max}, and the curves will attain steady-state after the first five cycles in this study.

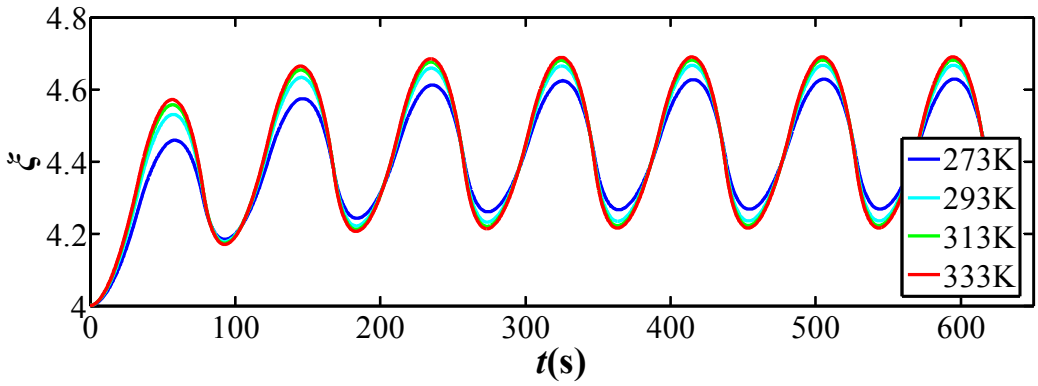


Figure 9. The inelastic stretch as a function of time at different temperatures.

Substitute the time-dependent stretch and inelastic stretch into Eq. (7), the nominal stress as a function of time can be depicted at different temperatures in Fig. 10. The curves show that the membrane of DE necessitates a larger equal-biaxial force at a lower temperature to satisfy the required operating condition. This phenomenon can be explained by the smaller modulus of DE induced by the higher temperature. Similarly, the nominal stress will attain steady-state after the first five cycles.

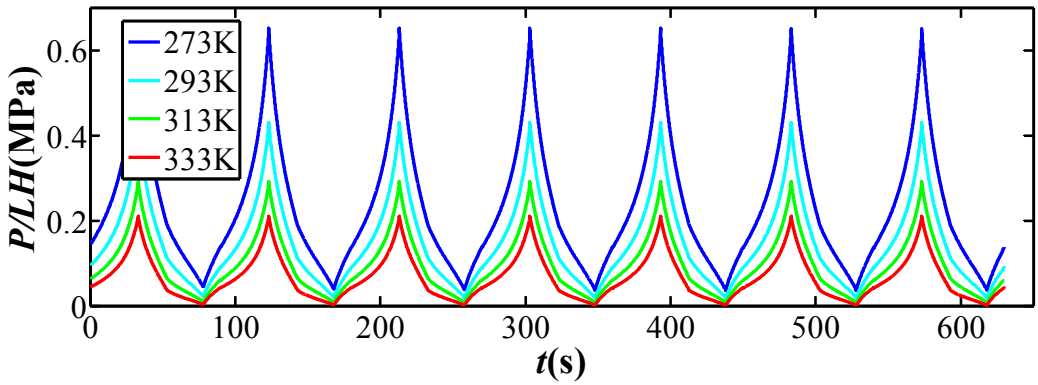


Figure 10. The nominal stress as a function of time at different temperatures.

The amount of the charges on the electrodes $\pm Q_p$ can be calculated from Eq. (10) and the temperature will affect the charges via the permittivity of DE, the factor b/T. The amounts of the charges on the electrodes and the leaked charges at different temperatures are illustrated in Fig. 11. It can be found from the curves that a higher temperature corresponds to a lower level of Q_p due to the negative effect of the temperature on the permittivity, and the amounts of the charges in process B→C and D→A are almost constant for the open circuit conditions. The amount of the leaked charges will increase after every single cycle due to the periodicity of the applied loads. The total amount of the charges through the conducting wire is obtained by the addition of Q_p and Q_{leak}, and will exhibit an increasing tendency as well.

Differentiate the charge-related parameters in Fig. 11 with respect to time, the magnitudes of currents can be solved with temperature variation. Similar to the tendencies of the charges, a higher temperature will make the magnitude of the current through the electrodes decline.

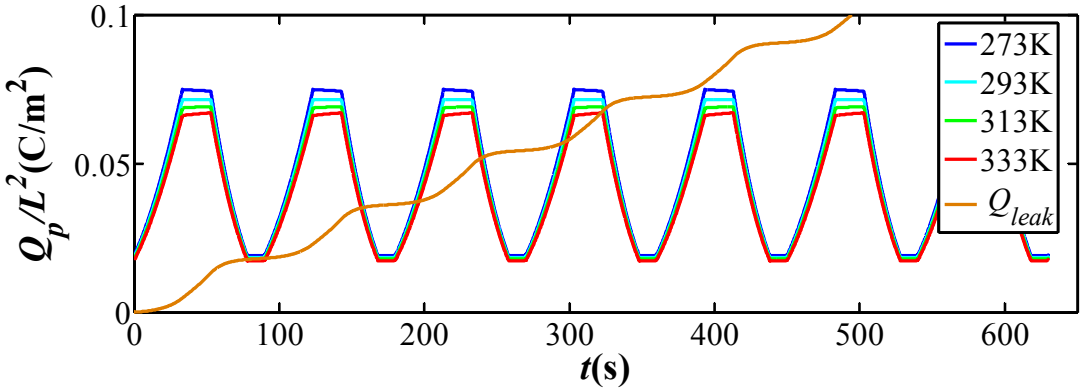


Figure 11. The amount of the charges on the electrodes as a function of time at different temperatures as well as the leaked charges.

For an ideal DEG, the energy input includes the mechanical work done by the equal-biaxial force and the electrical energy provided by the low-voltage battery, and the energy output is measured by the electrical energy pumped to the high-voltage battery. However, during the conversion, some parts of the energy will be lost, including the mechanical energy dissipated by the inelastic stretch and the electrical energy lost by the leaked current. The mathematical expressions of different kinds of energies are described as follows:

The mechanical work done per cycle depends on the stretch of the membrane λ and the magnitude of the equal-biaxial force P, which will attain steady-state after several cycles, as shown in Fig. 10. The total mechanical energy in a single cycle can be calculated by $2\int P\mathrm{d}(\lambda L)$. Since no specific dimensions of the original configuration is defined in this study, it is of great convenience to adopt the energy density to characterize the amplitude of energy. Thus, the mechanical energy density can be presented in Eq. (16).

$$w_{mech}=2\int (P/LH)\,\mathrm{d}\,\lambda \tag{16}$$

The membrane of DE will dissipate the mechanical energy partially through the work done by the dashpot during the deformation. As mentioned before, the stress acting on the dashpot equals the stress in spring B, and the nominal stress of the dashpot is presented by $s^B=\sigma^B/\xi$. Similarly, the energy density of viscous loss w_{visc} is given in Eq. (17). Moreover, the electrical energy dissipated by the current leakage which relies on the magnitude of voltage Φ and the amount of charges that leak through the electrodes Q_{leak} is described in Eq. (18).

$$w_{visc}=2\int s^B\,\mathrm{d}\,\xi \tag{17}$$

$$w_{leak}=\int (\Phi/H)\,\mathrm{d}(Q_{leak}/L^2) \tag{18}$$

Based on the condition of energy flow, the mechanical work done by the equal-biaxial force makes up the absolute input part, the energy dissipated by the inelastic stretch and the leaked current constitutes the loss part, and the difference between the electrical energy pumped to the high-voltage battery and the energy absorbed from the low-voltage battery stands for the absolute output part.

The mechano-electrical conversion efficiency described by α can be obtained from Eq. (19). It is obvious that the efficiency of the DEG operates at α=1 without any losses. If the amount of the energy lost is more than the mechanical energy, the generator will operate at a negative efficiency with no energy generated.

$$\alpha=1-(w_{visc}+w_{leak})/w_{mech} \tag{19}$$

Meanwhile, the electrical energy that contains the electrical energy generated and the energy lost by the leaked current is also figured out to examine the correctness of the energies above. This electrical energy can be calculated from Eq. (20).

$$w_{pele}=-\int(\Phi/H)\,\mathrm{d}(Q_p/L^2) \tag{20}$$

The relevant variables are taken from the seventh cycle (steady-state) in the response curves in Figs. 9-11 to calculate the integrals above. Specifically, the mechanical energy density is integrated over the cycle on the force (P/LH)-length (λ) diagram, as shown in Fig. 12a. The density of the viscous loss is integrated over the cycle on the force (s^B)-length (ξ) diagram, as shown in Fig. 12b. The density of the electrical energy containing the electrical loss and the generated part is integrated over the cycle on the voltage (Φ/H)-charge (Q_p/L^2) diagram, as shown in Fig. 12c. The density of the electrical loss is integrated over any single cycle of the voltage (Φ/H)-charge (Q_{leak}/L^2) diagram, as shown in Fig. 12d.

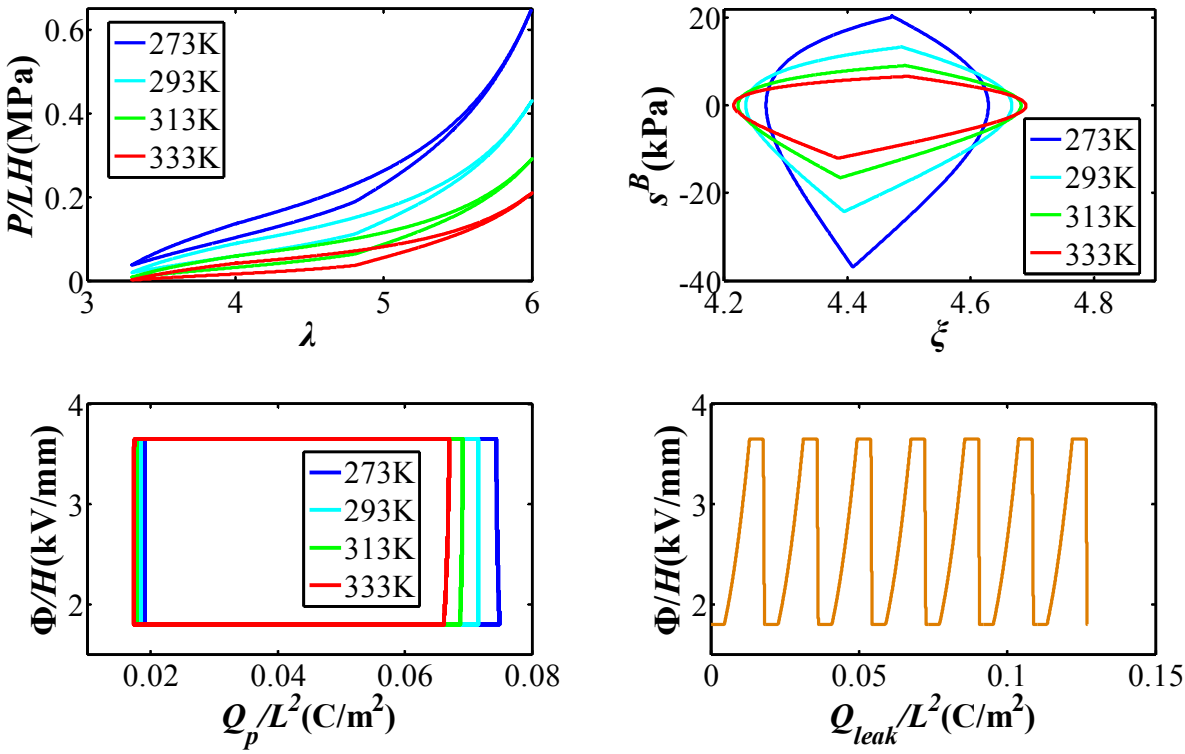


Figure 12. The curves characterize the energies of different kinds. (a) The mechanical energy produced by the force; (b) The energy dissipated by viscoelasticity; (c) The electrical energy involving the electrical loss and the generated part ; (d) The energy loss due to current leakage;

The values of these energy densities at different sampling temperatures are outlined in the following table. The second column of Tab. 2 indicates that the mechanical energy decreases with a higher temperature, and the temperature tends to have less influence on the mechanical energy (the difference of w_{mech} is 9800 J/m^3 from 273 K to 293 K and 7000 J/m^3 from 313 K to 333 K). The third column also shows that the DEG dissipates less mechanical energy at a higher temperature. Due to the direct control on the stretch in this study, the electrical energy lost by current leakage is independent of temperature. The conversion efficiency is listed in the fifth column and it is found that a higher conversion efficiency can be realized if the DEG is operated at a low temperature. The reason is that the reduction in temperature gives rise to relatively more mechanical energy than the viscous loss. However, the temperature effect on the efficiency will get smaller when the temperature keeps rising. The electrical energy which has the same trend as the mechanical energy is also listed in the last column. Furthermore, it is observed that the mechanical energy w_{mech} almost equals the sum of the viscous loss w_{visc} and the electrical energy w_{pele} involving the electrical loss w_{leak} and the absolute output part, satisfying the conservation of energy.

Table 2. Energy densities and conversion efficiency at different temperatures.

T (K)	w_{mech} (J/m^3)	w_{visc} (J/m^3)	w_{leak} (J/m^3)	α	w_{pele} (J/m^3)
273	1.287e+05	2.598e+04	4.842e+04	42.19%	1.027e+05
293	1.189e+05	2.052e+04	4.842e+04	42.01%	9.835e+04
313	1.095e+05	1.521e+04	4.842e+04	41.89%	9.427e+04
333	1.025e+05	1.124e+04	4.842e+04	41.78%	9.124e+04

6. Conclusions

In this study, with combination of the temperature-dependent permittivity and shear modulus and major dissipation processes including viscoelasticity and current leakage, an integrated model focusing on the influences of temperature on the performance of DEGs is established. On the basis of a specific energy conversion cycle, the performance parameters including the energy density and conversion efficiency can be all figured out at various temperatures. It is noticed that the DEG operates more efficiently at a lower temperature owing to the relatively more enhancement in the input mechanical energy than the viscous loss. In the meantime, the failure modes are considered at different temperatures to ensure the normal operation of the device. It is observed that the generator is more likely to suffer from LT and EMI at a higher temperature. The simulation in this study may offer great help and guideline in the design and optimization of energy harvesting with different temperature conditions, which can contribute to a more efficient dissipative DEG.

Acknowledgements

The authors wish to thank the support of National Natural Science Foundation of China (Grant No.61232014) and Project funded by China Postdoctoral Science Foundation. The authors also wish to thank Research Project of State Key Laboratory of Structural Analysis for Industrial Equipment (Grant No.GZ1403), the Science Fund of State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body (Grant No.51375001), and Research Project of State Key Laboratory of Mechanical Systems and Vibration MSV201711.

References

[1]. O'Halloran, A., O'Malley, F. and McHugh, P. (2008) A review on dielectric elastomer actuators, technology, applications, and challenges. *Journal of Applied Physics* **104**, 071101.

[2]. Suo, Z. G. (2010) Theory of dielectric elastomers. *Acta Mechanics Solids Sinica* **23**, 549-578.

[3]. Kornbluh, R. D., Pelrine, R., Prahlad, H., Wong-Foy, A., Mccoy, B. and Kim, S., et al. (2011) From boots to buoys: promises and challenges of dielectric elastomer energy harvesting. *Proceedings of SPIE* **7976**, 67-93.

[4]. Pelrine, R., Kornbluh, R., Pei, Q. and Joseph, J. (2000) High-speed electrically actuated elastomers with strain greater than 100%. *Science* **287**, 836-839.

[5]. Brochu, P. and Pei, Q. (2010) Advances in dielectric elastomers for actuators and artificial muscles. *Macromolecular Rapid Communications* **31**, 10-36.

[6]. Tsuchitani, S., Sunahara, T. and Miki, H. (2015) Dielectric elastomer actuators using Slide-Ring Material® with increased permittivity. *Smart Materials & Structures* **24**, 065030.

[7]. Giousouf, M. and Kovacs, G. (2013) Dielectric elastomer actuators used for pneumatic valve technology. *Smart Materials & Structures* **22**, 104010.

[8]. Adrian Koh, S. J., Keplinger, C., Li, T. F. and Bauer, S. (2011) Dielectric Elastomer Generators: How Much Energy Can Be Converted? *IEEE/ASME Transactions On* **16**, 33-41.

[9]. Lv, X. F., Liu, L. W., Liu, Y. J. and Leng, J. S. (2015) Dielectric elastomer energy harvesting: maximal converted energy, viscoelastic dissipation and a wave power generator. *Smart Materials & Structures* **24**, 115036.

[10]. Kornbluh, R. D. and Pei, Q. (2001). Dielectric elastomers: generator mode fundamentals and applications. *Proceedings of SPIE* **4329**, 148-156.

[11]. Rome, L. C., Flynn, L., Goldman, E. M. and Yoo, T. D. (2005) Generating electricity while walking with loads. *Science* **309**, 1725-1728.

[12]. Jean-mistral, C. (2008). Dielectric polymer: scavenging energy from human motion. *Proceedings of SPIE* **6927**, 692716-692716-10.

[13]. Chiba, S., Waki, M., Kornbluh, R. and Pelrine, R. (2009) Innovative wave power generation system using electroactive polymer artificial muscles. *Oceans IEEE*, 1-3.

[14]. Chiang Foo, C., Cai, S., Koh, S. J. A. and Bauer, S. (2012) Model of dissipative dielectric elastomers. *Journal of Applied Physics* **111**, 836-382.

[15]. Liu, L. W., Zhang, Z., Liu, Y. J. and Leng, J. S. (2014) Failure modeling of folded dielectric elastomer actuator. *Science China Physics, Mechanics & Astronomy* **57**, 263-272.

[16]. Zhao, X. H. and Suo, Z. G. (2007) Method to analyze electromechanical stability of dielectric elastomers. *Applied Physics Letter* **91**, 061921.

[17]. Zhou, J., Jiang, L. and Khayat, R. E. (2014) Electromechanical response and failure modes of a dielectric elastomer tube actuator with boundary constraints. *Smart Materials & Structures* **23**, 045028.

[18]. Sheng, J. J., Chen, H. L. and Li, B. (2011) Effect of temperature on the stability of dielectric elastomers. *Journal of Physics D Applied Physics* **44**, 365406.

[19]. Liu, L. W., Liu, Y. J., Yu, K. and Leng, J. S. (2014) Thermoelectromechanical stability of dielectric elastomers undergoing temperature variation. *Mechanics of Materials* **72**, 33-45.

[20]. Plante, J. S. and Dubowsky, S. (2006). Large-scale failure modes of dielectric elastomer actuators. *International Journal of Solids & Structures* **43**, 7727-7751.

[21]. Keplinger, C., Li, T. F., Baumgartner, R., Suo, Z. G. and Bauer, S. (2012) Harnessing snap-through instability in soft dielectrics to achieve giant voltage-triggered deformation. *Soft Matter* **8**, 285-288.

[22]. Leng, J. S., Liu, L. W., Liu, Y. J., Yu, K. and Sun, S. H. (2009) Electromechanical stability of dielectric elastomer. *Applied Physics Letters* **94**, 211901.

[23]. Zhao, X. H., Adrian Koh, S. J. and Suo, Z. G. (2011) Nonequilibrium thermodynamics of dielectric elastomers. *International journal of Applied Mechanics* **3**, 203-217.

[24]. Li, B., Zhang, J., Liu, L. and Chen, H. (2014). Modeling of dielectric elastomer as electromechanical resonator. *Journal of Applied Physics* **116**, 124509-124509-9.

[25]. Qiang, J. H., Chen, H. L. and Li, B. (2012) Experimental study on the dielectric properties of polyacrylate dielectric elastomer. *Smart Materials & Structures* **21**, 025006.

[26]. Jean-Mistral, C., Sylvestre, A., Basrour, S. and Chaillout, J. J. (2010) Dielectric properties of polyacrylate thick films used in sensors and actuators. *Smart Materials & Structures* **19**, 075019.

[27]. Zhao, X. H., Hong, W. and Suo, Z. G. (2007) Electromechanical coexistent states and hysteresis in dielectric elastomers. *Physics Review B* **76**, 1283.

[28]. Zhang, J. S., Tang, L. L., Li, B., Wang, Y. J. and Chen, H. L. (2015) Modeling of the dynamic characteristic of viscoelastic dielectric elastomer actuators subject to different conditions of mechanical load. *Journal of Applied Physics* **117**, 084902.

[29]. Wang, H. M., Lei, M., and Cai, S. Q. (2013) Viscoelastic deformation of a dielectric elastomer membrane subject to electromechanical loads. *Journal of Applied Physics* **113**, 213508-213508-7.
[30]. Zhang, J. S., Chen, H. L., Tang, L. L., Li, B., Sheng, J. J. and Liu, L. (2015) Modeling of spring roll actuators based on viscoelastic dielectric Elastomers. *Applied Physics A*. **119**, 825-835.
[31]. Zhou, J. Y., Jiang, L. Y. and Khayat, R. E. (2014) Viscoelastic effects on frequency tuning of a dielectric elastomer membrane resonator. *Journal of Applied Physics* **115**, 124106-124106-11.
[32]. Li, B., Chen, H. L., Qiang, J. H. and Zhou, J. X. (2012) A model for conditional polarization of the actuation enhancement of a dielectric elastomer. *Soft Matter* **8**, 311-317.
[33]. Gent, A. N. (1996) A new constitutive relation for rubber. *Rubber Chemistry and Technology* **69**, 59-61.
[34]. Michel, S., Zhang, X. Q., Wissler, M., Löwe, C. and Kovacs, G. (2010) A comparison between silicone and acrylic elastomers as dielectric materials in electroactive polymer actuators. *Polymer international* **59**, 391-399.
[35]. Vu-Cong, T., Jean-Mistral, C. and Sylvestre, A. (2013) New operating limits for applications with electroactive elastomer: effect of the drift of the dielectric permittivity and the electrical breakdown. *SPIE - the international Society for Optical Engineering* **30**, 793-806.
[36]. Sheng, J. J., Chen, H. L., Liu, L., Zhang, J. S., Wang, Y. Q., and Jia, S. H. (2013) Dynamic electromechanical performance of viscoelastic dielectric elastomers. *Journal of Applied Physics* **114**, 134101-134101-8.
[37]. Bergamini, A. (2011) Dielectric and insulating properties of an acrylic DEA material at high near-DC electric fields. *Proceedings of SPIE* **7976**, 79763B-79763B-6.
[38]. Gisby, T. A., Xie, S. Q., Calius, E. P. and Anderson, I. A. (2010) Leakage current as a predictor of failure in dielectric elastomer actuators. *Proceedings of SPIE* **7642**, 764213-764213-11.
[39]. Graf, C. and Maas, J. (2011) Evaluation and optimization of energy harvesting cycles using dielectric elastomers. *Electroactive Polymer Actuators & Devices* **7976**, 541-558.
[40]. Graf, C., Maas, J. and Schapeler, D. (2010) Energy harvesting cycles based on electro active polymers. *Proceedings of SPIE* **7642**, 579-595.
[41]. Graf, C., Eitzen, L. and Maas, J. (2011). Multilevel high voltage converter driving dielectric elastomer generators. *European Conference on Power Electronics and Applications IEEE Xplore*, 1-10.
[42]. Chiang Foo, C., Adrian Koh, S. J., Keplinger, C. and Kaltseis, R. (2012). Performance of dissipative dielectric elastomer generators. *Journal of Applied Physics* **111**, 836-839.
[43]. Adrian Koh, S. J., Zhao, X. H. and Suo, Z. G. (2009) Maximal energy that can be converted by a dielectric elastomer generator. *Applied Physics Letter* **94**, 262902-262902-3.
[44]. Bortot, E., Gei, M. and Debotton, G. (2015) Optimal energy harvesting cycles for load-driven dielectric elastomer generators under equibiaxial deformation. *Meccanica* **50**, 2751-2766.
[45]. Wissler, M. and Mazza, E. (2005) Modeling of a pre-strained circular actuator made of dielectric elastomers, *Sensors and Actuators A: Physical* **120**, 184-192.
[46]. Sommer-larsen, P., and Kofod, G. (2002). Performance of dielectric elastomer actuators and materials. *Proceedings of SPIE* 158-166.
[47]. Kofod, G., Sommer-larsen, P., Kornbluh, R. D. and Pelrine, R. (2003). Actuation response of polyacrylate dielectric elastomers. *Journal of Intelligent Material Systems & Structures* **14**, 787-793.
[48]. Tröls, A., Kogler, A., Baumgartner, R., Kaltseis, R., Keplinger, C. and Schwödiauer, R., et al. (2013). Stretch dependence of the electrical breakdown strength and dielectric constant of dielectric elastomers. *Smart Materials & Structures* **22**, 104012.

A Stabilized ES-FEM for Incompressible flow based on Quasi-implicit Characteristic-based Polynomial Pressure Projection

***† Chen Jiang [1], Xu Han [1], G.R. Liu[2,3]**

[1]State Key Laboratory of Advanced Technology of Design and Manufacturing for Vehicle Body, Hunan University, P.R. China, 410082

[2] Part-time Consultant Professor, State Key Laboratory of Advanced Technology of Design and Manufacturing for Vehicle Body, Hunan University, P.R. China, 410082

[3] Department of Aerospace Engineering and Engineering Mechanics, University of Cincinnati, 2851 Woodside Dr, Cincinnati, OH 45221, USA

*Presenting author: jiangchen2007@hotmail.com

†Corresponding author: hanxu@hnu.edu.cn

Abstract

In this paper, a stabilized edge-based smoothed finite element method (ES-FEM) is proposed to solve incompressible fluid flow problems. To stabilize the convection caused oscillation, characteristic-based method is employed. To stabilize the pressure caused oscillation, the polynomial pressure projection (P3) is used. The proposed stabilization is denoted as CBP3 scheme. One excellent merit of CBP3 scheme stabilization is that it suits for all types of element, even for the simplest 3-node triangular element. The CBP3 scheme implemented into ES-FEM with T3 element is denoted as CBP3/ES-FEM-T3. To verify proposed CBP3 scheme and check its stability, the benchmark, Taylor-Green vortex, is calculated. As a comparison, this benchmark is also calculated by CBP3/FEM-T3. From the solutions of CBP3/FEM-T3 and CBP3/ES-FEM-T3, CBP3 scheme is proved with ability to stabilize FEM and ES-FEM. The convergence studies carried out reveal that CBP3/ES-FEM-T3 beats CBP3/FEM-T3 on accuracies of both velocity and pressure.

Keywords: ES-FEM, Chracteristic-based method, Polynomial pressure projection, Triangular element, Incompressible laminar flow

Introduction

In solid mechanics, the edge-based smoothed finite element method (ES-FEM) has been proven by vast literature [1–5] with much better accuracy and convergence than FEM when using T3 element. Due to the gradient smoothing using edge-based smoothing domain, the strain distribution of ES-FEM in T3 mesh of whole system is smoother than FEM. Therefore, ES-FEM can soft the stiff behavior of T3 element to achieve better solution. Since T3 element can discretize complex geometry by well-developed unstructure mesh generators, one technique which can improve standard FEM T3 element, such as ES-FEM, will be welcomed by engineers.

As a result, it is also attractive to exploit the benefits of ES-FEM in fluid dynamics. The motivation of this paper is to extend ES-FEM using T3 element to solve laminar flow. But unlike the solid governing equation which is self-adjoint Laplace type equation, the Navier-Stokes equations are non self-adjoint. This non self-adjoint feature is derived from the nonlinear convection of fluid flow. Previous investigations about standard Galerkin weighted

residual method, such as FEM, for Navier-Stokes equations [6] have shown the spatial oscillation. The ES-FEM also has to face this convection caused oscillation. On the other hand, ES-FEM should also equip with suitable stabilization for pressure oscillation in incompressible flow. The reason why first calculates incompressible laminar flow using ES-FEM is that compressible flow is quite complicate in physics, such as shock waves and high Reynolds number. Usually, incompressible constraint is circumvented by selective S-FEM in solid mechanics. This time, the selective S-FEM approach is not applicable due to the fully incompressibility shown in fluid material behavior. Previous study made by the authors using selective reduced integration of 4-node quadrilateral elements has demonstrated bulk viscosity (analogous to bulk modulus in solid mechanics) must reach to 10e8 times of shear dynamic viscosity to obtain satisfied results. But selective S-FEM shown very violent oscillation in pressure under that high penalty factor.

Therefore, good convection and pressure stabilizations are the prerequisites for successfully using ES-FEM to solve fluid flows. Besides, the potential stabilizations must be friendly to T3 element. We do can successfully solve fluid flows by using higher order elements with easy-to-create unstructured mesh. However, higher order elements, such as 6-node triangular element, consume much more computation resources than T3 elements. Another consideration is that people will not have the impetus to add gradient smoothing to the quite accurate higher order elements. In this paper, we chose widely used characteristic-based method proposed by Zienkiewicz [7] as convection stabilization. On the pressure stabilization, the polynomial pressure projection proposed (P3) by Dohrmann and Bochev [8] is deployed. This P3 pressure stabilization has been demonstrated appropriate for any element type for any level of incompressiblity, including simplest T3 element. In its previous cooperation with S-FEMs for solid mechanics and Stokes flow, P3 stabilization can further soft the behaviors of FEM and S-FEM. Especially, ES-FEM with P3 stabilization outperformed FEM and other S-FEMs for incompressible solid problems using T3 element. Combine aforementioned two stabilizations, we proposed a new stabilization scheme named as Characteristic-based Polynomial Pressure Projection (CBP3). With this CBP3 scheme, FEM and ES-FEM should able to solve incompressible laminar flow.

The rest of the paper is structured as following. The second section introduces the governing equations of incompressible laminar flow. The third section is the brief derivations of CBP3 scheme. The following section describes the merge of CBP3 into ES-FEM. The forth section is the case studies to access the accuracy of proposed methods. In the last section, conclusions are drawn.

Governing equations

If flow is incompressible without heat transfer, the dimensional conservative N-S equations can be simplified to:

$$\begin{cases} \dfrac{\partial v_i}{\partial x_i} = 0, \\ \dfrac{\partial v_i}{\partial t} + \dfrac{\partial}{\partial x_j}(v_j v_i) = -\dfrac{1}{\rho}\dfrac{\partial p}{\partial x_i} + \dfrac{\partial \tau_{ij}}{\partial x_j} + g_i. \end{cases} \tag{1}$$

In which, p is the fluid pressure, v_i is the i-th component of fluid velocity, ρ is the fluid density which is a constant in incompressible flow, g_i is the gravity acceleration, and τ_{ij} is the deviatoric stresses in the fluid:

$$\tau_{ij} = \frac{\mu}{\rho}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\frac{\partial v_k}{\partial x_k}\delta_{ij} \right). \tag{2}$$

where, μ is the fluid dynamic (shear) viscosity and δ_{ij} is the Kronecker delta.

The boundary conditions of N-S equations are the Velocity Boundary Condition (VBC), the Pressure Boundary Condition (PBC) and Traction Boundary Condition (TBC) as below. Meanwhile, the Initial Condition (IC) is also presented below.

$$\begin{aligned} &VBC: v_i = \bar{v}_i, \text{ on } \Gamma_v. \\ &PBC: p = \bar{p}, \text{ on } \Gamma_p. \\ &TBC: t_i = n_j \sigma_{ij} = \bar{t}_i, \text{on } \Gamma_t. \\ &IC: v_i^0 = \bar{v}_i^0,\ p^0 = \bar{p}^0. \end{aligned} \tag{3}$$

where $\Gamma_v \cup \Gamma_t = \Gamma$ and $\Gamma_v \cap \Gamma_t = 0$.

Characteristic-based polynomial pressure projection scheme

Characteristic-based stabilization

When using Galerkin weighted residual numerical methods, like FEM, for N-S equation, the nonlinear non-self-adjoint convection term will cause the numerical oscillation. Since ES-FEM is also based on Galerkin weighted residual method, the characteristic-based stabilization is applied in this paper to cure the convection instability.

In this study, we select the Characteristic-based stabilization which leads to a Characteristic-Galerkin (CG) scheme. This method has a very solid mathematical background which will be demonstrated on one-dimensional scalar convection-diffusion equation,

$$\frac{\partial \phi}{\partial t} + \underbrace{U\frac{\partial \phi}{\partial x}}_{convective} - \underbrace{\frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x} \right)}_{diffusive} + Q = 0 \tag{4}$$

where ϕ is a scalar quantity being transported by the velocity U, and k is the diffusion coefficient.

If we change above equation into a new coordinate system x' which is defined as

$$dx' = dx - Udt \tag{5}$$

Because a physical quantity must not vary if we change our coordinate system, we can get

$$\phi(x,t) = \phi(x',t) \tag{6}$$

Then, we can get following relation

$$\frac{\partial \phi(x,t)}{\partial t} = \frac{\partial \phi(x',t)}{\partial t} + \frac{\partial \phi(x',t)}{\partial x'}\frac{\partial x'}{\partial t} = \frac{\partial \phi(x,t)}{\partial t} - U\frac{\partial \phi(x',t)}{\partial x'} \tag{7}$$

Substitute above relation into Eq.(1.6), we can simplify it at new coordinate system as

$$\frac{\partial \phi}{\partial t} - \frac{\partial}{\partial x'}\left(k\frac{\partial \phi}{\partial x'} \right) + Q(x') = 0 \tag{8}$$

In above equation, it has no convective term any more in this new coordinate system. Furthermore, the equation becomes self-adjoint.

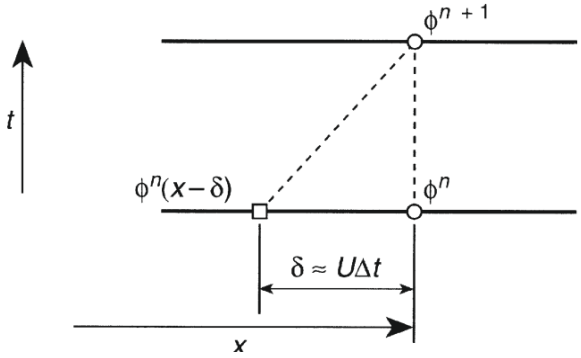


Figure 1. The simple explicit characteristic-Galerkin procedure [9].

We should also pay attention to the new coordinate system on characteristic direction is updated at every time step. The meaning of updating coordinate system is our Euler fluid mesh must moves like a Lagrangian solid mesh. However, this mesh updating is time consuming and will distort mesh. Thus, a simple explicit characteristic-Galerkin procedure was first proposed in 1984 [10] by using the local Taylor expansion on characteristic directions as Figure 1. The detail derivation of this explicit characteristic-Galerkin procedure can be found in reference [6]. We simply write down its final form with Euler time integration,

$$\begin{aligned}\Delta\phi = \phi^{n+1} - \phi^n = -\Delta t\left[U^n \frac{\partial \phi}{\partial x} - \frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right) + Q\right]^n \\ + \frac{\Delta t^2}{2} U^n \frac{\partial}{\partial x}\left[U^n \frac{\partial(\phi)}{\partial x} - \frac{\partial}{\partial x}\left(k\frac{\partial \phi}{\partial x}\right) + Q\right]^n .\end{aligned} \tag{9}$$

Its extension to multi-dimensional can be straightforwardly written as

$$\begin{aligned}\Delta\phi = \phi^{n+1} - \phi^n = -\Delta t\left[\frac{\partial U_j \phi}{\partial x_j} - \frac{\partial}{\partial x_i}\left(k\frac{\partial \phi}{\partial x_i}\right) + Q\right]^n \\ + \frac{\Delta t^2}{2} U_k^n \frac{\partial}{\partial x_k}\left[\frac{\partial(U_j\phi)}{\partial x_j} - \frac{\partial}{\partial x_i}\left(k\frac{\partial \phi}{\partial x_i}\right) + Q\right]^n .\end{aligned} \tag{10}$$

Similar procedures can be conducted for incompressible N-S equations as Eq.(1). Here, we only present the final form of explicit characteristic-Galerkin scheme for N-S equations with Euler time integration.

$$\begin{aligned}v_i^{n+1} - v_i^n = \Delta t\left[\underbrace{-\frac{\partial(v_j v_i)}{\partial x_j}}_{convective} \underbrace{-\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{1}{\rho}\frac{\partial \tau_{ij}}{\partial x_j}}_{diffusive} + g_i\right]^n \\ \underbrace{-\frac{\Delta t^2}{2} v_k \frac{\partial}{\partial x_k}\left[-\frac{\partial(v_j v_i)}{\partial x_j} - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{1}{\rho}\frac{\partial \tau_{ij}}{\partial x_j} + g_i\right]^n}_{stabilization} .\end{aligned} \tag{11}$$

If calculate the diffusive terms at n+1 time step to the left side of equation, we can achieve the quasi-implicit characteristic-Galerkin scheme as below

$$\begin{aligned} v_i^{n+1} + \underbrace{\frac{1}{\rho}\frac{\partial p^{n+1}}{\partial x_i} - \frac{1}{\rho}\frac{\partial \tau_{ij}^{n+1}}{\partial x_j}}_{diffusive} &= v_i^n + \Delta t \left[\underbrace{-\frac{\partial\left(v_j v_i\right)}{\partial x_j}}_{convective} \right]^n + g_i \\ \underbrace{-\frac{\Delta t^2}{2} v_k \frac{\partial}{\partial x_k}\left[-\frac{\partial\left(v_j v_i\right)}{\partial x_j} - \frac{1}{\rho}\frac{\partial p}{\partial x_i} + \frac{1}{\rho}\frac{\partial \tau_{ij}}{\partial x_j} + g_i \right]^n}_{stabilization} &. \end{aligned} \tag{12}$$

Due to only diffusive terms are calculated implicitly, this scheme is denoted as quasi-implicit scheme.

Characteristic-Galerkin semi-discretized form

Interpolate the velocity and pressure using shape function Φ,

$$v_i(\mathbf{x}) = \sum_{I=1}^{N_n} \Phi(\mathbf{x}_I) v_i(\mathbf{x}_I) = \sum_{I=1}^{N_n} \Phi_I v_{Ii},\ p(\mathbf{x}) = \sum_{I=1}^{N_n} \Phi(\mathbf{x}_I)\, p(\mathbf{x}_I) = \sum_{I=1}^{N_n} \Phi_I p_I \tag{13}$$

where N_n is total number of nodes in the model, $\Phi(\mathbf{x}_I)$ is the shape function of node *I*, $v_i(\mathbf{x}_I)$ is the velocity of node *I*.

Combine time discretized momentum equation Eq.(12) with the continuity equation in Eq.(1), the semi-discretized weak form is directly provided as below

$$\begin{bmatrix} \dfrac{M_{IJ}}{\Delta t} + K_{IJ} & -Q_{IiJ} \\ -G_{IJi} & 0 \end{bmatrix} \begin{bmatrix} v_{Ji}^{n+1} \\ p_J^{n+1} \end{bmatrix} = \begin{bmatrix} M_{IJ}\dfrac{v_{Ji}^n}{\Delta t} - C_{IJ}^n v_{Ji}^n - \dfrac{\Delta t}{2} H_{IJ}^n v_{Ji}^n + f_{Ii}^t + f_{Ii}^g \\ 0 \end{bmatrix} \tag{14}$$

where

$$\begin{aligned} &M_{IJ} = \int_\Omega \rho \Phi_I \Phi_J \mathrm{d}\Omega, [C]_{IJ}^n = \int_\Omega \rho \Phi_I \frac{\partial\left(v_j^n \Phi_J\right)}{\partial x_j} \mathrm{d}\Omega, \\ &[K]_{IJ} = \int_\Omega \mathbf{B}_I^T \mathbf{D}_\mu \mathbf{B}_J \mathrm{d}\Omega, [H]_{IJ}^n = \int_\Omega \frac{\partial(v_k^n \Phi_I)}{\partial x_k} \rho \frac{\partial(v_j^n \Phi_J)}{\partial x_j} \mathrm{d}\Omega, \\ &[Q]_{IiJ} = \int_\Omega \rho \frac{\partial\left(\Phi_I\right)}{\partial x_i} \Phi_J \mathrm{d}\Omega, [G]_{IJi} = \int_\Omega \rho \Phi_I \frac{\partial\left(\Phi_J\right)}{\partial x_i} \mathrm{d}\Omega, \\ &[f]_{Ii}^{\mathrm{g}} = \int_\Omega \Phi_I \rho g_i \mathrm{d}\Omega, [f]_{Ii}^t = \int_\Gamma \Phi_I \left(\tau_{ij}^n\right) n_j \mathrm{d}\Gamma. \end{aligned} \tag{15}$$

In above equation,

$$\mathbf{B}_I = \begin{bmatrix} \dfrac{\partial \Phi_I}{\partial x} & 0 \\ 0 & \dfrac{\partial \Phi_I}{\partial y} \\ \dfrac{\partial \Phi_I}{\partial y} & \dfrac{\partial \Phi_I}{\partial x} \end{bmatrix}, \ \mathbf{D}_u = \mu \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{16}$$

Although Eq.(14) is standard weak form for quasi-implicit characteristic-Galerkin scheme, its lower diagonal term of left side matrix is zero which doesn't circumvent the LBB restriction for incompressibility. Hence, directly using Eq.(14) will cause pressure locking or instability.

Polynomial pressure projection

Previous widely used technique to satisfy LBB condition is the Taylor-Hood element [11], using higher order element for velocity and one order lower element for pressure. However, Taylor-Hood elements introduced much more degree of freedoms than equal-order elements for both velocity and pressure. Another widely used approach is the selective reduced integration. This approach has limitation on the element type, usually not applicable for simplest linear triangles and tetrahedrons.

A relatively new pressure stabilization, so called polynomial pressure projection (P3), has been proposed by Dohrmann and Bochev [8]. This very potential technique has been demonstrated its capability to circumvent LBB restriction in incompressible Stokes equation using equal-order shape function for velocity and pressure.

The details of derivation of P3 stabilization can be found in reference [8] and [12]. The implementation of P3 stabilization is adding a stabilization term in weak form Eq.(14) as following

$$\begin{bmatrix} \dfrac{M_{IJ}}{\Delta t} + K_{IJ} & -Q_{IiJ} \\ -G_{IJi} & -V_{IJ} \end{bmatrix} \begin{bmatrix} v_{Ji}^{n+1} \\ p_J^{n+1} \end{bmatrix} = \begin{bmatrix} M_{IJ} \dfrac{v_{Ji}^n}{\Delta t} - C_{IJ}^n v_{Ji}^n - \dfrac{\Delta t}{2} H_{IJ}^n v_{Ji}^n + f_{Ii}^t + f_{Ii}^g \\ 0 \end{bmatrix} \tag{17}$$

where V_{IJ} is the P3 stabilization term. This term V_{IJ} is calculated on each element as below

$$V_{IJ} = \sum_{i=1}^{Ne} \int_{\Omega_e} \frac{\alpha}{\mu} \left(\Phi_I \Phi_J - h_I h_J \right) \mathrm{d}\Omega \tag{18}$$

where h_I is the one order lower polynomial interpolation than shape function Φ_I, α is the parameter of P3 stabilization. In the 3-node triangular element which is used in this paper, the pressure projection for an element is just the average pressure as below

$$\breve{p} = \frac{1}{3} \left(\tilde{p}_1 + \tilde{p}_2 + \tilde{p}_3 \right) \tag{19}$$

where $\breve{p}$ is the pressure projection, $\tilde{p}_i$ is the pressure of i-th node.

Correspondingly, the P3 stabilization term V_{IJ} for 3-node triangular element is given as below

$$V_{IJ} = \frac{\Delta\alpha}{12\mu}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} - \frac{\Delta\alpha}{9\mu}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{\Delta\alpha}{36\mu}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \tag{20}$$

Now, back to Eq.(17), its left side system matrix is positive definite leading to solution for any interpolation of velocity and pressure fields. Since the projected pressure is *k*-1 order lower than velocity of *k* order, the mixed patch condition is also satisfied.

Edge-based smoothed finite element method with CBP3 scheme

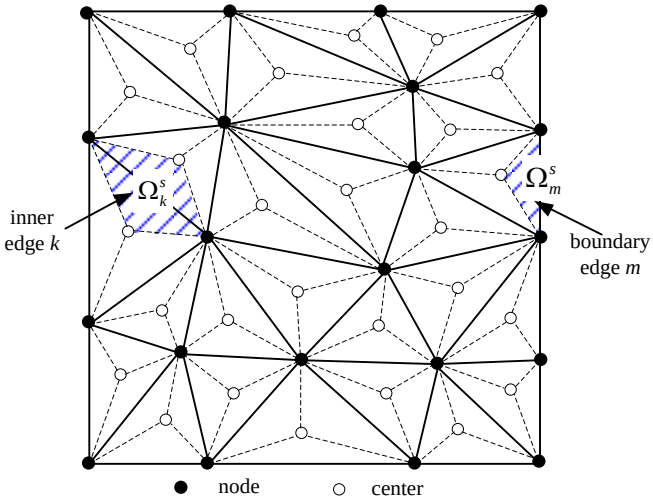


Figure 2. The edge-based gradient smoothing domain for T3 element.

Since ES-FEM for T3 element is highly recommended in solid mechanics, the ES-FEM is first implemented with CBP3 scheme. The gradient smoothing is the fundament of ES-FEM and other S-FEMs. In FEM, the field variables' gradients are calculated by

$$\frac{\partial v_i}{\partial x_j} = \sum_{I=1}^{N_n} \frac{\partial \Phi_I}{\partial x_j} v_{Ii}, \quad \frac{\partial p}{\partial x_j} = \sum_{I=1}^{N_n} \frac{\partial \Phi_I}{\partial x_j} p_I \tag{21}$$

In S-FEM, the gradients of field variables are calculated same with FEM but using smoothed derivatives of shape function

$$\frac{\partial v_i}{\partial x_j} = \sum_{I=1}^{N_n} \overline{\frac{\partial \Phi_I}{\partial x_j}} v_{Ii}, \quad \frac{\partial p}{\partial x_j} = \sum_{I=1}^{N_n} \overline{\frac{\partial \Phi_I}{\partial x_j}} p_I \tag{22}$$

The final equation of the smoothed derivatives of shape function is written as below, more detailed derivation can be found in reference [1]

$$\overline{\frac{\partial \Phi}{\partial x_j}} = \frac{1}{A_i} \sum_{J \in E_i} \Phi(\mathbf{x}_J)\mathbf{n} L_J . \tag{23}$$

where $J \in E_i$ means the *J*-th edge of smoothing domain Ω_i, L_J is the length of *J*-th edge, A_i is the area of smoothing domain.

In this paper, the edge-based smoothing domain is used. Its construction is illustrated in **Figure 2**. The edge-based smoothing domain crosses two connecting T3 element which makes the smoothed derivatives of shape function is more smooth than standard derivatives of shape function in whole system. As a result, the ES-FEM, respects to edge-based smoothing domain, for T3 element show softer behavior than FEM using T3 element.

The smoothed weak form of ES-FEM with CBP3 scheme using T3 element (CBP3/ES-FEM-T3) can be written as

$$\begin{bmatrix} \dfrac{M_{IJ}}{\Delta t}+\bar{K}_{IJ} & \bar{G}_{IJi} \\ \bar{G}_{IJi} & V_{IJ} \end{bmatrix} \begin{bmatrix} v_{Ji}^{n+1} \\ p_J^{n+1} \end{bmatrix} = \begin{bmatrix} M_{IJ}\dfrac{v_{Ji}^n}{\Delta t}-\bar{C}_{IJ}^n v_{Ji}^n-\dfrac{\Delta t}{2}\bar{H}_{IJ}^n v_{Ji}^n+f_{Ii}^t+f_{Ii}^g \\ 0 \end{bmatrix} \tag{24}$$

where $\overline{|\bullet|}$ means smoothed matrices calculated by ES-FEM

$$\begin{aligned}
&[\bar{C}]_{IJ}^n = \sum_{l=1}^{N_{ESD}} \int_{\bar{\Omega}_l} \rho\Phi_I \overline{\frac{\partial\left(v_j^n\Phi_J\right)}{\partial x_j}}\mathrm{d}\Omega, \\
&[\bar{K}]_{IJ} = \sum_{l=1}^{N_{ESD}} \int_{\bar{\Omega}_l} \bar{\mathbf{B}}_I^T \mathbf{D}_\mu \bar{\mathbf{B}}_J \mathrm{d}\Omega, [\bar{H}]_{IJ}^n = \sum_{l=1}^{N_{ESD}} \int_{\bar{\Omega}_l} \overline{\frac{\partial(v_k^n\Phi_I)}{\partial x_k}}\rho\overline{\frac{\partial(v_j^n\Phi_J)}{\partial x_j}}\mathrm{d}\Omega, \\
&[\bar{G}]_{IJi} = \sum_{l=1}^{N_{ESD}} \int_{\bar{\Omega}_l} \rho\Phi_I \overline{\frac{\partial\left(\Phi_J\right)}{\partial x_j}}\mathrm{d}\Omega, \\
&\bar{\mathbf{B}}_I = \begin{bmatrix} \overline{\dfrac{\partial\Phi_I}{\partial x}} & 0 \\ 0 & \overline{\dfrac{\partial\Phi_I}{\partial y}} \\ \overline{\dfrac{\partial\Phi_I}{\partial y}} & \overline{\dfrac{\partial\Phi_I}{\partial x}} \end{bmatrix}.
\end{aligned} \tag{25}$$

where N_{ESD} is the number of edge-based smoothing domains, $\bar{\Omega}_l$ is the l-th smoothing domain. Above terms' quadrature are approximated using one point area integration at the center of smoothing domain.

Numerical example

Taylor-Green vortex

The Taylor-Green vortex problem is a benchmark with following analytical solution [13],

$$\begin{cases} v_x = \sin(x)\cos(y)\exp(-2\dfrac{\mu t}{\rho}), \\ v_y = -\cos(x)\sin(y)\exp(-2\dfrac{\mu t}{\rho}), \\ p = \dfrac{\rho}{4}\left[\cos(2x)+\cos(2y)\right]\exp(-4\dfrac{\mu t}{\rho}). \end{cases} \tag{26}$$

The analytical solution of Taylor-Green vortex problem is a unsteady solution which is desire to verify the CBP3 scheme.

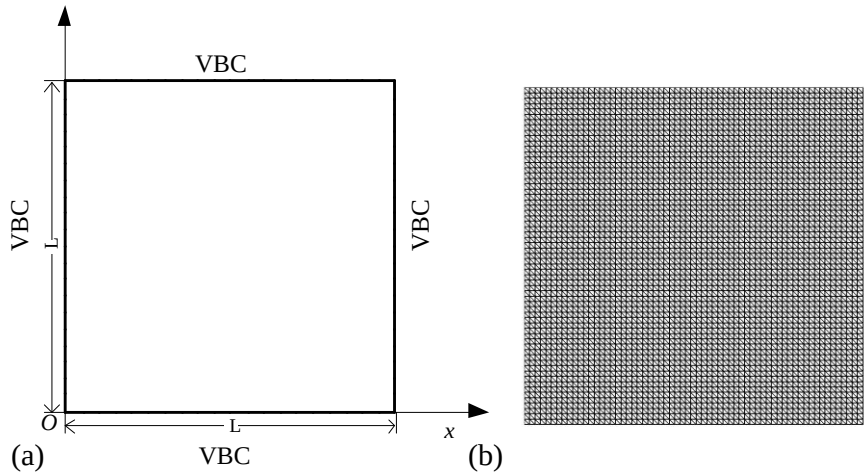


Figure 3. The edge-based gradient smoothing domain for T3 element.

The fluid domain for Taylor-Green vortex is a square with $\mathrm{L}=2\pi$ m , plotted in **Figure 3**(a). A set of uniform meshes, with characteristic element length h=0.4 m, h=0.3 m, h=0.2 m and h=0.1 m, is used to test the convergence of CBP3/FEM-T3 and CBP3/ES-FEM-T3. In (b), the mesh with h=0.1m is plotted. The total computation time is t=3s, the time step length is 0.1s for both methods. The fluid density is is 1 kg/m^3, the fluid dynamic viscosity μ is 0.1 kg/(m·s). The stabilization parameter $\alpha=1$, according to reference [8]. The initial conditions and boundary conditions are prescribed as the analytical solution in Eq.(26).

In **Figure 4**, the v_x contours of Taylor-Green vortex problem calculated by CBP3/FEM-T3 and CBP3/ES-FEM-T3 are presented. Two proposed methods give almost identical v_x contours without visible oscillation. The same situation is happened for the pressure contours , see

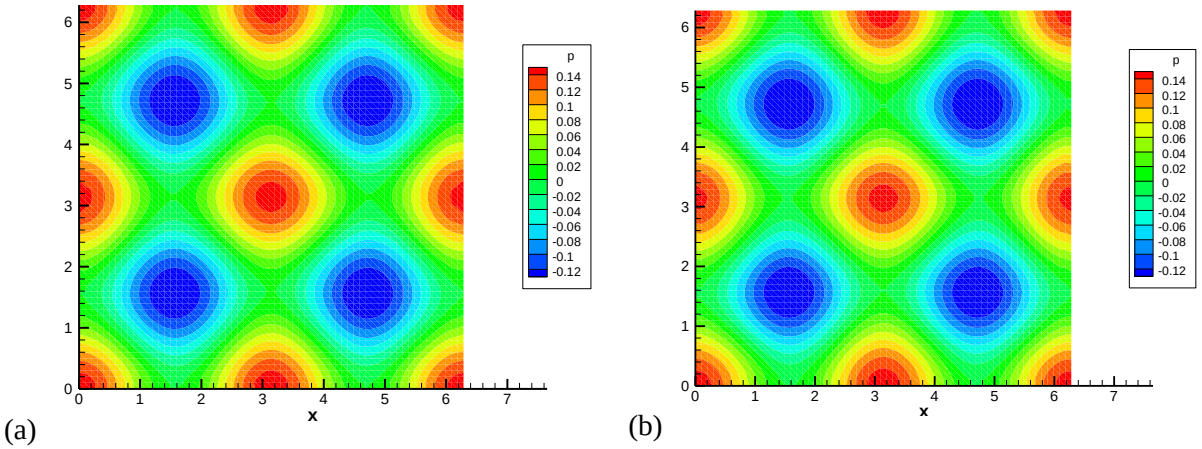
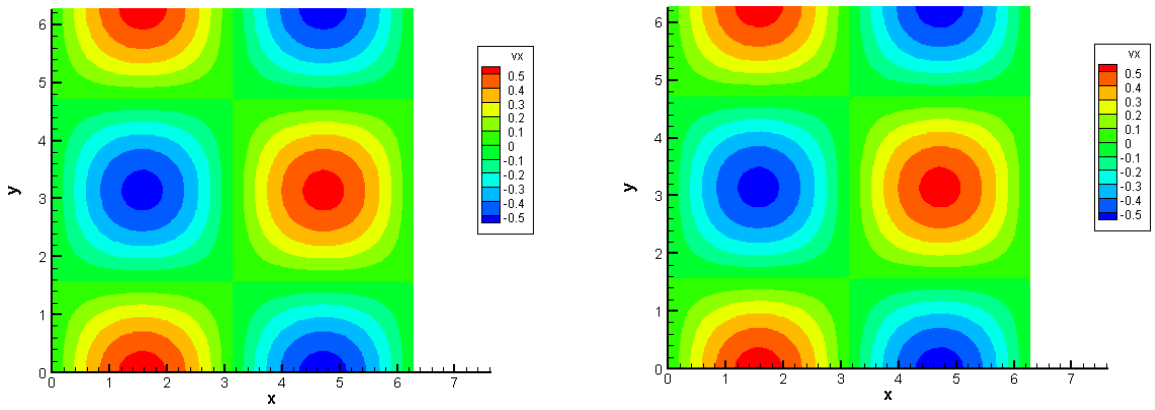
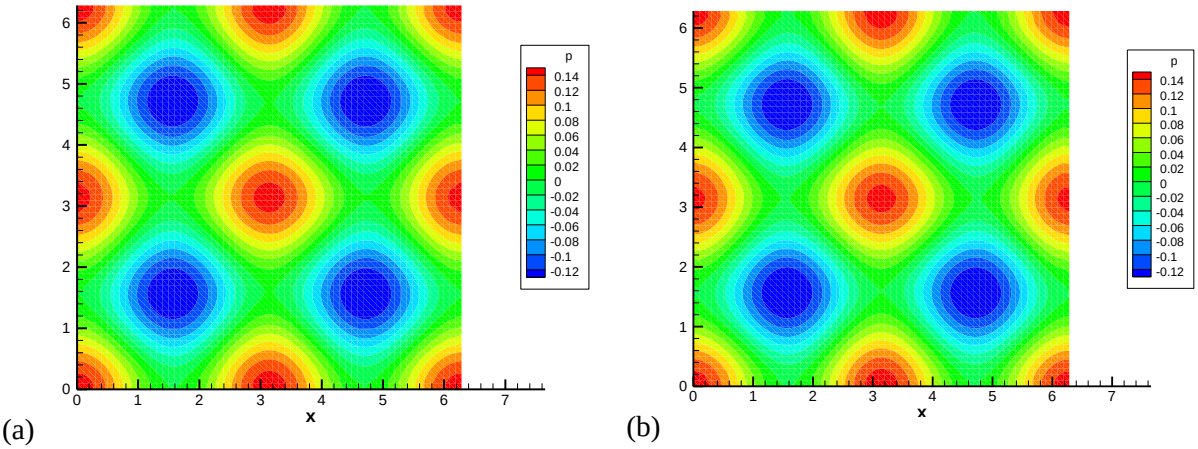


Figure 5.



(a) (b)

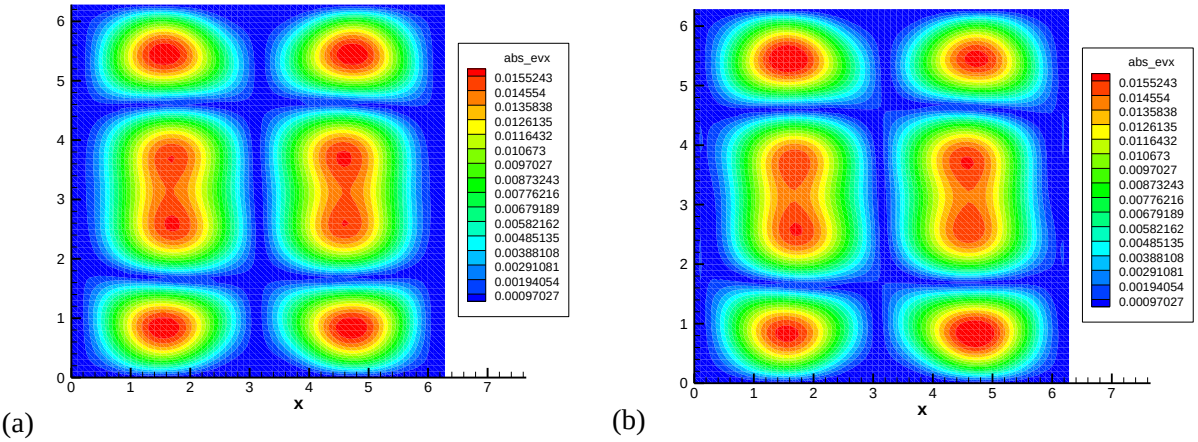
Figure 4. The contours of v_x for CBP3/FEM-T3 (a) and CBP3/ES-FEM-T3 (b).



(a) (b)

Figure 5. The contours of *p* for CBP3/FEM-T3 (a) and CBP3/ES-FEM-T3 (b).

As a further investigation, the absolute errors of v_x ($\left|v_x - v_x^{analytical}\right|$) on all node are calculated and plotted as contours in



(a) (b)

Figure 6. CBP3/ES-FEM-T3 exhibits slightly better results at areas near the line y=3m. Likewise, the absolute nodal pressure errors are also plotted as contours in

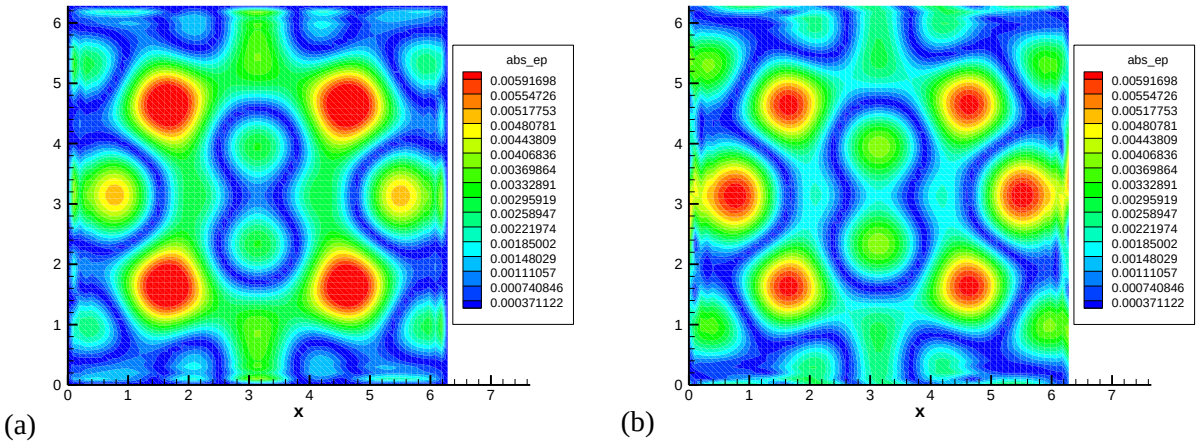


Figure 7. The errors distributed in problem domain share resemblance for CBP3/FEM-T3 and CBP3/ES-FEM-T3. But it is obvious that CBP3/ES-FEM-T3 has a better pressure accuracy.

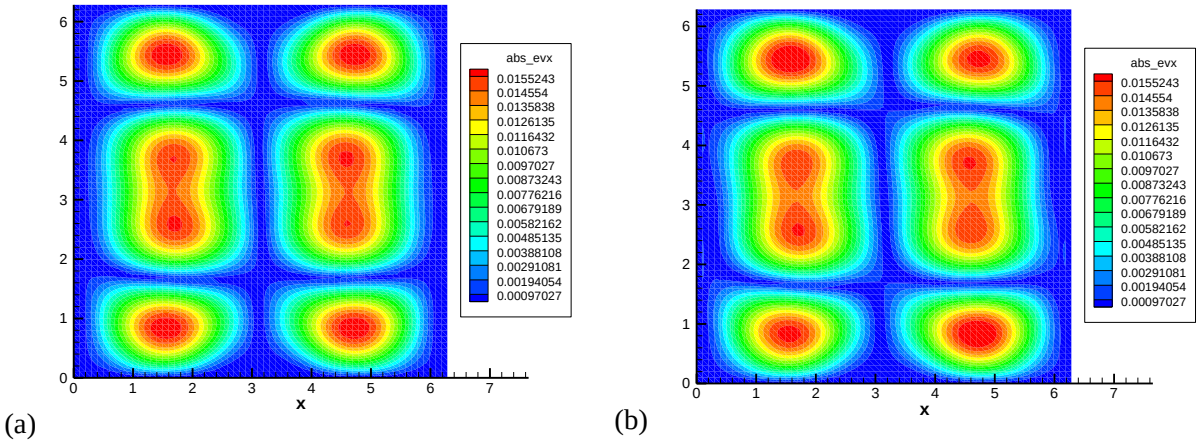


Figure 6. The contours of absolute error of v_x for CBP3/FEM-T3 (a) and CBP3/ES-FEM-T3 (b).

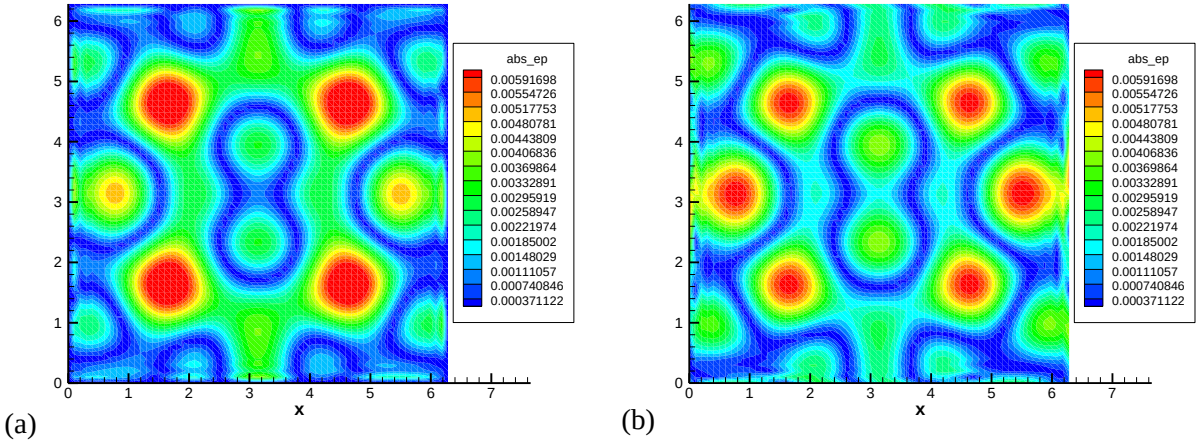
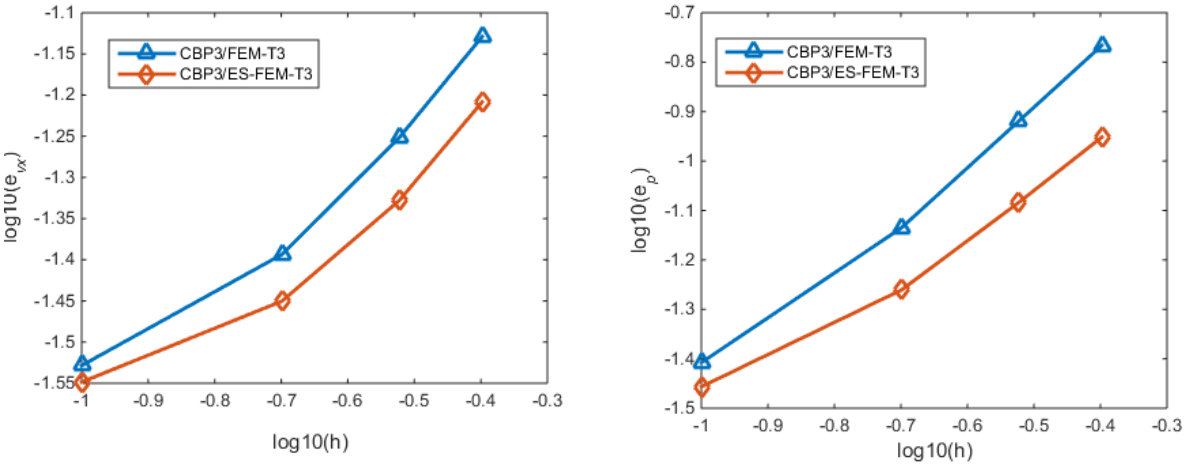


Figure 7. The contours of absolute error of p for CBP3/FEM-T3 (a) and CBP3/ES-FEM-T3 (b).

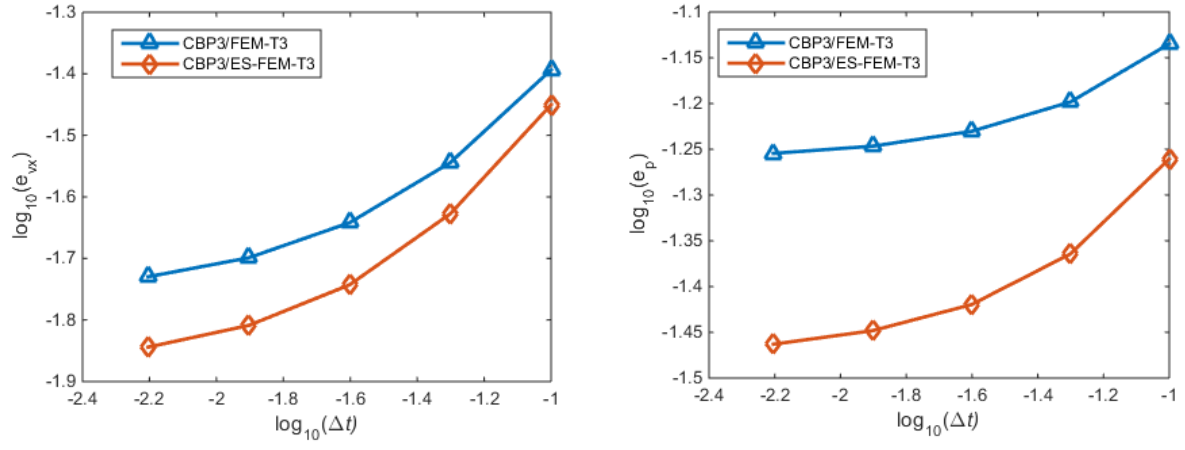


(a) (b)

Figure 8. The spatial convergence of v_x (a) and p (b) for CBP3/FEM-T3 and CBP3/ES-FEM-T3 ($\Delta t = 0.1s$, $\alpha = 1$).

The spatial convergence studies of CBP3/FEM-T3 and CBP3/ES-FEM-T3 are conducted on v_x and p. The L2 norms of errors of v_x and p are used as accuracy indicator. The convergence curves are obtained on meshes with characteristic element length h=0.4 m, h=0.3 m, h=0.2 m and h=0.1 m. In **Figure 8**, convergence curves of v_x and p are drawn. With the contribution of edge-based gradient smoothing, CBP3/ES-FEM-T3 has both better accuracies on velocity and pressure. It is consistent with the absolute errors distributions of CBP3/FEM-T3 and CBP3/ES-FEM-T3 in **Figure 6** and **Figure 7**.

The temporal convergence is also studied. Here, four time steps, $\Delta t = 0.1$s, $\Delta t = 0.05$s, $\Delta t = 0.025$s, $\Delta t = 0.0125$s and $\Delta t = 0.00625$s, are selected for CBP3/FEM-T3 and CBP3/ES-FEM-T3 using mesh with 1024 nodes and h=0.2m. The calculated L2 norms of errors of v_x and p are plotted in **Figure 9**. In all different time step circumstances, CBP3/ES-FEM-T3 is superior than CBP3/FEM-T3.



(a) (b)

Figure 9. The temporal convergence of v_x (a) and p (b) for CBP3/FEM-T3 and CBP3/ES-FEM-T3 (1024 nodes, $\alpha = 1$).

Conclusions

A new CBP3 stabilization scheme is developed in this paper. This CBP3 has also been implement into FEM and ES-FEM using T3 element. The convection oscillations of FEM and S-FEM are reduced by characteristic-based method. The pressure calculation in CBP3 for incompressible flow is no longer by solving the pressure Poisson equation of previous CBS algorithm. The CBP3 calculates pressure by polynomial pressure projection method whose equation only relates to shape function. The numerical example, Taylor-Green vortex, are employed here as a verification. Despite that the Reynolds number is only 10, the numerical example still demonstrated proposed CBP3 scheme is able to help FEM-T3 to solve incompressible laminar flows. Besides, with this proper stabilization, ES-FEM-T3 is proved its capability for computational fluid dynamics. Meanwhile, the edge-based gradient smoothing can boost the ES-FEM-T3 with better accuracy than FEM-T3. From this very primitive study, ES-FEM-T3 can be concluded as a better choice for incompressible flow simulation using unstructured T3 mesh.

Acknowledgements

The corresponding author also appreciates the support of the National Natural Science Foundation of China (No. 11232004) and Specialized Research Fund for the Doctoral Program of Higher Education, China (No. 20130161130001).

References

[1] Liu GR., Trung NT. (2010). Smoothed Finite Element Methods [Internet]. Boca Raton: CRC Press;

[2] Liu GR., Nguyen-Thoi T., Lam KY. (2009). An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analyses of solids. *J Sound Vib*. **320**(4–5):1100–30.

[3] He ZC., Li GY., Zhong ZH., Cheng a. G., Zhang GY., Liu GR., Li E., Zhou Z. (2012). An edge-based smoothed tetrahedron finite element method (ES-T-FEM) for 3D static and dynamic problems. *Comput Mech*. **52**(1):221–36.

[4] Chen H., Wang Q., Liu GR., Wang Y., Sun J. (2016). Simulation of thermoelastic crack problems using singular edge-based smoothed finite element method. *Int J Mech Sci*. **115**–**116**:123–34.

[5] Nguyen-Xuan H., Liu GRR. (2013). An edge-based smoothed finite element method softened with a bubble function (bES-FEM) for solid mechanics problems. *Comput Struct*. **128**:14–30.

[6] Zienkiewicz OC., Taylor RL., Nithiarasu P. (2014). The Finite Element Method for Fluid Dynamics [Internet]. Finite Elem. Method Fluid Dyn. Elsevier;

[7] Zienkiewicz OC., Codina R. (1995). A general algorithm for compressible and incompressible flow—Part I. the split, characteristic-based scheme. *Int J Numer Methods Fluids*. **20**:869–85.

[8] Dohrmann CR., Bochev PB. (2004). A stabilized finite element method for the Stokes problem based on polynomial pressure projections. *Int J Numer Methods Fluids*. **46**(2):183–201.

[9] Zienkiewicz OC., Taylor RL., Nithiarasu P. (2014). 03. The Characteristic-Based Split (CBS) Algorithm: A General Procedure for Compressible and Incompressible Flow [Internet]. Finite Elem. Method Fluid Dyn. Elsevier;

[10] Zienkiewicz, O.C., Loehner, R., Morgan, K., Nakazawa S. (1984). FINITE ELEMENTS IN FLUID MECHANICS: A DECADE OF PROGRESS. *Finite Elem Fluids*. **5**:1–26.

[11] F.Brezzi., Fortin M. (1991). Mixed and Hybrid Finite Element Methods. New York: Springer-Verlag;

[12] Zienkiewicz O., Taylor R., Zhu JZ. (2013). The Finite Element Method: Its Basis and Fundamentals. 7th ed. Oxford: Butterworth-Heinemann;

[13] Taylor, G. I. and Green AE. (1937). Mechanism of the Production of Small Eddies from Large OnesTitle. *Proc R Soc LondonSeries A,Mathematics Phys Sci*. **158**(895):499–521.

Numerical methods for structural dynamic responses based on radial basis functions approximation

***Y.T. Li[1,2], †J.Q. Xu[1,2], and X.B. Xu[1,2]**

[1]School of River & Ocean Engineering, Chongqing Jiaotong University, Chongqing 400074, China；
[2]National Engineering Technology Research Center for Inland Waterway Regulation,
Key Laboratory of Hydraulic & Waterway Engineering of the Ministry of Education,
Chongqing Jiaotong University, Chongqing 400074, China

*Presenting author: lijoan@outlook.com.com
†Corresponding author: jqxu11@sina.com

Abstract

A numerical method for solving structural dynamic response was proposed by combining the theory of radial basis functions (RBFs) approximation and the collocation point methods. To solve the problem that using basic RBFs point interpolation method will bring great numerical oscillations, a multivariate interpolation function with the linear combination of each order differential terms was developed and the arithmetic steps were given. Unlike other numerical methods, there were no theoretical supposes about changing rules of acceleration and load within time interval, so this method had an applicability to solve jerk and jerk (third-order) equations. Actual examples showed that RBFs approximation method had simple computational process and improved the convergency and stability effectively.

Keywords: Radial Basis Functions; Meshless Methods; Dynamic Response; Jerk; Jerk Equations; Initial Value Problems

Introduction

Problems of dynamic response of systems can often be come down to initial problems of second-order ordinary differential equations. At present the methods for structural dynamic response mainly include: mode superposition methods, direct integration methods and methods turning second-order into first [1], et al. Mode superposition methods are only used for integration methods of linear systems adopted only first several-order modes, so it is inapplicable to systems which the effects of high modes can't be neglected. Direct integration methods are appropriate for both linear systems and nonlinear systems, the most commonly used includes difference methods, linear acceleration methods and improved linear acceleration methods. Above Methods are all based on two following assumptions: (1) Continues time scale is divided into finite number of nodes where the motion differential equation is just satisfied and solutions of displacement, velocity and acceleration just are obtained; (2) There are some simple assumptions about change rules of acceleration or loads in time interval.

Theoretical defects of these methods made a lower precision with only first or second-order, and the calculating precision will poorer if actual acceleration belied these assumption in transient response phase. Furthermore, the uncontrollability of inherent algorithmic damping of direct integration method also causes great calculation errors. Because applicability of integrate methods depends on type of nonlinearity and load characteristics, etc, it can be hard to choose an appropriate differential scheme while solving a nonlinear problem. Precise integration methods [2] open up a new direction for solving dynamic responses, but when solving homogeneous equations under random loads, such as seismic waves and wind loads, it is necessary to make some assumption about change rules of loads in time interval.

Moreover, Jerk [3], the time rate-of-change of acceleration, has been increasingly applied in areas of chaos theory [4][5], nonlinear dynamics [6][7], mechanical design [8], and structural damage detection [9], etc. And jerk equation, third-order differential equation, of the form involving the third temporal derivative of displacement can describe some physical problems such as third-order mechanical oscillations [10][11]. Above methods are no longer able to be

used to solve jerk and jerk equation because of their inherent basic assumption. At present a few effective numerical methods can be used, e.g., fourth-order accurate Runge-Kutta method with sufficient-small step-sizes [12]. With deeper research about the role of jerk in mechanics and applications, jerk calculating will have more important implications.

Radial basis functions (RBFs) have advantages of simple form, isotropic and independent of space dimensions, etc. scholars, at home and aboard, have proposed a sea of methods based on radial basis functions which have been widely applied in scientific and engineering calculating areas of hydrodynamics, computational mechanics, picture processing, etc. Methless methods [13][14] based on RBFs have been used to solve boundary value problems, and it has acquired a great of achievements, but RBFs have not yet been used to solve initial problems up to now and we will try to do it.

1 Radial Basis Function Approximation Methods

Radial Basis Function (RBF) is a kind of basis function with a distance variable. It uses the simple function φ defined in $[0, +\infty)$ and Euclidean norm $\|\cdot\|_2$ in R^d to represent d-dimensional function $\varphi=\varphi(R_i)$, in which $R_i=\|x-x_i\|_2$, the distance from arbitrary point x to the point x_i, is independent variable. In essence, φ is a one-variable function —— function of distance, thus its simple form makes data convenient to store and calculate. Another important advantage of RBF is the powerful capacity of approximation that it can almost approximate all functions.

Depending on its scoped, RBFs can be divided into two categories: Globally supported RBFs (GS-RBFs) and compactly supported RBFs(CS-RBFs) [15][16][17]. It limits the former's use for calculating large structure that the calculation process will produce ill-conditioned matrixes. CS-RBFs can make the coefficient matrix has the characteristic of banded sparse.

The method combined RBFs approximation with collocation point methods has many advantages, such as meshless, simple form, no numerical integration and high calculation efficiency. But at present both domestic and overseas researches about solving fractional differential equations using RBFs are all related to boundary problems without time parameter. The main reason lies in the independent variable of radial basis function is spatial distance. Xu [18] have presented the concept of transformation from spatial distance to time interval, then we will try to solve initial problems by using RBFs.

First, we take a single-degree-freedom dynamic system (1) as an example to explain the numerical method.

$$\begin{cases} m\ddot{u}+c\dot{u}+ku=p(t) \\ u(0)=u_0, \dot{u}(0)=\dot{u}_0 \end{cases} \tag{1}$$

The time domain Ω can be discretisized with n nodes t_i, $i = 1, 2, \cdots, n$, then the approximate function $u^h(t)$ of displacement function $u(t)$ can use a linear combination of radial basis function $\varphi_i(t)$ which is taking t_i as the center to expresseitself as:

$$u^h(t)=\sum_{j=1}^{n} a_j\varphi_j(t)=\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{a} \tag{2}$$

Eq. (2) is a basic expression of RBFs interpolation, a_j denotes a series of unsolved coefficients, $\boldsymbol{a}=[a_1, a_2, \cdots, a_n]^{\mathrm{T}}$, $\boldsymbol{\Phi}(t)=[\Phi_1(t), \Phi_2(t), \cdots, \Phi_n(t)]^{\mathrm{T}}$. Put CS-RBF as interpolation cardinal function, in this paper, we use Wu function [15]:

$$\varphi(r)=(1-r)_+^5\left(1+5r+9r^2+5r^3+r^4\right) \tag{3}$$

In Eq. (3), For this method, we have $r=\|t-t_i\|_2 / R_{\max i}$. And $R_{\max i}$ is support radius of t_i , the maximum distance of t_i to any other points, which means that the effective region of t_i is the whole domain. $(1-r)_+$ can be defined as $(1\text{-}r)_+ = \begin{cases} 1-r, 0 \le r \le 1 \\ 0 \quad , other \end{cases}$.

Because traditional collocation methods have large numerical oscillation, for the characteristic that the objective solution of system (1) is the second derivative of u, thus we present the interpolation function combined displacement with velocity, as show in Eq. (4):

$$u^h(t) = \sum_{j=1}^{n} a_j \varphi_j(t) + b_1 \frac{\mathrm{d}\varphi_1(t)}{\mathrm{d}t} \tag{4}$$

In Eq. (4), there add a linear combination of first derivative of the initial time. According to the authors' solving experience [19], if we add a second derivative term on the basis of Eq. (4), then get the Eq. (5), the numerical oscillation can be diminished significantly. And adding the initial condition of the second derivative term as a new constrain, this initial constrain has explicit physical and mathematical interpretations which are the acceleration of initial time and that the second derivative satisfies the differential equation in the initial time.

$$u^h(t) = \sum_{j=1}^{n} a_j \varphi_j(t) + b_1 \frac{\mathrm{d}\varphi_1(t)}{\mathrm{d}t} + b_2 \frac{\mathrm{d}^2\varphi_1(t)}{\mathrm{d}t^2} \tag{5}$$

We can find Eq. (5) requires the RBF has high-order continuity, and taking high-order derivative of CS-RBFs will cause an ill-conditioned coefficient matrix. So we finally presented substituting helper function for high-order derivative term, as follows:

$$u^h(t) = \sum_{j=1}^{n} a_j \varphi_j(t) + b_1 \frac{\mathrm{d}\varphi_1(t)}{\mathrm{d}t} + b_2 \xi_1(t) = \boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{a} \tag{6}$$

In Eq. (6), b_1, b_2 denotes additional coefficient, $\xi_1(t)$ is helper function, and we can use other CS-RBF as an available helper function, for example,

$$\xi(t) = (1-r)_+^6 \left(6 + 36r + 82r^2 + 72r^3 + 30r^4 + 5r^5\right).$$

Then plugging t_i (i = 1, 2, ⋯, n, 1, 1) into the interpolation Eq. (6), we can obtain n+2 linear equations Eq. (7):

$$\boldsymbol{A}\boldsymbol{a} = \boldsymbol{u} \tag{7}$$

In Eq. (7),

$$\boldsymbol{A} = \begin{bmatrix} \varphi_1(t_1) & \varphi_2(t_1) & \cdots & \varphi_n(t_1) & \dfrac{\mathrm{d}\varphi_1(t_1)}{\mathrm{d}t} & \xi_1(t_1) \\ \varphi_1(t_2) & \varphi_2(t_2) & \cdots & \varphi_n(t_2) & \dfrac{\mathrm{d}\varphi_1(t_2)}{\mathrm{d}t} & \xi_1(t_2) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \varphi_1(t_n) & \varphi_2(t_n) & \cdots & \varphi_n(t_n) & \dfrac{\mathrm{d}\varphi_1(t_n)}{\mathrm{d}t} & \xi_1(t_n) \\ \dfrac{\mathrm{d}\varphi_1(t_1)}{\mathrm{d}t} & \dfrac{\mathrm{d}\varphi_2(t_1)}{\mathrm{d}t} & \cdots & \dfrac{\mathrm{d}\varphi_n(t_1)}{\mathrm{d}t} & \dfrac{\mathrm{d}^2\varphi_1(t_1)}{\mathrm{d}t^2} & \dfrac{\mathrm{d}\xi_1(t_1)}{\mathrm{d}t} \\ \dfrac{\mathrm{d}^2\varphi_1(t_1)}{\mathrm{d}t^2} & \dfrac{\mathrm{d}^2\varphi_1(t_1)}{\mathrm{d}t^2} & \cdots & \dfrac{\mathrm{d}^2\varphi_n(t_1)}{\mathrm{d}t^2} & \dfrac{\mathrm{d}^3\varphi_1(t_1)}{\mathrm{d}t^3} & \dfrac{\mathrm{d}^2\xi_1(t_1)}{\mathrm{d}t^2} \end{bmatrix},$$

$\boldsymbol{u}=[u_1, u_2, \cdots, u_n, v_1, v_2]^{\mathrm{T}}$ where v_1 、 v_2 denote additional unknowns which respectively represent the initial velocity and acceleration.

It is easy, from Eq. (7), to show that $\boldsymbol{a}=\boldsymbol{A}^{-1}\boldsymbol{u}$ and using $\boldsymbol{a}$ in Eq. (6) gives

$$u^h(t) = \boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{A}^{-1}\boldsymbol{u} = \boldsymbol{N}(t)\boldsymbol{u}. \tag{8}$$

Let $\boldsymbol{N}(t)=\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{A}^{-1}$. We could definite $\boldsymbol{N}(t)$ is time characteristic function, similar to finite element shape function, and $\boldsymbol{A}$ is dynamical characteristic matrix. Eq. (8) is an analytic equation, then differentiating Eq. (8) with respect to t gives the expression of velocity Eq. (9),

$$\dot{\boldsymbol{u}}(t)=\boldsymbol{N}'(t)\boldsymbol{u} \tag{9}$$

In Eq. (9),

$$\boldsymbol{N}'(t)=\begin{bmatrix}\varphi_1'(t_1) & \varphi_2'(t_1) & \cdots & \varphi_n'(t_1) & \varphi_1''(t_1) & \xi_1'(t_1)\\ \varphi_1'(t_2) & \varphi_2'(t_2) & \cdots & \varphi_n'(t_2) & \varphi_1''(t_2) & \xi_1'(t_2)\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots\\ \varphi_1'(t_n) & \varphi_2'(t_n) & \cdots & \varphi_n'(t_n) & \varphi_1''(t_n) & \xi_1'(t_n)\end{bmatrix}\boldsymbol{A}^{-1},\quad \varphi'(t)=\frac{\mathrm{d}\varphi}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}t},\quad \frac{\mathrm{d}r}{\mathrm{d}t}=\frac{t_i-t}{R_iR_{\max i}}.$$

Similarly, the expression of acceleration Eq. (10) as follows:

$$\ddot{\boldsymbol{u}}(t)=\boldsymbol{N}''(t)\boldsymbol{u}. \tag{10}$$

By substituting Eq. (8), (9) and (10) into dynamic system (1) and combining the external loads vector $\boldsymbol{p}$, we obtain $m\boldsymbol{N}''\boldsymbol{u}+c\boldsymbol{N}'\boldsymbol{u}+k\boldsymbol{u}=\boldsymbol{p}(t)$, it follows that

$$\left(m\boldsymbol{N}''+c\boldsymbol{N}'+k\boldsymbol{E}\right)\boldsymbol{u}=\boldsymbol{p}(t). \tag{11}$$

And according to the initial conditions of velocity and adding initial acceleration, we have

$$\left\{\varphi_1'(t_1)\quad \varphi_2'(t_1)\quad \cdots\quad \varphi_n'(t_1)\quad \varphi_1''(t_1)\quad \xi_1'(t_1)\right\}\boldsymbol{A}^{-1}\boldsymbol{u}=\dot{u}(0) \tag{12}$$

$$\left\{\varphi_1''(t_1)\quad \varphi_2''(t_1)\quad \cdots\quad \varphi_n''(t_1)\quad \varphi_1''(t_1)\quad \xi_1''(t_1)\right\}\boldsymbol{A}^{-1}\boldsymbol{u}=\ddot{u}(0) \tag{13}$$

Combining Eq. (11), (12) and (13), we get n+2 linear equations, thus the second-order differential equation is discretized into linear algebraic equations. Substituting the initial displacement conditions and solving the equations we obtain the solution of $\boldsymbol{u}$. Then plugging $\boldsymbol{u}$ back into Eq. (9) and (10) respectively gives the velocity and acceleration at every time note.

As demonstrated above, the radial basis function approximation method has no theoretical assumption except the interpolation, and $u^h(t)$ is an analytic expression, so taking the third derivative of Eq. (8) can solve the jerk effectively, we will elaborate on this in *Example 3.3*.

Next, for jerk equations of third-order dynamic system with the form of $\dddot{x}=J(x,\dot{x},\ddot{x})$, we can present the interpolation function Eq. (14)

$$u^h(t)=\sum_{j=1}^{n}a_j\varphi_j(t)+b_1\frac{\mathrm{d}\varphi_1(t)}{\mathrm{d}t}+b_2\xi_1(t)+b_3\frac{\mathrm{d}\xi_1(t)}{\mathrm{d}t}=\boldsymbol{\Phi}^{\mathrm{T}}(t)\boldsymbol{a}. \tag{14}$$

Using the similar process, the jerk equation is discretized into n+3 nonlinear algebraic equations

$$N'''x-J\left(x,N'x,N''x\right)=0,$$

where $\boldsymbol{N}$ denotes time characteristic matrix. And using initial displacement, velocity, acceleration and jerk as the constrains, by replacing functions at initial time with corresponding constrains, then using iterative method to calculate the nonlinear algebraic equations, we can obtain the solutions of the third-order equation. We will give a numerical example in *Example 2.4*.

2 Numerical Examples Analyses

2.1 Forced vibrations of a single degree of freedom

$$\begin{cases}\ddot{u}+4\dot{u}+5u=\sin(2t)\\ u(0)=\dfrac{57}{65},\dot{u}(0)=\dfrac{2}{65}\end{cases}. \tag{15}$$

Analytic solution of system (15) is $u = e^{-2t}(\cos t+2\sin t)-[8\cos(2t)-\sin(2t)]/65$. If time domain is t = 60.8 s， time interval is Δt = 0.1 s, using RBFs approximation methods, choosing Eq. (4) as interpolation function and using $u(0) = 57/65$ and $\dot{u}(0) = 2/65$ as initial conditions, then choosing Eq. (6) and adding a second-order differential initial condition $\ddot{u}(0) = -293/65$ to solve system (15), calculating the relative errors of solutions in first six notes as shown in Table 1, we are confident that the latter can diminish numerical oscillations obviously improve solving accuracy.

Using other traditional methods to solve system (15) with also Δt = 0.1 s, and relative errors at some moments are displayed in Table 2. It is seen that, RBFs approximation methods, as compared to Newmark method (NM) and Wilson-θ method (W-θ), improve the solving accuracy greatly and it has high stability of accuracy for long periods.

Table 1. The relative errors for RBFs approximation methods with different interpolation function and initial conditions

t(s)	Interpolation function (6) with $\ddot{u}(0)$			Interpolation function (4) without $\ddot{u}(0)$		
	displacement	velocity	acceleration	displacement	velocity	acceleration
0	0	2.59×10^{-15}	1.82×10^{-15}	0	3.94×10^{-4}	2.39×10^{-1}
0.1	1.46×10^{-4}	6.20×10^{-3}	3.16×10^{-3}	3.30×10^{-3}	1.08×10^{-1}	5.62×10^{-2}
0.2	3.09×10^{-4}	1.24×10^{-3}	2.58×10^{-3}	6.59×10^{-3}	3.19×10^{-2}	6.30×10^{-2}
0.3	4.14×10^{-4}	7.92×10^{-4}	5.75×10^{-3}	8.98×10^{-3}	1.62×10^{-2}	0.12×10^{-1}
0.4	5.00×10^{-4}	2.23×10^{-4}	1.12×10^{-1}	1.08×10^{-2}	5.36×10^{-3}	2.50
0.5	5.63×10^{-4}	4.27×10^{-4}	4.71×10^{-3}	1.22×10^{-2}	8.42×10^{-3}	9.85×10^{-2}

Table 2. The relative error of the solution of displacement, velocity and acceleration with several numerical methods

t(s)	displacement			velocity			acceleration		
	NM	W-θ	RBF	NM	W-θ	RBF	NM	W-θ	RBF
2	0.0051	0.0150	3.70×10^{-4}	0.0048	0.0192	2.94×10^{-4}	0.0444	0.2328	2.34×10^{-3}
4	0.0166	0.0189	4.40×10^{-6}	0.0016	0.0132	4.18×10^{-9}	0.0086	0.0187	5.30×10^{-6}
16	0.0003	0.0193	1.40×10^{-6}	0.0039	0.0004	2.20×10^{-6}	0.0070	0.0227	1.96×10^{-6}
34	0.0085	0.0029	8.33×10^{-6}	0.0029	0.0174	3.30×10^{-6}	0.0019	0.0004	4.00×10^{-7}
40	0.272	0.6670	5.21×10^{-5}	0.0005	0.0114	9.70×10^{-8}	0.2666	0.6656	5.10×10^{-7}
60	0.0005	0.0199	1.58×10^{-5}	0.0036	0.0012	1.27×10^{-5}	0.0072	0.0233	3.50×10^{-6}

2.2 Bending Vibration of Simply Supported Beam

A simply supported beam with constant section is illustrated in Fig. 1, and with length L=6 m, high of cross section h=0.02m, width b=0.02m, cross sectional area $A=bh$, section inertia $I=bh^3/12$, density $\rho=4\times10^4$ kg/m^3, elastic modulus E=210 GPa and poisson ratio μ=0.3. Suppose the beam is damping-free, and there is a lateral load $q(x,t)=F_0\sin(\omega_0 t)\,\delta(x-L/2)$. The theoretical solution of vibration displacement of this beam is

$$w(x,t)=\frac{2F_0}{\rho AL}\sum_{r=1,3,5,\cdots}^{\infty}\frac{(-1)^{(r-1)/2}}{\omega_r^2[1-(\omega_0/\omega_r)^2]}\sin\frac{r\pi x}{L}\times\left(\sin(\omega_0 t)-\frac{\omega_0}{\omega_r}\sin(\omega_r t)\right)$$

where $\omega_r=(r/\pi)^2\sqrt{EI/(\rho A)}$ is inherent frequency.

If F_0=1 kN, ω_0=4 Hz, t = 64 s, Δt = 0.2 s, dispersing the beam into ten cubic Hermite finite elements, using RBFs approximation method to solve this problem, some results are shown in Fig. 2. It is seen that, compared to the solutions [20] of newmark method and Wilson-Wilson-θ method, the solving accuracy of this method has improved and the errors do not accrue with time.

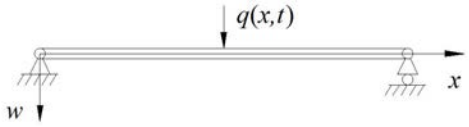
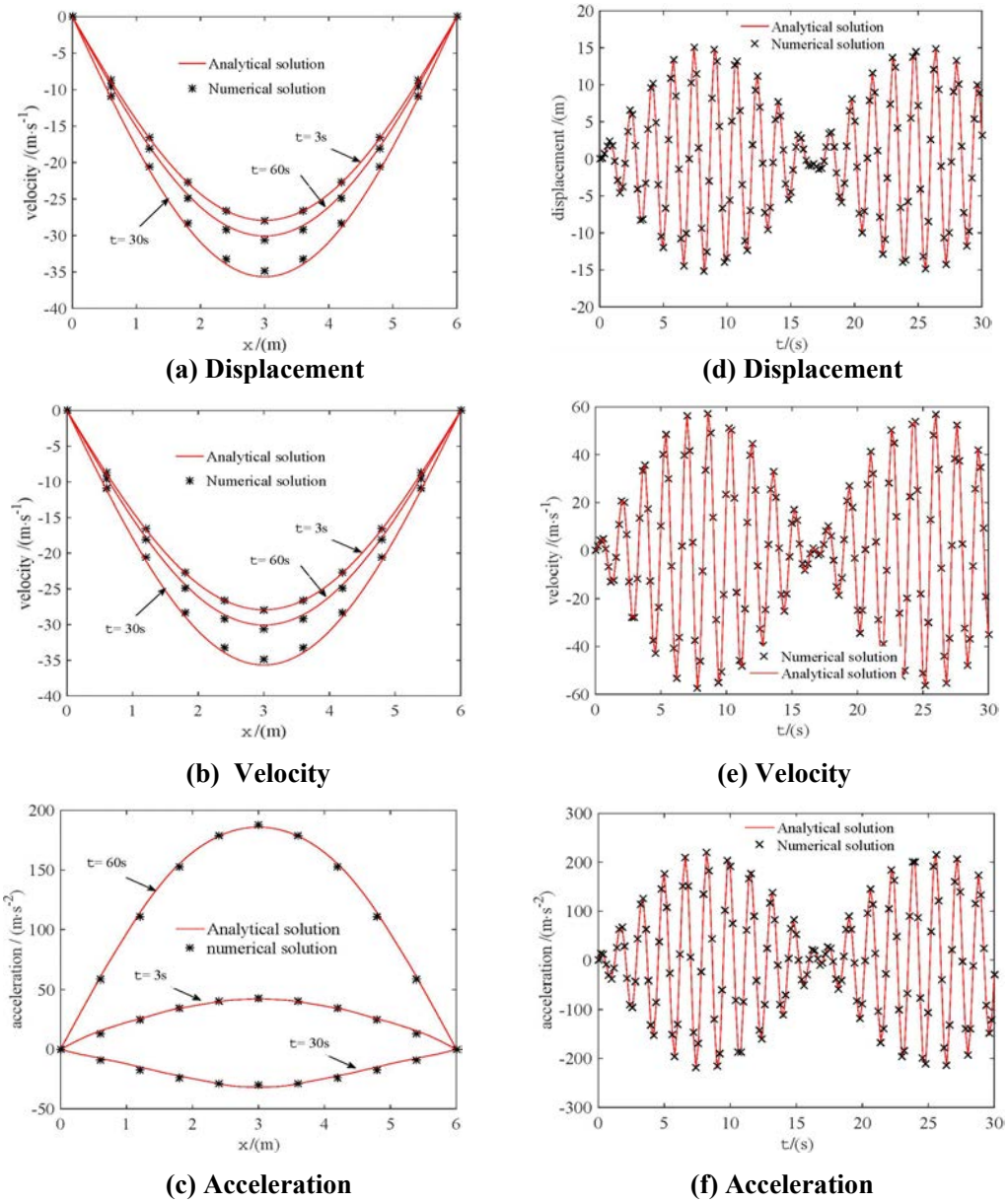


Figure 1. The sketch of simple supported beam bending vibration



(a) Displacement **(d) Displacement**

(b) Velocity **(e) Velocity**

(c) Acceleration **(f) Acceleration**

Figure 2. (a), (b) and (c) show the numerical solution and the exact solution at every node at t=3s,30s, 60s; (d), (e) and (f) show the numerical solution and the exact solution at the middle point in different time.

2.3 Three Stories Frame Structure

Fig. 3 shows a sketch of three-storey shear frame structure, the masses of each storey, including columns, are m_1=1.8×10^5 kg, m_2=2.7×10^5 kg, m_3=2.7×10^5 kg, respectively. Lateral stiffness are k_1=9.8×10^7 N/m, k_2=1.96×10^8 N/m, k_3=2.45×10^8 N/m. We will solve the dynamical responses and the jerk of this structure with Rayleigh damping under horizontal seismic excitation (as shown in Fig. 4), and the 1st and 2nd damping ratio are ξ_1=ξ_2=0.05. The initial conditions are $u(t)=0$, $\dot{u}(t)=0$, and adding $\ddot{u}(0)=0.014$, initial acceleration of ground vibration.

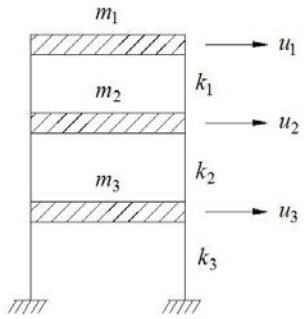


Figure 3. Sketch of three stories shear frame

The mathematical model can be expressed as

$$\boldsymbol{M}\ddot{\boldsymbol{U}}+\boldsymbol{C}\dot{\boldsymbol{U}}+\boldsymbol{K}\boldsymbol{U}=-\boldsymbol{M}\ddot{\boldsymbol{U}}_g, \tag{16}$$

Where $\ddot{\boldsymbol{U}}_g$ and $\ddot{\boldsymbol{U}}$ are respectively horizontal acceleration caused by seismic and relative acceleration, and the former also can be called carrier acceleration. For jerk, let us assume $\ddot{\boldsymbol{U}}_g$ is derivable, we have

$$\boldsymbol{M}\dddot{\boldsymbol{U}}+\boldsymbol{C}\ddot{\boldsymbol{U}}+\boldsymbol{K}\dot{\boldsymbol{U}}=-\boldsymbol{M}\dddot{\boldsymbol{U}}_g, \tag{17}$$

Then

$$\boldsymbol{M}\dddot{\boldsymbol{U}}_{abs}=-\boldsymbol{C}\ddot{\boldsymbol{U}}-\boldsymbol{K}\dot{\boldsymbol{U}}, \tag{18}$$

$\dddot{\boldsymbol{U}}_{abs}=\dddot{\boldsymbol{U}}+\dddot{\boldsymbol{U}}_g$ is absolute jerk of the structure. Using RBFs approximation method, we solved the solutions and some of them are shown in Fig. 5.

It is important to note that the Eq.(18) can't be used in traditional method, such as Newmark method, Wilson-θ method, and so on, because of the theoretical hypothesis. And in this example, although the jerk is also an object solving, we can't use the Eq.(14) because of a lack of initial jerk regarded as a constrain.

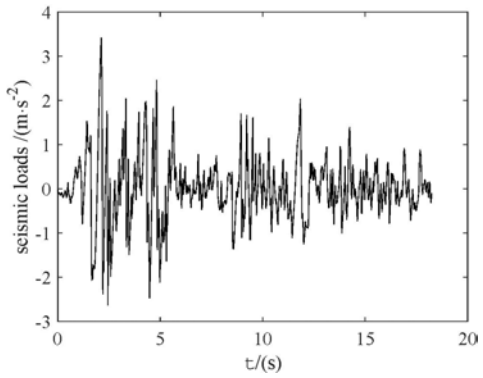
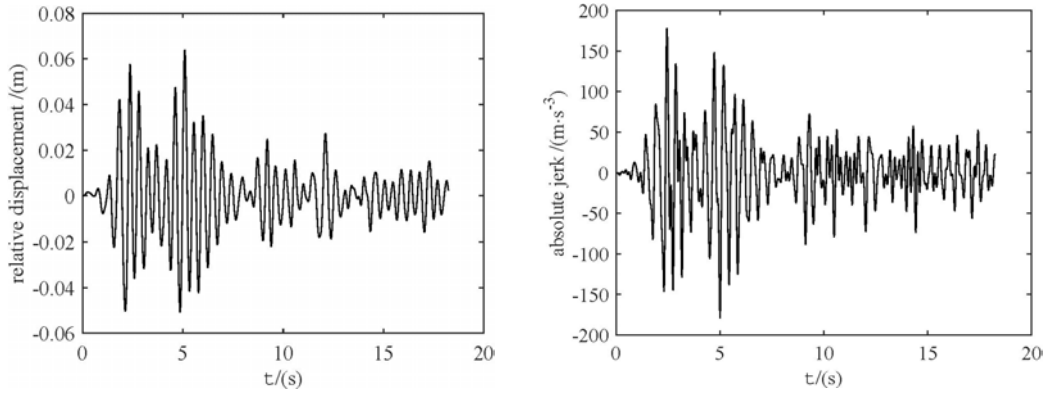


Figure 4. Transport acceleration causing by earthquake



(a) Displacement **(b) Absolute jerk**

Figure 5. the numerical results at the top storey

2.4 Third-order Mechanical Oscillations

Following Gottlieb [21], the most general jerk function which is invariant under time- and displacement- reversals is

$$\dddot{x} = -\gamma\dot{x} - \alpha\dot{x}^3 - \beta x^2\dot{x} + \delta x\dot{x}\ddot{x} - \varepsilon\dot{x}\ddot{x}^2 , \tag{19}$$

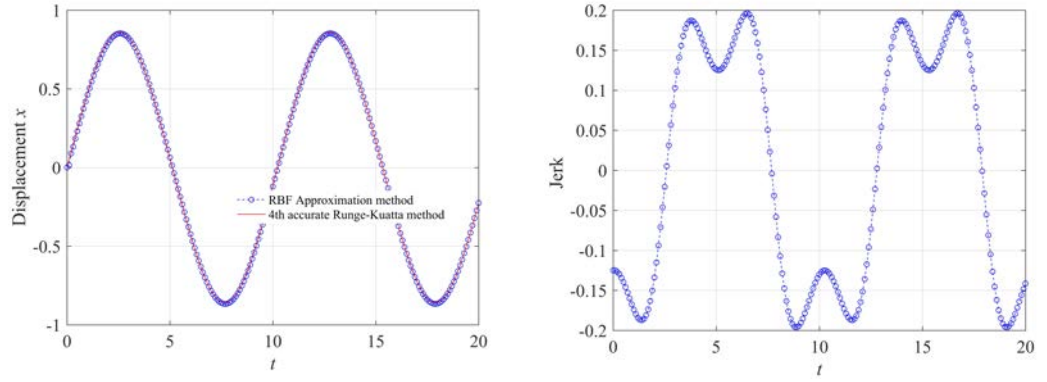
where α, β, γ, δ, and ε are all constants, and at least one of β, δ and ε should be different from zero. In addition, if ε=0, it is required that $\delta \neq -2\alpha$ such that the Eq. (19) is simply not the secon d-order ordinary differential equation. The corresponding initial conditions are

$$x(0) = 0 , \dot{x}(0) = B , \ddot{x}(0) = 0 .$$

In this paper, we consider the case for $\alpha = \beta$=1, $\gamma = \delta = \varepsilon$=0 and B=0.5. For this situation, Eq. (19) becomes

$$\begin{cases} \dddot{x} = -\dot{x}^3 - x^2\dot{x} \\ x(0) = 0, \dot{x}(0) = 0.5, \ddot{x}(0) = 0 \end{cases} . \tag{20}$$

Using RBFs approximation method with Δt=0.1s and choosing Eq. (14) as interpolation function, we get the period of the solution is t=10.210655 s. And the period using fourth-order accurate Runge-Kutta method with Δt=0.001s is 10.210761s [12][21]. The solution of displacement and jerk are shown in Fig. 6.



(a) Displacement **(b) Jerk**

Figure 7. Numerical results of Eq. (20) using RBFs approximation method with Eq. (14)

3 Conclusions

This paper developed a new approach for solving structural dynamic responses and jerk based on the powerful approximation capability of RBFs. The practical calculation examples show that the RBFs approximation method has great astringency and high solving accuracy. Furthermore, this method also has the following advantages:

1. RBFs approximation method is different from stepwise direct integration methods. it needn't numerical integration, has high calculation efficiency, and has no recursive formulae and error accumulation.

2. We proposed the combined interpolation expression of all-order derivatives and it is necessary to add initial condition that the order is same as the differential equation, which can decrease the numerical oscillation significantly.

3. This method has no assumption of load changes and acceleration in time interval, can solve jerk and jerk equation effectively, and it breaks the limitation that stepwise direct integration methods are difficult to solve jerk because of inherent assumptions.

The method of RBFs approximation has clear advantages and it may well become a common method.

References

[1] Yuan X. B., Dong J. F., Qiu C. C., et al. (2007) A new universal algorithm based on time FEM for dynamic response. *Chinese journal of applied mechanics* **27**, 579-583.

[2] Zhong W. X. (1994) On precise time-integration method for structural dynamics. *Journal of Danlian University of Technology* **32**,131-136.

[3] Schot S. H. (1978) Jerk: the time rate of change of acceleration. *America journal of physics* **46**, 1090-1094.

[4] J. C. Sprott. (1997) Some simple chaotic jerk functions. *American Association of Physics Teachers* **65**, 537-543.

[5] Wei Z. C., Zhang W., Yao M. H. (2015) On the periodic orbit bifurcating from one single non-hyperbolic equilibrium in a chaotic jerk systerm. *Nonlinear dynamics* **82**, 1251-1258.

[6] J.I.Ramos, C.M. García-López. (2011) On two classes of autonomous third-order nonlinear ordinary differential equations. *Applied Mathematics and Computation* **218**, 2011-2024.

[7] Linz S J. (1997) Nonlinear dynamical models and jerk motion. *American journal of physics* **65**, 523-526.

[8] Lan S, Anna B, Aleksandar S. (2012) Smooth control over jerk with displacement constraint. *Journal of mechanical engineer science* **226**, 2656-2673.

[9] An Yonghui, Jo Hongki, Spencer B F, et al. (2014) A damage localization method based on the “jerk energy”. *Smart material and structures* **23**, 025020.

[10] Y. A. Mitropolskii and N. V. Dao. (1997) *Applied Asymptotic Method in Nonlinear Oscillations*, Kluwer Academic Publishers, Dordrecht.

[11] H. Hu. (2008) Perturbation method for periodic solutions of nonlinear jerk equations. *Physics letter A* **372**, 4205-4209.

[12] J. I. Ramos. (2010) Analytical and approximate solutions to autonomous, nonlinear, third-order ordinary differential equations. *Nonlinear Analysis: Real World Applications* **11**, 1613-1626.

[13] Zhang X., Song K. Z., Lu M. W. (2000) Meshless methods based on collocation radial basis function. *Computational Mathematics* **26**, 333–343

[14] Wen L.， Ming L.， Chen C. S., Liu X. F. (2015) Compactly supported radial basis functions for solving certain high order partial differential equations in 3D. *Engineering Analysis with Boundary Elements* **55**, 2-9

[15] Wu Z. M. (1995) Compactly supported positive definite radial function. *Advances in Computational Mathematics* **4**, 283-292.

[16] Wendland H. (1995) Piecewise polynomial, positive definite and compactly supported Radial functions of minimal degree. *Advances in Computational Mathematics* **4**, 389–396.

[17] Buhmann M. D. (1998) Radial function on compact support. Proceedings of the Edinburgh Mathematical Society **41**, 33-46.

[18] Xu J. Q., Li Z. L., Wu L. J. (2014) A Calculation Method for Structural Dynamic Responses Based on the Approximation Theory of Radial Basis Function. *Applied Mathematics and Mechanics* **35**, 533-541.

[19] Li, Y. T., Xu, X. B., Zhou, S. L., Xu J. Q. (2016) A numerical approximation method for nonlinear dynamic system based on radial bases functions. *Applied Mathematics and Mechanics* **37**, 311--318.

[20] Jian K. L., Wen W. B., Luo S. M. (2014) Step by step integration method based on uniform septuple B-spline interpolation for solving structural dynamic response. *Journal of vibration and shock* **33**, 171-176.

[21] H. P. W. Gottlieb. (2004) Harmonic banlance approach to periodic solutions of nonlinear jerk equations. *Journal of sound and vibration* **271**, 671-683.

A Segmentation Method for Intracoronary Optical Coherence Tomography (OCT) Image Based on Least Squares Support Vector Machine: Vulnerable Coronary Plaque Cap Thickness Quantification

†Xiaoya Guo,[1] Dalin Tang,[1*,2] David Molony, [3] Chun Yang,[2] Habib Samady[3], Jie Zheng[4], Gary S. Mintz[5], Akiko Maehara[5], Liang Wang[2], Xuan Pei[6], Zhi-Yong Li[6], Genshan Ma[7*], Don P. Giddens[3,8]

[1] Department of Mathematics, Southeast University, Nanjing, 210096, China
[2] Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA
[3] Department of Medicine, Emory University School of Medicine, Atlanta, GA, 30307, USA
[4] Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO, 63110, USA
[5] The Cardiovascular Research Foundation, Columbia University, New York, NY 10022, USA
[6] School of Biological Science & Medical Engineering, Southeast University, Nanjing, China
[7] Department of Cardiology, Zhongda Hospital, Southeast University Nanjing, 210009, China
[8] The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA

† Presenting author: Xiaoya Guo, Southeast University, Nanjing China;
* Corresponding authors: Dalin Tang, Southeast University, Nanjing China, email: dtang@wpi.edu, fax number: 508 831-5824; Genshan Ma, Southeast University, Nanjing China, magenshan@hotmail.com.

Abstract

Accurate cap thickness quantification is of fundamental importance for vulnerable plaque research. A segmentation method for intracoronary optical coherence tomography (OCT) image based on least squares support vector machine (LS-SVM) to characterize plaque lumen surface, segment borders and fibrous cap for plaque cap thickness when image quality is not high enough, especially at the location of bifurcation.

In vivo intravascular ultrasound (IVUS) and OCT coronary plaque data were acquired from one patient with informed consent obtained. Manual segmentation in OCT images based on the combination of VH-IVUS image and OCT image were given by experts as the gold standard. Processed OCT images were trained and tested via LS-SVM by two methods (M1 and M2). In M1, 500 pixels were randomly selected from each lipid class and vessel tissue class for 9 OCT images. The training data set would be the feature vectors from 9000 pixels. In M2, a procedure similar to leave-one-out cross validation was employed as any 8 out of 9 images were used as training data while the remaining one as the testing data. Borders and lipid contours were extracted from prediction results for cap thickness. Virtual histology (VH) IVUS data were processed with minimum cap thickness set as 50 and 180 micron to generate IVUS50 and IVUS180 data sets, respectively. Cap thickness from manual segmentation, predictions from M1 and M2 based LS-SVM, IVUS50 and IVUS180 data sets were compared.

The accuracy of M1 and M2 were above 76%. Average of mean cap thickness (unit: mm) from 9 images was 0.561 (manual), 0.470 (M1), 0.463 (M2), 0.128 (IVUS50) and 0.204 (IVUS180). Average of minimum cap thickness (9 slices) was 0.390 (manual), 0.288 (M1), 0.282 (M2), 0.040 (IVUS50) and 0.165 (IVUS180). IVUS50 and IVUS180 underestimated cap thickness. The mean cap thickness from prediction were close to manual results (error<18%). The point-point cap thickness from five groups showed that the prediction based LS-SVM had agreement with manual segmentation.

Conclusion. The segmentation methods based on LS-SVM provided reasonable accuracy for plaque cap thickness quantification. More data sets and better gold standard are needed for further improvement.

Keywords: Vulnerable plaque; coronary; OCT; cap thickness; LS-SVM; segmentation.

1. Introduction

Coronary atherosclerotic plaques often rupture without warning and cause acute syndromes, such as heart attack. It is commonly believed that atherosclerotic plaques rupture is closely associated with thin fibrous cap (cap thickness < 65 μm) which covers a large lipid-rich necrotic core [1-2]. Advanced medical imaging technologies have been employed to identify the plaque morphology. Intravascular ultrasound (IVUS) is one of the major imaging tools to visualize coronary plaques, and is also extensively used for constructing computational coronary models. Virtual histology IVUS (VH-IVUS) segments the grayscale IVUS data into four tissue types: fibrotic, fibro-fatty, lipid necrotic core and dense calcified tissue [3-5]. Although VH-IVUS has provided unique perspectives on in vivo human atherosclerotic plaque and novel insights into plaque morphology and natural history, it still has limited resolution of 150-250 micron and is not able to detect thin plaque cap with thickness around 65 μm, the critical threshold value determined from histopathological studies [1, 6]. Quantifying cap thickness and its impact on stress/strain calculations remains to be a great challenge for plaque research.

In recent years, optical coherence tomography (OCT) has emerged as an imaging modality to identify vulnerable plaques from non-vulnerable plaques to enhance future risk prediction. This medical imaging technology provides a resolution of approximately 10 μm which is far better than what IVUS could achieve. OCT has the capacity of characterizing the superficial structure of the vessel in greater detail [7-8]. Traditionally, the segmentation of OCT images to detect lumen and plaque components has been performed manually [9-10]. Rabel et al. introduced a manual method combined IVUS/VH-IVUS and OCT images to characterize the plaque [11]. The combination of IVUS/VH-IVUS and OCT images provided more complete plaque information. However, Manual segmentation is a time-consuming procedure and it also suffers higher inter- and intra-user variability. To overcome these limitations, some approaches for semi-automatic and fully-automatic segmentations have been proposed. Wang et al. introduced a semiautomatic segmentation of calcified plaques which is signal-poor regions with sharply delineated borders in OCT images [12]. Wang et al. provided semiautomatic segmentation using a dynamic programming algorithm to quantified fibrous cap [13]. The methodology using attenuation coefficients to characterize plaques [14-15]. Athanasiou developed an automated methodology for the segmentation of the composition of the superficial plaque in OCT images [16]. Some researchers used support vector machine (SVM) to classify plaque components [17-18]. Most methods in previous literatures aimed at segmenting high quality images. However, image quality is often affected by bifurcation, thrombus, and residual blood, etc. Necrotic core is signal-poor regions which have poor diffuse borders in OCT images. Accurate detection of lipid-rich necrotic core (lipid in short) is still a challenging task.

While OCT has high resolution and can quantify thin cap thickness, it has limited penetration (1 to 2 mm) and is often unable to provide complete information of the whole vessel, especially at the

locations with large lumen, bifurcation and increased plaque burden [7]. In this paper, we introduced an OCT segmentation method to characterize plaque lumen surface, segment borders and fibrous cap for plaque cap thickness when image quality is not high enough, especially at the location of bifurcation. Support vector machine was used in separating lipid component and vessel tissue. Validation was performed based on the segmented results from experts on VH-IVUS and OCT images.

2. Data and Methods

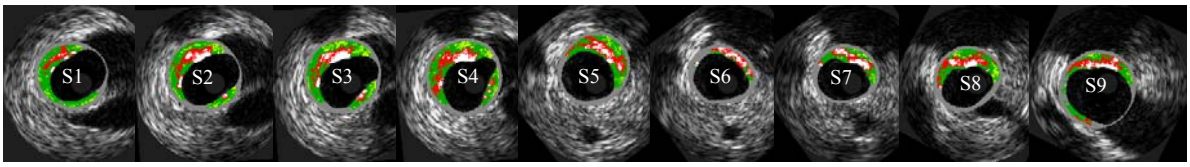
2.1 Data acquisition

In vivo OCT data were acquired from one patient (male, age 51) using ILUMIEN OPTIS System, and Dragonfly JP Imaging Catheter (St. Jude, Minnesota, MN) with informed consent obtained. During image acquisition, blood was displaces by the injection of Dextran, and the OCT catheter was traversed to the region of interest an automatic pullback at 20mm/s was performed. OCT raw data were logarithmically compressed grayscale images with pixel intensity recorded as a two-dimensional matrix in polar coordinates $I(r,\Theta)$, where r is the range dimension and Θ is the acquisition angle. The OCT image in polar coordinate $I(r,\Theta)$ was converted to Cartesian coordinates I(i,j) for later image processing, using the relations: $i = r \cos \Theta$, $j = r \sin \Theta$. OCT images in Cartesian coordinates consists of 1024 by 1024 pixels representing a real physical size of 10mm by 10mm (pixel size: 9.9 by 9.9 micron).

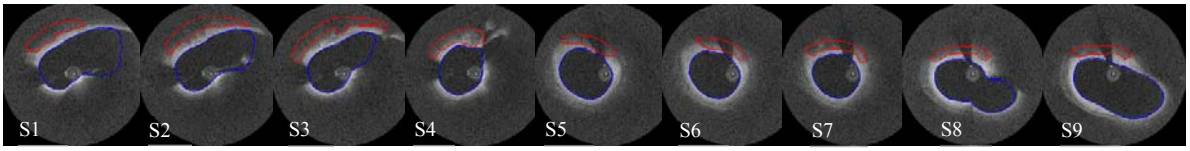
IVUS images were recorded for the same segment of coronary using a 20MHz, 2.9F phased-array catheter (Eagle Eye Gold, Volcano Therapeutics, Rancho Cordova, CA). Following the OCT image acquisition, the IVUS catheter was traversed distally though the artery to the region of interest at an automatic pullback speed of 0.5mm/s. VH-IVUS images were created based on IVUS images by using autoregressive models to differentiate four plaque types: fibrous, fibro-fatty, necrotic core (lipid) and dense calcified tissue (calcification, Ca in short). Each plaque type is indicated by different color on the IVUS-VH image. The positions of both catheters were recorded with angiography for the co-registration of OCT and IVUS images. All imaging data were digitally stored for offline analysis.

Nine OCT images were selected from OCT data set of this patient with co-registered VH-IVUS images. Images from the locations of bifurcation were also considered in this paper. Tiny calcium was neglected in OCT images since fibrous cap was our main interested region. Experts provided manual lipid contours in OCT images based on the combination of VH-IVUS and OCT images following the standard procedure described in [11]. The manual lipid delineation was considered as the standard of classification and saved as image masks for comparison with our segmentation. Figure 1 shows VH-IVUS and OCT images as well as manually segmented lipid and lumen contours.

(a) VH-IVUS Slices



(b) Matching OCT Slices and Segmented Contours



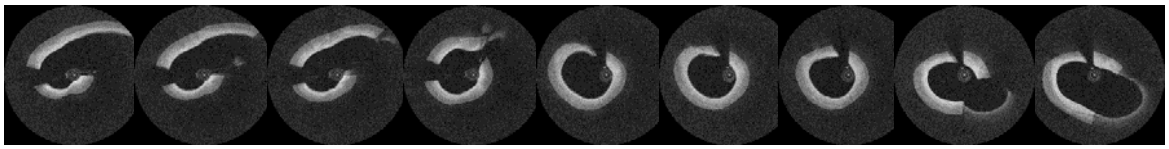
(c) Manual Segmented Contours

Figure 1. Nine matching VH-IVUS and OCT images and manual segmented contours. (a) Nine VH-IVUS slices; (b) Nine matching OCT images with the segmented contours; (c) Manual segmented contours combined OCT and VH-IVUS. Colors used in VH-IVUS: *Red* necrotic core, *White* dense calcium, *Dark Green* Fibrous, *Light Green* Fibro-Fatty.

2.2 Image processing

OCT images in polar coordinate were processing to detect lumen border by applying the following steps: filter, Otsu's thresholding method and removal of catheter and artifacts. More details of automatic lumen detection were given in [19-20]. Since fibrous cap thickness was an important focus in this study, the lumen border was expanded from center of catheter outward 1mm as the outer border to include enough space for fibrous cap characterization. At the locations of bifurcation, branches of the coronary without catheter passing through were removed along with its reduced artifacts. Thus the region of interest (ROI) was determined as the area bounded by lumen and outer border with artery branches and other artifacts removed. The detection of borders and the determination of ROI were showed in Figure 2.

(a) Region of interest



(b) Borders of ROI

Figure 2. Nine OCT images with ROI and borders of ROI. (a) ROI in OCT images. White color means ROI. (b) Borders of ROI.

2.3 Feature extraction and selection

Local binary patterns (LBPs), Gray level co-occurrence matrices (GLCMs), entropy and mean value were calculated as features in this work. All features were computed in an 11 by 11 pixels

neighborhood window in ROI. The window size had been tested to be the best in [16]. Gray level co-occurrence matrices (GLCMs) are a well-known method for texture analysis and 32 statistical features were extracted from GLCMs which were contrast, correlation, energy and homogeneity features for each of eight angles (Θ= 0, 45, 90, 135,180, 225, 270, 315 degree). LBPs is an image operator used to describe the local texture feature with rotating invariance and grayscale invariance. Rotation invariant LBPs with P=8, R=1 was used to extract ten features. Details and parameters can be found from [16, 21].

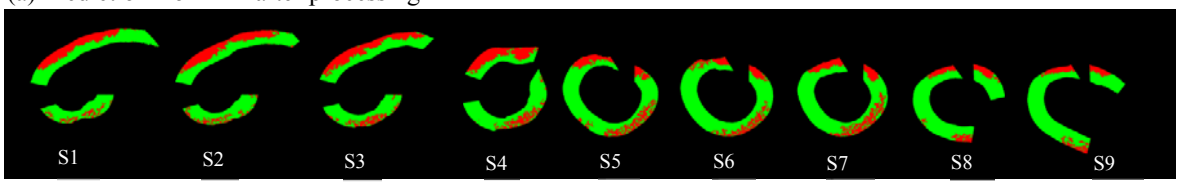
Superabundant features would affect the classification accuracy of classifier. Twenty-eight features were selected to improve the accuracy which contains 16 GLCMs (The four aforementioned features at each angle Θ= 0, 45, 90, 135 degree), 10 LBPs, entropy and mean value. These 28 features were the optimal feature combination to reach the highest accuracy for our classification algorithm based on this patient OCT images.

2.4 Classification

The data base to feed support vector machine (SVM) were assembled by all selected features extracted from all pixels in ROI into a data matrix with dimension n by m, where n is equal to the number of pixels and m is the length of the feature vector, more specifically 28 features in this work. Since fibrous cap could be identified once the lipid area is determined, only two classes are needed to be characterized indicating two different tissue types: lipid and vessel tissue. Since the recognitions of fibrous tissue and fibro-fatty tissue have no impact on the measurement of fibrous cap, so they are considered as vessel tissue. SVM separates feature patterns for two classes by defining one hyperplane that maximizes the separating margin between two classes. Least squares support vector machine (LS-SVM) follows the structural risk minimization principle of kernel function as learning machine which is an improved the conventional SVM. LS-SVM classifier was chosen to classify the lipid and vessel tissue with Gaussian Radial Basis Function as the kernel function and steepest descent method for searching optimal parameters. More details about LS-SVM can be found in [22-23].

Two methods were proposed to select the training data set to fit SL-SVM classifier. In the first method (denoted as M1), 500 pixels were randomly selected from each lipid class and vessel tissue class for 9 OCT images. The training data set would be the feature vectors from 9000 pixels. The rest of non-selected pixels in ROI of each image would be treated as testing data set, The process was repeated ten times to stabilize the classification error estimation. While in the second method (denoted as M2), a procedure similar to leave-one-out cross validation was employed as any 8 out of 9 images were used as training data while the remaining one as the testing data. For each group of 8 OCT images, 500 pixels were random selected from lipid class and vessel tissue class each to give 8000 pixels. The testing set consists of all pixels in ROI of the rest one OCT image. Nine times training and testing were done in this method. Figure 3 showed nine images of prediction from LS-SVM after processing using M1 and M2.

(a) Prediction from M1 after processing



(b) Prediction from M2 after processing

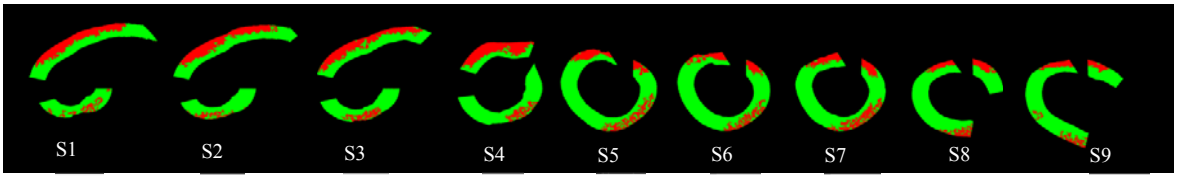


Figure 3. Prediction by LS-SVM from M1 and M2. (a) Prediction of 9 images from M1 after processing; (b) Prediction of 9 images from M2 after processing. Colors used in Prediction images: *Red* necrotic core, *Green* Fibrous and Fibro-Fatty.

2.5 Contours extraction

Lipid contours were extracted and its distance with lumen contour were calculated as cap thickness. The fibrous cap thickness was the greatest concern in this study, so lipid contours were extracted to quantify the cap thickness. According to the prediction of LS-SVM, the prediction images were processed to extract contours. Some scattered pixels in the images of prediction were filter to find the clear border in fibrous cap. Parametric active contour model was used for the lipid border detection. Figure 4 showed contours from active contour model.

(a) OCT contours from M1

(b) OCT contours from M2

Figure 4. Lipid contours and borders. Color contours: *Red* lipid contours, *Blue* borders.

2.6 Co-registered VH-IVUS segmentation

Cap thickness was the concerned part in this study. IVUS resolution limitation (150 μm), and original VH-IVUS data often had lipid core exposed directly to lumen, i.e., there is no fibrous cap covering the lipid core. A minimum cap thickness has to be added to VH-IVUS data so that the lipid cores would be covered. In our previous publications, 50 μm and 180 μm cap thickness were often added in VH-IVUS segmentation where 50 μm was chosen as it is a reasonable representative thin cap number < 65 μm threshold. VH-IVUS images were segmented by the custom-made software Atherosclerostic Plaque Segmentation software. In this work, minimum cap thickness value (50 μm and 180 μm) were set in our segmentation software. The segmented contours denoted as IVUS50 and IVUS180 (see Figure 5).

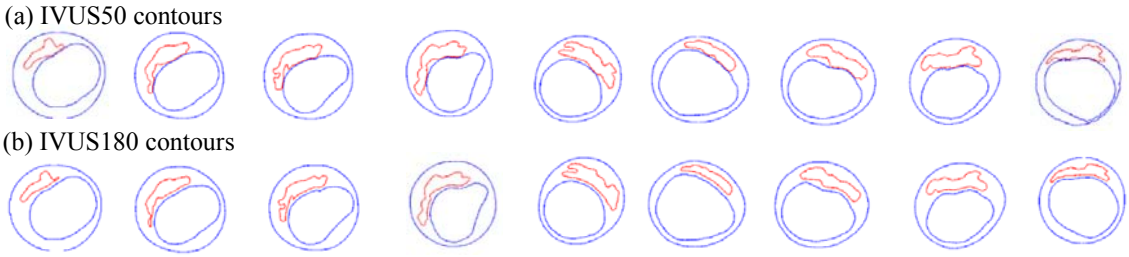


Figure 5. Segmented contours from IVUS50 and IVUS180. Color contours: *Red* lipid contours, *Blue* borders.

3. Results

3.1 Accuracy

The features data were used to train LS-SVM classifier which, in turn, would predict each data point in the testing data set into one of the classes (lipid and vessel-tissue) using the methods M1 and M2 described in Section 2.4. The prediction results was compared to manual segmentation to evaluate the prediction accuracy (denoted as Acc) defined as in the following formula:

$$Acc = \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

Lipid and vessel tissue were set positive and negative pattern, respectively. Here TP is the number of true positive outcomes, FP is the number false positive outcomes, TN is the number of true negative outcomes and FN is the false negative outcomes. Sensitivity (Sen) and specificity (Spe) were also calculated using the formulas below to avoid the reliance of the uneven distributed data in two classes.

$$Sen = \frac{TP}{TP+FN} \tag{2}$$

$$Spe = \frac{TN}{TN+FP} \tag{3}$$

Table 1 gives the prediction results from ten times repeated experiment using M1. Table 1 gives the prediction results from ten times repeated experiment using M1. Sensitivity was about 10% higher than specificity for most OCT images. For the extreme worst case, sensitivity and specificity would still above 70%, while will achieve up to 97% for good cases during the ten times of trial. The average values of sensitivity and specificity are 89.6% and 78.6% respectively.

Table 2 compares the predictions accuracy of M1 and M2. Based on the average values of 9 OCT images, M1 showed a slightly better overall prediction accuracy than M2 (M1 (79.7%) vs M2 (76.1%)). For Sensitivity, M1 provided much higher value than M2 (M1 (89.6%) vs M2 (76.0%)) while the specificity values are quite close (M1 (78.6%) vs M2 (76.6%)). The good agreement between two methods demonstrated the stability of two automatic segmentation method proposed in this paper.

Table 1. Method 1 sensitivity and specificity from 10 tests.

Sensitivity	S1	S2	S3	S4	S5	S6	S7	S8	S9
T1	0.952	0.951	0.860	0.790	0.970	0.910	0.886	0.829	0.930
T2	0.940	0.930	0.851	0.791	0.974	0.922	0.905	0.851	0.927
T3	0.959	0.929	0.852	0.796	0.963	0.919	0.913	0.838	0.937
T4	0.941	0.943	0.859	0.834	0.979	0.906	0.922	0.851	0.941
T5	0.970	0.967	0.866	0.810	0.971	0.897	0.892	0.816	0.938
T6	0.948	0.935	0.862	0.777	0.968	0.915	0.919	0.855	0.941
T7	0.948	0.952	0.856	0.783	0.959	0.906	0.882	0.825	0.908
T8	0.945	0.943	0.874	0.782	0.965	0.878	0.885	0.826	0.938
T9	0.952	0.949	0.843	0.804	0.968	0.865	0.883	0.798	0.910
T10	0.945	0.914	0.817	0.805	0.947	0.896	0.894	0.821	0.880
Ave.	0.950	0.941	0.854	0.797	0.966	0.901	0.898	0.831	0.925
Average sensitivity from all slices:0.896									
Specificity	S1	S2	S3	S4	S5	S6	S7	S8	S9
T1	0.809	0.743	0.787	0.808	0.720	0.769	0.715	0.854	0.847
T2	0.813	0.741	0.790	0.804	0.721	0.770	0.722	0.852	0.845
T3	0.813	0.736	0.786	0.799	0.730	0.766	0.714	0.850	0.842
T4	0.828	0.760	0.822	0.812	0.729	0.776	0.724	0.860	0.852
T5	0.804	0.732	0.785	0.788	0.736	0.789	0.731	0.869	0.858
T6	0.820	0.749	0.801	0.813	0.748	0.746	0.691	0.844	0.843
T7	0.817	0.744	0.795	0.807	0.726	0.775	0.730	0.860	0.847
T8	0.817	0.762	0.805	0.827	0.750	0.797	0.742	0.856	0.854
T9	0.821	0.753	0.798	0.820	0.740	0.786	0.732	0.852	0.852
T10	0.770	0.715	0.705	0.745	0.739	0.771	0.728	0.849	0.834
Ave.	0.811	0.743	0.787	0.802	0.734	0.774	0.723	0.855	0.847
Average specificity from all slices:0.786									

Table 2. Segmentation accuracy for lipid and fibrous tissue using M1 and M2. Sen: Sensitivity; Spe: Specificity; Acc: Accuracy.

	M 1			M 2		
Slice#	Sen	Spe	Acc	Sen	Spe	Acc
1	0.950	0.811	0.834	0.829	0.789	0.795
2	0.941	0.743	0.771	0.792	0.714	0.725
3	0.854	0.787	0.799	0.720	0.761	0.754
4	0.797	0.802	0.800	0.648	0.785	0.743
5	0.966	0.734	0.743	0.900	0.685	0.693
6	0.901	0.774	0.782	0.701	0.762	0.758
7	0.898	0.723	0.740	0.787	0.717	0.724
8	0.831	0.855	0.852	0.646	0.844	0.820
9	0.925	0.847	0.854	0.817	0.835	0.834
Ave.	0.896	0.786	0.797	0.760	0.766	0.761

3.2 Cap thickness

For all 9 co-registered OCT/VH-IVUS images, five different segmentation methods (manual segmentation (denoted as Ma), M1, M2, IVUS50 and IVUS180) were applied to detect the lipid contours and lumen borders. The cap thickness between lumen and a particular lipid was calculated for many locations on the lumen. The average value and minimum over the cap is recorded in Table 3 as mean cap thickness and min cap thickness. Using manual segmentation as standard, M1 and M2 generally gave better results than IVUS50 and IVUS180. The average of mean cap thickness (9 slices, unit: mm) from 5 groups were 0.561, 0.470, 0.463, 0.128, and 0.204, respectively. Average values from M1 and M2 were about same value. The average values of M1, M2, IVUS50 and IVUS180 were 16.3%, 17.4%, 77.2% and 63.6% less than Ma, respectively. Average of minimum cap thickness (9 slices) from 5 groups were 0.390, 0.288, 0.282, 0.040, and 0.165, respectively. The average min values of M1, M2, IVUS50 and IVUS180 were 26.0%, 27.5%, 89.6% and 57.6% less than Ma, respectively.

Table 3. Mean cap thickness and minimum cap thickness for 9 slices from five groups.

Mean Cap Thickness(mm)									
Slice#	Ma	M1	error	M2	error	IVUS 50	error	IVUS 180	error
1	0.597	0.484	18.9%	0.459	23.1%	0.166	72.2%	0.239	60.1%
2	0.683	0.483	29.2%	0.446	34.8%	0.123	82%	0.208	69.5%
3	0.765	0.492	35.7%	0.478	37.5%	0.094	87.8%	0.196	74.4%
4	0.337	0.383	13.5%	0.350	3.9%	0.120	64.5%	0.203	39.8%
5	0.575	0.399	30.6%	0.455	20.9%	0.173	69.9%	0.216	62.5%
6	0.491	0.484	1.3%	0.459	6.5%	0.087	82.3%	0.178	63.6%
7	0.477	0.412	13.6%	0.412	13.7%	0.139	70.9%	0.209	56.2%
8	0.539	0.572	6.2%	0.549	1.9%	0.144	73.3%	0.210	61.1%
9	0.584	0.515	11.9%	0.561	4.7%	0.103	82.3%	0.181	69%
Ave.	0.561	0.470	16.3%	0.463	17.4%	0.128	77.2%	0.204	63.6%
Min Cap Thickness(mm)									
1	0.497	0.308	38%	0.314	36.9%	0.050	89.8%	0.172	65.4%
2	0.429	0.330	23.1%	0.225	47.6%	0.042	90.2%	0.165	61.5%
3	0.539	0.285	47.2%	0.285	47.2%	0.035	93.5%	0.170	68.5%
4	0.158	0.150	5.0%	0.163	3.0%	0.033	78.9%	0.159	0.5%
5	0.331	0.163	50.8%	0.227	31.6%	0.042	87.4%	0.165	50.3%
6	0.326	0.285	12.7%	0.274	16.0%	0.039	88.1%	0.162	50.5%
7	0.392	0.338	13.7%	0.295	24.7%	0.043	89.2%	0.167	57.3%
8	0.445	0.413	7.3%	0.427	4.0%	0.050	88.7%	0.164	63.1%
9	0.390	0.324	16.8%	0.333	14.5%	0.030	92.4%	0.163	58.2%
Ave.	0.390	0.288	26.0%	0.282	27.5%	0.040	89.6%	0.165	57.6%

3.3 Point-point cap thickness

The cap thickness were compared by point to point for from five segmentation methods. Some points from IVUS contours were matched one-by-one by sharing the same location with some points from OCT contours. Slice 4 had the minimum of min cap thickness less than 180μm

according to the manual segmentation (0.158 mm), and the slice 1 had the maximum of min cap thickness in 9 slices. So slice 1 and slice 4 were selected to match points for comparing the point-point cap thickness. The matching points were showed in Figure 6. And the cap thickness of matched points were given in Table 4 and Table 5. Point 1-Point 4 from slice 4 were close to 180 μm, and the cap thickness difference between M1 and M2 were less than 14%. The cap thickness values of IVUS180 were overestimated about 10% than the value from OCT. The cap thickness values from IVUS50 were underestimated about 70% than the cap thickness from OCT. Point 10-14 from slice 4 were about 300μm. The cap thickness of IVUS50 and IVUS180 were underestimated by about 80% and 50%. The cap thickness values of two segmentation methods (IVUS50 and IVUS180) were largely determined by the previously educated guesses.

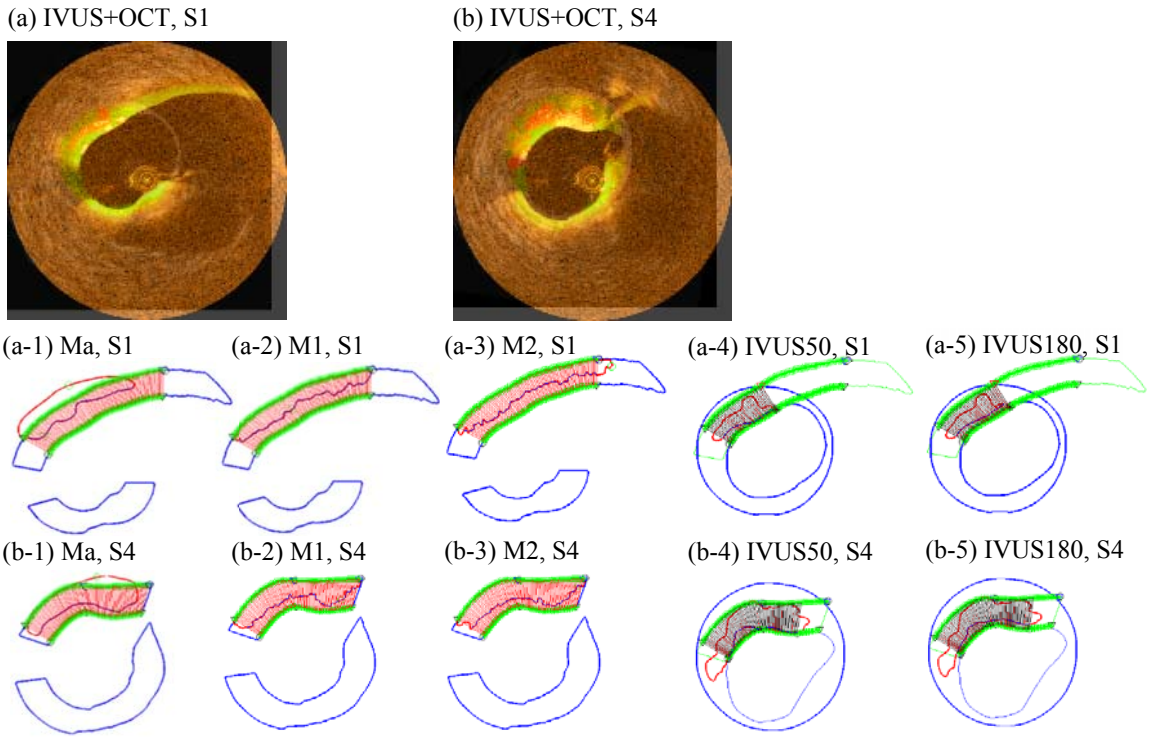


Figure 6. The point-point cap thickness in slice 1 and slice 4. (a) S1 OCT image overlaid with S1 IVUS image. (b) S4 OCT image overlaid with S4 IVUS image. (a-1) cap thickness from Manual segmentation in slice 1. (a-2) cap thickness from M1 in slice 1. (a-3) cap thickness from M2 in slice 1. (a-4) cap thickness from IVUS50 in slice 1. (a-5) cap thickness from IVUS180 in slice 1. (b-*) is in slice 4, and similar to (a-*).

Table 4. Point- point cap thickness in slice 1 from five groups.

Point to Point Cap thickness(mm)					
Slice 1 point	Ma	M1	M2	IVUS50	IVUS180
1	0.642	0.538	0.539	0.057	0.171
2	0.579	0.542	0.551	0.054	0.183
3	0.543	0.443	0.460	0.054	0.171
4	0.522	0.405	0.428	0.054	0.176
5	0.507	0.398	0.376	0.053	0.179
6	0.500	0.411	0.384	0.057	0.171
7	0.497	0.422	0.387	0.053	0.175
8	0.501	0.416	0.351	0.054	0.172
9	0.513	0.434	0.350	0.049	0.178
10	0.544	0.456	0.467	0.052	0.180
11	0.591	0.434	0.406	0.053	0.186
12	0.644	0.440	0.346	0.076	0.191
13	0.675	0.442	0.328	0.082	0.198
14	0.681	0.395	0.321	0.105	0.195
15	0.669	0.400	0.352	0.147	0.191
16	0.650	0.411	0.314	0.149	0.197
17	0.623	0.405	0.347	0.159	0.186
18	0.596	0.451	0.364	0.202	0.203
19	0.572	0.497	0.368	0.224	0.226
20	0.552	0.507	0.532	0.242	0.244
21	0.536	0.485	0.472	0.253	0.254
22	0.524	0.524	0.415	0.285	0.287
23	0.515	0.537	0.571	0.296	0.296
24	0.515	0.526	0.608	0.310	0.312
25	0.528	0.535	0.631	0.326	0.327
26	0.568	0.563	0.565	0.339	0.339
27	0.642	0.591	0.540	0.345	0.345

Table 5. Point- point cap thickness in slice 4 from five groups.

Point to Point Cap thickness(mm)					
Slice 4 point	Ma	M1	M2	IVUS50	IVUS180
1	0.158	0.150	0.165	0.052	0.180
2	0.159	0.171	0.170	0.048	0.179
3	0.164	0.189	0.173	0.045	0.172
4	0.172	0.191	0.202	0.044	0.169
5	0.182	0.192	0.223	0.038	0.166
6	0.190	0.246	0.204	0.038	0.167
7	0.303	0.314	0.290	0.050	0.173
8	0.307	0.300	0.289	0.055	0.178
9	0.312	0.321	0.279	0.051	0.171
10	0.319	0.430	0.283	0.052	0.173
11	0.329	0.427	0.298	0.056	0.177
12	0.337	0.320	0.311	0.060	0.179
13	0.340	0.326	0.311	0.076	0.181
14	0.346	0.340	0.296	0.117	0.180
15	0.350	0.340	0.346	0.284	0.284
16	0.351	0.337	0.369	0.283	0.283
17	0.355	0.321	0.361	0.292	0.292
18	0.359	0.344	0.353	0.305	0.305
19	0.391	0.396	0.388	0.330	0.330
20	0.403	0.387	0.326	0.285	0.285
21	0.416	0.382	0.347	0.239	0.239
22	0.428	0.374	0.375	0.216	0.216

4. Discussion

4.1 Significance of OCT images for cap thickness

Imaging resolution has been a major limitation for vulnerable plaque research (and other areas in a broader sense) since the introduction of medical imaging. With image resolutions at 150 μm (IVUS) – 300 μm (MRI) or even worse, and plaque vessel wall thickness changes were normally under 200 μm, and "vulnerable plaque" cap thickness threshold value defined as 65 μm, many published results were educated guesses by segmentation software. Possessing the capability of detecting thin fibrous cap, OCT is able to provide more accurate cap thickness information to promote both the morphological and mechanical analyses in vulnerable plaque research.

4.2 Prediction for cap thickness

Using manual segmentation as standard, IVUS50 and IVUS180 would underestimate the real cap thickness. Particularly, IVUS50 demonstrated an overall worse ability in estimating the cap thickness. Regarding to some particular thickness values, IVUS180 provided relatively accurate cap thickness measurement when the cap thickness is about 180 μm. While IVUS50 and IVUS180 showed no ability in capturing reasonable cap thickness.

Among the different segmentation methods, predictions from LS-SVM (M1 and M2) showed higher agreements with manual segmentation in measuring mean cap thickness, comparing to IVUS50 and IVUS180. The relative error between M1/M2 and manual segmentation is less than 17.2% for mean cap thickness, and less than 26.4%, if we look at the minimum cap thickness.

4.3 Limitations

One major limitation of this study is lack of histology data as the golden standard. Manual segmentation results based on IVUS and OCT images were considered as the alternative to the golden standard. Another limitation is the small sample size of OCT image studied. Large-scale studies with more OCT image are needed to validate and improve the significance of prediction method.

Funding:

This research was supported in part by NIH grant R01 EB004759 and a Jiangsu Province Science and Technology Agency grant BE2016785.

Conflict of Interest:

There are no conflicts of interest in this study.

References

1. Virmani R, Kolodgie, FD, Burke AP, Farb A, Schwartz SM. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arteriosclerosis Thrombosis, and Vascular Biology* 2000; 20(5), 1262-1275.
2. Finn AV, Nakano M, Narula J, Kolodgie FD, Virmani R. Concept of vulnerable/unstable plaque. *Arteriosclerosis, thrombosis, and vascular biology* 2010; 30(7):1282-92.
3. Tang D, Kamm RD, Yang C, Zheng J, Canton G, Bach R, Huang X, Hatsukami TS, Zhu J, Ma G, Maehara A. Image-based modeling for better understanding and assessment of atherosclerotic plaque progression and vulnerability: Data, modeling, validation, uncertainty and predictions. *Journal of biomechanics* 2014; 47(4):834-46.
4. Wang L, Wu Z, Yang C, Zheng J, Bach R, Muccigrosso D, Billiar K, Maehara A, Mintz GS, Tang D. IVUS-based FSI models for human coronary plaque progression study: components, correlation and predictive analysis. *Annals of biomedical engineering* 2015; 43(1):107-21.
5. Kubo T, Maehara A, Mintz GS, et al. The dynamic nature of coronary artery lesion morphology assessed by serial virtual histology intravascular ultrasound tissue characterization. *Journal of the American College of Cardiology* 2010; 55(15): 1590-1597.
6. Jang I K, Bouma B E, Kang D H, et al. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. *Journal of the American College of Cardiology* 2002; 39(4): 604-609.
7. Tearney GJ, Regar E, Akasaka T, et al. Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies. *Journal of the American College of Cardiology* 2012; 59(12): 1058-1072.
8. Bezerra HG, Costa MA, Guagliumi G, et al. Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications. *JACC: Cardiovascular Interventions* 2009; 2(11): 1035-1046.
9. Jang IK, Tearney GJ, MacNeill B, et al. In vivo characterization of coronary atherosclerotic plaque by use of optical coherence tomography. *Circulation* 2005; 111(12): 1551-1555.
10. Yabushita H, Bouma BE, Houser SL, et al. Characterization of human atherosclerosis by optical coherence tomography. *Circulation* 2002; 106(13): 1640-1645.
11. Räber L, Heo J H, Radu M D, et al. Offline fusion of co-registered intravascular ultrasound and frequency domain optical coherence tomography images for the analysis of human atherosclerotic plaques. *EuroIntervention* 2012; 8(1): 98-108.

12. Wang Z, Kyono H, Bezerra HG, et al. Semiautomatic segmentation and quantification of calcified plaques in intracoronary optical coherence tomography images. *Journal of biomedical optics* 2010; 15(6): 061711-061711-10.
13. Wang Z, Chamie D, Bezerra HG, et al. Volumetric quantification of fibrous caps using intravascular optical coherence tomography. *Biomedical optics express* 2012; 3(6): 1413-1426.
14. Xu C, Schmitt JM, Carlier SG, et al. Characterization of atherosclerosis plaques by measuring both backscattering and attenuation coefficients in optical coherence tomography. *Journal of biomedical optics* 2008; 13(3): 034003-034003-8.
15. Van Soest G, Regar E, KoljenoviÄ S, et al. Atherosclerotic tissue characterization in vivo by optical coherence tomography attenuation imaging. *Journal of biomedical optics* 2010; 15(1): 011105-011105-9.
16. Athanasiou LS, Bourantas CV, Rigas G, et al. Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of biomedical optics* 2014; 19(2): 026009-026009.
17. Shalev R, Prabhu D, Tanaka K, et al. Intravascular optical coherence tomography image analysis method. Biomedical Engineering Conference (NEBEC), 2015 41st Annual Northeast. IEEE, 2015: 1-2.
18. Gargesha M, Shalev R, Prabhu D, et al. Parameter estimation of atherosclerotic tissue optical properties from three-dimensional intravascular optical coherence tomography. *Journal of medical imaging* 2015; 2(1): 016001-016001
19. Ughi GJ, Adriaenssens T, Sinnaeve P, et al. Automated tissue characterization of in vivo atherosclerotic plaques by intravascular optical coherence tomography images. *Biomedical optics express* 2013; 4(7): 1014-1030.
20. Tsantis S, Kagadis G C, Katsanos K, et al. Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography. *Medical physics* 2012; 39(1): 503-513.
21. Haralick RM, Shanmugam K. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* 1973; 3(6): 610-621.
22. Suykens J A K, Vandewalle J. Least squares support vector machine classifiers. *Neural processing letters* 1999; 9(3): 293-300.
23. Suykens J A K, Van Gestel T, De Brabanter J. Least squares support vector machines. *World Scientific* 2002.

A four-noded triangular (Tr4) element for solid mechanics problems with curved boundaries

***J.H. Yue[1]; G.R. Liu[2]; R.P. Niu[a]; †M. Li[a]**

[1] College of Mathematics, Taiyuan University of Technology, Taiyuan, China.

[2] Consultant, Taiyuan University of Technology, Taiyuan, China; School of Aerospace Systems, University of Cincinnati, Cincinnati, USA.

*Presenting author: woyuejunhong@163.com

†Corresponding author: liming04@gmail.com

Abstract

Linear triangular elements with 3 nodes (Tr3) were the earliest, simplest and most widely used in finite element (FE) developed for solving mechanics and other physics problems. The most important advantages of the Tr3 elements are the simplicity, ease in generation, and excellent adaptation to any complicated geometry with straight boundaries. However, it cannot model well the geometries with curved boundaries, which is known as one of the major drawbacks. In this paper, a four-noded triangular (Tr4) element with one curved edge is first used to model the curved boundaries. Two types of shape functions of Tr4 elements have been presented, which can be applied to FEM models based on the isoparametric formulation. FE meshes can be created with mixed linear Tr3 and the proposed Tr4 (Tr3-4) elements, with Tr3 elements for interior and Tr4 elements for the curved boundaries. Compared to the pure FEM-Tr3, the FEM-Tr3-4 can significantly improve the accuracy of the solutions on the curved boundaries because of accurate approximation of the curved boundaries. Several solid mechanics problems are conducted, which validates the effectiveness of FEM models using mixed Tr3-4 meshes.

Keyword: Finite element method; Curved boundaries; Four-noded triangular element; Gaussian integration

1 Introduction

The earliest version of Finite Element Method (FEM) was proposed by Courant [1] in 1943. Since then the FEM has been studied intensively and it is now well developed [2]. It is employed to analyze various engineering systems, such as solid mechanics [3][4], fluid dynamics [5], heat transfer [6][7] and electromagnetics [8] and so on. The FEM has become a powerful and versatile numerical technique to effectively find approximate solutions to practical engineering problems. The general procedure using FEM to solve problems [9]-[10] includes: domain discretization, shape function construction, strain evaluation, element-stiffness-matrix computation, global-matrix assembly, displacement computation, and strain or stress retrieval. In domain discretization, the three-noded linear triangular (Tr3) element is the simplest and most popular element for 2D problems. Some typical triangular elements are shown in Fig. 1a. The triangular elements has a very attractive feature: problem domain with complicated geometries boundaries can be automatically discretized using Tr3 elements, which is essential for automation in modeling and simulation and for adaptive analysis. There are lots of triangular mesh generation algorithms readily available, such as the

Delaunay triangulation algorithm [11][12] proposed by Boris Delaunay, the advancing front technique (AFT) [13][14] developed from the iterative algorithm by Tracy. Owing to its simplicity and ease in generation for solving engineering problems, the Tr3 elements are currently most widely used in practice. However, the accuracy of the FEM using the linear Tr3 elements is poor, especially in stress solutions on the curved boundary, because the stress or strain in the triangular elements is constant.

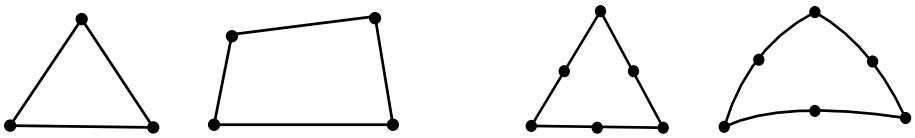


Fig. 1 Traditional finite elements for two dimension: (a) three-noded triangular; (b) four-noded quadrilateral elements; (c) six-noded triangular element with straight; (d) six-noded triangular element curved edges.

To improve stress accuracy, four-noded quadrilateral (Q4) element shown in Fig. 1b are proposed. It is well known that FEM using the Q4 element mesh (FEM-Q4) can obtain more accurate stress solutions than FEM using the linear Tr3 element mesh (FEM-Tr3) with the same node. However, the Q4 element mesh cannot generate a good quality mesh automatically for arbitrarily irregular problem domains, which is its biggest shortcoming. Except the Q4 elements, the higher order element [15]-[17] can also be used for better approximation of the curved boundaries [18][19]. For example, the six-noded isoparametric triangular elements shown in Fig. 1c,d can be automatically generated for complicated geometries and model the curved boundaries. However, the higher order elements are more complicated, and can produce less well-conditioned stiffness matrix with much higher condition numbers. This has been observed clearly from the significantly over-estimated eigenvalues of higher modes. Therefore, in engineering practices, linear elements are often preferred. For the problems with curved boundaries, however, linear elements cannot stimulate well the geometry of the curved boundaries, because the Tr3 element edges are straight lines deviating from the curved boundaries. In this paper, we presents a novel Tr4 elements with an additional node on the curved edge. In a FEM model with curved edges, we use linear Tr3 elements for the most interior part of the problem domain, and Tr4 elements on the curved boundaries. Such a mixed mesh ensures the simplicity and effectiveness for solving problems with curved boundaries.

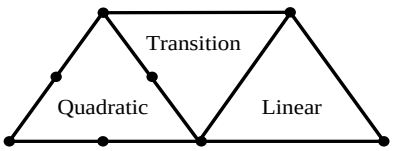


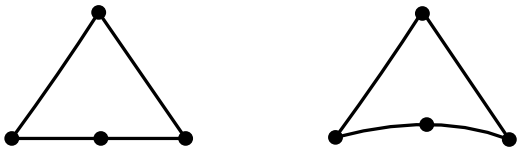
Fig. 2 A transition element with four nodes connecting a linear triangular element and a quadratic triangular element to ensure mesh compatibility.

Tr4 elements are first proposed as a transition element (shown in Fig. 2) to connect quadratic and linear elements to ensure the compatibility of meshes [10]. In our work, Tr4 elements are applied to FEM models for improving the accuracy of the solutions on the curved boundaries. This effectively enables the use of the simplest Tr3 elements for problems with curved boundaries by attaching a layer of Tr4 elements on the curved boundary. We firstly introduce two types of shape functions of the Tr4 element for FEM models. The compatible strain matrices and the element stiffness matrices are then evaluated using the coordinate mapping process and the Gaussian integration [20]-[23] for FEM using Tr4 elements. The stain or stress

solutions on Tr4 elements are recovered by extrapolation way. Finally, the effectiveness of FEM models using Tr3-4 meshes (FEM-Tr3-4) is validated through intensive solid mechanics examples, where Tr3 and Tr4 elements are both used in a mixed fashion. Through these numerical experiments, significant improvements are found on stress solutions on the curved boundaries for the problem domain with curved boundaries for FEM-Tr3-4 models.

2 The proposed idea: a Tr4 element with one curved edge

To enable the Tr3 element for curved boundary, an additional node is added on the middle of one edge, which results in a Tr4 element, as shown in Fig. 3. Through an isoparametric formulation, the edge with 3 nodes can now be curved to fit the curved boundaries. This can overcome the poor stress accuracy caused by inaccurate approximation on the curved boundary. Based on a Tr3 mesh (that can be automatically generated), the Tr4 elements can be created by adding simply a middle-edge node on the curved edge.

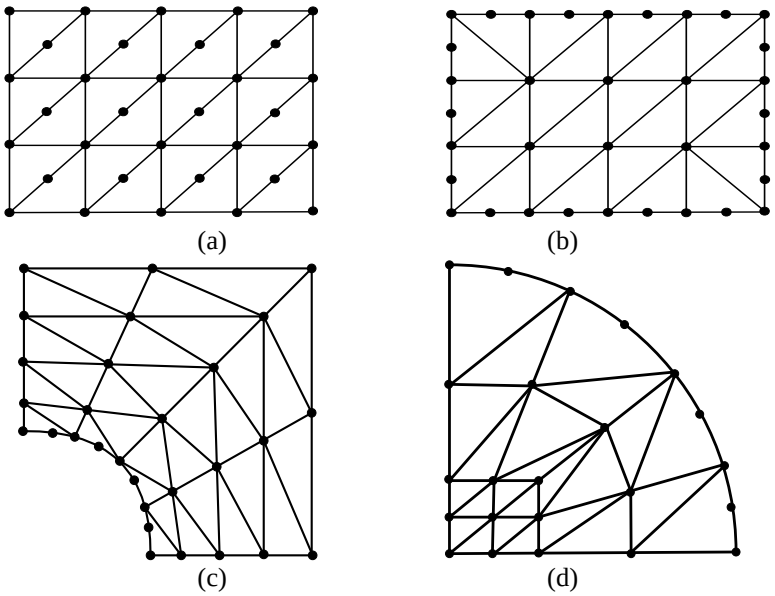


(a) (b)

Fig. 3 Tr4 elements: (a) four-noded triangular element with all straight edges; (b) four-noded triangular element with one curved edge.

2.1 Tr3-4 meshes

There are a few existing works on Tr4 elements are available for the purpose of mesh transition between linear and quadratic elements. In this paper, we consider several possible situations, where Tr4 elements are can be used on the basis of the linear Tr3 element meshes. Fig. 4 shows three meshes that contain Tr4 elements: 1) Tr4 elements with additional nodes on the longer edge, as shown in Fig. 4a; 2) a mesh with mixed Tr3 and Tr4 elements (Fig. 4b), where the additional nodes is added on each boundary (straight) edge; 3) a mesh with mixed Tr3 and Tr4 elements Fig. 4c,d , where a middle node is added on each curved edge. For convenience in discussion later, the FEM only using all Tr4 elements is called FEM-Tr4. The FEM using a mixed Tr3-4 elements mesh is denoted as FEM-Tr3-4.



(a) (b)

(c) (d)

Fig. 4 Possible uses of mixed Tr3-4 meshes: (a) Tr4 element mesh with Tr4 elements only; (b) Tr3-4 element mesh in which Tr4 elements are used for boundary elements and Tr3 elements are used for the interior elements; (c) Tr3-4 element mesh in which Tr4 elements are used for the inner curved boundary and the other part still use Tr3 elements; (d) Tr3-4 element mesh in which T4 elements are used for the outer curved boundary only.

2.2 Shape functions construction

Two types of shape functions of Tr4 elements for FEM models based on isoparametric formulation are constructed. Our formulation for these shape functions will be based on the Cartesian natural coordinates (ξ, η), using the following assumptions:

Case I: Assuming the displacements change linearly in the radial direction as shown in Fig. 5a:

1) In the radial direction originated from node 3, the displacement varies linearly;

2) In the η-direction along with nodes 1, 4, and 2, the displacement varies quadratically.

Case II: the following assumptions are used:

1) In the η-direction the displacement varies linearly, as shown in Fig. 5b;

2) In the ξ-direction along with nodes 1, 4, and 2, the displacement varies quadratically.

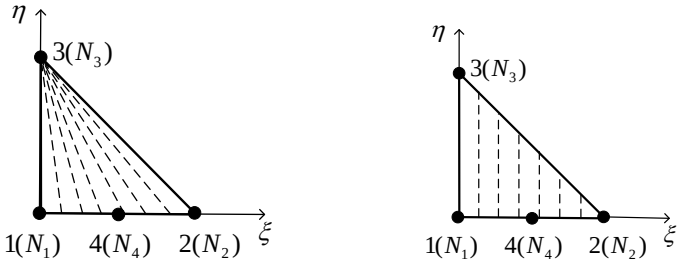


Fig. 5 Displacement variation in a Tr4 element defined in the Cartesian natural coordinates: (a) linear along the radial direction (dash line); (b) linear along the vertical direction (dash line).

Consider first Case I. The shape functions of Tr4 element can be constructed using the following two methods:

Method 1: Direct construction method

First, we perform simple interpolation for a displacement component *u*. Consider any point *P* in the Tr4 element, through which point, a curved segment, $\alpha-\beta-\gamma$ being 'parallel' to the edge 1-4-2, can be drawn, as shown in Fig. 6a. The displacement component at the point is then evaluated using interpolation:

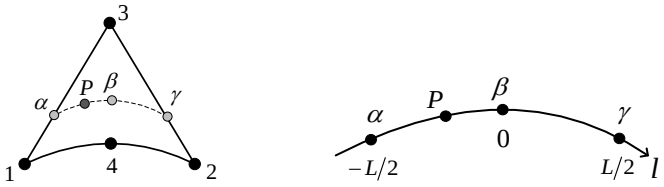
$$u = u_\alpha \phi_\alpha^l + u_\beta \phi_\beta^l + u_\gamma \phi_\gamma^l \tag{1}$$

where $\phi_\alpha^l, \phi_\beta^l, \phi_\gamma^l$ are, respectively, the shape functions of point α, β, γ, and can be written as

$$\phi_\alpha^l = \zeta(\zeta-1)/2,\ \phi_\beta^l = (1-\zeta)(1+\zeta),\ \phi_\gamma^l = \zeta(\zeta+1)/2 \tag{2}$$

in which $\zeta = 2l/L, \zeta \in [-1,1]$, L is the (arc) length of the curved segment $\alpha-\gamma$, and l

represents local curve-linear coordinate value (with the original point β) at point *P* on the segment $\alpha-\beta-\gamma$ (see Fig. 6b)



(a) (b)

Fig. 6 (a) Interpolation over Tr4 element with one curved edge. (b) Local curve-linear coordinate.

and $l=l_{\alpha-p}-l_{\alpha-\beta}$, $l_{\alpha-p}$ and $l_{\alpha-\beta}$ represent, respectively, the (arc) lengths of the curved segment $\alpha-p$ and the curved segment $\alpha-\beta$.

Now, the displacement in the Tr4 element can be evaluated using the following formulation

$$\begin{aligned} u &= u_\alpha \phi_\alpha^l + u_\beta \phi_\beta^l + u_\gamma \phi_\gamma^l \\ &= \left((1-\kappa)u_3 + \kappa u_1\right)\phi_\alpha^l + \left((1-\kappa)u_3 + \kappa u_4\right)\phi_\beta^l + \left((1-\kappa)u_3 + \kappa u_2\right)\phi_\lambda^l \\ &= \kappa\phi_\alpha^l u_1 + \kappa\phi_\gamma^l u_2 + (1-\kappa)u_3 + \kappa\phi_\beta^l u_4 \\ &= N_1 u_1 + N_2 u_2 + N_3 u_3 + N_4 u_4, \end{aligned} \tag{3}$$

where the displacements at α, β and γ are calculated using:

$$u_\alpha = (1-\kappa)u_3 + \kappa u_1, u_\beta = (1-\kappa)u_3 + \kappa u_4, u_\gamma = (1-\kappa)u_3 + \kappa u_2,$$

in which using the fact that

$$l_{3-\alpha}/l_{3-1} = l_{3-\beta}/l_{3-4} = l_{3-\gamma}/l_{3-2} = \kappa, \tag{4}$$

and l_{i-j} represents the length of segment $i-j$.

Therefore, the general form of shape functions $N_i\,(i=1,2,3,4)$ are expressed as

$$\begin{cases} N_1(\zeta,\kappa) = 1/2\kappa\zeta(\zeta-1) \\ N_2(\zeta,\kappa) = 1/2\kappa\zeta(\zeta+1) \\ N_3(\zeta,\kappa) = 1-\kappa \\ N_4(\zeta,\kappa) = \kappa(1-\zeta)(1+\zeta) \end{cases}, \tag{5}$$

where the parameter $\kappa\in[0,1]$, $\zeta\in[-1,1]$.

Finally, using Eq. (5) and the following relations:

$$\kappa = 1-\eta,\ \xi = 2\xi/(1-\eta)-1. \tag{6}$$

we obtain the shape functions of the Tr4 element in the natural coordinate system:

$$
\begin{aligned}
N_1 &= -r(1-r)(1-s)/4 = (1-2\xi-\eta)(1-\xi-\eta)/(1-\eta),\\
N_2 &= r(1+r)(1-s)/4 = \xi(2\xi+\eta-1)/(1-\eta),\\
N_3 &= (1+s)/2 = \eta,\\
N_4 &= (1-r^2)(1-s)/2 = 4\xi(1-\xi-\eta)/(1-\eta).
\end{aligned}
\tag{7}
$$

Method 2: Indirect construction method

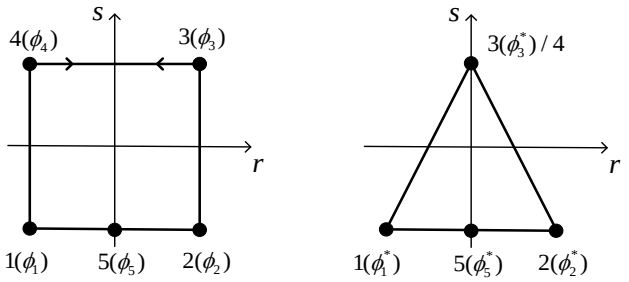


Fig. 7 Collapsing a five-noded quadrilateral element into a four-noded triangular element; (a) Five-noded quadrilateral element; (b) Four-noded triangular element.

The shape functions of five-noded quadrilateral element shown in Fig. 7a are found in Ref. [10]:

$$
\begin{aligned}
\phi_1 &= 1/4(1-r)(1-s) - 1/2\phi_5,\\
\phi_2 &= 1/4(1+r)(1-s) - 1/2\phi_5,\\
\phi_3 &= 1/4(1+r)(1+s),\\
\phi_4 &= 1/4(1-r)(1+s),\\
\phi_5 &= 1/2(1-r^2)(1-s).
\end{aligned}
\tag{8}
$$

The four-noded triangular element can be obtained by collapsing the side 3-4 of the five-noded element as shown in Fig. 7. The shape functions of Tr4 element shown in Fig. 7b are obtained by simply amending shape functions in Eq. (1):

$$
\begin{aligned}
\phi_1^* &= 1/4r(1-r)(1-s)\\
\phi_2^* &= 1/4r(1+r)(1-s)\\
\phi_3^* &= 1/2(1+r)(1+s)\\
\phi_5^* &= 1/2(1-r^2)(1-s)
\end{aligned}
\tag{9}
$$

Mapping the Tr4 element in Fig. 7b into the standard Tr4 element in Fig. 5 using the following relationship:

$$
r = 2\xi/(1-\eta) - 1, \quad s = 2\eta - 1 \tag{10}
$$

which gives the same shape functions given in Eq.(7) for the Tr4 element.

Under the assumptions of case II, shape functions of Tr4 element shown in Fig. 5 are constructed as follows:

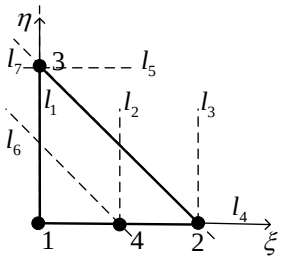


Fig. 8 Standard Tr4 element in natural coordinate system for shape function construction.

Fig. 8 shows a standard Tr4 element in natural coordinates. The equations of lines $l_1, l_2, l_3, l_4, l_5, l_6$ and l_7 shown in Fig. 8 are :

$$\begin{aligned} &l_1 : \xi = 0, \quad l_2 : \xi = 1/2, \quad l_3 : \xi = 1, \\ &l_4 : \eta = 0, \quad l_5 : \eta = 1, \\ &l_6 : 1/2 - \xi - \eta = 0, \quad l_7 : 1 - \xi - \eta = 0. \end{aligned} \tag{11}$$

Based on Delta function properties, Shape functions of the Tr4 elements, satisfies that

$$N_1\big|_{l_2, l_7} = 0, \quad N_2\big|_{l_1, l_6} = 0, \quad N_3\big|_{l_4} = 0, \quad N_4\big|_{l_1, l_7} = 0, \tag{12}$$

and

$$N_1\big|_{(0,0)} = 1, \quad N_2\big|_{(1,0)} = 1, \quad N_3\big|_{(0,1)} = 1, \quad N_4\big|_{(1/2,0)} = 1. \tag{13}$$

It is from Eq.(11) that the shape functions can be assumed as:

$$\begin{aligned} N_1 &= a_1 (\xi - 1/2)(1 - \xi - \eta), \\ N_2 &= a_2 \xi (1/2 - \xi - \eta), \\ N_3 &= a_3 \eta, \\ N_4 &= a_4 \xi (1 - \xi - \eta). \end{aligned} \tag{14}$$

and from Eq.(13), we have

$$a_1 = -2, \quad a_2 = -2, \quad a_3 = 1, \quad a_4 = 4. \tag{15}$$

Therefore, we have

$$\begin{aligned} N_1 &= (1 - \xi - \eta)(1 - 2\xi), \\ N_2 &= -\xi (1 - 2\xi - 2\eta), \\ N_3 &= \eta, \\ N_4 &= 4\xi (1 - \xi - \eta). \end{aligned} \tag{16}$$

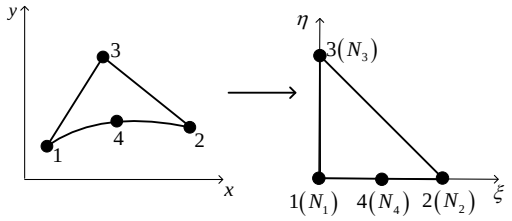
It is note that the shape functions of Tr4 element given in Eq.(16) are consistent with the report in [10], in which the shape functions of Tr4 element are constructed using other way.

It is clear that shape functions $N_i \left(i = 1, 2, 3, 4\right)$ satisfy: (1) Linear independence; (2) Delta function properties; (3) Partitions of unity property. It can be implemented in the standard FEM formulations. For convenience in discussion later, the FEM-Tr3-4 model using the shape

functions in case I and case II are, respectively, denoted as FEM-Tr3-4-N1 and FEM-Tr3-4-N2 models. Similarly, we have the FEM-Tr4-N1 and FEM-Tr4-N2 models for FEM-Tr4 model.

2.3 Strain matrix evaluation

The compatible strain matrix can be evaluated based on isoparametric mapping formulation for FEM models. Fig. 9 shows that a coordinate mapping from physical coordinates to natural coordinates, which is also indispensable for evaluating the strain matrix.



(a) (b)

Fig. 9 Coordinate mapping from physical coordinate system to natural coordinate system; (a). A Tr4 element in the physical coordinate system; (b). A standard Tr4 element in the natural coordinate system.

The isoparametric mapping formulation between two coordinate systems can be expressed as

$$x=\sum_{i=1}^{4}N_i(\xi,\eta)x_i,\quad y=\sum_{i=1}^{4}N_i(\xi,\eta)y_i. \tag{17}$$

where x_i and y_i represent coordinates of the *i*th node.

Hence, the strain matrix $\mathbf{B}_i^e$ for the Tr4 element can be evaluated using the formula

$$\mathbf{B}_i^e=\mathbf{LN}=\begin{bmatrix}\partial N_i/\partial x & 0 & \partial N_i/\partial y\\ 0 & \partial N_i/\partial y & \partial N_i/\partial x\end{bmatrix}^T. \tag{18}$$

where the partial derivatives $\partial N_i/\partial x,\partial N_i/\partial y\left(i=1,2,3,4\right)$ can be evaluated using

$$\begin{bmatrix}\partial N_i/\partial x\\ \partial N_i/\partial y\end{bmatrix}=\begin{bmatrix}\partial x/\partial\xi & \partial y/\partial\xi\\ \partial x/\partial\eta & \partial y/\partial\eta\end{bmatrix}^{-1}\begin{bmatrix}\partial N_i/\partial\xi\\ \partial N_i/\partial\eta\end{bmatrix}=\mathbf{J}^{-1}\begin{bmatrix}\partial N_i/\partial\xi\\ \partial N_i/\partial\eta\end{bmatrix}. \tag{19}$$

where $\mathbf{J}$ is a Jacobian matrix between the two coordinate systems.

2.4 linear system of equations

For general FEM models, the linear system of equations has the form:

$$\mathbf{Kd}=\tilde{\mathbf{f}}. \tag{20}$$

where $\mathbf{K}$ is the global system stiffness matrix consisting of element stiffness matrices. The element stiffness matrix of the Tr4 elements can be computed using Gaussian integration:

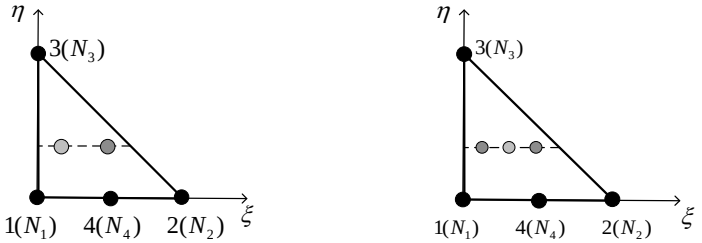
$$\mathbf{k}_e=\int_{A_e}\mathbf{B}_i^{eT}\mathbf{c}\mathbf{B}_i^e dA=\int_0^1\int_0^{1-\xi}\left(\mathbf{B}_i^e\right)^T\mathbf{c}\mathbf{B}_i^e\det\left|\mathbf{J}\right|d\eta d\xi=\sum_{j=1}^{n_G}w_j\left(\left(\mathbf{B}_i^e\right)^T\mathbf{c}\mathbf{B}_i^e\det\left|\mathbf{J}\right|\right)\Big|_{(\xi_j,\eta_j)}. \tag{21}$$

where $\det|\mathbf{J}|$ is determinant of the Jacobian matrix $\mathbf{J}$; (ξ_j,η_j) and w_j, respectively, represent the coordinates and weights of the *j*th Gauss-point; n_G represents the number of Gauss-points. Table 1 lists the number, locations, and corresponding weights of Gauss-points for satisfying different accuracy order of the integrand. For example, three Gauss-points are at least needed for quadratic triangular elements [18], and two Gauss-points are not sufficient to two order of accuracy.

Table 1. The coordinates of Gauss-points and corresponding weights for the standard Tr3 element.

Number of Gauss-points	(ξ,η)	Weights	Order of accuracy
1	(1/3,1/3)	1/2	1
3	(1/6,1/6)	1/6	2
	(4/6,1/6)	1/6	
	(1/6,4/6)	1/6	
7	0.1012865073235	0.1012865073235	5
	0.7974269853531	0.1012865073235	
	0.1012865073235	0.7974269853531	
	0.4701420641051	0.0597158717898	
	0.4701420641051	0.4701420641051	
	0.0597158717898	0.4701420641051	
	0.3333333333333	0.3333333333333	

It is worthwhile to note that the locations of Gauss-points are different for case I and case II in Section 2.2. For case I, the locations and corresponding weights of Gauss-points on the Tr4 elements are listed in Table 2. For case II, the locations and corresponding weights of Gauss-points on the Tr4 elements are listed in Table 1, which is shown in Fig. 10. The effectiveness of Gaussian integration on the locations of Gauss-points listed in Table 1 and Table 2 has been validated on standard Tr4 element.



(a) (b)

Fig. 10. The locations of Gauss-points for the standard Tr4 element: (a) two Gauss-points; (b) three Gauss-points.

Table 2. The coordinates of Gauss-points and corresponding weights for the standard Tr4 element.

Number of Gauss Points	(ξ,η)	weights
2	$\left(1/3-\sqrt{3}/9,1/3\right)$	1/6
	$\left(1/3+\sqrt{3}/9,1/3\right)$	1/6
3	$\left(1/3-\sqrt{1/15},1/3\right)$	5/18
	$(1/3,1/3)$	8/18
	$\left(1/3+\sqrt{1/15},1/3\right)$	5/18

2.5 Retrieval of strains for the Tr4 elements

For the Tr3 elements, the nodal strain, $\tilde{\boldsymbol{\varepsilon}}(\mathbf{x}_k)$, at the node k is evaluated using the following formula, which is an area-weighted averaged strain of elements connecting to the node k [24],

$$\tilde{\boldsymbol{\varepsilon}}(\mathbf{x}_k) = \frac{1}{A_k^{ne}} \sum_{i=1}^{n_e^k} \tilde{\boldsymbol{\varepsilon}}_i A_i^e \,. \tag{22}$$

where n_e^k and A_k^{ne}, respectively, represent the number and total area of elements connecting to the node k. $\tilde{\boldsymbol{\varepsilon}}_i$ and A_i^e, respectively, represent the strain and area of the *i*th element.

Note that the strain of the Tr3 element is a constant, while the strain of the Tr4 element is not a constant. Hence nodal strain $\tilde{\boldsymbol{\varepsilon}}(\mathbf{x}_k)$ of the Tr4 elements is calculated using the following manner: (1) The nodal strains are evaluated at the sample points in the Tr4 elements. (2) The strains of element-nodes can be obtained using the strains of these sample points in the element by the extrapolation technique. (3) Similarly to Tr3 elements, the extrapolated strains are generally different from those at the same nodes from the adjacent elements, because of the discontinuity of the strains in the FEM-Tr3-4. Hence, the nodal strains are calculated using these extrapolated strains through the area-weighted averaged schedule.

The further elaboration of the steps (1) and (3) does not require, since they can be easily implemented. However, it is necessary to elaborate the extrapolation process in the step (2). Consider a Tr4 element for Case I, and a set of sample points including Gauss-points on a standard Tr4 element is chosen to evaluate the strains at nodes of the element. These sample points are marked in the same order as the node-numbering of element as $\#1^g$, $\#2^g$, $\#3^g$ and $\#4^g$ in Fig. 11a. In the coordinate system $O\xi\eta$, coordinates of element-nodes and these sample points are listed in Table 3.

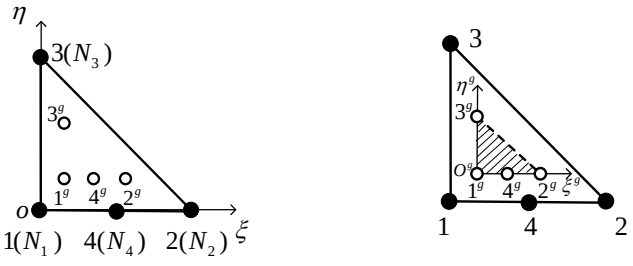


Fig. 11 Extrapolation from sample points to element-nodes: (a). the natural coordinate system $O\xi\eta$; (b) the coordinate system $O^g\xi^g\eta^g$ for the "sample element".

Fig. 11b shows a "standard element" consisting of four sample points. The relationship between the natural coordinate system $O\xi\eta$ and $O^g\xi^g\eta^g$ can be written as

$$\xi^g = (\sqrt{15}/2)\xi - \sqrt{15}/6 + 1/2\,, \quad \eta^g = 3\eta - 1\,. \tag{23}$$

In the coordinate system $O^g\xi^g\eta^g$, coordinates of element-nodes and these sample points are listed in

Table 4, which are obtained by using the relationship in Eq.(23).

Table 3. Coordinates of element-nodes and Gauss-points in the natural coordinate system $O\xi\eta$ for a Tr4 element of Case I.

element-nodes	#1	#2	#3	#4				
Gauss-points					$\#1^g$	$\#2^g$	$\#3^g$	$\#4^g$
ξ	0	1	0	0.5	$\frac{1}{3}-\frac{1}{\sqrt{15}}$	$\frac{1}{3}+\frac{1}{\sqrt{15}}$	$\frac{1}{3}-\frac{1}{\sqrt{15}}$	1/3
η	0	0	1	0	1/3	1/3	2/3	1/3

Table 4. Coordinates of element-nodes and Gauss-points in the natural coordinate system $O^g\xi^g\eta^g$ for a Tr4 element of Case I.

element-nodes	#1	#2	#3	#4				
Gauss-points					$\#1^g$	$\#2^g$	$\#3^g$	$\#4^g$
ξ^g	$\frac{1}{2}-\frac{\sqrt{15}}{6}$	$\frac{1}{2}+\frac{\sqrt{15}}{3}$	$\frac{1}{2}-\frac{\sqrt{15}}{6}$	$\frac{1}{2}+\frac{\sqrt{15}}{12}$	0	1	0	1/2
η^g	-1	-1	2	-1	0	0	1	0

For a Tr4 element, the strains at the sample points, $\tilde{\boldsymbol{\varepsilon}}_j$ ($j=1^g,...,4^g$), can be computed. Using strains of sample points and the nodal shape functions in the coordinate system $O^g\xi^g\eta^g$, strains $\tilde{\boldsymbol{\varepsilon}}$ at nodes of the Tr4 element can be obtained:

$$\tilde{\boldsymbol{\varepsilon}}\left(\xi^g,\eta^g\right)=\tilde{\boldsymbol{\varepsilon}}_g\mathbf{N}_g, \tag{24}$$

where $\tilde{\boldsymbol{\varepsilon}}_g=\left[\tilde{\boldsymbol{\varepsilon}}_{1^g}\quad\tilde{\boldsymbol{\varepsilon}}_{2^g}\quad\tilde{\boldsymbol{\varepsilon}}_{3^g}\quad\tilde{\boldsymbol{\varepsilon}}_{4^g}\right],\mathbf{N}_g=\left[N_1^g\quad N_2^g\quad N_3^g\quad N_4^g\right]^T_{\left(\xi^g,\eta^g\right)}$, and

$$\begin{cases}N_1^g\left(\xi^g,\eta^g\right)=2(\xi^g)^2/(1-\eta^g)-3\xi^g+\left(1-\eta^g\right)\\ N_2^g\left(\xi^g,\eta^g\right)=2(\xi^g)^2/(1-\eta^g)-\xi^g\\ N_3^g\left(\xi^g,\eta^g\right)=\eta^g\\ N_4^g\left(\xi^g,\eta^g\right)=4\xi^g-4(\xi^g)^2/(1-\eta^g)\end{cases}. \tag{25}$$

Let $\xi^g=1/2-\sqrt{15}/6$, $\eta^g=-1$ in Eq.(26), and we can obtain the strain $\tilde{\boldsymbol{\varepsilon}}_1$ at the node 1. Similarly, we can obtain the strain at nodes 2, 3 and 4, and have:

$$\begin{bmatrix}\tilde{\boldsymbol{\varepsilon}}_1\\ \tilde{\boldsymbol{\varepsilon}}_2\\ \tilde{\boldsymbol{\varepsilon}}_3\\ \tilde{\boldsymbol{\varepsilon}}_4\end{bmatrix}=\begin{bmatrix}\frac{7}{6}+\frac{\sqrt{15}}{3} & \frac{1}{6} & -1 & \frac{2}{3}-\frac{\sqrt{15}}{3}\\ \frac{29}{12}-\frac{2}{3}\sqrt{15} & \frac{17}{12} & -1 & -\frac{11}{6}+\frac{2\sqrt{15}}{3}\\ -\frac{13}{3}+\frac{5\sqrt{15}}{6} & -\frac{11}{6}+\frac{\sqrt{15}}{2} & 2 & \frac{14}{3}-\frac{4\sqrt{15}}{3}\\ \frac{41}{48}-\frac{\sqrt{15}}{6} & -\frac{7}{48} & -1 & \frac{31}{24}+\frac{\sqrt{15}}{6}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{\varepsilon}}_{1^g}\\ \tilde{\boldsymbol{\varepsilon}}_{2^g}\\ \tilde{\boldsymbol{\varepsilon}}_{3^g}\\ \tilde{\boldsymbol{\varepsilon}}_{4^g}\end{bmatrix}. \tag{26}$$

Similarly, when a Tr4 element for Case II is considered, we have the relationship between the natural coordinate system $O\xi\eta$ and $O^g\xi^g\eta^g$ can be written as

$$\xi^g = 2\xi - 1/3, \eta^g = 2\eta - 1/3, \tag{27}$$

Using above relationship, Table 5 and Table 6 can be obtained, and we have the extrapolated matrix:

$$\begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_1 \\ \tilde{\boldsymbol{\varepsilon}}_2 \\ \tilde{\boldsymbol{\varepsilon}}_3 \\ \tilde{\boldsymbol{\varepsilon}}_4 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 25 & 7 & -3 & -20 \\ 7 & 25 & -3 & -20 \\ -5 & -5 & 15 & 4 \\ -2 & -2 & -3 & 16 \end{bmatrix}\begin{bmatrix} \tilde{\boldsymbol{\varepsilon}}_{1^g} \\ \tilde{\boldsymbol{\varepsilon}}_{2^g} \\ \tilde{\boldsymbol{\varepsilon}}_{3^g} \\ \tilde{\boldsymbol{\varepsilon}}_{4^g} \end{bmatrix}, \tag{28}$$

Table 5. Coordinates of element-nodes and Gauss-points in the natural coordinate system $O\xi\eta$ for a Tr4 element of Case II.

element-nodes	#1	#2	#3	#4				
Gauss-points					$\#1^g$	$\#2^g$	$\#3^g$	$\#4^g$
ξ	0	1	0	0.5	1/6	4/6	1/6	5/12
η	0	0	1	0	1/6	1/6	4/6	1/6

Table 6. Coordinates of element-nodes and Gauss-points in the natural coordinate system $O^g\xi^g\eta^g$ for a Tr4 element of Case II

element-nodes	#1	#2	#3	#4				
Gauss-points					$\#1^g$	$\#2^g$	$\#3^g$	$\#4^g$
ξ^g	-1/3	5/3	-1/3	2/3	0	1	0	1/2
η^g	-1/3	-1/3	5/3	-1/3	0	0	1	0

Using the above way, the recovery strains and stresses for FEM-Tr3-4 models can be obtained, and they possess high accuracy, compared to those obtained using the FEM-Tr3 model.

3 Numerical examples and discussions

In this section, to validate the effectiveness of the FEM-Tr3-4 or FEM-Tr4 models, three solid mechanics problems are conducted using different numerical models including the standard FEM-Tr3 and the well-known FEM-Q4 models. Relative errors in the displacement and nodal stress defined in the following formulas are used to analyze the results.

$$e_r = \left(\sum_{i=1}^{N} \left(u_i^e - u_i^{num} \right)^2 \Big/ \sum_{i=1}^{N} \left(u_i^e \right)^2 \right)^{1/2}, e_r = \left(\sum_{i=1}^{N} \left(s_i^e - s_i^{num} \right)^2 \Big/ \sum_{i=1}^{N} \left(s_i^e \right)^2 \right)^{1/2}. \tag{29}$$

where u_i^e and s_i^e, respectively, represent the exact solutions in the displacement and stress of the *i*th node; u_i^{num} and s_i^{num}, respectively, represent numerical solutions in the displacement and stress of the *i*th node.

3.1 A Rectangular Cantilever loaded at the end

Then, we consider a rectangular cantilever, which is a plane stress problem. Fig. 12 shows a cantilever with height D and length L, and a parabolic traction is imposed on its free end.

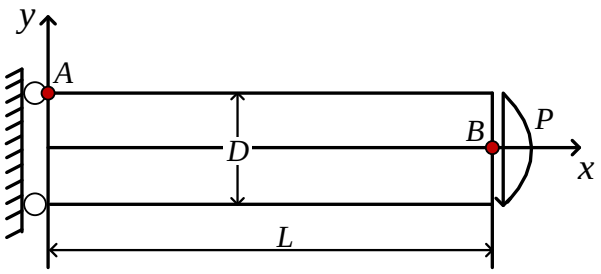


Fig. 12 A rectangular cantilever loaded at the free end.

The analytical solutions for the displacement and stresses are found in [25]:

$$
\begin{aligned}
u_x &= \frac{Py}{6EI}\left[(6L-3x)x+(2+\nu)(y^2-\frac{D^2}{4})\right],\\
u_y &= -\frac{P}{6EI}\left[3\nu y^2(L-x)+(4+5\nu)\frac{D^2x}{4}+(3L-x)x^2\right],\\
\sigma_{xx}(x,y) &= \frac{P(L-x)y}{I}, \sigma_{yy}(x,y)=0, \tau_{xy}(x,y)=-\frac{P}{2I}\left(\frac{D^2}{4}-y^2\right),
\end{aligned}
\tag{30}
$$

where I represents the moment of inertia and can be evaluated using $I = D^3/12$. E and ν, respectively, represent Young modulus and Poisson ratio of the problem.

In the calculation, the above parameters are taken as $D = 12\,\text{m}$, $L = 48\,\text{m}$, $P = 1000\,\text{N}$, $E = 3.0\times10^7\ \text{N/m}^2$ and $\nu = 0.3$. Boundary conditions in the left end and right end of the cantilever are imposed using Eq.(30).

To validate the effectiveness of FEM-Tr3-4-N1 and FEM-Tr3-4-N2 models, we consider the effect of the boundary refinement on the results for the rectangular cantilever. The domain of the cantilever is discretized using linear Tr3 elements for inner elements and Tr4 elements for boundary elements, as shown in Fig. 13.

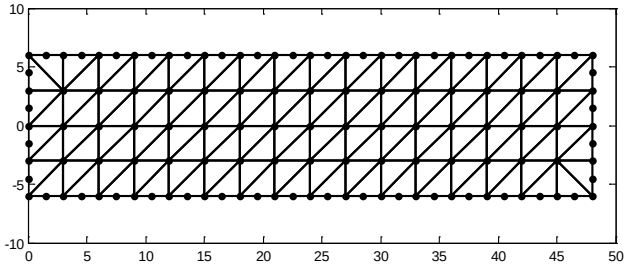


Fig. 13 Mesh using Tr3 elements and Tr4 elements for the rectangular cantilever.

Table 7 gives the relative errors in displacement on the middle points of upper boundary edges using different methods and mesh sizes. The mesh size S is defined as

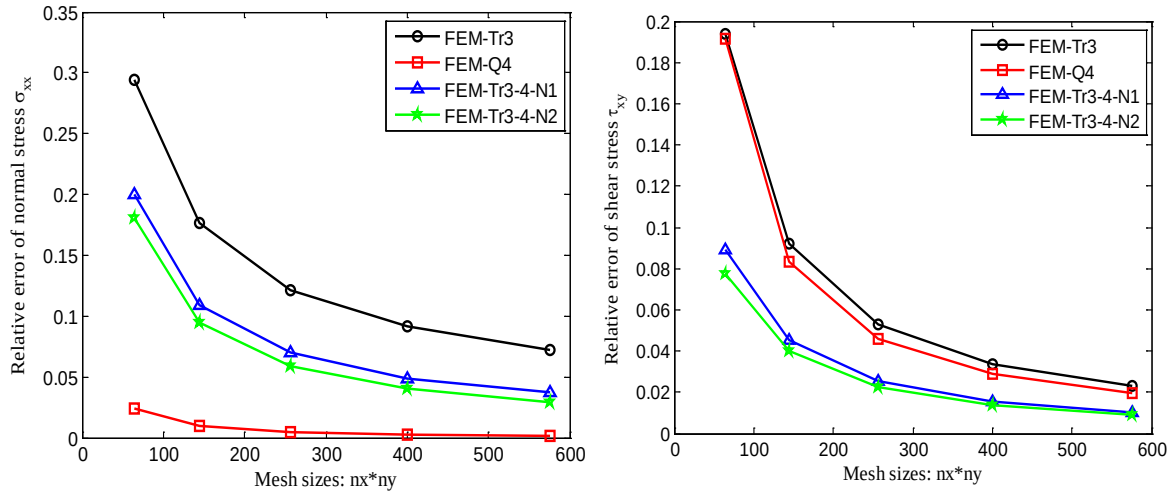
$$S = n_x \times n_y \tag{31}$$

where n_x and n_y ,respectively, represent the number of mesh in *x* direction and *y* direction. The results show that FEM-Tr3-4-N1 and FEM-Tr3-4-N2 models only slightly improve the accuracy in displacement on the boundaries, compared to FEM-Tr3 model.

Table 7. Relative errors in displacement component *u* on the middle nodes of upper boundary edges.

Method / Mesh size	16×4	24×6	32×8	40×10	48×12
FEM-Tr3	1.746e-01	8.698e-02	5.118e-02	3.349e-02	2.355e-02
FEM-Tr3-4-N1	1.734e-01	8.672e-02	5.113e-02	3.342e-02	2.350e-02
FEM-Tr3-4-N2	1.712e-01	8.590e-02	5.073e-02	3.327e-02	2.344e-02
FEM-Q4	3.050e-02	1.381e-02	7.820e-03	5.021e-03	3.493e-03

Consider middle-points of closest to the point A and point B shown in Fig. 12 on boundary edges, which are marked as A' and B' respectively. It is from Eq. (30) that the normal stress σ_{xx} of point A and the shear stress τ_{xy} of point B is maximum for the rectangular cantilever problem. Fig. 14a and Fig. 14b show, respectively, the relative errors in normal stress σ_{xx} of the point A' and in shear stress τ_{xy} of point B' against mesh sizes using different methods. It is clearly observed that FEM-Tr3-4 models stand out in the normal and shear stresses, comparing with FEM-Tr3 models; The results of FEM-Tr3-4-N2 model are the most excellent in the terms of the shear stress solutions, comparing with the FEM-Tr3 and FEM-Q4 models.



(a) (b)

Fig. 14 Relative errors in stresses against mesh sizes using different methods for the rectangular cantilever: (a) normal stress of Point A' (b) shear stress of Point B'.

3.2 Infinite plate with a circular hole

Next, we consider a square plate with the side length *D* and a circular hole with a radius *a* in the center of plate, which is a plane strain problem. The unidirectional tensile load σ is imposed on the right and left boundary along the *x*-direction, and owing to its symmetry, only quarter of the square plate is modeled, as shown in Fig. 15.

The analytical solution for the displacements u_x, u_y and stresses $\sigma_{xx}, \sigma_{yy}, \tau_{xy}$ in the polar coordinates and can be found in [25],

$$
\begin{aligned}
u_x &= \frac{a}{8\mu}\left[\frac{r}{a}(\kappa+1)\cos\theta + 2\frac{a}{r}\left((1+\kappa)\cos\theta+\cos 3\theta\right) - 2\frac{a^3}{r^3}\cos 3\theta\right],\\
u_y &= \frac{a}{8\mu}\left[\frac{r}{a}(\kappa-3)\sin\theta + 2\frac{a}{r}\left((1-\kappa)\sin\theta+\sin 3\theta\right) - 2\frac{a^3}{r^3}\sin 3\theta\right],\\
\sigma_{xx} &= 1-\frac{a^2}{r^2}\left[\frac{3}{2}\cos 2\theta+\cos 4\theta\right]+\frac{3a^4}{2r^4}\cos 4\theta,\\
\sigma_{yy} &= -\frac{a^2}{r^2}\left[\frac{1}{2}\cos 2\theta-\cos 4\theta\right]-\frac{3a^4}{2r^4}\cos 4\theta,\\
\tau_{xy} &= -\frac{a^2}{r^2}\left[\frac{1}{2}\sin 2\theta+\sin 4\theta\right]+\frac{3a^4}{2r^4}\sin 4\theta,
\end{aligned}
\tag{32}
$$

where r and θ represent the polar coordinate values; The parameters μ and κ are evaluated using

$$\mu = E/\left(2(1+\nu)\right),\ \kappa = 3-4\nu. \tag{33}$$

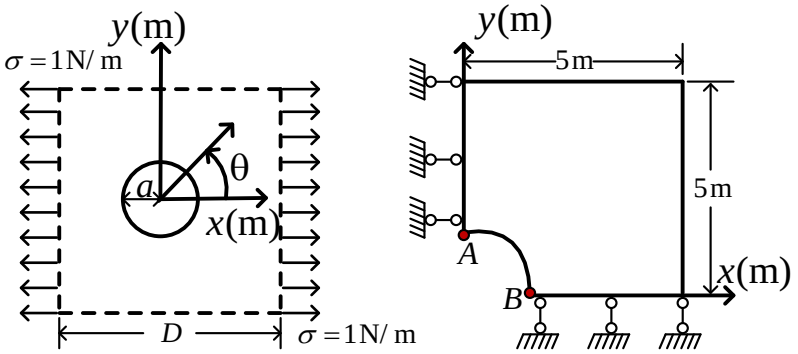


Fig. 15 Infinite plate with a circular hole subjected to unidirectional tension and its quarter model with symmetric conditions imposed on the left and bottom edges.

In the calculation, the above parameters are taken as $D=10\text{m}$, $a=1\text{m}$, $\sigma=1\text{N/m}$ $E=1.0\times10^3\ \text{N/m}^2$ and $\nu=0.3$. The traction boundary conditions and symmetric condition, respectively, are imposed on the right edges ($x=5$) and the left-bottom edges ($x=0$ and $y=0$). The inner edges ($r=1$) and the top edges ($y=5$) are free of traction. The domain of the square plate is discretized using Tr4 elements for the curved boundary and linear Tr3 elements for other parts, which is shown in Fig. 16.

Table 8. Relative errors in displacement component *u* on the middle nodes of all curved edges.

Method / Mesh size	8×8	12×12	16×16	20×20	24×24
FEM-Tr3	9.304e-02	4.444e-02	2.511e-02	1.601e-02	1.106e-02
FEM-Tr3-4-N1	8.560e-02	4.127e-02	2.362e-02	1.518e-02	1.055e-02
FEM-Tr3-4-N2	8.167e-02	3.991e-02	2.309e-02	1.494e-02	1.043e-02
FEM-Q4	3.796e-02	1.525e-02	7.933e-03	4.828e-03	3.246e-03

Table 8 gives the relative errors in displacement with mesh sizes using different numerical methods on the middle points of all curved edges. It is observed that FEM-Tr3-4-N1 and FEM-Tr3-4-N2 models can slightly improve the accuracy in displacement, compared to

FEM-Tr3 model.

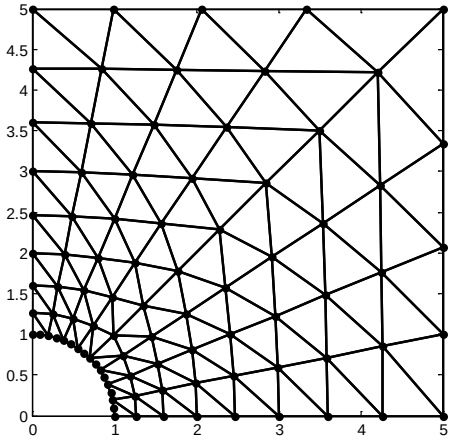
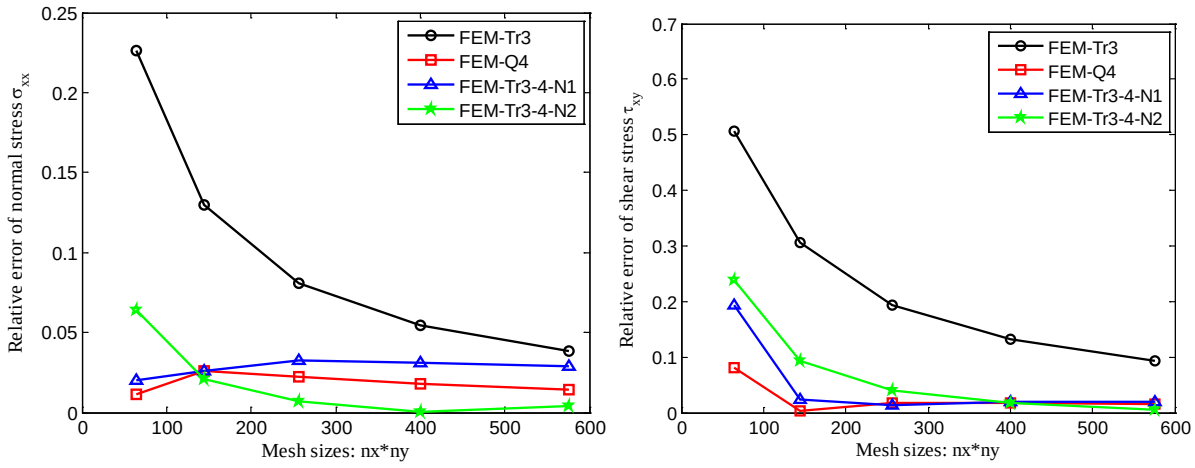


Fig. 16 Mesh with Tr3 elements and Tr4 elements for infinite plate with a circular hole.

According to the analytical solution in Eq.(32), the normal stress is maximum at the point A, and the shear stress is maximum at the point B. Fig. 17a shows relative errors in normal stresses σ_{xx} of middle-point A' (closest to Point A) on the boundary edges with mesh sizes using different methods. We find that the results of FEM-Tr3-4-N2 model are the most excellent in the terms of the normal stress σ_{xx} as the mesh size is greater to 144, comparing with the FEM-Tr3 and FEM-Q4 models.



(a) (b)

Fig. 17 Relative errors in stresses against Mesh sizes using different methods for infinite plate with a circular hole: (a) normal stress of point A'; (b) shear stress of Point B'.

Fig. 17b shows relative errors in shear stresses τ_{xy} of middle-point B' (closest to point B) with mesh sizes using different methods. we again find that FEM-Tr3-4 models stand out in the normal and shear stresses, comparing with FEM-Tr3 models, which also confirms that Tr4 elements effectively improve the accuracy of results, especially in stresses on the curved boundaries.

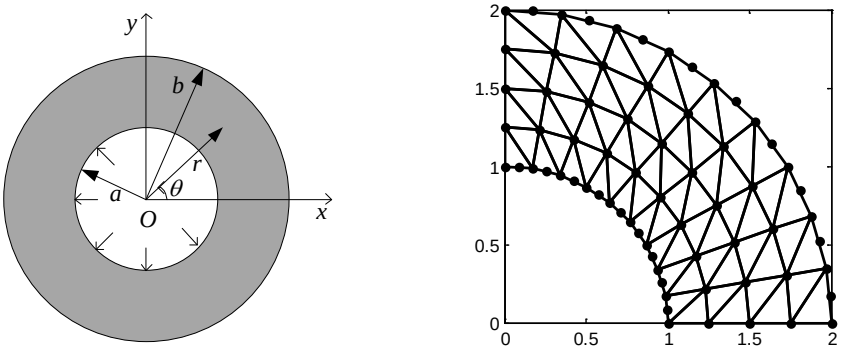
3.3 Axisymmetric ring subjected to internal pressure

Consider a 2D axisymmetric ring with inner radius *a*, outer radius *b* and subjected to internal pressure p_a, as shown in Fig. 18a. This problem is assumed to have a unit thickness so that

the plane stress condition is valid. The analytical solution of displacement and stresses of this problem are available in polar coordinate system [25]:

$$u_r = \frac{r}{E}\frac{a^2 q_a}{b^2 - a^2}\left(1-\nu+\frac{b^2}{r^2}(1+\mu)\right),$$
$$\sigma_r = \frac{a^2 q_a}{b^2 - a^2}\left(1-\frac{b^2}{r^2}\right), \sigma_\theta = \frac{a^2 q_a}{b^2 - a^2}\left(1+\frac{b^2}{r^2}\right), \tag{34}$$

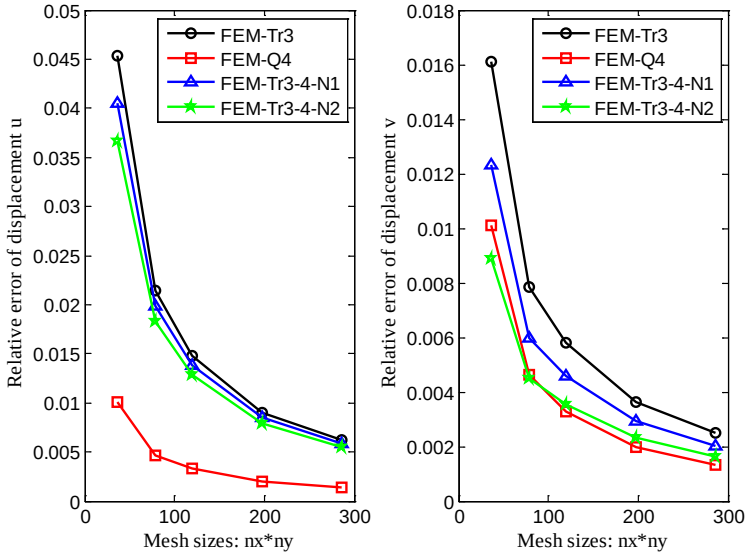
where (r,θ) are the polar coordinates and θ is measured counterclockwise from the positive *x*-axis.



(a) (b)

Fig. 18 (a) Axisymmetric ring subjected to inner pressure. (b) Domain discretization of the quarter ring.

Due to its symmetry, only one-fourth of the ring is modeled. Fig. 18b gives the discretization of the domain using Tr3 and Tr4 elements. The related parameters are taken as $E = 1.0\times10^3\,\mathrm{N/m^2}$, $\nu = 0.3$, $a = 1\mathrm{m}$, $b = 2\mathrm{m}$, $p_a = 100\mathrm{N/m}$. Symmetric conditions are imposed on the bottom and left edges, and the outer boundary of the ring is traction free.



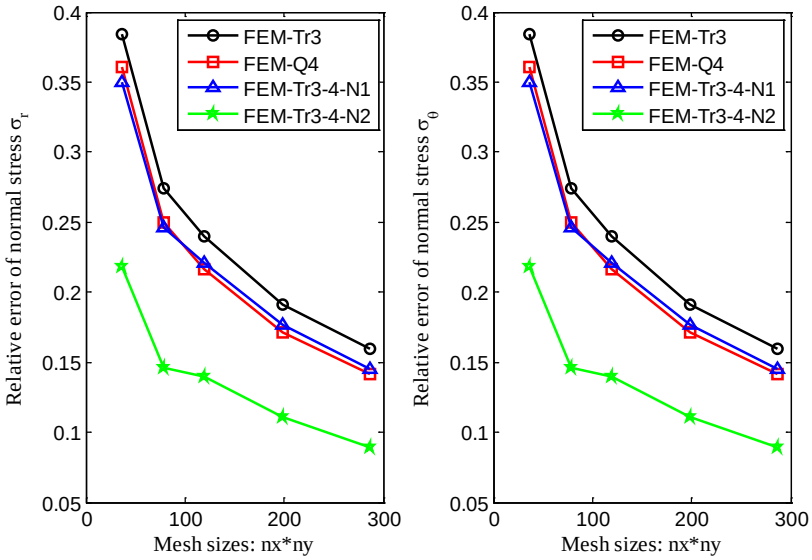
(a) (b)

Fig. 19 Relative errors in displacement components *u* and *v* against Mesh sizes on the inner curved boundary using different methods.

Fig. 19 presents the relative errors in displacement against mesh sizes using different

numerical methods on the middle points of inner curved edges, from which we observed again that FEM-Tr3-4 (FEM-Tr3-4-N1 and FEM-Tr3-4-N2) models improve slightly the accuracy in displacement component *u*, and improve distinctly the accuracy in displacement component *v*, compared to FEM-Tr3 model.

It is from Eq.(34) that maximum normal stresses (σ_r and σ_θ) in polar coordinate are on the inner curved boundary of ring. Fig. 20 shows relative errors in normal stress on inner curved edges against mesh sizes using different methods. We again find that FEM-Tr3-4 models is the most excellent, comparing with FEM-Tr3 models, and the results of FEM-Tr3-4-N2 model are the most accuracy in all models. This example again confirms Tr4 elements effectively improve the accuracy of results in stresses on the curved boundaries.



(a) (b)

Fig. 20 Relative errors in normal stresses against Mesh sizes on the inner curved boundary using different methods

3.4 A Rectangular Cantilever using Tr4 mesh

The Tr4 elements can be also used in entire problem shown in Fig. 4a. Now, consider the same cantilever problem as Example 3.1. The domain of the cantilever is discretized only using Tr4 elements, as shown in Fig. 21.

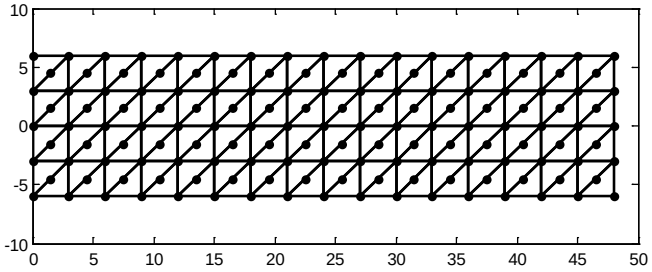


Fig. 21 The mesh using Tr4 elements for the rectangular cantilever.

Fig. 22 is the numerical results of strain energies against mesh sizes using different methods, which indicates that the strain energy solutions of FEM-Tr4 models can stably converge to the analytical strain energy from below. It is also clearly seen that the energy results of FEM-Tr4-N1 and FEM-Tr4-N2 models are more accurate, compared to FEM-Tr3 model and even FEM-Q4 model.

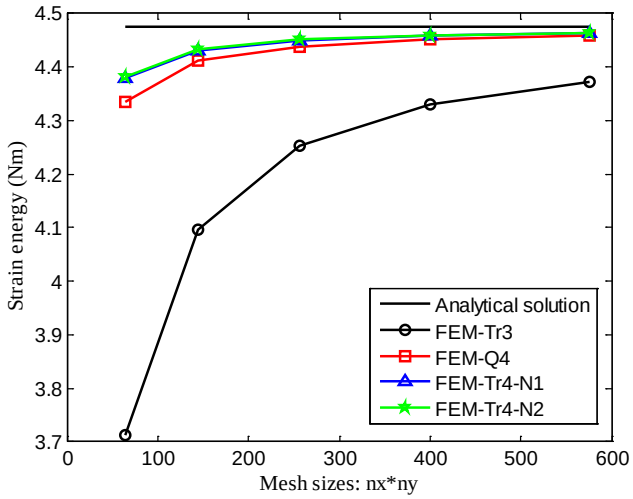
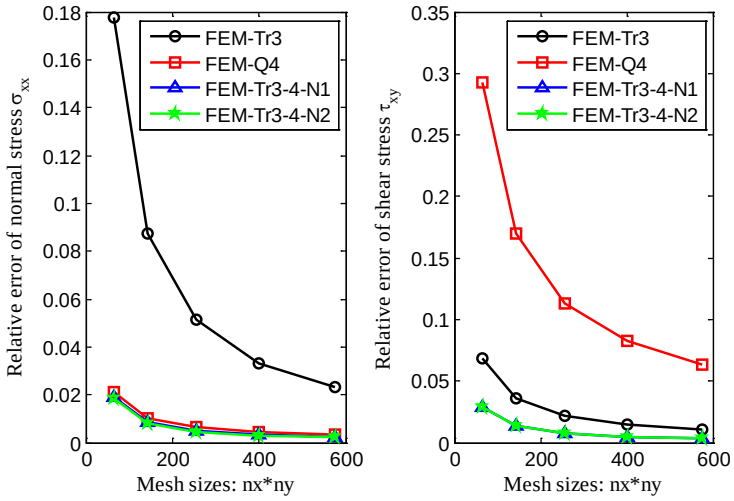


Fig. 22 Strain energies against mesh sizes using different methods for the rectangular cantilever.

Because the analytic solution of normal stress σ_{yy} in *y* direction is always zero, the relative errors in normal stress σ_{xx} and shear stress τ_{xy} are only considered, which are plotted in Fig. 23 against mesh sizes using different methods. From Fig. 23, we find the relative errors in stress of FEM-Tr4 models are obviously better than those of FEM-Tr3 model, and slightly better than those of FEM-Q4 model which is well known for its high accuracy. The proposed FEM-Tr4-N2 model can produce the most accurate solution among those models using the same mesh.



(a) (b)

Fig. 23 Relative errors in stress against mesh sizes for the rectangular cantilever using different methods: (a) normal stress σ_{xx}; (b) shear stress τ_{xy}.

4 conclusion

In order to using linear elements to approximate the curved boundaries of problem domain, we present a novel Tr4 element in the paper. . Based on two types of shape functions of the Tr4 element, we put forward to the FEM-Tr3-4-N1 and FEM-Tr3-4-N2 models. Comparing to the FEM-Tr3 model, the FEM-Tr3-4 models can obtain higher accuracy in stress solutions.

Especially the solution of FEM-Tr3-4-N2 model is more stable and high-accurate, which are even better than those obtained using the well-known FEM-Q4 model on the curved boundaries. Through intensive numerical examples, it is concluded that our novel Tr4 elements can model the curved boundaries efficiently and accurately.

Acknowledgments

In this paper, the work was supported by the NSF under the Award No. DMS-1214188 and the National Natural Science Foundation of China (Grant No. 11472184).

References

[1] Courant, R. (1943) Variational methods for the solution of problems of equilibrium and vibrations, *Bulletin of the American Mathematical Society*, **49**, 1-23.
[2] Liu, G. R. and Quek, S. S. (2003) *The Finite Element Method: A Practical Course*, 2nd edn, Butterworth Heinemann Ltd, Oxford.
[3] Liu, G. R., Xu, Y. G. and Wu, Z. P. (2001) Total solution for structural mechanics problems, *Computer Methods in Applied Mechanics & Engineering* **191**, 989-1012.
[4] Barsoum, R. S. (1976) On the use of isoparametric finite elements in linear fracture mechanics, *International Journal for Numerical Methods in Engineering* **10**, 25-38.
[5] Daily, J. W. and Harleman, D. R. F. (1966) *Fluid Dynamics*, Addison-Wesley Publishing Co., Reading, Mass.
[6] Gokul, K. C., Gurung, D. B. and Adhikary, P. R. (2013) FEM Approach for Transient Heat Transfer in Human Eye, *Applied Mathematics* **4**, 30-36.
[7] Hadda, H., Guessasma, M. and Fortin, J. (2014) Heat transfer by conduction using DEM-FEM coupling method, *Computational Materials Science* **81**, 339-347.
[8] Silvester, P. (1969) Finite-Element Solution of Homogeneous Waveguide Problems, *Alta Frequenza* **38**, 313-317.
[9] Zienkiewicz, O.C. and Taylor, R.L. (2000) *The Finite Element Method*, 5th edn, Butterworth Heinemann, Oxford.
[10] Bathe, K. J. (1996) *Finite element procedures*, Englewood Cliffs, NJ: Prentice-Hall.
[11] O'Rourke, J. (1998) *Computational Geometry in C*, 2nd edn, Cambridge Univ. Press.
[12] Si, W. and Zomaya, A. Y. (2011) New Memoryless Online Routing Algorithms for Delaunay Triangulations, *IEEE Transactions on Parallel & Distributed Systems* **23**, 1520-1527.
[13] Löhner, R. and Parikh, P. (1988) Generation of three-dimensional unstructured grids by the advancing-front method, *International Journal for Numerical Methods in Fluids* **8**, 1135-1149.
[14] Pirzadeh, S. (1993) Structured background grids of generation of unstructured grid by advancing front method, *Aiaa Journal* **31**, 257-265.
[15] Hoppe, V. (1980) High Order Polynomial Elements with Isoparametric Mapping, *International Journal for Numerical Methods in Engineering* **15**, 1747-1769.
[16] Yosibash, Z., Weissa, D. and Hartmannb, S. (2014) High-order FEMs for thermo-hyperelasticity at finite strains, *Computers and Mathematics with Applications* **67**, 477-496.
[17] Willberg, C., Duczek, S., Perez, J. M. V., Schmicker, D. and Gabbert, U. (2012) Comparison of different higher order finite element schemes for the simulation of Lamb waves, *Computer Methods in Applied Mechanics & Engineering* **241**, 246-261.
[18] Andersen, C. M. and Mcleod, R. J. Y. (1979) Integration Techniques for Isoparametri and Higher Order Bases on Finite Elements with a Curved Side, *Computers & Mathematics with Applications* **5**, 285-295.
[19] Naidu, V. K. and Nagaraja, K. V. (2013) Advantages of cubic arcs for approximating curved boundaries by subparametric transformations for some higher order triangular elements, *Applied Mathematics and Computation* **219**, 6893-6910.
[20] Eisenberg, M. A. and Malvern, L. E. (1973) On finite element integration in natural coordinates, *International Journal for Numerical Methods in Engineering* **7**, 574-575.
[21] Cowper, G. R. (1973) Gaussian quadrature formulas for triangles, *International Journal for Numerical Methods in Engineering* **7**, 405–408.
[22] Dunavant, D. A. (1985) High degree efficient symmetrical Gaussian quadrature rules for the triangle, *International Journal for Numerical Methods in Engineering* **21**, 1129-1148.
[23] Xiao, H., Gimbutasb, Z. (2010) A numerical algorithm for the construction of efficient quadrature rules in two and higher dimensions, *Computers and Mathematics with Applications* **59**, 663-676.
[24] Liu, G.R. and Trung, N. T. (2010) *Smoothed Finite Element Methods*. CRC Press, USA.
[25] Timoshenko, S.P. and Goodier, J.N. (1970) *Theory of Elasticity*, 3rd edn, McGraw-Hill, New York.

A Correspondence between Errors and Pseudo-errors of Approximate Computations with Similar Rates of Convergence

Aram Soroushian

Structural Engineering Research Centre, International Institute of Earthquake Engineering and Seismology, No. 21, West Arghavan, North Dibajee, S. Lavasani (Farmaniyeh), Tehran 19537, Iran

a.soroushian@iiees.ac.ir, aramsoro@yahoo.com

Abstract

A major requirement of arbitrary approximate computation is convergence to the exact solution. In view of the definition of error, problems with analytical closed from solutions generally need to be at hand, for numerical study of convergence, as well as verification and validation issues. In order to eliminate this necessity, the concepts of pseudo-error and pseudo-convergence are introduced in 2010. Towards better control of accuracy, it is herein explained that, if two solutions, obtained from two different methods, for one problem, converge to the exact solution with similar rates, and in intervals of proper convergence, and the errors of the first solution are more/less than the errors of the second solution, the same can be claimed about the two pseudo-errors, and vice versa. Extension to several computations is concluded and the implementation in practice is briefly discussed.

Keywords: Convergence, Exact solution, Error, Pseudo-error, Rate of convergence, Proper convergence

Introduction

Convergence is the most important essentiality of arbitrary approximate computations [1, 2]. With the purpose to study convergence and its trend numerically and without the exact solutions, the concepts of pseudo-error and pseudo-convergence are introduced and later extended to non-geometric changes of the algorithmic parameters [3-5]. In brief, for an arbitrary approximate computation $\mathbf{U}^a$ defined in terms of the algorithmic parameter λ, by defining the pseudo-error D, as (q stands for the rate of convergence [1-5], the right subscript i implies the result of the computation when using λ_i as the value of the algorithmic parameter, and $\| \ \|$ represents an arbitrary norm [6]):

$$D_i = \frac{\left\| \mathbf{U}^a{}_i - \mathbf{U}^a{}_{i-1} \right\|}{\left(\frac{\lambda_{i-1}}{\lambda_i} \right)^q - 1} \quad , \quad \lambda_i < \lambda_{i-1} \tag{1}$$

the equivalence addressed in Fig. 1 holds, i.e. either both the errors, E, and the pseudo-errors, D, or neither of them, converge properly. Proper convergence [7] is defined as the decrease of the errors in the log-log convergence plot as a straight line sloped q (q is defined just before

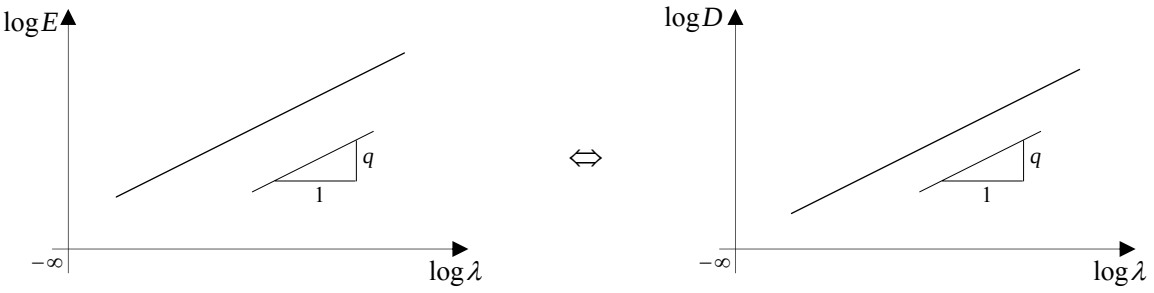


Figure 1. An existing equivalence between errors and pseudo-errors [4, 5]

Eq. (1)), and, considering **U** as the exact solution, the error, *E*, is defined as [8]:

$$E = \left\| \mathbf{U}^a - \mathbf{U} \right\| \tag{2}$$

In view of the academic need in an in-progress research, the main objective in this paper is to display that for two approximate computations for one problem converging properly [7] with identical rates, the computation with more/less error is associated with more/less pseudo-error, and vice versa; see Fig. 2; extension to several computations and even beyond the proper convergence region is considered as the conclusion.

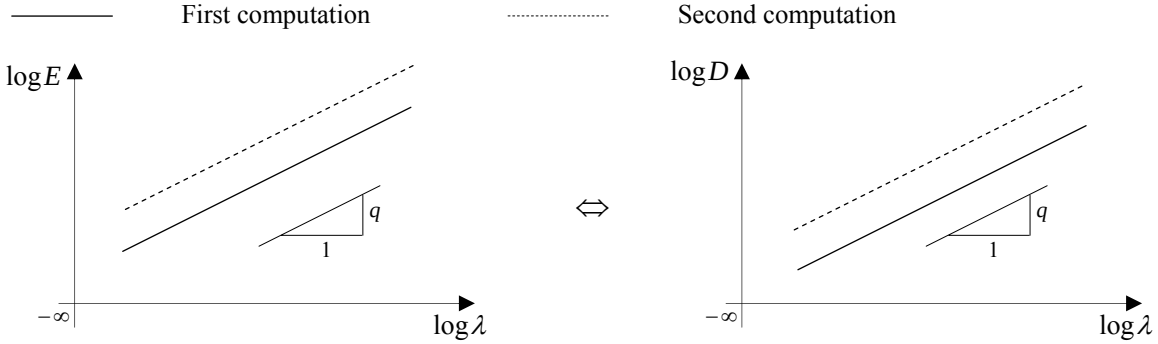


Figure 2. A new equivalence between convergence and pseudo-convergence

Theory

In order to explain the validity of the equivalence addressed in Fig. 2, consider three arbitrary different points on each of the two convergence trends in the left plot in Fig. 2, and address them such that

$$\lambda_1 > \lambda_2 > \lambda_3 > 0 \tag{3}$$

By addressing the two computations with U and $\overline{U}$, and addressing the information about the three points, with the right subscripts "1", "2", and "3, in view of the proper convergence assumption and from simple geometry,

$$\begin{aligned} &\text{For the first computation :} \quad \frac{\log E_1 - \log E_2}{\log \lambda_1 - \log \lambda_2} = \frac{\log E_2 - \log E_3}{\log \lambda_2 - \log \lambda_3} = q \\ &\text{For the second computation :} \quad \frac{\log \overline{E}_1 - \log \overline{E}_2}{\log \lambda_1 - \log \lambda_2} = \frac{\log \overline{E}_2 - \log \overline{E}_3}{\log \lambda_2 - \log \lambda_3} = q \end{aligned} \tag{4}$$

from which,

$$\frac{\log E_1 - \log E_2}{\log \overline{E}_1 - \log \overline{E}_2} = \frac{\log E_2 - \log E_3}{\log \overline{E}_2 - \log \overline{E}_3} = 1 \tag{5}$$

and hence,

$$\begin{aligned} \frac{E_1}{E_2} &= \frac{\overline{E}_1}{\overline{E}_2} \\ \frac{E_2}{E_3} &= \frac{\overline{E}_2}{\overline{E}_3} \end{aligned} \tag{6}$$

Consequently,

$$\frac{E_1}{\overline{E}_1} = \frac{E_2}{\overline{E}_2} = \frac{E_3}{\overline{E}_3} \tag{7}$$

which can be simply extended to (with more points in the left plot in Fig. 2):

$$\frac{E_1}{\overline{E}_1}=\frac{E_2}{\overline{E}_2}=\cdots=\frac{E_n}{\overline{E}_n}=\cdots=\alpha \quad , \quad n\in Z^+ ,\ \alpha\in R^+ \tag{8}$$

By considering different components of the solution separately, and taking into account that, for each component, the proper convergence occurs from one side [9], i.e. with values always more or always less than the exact value, Eq. (1) leads to

$$\begin{aligned} D_i &= \frac{\left|U^a{}_i - U^a{}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} \\ &= \frac{\left|\left(U^a{}_i - U\right) - \left(U^a{}_{i-1} - U\right)\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \frac{\left|\pm E_i \mp E_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \frac{\left|E_i - E_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} \end{aligned} \tag{9}$$

for arbitrary component of the solution. From Eqs. (8) and (9), we can conclude that

$$\begin{aligned} D_i &= \frac{\left|E_i - E_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} \\ &= \frac{\left|\alpha\overline{E}_i - \alpha\overline{E}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \alpha\frac{\left|\overline{E}_i - \overline{E}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} \end{aligned} \tag{10}$$

and since, according to the definition of pseudo-errors (see Eq. (1)), and with attention to [9],

$$\overline{D}_i = \frac{\left|\overline{U}^a{}_i - \overline{U}^a{}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \frac{\left|\left(\overline{U}^a{}_i - U\right) - \left(\overline{U}^a{}_{i-1} - U\right)\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \frac{\left|\pm \overline{E}_i \mp \overline{E}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} = \frac{\left|\overline{E}_i - \overline{E}_{i-1}\right|}{\left(\frac{\lambda_{i-1}}{\lambda_i}\right)^q - 1} \tag{11}$$

from Eqs. (8), (10) and (11), we can conclude

$$D_i = \alpha\overline{D}_i \quad \equiv \quad \frac{D_i}{\overline{D}_i} = \alpha , \qquad \alpha\in R^+ \tag{12}$$

Comparing Eqs. (8) and (12) leads to the fact that, when E_i is larger/smaller than $\overline{E}_i$, a similar relation exists between D_i and $\overline{D}_i$, and since this is for an arbitrary component of the solution, it can also be considered valid for the total solutions $\mathbf{U}^a$ and $\overline{\mathbf{U}}^a$ in the regions of proper convergence, i.e. we have succeeded to arrive at the right plot in Fig. 2 from the left plot.

In order to arrive at the left plot in Fig. 2 from the right plot, by considering *n* separate points in the right plot, we can arrive at

$$\frac{D_1}{\overline{D}_1}=\frac{D_2}{\overline{D}_2}=\cdots=\frac{D_n}{\overline{D}_n}=\cdots=\beta \quad , \quad n\in Z^+ ,\ \beta\in R^+ \tag{13}$$

Again restricting the discussion to arbitrary specific component of the solution and considering the fact that properly converging solutions not affected by round off (as displayed in Fig. 2) converge to the exact solution from one side [9], from Eqs. (1) and (13), we can conclude

$$\frac{U^a{}_0 - U^a{}_1}{\overline{U}^a{}_0 - \overline{U}^a{}_1} = \frac{U^a{}_1 - U^a{}_2}{\overline{U}^a{}_1 - \overline{U}^a{}_2} = \frac{U^a{}_2 - U^a{}_3}{\overline{U}^a{}_2 - \overline{U}^a{}_3} = \cdots = \frac{U^a{}_{n-1} - U^a{}_n}{\overline{U}^a{}_{n-1} - \overline{U}^a{}_n} = \cdots = \pm\beta \quad , \quad n \in Z^+ , \ \beta \in R^+ \tag{14}$$

and since theoretically, when disregarding round-off

$$\overline{U}^a{}_\infty = U^a{}_\infty = U \qquad \text{(the exact solution)} \tag{15}$$

from simple mathematics and Eq. (14), we will arrive at

$\forall\ m = 0,1,2,3,\ldots$

$$\frac{U^a{}_m - U}{\overline{U}^a{}_m - U} = \frac{\left(U^a{}_m - U^a{}_{m+1}\right) + \left(U^a{}_{m+1} - U^a{}_{m+2}\right) + \ldots}{\left(\overline{U}^a{}_m - \overline{U}^a{}_{m+1}\right) + \left(\overline{U}^a{}_{m+1} - \overline{U}^a{}_{m+2}\right) + \ldots} = \frac{\sum_{i=0}^{\infty}\left(U^a{}_{m+i} - U^a{}_{m+i+1}\right)}{\sum_{i=0}^{\infty}\left(\overline{U}^a{}_{m+i} - \overline{U}^a{}_{m+i+1}\right)} = \pm\beta \tag{16}$$

leading to

$$\frac{\left|U^a - U\right|}{\left|\overline{U}^a - U\right|} = \beta \tag{17}$$

or

$$\frac{E}{\overline{E}} = \beta \tag{18}$$

implying the completion of the proof for an arbitrary component of the solution and accordingly for the total solution in the proper convergence region. Thus, the proof is complete. As a direct consequence, when a problem can be solved approximately with several methods, and the computational methods provide similar rates of convergence and converge properly, the method with the i th size of error is also of the i th size of pseudo-error; see Fig. 3 for five methods; specifically the method displaying the least pseudo-error leads to the most accurate solution. Furthermore, from Eqs. (8), (12), (13), and (18), for any two of these computations, in the region of proper convergence,

$$\frac{E}{\overline{E}} = \frac{D}{\overline{D}} = \text{Const.} > 0 \tag{19}$$

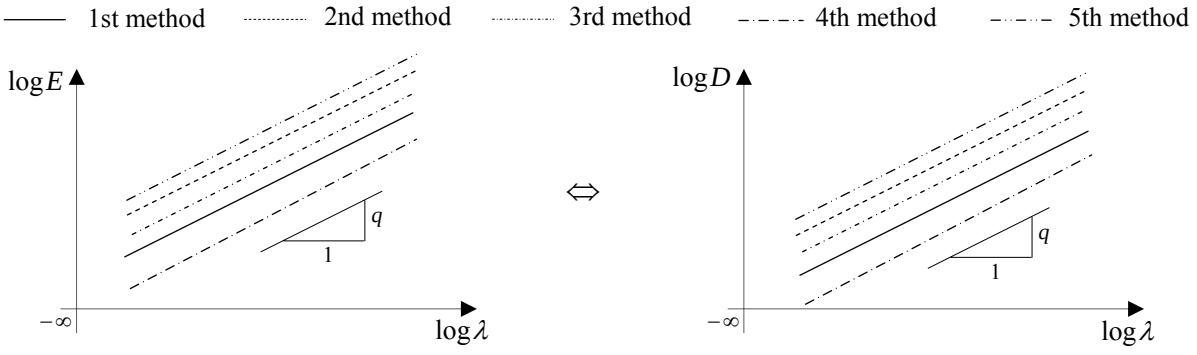


Figure 3. An extension of Fig. 2

And finally, since the basis of convergence plots and specifically the proper convergence regions is the Taylor series expansion [10] of approximate computations with respect to the algorithmic parameters [11], i.e.

$$\mathbf{U}^q = \mathbf{U} + \mathbf{C}_0\lambda^q + \mathbf{C}_1\lambda^{q+1} + \cdots \quad , \quad \mathbf{C}_0 \neq \overline{\mathbf{O}} \tag{20}$$

in view of the continuity of the Taylor series expansion addressed in Eq. (20) with respect to λ [10], it is also reasonable to expect the validity of the claim implied in Eq. (19) and Fig. 3 for values of the algorithmic parameter slightly larger than those corresponding to the proper convergence regions.

An Example

Since the previous section is carried out in a mostly rigorous manner, for the sake of brevity, only one example is presented here (and this example is also studied for other purposes in [12, 13]). Consider the shear frame structural system defined in Fig. 4 and Table 1 ($\ddot{u}_g$ stands for the ground acceleration and $_f\Delta t$ implies the size by which $\ddot{u}_g$ is digitized). Transient analysis of the structural behavior by direct time integration [14-17] is the approximate computation under consideration. The time integration methods are the Houbolt, the average acceleration Newmark, the C-H ($\rho_\infty = 0.8$), and the C-H ($\rho_\infty = 0.5$) methods [18-22], all providing second order convergence, i.e. $q = 2$ [14, 15, 23]. The peak lateral displacement of the top floor and the final floors shear forces are the target solutions. In the study of the convergence trend, the integration steps are the algorithmic parameters [11, 14, 15, 24, 25], and as conventional [17, 26, 27], the steps of direct time integration are halved sequentially,

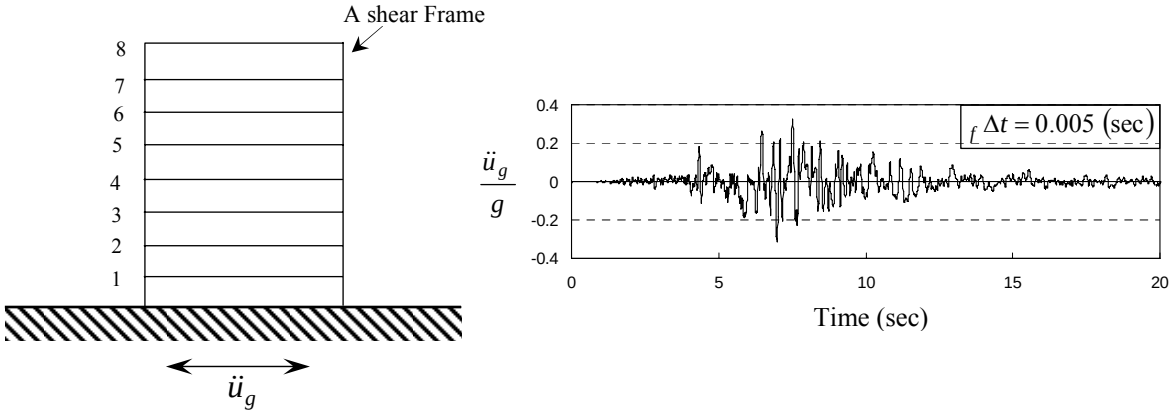


Figure 4: Structural system under consideration: (a) Structural model, (b) Excitation

Table 1: Some properties of the shear frame in Fig. 4(a)

Storey	Mass (Kg)	Stiffness (N/m^2)	Damping (N/m/sec)
1	1036E4	860E7	0
2	1034E4	840E7	
3	1032E4	820E7	
4	1030E4	700E7	
5	1028E4	680E7	
6	1026E4	660E7	
7	1024E4	640E7	
8	1022E4	620E7	

while in the first analysis (computation),

$$\Delta t = \lambda =_f \Delta t = 0.005 \quad \text{(sec)} \tag{21}$$

In a special time integration analysis (computation), carried out with the purpose to determine the errors with high precision, the steps of the direct time integration are considered equal to the very small values stated below:

$$\Delta t = \lambda = 0.005\left(2^{-12}\right) \quad \text{(sec)} \tag{22}$$

The convergence and pseudo-convergence plots are depicted in Figs. 5 and 6, while for the second target solution, i.e. the floors final shear forces, the L_2 norm is implemented for computing the errors and pseudo-errors. Figs. 5 and 6 clearly display the validity of the claims discussed in the previous section, i.e. (1) larger/smaller errors imply larger/smaller pseudo-errors, in the proper convergence regions, and vice versa, (2) validity of Eq. (19) in the proper convergence regions, (3) possibility to extend the previous two points to values of the algorithmic parameter slightly larger than those corresponding to the proper convergence regions.

Several other examples concentrating on different approximate computations, including structural analysis by finite elements, nonlinearity solutions, and different ways of computing π are under study.

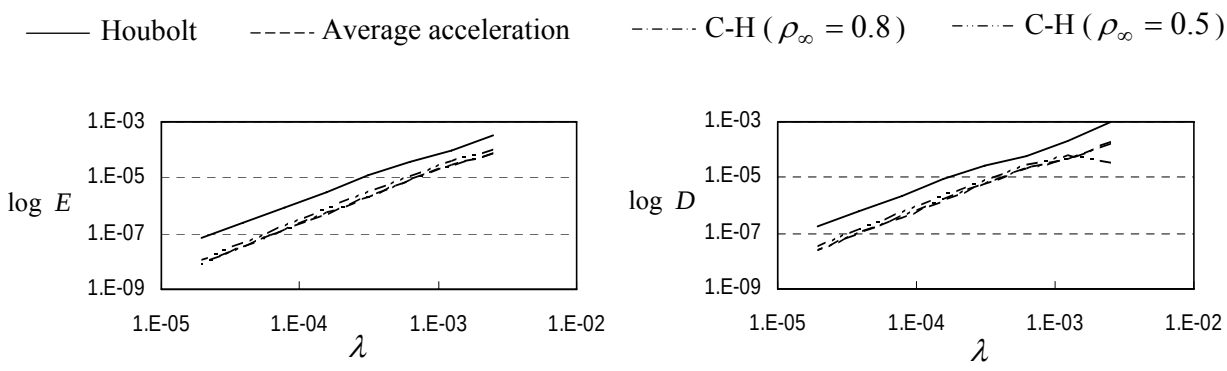


Figure 5. Convergence and pseudo-convergence plots for the peak top displacements

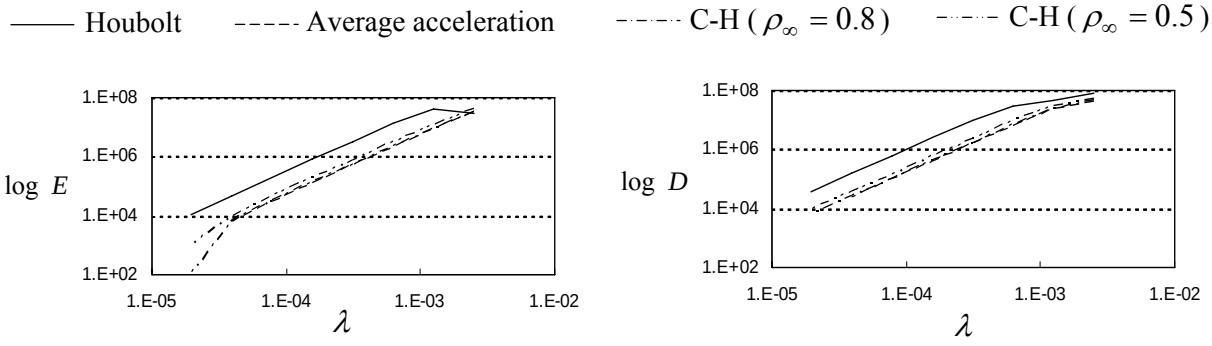


Figure 6. Convergence and pseudo-convergence plots for the final floor shear forces with the L_2 norm

Discussion

Recognition of the most accurate computation from several approximate computations, though of high importance, is generally not an easy task. From the other side of view, arriving at solutions very close to the exact solutions, in order to lead to the computational errors, is computationally expensive. Even the existing error estimations are not reliable in many cases. In these cases, the discussions presented in this paper can be significantly effective.

In this section, some complementary explanation is stated about proper convergence, how to check proper convergence with small computational effort, and meanwhile extension of the discussion to vector and matrix solutions. The equivalence addressed in Fig. 1 is a simple way to check the proper convergence (a simpler way based on purification of convergence [28] is yet not finalized). Nevertheless, for locating each point in the pseudo convergence plot two approximate computations should be carried out and for checking the proper convergence, at least two points should be located. This means three computations, and sounds entailing considerable additional computational cost, especially when taking into account that approximate computations with smaller algorithmic parameters are more costly. This is however not correct. For many approximate computations, e.g. solution of ordinary differential equations, finite element analysis, time integration analysis and nonlinear time history analysis against seismic excitations [17, 26, 27, 29], repetition of the computation after assigning smaller (mostly half) values to the algorithmic parameters (and even repetition of the computation by times) is strongly recommended and in cases considered as an obligatory requirement; see [27]. Therefore, the above-mentioned additional computational effort at most corresponds to one computation, and even the additional cost can be obviated by different approaches, from which, two are (also see [30]): (1) assigning slightly larger values to the algorithmic parameter in the first computation, and (2) while repeating the computation, considering smaller values for r in

$$\frac{\lambda_1}{\lambda_2} = \frac{\lambda_2}{\lambda_3} = \cdots = r \qquad , \qquad r > 1 \tag{23}$$

Consequently, even in cases that additional computational cost basically exists, for practical implementation of the achievements, the additional costs can mostly be hesitated or lessened.

Finally, for non-scalar solutions, the requirement of proper convergence for all components of the solution can be merely sufficient, not necessary (based on the norm). Components of the solution with small contribution in the error/pseudo-error need less to converge properly.

Conclusions

In this paper, it is displayed via theoretical discussion and an example, that for arbitrary approximate computation, solution, and problem:

(1) From several solutions converging properly to the exact solution with similar rates, the most/least accurate solution converges with least/most pseudo-errors, and vice versa.
(2) The ratio of the errors of a computation to the errors of another computation when both computations converge properly and with the same rate remains unchanged, if instead of the errors we compare the pseudo-errors.
(3) The above two points persist for values of the algorithmic parameter slightly larger than those corresponding to the proper convergence regions.

Considering the unavailability of exact solutions and the high computational effort associated with highly precise solutions, implementation of the achievements in practice is briefly discussed; more investigation is essential and strongly recommended.

References

[1] Henrici, P. (1962) *Discrete Variable Methods in Ordinary Differential Equations*, Prentice-Hall, USA.
[2] Srikwerda, J. C. (1989) *Finite Difference Schemes and Partial Differential Equations*, Wadsworth & Books/Cole, USA.
[3] Soroushian, A., *New methods to maintain responses 'convergence and control responses' errors in the analysis of nonlinear dynamic models of structural systems*, PhD Thesis, University of Tehran, Iran, 2003 (in Persian).
[4] Soroushian, A., Pseudo convergence and its implementation in engineering approximate computations, *Proceedings of 4th International Conference from Scientific Computing to Computational Engineering (IC-SCCE 2010)*, Athens, Greece, July 7-10, 2010.
[5] Soroushian, A., Equivalence between convergence and pseudo convergence when algorithmic parameters do not change geometrically, *Proceedings of 6th International Conference from Scientific Computing to Computational Engineering (IC-SCCE 2014)*, Athens, Greece, July 9-12, 2014.
[6] Noble, B., Daniel, J. W. (1977) *Applied Linear Algebra*, Prentice Hall, USA.
[7] Soroushian, A., Proper convergence a concept new in science and important in engineering, *Proceedings of 4th International Conference from Scientific Computing to Computational Engineering (IC-SCCE 2010)*, Athens, Greece, July 7-10, 2010.
[8] Ralston, A., Rabinowitz, P. (1978) *First Course in Numerical Analysis*, McGraw-Hill, USA.
[9] Soroushian, A., Wriggers, P., and Farjoodi, J., A statement for the convergence of approximate responses and its application in structural dynamics, *Proceedings of 2nd International Conference from Scientific Computing to Computational Engineering (IC-SCCE 2006)*, Athens, Greece, July 5-8, 2006.
[10] Apostol, T. M. (1967) *Calculus*, *Vol. I*., Wiley, USA.
[11] Soroushian, A., Wriggers, P., Farjoodi, J. (2009) Asymptotic upper bounds for the errors of Richardson extrapolation with practical application in approximate computations, Int J Numer Meth Eng **80**, 565-595.
[12] Soroushian, A., Transient analysis with steps larger than the excitation steps independent from the frequency contents, *Proceedings of 6th ECCOMAS Thematic Conference on Computational Methods in Earthquake Engineering and Structural Dynamics (COMPDYN 2017)*, Island of Rhodes, Greece, June 15-17, 2017.
[13] Soroushian, A., Zarabimanesh, Y., Soleymani, K., Khalkhali, S. M., A new technique for fractional enlargement of integration steps in transient analysis against digitized excitations, *Proceedings of the International Conference on Structural Engineering Dynamics (ICEDyn 2017)*, Ericeira, Portugal, July 3-5, 2017.
[14] Belytschko, T., Hughes, T. J. R. (1983) *Computational Methods for Transient Analysis*, Elsevier, 1983.
[15] Wood, W. L. (1990) *Practical Time Stepping Schemes*, Oxford, 1990.
[16] Paultre, P. (2010) *Dynamics of Structures*, Wiley, USA.
[17] Clough, R. W., Penzien, J. (1993) *Dynamics of Structures*, McGraw-Hill, 1993.
[18] Houbolt, J. C. (1950) A recurrence matrix solution for the dynamic response of elastic aircraft, J Aeronaut Sci **17**, 540–550.
[19] Katona, M. G., Zienkiewicz, O. C. (1985) A unified set of single step algorithms part 3: the beta-m method, a generalization of the Newmark scheme, Int J Numer Meth Eng **21**, 1345–1359.
[20] Soroushian, A., Farjoodi, J. (2008) A unified starting procedure for the Houbolt method, Commun Numer Meth Eng **24**, 1–13.
[21] Newmark, N. M. (1959) A method of computation for structural dynamics, J Eng Mech-ASCE **85**, 67–94.
[22] Chung, J., Hulbert, G. M. (1993) A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-a method, J Appl Mech-T ASME **60**, 371–375.
[23] Hughes, T. J. R. (1987) *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Prentice-Hall, USA.
[24] Soroushian, A. (2017) Integration step size and its adequate selection in analysis of structural systems against earthquakes. In Eds. M. Papadrakakis, V. Plevris, N.D. Lagaros, eds., *Computational Methods in Earthquake Engineering Vol 3*, Springer, Norway.
[25] Soroushian, A. (2008) A technique for time integration with steps larger than the excitation steps, Commun in Numer Meth Eng **24**, 2087-2111.
[26] Hairer, E., Wanner, G. (1996) *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, Springer, USA.
[27] Standards New Zealand (2004) *Structural Design Actions Part 5: Earthquake Actions–New Zealand Commentary (Supplement to NZS 1170.5:2004) (2004) NZS 1170.5 Supp 1*, New Zealand.
[28] Soroushian, A., Purification of convergence an approach towards reliable error evaluation, *Proceedings of 11th World Congress on Computational Mechanics (WCCM XI)*, Barcelona, Spain, July 20-25, 2014.
[29] Fish, J., Belytschko, Y. (2009) *A First Course in Finite Elements*, Wiley, UK.
[30] Soroushian, A., A simple approach towards further accuracy in structural dynamic analysis, *Proceedings of 4th ECCOMAS Thematic Conference on Computational Methods in Structural Dynamics and Earthquake Engineering (COMPDYN 2013)*, Kos Island, Greece, June 12-14, 2013.

Carotid Atherosclerotic Plaque Vulnerability Assessment Using 3D Thin-Layer Model with In Vivo Patient-Specific Vessel Material Properties

***Qingyu Wang,[1] Gador Canton,[2] Thomas S. Hatsukami,[3] Kristen L Billiar,[4] Chun Yuan,[5] Zheyang Wu,[6] †Dalin Tang[1,6]**

[1] School of Biological Science and Medical Engineering, Southeast University, Nanjing, 201196, China
[2] Department of Mechanical Engineering, University of Washington, Seattle, WA 98195
[3] Division of Vascular Surgery, University of Washington, Seattle, WA. 98195
[4] Biomedical Engineering Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA
[5] Department of Radiology, University of Washington, Seattle, WA 98195
[6] Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA

* Presenting author: Qingyu Wang, Southeast University, Nanjing China, email: 1263800783@qq.com.
†Corresponding author: Dalin Tang, Southeast University, Nanjing China; Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609, email: dtang@wpi.edu

Abstract

Patient-specific vessel material properties are in general lacking in image-based computational models, limiting the accuracy of their stress/strain calculations. A noninvasive approach of combining in vivo 3D multi-contrast and Cine magnetic resonance imaging (MRI) and computational modeling was introduced to quantify patient-specific carotid plaque material properties for potential plaque model improvements. Carotid plaque stress and strain conditions with in vivo and old material model were investigated. A computational plaque stress index (CPVI) was proposed to combine mechanical analysis, plaque morphology and compositions for more complete carotid plaque vulnerability assessment.

In vivo 3D multi-contrast and Cine MRI carotid plaque data were acquired from 8 patients with follow-up (18 months) with written informed consent obtained. 3D thin-layer model and an established iterative procedure were used to determine parameter values in the Mooney-Rivlin models for the 81slices from 16 plaque samples. Effective Young's Modulus (YM) values were calculated for comparison and analysis.

The average YM, circumferential shrink (C-Shrink) and lumen circumference variation measure by Cine MRI of the 81 slices was 411 kPa, 5.62%, and 8.91%, respectively. Average YM values by vessel varied from 109 kPa (softest) to 922 kPa (stiffest), a 746% difference. The average absolute variation of average stress values from 16 carotid plaques were 16.42% between in vivo material model and old material model, where 8 cases had in vivo material model stress greater than old material model stress and 8 cases had old material model stress greater than in vivo material model stress. The range of absolute variation values was [0.29%, 30.98%]. The average absolute variation of average strain values from 16 carotid plaques were 71.99% between in vivo material model and old material model, where 9 cases had in vivo material model strain greater than old material model strain and 7 cases had old material model strain greater than in vivo material model strain. The range of absolute variation values was [2.82%, 377.34%]. YM values showed positive correlation with Max stress ($r=0.3531$, $p=0.1797$) and critical stress ($r=0.5733$, $p=0.0202$). YM values also showed strong negative correlation with Max strain ($r=-0.8246$, $p<0.0001$) and critical strain ($r=-0.7376$, $p=0.0011$). The five intervals (unit: kPa) [0, 46.8), [46.8,

80), [80, 92), [92, 103), and [103, +∞) from in vivo material models were used for CPVI values of 0, 1, 2, 3 and 4, respectively. The optimized agreement rate was 85.19%.

In vivo carotid vessel material properties have large variations from patient to patient. The use of patient-specific material properties in plaque models could potentially improve the accuracy of model stress/strain calculations. Large-scale studies are needed to further demonstrate that CPVI has the potential to improve the current image-based screening and plaque vulnerability assessment schemes.

Keyword: Atherosclerotic plaque; magnetic resonance imaging (MRI); material properties; stress/strain calculation; carotid artery modeling.

1. Introduction

Cardiovascular diseases are the major cause of death in the world [1]. Atherosclerotic plaques may rupture without warning and cause fatal clinical events such as heart attack and stroke. Mechanisms for plaque progression and rupture are not well understood. It has been accepted that mechanical forces may play an important role in plaque rupture process and should be considered in an integrated way with plaque morphology and composition for possible improvement of plaque assessment schemes [2]. Currently, screening and diagnosis of patients with atherosclerotic plaques are based on medical images such as magnetic resonance image (MRI), ultrasound, intravascular ultrasound (IVUS), computerized tomography (CT), or optical coherence tomography (OCT) . Increasing evidences showed that such medical imaging technologies are not enough to identify those victims before the event occurs [3]-[4]. It has been hypothesized that image-based plaque models and mechanical stress and strain conditions may be useful for more accurate plaque vulnerability assessment and prediction of future clinical events.

In recent years, MRI techniques have shown great potential to non-invasively quantify plaque size, shape and components (fibrous cap, lipid-rich necrotic core and calcification/inflammation) [5]. Yuan et al. developed multi-contrast techniques to improve the quality of MR-images and to better differentiate various components of the plaque [6]-[7]. With the advances of medical imaging technologies [6]-[9], image-based computational models have been introduced to calculate plaque stress/strain conditions and investigate their association with plaque progression and rupture [10]-[21]. However, the accuracy of the computational results is heavily dependent on the data and assumptions used by those models. Data needed for image-based plaque computational models include: a) plaque morphology and components; b) vessel and plaque component material properties; and c) blood flow and pressure conditions [16]. While many image-based models used patient-specific plaque morphology data, patient-specific vessel material properties are lacking in those models [10]-[24]. Non-invasive techniques to obtain in vivo patient-specific vessel material properties are needed to further improve in vivo image-based plaque models [25]-[27].

Considerable efforts have been made by several research groups to quantify mechanical material properties of atherosclerotic vessels. Smoljkić et al. proposed a non-invasive, energy-based assessment of patient-specific material properties of arterial tissue [26]. Their results showed that imposing conditions on strain energy can provide a good estimation of carotid material properties from the non-invasively measured pressure and diameter data. Czernuszewicz et al. performed some preliminary study of non-invasive in vivo characterization of human carotid plaques with acoustic radiation force impulse ultrasound. Their method was able to differentiate soft tissues from

stiffer tissues with histological validations [27]. Liu et al. introduced a non-invasive approach to quantify patient-specific vessel material properties and plaque circumferential shrinkage rate between in vivo and "no-load" vessel geometries [28]. Their material properties and circumferential shrinkage rate were calculated by 2D plaque models. Their results showed that effective Young's Modulus (YM) from the 12 human carotid arteries varied from 137 kPa to 1435 kPa and vessel circumferential shrinkage to "no-load" condition varied from 6% to 32%. Overall, quantified patient-specific shrinkage rate using in vivo data are rare in the current literature.

The foundation for plaque classifications was established by Stary et al. in a series of American Heart Association (AHA) committee reports on vascular lesions of Council on Arteriosclerosis which provided a histological lesion classification scheme [29]-[31]. Among the AHA Type I – VIII lesions, Type I-IV are considered stable (I-III) or minimally unstable (IV). Types V (lipid-rich), VI (complex), VII (calcified), and VIII (fibrotic) are the advanced plaques capable of rupture. Using non-invasive MRI techniques, Cai et al. developed a classification system for carotid plaques based on in vivo MRI [32]. Tang et al. introduced a "local maximum stress hypothesis" to identify the critical site and stress conditions in the plaque and proposed an ex vivo MRI-based computation plaque vulnerability index (CPVI) to access plaque vulnerability [14][33]-[34].

In this paper, a non-invasive approach [28] of combining 3D multi-contrast MRI, in vivo Cine MRI and computational 3D thin-layer model [35] was used to quantify patient-specific carotid plaque material properties and circumferential shrinkage rates. A thin-layer modeling method was used to quantify the impact of patient-specific vessel material properties on stress and strain predictions [17]. These methods and results will fill a gap in the current literature. For plaque vulnerability assessment, a stress-based computational plaque vulnerability index (CPVI) was introduced and its value was assigned for all 81 slices using stress values from models with patient-specific material data and old material data. The resulting CPVI classifications were compared with a morphological plaque severity index (MPSI) which was defined based on plaque morphological characteristics known to correlate with plaque vulnerability [32].

2. Methods

2.1 MRI Data Acquisition.

Serial MRI data of carotid atherosclerotic plaques from 8 patients (5 male, 3 female; age: 62–83, mean=71) were acquired at the University of Washington (UW), Seattle by the Vascular Imaging Laboratory (VIL) using protocols approved by the UW Institutional Review Board and with written informed consent obtained. For each patient, MRI slices at baseline (Time 1, T1) and follow-up (Time 2, T2, Scan time intervals were about 18 months) were matched up using vessel bifurcation, stenosis features and with careful review by the MRI group. Cuff systolic and diastolic arm pressure was recorded for modeling use. In vivo Cine and 3D multi-contrast MR images of the carotid arteries were acquired using a 3.0T whole-body scanner (Philips Achieva, R2.6.1, Best, The Netherlands) and a dedicated 8-channel, phased array carotid coil. The carotid bifurcation was located on 2D TOF (Time of Flight) and oblique black blood MR images. A 3.5cm region centered on the carotid bifurcation was imaged by high-resolution axial bright and black blood imaging. Detailed data acquisition and segmentation procedures were published before and are omitted here [11][28]. For each patient, locations with Cine sequence and nearly-circular lumen cross-section were selected for calculating the material parameter values in the modified Mooney-Rivlin model [11][28]. Figure 1 gives 5 selected MRI slices with segmented contour plots of the plaque.

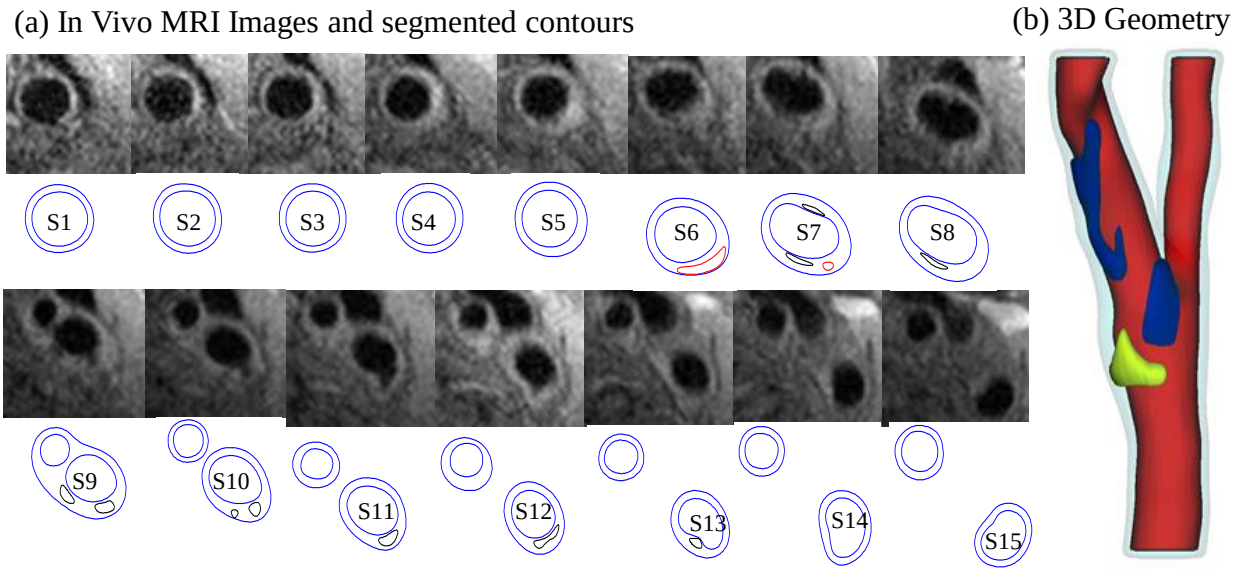


Figure 1. A plaque example showing MRI slices, segmented contours, and re-constructed geometry. (a) In vivo MR-images and segmented contour plots showing plaque components (Blue: lumen, vessel wall; black: calcification; red: lipid core;); (b): 3D reconstructed geometry (Red: lumen; light blue: vessel; blue: calcification; yellow: lipid core).

2.2 Computational Models, Mesh Generation and Solution Methods.

A 3D thin-layer modeling approach introduced by Huang et al. [35] was used to determine material parameter values in our selected material model. For every slice that Cine data was available, a thin slice thickness (0.5 mm) was added to make a 3D thin-layer model (Figure 2). The carotid artery was assumed to be hyperelastic, isotropic, incompressible and homogeneous. The nonlinear modified Mooney-Rivlin (M-R) model was selected to describe the material properties of the vessel wall [36][37]. The strain energy function was given by:

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1\left[\exp\left(D_2(I_1 - 3)\right) - 1\right], \quad (1)$$

$$I_1 = \sum C_{ii}\,,\ I_2 = \frac{1}{2}\left[I_1^2 - C_{ij}C_{ij}\right], \quad (2)$$

where $C = [C_{ij}] = X^T X$ is the right Cauchy-Green deformation tensor; I_1 and I_2 are the invariants of C; $X = [X_{ij}] = \left[\frac{\partial x_i}{\partial a_j}\right]$ is the deformation gradient; c_1, c_2, D_1 and D_2 form the material parameter set. The modified Mooney-Rivlin model was selected because it was able to fit carotid artery vessel properties measured by uniaxial and biaxial mechanical testing data and good agreement was obtained [38]. According to our previous literatures [11][17], material parameters c_i and D_i (i=1,2) were chosen to match experimental measurements: old vessel material/fibrous cap, c_1=36.8 kPa, D_1=14.4 kPa, D_2=2; lipid core/hemorrhage, c_1=2 kPa, D_1=2 kPa, D_2=1.5; calcification, c_1=368 kPa, D_1=144 kPa, D_2=2.0; loose matrix, c_1=18.4 kPa, D_1=7.2 kPa; D_2=1.5. c_2 = 0 for all materials [17].

(a) Created Lines connecting data points and dividing lines for a slice.

(b) Add slice thickness to make 3D thin-wall model

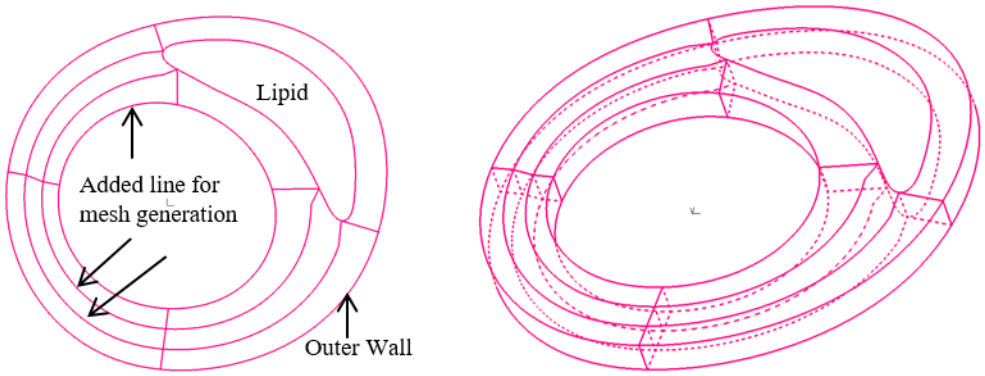


Figure 2. Thin-layer model construction.

For each 3D thin-layer model, a 10% axial shrinkage rate was applied. Then an iterative procedure [28] was followed to adjust the parameter values in the modified M-R model and the circumferential shrinkage rate to match both maximum and minimum Cine lumen circumferences corresponding to systolic and diastolic pressures. The details of the iteration procedure were described in our previous paper [28]. The 3D thin-layer model for each iteration was solved by ADINA (ADINA R & D, Watertown, MA). The stress-stretch relationship for the Mooney-Rivlin model is given by:

$$\sigma = \lambda\frac{dW}{d\lambda} = 2\lambda(\lambda - \lambda^{-2})\left(c_1 + D_1 D_2 e^{D_2\left(\lambda^2+\frac{2}{\lambda}-3\right)}\right), \quad (3)$$

where σ is Cauchy stress, and λ is stretch ratio. In order to facilitate comparison, it is easier to use a single parameter to compare vessel stiffness from different patients or slices. The effective Young's modulus (YM) E for the stretch ratio interval [1.0, 1.3] is defined as:

$$\sigma = E(\lambda - 1), \quad (4)$$

The least-squares technique was used to calculate the YM values that best fit the M-R model.

2.3 Definition and Calculation of Critical Stress.

It is known that thin plaque cap is closely related to plaque rupture. Thus all locations where a thin region covers a plaque component were considered as candidate critical sites. It should be noted that our "thin region" includes fibrous cap over a lipid core, as well as "cap" over calcification and other plaque components. The site with the maximum Stress (maximum principal stress) value among all the candidate sites was defined as the critical site, and the stress value at this site was defined as the critical stress [33][34]. For slices without any components, critical stress was defined as zero since these slices are very stable.

2.4 Assignment of Morphological Plaque Severity Index (MPSI).

Since histological data is in general not available for in vivo studies, a morphological plaque severity index (MPSI) was introduced (Table 1) and assigned to each segmented MRI slice based on plaque morphological features known to correlate with plaque vulnerability from histopathological studies [39]-[41]. These features include: 1) the size and distribution of the soft lipid rich necrotic core (LRNC); 2) the fibrous cap thickness (which correlates with plaque

stability); and 3) the presence of ulcer, intraplaque hemorrhage and thrombi. MPSI values (0, 1, 2, 3 to 4) indicate the level of increasing severity. The MPSI definitions are closely associated with the AHA (American Heart Association) lesion type classifications (see Table 1). Figure 3 presents 5 representative slices with above described morphological characteristics.

Table 1. Morphological plaque vulnerability index (MPSI) classifications and comparison with AHA classifications.

MPSI Category	Corresponding AHA lesion types (modified)	Description	Level of vulnerability
0	I or II	Normal or nearly normal wall.	Very stable
1	III	Moderate intimal thickening, no extracellular lipid, calcification or significant inflammation.	Stable
2	IV/V with less than 30% NC by area; or VII; or VIII	Advanced lesion with small necrotic core (<30% of plaque size), or can be fibrotic or calcified, thick fibrous cap (> 200 μm).	Slightly unstable
3	IV/V with 30-40% NC by area	Advanced lesion with Moderate lipid core (30-40% of plaque size) and fibrous cap (150-200μm).	Moderately unstable
4	IV/V with > 40% NC by area; or VI	Advanced lesion with a very large necrotic core (>40%), thin fibrous cap (<150 μm), or with fibrous cap rupture, ulceration, or intraplaque hemorrhage.	Very unstable

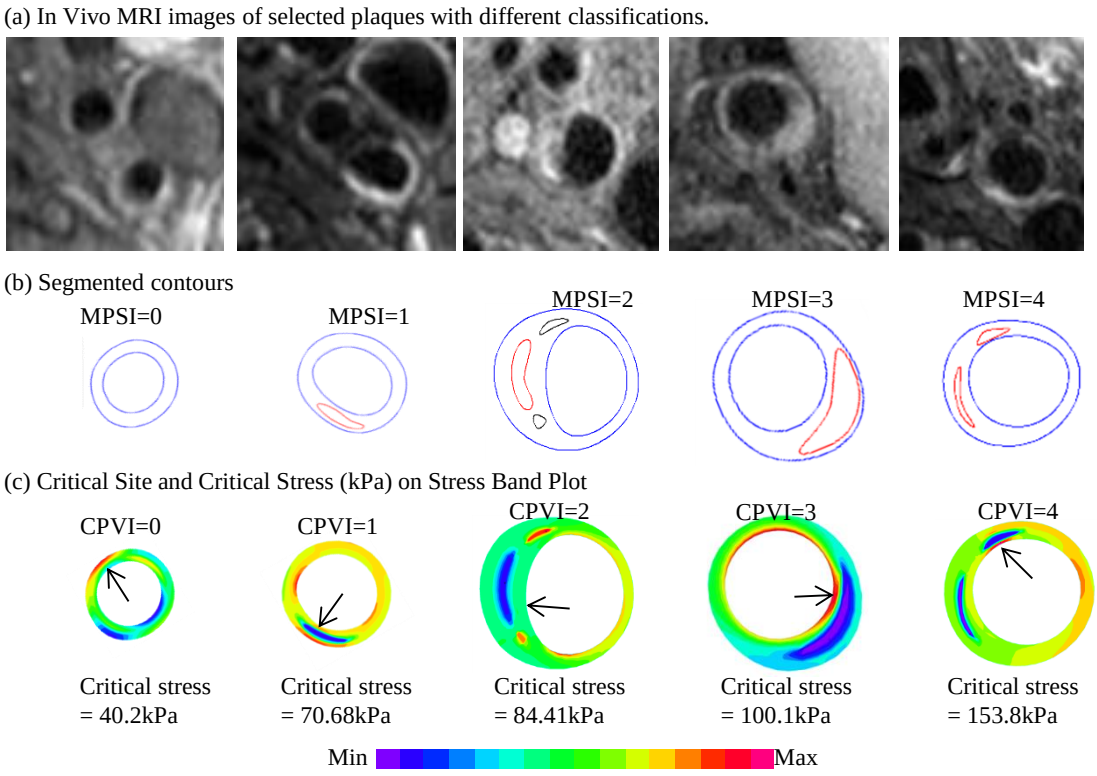


Figure 3. Plaque samples showing morphological features and critical stress values for plaque classifications. (a) In vivo MR-images; (b) segmented contour plots showing plaque components(Black: calcification; Red: Lipid core); (c) stress plots showing critical stress of corresponding slices; CPVI values were calculated based on critical stress values at critical sites.

2.5 CPVI Assignment and Data Analysis.

Each slice was assigned a CPVI value (0, 1, 2, 3 or 4) according to its critical stress by using five stress intervals, which were determined to have best match rate with MPSI. Correlations between CPVI values and plaque morphological features including lipid core size, cap thickness and normalized wall index were analyzed. Average stress/strain and critical stress/strain on the lumen and all the cap nodes covering the lipid-rich pool were recorded for comparison. To be clear, since stress and strain are tensors, maximum principal stress and maximum principal strain were taken as the scalar representatives of stress and strain to present our results, respectively. Correlation studies were performed using standard student t-test method.

3. Results

3.1 In vivo patient-specific vessel material properties based on Cine MRI data.

Stress-Stretch Ratio curves from Mooney-Rivlin models for the 16 plaque samples are presented by Figure 4 using average parameter values of slices with Cine data. Average YM values and circumferential shrinkage [28] (C-Shrink) values from 16 plaque samples were given in Table 2. The average YM values for the stiffest plaque sample (P16) was 922 kPa, 746% higher than that for the softest plaque (P13, YM=109 kPa). This showed that plaque material properties have large variations from patient to patient and patient-specific material properties should be used in plaque models. Average C-shrink value from the 16 samples was 6.51%. The softest sample had 21.7% C-shrink value, while the stiffest sample had a negative C-Shrink value (-1.78%). Negative C-Shrink value means the in vivo slice lumen needed to expand slightly to obtain the zero-load geometry of the 3D thin-layer model so that it could regain the in vivo circumference when 10% axial stretch and pressure were applied. Axial stretch makes the vessel to shrink in radial direction.

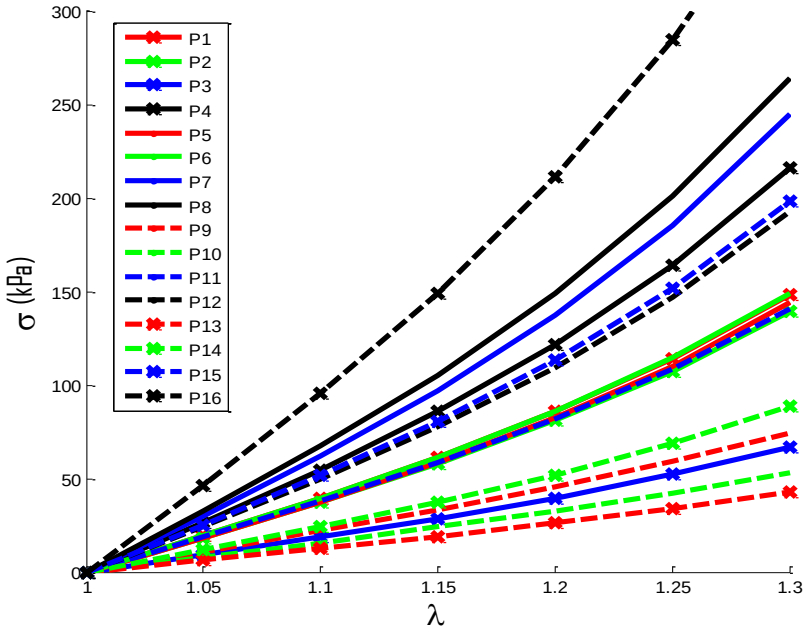


Figure 4. Stress-Stretch curves from Mooney-Rivlin Models using parameter values determined from Cine MRI for the 16 plaque samples studied.

Table 2. Average material parameter values and circumferential shrinkage for 16 human carotid plaque samples based on Cine MRI data. Due to axial shrink applied to the 3D thin-layer model, some C-Shrink values in 3D thin-layer model were negative.

Plaque	$Cir._{Max}$ (cm)	$Cir._{Min}$ (cm)	δ_{Cir} (%)	P (mmHg)	C_1 (kPa)	D_1 (kPa)	YM (kPa)	C-Shrink (%)
P1	2.564	2.350	8.35	(120,80)	34.6	14.6	370	7.40
P2	2.328	2.163	7.11	(120,80)	34.8	12.9	348	6.26
P3	2.415	2.119	12.3	(120,70)	19.5	5.31	169	12.3
P4	2.208	2.050	7.15	(141,72)	37.5	25.3	531	1.05
P5	2.089	1.903	8.90	(130,70)	30.0	15.3	357	4.51
P6	2.130	1.950	8.44	(143,80)	34.3	14.7	371	4.11
P7	2.952	2.747	6.93	(146,81)	41.7	28.9	601	2.45
P8	2.073	1.942	6.33	(146,81)	48.3	30.1	650	1.07
P9	2.341	2.094	10.6	(100,60)	28.8	3.66	191	7.48
P10	1.301	1.150	11.6	(100,60)	20.6	2.61	137	9.88
P11	2.136	1.906	10.8	(143,73)	35.0	13.1	352	4.56
P12	2.658	2.389	10.1	(143,73)	38.2	21.1	476	3.96
P13	1.537	1.306	15.0	(143,90)	16.3	2.16	109	21.7
P14	2.018	1.801	10.8	(143,91)	23.3	7.84	222	14.7
P15	2.127	1.933	9.11	(143,65)	41.5	21.0	492	0.92
P16	2.191	2.075	5.28	(143,65)	66.4	43.4	922	-1.78
Ave	2.192	1.992	9.3	(133,74)	34.4	16.4	394	6.51

3.2 Impact of patient-specific material properties on stress/strain calculations.

To show impact of patient-specific material properties on stress/strain calculations, we compare models using patient-specific material properties with models using material parameters from the literature (referred to as old material) [17]. Figure 5 gives an example to show the stress/strain differences from the two models. Average lumen stress and strain values for the 16 plaque samples calculated by the 3D thin-layer model using in vivo patient-specific material and old material are given in Table 3. The differences of average stress ($\Delta_{\text{-stress}}$) and average strain ($\Delta_{\text{-strain}}$) between in vivo material and old material models are given by:

$$\Delta_{\text{-stress}} = |\text{stress}_1 - \text{stress}_2|/\text{stress}_2, \quad (5)$$

$$\Delta_{\text{-strain}} = |\text{strain}_1 - \text{strain}_2|/\text{strain}_2, \quad (6)$$

where stress_1 and strain_1 are the average stress and strain values, respectively, calculated using the in vivo material model, and stress_2 and strain_2 are the average stress and strain values, respectively, calculated using the old material model. The average absolute variation of average stress values from 16 carotid plaques were 16.42%. The range of $\Delta_{\text{-stress}}$ values was [0.29%, 30.98%]. The average absolute variation of average strain values from 16 carotid plaques were 71.99%. The range of $\Delta_{\text{-strain}}$ values was [2.82%, 377.34%]. It shows using patient-specific material properties in computational models would lead to significant improvement on accuracy of plaque stress and strain calculation. Strain calculation is more sensitive to material stiffness changes.

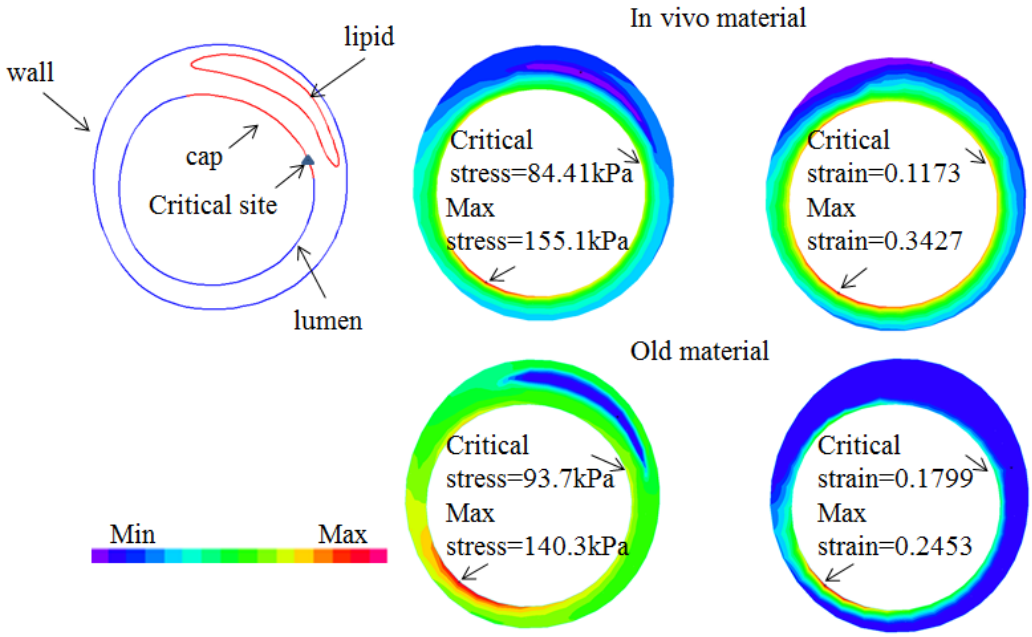


Figure 5. Stress and strain differences from models using patient-specific material and old material (in vivo material: C_1=20kPa, D_1=1.857kPa, D_2=2; old material: c_1=36.8kPa, D1=14.4kPa, D2=2; lipid core: c1=2kPa, D1=2kPa, D2=1.5; c2 = 0 for all materials.)

Table 3. The average stress and strain values for 16 carotid plaque samples based on MRI data by using in vivo material model and old material [17] model.

Plaque	p	In vivo material		old material		Variation	
		stress	strain	stress	strain	($\Delta_{\text{-stress}}$)	($\Delta_{\text{-strain}}$)
	(mmHg)	(kPa)		(kPa)		(%)	(%)
P1	(120,80)	92.30	0.2745	109.78	0.1952	15.92	40.63
P2	(120,80)	81.75	0.1862	101.34	0.1811	19.33	2.82
P3	(120,70)	78.25	0.3540	69.09	0.1332	13.26	165.77
P4	(141,72)	94.31	0.1345	87.31	0.1657	8.02	18.83
P5	(130,70)	89.82	0.2076	89.56	0.1683	0.29	23.35
P6	(143,80)	81.89	0.1644	93.64	0.1774	12.55	7.33
P7	(146,81)	113.16	0.1295	100.68	0.1865	12.40	30.56
P8	(146,81)	111.71	0.1375	98.27	0.1817	13.68	24.33
P9	(100,60)	55.95	0.2499	76.42	0.1408	26.79	77.49
P10	(100,60)	43.92	0.3107	58.67	0.1196	25.14	159.78
P11	(143,73)	97.00	0.1975	101.86	0.1900	4.77	3.95
P12	(143,73)	124.88	0.1860	133.75	0.2283	6.63	18.53
P13	(143,90)	96.27	0.6721	73.50	0.1408	30.98	377.34
P14	(143,91)	113.94	0.3978	91.32	0.1686	24.77	135.94
P15	(143,65)	92.96	0.1661	117.40	0.2155	20.82	22.92
P16	(143,65)	139.63	0.1173	109.59	0.2030	27.41	42.22
Ave	(133,74)	94.23	0.2429	94.51	0.1747	16.42	71.99
Min	(100,60)	43.92	0.1173	58.67	0.1196	0.29	2.82
Max	(146,81)	139.63	0.6721	133.75	0.2283	30.98	377.34

3.3 Critical stress and strain using patient-specific vessel material data.

The average maximum and critical stress and strain values from 16 plaque samples were given in Table 4. For 16 carotid plaques, systolic blood pressure shows positive correlation with maximum (Max) stress (r=0.5589, p=0.0244) and critical stress (r=0.6168, p=0.0109). It indicates that systolic blood pressure has significant impact on stress calculations. YM values showed positive correlation with critical stress (r=0.5733, p=0.0202), but its correlation with maximum stress was not significant (r=0.3531, p=0.1797). The YM value shows strong negative correlation with maximum strain (r=-0.8246, p<0.0001) and critical strain (r=-0.7376, p=0.0011).

Table 4. Summary of stress and strain values and other risk factors for 16 human carotid plaque samples from 3D thin-layer model with patient-specific vessel material data.

Plaque	PB	WT	P	YM	Stress (kPa)		Strain	
	(%)	(mm)	(mmHg)	(kPa)	Max	Critical	Max	Critical
P1	32	0.911	(120,80)	370	105.22	72.51	0.2904	0.1865
P2	34.2	0.883	(120,80)	348	105.41	62.6	0.2575	0.1425
P3	49.4	1.337	(120,70)	169	153.16	78.09	0.5491	0.3698
P4	46.5	1.145	(141,72)	531	127.07	84.21	0.2003	0.1385
P5	47.3	1.126	(130,70)	357	131.23	86.97	0.3001	0.2268
P6	42.1	1.059	(143,80)	371	113.49	67.91	0.2239	0.1325
P7	43.9	1.364	(146,81)	601	159.08	114.73	0.1619	0.1331
P8	49.8	1.327	(146,81)	650	144.46	116.33	0.1607	0.1256
P9	39.4	1.077	(100,60)	191	103.09	61.64	0.4479	0.2794
P10	64.1	1.293	(100,60)	137	67.88	57.89	0.4507	0.407
P11	39	0.872	(143,73)	352	111.25	82.05	0.2352	0.1641
P12	38	1.016	(143,73)	476	227.93	168.45	0.3046	0.2292
P13	49.7	0.98	(143,90)	109	127.26	87.59	0.7809	0.6467
P14	44.3	1.159	(143,91)	222	160.53	111.51	0.4899	0.3794
P15	40.2	1.098	(143,65)	492	114.23	82.47	0.2037	0.1465
P16	38.4	0.873	(143,65)	922	149.73	132.05	0.1173	0.1173
Ave	43.6	1.095	(133,74)	411	131.31	91.69	0.3234	0.2391
Min	32	0.872	(100,60)	109	67.88	57.89	0.1173	0.1173
Max	64.1	1.364	(146,81)	922	227.93	168.45	0.7809	0.6467

3.4 CPVI using in vivo material and old material and agreement with MPSI.

Figure 3 (b) and (c) give an example for the MPSI and CPVI groups, respectively. Figure 6 shows that critical stress values correlate much better with MPSI than the maximum (Max) stress values. Figure 6 also shows that critical stress values from in vivo material models correlate much better with MPSI than that from the old material model. A simple numerical code was used to determine five equal stress intervals [0, a), [a, 2a), [2a, 3a), [3a, 4a), and [4a, +∞) corresponding to CPVI values 0-4 to reach the best agreement between CPVI and MPSI. The five intervals (unit: kPa) [0, 46.8), [46.8, 80), [80, 92), [92, 103), and [103, +∞) from in vivo material models were used for

CPVI values of 0, 1, 2, 3 and 4, respectively. And the five intervals (unit: kPa) [0, 50.4), [50.4, 82), [82, 91), [91, 140), and [140, +∞) from old material models were used for CPVI values of 0, 1, 2, 3 and 4, respectively. The optimized agreement rate was 85.19% and 83.95%, respectively. The Pearson correlation coefficient between CPVI and MPSI was 0.9103 ($p<0.0001$) and 0.8661 ($p<0.0001$), respectively. Table 5 lists number of cases and agreement rate for each MPSI grade group. The ones with MPSI being 3 have the lowest match rates, which are 57.14% and 42.86%, respectively. From the agreement rates, the in vivo results more than 50%. According to the CPVI stress intervals, a plaque will be considered unstable (risk) if its critical stress is higher than 100 kPa from in vivo material models. And a plaque will be considered highly vulnerable (high risk) if its critical stress is higher than 140 kPa.

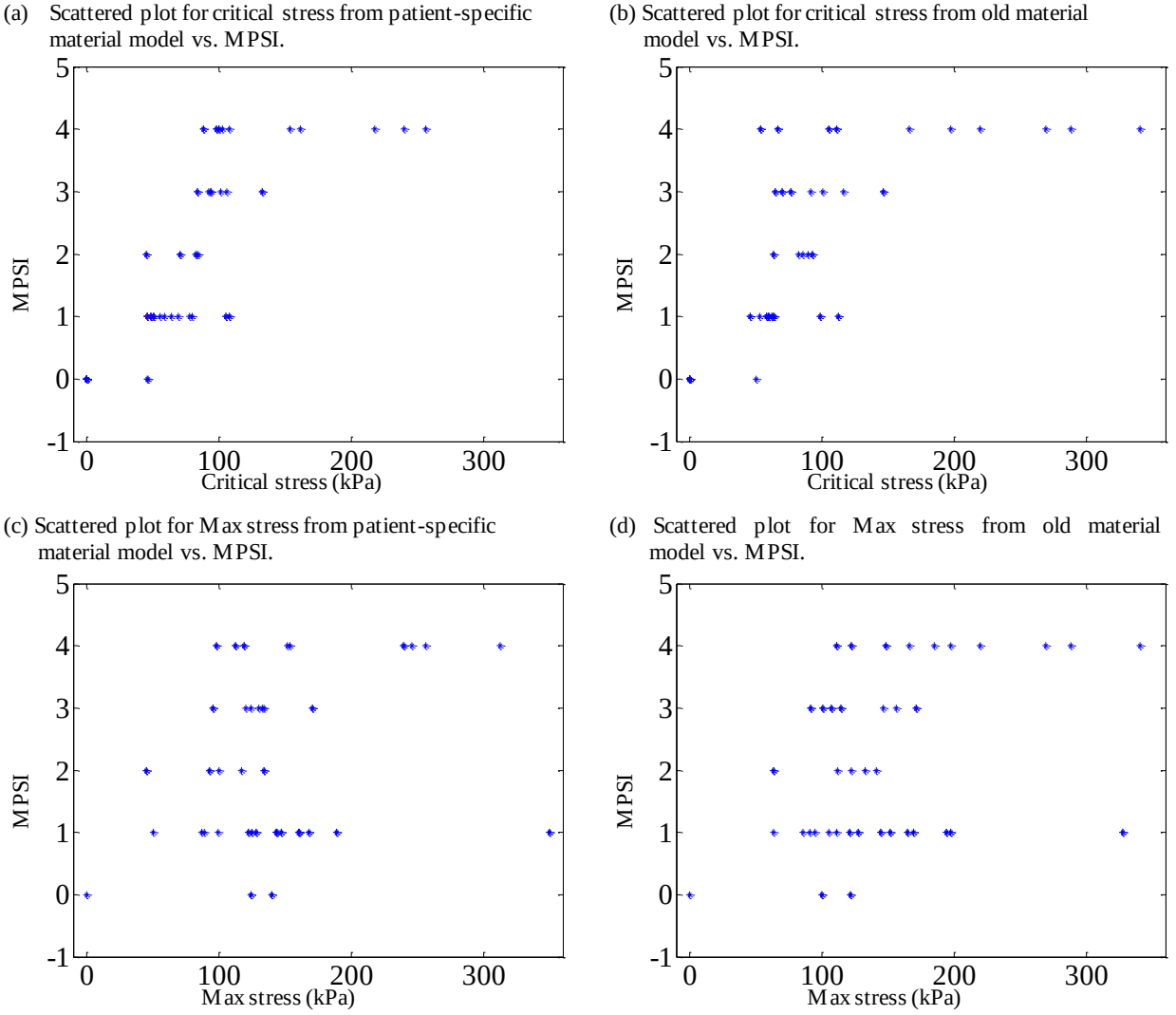


Figure 6. Critical stress shows much better correlation with MPSI from 16 carotid plaque samples. (a) Scattered plot for critical stress from patient-specific material model vs. MPSI; (b) Scattered plot for critical stress from old material model vs. MPSI; (c) Scattered plot for Max stress from patient-specific material model vs. MPSI; (d) Scattered plot for Max stress from old material model vs. MPSI.

Table 5. Case distributions according to MPSI and agreement rate between CPVI and MPSI

MPSI	Number of slices	Percentage (%)	Agreement Rate (%)	
			in vivo	old
0	44	54.32	100.00	100.00
1	15	18.52	73.33	80.00
2	5	6.17	60.00	60.00
3	7	8.64	57.14	42.86
4	10	12.35	70.00	60.00
All	81	100	85.19	83.95

4. Discussion

4.1. Significance of in vivo patient-specific vessel material properties.

Most of the research on determining arterial wall material properties has been performed using ex vivo specimens and in vitro experimental techniques. In vivo estimation of patient-specific material properties is scarce, which is a serious limitation for patient-specific plaque models. A noninvasive approach of combining in vivo Cine and 3D MRI and simple 3D thin-layer modeling was introduced to quantify patient-specific vessel material properties and improve model prediction accuracies. Our results from 16 plaques showed that slice YM values could vary from 109 kPa to 922 kPa, 7 times of the lowest YM value. Future studies should render plaque models using patient-specific material properties to quantify their impact on stress/strain calculations.

4.2 Vessel material has greater impact on strain predictions.

Using the in vivo material models, the average strain values from 16 plaque samples were 71.99% higher than that from the old material model, while average strain values were only 16.42% higher. Considering that most research reports have been focused on critical stress conditions, our results indicated that plaque mechanical investigations should include both critical stress and strain conditions when the accurate in vivo vessel material properties become available.

4.3 Threshold Critical Stress Value for Highly Vulnerable Plaques.

It should be noted that our threshold critical stress value from in vivo material model (103 kPa for CPVI=4) are lower than the threshold value from old material model (140 kPa for CPVI=4) for several reasons: a) Our models were based on in vivo material models could led to different stress predictions; b) Our 81 slices from 16 plaque samples included cases from stable to unstable and the number of CPVI=4 were only 7 slices.

4.4 Purpose of Introducing CPVI and Modeling Considerations.

The purpose of introducing CPVI is to have a more complete plaque assessment scheme which includes mechanical factors, plaque morphological features and tissue compositions for possible patient-screening applications. Results from 81 slices suggested that CPVI and MPSI had good agreement on plaque classifications. At the same time, the disagreement cases suggested that CPVI scheme may complement image-only assessment schemes and lead to potential improvements. The present study is the first in vivo case studies quantifying differences between mechanics-image combined and morphology-only assessment schemes.

It should be understood that plaque rupture is a multi-faceted process. CPVI covers only mechanical and morphological factors. We hope CPVI could provide complementing information for plaque assessment that image alone could not provide. Multiple biomarkers from different channels such as cell activities, lumen surface conditions, inflammation, blood conditions (cholesterol level and diabetes, for example) should be jointly considered for more complete and accurate vulnerability assessment.

4.5 Model limitations.

Cine MRI was used to determine vessel material parameter values, matching in vivo plaque geometries under both systolic and diastolic pressure conditions. Cine MRI is widely accepted to acquire time-dependent vessel motion and deformation. Multi-layer structure and anisotropic material properties of arteries were not considered since MRI does not provide layer information. Cine data provided only circumference variations under cardiac pressure. Another limitation was that location-specific pressure measurement was not available. Currently, arm cuff pressure values are used in most image-based studies. Noninvasive acquisition of intraplaque pressure data remains a challenge.

In computational models, contours for plaque components are generated based on segmentation data, the limitation of MRI resolution have impact on our calculated results. Furthermore, larger patient size will potentially lead to better plaque vulnerability prediction result.

5. Conclusion

Our preliminary results indicated that in vivo carotid vessel material properties have large variations from patient to patient, and vessel stiffness have impact on stress and strain calculations. In vivo material plaque model show significant difference with old material plaque model on stress and strain calculations. These differences showed that using in vivo material model to replace old material model would improve the accuracy of stress and strain calculation.

Funding:

This research was supported in part by NSF grant DMS-0540684 and NIH grant R01 EB004759. Tang's research was also supported in part by National Sciences Foundation of China grant 11672001 and a Jiangsu Province Science and Technology Agency grant BE2016785.

Conflict of Interest:

The authors declare that they have no conflict of interest.

Ethical approval:

This study was approved by the University of Washington Institutional Review Board and with written informed consent obtained. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Reference

1. World Health Organization. Global Atlas on cardiovascular disease prevention and control. Geneva: World Health Organization; 2011.
2. Fuster, V. The Vulnerable Atherosclerotic Plaque: Understanding, Identification, and Modification, Edited by V. Fuster, J. F. Cornhill, R. E. Dinsmore, J. T. Fallon,W. Insull, P. Libby, S. Nissen, M. E. Rosenfeld, and W. D. Wagner. AHA Monograph Series. Armonk, NY: Futura Publishing, 1998.
3. Naghavi, M., P. Libby, E. Falk, S. W. Casscells, S. Litovsky, J. Rumberger, J. J. Badimon, C. Stefanadis, P. Moreno, G. Pasterkamp, Z. Fayad, P. H. Stone, S. Waxman, P. Raggi, M. Madjid, A. Zarrabi, A. Burke, C. Yuan, P. J. Fitzgerald, D. S. Siscovick, C. L. de Korte, M. Aikawa, K. E. Juhani Airaksinen, G. Assmann, C. R. Becker, J. H. Chesebro, A. Farb, Z. S. Galis, C. Jackson, I. K. Jang, W. Koenig, R. A. Lodder, K. March, J. Demirovic, M. Navab, S. G. Priori, M. D. Rekhter, R. Bahr, S. M. Grundy, R. Mehran, A. Colombo, E. Boerwinkle, C. Ballantyne, W. Jr. Insull, R. S. Schwartz, R. Vogel, P. W. Serruys, G. K. Hansson, D. P. Faxon, S. Kaul, H. Drexler, P. Greenland, J. E. Muller, R. Virmani, P. M. Ridker, D. P. Zipes, P. K. Shah, J. T. Willerson. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: Part I. Circulation. 108(14):1664-72, 2003.
4. Naghavi, M., P. Libby, E. Falk, S. W. Casscells, S. Litovsky, J. Rumberger, J. J. Badimon, C. Stefanadis, P. Moreno, G. Pasterkamp, Z. Fayad, P. H. Stone, S. Waxman, P. Raggi, M. Madjid, A. Zarrabi, A. Burke, C. Yuan, P. J. Fitzgerald, D. S. Siscovick, C. L. de Korte, M. Aikawa, K. E. Juhani Airaksinen, G. Assmann, C. R. Becker, J. H. Chesebro, A. Farb, Z. S. Galis, C. Jackson, I. K. Jang, W. Koenig, R. A. Lodder, K. March, J. Demirovic, M. Navab, S. G. Priori, M. D. Rekhter, R. Bahr, S. M. Grundy, R. Mehran, A. Colombo, E. Boerwinkle, C. Ballantyne, W. Jr. Insull, R. S. Schwartz, R. Vogel, P. W. Serruys, G. K. Hansson, D. P. Faxon, S. Kaul, H. Drexler, P. Greenland, J. E. Muller, R. Virmani, P. M. Ridker, D. P. Zipes, P. K. Shah, J. T. Willerson. From vulnerable plaque to vulnerable patient: a call for new definitions and risk assessment strategies: Part II. Circulation. 108(15):1772-8, 2003.
5. Saam, T, M. S. Ferguson, V. L. Yarnykh, N. Takaya, D. Xu, N. L. Polissar, T. S. Hatsukami, C. Yuan. Quantitative evaluation of carotid plaque composition by in vivo MRI. Arterioscler Thromb. Vasc. Biol. 25(1):234-9, 2005.
6. Yuan, C., L. M. Mitsumori, K. W. Beach, K. R. Maravilla. Special review: Carotid atherosclerotic plaque: Noninvasive MR characterization and identification of vulnerable lesions. Radiology. 221:285–99, 2001.
7. Yuan, C., L. M. Mitsumori, M. S. Ferguson, N. L. Polissar, D. E. Echelard, G. Ortiz, R. Small, J. W. Davies, W. S. Kerwin, T. S. Hatsukami. In vivo accuracy of multispectral MR imaging for identifying lipid-rich necrotic cores and intraplaque hemorrhage in advanced human carotid plaques. Circulation. 104:2051–2056, 2001.
8. Underhill HR, Hatsukami TS, Fayad ZA, Fuster V, Yuan C. MRI of carotid atherosclerosis: Clinical implications and future directions. Nat Rev Cardiol.2010; 7(3):165-73.
9. Yuan C, Zhang SX, Polissar NL, Echelard D, Ortiz G, Davis JW, Ellington E, Ferguson MS, Hatsukami TS. Identification of fibrous cap rupture with MRI is highly associated with recent transient ischemic attack or stroke. Circulation.2002; 105:181–5.
10. Friedman MH, Krams R, Chandran KB. Flow interactions with cells and tissues: Cardiovascular flows and fluid-structure interactions. Ann Biomed Eng. 2010; 38(3):1178-87.
11. Tang D, Teng Z, Canton G, Yang C, Ferguson M, Huang X, Zheng J, Woodard PK, Yuan C. Sites of rupture in human atherosclerotic carotid plaques are associated with high structural stresses: an in vivo MRI-based 3D fluid-structure interaction study. Stroke.2009; 40:3258-3263. Featured article on MDlinx.com.
12. Holzapfel GA, Stadler M, Schulze-Bause CA. A layer-specific three dimensional model for the simulation of balloon angioplasty using magnetic resonance imaging and mechanical testing. Ann Biomed Eng. 2002; 30(6):753-767.
13. Bluestein D, Alemu Y, Avrahami I, Gharib M, Dumont K, Ricotta JJ, Einav S. Influence of microcalcifications on vulnerable plaque mechanics using FSI modeling. J Biomech.2008; 41(5):1111-1118.
14. Teng Z, Canton G, Yuan C, Ferguson M, Yang C, Huang X, Zheng J, Woodard PK, Tang D. 3D critical plaque wall stress is a better predictor of carotid plaque rupture sites than flow shear stress: an in vivo MRI-based 3D FSI study. J Biomech Eng.2010; 132(3):031007.
15. Kaazempur-Mofrad MR, Isasi AG, Younis HF, Chan RC, Hinton DP, Sukhova G, Lamuraglia GM, Lee RT, Kamm RD. Characterization of the atherosclerotic carotid bifurcation using MRI, finite element modeling, and histology. Ann Biomed Eng. 2004; 32(7):932–946.
16. Tang D, Kamm RD, Yang C, Zheng J, Canton G, Bach RG, Huang XY, Hatsukami TS, Zhu J, Ma G, Maehara A, Mintz GS, Yuan C. Image-based modeling for better understanding and assessment of atherosclerotic plaque

progression and vulnerability: Data, modeling, validation, uncertainty and predictions. Journal of Biomechanics. 2014; 47(4):834–846.
17. Tang D, Yang C, Zheng J, Woodard PK, Sicard GA, Saffitz JE, and Yuan C. 3D MRI-based multi-component FSI models for atherosclerotic plaques: a 3-D FSI model. Ann. Biomed. Eng.2004; 32(7): 947-960.
18. Li ZY, Howarth S, Trivedi RA, U-King-Im JM, Graves MJ, Brown A, Wang LQ, Gillard JH. Stress analysis of carotid plaque rupture based on in vivo high resolution MRI. J Biomech.2006; 39(14):2611-2622.
19. Gao H, Long Q, Das SK, Sadat U, Graves M, Gillard JH, Li ZY. Stress analysis of carotid atheroma in transient ischemic attack patients: evidence for extreme stress-induced plaque rupture. Ann Biomed Eng. 2011 Aug;39(8):2203-12.
20. Gao H, Long Q, Kumar Das S, Halls J, Graves M, Gillard JH, Li ZY. Study of carotid arterial plaque stress for symptomatic and asymptomatic patients. J Biomech. 2011 Sep 23;44(14):2551-7. doi: 10.1016/j.jbiomech.2011.07.012. Epub 2011 Aug 6.
21. Gallo D, Steinman DA, Morbiducci U. An insight into the mechanistic role of the common carotid artery on the hemodynamics at the carotid bifurcation. Ann Biomed Eng. 2015 Jan;43(1):68-81. doi: 10.1007/s10439-014-1119-0. Epub 2014 Sep 19.
22. Yang C, Tang D, Atluri S. Three-dimensional carotid plaque progression simulation using meshless generalized finite difference method based on multi-year MRI patient-tracking data. Computer Modeling in Engineering and Sciences.2010; 57(1):51-76.
23. Yang C, Tang D, Yuan C, Kerwin W, Liu F, Canton G, Hatsukami TS, Atluri S. Meshless generalized finite difference method and human carotid atherosclerotic plaque progression simulation using multi-year MRI patient-tracking data. Computer Modeling in Engineering and Sciences.2008; 28(2):95-107.
24. Yang C, Canton G, Yuan C, Ferguson M, Hatsukami TS, Tang D. Advanced human carotid plaque progression correlates positively with flow shear stress using follow-up scan data: an in vivo MRI multi-patient 3D FSI study. J Biomech. 2010; 43(13):2530-8.
25. Nieuwstadt HA, Fekkes S, Hansen HH, de Korte CL, van der Lugt A, Wentzel JJ, van der Steen AF, Gijsen FJ. Carotid plaque elasticity estimation using ultrasound elastography, MRI, and inverse FEA - A numerical feasibility study. Med Eng Phys. 2015 Aug;37(8):801-7.
26. Smoljkić M, Vander Sloten J, Segers P, Famaey N. Non-invasive, energy-based assessment of patient-specific material properties of arterial tissue. Biomech Model Mechanobiol. 2015 Oct;14(5):1045-56.
27. Czernuszewicz TJ, Homeister JW, Caughey MC, Farber MA, Fulton JJ, Ford PF, Marston WA, Vallabhaneni R, Nichols TC, Gallippi CM. Non-invasive in vivo characterization of human carotid plaques with acoustic radiation force impulse ultrasound: comparison with histology after endarterectomy. Ultrasound Med Biol. 2015 Mar;41(3):685-97.
28. Liu H, Canton G, Yuan C, Yang C, Billiar K, Teng Z, Hoffman AH, Tang D. Using in vivo cine and 3D multi-contrast MRI to determine human atherosclerotic carotid artery material properties and circumferential shrinkage rate and their impact on stress/strain predictions. J Biomech Eng. 2012; 134(1):223-233.
29. Stary, H. C., D. H. Blankenhorn, A. B. Chandler, S. Glagov, Jr. W. Insull, M. Richardson, M. E. Rosenfeld, S. A. Schaffer, C. J. Schwartz, W. D. Wagner, R. W. Wissler. A definition of the intima of human arteries and of its atherosclerosis-prone regions. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, AHA. Circulation. 85: 391–405, 1992.
30. Stary, H. C., A. B. Chandler, S. Glagov, J. R. Guyton, Jr. W. Insull, M. Richardson, M. E. Rosenfeld, S. A. Schaffer, C. J. Schwartz, W. D. Wagner, R. W. Wissler. A definition of initial, fatty streak and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, AHA. Circulation. 89: 2462–2478, 1994.
31. Stary, H. C., A. B. Chandler, M. D. Dinsmore, V. Fuster, S. Glagov, Jr. W. Insull, M. E. Rosenfeld, C. J. Schwartz, W. D. Wagner, R. W. Wissler. Definitions of advanced types of atherosclerostic lesions and the histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, AHA. Circulation. 92:1355–1374, 1995.
32. Cai, J. M., T. S. Hatsukami, M. S. Ferguson, R. Small, N. L. Polissar, C. Yuan. Classification of human carotid atherosclerotic lesions with in vivo multicontrast magnetic resonance imaging. Circulation. 106:1368–1373, 2002.
33. Tang, D., C. Yang, J. Zheng, P. K. Woodard, J. E. Saffitz, J. D. Petruccelli, G. A. Sicard, C. Yuan. Local maximal stress hypothesis and computational plaque vulnerability index for atherosclerotic plaque assessment. Annals of Biomed Engineering. 33(12):1789-1801, 2005.
34. D Tang,Z Teng,G Canton,TS Hatsukami,L Dong, XY Huang, C. Yuan. Local critical stress correlates better than global maximum stress with plaque morphological features linked to atherosclerotic plaque vulnerability: an in vivo multi-patient study. BioMedical Engineering OnLine. 8(1):15, 2009

35. Huang X, Yang C, Zheng J, Bach R, Muccigrosso D, Woodard PK, Tang D. 3D MRI-based multicomponent thin layer structure only plaque models for atherosclerotic plaques. J Biomechanics.2016; accepted.
36. Bathe KJ, 1996. Finite Element Procedures, Prentice Hall, Inc. New Jersey.
37. Bathe KJ, Editor. 2002. Theory and Modeling Guide, Vol I & II: ADINA and ADINA-F, ADINA R & D, Inc., Watertown, MA.
38. Kural MH, Cai M, Tang D, Gwyther T, Zheng J, Billiar KL. Planar biaxial characterization of diseased human coronary and carotid arteries for computational modeling. J Biomech. 2012; 45(5):790-798.
39. Cheng, G. C., H. M. Loree, R. D. Kamm, M. C. Fishbein, R. T. Lee. Distribution of circumferential stress in ruptured and stable atherosclerotic lesions, a structural analysis with histopathological correlation. Circulation. 87:1179–1187, 1993.
40. Kerwin, W., D. Xu, F. Liu, T. Saam, H. Underhill, N. Takaya, B. Chu, T. Hatsukami, C. Yuan. Magnetic resonance imaging of carotid atherosclerosis: Plaque analysis. Topics in Magnetic Resonance Imaging. 18: 371-378, 2007.
41. Long, Q., X. Y. Xu, B. Ariff, S. A. Thom, A. D. Hughes, A. V. Stanton, Reconstruction of blood flow patterns in a human carotid bifurcation: A combined CFD and MRI study, Journal of Magnetic Resonance Imaging. 11:299-311, 2000.

Numerical simulation of gasoline blending based on RJM system

***J. Y. Pan, Y. P. Qu, †S. Y. Chen, D. D. Wan, X. X. Tang**

Key Laboratory of High Efficiency and Clean Mechanical Manufacture
School of Mechanical Engineering, Shandong University, China

*Presenting author: 1933651144@qq.com
†Corresponding author: chensy66@sdu.edu.cn

Abstract

In order to improve the mixing efficiency and reduce the mixing time, a plurality of nozzles with different installation angles should be selected in the process of gasoline blending. Based on the gasoline physical parameters requirements, nozzle jet model and five kinds of gasoline blending schemes are built to simulate the mixing flow field. Combining the standard $k-\varepsilon$ turbulence model and the sliding grid technique, the numerical simulation and process optimization is done with the commercial Fluent software. The attenuation curves of the nozzle axis radial velocity at the landscape orientation, of the dynamic pressure at the transverse jet center, the velocity distribution curve of the interface and the velocity contours of every mixing flow field are studied. The injection performance of the nozzles and mixed characteristics of the RJM system is finally worked out to get the density- mixing time rules.

Keywords: Gasoline blending; Numerical simulation; Swirling jet

0 Preface

With the development of science and technology, people pay more and more attention to energy conservation and environmental protection, which improves the various industries' quality requirements for petroleum products. Because of the limitation of the processing technology, many mono-component products can't meet the needs, so two or multi component products are mixed and stirred in different proportions. These different components can give full play to their excellent performance in order to meet the product quality requirements of the consumers. Therefore, oil blending is a necessary part of the oil production process [1]-[2].

The research work of jet mixing began in the 50s of last century. In the past 20 years, with the development of CFD technology, the research of jet mixing system has been further developed [3]-[7]. In 2002, A. W. Patwarhan took the jet velocity, the nozzle angle and the geometric dimensions as variables to predict the mixing time and the concentration distribution of jet mixing system, the total mixing time of simulation is consistent with the actual test result [8]. In 2004, Zughbi and Rakib studied the effects of jet angle and jet number on mixing time [9]. In 2006, Sun Wei simulated the core components in the rotary jet mixing system- the axial flow turbine, which was used to clean up the industrial oil sludge, and the author proved the influence factors and control methods of the power turbine's hydraulic performance [10]. In 2007, Rahimi and Parvareh [11] studied the mixing process of the crude oil storage tanks, which simultaneously installed the nozzles and the impellers inside, the conclusion was that the angle

between the nozzle and the impeller had a significant effect on the mixing time. In 2008, Tian Yanli and others numerically simulated the three-dimensional flow field of the injection flow in the oil tank, they preliminarily studied the influence rule and the optimal combination of the nozzle structure, fluid properties, inlet conditions and the other factors [12]-[13]. In 2009, Parvareh and Rahimi et al proved that the nozzle position had a significant effect on the neutralization reaction rate in the mixing tank by means of experiments and numerical simulation [14]. Since 2010, Chen Songying [15]-[17] et al used the CFD software to simulate the flow field change of gasoline component, whose initial state was five-layer distributions under the action of rotary jet system, the simulation was established based on the three-dimensional gasoline components Mixture multiphase flow model, they investigated the change rule of density- mixing time under the different system speeds and nozzle angles. In recent years, more and more people work on the numerical simulation of rotary jet mixing system [18-21], which has a positive reference on the actual production.

Rotary jet mixing system (RJM) is a new attempt and application in the field of oil blending. The author designed the numerical calculation model of RJM system, simulated the flow field mixing of different nozzle angles and rotational speeds, got the flow field velocity contours, the axial dynamic pressure attenuation curve and the axial velocity and the range distribution curve. The author also obtained the density- mixing time curve of the specified cross-section and so on, got the change rules of dynamic characteristics, compared and analyzed the different mixing effect, conducted a preliminary analysis of optimization and comparison.

1 RJM system model

In order to research the dynamic characteristics of the gasoline blending, without considering the volatile oil and the electrostatic accumulation, the research was taken from the mixed flow field, found a jet nozzle model (Figure 1) to reflect some characteristics of the flow field in the actual operation system. Considering the important role of the jet shear force and the entrainment in the mixed flow, ignoring the influence of the pipeline in the system, the model is a simple open-loop system, with the import and the export boundary conditions. The flow field in the tank was the outflow field of RJM system. Because of the outlet boundary conditions of this model, that needed to increase the height of the tank model in order to avoid the effect to the phase proportion of the final mixture, ignored the free surface fluctuation effect on mixed flow at the top of gasoline blending tank [12].

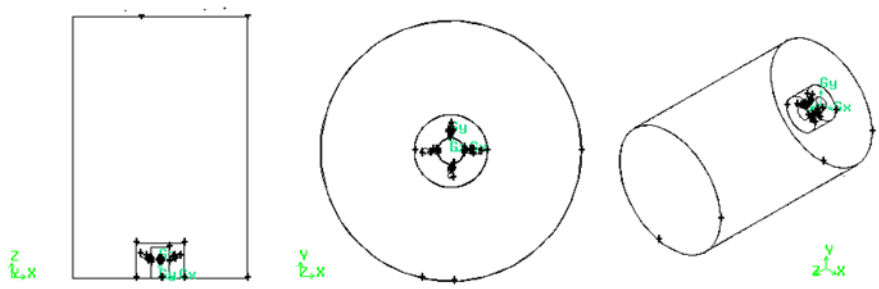


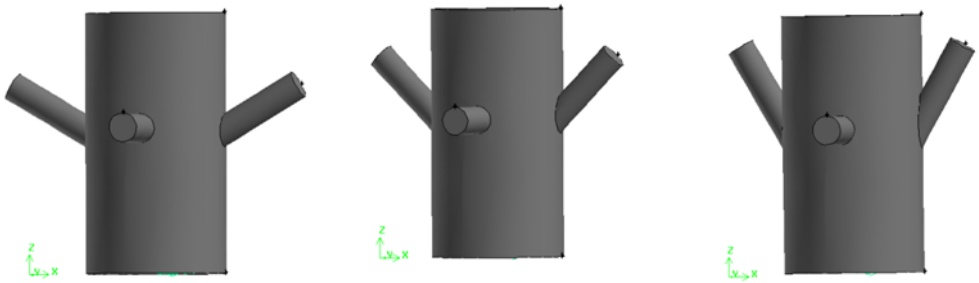
Figure 1. Nozzle jet model

2 Numerical simulation and optimization of mixed flow field

In order to improve the mixing efficiency and reduce the mixing time, it must have highly requirement of the flow field of the whole tank in the process of the gasoline blending. The author established five different kinds combination schemes $\emptyset 1\text{m} \times 1.5\text{m}$ of gasoline blending to do numerical simulation analysis and optimization of the flow field based on the blending tank and the real physical parameters of the gasoline ingredients.

2.1 RJM entity model

The inlet and outlet boundary conditions existed in the model of open loop system. In order to avoid the outlet boundary to destroy the various components of the gasoline, the height of the cylinder model should be increased properly, and an additional component was added as the main phase of the outlet boundary.



(a) $30°$ Angle of RJM system (b) $45°$ Angle of RJM system (c) $60°$ Angle of RJM system

Figure 2. Three RJM system appearances with different angles

In order to ensure a nozzle group arrangement with $180°$ distribution under the premise of constant horizontal position, five schemes are used in the different RJM models, the oblique installation angle of nozzles selected from $30°$, $45°$ and $60°$ (Figure 2) and the system speed selected from 0.1rad/s, 0.5 rad/s and 1 rad/s. In this paper, the author used the nozzle name code to distinguish the different schemes. For example, $0.1-45°$ nozzle is the model with a group of nozzles arranging at $180°$ in horizontal position, and the other group of oblique nozzles has $45°$ dip angle comparing with the horizontal plane (XY plane), and the speed of the main system is 0.1 rad/s. Following Table 1, the jet velocity at the exit of the five nozzles is the absolute velocity, and the velocity direction is perpendicular to the nozzle exit plane.

Table 1. Parameters of the five different schemes

Nozzle name	Jet velocity (m/s)	X direction angle (°)	Y direction angle (°)	RJM rotation rate (rad/s)
$0.1-45°$	3	0	45	0.1
$0.1-30°$	3	0	30	0.1
$0.1-60°$	3	0	60	0.1
$0.5-45°$	3	0	45	0.5
$1-45°$	3	0	45	1

2.2 Gridding and boundary conditions

Because the RJM system has a certain rotation speed, the magnitude and direction of the speed is constant, the author decided to use the sliding grid technique. The outer flow field of the whole system is divided into two parts, one is the moving region including the RJM system, and the other is the static region. The two regions are connected by the outer surface of the cylinder, and the type is defined as the Interface. The establishment of the interface was finished in the Fluent, so it can ensure the mutual transfer and exchange of each parameter of the flow field. The wall of the RJM system is set to the Moving Wall, and its rotational speed is 0 rad/s related to the moving region.

The size of the RJM system is different from the tank. Dividing the grid in block is to ensure the grid's quality. Using the Tet/Hybrid three-dimensional grid unit is to make full use of the established static and dynamic region. The grid sizes of two regions are set respectively to 0.0125mm and 0.015mm, which are divided totally into 580484 grids.

In the actual oil refining industry, a brand of gasoline was composed of the following four components: reformulated gasoline, MTBE, catalytic reforming gasoline and gasoline alkylate, plus the material of export spillover previously mentioned, so, this model was composed of five parts, as shown in Table 2.

Table 2. Physical parameters and distribution area of the blended gasoline components

Name	Density (kg/m^3)	Viscosity ($Pa \cdot s$)	Distribution by Z (m)
Reformulated gasoline	860.7	0.0004	0-0.22
MTBE	741.3	0.00036	0.22-0.32
Catalytic gasoline	733.2	0.00035	0.32-0.9
Gasoline alkylate	699.3	0.00034	0.9-1.0

The working pressure was set on the top of the blending tank at Z=1.5m plane, and the pressure value was the standard atmospheric pressure 101325Pa. Opened the gravity option, the size of the gravity was 9.8m/s^2, and the direction was the negative direction of the Z axis.

It should be noted that, in the initial state, all five components were assumed to be filled the whole tank in the form of stratification. Under the effect of gravity, the first phase with the maximum density (reforming gasoline) was located at the bottom of the tank, and the fifth material was distributed to the top. The Reynolds number at the nozzle exit (full of the first phase) was:

$$\mathrm{Re} = \frac{\rho v d_0}{\mu} = 6.14 \times 10^4 \tag{1}$$

The Reynolds number of the other regions with larger density was larger, so the mixed flow field was in a turbulent state. This model used the standard k-ε model, the boundary

conditions can be summarized as: the four nozzles were set to VELOCITY INLET, the absolute speed was 3m/s, the direction was perpendicular to the nozzle section and pointed to the external flow field.

2.3 Numerical simulation results

Because the five models had the same arrangements with the transverse nozzles, the nozzle structures and exit velocities were also the same, they had the same axial velocity attenuation law and section velocity distribution law. Taking the $0.1-45°$ nozzle as an example, the injection performance of the nozzles was analyzed from the radial velocity attenuation curve of the nozzle axis and the center pressure attenuation curve of the nozzle jet.

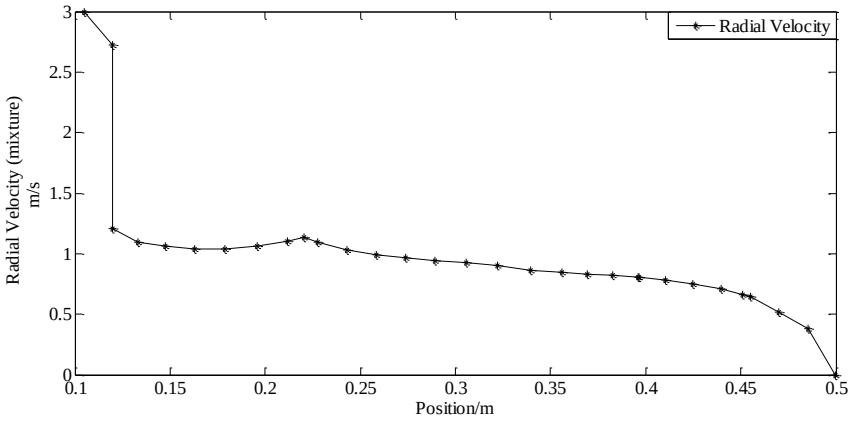


Figure 3. Radial velocity attenuation curve of the $0.1-45°$ nozzle axis

Figure 3 shows the radial velocity attenuation curve of the transverse nozzle of the system, which eliminates the influence of the axial rotation of the system. It can be seen that the horizontal nozzle in the most of flow field can maintain a moderate speed, about 1m/s, which is very important to ensure the uniformity of oil blending. In the vicinity of the wall, Y=0.5m, the axial velocity decay to about 0.35m/s, which is the lowest rate to ensure the operation of oil blending. Noted that the nozzle axis radial velocity in the vicinity of X=0.105m, the speed decays quickly, it is because that the position is the interface of dynamic and static area during the numerical simulation time, which had a greater resistance to the interface, on the other hand, the grid was in constant slip state, that affected the normal transmission of velocity field.

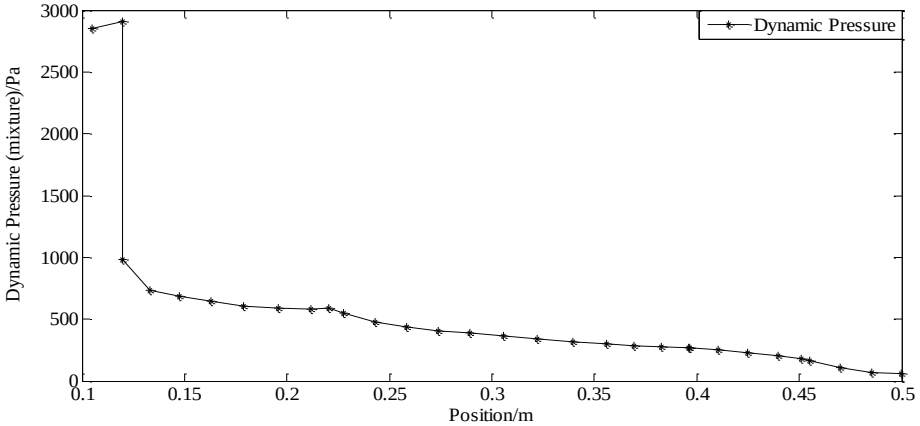


Figure 4. Center dynamic pressure attenuation curve of the $0.1-45°$ nozzle

Figure 4 shows the center dynamic pressure attenuation curve of the nozzle in this system. The jet dynamic pressure reached the maximum value at the nozzle exit, the jet dynamic pressure and total pressure were reduced after the mixture released from the nozzle, then it decayed to a small level near the wall after remaining a certain distance. Therefore, the shock force of the tank wall was small, which helped to reduce the accumulation of static electricity in the mixed flow field. The sharp drop of dynamic pressure at X=0.105m was also caused by the interface, the reason was similar to the speed decay.

The diffusion effect of the nozzle was studied by analyzing the distribution of the cross section velocity perpendicular to the axial direction of the nozzle jet.

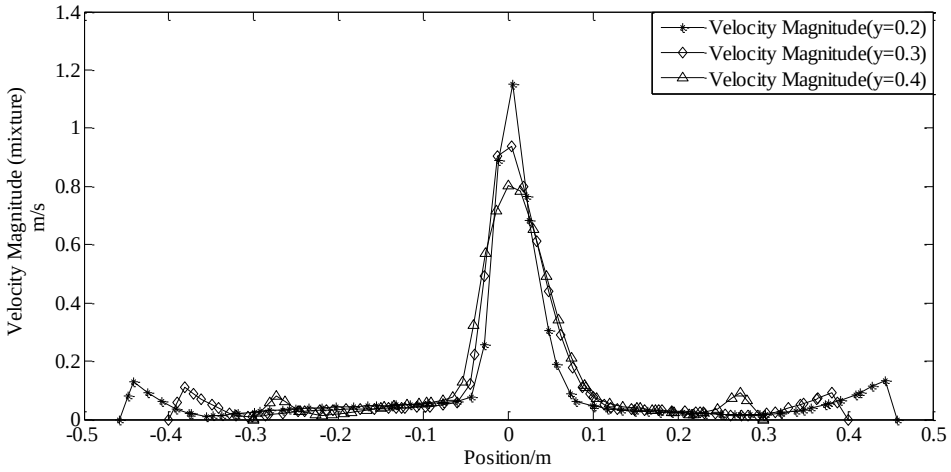


Figure 5. Velocity profile of the three ranges on the Z=0.1 plane

Figure 5 shows the mixing of the jet-flow to the same plane (Z=0.1). The velocity profile of the three different ranges of Y=0.2m, Y=0.3m and Y=0.4m plane can be seen, the maximum speed of the axis decreases with the increase of range, namely 1.15m/s, 0.94m/s and 0.8m/s. This is consistent with the axial velocity decay law of the nozzle as shown in Figure 3. The velocity distribution of the section has a self - mode, that is, the maximum velocity on the axis, the smaller the farther away from the axis, and the symmetrical distribution on the axis, the cross section velocity on different ranges is similar. The velocity in the decay process, appeared a couple symmetrical angle at the high speed on the both ends, this was due to the influence of the axial velocity of the flow field, and the velocity was mainly caused by the deflection of the flow direction at the end of the jet-flow.

(1) Simulation results of different nozzle angles at the same rotating speed

In this model, the numerical simulations of the mixing characteristics of the RJM system with three kinds of oblique nozzle inclination angle of $30°$, $45°$ and $60°$ were carried out. The results of simulation and analysis were as follows.

Figure 6, Figure 7, Figure 8 clearly show the core area and the main body of the jet-flow, as well as the velocity gradient in the process of oblique nozzle jet attenuation. For the RJM system with the $0.1-30°$ nozzle, the velocity was about 0.3m/s - 0.35m/s in the axial direction of the nozzle when the exit velocity was 3m/s, the velocity met the requirements of petroleum

blending. In addition, the variations of the RJM system rotates counterclockwise along the Z axis and the jet-flow region under the effect of gravity can also be seen. For different nozzle inclination angles, the main affected areas of the jet-flow were also different. For the $0.1-30°$ nozzle, stir at the bottom of the tank was violent, conducive to a high density components upward movement to full mixture. For the $0.1-60°$ nozzle RJM system, the range of the oblique nozzle was far, focusing on the mixing in the upper and middle blending tank, relatively easy to cause the volatilization of the gasoline, while it was not conducive to the rapid mixing of the larger density phase at the bottom of the tank.

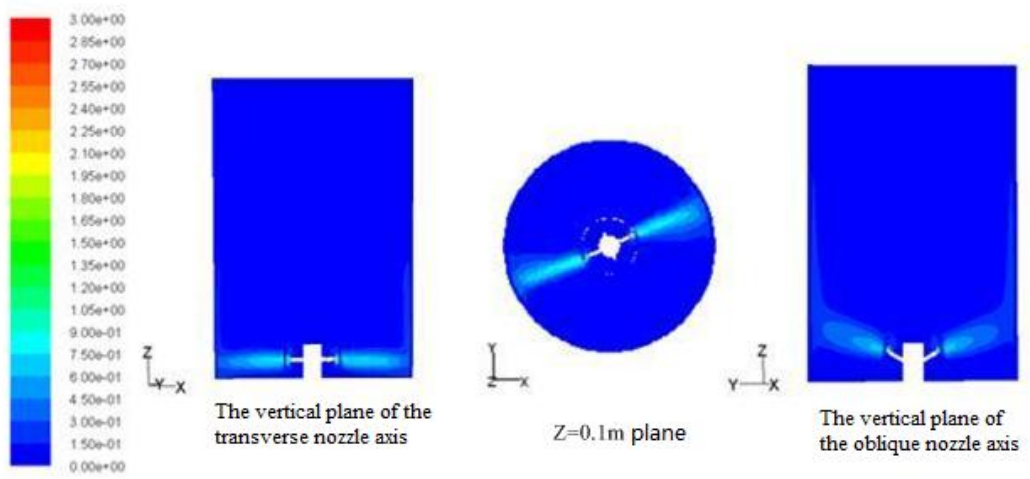


Figure 6. The velocity nephogram in the three plane mixed phase of the $0.1-30°$ nozzle RJM system at 460s

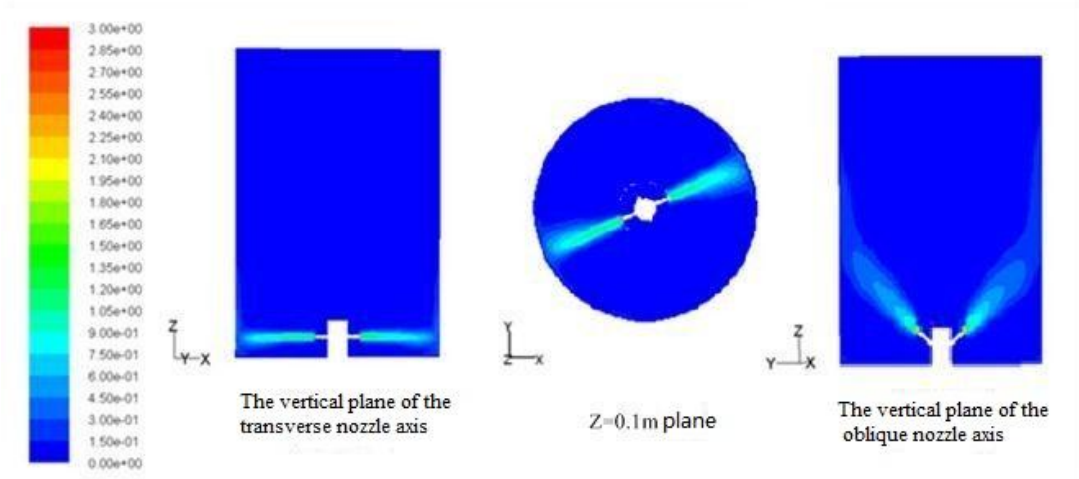


Figure 7. The velocity nephogram in the three plane mixed phase of the $0.1-45°$ nozzle RJM system at 460s

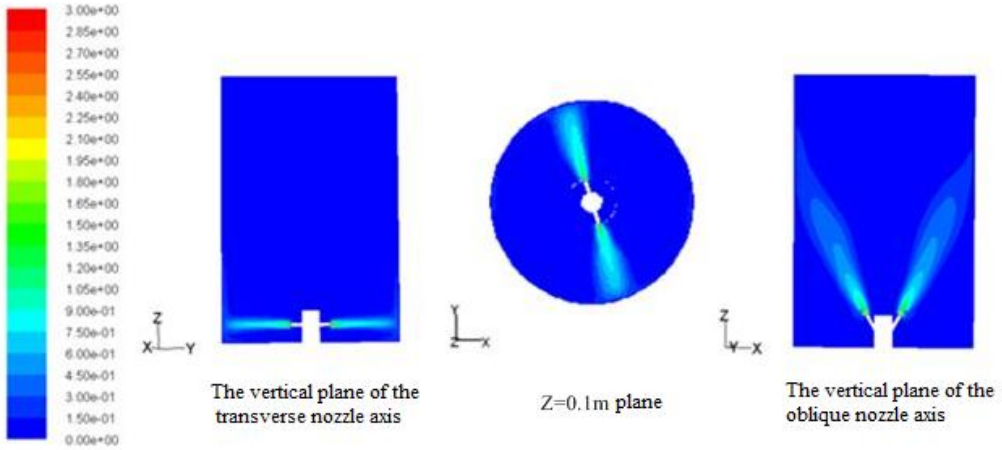


Figure 8. The velocity nephogram in the three plane mixed phase of the $0.1-60°$ nozzle RJM system at 460s

The axial flow in the mixed flow field was necessary to make the initial state of the multiphase flow in a stratified arrangement, which was mainly dependent on the jet flow of the oblique nozzle in the RJM system. In the Y=0 plane, the velocity distributions in the range X=0.2m, X=0.3m and X=0.4m were taken to investigate the jet influence of the mixing of the upper and lower fluid to the tank, as shown in Figures 9, 10 and 11. Figures 9, 10 and 11 illustrated the velocity distribution of an oblique nozzle in a straight line in the Z direction, the velocity profile of the nozzle was similar to that of the transverse nozzle, and there were also high speed and low speed zones. But the difference was that, due to the influence of oblique nozzle angle, for different values of X, high speed area appearing in the Z direction was not consistent, and with the increase of X value, high speed zone shifted to Z positive axis, maximum velocity also showed the attenuation trend. For example, for the $0.1-45°$ nozzle RJM system, the maximum speed of the X=0.2m section was in the vicinity of Z=0.25m, the maximum value was 0.7m/s, while the maximum speed of the X=0.4m section was in the vicinity of Z=0.45m, the maximum was attenuated to 0.46m/s.

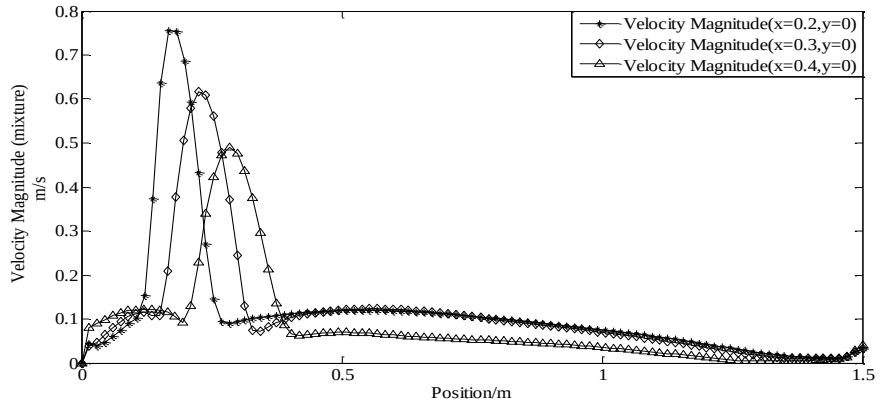


Figure 9. The velocity profile of the three range of the $0.1-30°$ nozzle RJM system

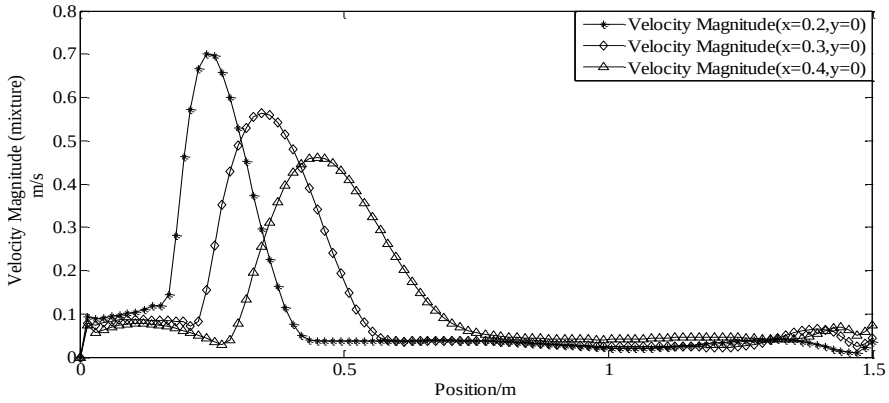


Figure 10. The velocity profile of the three range of the $0.1-45°$ nozzle RJM system

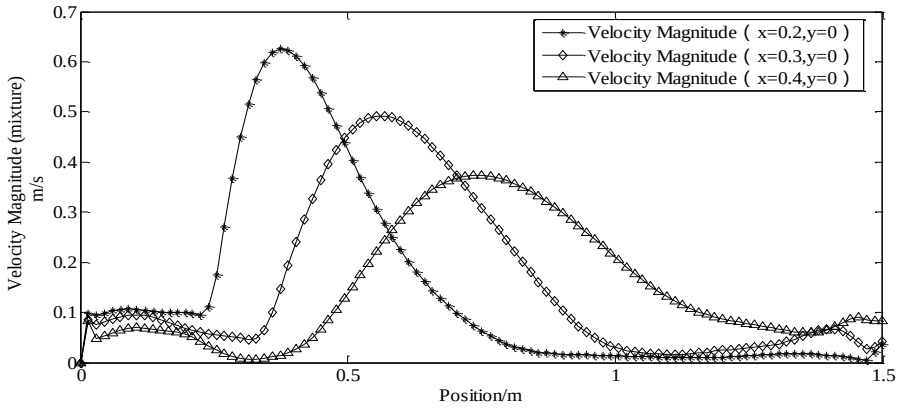


Figure 11. The velocity profile of the three range of the $0.1-60°$ nozzle RJM system

It was different from the transverse velocity distribution of the transverse nozzle (Figure 5), the above three diagrams also showed that the velocity distribution on the cross section was not symmetrical under gravity. In the flow field at the bottom of the tank, the velocity could be kept constant by the influence of the circumferential flow formed by the transverse nozzle. So in the [0, 0.2] region of the abscissa, a constant speed was independent of the range X. Similarly, affected by the outlet pressure, the upper oil tank area would develop a temporary speed recovery area, which was the direct result of the abnormal velocity distribution in Figure 11 at the X=0.4 cross section.

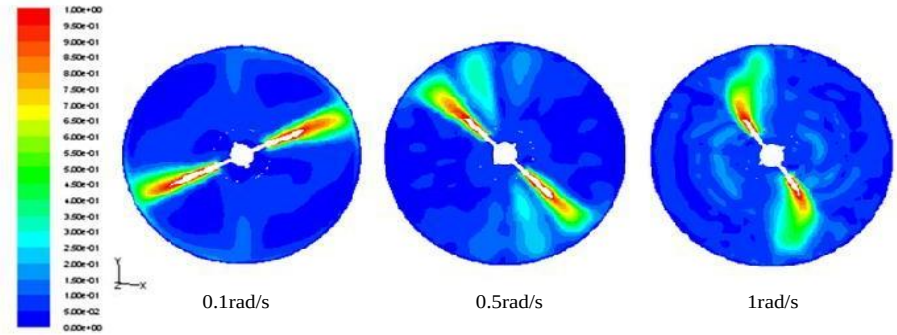


Figure 12. Velocity distribution of Z=0.1 section with different rotating speed of the nozzle RJM system

(2) Simulation results of the same nozzle inclination at different speeds

In order to explore the influence of the rotation speed of RJM system around the Z axis on the mixing effect, the mixing characteristics of RJM system with three different speeds (0.1rad/s, 0.5rad/s and 1rad/s) were simulated. The most direct effect on the mixing flow field was the change of circumferential velocity in the XY plane. As shown in Figure 12, the 0.1rad/s speed of the system will appear in the vicinity of the nozzle 0.1m/s - 0.15m/s low velocity correlation, but the overall impact on the flow field is very small, it can be regarded as static. As shown in Figure 12, the 0.1m/s - 0.15m/s low velocity correlation would appear in the vicinity of the transverse nozzle under the 0.1rad/s speed of the system, but the overall impact on the flow field was very small, that can be regarded as static. The system speed of the 0.5rad/s was obviously wider than that of the 0.1rad/s, and the range of the high speed area was larger, and it was obvious that the jet traces left by the transverse nozzle at the last moment. In the Z=0.1 plane, the mixed flow field of 1rad/s system was more variable, the relatively large rotation speed caused a circumferential velocity, the velocity was larger and larger along the radial direction, which was proportional to the distance from the point to the center of the circle. Finally, the direction of the jet was deflected along the opposite direction of rotation, and the terminal velocity of the whole flow field was up to 0.3m/s, that was beneficial to the mixing of multiphase fluids.

The above results do not show that with the increase of the rotating speed of the system, the mixing time will be shorter. This is because a mixture of multiphase flow mainly depends on the circumferential velocity of jet, the increasing rotation speed will significantly increase the circumferential velocity of flow, and that may format the vortex in the XY plane, which has little effect on mixing multiphase flow, sometimes plays an opposite role. On the other hand, the large circumferential velocity has a great influence on the range of the transverse nozzle, so, it is easy to form a dead zone in the tank wall, which affects the final mixing time.

3 Conclusion

This article based on the reduced gasoline tank model without considering the nozzle structure, ignoring the premise of volatile oil and electrostatic accumulation. Mixing model was established from the jet mixing angle, and the author studied on the gasoline blending dynamics by using the Fluent software, obtained the following conclusions:

(1) The transverse nozzle can maintain a moderate speed about 1m/s in most regions of the flow field, which was very important for ensuring the uniformity of oil blending. In the vicinity of the wall, the axial velocity was reduced to about 0.35m/s, which ensures the smooth operation of the lowest rate of oil blending. It can be seen that the RJM jet mixer can reduce the volatile gasoline at the same time it can ensure the uniform mixing of the most flow field.

(2) The jet dynamic pressure reached maximum at the nozzle outlet. When the mixed fluid left the nozzle, the jet dynamic pressure and total pressure were decreased, after maintaining a distance, finally were attenuated to a very small level near the wall. Therefore, the shock force of the tank wall was small, which helped to reduce the accumulation of static electricity in the mixed flow field.

(3) For the 30° nozzle RJM system, stir on the bottom of the tank was violent, which was conducive to the phase with larger density moving upwards to fully mix. For the 60° nozzle RJM system, the oblique nozzle had a long range, it focused on the upper mixed in the blending tank, which may easily cause the gasoline volatilization, while not conducive to rapidly mix of the high density phase at the bottom of the tank.

Acknowledgement

This project is supported by Science and Technology Development Planning of Shandong Province (2014GGX108001), Key research and development program of Shandong Province (2016GGX104018).

Reference

[1] Li, Z. X. and Xu, S. W. (1997) *Oil Storage and Transportation Design Manual*, Petroleum Industry Press, Beijing, China.

[2] Chang, S. R. and Liang, W. (1999) Progress and prospect of gasoline, *Natural Gas and Petroleum* **04**, 19-22.

[3] Ranade, J. J. (1996) Towards better mixing protocols by designing spatially periodic flows: the case of a jet mixer, *Chemical Engineering Science* **51**, 2637-2642.

[4] Masoud, R. and Arsalan, P. (2005) Experimental and CFD investigation on mixing in a semi-industrial stirred tank, *Chemical Engineering Journal* **115**, 82-92.

[5] Wang, X.P. (2008) Research on the jet mixing system used for oil tank anti-sedimentation, Shandong University, Jinan, China.

[6] Liu, L. Y., Meng, F. L., Wang, W. Q. and Song, B. (2015) Application and development of agitator in crude oil storage tank, *Contemporary Chemical Industry* **02**, 386-388+396.

[7] Wang, J. L. (2007) Numerical simulation and hydraulic characteristics of internal flow in hydraulic driven rotary jet mixer, Zhejiang University, Hangzhou, China.

[8] Patwardhan, A. W. (2002) CFD modeling of jet mixed tanks, *Chemical Engineering Science* **57**, 1307-1318.

[9] Zughbi, H. D. and Rakib. M. A. (2004) Mixing in a fluid jet agitated tank: effects of jet angel and elevation and number of jets, *Chemical Engineering Science* **59**, 829-842.

[10] Sun, W. (2006) The hydraulic performance study and structure optimization design of power turbine in industrial sludge jet mixing system, Zhejiang University, Hangzhou, China.

[11] Rahimi, M. and Parvareh, A. (2007) CFD study on mixing by coupled jet-impeller mixers in a large crude oil storage tank, *Computers and Chemical Engineering* **31**, 737-744.

[12] Tian, Y. L., Hang, L. Q. and Pan, D. J. (2008) Full numerical simulation analysis of rotary jet mixer, *Electromechanical Engineering* **25**, 86-90.

[13] Wang, L. Q., Lin, S. D. and Tian, Y. L. (2008) Study on jet performance of large flow nozzle based on CFD, *Fluid Machinery* **11**, 17-22.

[14] Parvareh, A., Rahimi, M. and Alsairafi, A. A. (2009) Experimental and CFD studies on the effect of the jet position on mixing performance, *Iranian Journal of Chemical Engineering* **6**, 3-10.

[15] Chen, S. Y., Xie, F. C. and Qu, Y. P. (2010) Research on the jet mixing system used for oil blending, *Journal of Engineering Thermo-physics* **31**, 229-232.

[16] Xie, F. C. (2011) Study on the dynamics peculiarity of gasoline blending in the process of rotary jet mixing, Shandong University, Jinan, China.

[17] Liu, C. W. (2015) Research on the jet stirring dynamics of lime slurry pond of wet flue gas desulfurization system, Shandong University, Jinan, China.

[18] Wang, M. R., Zhang, Y. F. and Cai, Y. B. (2014) Numerical simulation of flow field of omni-directional rotate jet mixer, *Chemical Industry and Engineering Progress* **S1**, 79-84.

[19] Yu, B. B., Yang, J. Y. and Liu, Q. (2015) Numerical simulation on rotary jet mixer of oil tank at different jet velocity, *Contemporary Chemical Industry* **06**, 1412-1414+1418.

[20] Fan, D. W. (2016) Numerical simulation on rotary jet mixer of large storage tank, *Oil-gas field Surface Engineering* **09**, 4-5+9.

[21] Hu, J. G. (2017) Study on CFD for coal powder slot anti-settling jet mixing system, *Shanghai Chemical Industry* **01**, 16-20.

Research on the features of gasoline mixture flow field with rotary jet mixing

***D. D. Wan, Y. P. Qu, †S. Y. Chen, J. Y. Pan, X. X. Tang**

Key Laboratory of High Efficiency and Clean Machinery Manufacturing
School of Mechanical Engineering, Shandong University, China

*Presenting author: 1242855830@qq.com
†Corresponding author: chensy66@sdu.edu.cn

Abstract

Numerical simulation of the gasoline flow features and mixing efficiency in a gasoline mixture tank with a rotary jet mixing (RJM) system installed at the bottom center has been studied applying the standard turbulent model and slipping grid technique. The result shows that the RJM does well at mixing various components with no blind corner and high mixing efficiency. The mixing density difference met the mixing requirement for the first time at 31.2s and then showed a tendency of deterioration. It met the requirement again at 58.2s with the mixing density difference keeping in the mixing criterion of 3‰.

Keywords: Gasoline mixing; Rotary jet mixing; CFD simulation; Mixing quality

Introduction

Refinery enterprises usually adopt different processing technology to attain sorts of gasoline with different densities through various steps like atmospheric distillation, hydrogenation and etc. In order to meet the national petroleum products standards, the various line components should be mixed to make the gasoline physical and chemical properties more uniform. Thereby, gasoline mixing is a necessary step in the production of petroleum, the mixing efficiency directly corresponds to the quality of petroleum [1].

Gasoline mixing in refinery enterprises consists two main categories: tank mixing and pipe mixing. Tank mixing approaches include compressed air mixing [2], mechanical agitation [3]-[4] and nozzle mixing with pump circulation [6]-[8]. The first usually leads to the gasoline oxidation because of the air in the tank. Moreover, the compressed air will produce strong vortex which probably cause static electricity and this immensely threatens the tank safety. Therefore, the compressed air method has a tremendous limitation in production. Mechanical agitation is also a common method in gasoline mixing, but no matter the axial flow or the radial flow is adopted, it still causes blind corners easily which will lower the stirring efficiency with high energy consuming. High-speed nozzle jet mixing with pump circulation works in this way: the gasoline enters the tank again through the nozzle jet and the submerged jet flow will promote the motion of the static fluid, then a plenty of vortexes will generate in the boundary of jet flow, which in turn trap surrounding fluids into the jet to improve the mixing of fluids. Nozzle mixing with pump circulation is applied wider gradually for its simple structure, high safety, convenient operation and etc.

Date up to 1951, Fosset [9] had already conducted the study on jet mixing and found that nozzle mixing has higher mixing efficiency than traditional mechanical agitation; In 1982, Maruyama, Ban and Mizushina [10] found that the mixing time was up to the depth of fluid and nozzle length; In 1983, Zhu and Chang [11] introduced the principle and effect of nozzle mixing with pump circulation; In 2004, Yu [2] analyzed the features of fluid filed in a tank with a rotary nozzle and the result showed that the distribution of nozzles had an obvious effect on the fluid filed. With the development of CFD, jet stirring gets a further promotion [12]-[14]. Wang [15] studied the performance of large flux nozzle based on CFD; In 2007, Wang [16] simulated the inner flow features of jet agitator numerically; In 2012, Zhang [17] et al researched the rotary nozzle for gasoline mixing. Barekatain, H [18] et al improved the mixing by submerged rotary jet system with CFD software in a large storage tank; Neyestanak [19] et al introduced a new relation of estimating the mixing time of crude oil tank with a submerged rotary jet mixer. Zhong [20] et al studied the gas-liquid two phase flow in a slurry pool with rotary jet mixing. In this paper, CFD is used for studying the flow features and mixing efficiency in a gasoline mixing tank with RJM system.

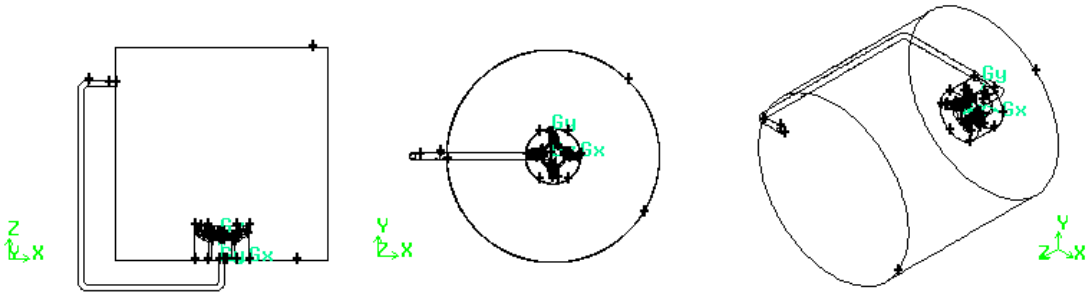


Figure 1. Sketch of the mixing tank model

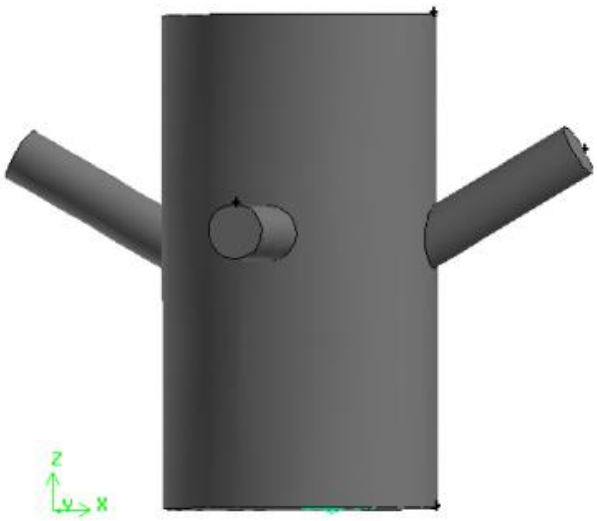


Figure 2. The RJM system with 30 °inclined nozzle

1. The geometric model and meshing

1.1 The geometric model

The geometry structure of the mixing tank is illustrated as Figure 1. Under an assumption of no pipe leak and loss of flow, the system can be assumed as closed. Thus, there does not exist inlet and outlet boundary conditions. The motivation of the whole in-tank system is provided by the source term nearby the outlet pipe, which approximates the function of circulating pump. For the convenience of calculation, the diameter and the height of the tank is set to be 1m and 2m, respectively. The diameter of the rotary jet nozzle is 26mm, of which the nozzle number is 4 in a uniform distribution across the 360 °circumferential directions. One group is horizontal and another inclined upward, of which the axis is 30 °to horizontal level. The detailed structure is showed as Figure 2.

1.2 Meshing and boundary conditions

Software Gambit 6.3 is used to mesh the model. Because the RJM system has a fixed rotating speed with constant magnitude and direction, sliding mesh is adopted to divide the whole flow zone into four parts: moving zone, static zone, source zone and pipe zone except source term. In the model, the pipe diameter is 0.03m and the interface between zones is defined as interface and the wall of RJM system is moving wall with a rotary speed of 0 rad/s relative to the moving zone showing in Figure 3. These four zones are all meshed with Tet/Hybrid 3D element. In order to assure the grid quality, the grid sizes are: 0.0125mm in moving zone, 0.02mm in static zone, 0.01mm in both source zone and pipe zone, respectively. Verified the grid independence, the total number of grids is 732205.

The operating pressure, whose value is standard atmospheric pressure 101325 Pa, is set to act on the top plane of the tank with z=1m. Gravity term with a magnitude of 9.8 m/s^2 and a direction pointing to minus z axis is chosen. Since the mixing flow field is turbulent, the standard $k-\varepsilon$ model is applied. As for the phase, four components, which can reflect the mixing state of a certain kind of gasoline, are chosen as Table 1.

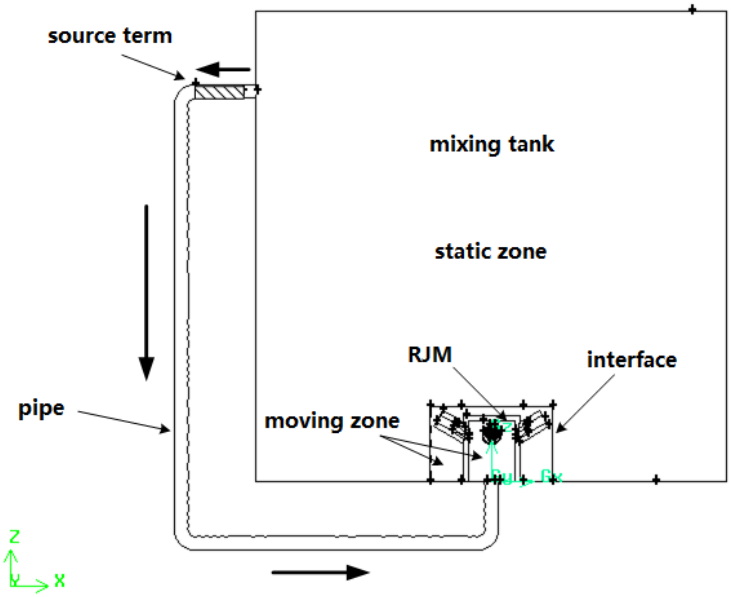


Figure 3. Divided zones in the tank

Table 1. The physical property and distribution zones of the main components of a certain brand mixing gasoline

Name	Density (kg/m^3)	Viscosity (Pa s)	Z direction distribution zones(m)	Phase number
Reforming gasoline	860.7	0.00040	0-0.225+source term zone	first
MTBE	741.3	0.00036	0.225-0.315	second
Catalytic gasoline	733.2	0.00035	0.315-0.912+pipe zone	third
Gasoline alkylate	699.3	0.00034	0.912-1.0	fourth

Note: The density and viscosity of the kinds of gasoline listed above all was measured at 20℃.

1.3 Calculation strategy

The continuum equation, turbulent equation and slipping velocity equation in constant flow are solved then the volume distribution function is calculated in unsteady flow state. Therefore, the convergence can be accelerated and a convergent density field can be attained.

According to GB/T 4756-1998 manual sampling of gasoline liquid, three points, A (0.3, 0, 0.1), B (0.3, 0, 0.5), C (0.3, 0, 0.9), are chosen as density monitoring points in the mixing tank. More serious mixing time criterion is put forward:

$$t_{95\%} = \text{time for } \left|\frac{d-\bar{d}}{\bar{d}}\right| \leq 0.003 \quad (1)$$

Where d is the density of monitoring points with mean value of the whole flow field. It can be regarded as uniform mixing when the relative density value between the point A, B and C equidistant in vertical direction becomes smaller than 3‰.

2. Results and discussion

Figure 4 shows the axial velocity distribution of the RJM, in which the axial velocity in zone [-0.5, 0.5] is exactly caused by the fluid in the inlet pipe and this conforms to the velocity distribution law in pipe flow. The area nearby the wall of RJM has a minus value of velocity and this is caused by the fluid turning around after crashing the top plane of the RJM system.

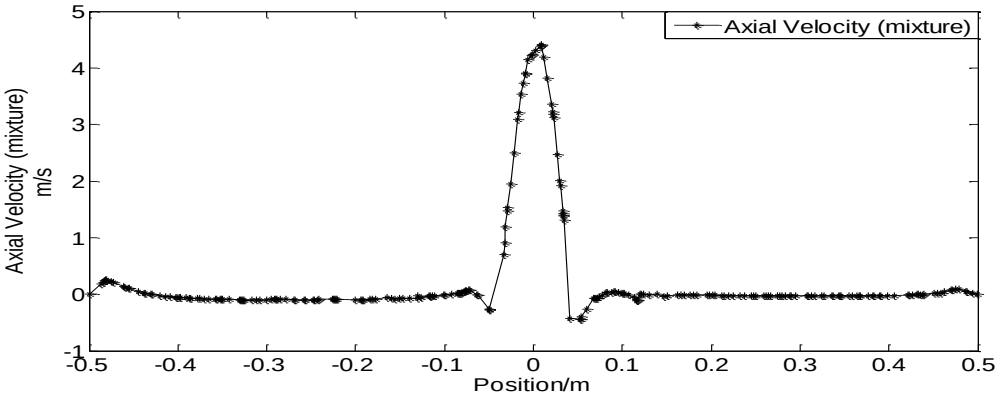


Figure 4. The axial speed distribution of the horizontal nozzles in x=0 plane

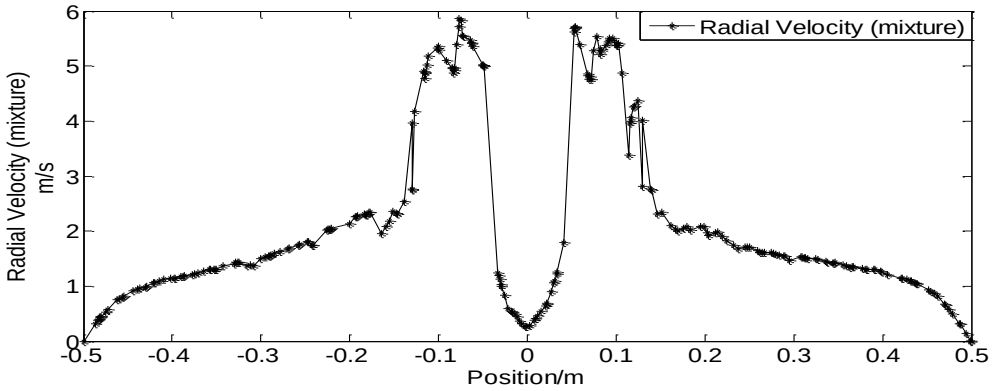


Figure 5. The radial speed distribution of the horizontal nozzles

From Figure 5, an obvious acceleration function of the nozzle can be seen. In the plane of x=0, two nozzles almost distribute equal flow flux, which thus produces the approximately same outlet velocity. Because of the interface of the moving zone and the static zone, the velocity at $y=\pm 0.13$ decreases suddenly and then the velocity declines to 0.35 m/s nearby the wall gradually, which conforms to the velocity attenuation law. One inclined nozzle group distributes flow flux and velocity according to the analogous law in the y=0 plane.

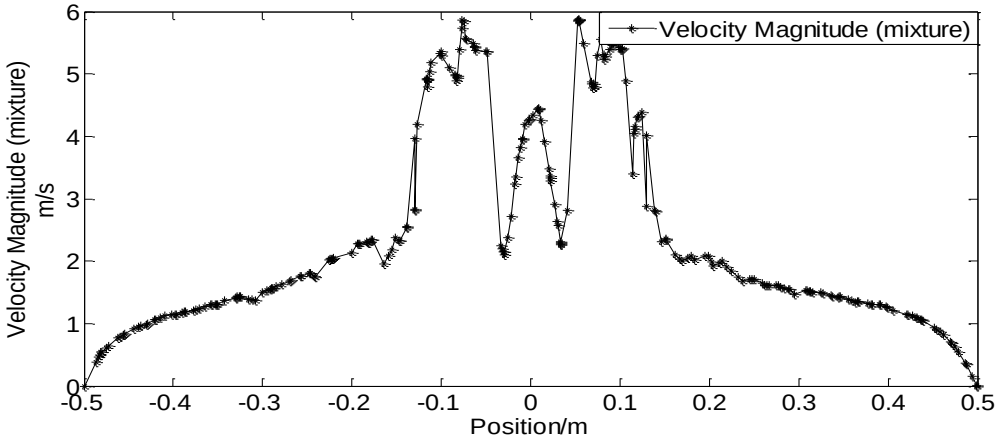


Figure 6. The absolute speed distribution of the horizontal nozzles

In the RJM system, the axial flow and circumferential flow with an order of magnitude of dominate and the radial velocity only appears at the surrounding of the nozzle inlet. The absolute velocity is obtained by combining the three velocities. As illustrated in Figure 6, a minimum velocity of 0.3 m/s can be kept near the wall of the tank, which rightly meets the requirement of gasoline mixing. Across the interface, the velocity decreases about$2/3$, and this is a factor that cannot be ignored in the numerical simulation and that is why the outlet speed needs to be larger than the theoretical calculation values.

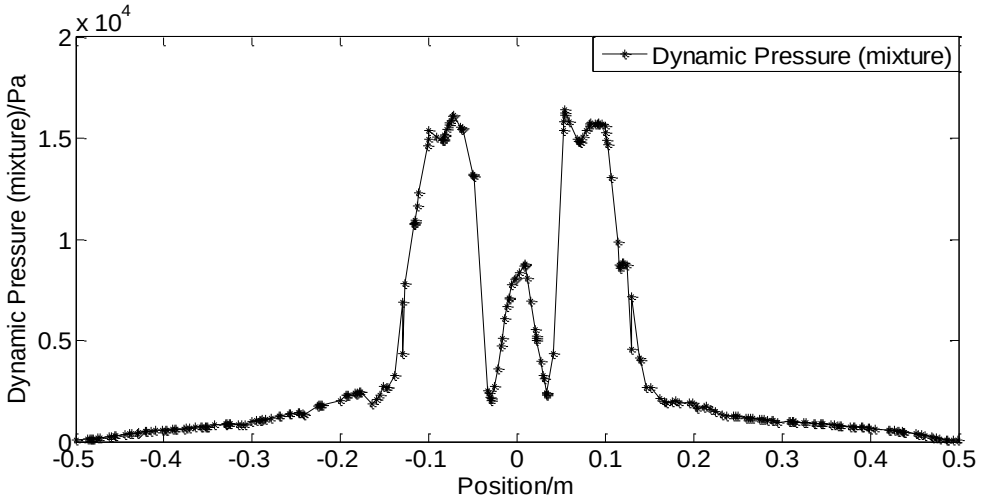


Figure 7. The dynamic pressure distribution of the horizontal nozzles

Figure 7 is quite similar to the absolute velocity distribution in their tendencies and what is different is that the effect of interface on dynamic pressure is more obvious than that on velocity. The dynamic pressure in [-0.05, 0.05] zone is mainly generated by the axial speed of the RJM while the dynamic pressure in other zones is a result of the jet speed of two horizontal nozzles, which justly verified the function of gathering energy and improving pressure of the nozzle.

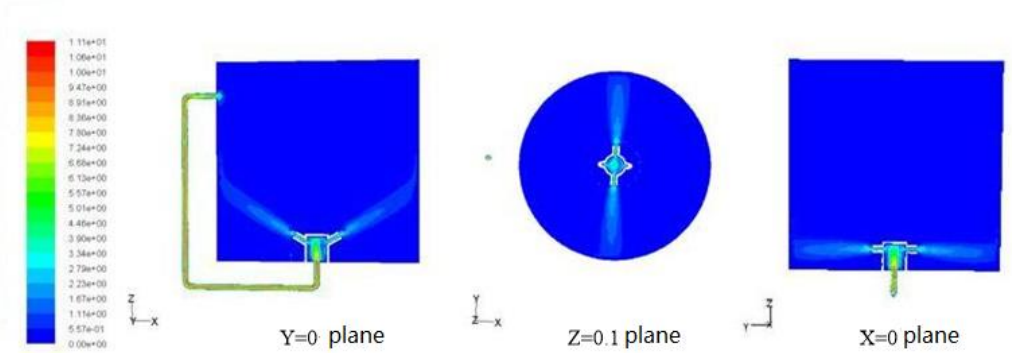


Figure 8. The velocity distribution of the mixing phase in three planes

Figure 8 is the cloud chart of speed attenuation of the horizontal nozzle in x=0 plane and of the inclined nozzle in y=0 plane. The outlet speed of the nozzle is about 5.5 m/s and the speed declined to 0.5 m/s at the wall, which is slightly larger than the required value. Thus, the source term need to be decreased. In Figure 9, the cloud chart of velocity in x=0 plane also showed the effect of gravity on jet speed and the gravity can make the jet trajectory incline to

the bottom of the tank. The speed in the pipe declined slightly and the axial speed declined from 10.5 m/s at the outlet of the source term to 10 m/s at the inlet of the agitator.

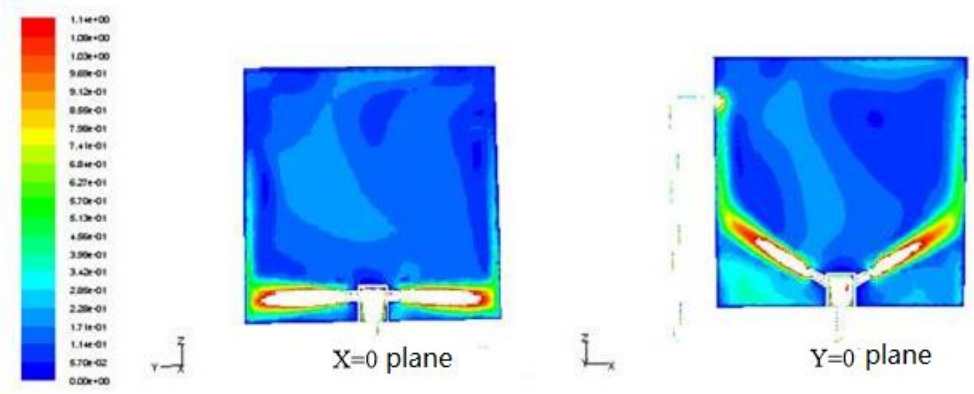


Figure 9. The cloud chart of velocity distribution in two planes

The whole flow field seems to be ideal, but there exists two low-speed zones with narrow regions in the opposition of the outlet pipe and around the RJM system. However, the so-called low-speed zones will disappear with the continuous velocity superposition with the rotation of the RJM system.

In the simulation process, besides the A, B, C monitoring points, plane z=0.1, z=0.5 and z=0.9 are monitored as well. After launching the RJM system, the third phase and the fourth phase began to enter the agitator through the pipe under the action of outlet pipe and the source term. Then these two phases jet into the first phase zone through the nozzle. Before these two phases entered the pipe, they mixed in a certain region in virtue of the speed change. Therefore, among the three planes, the density change firstly appeared in z=0.9 plane and mixing started in the other two planes in 4.4s. From the point of phase, the third and fourth phase mixed in a certain region before entering in the pipe. But in the initial time, the main mixing still happens between the main phase and the third phase with the maximum volume fraction. Finally, the density of the mixing phase approached the equilibrium density 764.5 kg/m ^{3}.

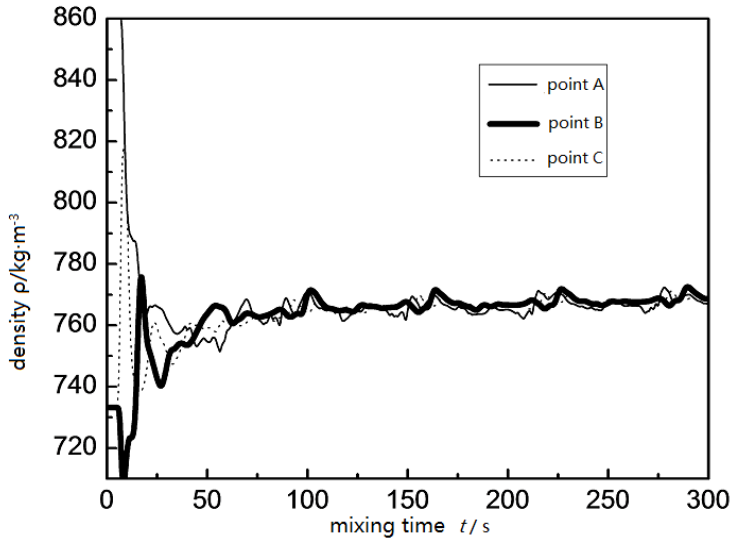


Figure 10. The density-mixing time curves of the three points

In Figure 10, the time of meeting the mixing criterion Formula 1 for the first time is at 72.4s and this is when the density of A, B C is 764.21057 kg/m ³; 763.54425 kg/m ³; 762.06708 kg/m ³; respectively. The density difference of the three is 2.81‰ and soon exceeded the limit within 1%, which is mainly caused by the sensitivity of points to value in 3D space. Hence, the mean density distribution in the three planes need to be checked.

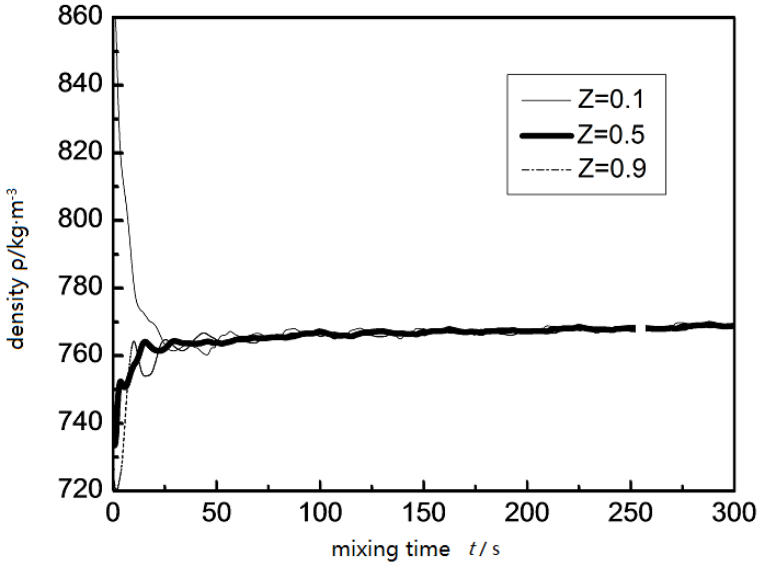


Figure 11. The density-mixing time curves of the three planes

In Figure 11, the density values in the three planes is obtained by averaging the density of all points in z=0.1, z=0.5 and z=0.9 planes respectively, which could eliminate the sensitivity of points to the result errors and could actually reflect the mixing effect of the flow field better.

The mean density value of z=0.1 plane is 761.6552 kg/m ³at 31.2s and z=0.5 is 763.9341 kg/m ³ and z=0.9 is 762.173 kg/m³. The density difference is 2.992‰. And this is the time which met the mixing criterion for the first time. But after 38.8s, it went up and exceeded 3‰. The difference decreased to 2.923‰ (Table 2) again at 58.2s and then kept within 3‰. Therefore, the mixing time of the model could be recognized to be 58.2s.

Table 2. The mean value of density in three planes at 58.2s

Z=0.1 plane (kg/m ³)	Z=0.5 plane (kg/m ³)	Z=0.9 plane (kg/m ³)	Density difference (‰)
766.4078	764.1741	764.4686	2.923

3. Conclusion

The paper analyzed the flow field features of the RJM system agitating in a tank with two 30 ° inclined nozzles and two horizontal nozzles by numerical simulation and three conclusions came to as:

a) A minimum speed of about 0.3m/s nearby the wall of the tank can meet the gasoline mixing

requirements.

b) The whole flow field seems to be ideal, but there exists two low-speed zones with narrow regions in the opposition of the outlet pipe and around the RJM system. However, the so-called low-speed zones disappear with the continuous velocity superposition with the rotation of the RJM system. There is no blind corner in the agitation.

c) By analyzing the monitoring planes, it can be concluded that the mixing time of the RJM system is 58.2s and the mixing efficiency is higher than traditional methods.

Acknowledgement

This project is supported by Science and Technology Development Planning of Shandong Province (2014GGX108001), Key research and development program of Shandong Province (2016GGX104018).

References

[1] Li, Z. X. and Xu, S. W. (1997) *Oil Storage and Transportation Design Manual*, Petroleum Industry Press, Beijing, China.

[2] Yu, Z. B. (2006) Comparison of petrochemical enterprise tank mixing way, *Petroleum Products Application* **4**, 80-83.

[3] Zhang, L. J., Chen, G. G. and Bai, Y. (2013) Numerical simulation on field characteristics of multiphase flow in side-entering stirred tank, *Chinese Journal of Environmental Engineering* **7,** 1594-1600.

[4] Song, J. L., Chen, G. J. and Wang, J. (2015) Numerical simulation of solid-liquid two-phase flow in fermentation tank, *Energy Conservation* **5**, 22-25.

[5] Peng, J. H. and Min, S. Z. (1983) Side agitator in the application of the tank, *Petro-chemical Equipment Technology* **1**, 1-11.

[6] Hou, X. L (2001) *China's oil refining technology*, China Petrochemical Press, Beijing, China.

[7] Zhang, H. L., Cui, F. Y. and Wu, Y. B. (2010) Mechanical mixing pool three dimensional flow field numerical simulation research, *China Water & Wastewater* **26**, 65-67.

[8] Tian, Y. L., Huang, L. Q. and Pan, D. J. (2008) Full flow numerical simulation and analysis of rotary jet mixing system, *Electrical and Mechanical Engineering* **25,** 86-90.

[9] Fossett H. (1951) The action of free jets in mixing of fluids, *Transactions of the Institution of Chemical Engineers* **29a,** 322-332.

[10] Maruyama T, Ban Y. and Mizushina T. (1982) Jet mixing of fluids in Tanks, *Journal of Chemical Engineering Japan* **15,** 342-348.

[11] Zhu, Y. S. and Chang, D. Z. (1983) Multi-nozzle tank blending of gasoline products, *Oil & Gas Storage and Transportation* **2**, 22-25.

[12] Ranade J J. (1996) Towards better mixing protocols by designing spatially periodic flows, *Chemical Engineering Science* **51**, 2637-2642.

[13] Masoud Rahimi and Arsalan Parvareh. (2005) Experimental and CFD investigation on mixing in a semi-industrial stirred tank, *Chemical Engineering Journal* **115**, 82-92.

[14] Meng, H. B., Wang, Y. F. and Yu, Y. F. (2012) Research progress of mixing time in jet mixing equipment, *Chemical Industry and Engineering Progress* **31**, 2615-2625.

[15] Wang, L. Q., Lin, S. D. and Tian, Y. L. (2008) Study on great flow nozzle performance with CFD method, *Fluid Machinery* **36**, 17-22.

[16] Wang, J. L. (2007) Hydraulic drive rotary jet mixer internal flow numerical simulation and research on

hydraulic characteristics, Zhejiang University, Hangzhou, China.

[17] Zhang, Q. H., Fu, Z. Q. and Wan, S. T. (2012) Research of oil product blending spinning nozzle, *Process Equipment& Piping* **49**, 39-43.

[18] Barekatain H and Hashemabadi S H. (2011) Improving of mixing by submerged rotary jet (SRJ) system in a large industrial storage tank by CFD techniques, American Institute of Physics, 123-126.

[19] AAL Neyestanak, G Asadi and S Daneshmand. (2016) Introducing new relation to estimate mixing time of crude oil tank having submerged rotary jet mixer, *Revue Roumaine De Chimie* **61** 37.

[20] Zhong, Y. P., Chen S. Y. and Wang R. Y. (2016) Research on the gas-liquid two-phase flow in the lime slurry pool based on rotary jet mixing, *Journal of Engineering Thermo-physics* **37**, 1875-1883.

Pattern Matching for Industrial Object Recognition Using Geometry Based Vector Mapping Descriptor

D.S. Pae[1]*, Y.T. You[2], T. K. Kang[3] and M.T. Lim[1]†

[1]School of Electrical Engineering, Korea University of Seoul, Republic of Korea.
[2]Department of Mechanics, Korea University of Seoul, Republic of Korea.
[3]Department of Human Intelligence and Robot Engineering University, Cheonan, Republic of Korea.

*Presenting author: paeds915@korea.ac.kr
†Corresponding author: mlim@korea.ac.kr

Abstract

This paper introduces pattern matching algorithm from industrial camera based on geometric features. In industry production line, object recognition using vision is a challenging problem. In the object recognition, feature is an important element that represents object's state. Although its large amount of information contains location, rotation and scale difference, geometric feature is hard to get because it is sensitive to noise. To overcome the weakness, we propose two types of geometric features. Then geometric features are detected in order to construct descriptor. The geometry based Vector Mapping Descriptor (VMD) for pattern matching is proposed to effectively match salient feature points between different images under geometric transformation regardless of missing or additional feature points. VMD represents the correlation of features that includes Euclidean distance and angle. The group of one to one corresponding feature points on different images results in the completed object matching. The proposed algorithm is invariant to translation, rotation and scaling difference. To demonstrate the performance of the proposed algorithm, we conducted an experiment with both reference image data and real-time industrial camera. The result provides accurate and robust feature matching.

Keywords: Pattern Matching, Industrial Camera, Geometric Feature, Vector Mapping Descriptor

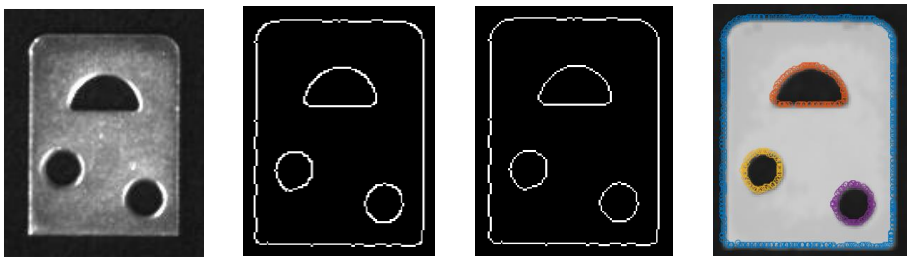
Introduction

Recently, numerous studies on pattern matching have been conducted in the industrial area. Some researches focus on complex image scene like crowded place and many buildings in city. These algorithms are based on obtaining certain point (feature) from image pyramid. Therefore the features are robust for each image pyramid. Lowe [1] suggested Scale Invariant Feature Transform (SIFT) algorithms. Ker et al. [2] improved SIFT algorithm by implementing the Principal Components Analysis (PCA) to the normalized gradient patch. Mikolajczyk et al. [3] proposed an extension of the SIFT descriptor and named Gradient Location and Orientation Histogram (GLOH). Other previous studies performed the object recognition through the shape based retrieval approach. Ling et al. [4] suggested the inner distance shape context (IDSC) which is extension of shape context (SC) [5]. SC is defined as comparing similarity of shape contexts for corresponding points. The inner distance is defined as the shortest path within edge of object shape and results in invariant shape articulation.
Matching can be performed by image segmentation method. Zhang and Ji [6] proposed image segmentation using a Bayesian network to object detection. Ferrari [7] also proposed contour segments for object detection. Dynamic Programming was used for distorted and occluded object retrieval by Petrakis [8].
However the previous methods are disadvantage to simple image like industrial component recognition where its background is mono-color and simple object. Firstly the current algorithms are too heavy. In industrial, the speed is one of the most important issues.

However the previous works did not satisfy the demand. The background is mono-color and the object is simple which indicates that there is less salient feature information. This causes the problem when the objects are occluded. They cannot build the relation between occluded objects due to the lack of number of features. Occlusion can cause building the insufficient relation. Therefore it leads to miss object matching.
In this paper, we address the challenge of improving the efficiency and reliability of object matching in image processing. Pre-processing for object detection is applied on input images at a lower level of abstraction and its purpose is to reduce undesired outliers and enhance the useful image data which is important for further processing. The feature extraction process is an important procedure of object matching system. The feature contains unique, relevant information relating to the model object and target object. The geometry based Vector Mapping Descriptor (VMD) for pattern matching is proposed for object matching. In the proposed method, VMD for each feature is developed in order to obtain corresponding feature points between the model image and target image. Object matching is performed by the constructed descriptor by finding corresponding feature points that cope with both distortion and occlusion and it is invariant to geometric transformations.
The rest of this paper is organized as follows. Section 2 introduces the preprocessing by edge-enhancement using image sub-pixeling. Section 3 proposes two types of geometric features. Section 4 represents the VMD algorithm between model and target. In section 5, our experimental result is conducted with three different cases on both image data set and real-time camera. Section 6 discusses the result on previous experimental section and concludes proposed algorithm with the future works of what is left and how to improve our system.

2. Contour Detection in Sub-Pixel Range

The geometric feature is obtained from the contour of object. Thus, we need to detect contour to get geometric information. Contour detection process has three steps. First step is edge extraction. In edge extraction, we use canny edge algorithm. Second step is edge thinning. Result of edge extraction has shape of thick line. It is hard to get geometric information from thick line. Edge thinning method makes thick line to thin line. The third step is edge linking. Edge linking method links the edge and neighbors. Linked edges are contours of the object.
General method, explained earlier, has low accuracy and hard to extract geometric feature from contour. So we add two processes for improving performance. We apply sub-pixel approximation to increase accuracy. Afterward, we take a process called edge enhancement. Edge enhancement is required to get smoother edge for later stage's geometric feature extraction.



(a) Input image (b) Edge extraction (c) Thinning (d) Edge linking

Figure 1. Contour detection progress

2.1 Sub-pixel

Edge detection algorithms such as Sobel Operator, Robert's cross operator, Prewitt's operator calculate the edge's position information as pixel units. However pixel units have limitations in detecting accurate features such as circles and lines. The sequence of pixels becomes straight instead of becoming smooth. Application of sub pixel [9] units in image processing algorithms is suggested to improve the accuracy of position information.

2.2 Edge Enhancement

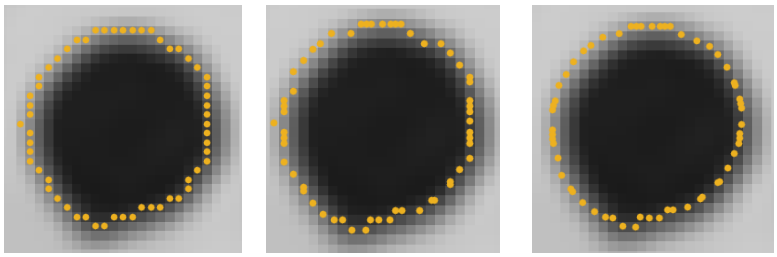
The detected sub pixel unit edge is more suitable than pixel unit edge but it can be processed further to smoother edge. Therefore the edge is improved for geometric feature extraction. We

proposed the method which revises the location of the edge point by using the relation of the neighbor points. In this case, we use two fitting methods that are line fitting and circle fitting [10]. Each fitting equation is shown as follows in (1) and (2).

$$E_l = \frac{1}{N}\sum_{i=1}^{N}\frac{(ax_i + by_i - 1)^2}{a^2 + b^2}. \tag{1}$$

$$E_c = \frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{(x_i - c_x)^2 + (y_i - c_y)^2} - r\right)^2. \tag{2}$$

where N is number of edge points, edge point pixel position (x_i, y_i), secondary line equation parameters a and b, circle center position (c_x, c_y) and circle radius r. Calculate fitting errors and choose smaller one. If revise distance is longer than one pixel, the edge point is fixed.



(a) result of edge linking (b) sub-pixeling (c) edge enhancement

Figure 2. Results on sub-pixeling and edge enhancement

3. Geometric Feature Extraction

Geometric features are extracted from contour of the object. One object has contour set $\mathbf{C} = \{C_1, \cdots, C_k\}$. Contour element C_i has edge points set $\mathbf{E}_{c_i} = \{C_i(x_1, y_1), \cdots, C_i(x_n, y_n)\}$. To present the object, we need the features that are accurate and robust at environment. In this section we define the geometric features that are used for creating Vector Mapping Descriptor. We define two types of geometric features: (a) circle and its center, and (b) line segments and their intersection points. This section is divided into two parts as two main geometric features are employed. We begin with detecting the circle information (a). Once this step is completed, remaining contours are examined to determine whether it is a line or not. The defined line segments are extended imaginarily to produce the intersections among the lines used for salient feature points.

3.1 Circle Feature Detection

The least square circle fitting method is used to represent the circle along the contour. According to the following equation (2) we calculate parameter to minimize the circle fitting error. If circle fitting error is smaller than user parameter, we define the center of circle as geometric feature point. If not, we consider the contour non-circular regions.

3.2 Line Feature Detection

Once inspection of circles from the object is completed, the remaining contour regions are indicated as lines. The least square line fitting algorithm (1) is used to detect line segments. Firstly one point $C_i(x_p, y_p)$ is selected from the remainder contour. The selected point is the starting point of the line segments if the point and its sequence of the neighbor points $\{C_i(x_p, y_p), C_i(x_{p+1}, y_{p+1}), \cdots, C_i(x_{p+q}, y_{p+q})\}$ construct the line, L_k by using equation (1). The

line needs to be formed by satisfying equation (1) within condition of initial number (q) of neighbor point. Once the starting point $C_i\left(x_p, y_p\right)$ is established, the initial line, L_k is formed and its seed starts to grow, adding neighbor points, until the equation (1) is not satisfied. This procedure is repeated until all of the remaining contours are inspected. The noise causes the separation of one long line segments into several line segments. The line is compared with neighbor lines by equation (1) so that it can be combined or one left to be separated. Therefore the line segments are completely detected. However the line fitting still contains error. We need to minimize the errors by shortening the length of the line until it has least amount of errors. The shortening process is done by cutting both side of line segment points until it has minimum error. The parameter can be determined to minimum range of cutting of line segment points.

Now we have line segments with least error. However, the shortening process causes the line segments to have different length which means they have different starting point and ending point. These different results in line segments cause the instability in vertex and these features are unable to be used as salient feature. Thereby we apply the intersection of lines as salient feature. This creates another concept of vertex. The defined extended line segments cause the intersections. The intersection points are robust regardless of differences in length of line segments thus can be used as salient feature.

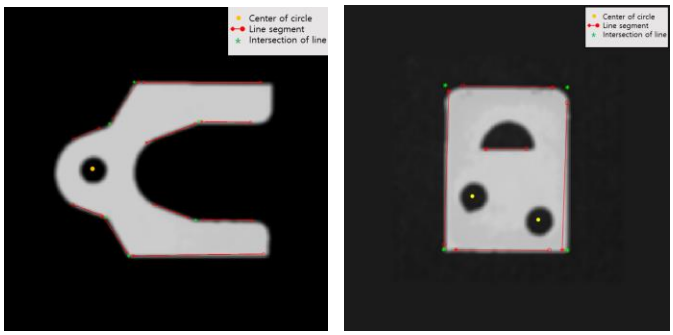


Figure 3. Results on geometric feature extraction

4. Constructing Vector Map Descriptor and Matching Algorithm

In this section we describe the three steps of our proposed geometric matching method: constructing the descriptor, one to one feature matching and finding desired point with completed object matching. Firstly we need the descriptor that is unique identity for an object. Then we compare each descriptors from one feature to the other features. Lastly, the position of objects in target is calculated according to the user's setting point from object on model.

4.1 Constructing Vector Mapping Descriptor

Descriptor is defined as containing unique information of the object. This descriptor is used as recognizing the objects in different scene. The target objects could be single or multi. Our descriptors are created from geometrical feature F which defined at the Chapter 3. The relationship between the defined features is represented as vectors which are separated to form of distance descriptors and angle descriptors. Figure 4 represents the feature locations and feature vectors contain distances and angles.

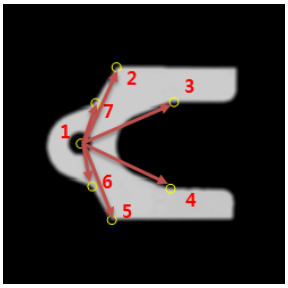


Figure 4. Vector mapping descriptor

Vector between two feature points can represent as follows:

$$\vec{v}_{ij} = F_i - F_j. \tag{3}$$

where $i \in (1,\cdots,n)$ and $j \in (1,\cdots,n)$, and n represents the number of features. Its distance descriptor D and angle descriptor Θ have n x n matrix size. The element of distance descriptor d_{ij} and angle descriptor θ_{ij} are denoted as follows:

$$d_{ij} = \left\|\vec{v}_{ij}\right\|,\ \theta_{ij} = \angle\vec{v}_{ij}. \tag{4}$$

4.2 One to One Feature Matching Using VMD

Object matching indicates that one feature in object can only have one corresponding feature in another scene. Thus, the corresponding feature set should only contain unique correspondences between features. Object matching indicates that one feature in an object can only have one corresponding feature in another object. Thus, the corresponding feature set should only contain unique correspondences between features. To achieve one to one feature matching, we use the row matrices of object descriptor that are feature descriptor distance $\bar{D}_i$ and feature descriptor angle $\bar{\Theta}_i$. In this paper, we use subscript M to state model variable, e.g. I_M, subscript T to state target variable, I_T. The rotation and scaling relationship for arbitrary features F_M from model to feature F_T in target are denoted respectively, by

$$\sigma = d_{M_{ij}} / d_{T_{ij}},\quad \Delta\theta = \theta_{M_{ij}} - \theta_{T_{ij}}. \tag{5}$$

Once the arbitrary scale factor and rotation angle are calculated, assume that the target image I_T is enlarged by σ and rotated by $\Delta\theta$.

The feature descriptor distance and the feature descriptor angle are similar to the enlarged and rotated feature descriptor distance and feature descriptor angle.

Vector Mapping Descriptor algorithm

for i from 1 to N
  for j from 1 to N
    $\sigma = d_{M_{ij}} / d_{T_{ij}},\ \Delta\theta = \theta_{M_{ij}} - \theta_{T_{ij}}$
    $E_{matching} = \sum_{k=1}^{N} \sqrt{\left(d_{M_{ik}} - \sigma d_{T_{ik}}\right)^2 + \left(\theta_{M_{ik}} - \theta_{T_{ik}} - \Delta\theta\right)^2}$
    save feature pair in array
  end
  Find feature pair (F_P) having minimum matching error
  if F_P < Threshold
    Calculate matching position, angle and scale.
  end
end

4.3 Completed Object Matching and Desired Detection Point

The geometric relations between the set of corresponding feature points in two different images are derived from the rotational angle and distance. Once all sets of corresponding feature points are obtained, it is compared with number of feature points in model image to the number of feature points in target image. The features in the model image are standard. The desired point (x_T, y_T) at target image can be calculated from each corresponding feature points; the x direction and y direction vectors of model, $\vec{v}_{xi}$ and $\vec{v}_{yi}$, rotation angle, $\Delta\theta$ and feature points from target $\left(F_{T_x}, F_{T_y}\right)$. The desired point from the original image can be detected in transformed image respectively by,

$$\begin{bmatrix} x_T \\ y_T \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{bmatrix} \begin{bmatrix} \vec{v}_{xi} \\ \vec{v}_{yi} \end{bmatrix} + \begin{bmatrix} F_{T_x} \\ F_{T_y} \end{bmatrix}. \tag{6}$$

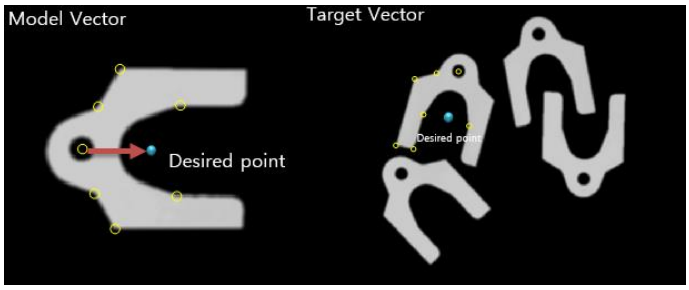


Figure 5. Desired detection point in target object

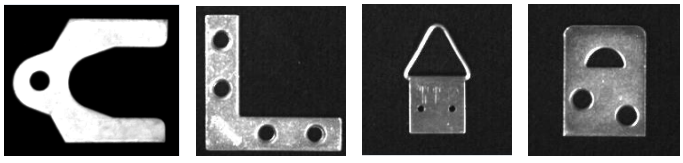
5. Experimental Result

5.1 Experimental Environment

The experimental environment is set as industrial inspection system; mono-color background is used and simple object is recognized. Mono-color background leads the clear separation between the background and foreground. The clear separation also helps the clear edge detection in preprocessing on section 2. The clear edge has close relationship on feature accuracy on section 3 and this results from the desired feature, center of circle and intersection of lines, detection as mentioned on section 3. To reduce the time consuming we limited the intersection of line feature to be formed on the edge and its neighbor only. Finally the Vector Mapping Descriptor algorithm is performed by using the relation between the geometric features.

5.2 Experimental Matching Result

We performed the experiment with four different model images and tested with 20 different target images. The average errors for each models are calculated to evaluate the proposed algorithm that is provided in Table 1. The total process times are shown in Table 2.



(a) Model 1 (b) Model 2 (c) Model 3 (d) Model 4

Figure 6. Model image

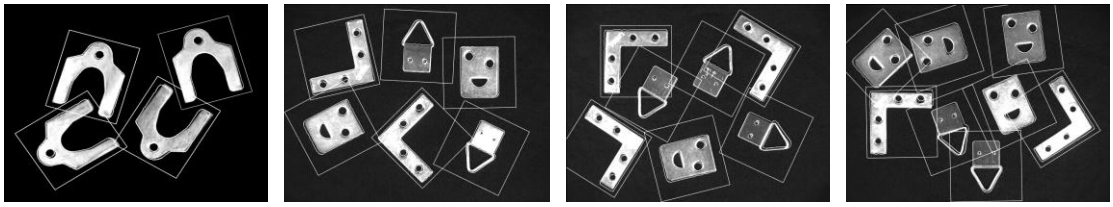


Figure 7. Results on target image

Table 1. Average error in geometric accuracy

Model	*X-axis (pixel)*	*Y-axis (pixel)*	*Scale (rate)*	*Angular (deg)*
Model 1	0.4838	1.0302	0.0132	1.0135
Model 2	0.5625	0.6150	0.0240	5.4371
Model 3	1.1881	1.4383	0.0094	0.9606
Model 4	0.8402	0.7481	0.0072	0.4835

Table 2. Processing Time

Processing step	*Average Computation Time (s)*
Filter	0.1062
Edge detection	0.0464
Edge linking	0.4586
Feature extraction	0.0842
Matching	0.0536
Total	0.749

The result of Table 1 shows efficiency of our method. In X-axis and Y-axis location accuracy, all average errors are lower than 1.5 pixel. Also, total process time is shorter than 1 second as shown in table 2.

6. Conclusion

In this paper, we start from demonstrating how to obtain fine images from coarse images. It results in obtaining robust geometric features. The strong geometric features are introduced for creating VMD algorithm which is proposed to solve the object recognition problem. Based on the results, it is shown that the proposed matching method is invariant to arbitrary geometric transformation: translations, rotations, scale changes, extra or missing feature points and the occlusion. The speed of matching algorithm is also fast enough for industrial recognition system. Moreover the VMD algorithm, proposed algorithm can be performed with any case if the feature of object in model and target are similar.

In future, we have a plan to improve the performance of VMD algorithm. We will find the optimal parameters for recognition system. Our research will focus in 3D data sets and their occlusion as well. The distance and angle descriptor extend to x, y and z axis can be one of the solution. Representing 3D image to 2D image by projection is also considered as solution. Image matching and registration are the foundation for many computer vision and its application such as navigation and security surveillance by recognizing the desired objects.

References

[1] Lowe, David G. "Distinctive image features from scale-invariant keypoints." International Journal of Computer Vision, 2004, Vol. 60, No.2, pp. 91-110.

[2] Ke, Yan, and Rahul Sukthankar. "PCA-SIFT: A more distinctive representation for local image descriptors." Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, 2004, Vol. 2, pp. 506-513.

[3] Mikolajczyk, Krystian, and Cordelia Schmid. "A performance evaluation of local descriptors." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, Vol. 27, No. 10, pp. 1615-1630.

[4] Ling, Haibin, and David W. Jacobs. "Shape classification using the inner-distance." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, Vol. 29, No. 2, pp. 286-299.

[5] Belongie, Serge, Jitendra Malik, and Jan Puzicha. "Shape matching and object recognition using shape contexts." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, No. 4, pp. 509-522.
[6] Zhang, Lei, and Qiang Ji. "Integration of multiple contextual information for image segmentation using a bayesian network." Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, 2008, pp. 1-6.
[7] Petrakis, Euripides G. M., Aristeidis Diplaros, and Evangelos Milios. "Matching and retrieval of distorted and occluded shapes using dynamic programming." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, No .11, pp. 1501-1516.
[8] Ferrari, Vittorio, et al. "Groups of adjacent contour segments for object detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, Vol. 30, No. 1, pp. 36-51.
[9] Canny, John. "A computational approach to edge detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, Vol. 6, pp. 679-698.
[10] Gander, Walter, Gene H. Golub, and Rolf Strebel. "Least-squares fitting of circles and ellipses." BIT Numerical Mathematics, 1994, Vol. 34, No. 4, pp. 558-578.

Hemodynamic effects of lesion length on competitive flow with internal mammary artery bypass

Boyan Mao[1], Huixia Zhang[1], Wenxin Wang[1], Bao Li[1], Xiaoyan Zhang[1], * Youjun Liu[1]

[1] College of Life Science and Bio-Engineering, Beijing University of Technology ,China.

*Corresponding author: lyjlma@bjut.edu.cn

Abstract

Background: The left internal mammary artery (LIMA) is the most selected vessel in coronary artery bypass graft, because of its long-term patency. But the influence of flow coming from incomplete stenosis of native coronary artery which is called competitive flow is an important factor for graft failure. There are many factors that affect the competitive flow such as the stenosis severity, the amount of myocardial tissue subtended by the stenosis, the length of lesion, as well as the afterload etc. This article is aimed to discuss the hemodynamic influence of different length of intermediate lesion to competitive flow.

Materials and Methods: A patient-special 3-dimensional model with internal mammary artery graft was constructed. The research made a 50% stenosis of six different lengths in the left anterior descending (LAD), including 3mm, 5mm,10mm,15mm,20mm,25mm. Geometric multiscale analysis method which couples the lumped parameter model(LPM) and three-dimensional (3D) model was used in the numerical simulations.

Results: From the results of numerical simulation, graft flow, coronary flow, wall shear stress(WSS) and the oscillatory shear index(OSI) of graft were extracted. With the increase of the lesion length, the mean graft flow and graft WSS increased, in addition, the coronary flow and graft OSI decreased, but when the stenosis reached 15 mm, these variables were stable and almost had no change.

Conclusion: Apart from the stenosis rate, competitive flow is also affected by the length of lesion. Because of the presence of competitive flow, it should be put more attention to choose bypass surgery for moderate stenosis. But below 15mm, with the increase of the length of stenosis, the hemodynamic environment in bypass surgery is becoming better. In summary, for moderate stenosis, stenosis length is a very important factor in the choice of treatment options.

Keywords: competitive flow; coronary artery; the lumped parameter model; WSS; OSI

Introduction

In anatomy, there is no significant correlation between coronary artery stenosis and

myocardial ischemia. Moderate severity of coronary artery disease may lead to ischemia and coronary heart disease. When LIMA is selected, the string sign is the most common phenomenon appearing in the early postoperative period. Most scholars believe that the string sign is caused by the not serious stenosis [1]-[5]. It's reported that the string sign which is described as diffuse distal narrowing of the graft and usually occurs early in the postoperative period has relationship with competitive flow [1][6][7]. As early as 1972, Bousvaros et al [8] proposed the competitive flow. Since then, many scholars studied the influence of competitive flow to graft patency. Nordgaard et al [9] defined the competitive flow as the flow competing each flow from the partially stenosed native coronary artery and from the arterial graft. Numerous studies have shown that the blood flow from native coronary artery reduces arterial graft flow and its long-term patency [10]-[13]. Thus, competitive flow is an important factor affecting early arterial graft patency. [14][15] Many researchers have found that competitive flow may lead to adverse hemodynamic environment, such as low WSS, high OSI which can induce the intimal hyperplasia and atherosclerosis [16]-[18]. Currently, there are a number of researchers dedicated to the competitive flow evaluation. For research of competitive flow, there are three kinds of common methods: statistical analysis of clinical data, animal experiment and numerical simulation analysis. Nakajima H et al [19] believe that a minimum of coronary artery stenosis can be used to measure the intensity of competitive flow. Joseph F.Sabik et al [20] found that the minimum diameter and the coronary artery graft patency has a good correlation. In addition, through clipping the target vessel, graft blood flow changes represent the intensity of the competitve flow [21]. Li lanlan et al [22] studied the hemodynamic effects of competitive flow at varying degree of stenosis. They found that a stenosis rate of less than 50% will cause bad hemodynamic effects, such as low WSS, high OSI. In addition, when the stenosis percentage reaches 50%, there will be inverse flow. The severity of native coronary artery stenosis is not the only factor affecting the hemodynamic parameters, the anastomotic configuration, morphometric parameters and afterload are other major factors affecting the hemodynamics. The stenosis length is also common.

The length is one of determinants of reclusion [23]. Brosh D [24] et al hold the view that lesion length has a significant impact on the physiological significance of intermedia-grade coronary lesions. Fractional Flow Reserve (FFR) is the "gold standard" which determines whether coronary artery stenosis causes myocardial ischemia. According to statistical analysis, it is found that the area of stenosis and lesion length are the most influential factors [25]. Other scholars hold the view that with the degree of moderate stenosis and lesion length increases, FFRCT value will gradually decline to 0.80 or less, indicating that moderate degree of disease and lesion length can also cause myocardial ischemia, which is consistent with clinical results. Even some scholars have shown that when the lesion length is greater than 10mm, there is a strong collection between FFRCT and lesion length [26]. Apart from the degree of lesion stenosis, lesion length is another considerable geometric parameter in the morphological assessment of coronary ischemia, the impact of lesion length on competitive flow has never been adequately assessed. The purpose of this study was to investigate the relationships between competitive flow and lesion length.

Methods

Reconstruction of 3D models

The patient's CT data obtained from a 55-year-old male. A 50% diameter stenosis with different length (3mm,5mm,10mm,15mm,20mm,25mm) was set in the LAD, respectively, by using 'Freeform' (software of the 3D modeling system) and 'PHANTOM DESKTOP' (a kind of force feedback device). In addition, the LIMA was constructed in the process of 3D model establishment. One of the final 3D models are illustrated in Figure 1.

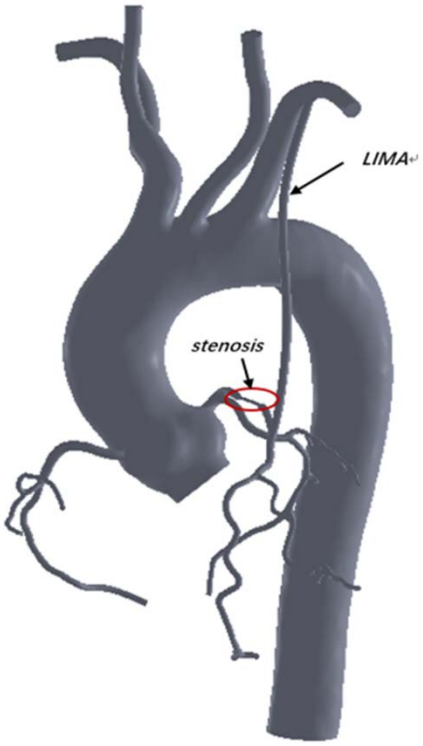


Figure 1. The 10mm stenosis 3D model with LIMA which was anastomosed to the LAD. Seven different length of stenosis were made by hands in the red mark location.

Meshing is the basis of numerical simulation of 3D model, high-quality meshing can guarantee the accuracy of simulation. The computational models were generated by meshing the 3D sub-models with hexahedral mesh. This process is done by the CFX module of commercial software ANSYS. The nodes and elements numbers of the 3D models are shown in Table 1.

Table1. The nodes and elements numbers of the 3D models

Lesion length	***3mm***	***5mm***	***10mm***	***15mm***	***20mm***	***25mm***
nodes	994196	999085	984951	1018705	977169	1017709
elements	1296410	1308740	1290611	1327942	1278619	1334527

Lumped parameter model

Construction of Lumped Parameter Model

Individualized Lumped parameter model can provide physiological boundary conditions for 3D simulation calculation, which cannot be obtained from clinical data such as MR and CT etc. The most common lumped parameter model is the Windkessel model, which simulates downstream resistance and compliance of vessel [27]. The three-element Windkessel model consists of proximal resistance(R_p), a capacitor(C), and a remote resistor(R_d), as shown in the following Figure 2.

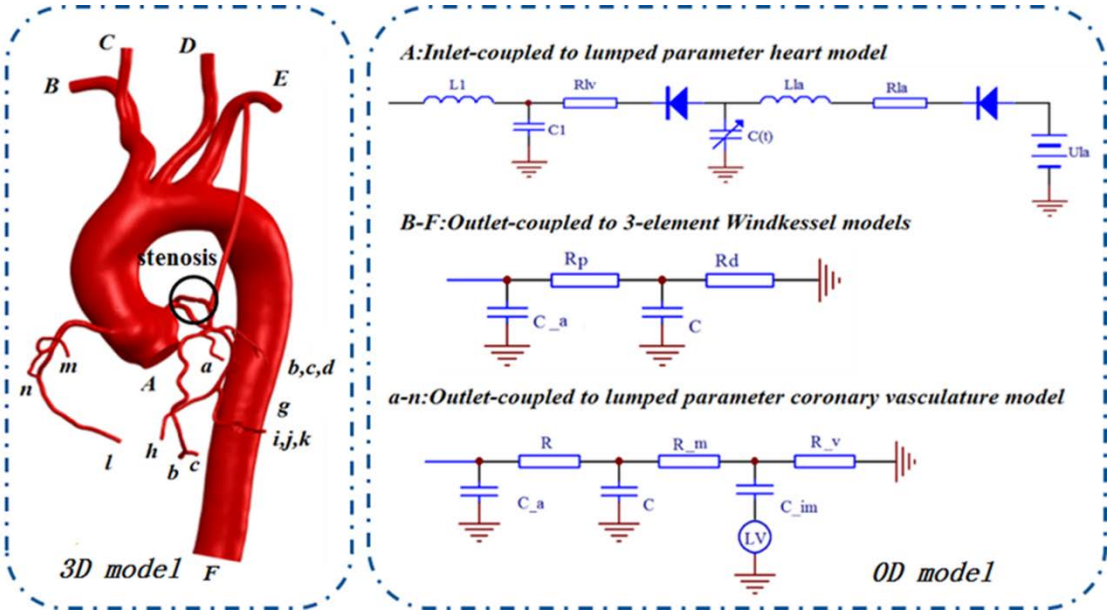


Figure 2. Part A which was coupled with the right aorta lumped parameter model is the inlet. Part B to F are the outlets, and they were coupled to the right middle 3-element Windkessel model. From a to n are also outlets, and they were coupled to the right lumped parameter coronary vasculature model

In this circuit model, the relationship between pressure and flow is shown as Eq.(1), where Q_t is the outlet flow and P_t is the outlet pressure.

$$\mathrm{P}(t) = \left[P(0) - R_p Q(0)\right] e^{-\frac{t}{R_d C}} + R_p Q(t) + \int_0^t \frac{-\frac{t-\tilde{t}}{R_d C}}{C} Q(\tilde{t}) d\tilde{t} \qquad (1)$$

The lumped parameter model of the heart is also shown in the following Figure 2. In the model, the heart valve has a diode to ensure its one-way flow, and the variable capacitance simulates the contraction and relaxation of the heart. Rlv and Rla represent left ventricular blood flow resistance and left atrium resistance, respectively. In the process of cardiac contraction, the coronary artery wall is squeezed, and blood flow decreased significantly, in the diastolic, coronary return to natural state. Based on this factor, a coronary lumped parameter model is created as shown Figure 2, where R, R_m, C, C_{im}, R_v, U_{lv} indicate the coronary artery resistance, microcirculation resistance, the compliance of the vessel, compliance of myocardial, coronary vein and its microcirculation resistance and myocardial pressure, respectively. The complete lumped parameter model is shown in Figure 3.

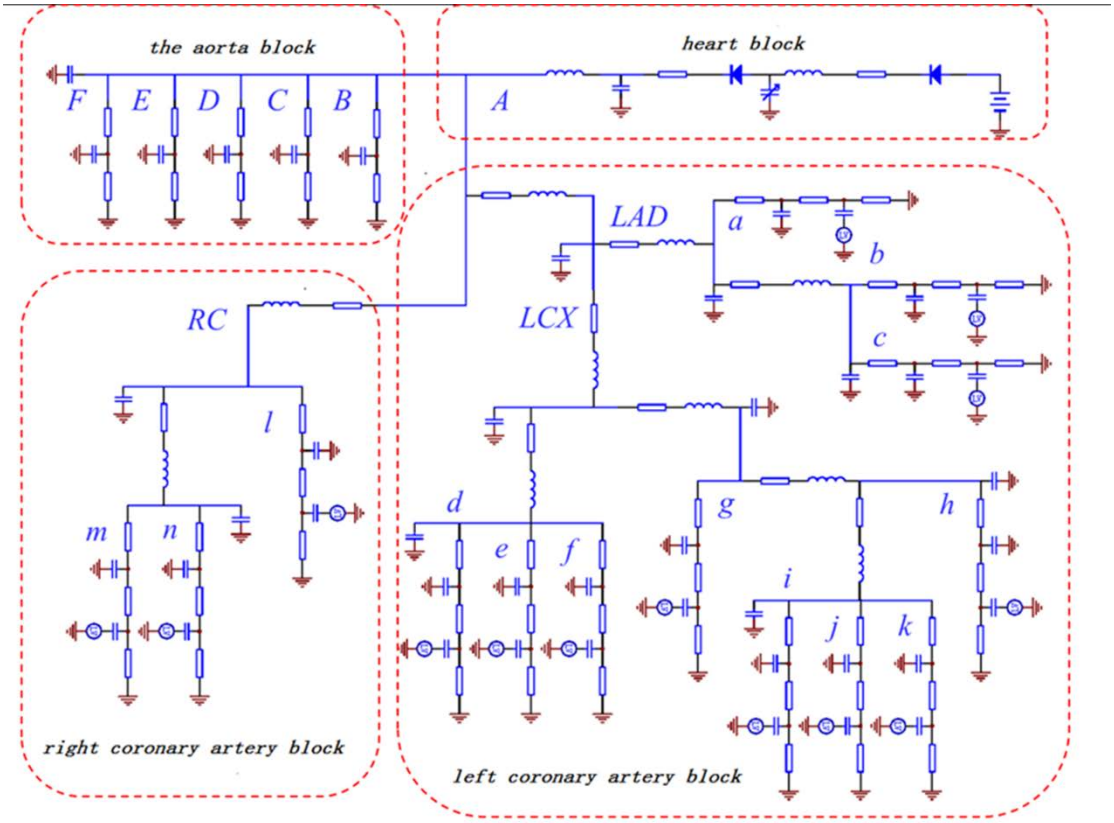


Figure 3. Lumped parameter cardiovascular system model

Parameter Determination of Lumped Parameter Model

In the compartment of the ventricle, the cardiac cycle of the left ventricles was demonstrated by the function of pressure-volume relationship which has been determined by several animal experiments (shown as Eq. (2))

$$\mathrm{E}(t) = \frac{P(t)}{V(t)-V_0} \tag{2}$$

where $\mathrm{E}(t)$ is the time-varying elastance $(mmHg/ml)$, $\mathrm{V}(t)$ and $\mathrm{P}(t)$ are the ventricle volume (ml) and pressure $(mmHg)$, respectively , with the reference volume (ml). Mathematically, the function is used as the approximation (shown as Eq. (3) and Eq. (4))

$$\mathrm{E}(t) = (E_{max} - E_{min}) \cdot E_n(t_n) + E_{min} \tag{3}$$

$$E_n(t_n) = 1.55\left[\frac{\left(\frac{t_n}{0.7}\right)^{1.9}}{1+\left(\frac{t_n}{0.7}\right)^{1.9}}\right]\left[\frac{1}{1+\left(\frac{t_n}{1.17}\right)^{21.9}}\right] \tag{4}$$

where $E_n(t_n)$ is the normalized time varying elastance and $t_n = t/t_{max}$, $t_{max} = 0.2 + 0.15tc$, and tc is one cardiac cycle interval(s). In this paper, we set $E_{max} = 2.0$, $E_{min} = 0.002458$ and $\mathrm{tc} = 0.8\mathrm{s}$.

E_{max} and E_{min}, respectively , are related to end of cardiac systole and end of diastole pressure and volume. Tayor et al [28] have shown that there is a certain relationship between flow distribution and vessel diameter. Based on this theory, by using the genetic algorithm, these values were adjusted to make sure that the cardiac output, the systolic pressures

matched the patient's data. (Table 2)

Table2. The list of model predictions compared against clinical data

	Model Predictions	***Clinical Data***
Systolic Pressure	147.69*mmHg*	147*mmHg*
Diastolic Pressure	103.48*mmHg*	103*mmHg*
Cardiac Output	4.58*L/min*	4.6*L/min*

0D/3D coupling calculations

The algorithm of coupling used in this study has been used in previous studies [29][30]. In the coupling process, the 0D model provides the outlet pressure and the inlet flow rate for the 3D model, and the 3D model provides the inlet pressure and the outlet flow for the 0D model. The coupling problem for each time step between the two dimensional models is calculated using the formula (5) (6). This cycle is calculated at each time step until the result is determined to be convergence.

$$\bar{P}_{3D,in} = \frac{1}{A_{3D,in}} \int_{\Gamma_{in}} P d\gamma = P_{0D,in} \tag{5}$$

$$Q_{3D,out} = \rho \int_{\Gamma_{out}} u \cdot n_i \, d\gamma = Q_{0D,out} \tag{6}$$

A_{3D}, Q_{3D}, Q_{0D} is the interface area, flow of 3D model and interface flow of 0D model, respectively whose unit is cm^2, ml/s, ml/. μ is the interface velocity and P is the average pressure at the interface with their units are cm/s, Pa. Γ_{in} and Γ_{out} is defined of the import and export interface of 0D and 3D.

The convergence judgment condition is as shown in Eq. (7) (8).

$$\left| P_{out}^{(k)} - P_{out}^{(k-T_k)} \right| \leq \varepsilon \tag{7}$$

$$\left| P_{in}^{(k)} - P_{in}^{(k-T_k)} \right| \leq \varepsilon \tag{8}$$

In the 3D simulation, the vessel is assumed to be rigid. The blood flow was treated as an incompressible viscous Newtonian fluid. The density of the blood flow was 1050 kg/m^3, the dynamic viscosity is 0.0035Pa.s.

Results

Graft flow

In coronary artery bypass surgery, the graft flow is a very important indicator that determines whether the operation is successful. The time-averaged graft flow of different stenosis length

in one cycle which were extracted from the middle of the graft are listed in Figure 4.

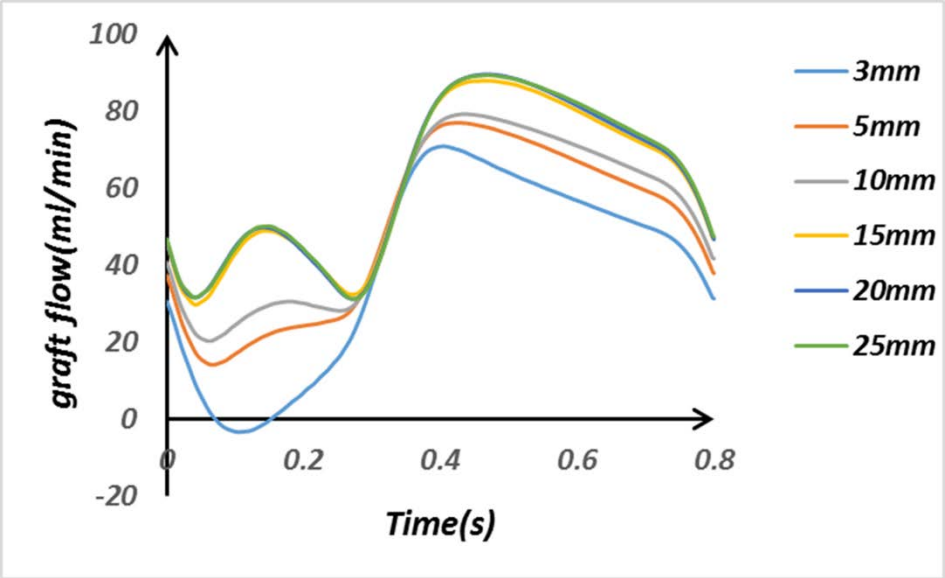


Figure 4. The flow of LIMA with different stenosis length in one cycle

The average flow of graft in one cycle varies with the length of stenosis as shown in Figure 6. It can be seen that the inverse flow appears during the systole in the 3mm stenosis. But with the increase of stenosis length, the inverse flow disappeared. In addition, as the degree of stenosis increases, the graft flow increases, but when the stenosis reaches 15mm, the flow becomes no longer change basically.

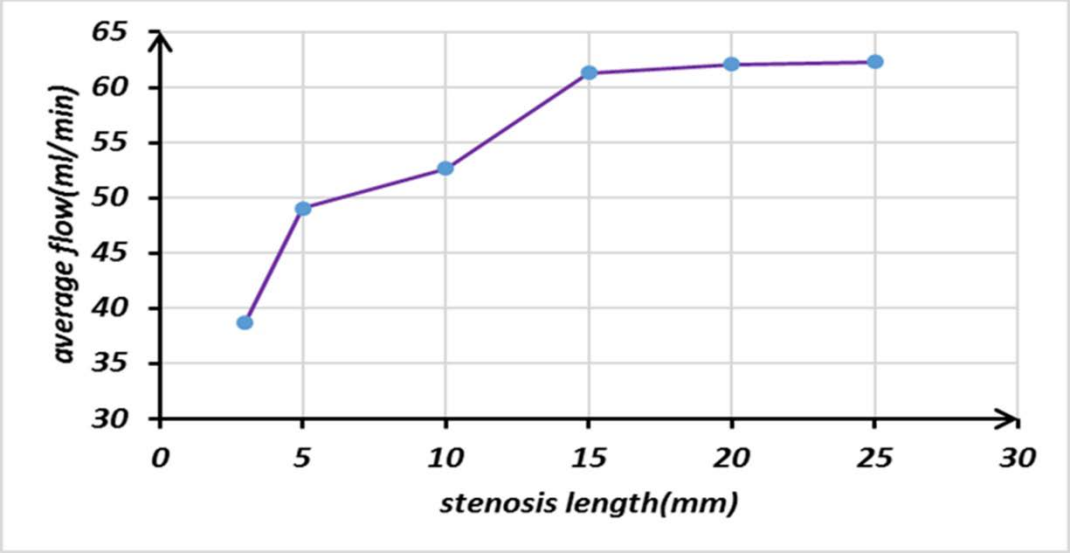


Figure 5. The changing trend of time-average graft flow with different length

Competitive flow

The average flow of the graft and stenosis are shown in the following Table 3. $R_{C/G}$ (Eq.9) is a quantity that characterizes the degree of the competitive flow. The trend is shown in Figure 6. At 3 mm, the graft flow is almost the same as that at the stenosis, indicating that the competitive flow is obvious. This situation is likely to result in the string sign, that is, the failure of the graft. With the increase of the stenosis length, the competitive flow declined,

Table 3. Time-average graft flow, coronary flow and $R_{C/G}$

Lesion length	*3mm*	*5mm*	*10mm*	*15mm*	*20mm*	*25mm*
Q_{LIMA} *(ml/min)*	38.729	49.039	52.633	61.333	62.061	62.289
Q_{LAD} *(ml/min)*	41.375	30.981	27.345	17.907	17.693	17.413
$R_{C/G}$	1.068	0.632	0.520	0.292	0.285	0.280

but when the length reaches 15mm, the competitive flow becomes basically no longer change.

$$R_{C/G} = \frac{Q_{LAD}}{Q_{LIMA}} \tag{9}$$

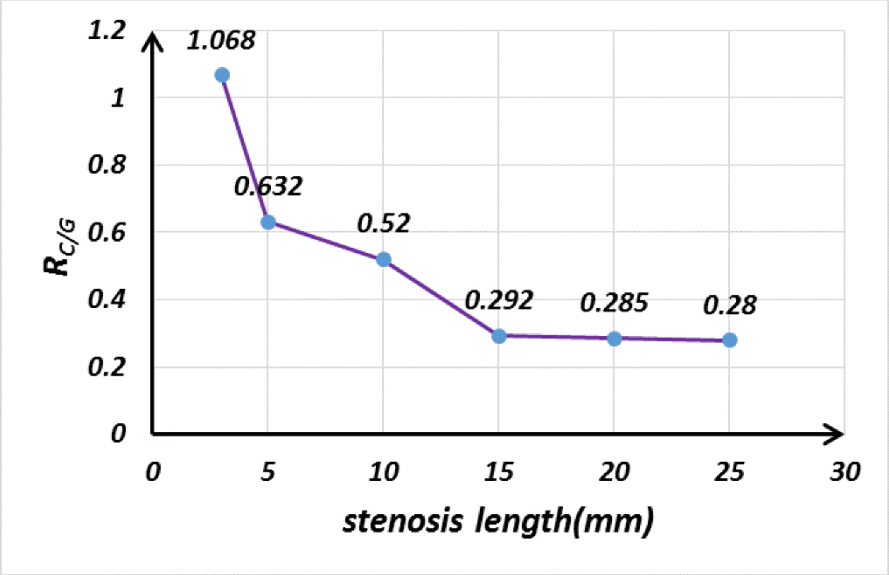


Figure 6. $R_{C/G}$ for LIMA in each model

Wall shear stress

Many studies have found that wall shear stress is an important factor that relates the blood flow to the localization of atherosclerosis which is characterized by localized plaques that form within the artery wall and the plaque rupture. According to the Poiseuille theory, WSS is proportional to blood viscosity and blood flow, inversely proportional to the diameter, the formula is as follows Eq. (10) and Eq. (11). The average wall shear stress for graft vessel in one cycle is shown below Figure 7. The wall shear stress contour at the peak systolic flow rate were extracted (Figure 8).

$$\text{wss} = \tau_{\omega} = -\mu \frac{du}{dr} \tag{10}$$

$$\text{wss} = -\mu \left[\frac{2Q}{\pi R^4} (-2r)|_{r=R} \right] = \frac{4\mu Q}{\pi R^3} \tag{11}$$

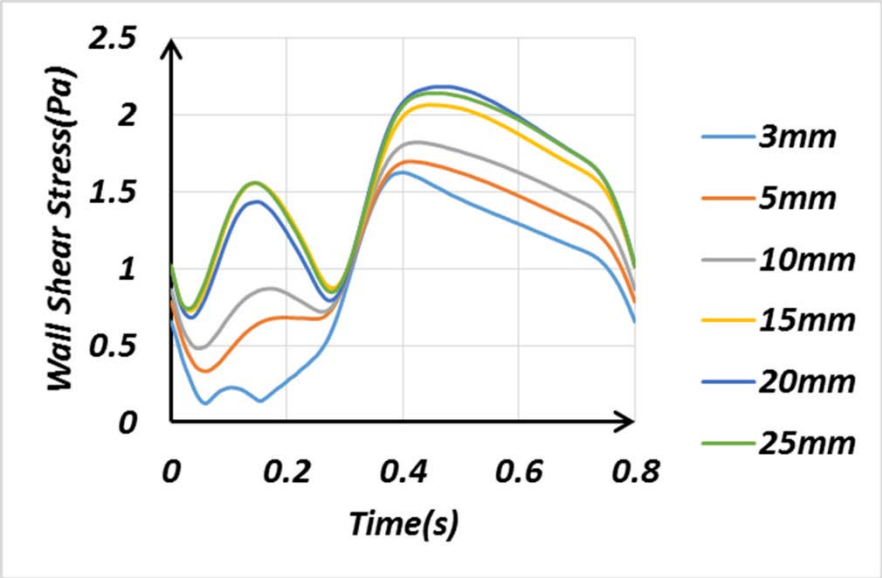


Figure 7. The graft WSS of different stenosis length in one cycle

Oscillatory shear index

The OSI is a known predictor of formation of atherosclerosis and vulnerability for plaque in coronary arteries. The OSI for the pulsatile low simulation is calculated as [29]

$$\mathrm{OSI} = 0.5\left(1 - \frac{\left|\int_0^T \overrightarrow{\tau_\omega} dt\right|}{\int_0^T \left|\overrightarrow{\tau_\omega}\right| dt}\right) \tag{12}$$

where $\overrightarrow{\tau_\omega}$ represents wall shear stress, dt means differential time, and T is upper limit of integral. Table 4 shows the average OSI value of the LIMA surface. It can be seen that when the stenosis length increases from 3mm to 25mm, the OSI decreases from 0.0190 to 0.0042. As the degree of stenosis increases, the OSI gradually decreases. When the stenosis reaches 15mm, OSI changes very little.

Table 4. The OSI values of the six grafts

Lesion length	*3mm*	*5mm*	*10mm*	*15mm*	*20mm*	*25mm*
OSI	***0.019***	***0.0076***	***0.0054***	***0.0045***	***0.0043***	***0.0042***

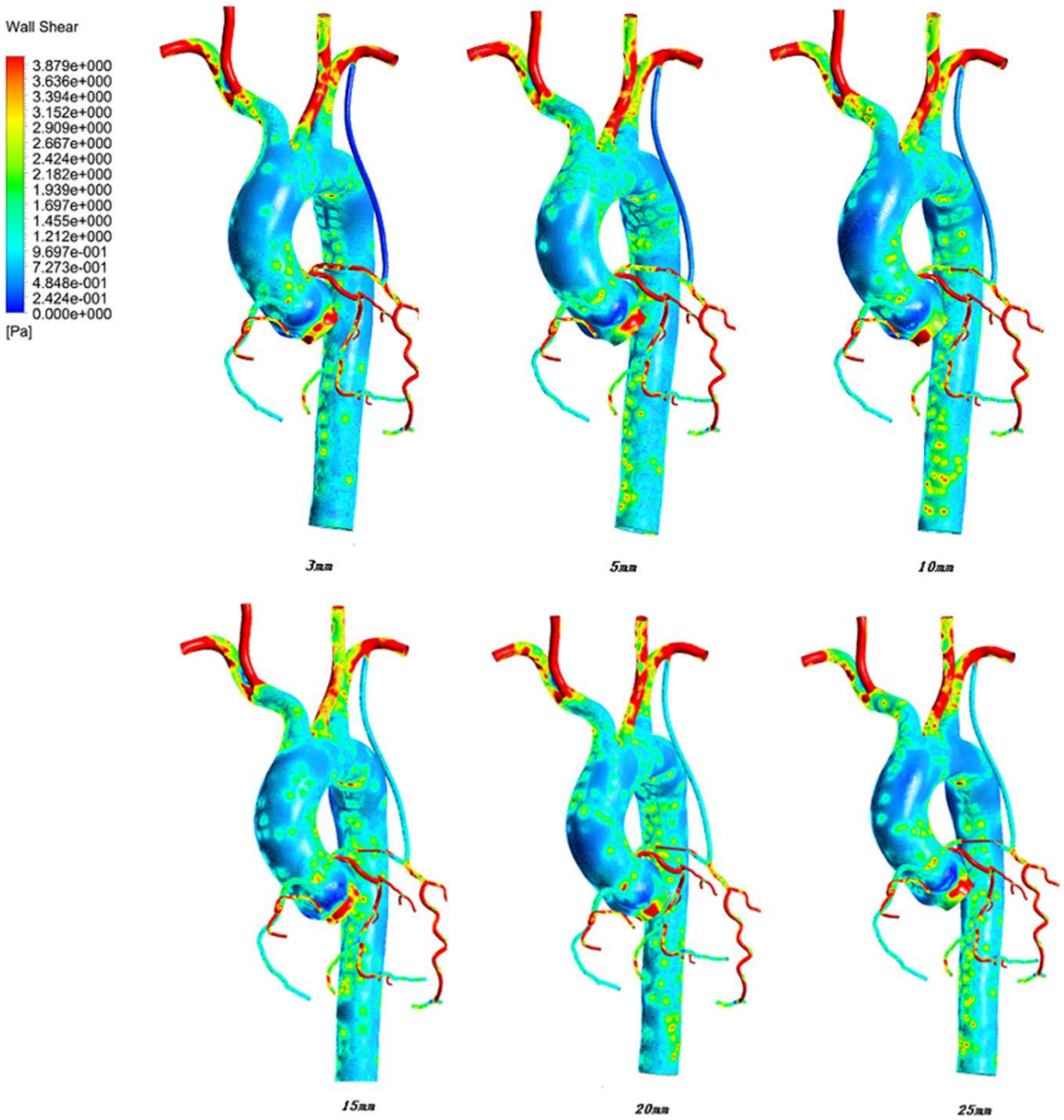


Figure 8. The WSS contours at the moment of maximum graft flow.

Discussion

Competitive flow

Inverse flow is a bad hemodynamic phenomenon that may lead to string sign. In clinical surgery, the severity of inverse flow is a factor for doctors to judge the surgical results. The inverse flow appeared in the 3mm stenosis, but in other lengths, the inverse flow disappeared. With the increase of the stenosis length, the average blood flow in the upstream of the anastomosis decreased continuously, but when the stenosis reached 15 mm, the blood flow remained at a steady state basically. Compared with the results of Li lanlan [22], when the lesion length reached 15mm, the graft flow is similar to the degree of 75% stenosis. In other words, when the lesion length increases to a certain value, it can also produce the same degree

of severe stenosis results. Because of the differences in the coronary of each individual, the correspondence between the lesion length and the degree of stenosis is also different in terms of the competitive flow. Therefore, facing with the treatment choice of moderate narrow, the narrow length is a decisive factor.

WSS&OSI

Zhao et al. hold the view that the high OSI might be the major hemodynamic characteristic leading to the string sign in an LIMA graft. In this study, we found that the wall shear stress has a peak in the systolic period. When the length is less than 15mm, the wall shear stress increases with the increase of the stenosis length, but it does not change after 15mm. So in a certain range of stenosis length, the longer the stenosis, the better the long-term patency of the graft. From the OSI point of view, when the lesion length reaches 15 mm, the OSI value does not change, but before 15mm, the longer the length, the graft OSI is smaller.

Limitation

The limitations of the article are mainly reflected in the following aspects. Firstly, for the boundary conditions of 3D model, there is a gap between model and the physical reality. For example, in order to reduce the calculation time, the vessel wall is set to be rigid, in fact, the properties of blood vessels is very complex, and fluid-structure interaction (FSI) was not considered due to the complex 3D models.

Conclusion

Except for the stenosis rate, the stenosis length is also an important factor influencing the competitive flow, but not linearly. As the lesion length increases, the competitive flow decreases, but when the narrow length reaches 15 mm, the competitive flow does not change.

Acknowledgements

This work is supported by National Science Foundation of China (11172016,11472022).

References

[1] Geha A S. Early and late results of coronary revascularization with saphenous vein and internal mammary artery grafts[J]. American Journal of Surgery, 1979, 137(4):456-463.

[2] Hashimoto H, Isshiki T, Ikari Y, et al. Effects of competitive blood flow on arterial graft patency and diameter - The Journal of Thoracic and Cardiovascular Surgery[J]. Journal of Thoracic & Cardiovascular Surgery, 1996, 111(2):399-407.

[3] Kolozsvari R, Galajda Z, Ungvari T, et al. Various clinical scenarios leading to development of the string sign of the internal thoracic artery after coronary bypass surgery: the role of competitive flow, a case series[J]. Journal of Cardiothoracic Surgery, 2012, 7(1):1-9.

[4] Susumu M, Toshihiro F, Tomoki S, et al. Increased graft occlusion or string sign in composite arterial grafting for mildly stenosed target vessels. [J]. Annals of Thoracic Surgery, 2010, 89(3):683-687.

[5] Beyersdorf F. The internal mammary artery 'string phenomenon.' Analysis of 10 cases[J]. European Journal of Cardio-Thoracic Surgery, 1993, 7(5):235-238.

[6] Barner H B. Double Internal Mammary-Coronary Artery Bypass[J]. Archives of Surgery, 1974, 109(5):627-630.

[7] Ivert T, Huttunen K, Landou C, et al. Angiographic studies of internal mammary artery grafts 11 years after coronary artery bypass grafting. [J]. Journal of Thoracic & Cardiovascular Surgery, 1988, 96(1):1-12.

[8] Bousvaros G, Piracha A R, Chaudhry M A, et al. Increase in severity of proximal coronary disease after successful distal aortocoronary grafts. Its nature and effects. [J]. Circulation, 1972, 46(5):870-879.

[9] Nordgaard H, Swillens A, Nordhaug D, et al. Impact of competitive flow on wall shear stress in coronary surgery: computational fluid dynamics of a LIMA-LAD model. [J]. Cardiovascular Research, 2010, 88(3):512-519.

[10] Nordgarrd H, Nordhaug D, Kirkeby-Garstad I, et al. Different graft flow patterns due to competitive flow or stenosis in the coronary anastomosis assessed by transit-time flowmetry in a porcine model [J]. European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery, 2009, 36(1):137-142.

[11] Run-Wei M A, Ruo-Bin W U, Guo H M, et al. Relation Between Competitive Flow and Graft Flow in Coronary Artery Bypass Grafting[J]. Chinese Journal of Clinical Thoracic & Cardiovascular Surgery, 2005.

[12] Kawamura M, Nakajima H, Kobayashi J, et al. Patency rate of the internal thoracic artery to the left anterior descending artery bypass is reduced by competitive flow from the concomitant saphenous vein graft in the left coronary artery[J]. European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery, 2008, 34(4):833-838.

[13] Berger A, Maccarthy P A, Siebert U, et al. Long-term patency of internal mammary artery bypass grafts: relationship with preoperative severity of the native coronary artery stenosis. [J]. Circulation, 2004, 110(11 Suppl 1):36-40.

[14] Shumway S J, Sabik, Tatoulis J, et al. Does competitive flow reduce internal thoracic artery graft patency? Discussion[J]. Annals of Thoracic Surgery, 2003, 76(5):1497-1497.

[15] Hashimoto H, Isshiki T, Ikari Y, et al. Effects of competitive blood flow on arterial graft patency and diameter - The Journal of Thoracic and Cardiovascular Surgery[J]. Journal of Thoracic & Cardiovascular Surgery, 1996, 111(2):399-407.

[16] Lytle B W, Loop F D, Cosgrove D M, et al. Long term serial studies of internal mammary artery and saphenous vein coronary bypass grafts[J]. Journal of Thoracic & Cardiovascular Surgery, 1985, 89(2):248-58.

[17] Fh. S. A comparison of coronary and internal mammary arteries and implications of the results in the etiology of arteriosclerosis[J]. American Heart Journal, 1983, 105(4):560-6.

[18] Villareal R P, Mathur V S. The string phenomenon: an important cause of internal mammary artery graft failure.[J]. Texas Heart Institute Journal, 1999, 27(4):346-9.

[19] Nakajima H, Kobayashi J, Toda K, et al. Angiographic evaluation of flow distribution in sequential and composite arterial grafts for three vessel disease.[J]. European journal of cardio-thoracic surgery: official journal of the European Association for Cardio-thoracic Surgery, 2012, 41(41):763-769.

[20] Iii J F S, Blackstone E H. Coronary artery bypass graft patency and competitive flow.[J]. Journal of the American College of Cardiology, 2008, 51(2):126-128.

[21] Gil B, Kypson A P, L Wiley N, et al. A technique for evaluating competitive flow for intraoperative decision making in coronary artery surgery.[J]. Annals of Thoracic Surgery, 2003, 76(6):2118-20.

[22] Lanlan L I, Liu Y, Zhao X I, et al. IMPACT OF COMPETITIVE FLOW ON HEMODYNAMICS OF LIMA-LAD GRAFTING WITH DIFFERENT STENOSIS: A NUMERICAL STUDY[J]. Journal of Mechanics in Medicine & Biology, 2016.

[23] de Cesare N B, Ellis S G, Williamson P R, et al. Early reocclusion after successful thrombolysis is related to

lesion length and roughness.[J]. Coronary Artery Disease, 1993, 4(2):159-166.

[24] Brosh D, Higano S T, Lennon R J, et al. Effect of lesion length on fractional flow reserve in intermediate coronary lesions[J]. American Heart Journal, 2005, 150(2):338-343.

[25] Kristensen T S, Engstrøm T, Kelbæk H, et al. Correlation between coronary computed tomographic angiography and fractional flow reserve ☆[J]. International Journal of Cardiology, 2010, 144(2):200-5.

[26] AliAlghamdi, MohammedBalgaith, AbdulazizAlkhaldi. Influence of the length of coronary artery lesions on fractional flow reserve across intermediate coronary obstruction[J]. European Heart Journal Supplements, 2014, 16(suppl_b):B76-B79.

[27] Vignon-Clementel I E, Figueroa C A, Marsden A L, et al. Outflow boundary conditions for three-dimensional simulations of non-periodic blood flow and pressure fields in deformable arteries[J]. Journal of Biomechanics, 2006, 39(06):S431.

[28] Taylor C A, Fonte T A, Min J K. Computational Fluid Dynamics Applied to Cardiac Computed Tomography for Noninvasive Quantification of Fractional Flow Reserve : Scientific Basis[J]. Journal of the American College of Cardiology, 2013, 61(22):2233-2241.

[29] Mori D, Hayasaka T, Yamaguchi T. Modeling of the Human Aortic Arch with Its Major Branches for Computational Fluid Dynamics Simulation of the Blood Flow[J]. Jsme International Journal, 2002, 45(45):997-1002.

In Vivo Intravascular Ultrasound-Based 3D Thin-Walled Model for Human Coronary Plaque Progression Study: Transforming Research to Potential Commercialization

***Jian Guo,[1] Liang Wang,[2] David Molony,[3] Habib Samady,[3] Jie Zheng,[4], Xiaoya Guo,[5] Akiko Maehara[6], Gary S. Mintz[6], Jian Zhu,[7] †Genshan Ma,[7] †Dalin Tang[1,2]**

[1]School of Biological Science & Medical Engineering, Southeast University, Nanjing, China
[2]Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA
[3]Department of Medicine, Emory University School of Medicine, Atlanta, GA, 30307, USA
[4]Mallinckrodt Institute of Radiology, Washington University, St. Louis, MO, 63110, USA
[5]Department of Mathematics, Southeast University, Nanjing, 210096, China
[6]The Cardiovascular Research Foundation, Columbia University, New York, NY 10022, USA
[7]Department of Cardiology, Zhongda Hospital, Southeast University, Nanjing, 210009, China

* Presenting author: Jian Guo, School of Biological Science & Medical Engineering, Southeast University, Nanjing China, email: wyguojian@126.com.

† Corresponding authors. (1) Dalin Tang, Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609, email: dtang@wpi.edu; (2) Genshan Ma, Department of Cardiology, Zhongda Hospital, Southeast University, Nanjing, 210009, China, magenshan@hotmail.com.

Abstract

Cardiovascular disease (CVD) is the leading cause of death in the world. Considerable research has been done linking various risk factors to plaque progression and rupture which often lead to drastic clinical events such as heart attack and stroke. However, methods transforming research results to clinical implementation are limited. There has been evidence indicated that mechanical stress and strain may be linked to plaque progression. However, 3D plaque model construction is extremely time consuming making is near impossible for clinical implementations. 2D structure-only model is easy to make, but its stress/strain predictions are not good enough to serve as approximation to 3D solutions. In this study, an in vivo IVUS based 3D Thin-Wall model was developed to approximate 3D FSI model for clinical implementations. Results from one patient data (100 TW models) indicated that mean value of maximum plaque wall stress (MPWS) and average plaque wall stress (APWS) from TW model were 1.9% and 3.0% higher than those from FSI model (MPWS: 127.2±74.3 kPa vs. 124.8±65.7 kPa , APWS: 65.9±18.8 kPa vs. 64.0±21.4 kPa). Mean value of maximum plaque wall strain (MPWSn) and average plaque wall strain (APWSn) from TW model were 1.3% and 1.6% lower than those from FSI model (MPWSn: 0.0760±0.0221 vs. 0.0770±0.0161, APWSn: 0.0572±0.0056 vs. 0.0581±0.0070). Wall thickness (WT) from TW model was 1.7% lower than those from FSI model (0.0630±0.0126 cm vs. 0.0640±0.0126). At baseline, PWS and plaque progression measured by wall thickness increase (WTI) had no significant correlations using either FSI or TW models (FSI: r = 0.0446, p = 0.7684; TW: r = -0.0414, p = 0.7848). WT had significant negative correlation with WTI (FSI: r = -0.3337, p = 0.0234; TW: r = -0.3559, p = 0.0152). PWSn had weak positive correlation with WTI by FSI model (r = 0.2830, p = 0.0506), while this cannot be predicted by TW model (r = -0.1493, p = 0.3219). Large scale studies are needed to further investigate the feasibility of using TW models as approximation for FSI models in potential clinical implementations.

Key words: Coronary, fluid-structure interactions, plaque progression, vulnerable plaque, IVUS.

1. Introduction

Cardiovascular disease (CVD) is the leading cause of death in the world. It is well known that most of the heart attacks are related to the rupture of vulnerable plaques [1][12]. However, it has been a great challenge to understand the mechanisms governing plaque progression and rupture and predict them before tragical clinical events happen. Interventional surgical treatments such as stenting, carotid endarterectomy (CEA), and graft bridging are routinely used in CVD treatment with artery stenosis severity used as the primary guidance. On the other hand, it has been well documented that stenosis severity does not correlate well with occurrence of CVD clinical events calling for surgical interventions [16][24]. Considerable effort has been devoted to identify possible linkage between mechanical stress/strain conditions and plaque progression and rupture. Since 3D model construction is time consuming, that became a hinder in bringing modeling research results to real clinical applications. In this paper, we demonstrate that 3D thin-walled models could serve as a tool to overcome that difficulty.

Plaque progression and rupture are complex processes involving many factors including mechanical forces, plaque morphology, hemorrhage, cell activities, blood conditions and many others. And it is well believed that the mechanical factors play an important role. With the improvement of medical imaging, many research groups have proposed and improved the computational models based on the clinical image data to obtain more accurate and reliable plaque stress/stain and flow shear stress conditions. The mature one was the multi-component Fluid-Structure Interaction (FSI) model based on the in vivo data [28], which was validated by experimental data on its accuracy and reliability [4][19][22]. As mentioned above, the model construction time limited its application for clinical implementations. Considering other alternatives, 2D structure model could not provide more reliable result, while 3D wall-only model did not reduce the model constructing and solving time [8][26]. Recently, a new thin layer structure only model (TLS) based on ex vivo MR-Image data was introduced, the results from TLS model showed good agreement with FSI model, and the time of model constructing and solving was controlled within 2 hours per plaque, which provided the possibility of its use for clinical plaque assessment [7].

In this paper, thin-wall (TW) and FSI models were constructed based on patient follow-up intravascular ultrasound (IVUS) for a patient. Plaque stress, strain and wall thickness data from both models were compared to find the differences and seek the feasibility of using TW models to replace 3F FSI models for potential clinical applications. Results for plaque progression correlation analysis from both models were also compared.

2. Method

2.1 IVUS data acquisition

Patient follow-up IVUS data of coronary plaques from one patient (Gender: M, Age: 52) were provided by Cardiovascular Research Foundation (Cardiovascular Research Foundation, New York, United States) with informed consent obtained at time of data acquisition. Vessel and multi-components detection was performed to obtain IVUS-VH (Virtual Histology) data using automated Virtual Histology software (version 3.1) on a Volcano s5 Imaging System (Volcano Crop., Rancho Cordova, CA). IVUS-VH slices at the baseline (T1) and follow-up (T2) were processed to obtain contours for vessel lumen, out-boundary and plaque components for model construction using established procedures described in *Yang et al*. [28].

X-ray angiogram (Allura Xper FD10 System, Philips, Bothel, WA) was obtained at both baseline scan (T1) and follow-up scan (T2) prior to the pullback of the IVUS catheter to determine the location of the coronary artery stenosis, vessel curvature and cyclic bending caused by heart contraction. Fusion of angiography and IVUS slices was done using method similar to that described in *Wahle et al.* [25]. The view angle of the angiography was chosen so that the angiography plane was close to the principal tangent plane of the chosen coronary segment. Co-registration (both longitudinal and circumferential) of baseline and follow-up IVUS data were performed by IVUS expert using angiography movie, location of the myocardium, vessel bifurcation, stenosis and plaque component features. Figure 1 gives the patient follow-up IVUS images, contours and reconstructed 3D geometry of the plaque.

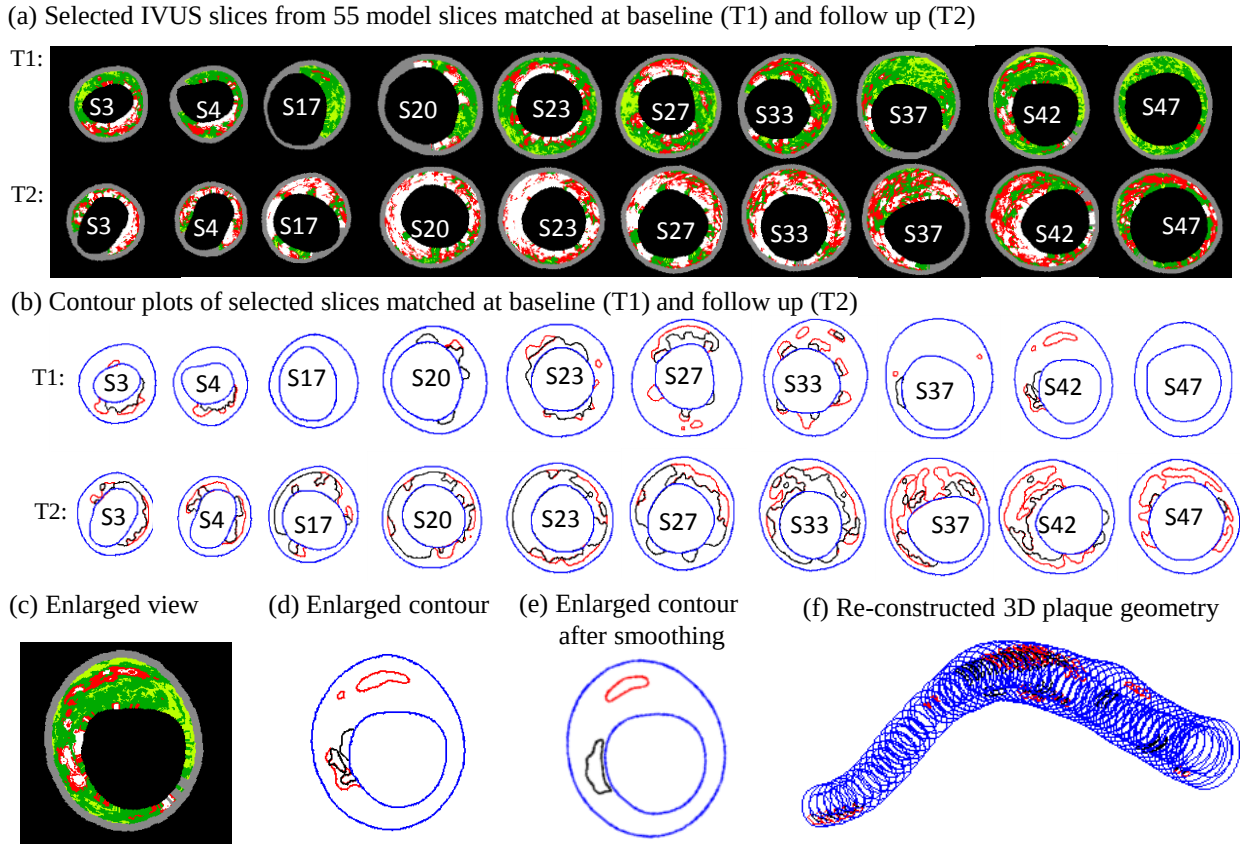


Figure 1. Patient follow-up IVUS data from a patient. (a) Baseline (T1) and follow up (T2) co-registered IVUS-VH and (b) segmented contours; (c) Enlarged view of S42; (d) Enlarged contour of S42; (e) Enlarged contour after smoothing; (f) Re-constructed 3D plaque geometry.

2.2 Pressure condition and material parameters

Patient-specific systolic and diastolic pressure was used to scale a typical coronary blood pressure profile obtained from catheterization and adopted in our model. Figure 2 shows the upstream pressure (Pin) and the downstream pressure (Pout) used in the 3D FSI model. Axial pressure drop was not needed in TW model which is structure-only and had no flow. The uniform pressure (Pin) was applied over the lumen surface of TW model.

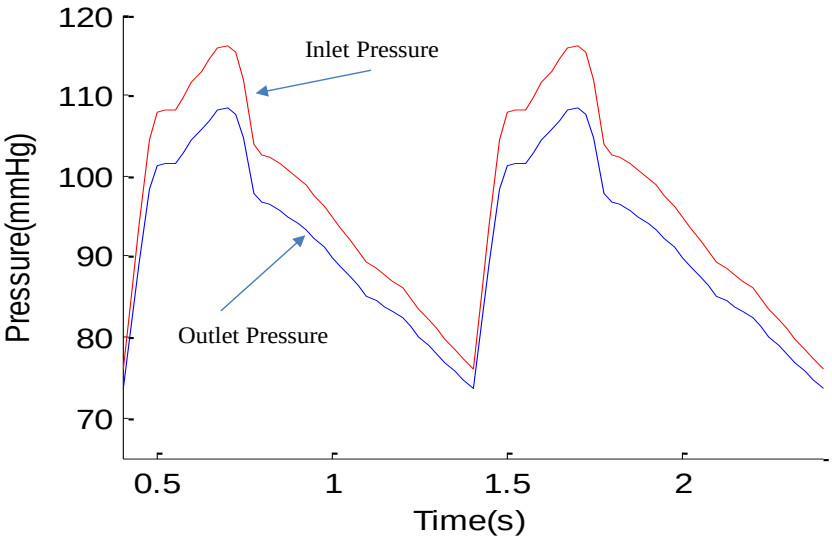


Figure 2. The cardiac pressure profile used for the FSI and TW models.

Vessel material was assumed to be hyperelastic, anisotropic, nearly-incompressible and homogeneous. Plaque components were assumed to be hyperelastic, isotropic, nearly incompressible and homogeneous for simplicity. No-slip conditions and natural traction equilibrium conditions were assumed at all interfaces. Our complete FSI model can be found from *Yang et al.* and are omitted here [28].

Biaxial testing was performed using eight coronary arteries from 4 cadavers (age: 50-81) to obtained realistic vessel material data for our model [9]. A modified Mooney-Rivlin model was used for the vessel fitting our biaxial data: [9][28]

$$W = c_1(I_1 - 3) + c_2(I_2 - 3) + D_1\left[exp\left(D_2(I_1 - 3)\right) - 1\right] + K_1/(2K_2)\, exp[K_2(I_4 - 1)^2 - 1] \quad (1)$$

$$I_1 = \sum C_{ii}, \quad I_2 = 1/2[{I_1}^2 - C_{ij}C_{ij}] \quad (2)$$

where I_1 and I_2 are the first and second invariants of right Cauchy-Green deformation tensor $\boldsymbol{C}$ defined as $\boldsymbol{C} = [C_{ij}] = X^T X$, $\boldsymbol{X} = [X_{ij}] = [\partial x_i/\partial a_j]$ ((x_i) is current position, (a_i) is original position), $I_4 = C_{ij}(n_c)_i(n_c)_j$, n_c is the unit vector in the circumferential direction of the vessel, *c1*, *D1*, *D2*, and *K1* and *K2* are material constants. The parameter values used in this paper were: *c1* = -1312.9 kPa, *c2* = 114.7 kPa, *D1* = 629.7 kPa, *D2* = 2.0, *K1* = 35.9 kPa, *K2* = 23.5. Our measurements are also consistent with data available in the literature [2][3][23][26][27].

2.3 Plaque geometry reconstruction and computational models

About plaque geometry reconstruction, several important techniques and details should be explained. First, since arteries were axially stretched and pressurized under in vivo condition, a preshrink ratio should be found to shrink the in vivo shape to a no-load shape so that when axial stretch and pressure were added, the artery could re-gain its in vivo shape [4]. Second, cyclic bending of coronary arteries caused by cardiac contraction/expansion should be involved in the 3D FSI model to obtain more accurate result. Another major step to construct the computational model was the mesh generation, and the Volume Component-Fitting Method was applied to

generate mesh for the plaque models [5][28]. In that process, the plaque was divided into hundreds of small volumes to curve-fit the complex plaque geometry, and to generate mesh for the finite element model which will be solved by ADINA (Adina R & D, Watertown, MA).

The think-wall (TW) model was made by adding a thin wall thickness to a plaque slice to form a 3D model. Since the TW model does not really have 3D geometry complications, mesh generation is practically the same as 2D models. There were no cyclic bending and fluid part for it. However, compared to 2D structure only models, axial stretch was included in TW model. In this way, TW model reserved some characteristics of 3D model and greatly simplified the complexity of model construction. Figure 3 showed the construction process of TW model.

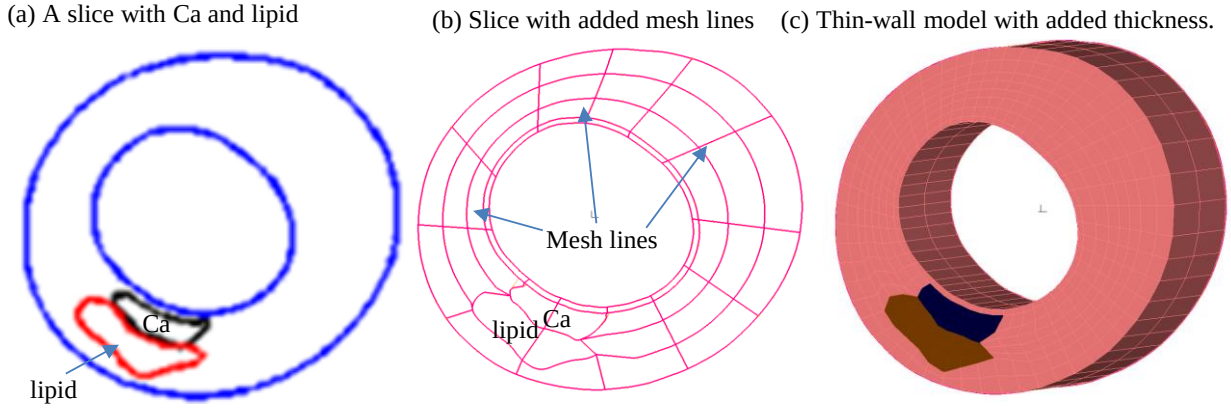


Figure 3. The process of constructing a TW model. (a) A slice with Ca and lipid; (b) Slice with added mesh line; c) Thin-wall model with added thickness.

2.4 Finite element simulation and solution method

The computational models were solved using the commercial finite package ADINA (ADINA R&D Inc., USA), which offers a wide range of capabilities based on reliable and efficient finite element procedures. The governing finite element equations were solved by the modified Newton-Raphson iteration method. More details of the computational models and solution methods can be found in *Tang et al.* [4][17].

2.5 Data analysis

Plaque wall stress (PWS), strain (PWSn), and wall thickness (WT) values were extracted from the solutions from 3D FSI and TW models for analysis. The maximum principal stress and maximum principal strain were used as the scalar values of plaque stress and strain for simplicity. Results on matched lumen nodes from TW model and FSI model were compared to investigate their differences. Using the FSI model results as the bench mark, the relative error of the TW model is defined as:

$$Error = \frac{1}{n} \left(\sum_{i=1}^{n} |a_i - b_i|/b_i\right) \quad (3)$$

where a_i was the values from TW models, b_i was the values from FSI models, and n was the number of data points.

For progression analysis, slices at baseline (T1) and follow up (T2) were paired. Plaque progression was measured by wall thickness increase (WTI) defined as WT at T2 – WT at T1. Correlation results using FSI models and TW models were compared.

3. Results

3.1 Construction time

Technically, TW 3D model is expanded from 2D model. The cost of time on model constructing and solving for a TW model is almost same as that for a 2D model. However, when it comes to FSI model, 3D wall only-model or TLS model, because all of them are 3D models in nature, the cost of time would obviously be more than that for TW model.

For an experienced researcher, it takes less than 10 minutes to perform all steps for one TW model. For the case we are using, there are 50 slices. So 50 TW models took less than 10 hours to construct and obtain all the results. On the other hand, construction for one FSI model or 3D wall-only model normally takes 1-2 week (for a trained operator), which makes it impossible for potential clinic implementations.

3.2 Comparisons of TW models and FSI models on plaque wall stress/strain, and wall thickness

Figure 4 gives mean value bar plots of plaque wall strain (PWSn) at T1 and T2 from FSI and TW models showing comparisons. Table 1 shows plaque wall stress (PWS), plaque wall strain (PWSn), and wall thickness (WT) from FSI and TW models. Mean value of MPWS and APWS from all TW models (100 models) were 1.9% and 3.0% higher than those from FSI model (MPWS: 127.2±74.3 kPa vs. 124.8±65.7 kPa , APWS: 65.9±18.8 kPa vs. 64.0±21.4 kPa). with relative errors were 11.9% and 6.7%, respectively. And mean value of MPWSn and APWSn from all TW models (n=100) were 1.3% and 1.6% lower than those from FSI model (MPWSn: 0.0760±0.0221 vs. 0.0770±0.0161, APWSn: 0.0572±0.0056 vs. 0.0581±0.0070). Their relative errors were 10.5% and 4.9%.

With regard to WT, TW models showed good agreement with FSI models. The mean value of WT obtained from TW models was 1.7% lower than that from FSI model (0.0630 ± 0.0126 cm vs. 0.0640±0.0126), and the relative errors was 1.7%.

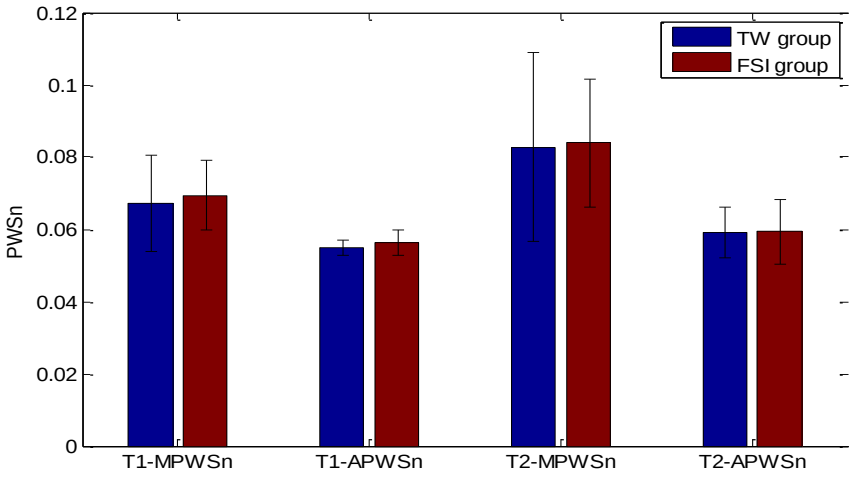


Figure 4. Comparison of PWSn between TW model and FSI model at baseline (T1) and follow up (T2).

Overall results suggested that the values of various factors provided by TW model had slight differences with those by FSI model, most of them kept the relative error lower than 10%.

Table.1 The comparison of PWS, PWSn and WT between FSI model and TW model

Factors	Period	Slices	TW	FSI	Error(%)
MPWS(kPa)	T1	49	97.4±37.2	94.5±24.4	11.3
	T2	50	150.5±89.3	150.7±78.1	13.3
	combined	92	127.2±74.3	124.8±65.7	11.9
APWS(kPa)	T1	49	59.9±6.8	57.5±6.3	5.6
	T2	50	70.0±24.9	68.6±28.5	8.3
	combined	92	65.9±18.8	64.0±21.4	6.7
MPWSn	T1	49	0.0671±0.0134	0.0694±0.0096	9.6
	T2	50	0.0828±0.0261	0.0840±0.0177	12.2
	combined	92	0.0760±0.0221	0.0770±0.0161	10.5
APWSn	T1	49	0.0549±0.0021	0.0564±0.0036	4.3
	T2	50	0.0591±0.0070	0.0593±0.0090	5.5
	combined	92	0.0572±0.0056	0.0581±0.0070	4.9
WT(cm)	T1	49	0.0636±0.0131	0.0640±0.0127	0.7
	T2	50	0.0628±0.0120	0.0645±0.0121	2.7
	combined	92	0.0630±0.0126	0.0640±0.0126	1.7

Combined data was based on the paired slice (92). Values are expressed as mean±standard deviation.

3.3 Correlation results between plaque progression (WTI) and risk factors using TW models and FSI models

Table 2 summarized the correlation results between plaque progression and three risk factors (PWS, PWSn, & WT) at baseline and follow up. At baseline, PWS and WTI had no significant correlations (FSI: r = 0.0446, p = 0.7684; TW: r = -0.0414, p = 0.7848). WT had significant negative correlation with WTI (FSI: r = -0.3337, p = 0.0234; TW: r = -0.3559, p = 0.0152). PWSn had weak positive correlation with WTI by FSI model (r = 0.2830, p = 0.0506), while this cannot be predicted by TW model (r = -0.1493, p = 0.3219).

The correlations by TW models were coincident with those by FSI models at follow up, there were no correlations between WT and WTI in both models (FSI: r = 0.0806, p = 0.5945; TW: r = 0.0423, p = 0.7804). PWS showed negative correlation with WTI in both models, but the TW model was weaker and not that significant (FSI: r = -0.3718, p = 0.0109; TW: r = -0.2246, p = 0.1335). The PWSn was in the same situation as PWS (FSI: r = -0.3904, p = 0.0073; TW: R = -0.2358, p = 0.1147).

Table 2. Correlation results between plaque progression (WTI) and the risk factors (PWS, PWSn, & WT)using FSI model and TW model at baseline(T1) and follow up(T2)

Model	Factors	T1(baseline)		T2(follow up)	
		r	p	r	p
FSI	**PWS**	0.0446	0.7684	-0.3718	0.0109
	PWSn	0.2830	0.0506	-0.3904	0.0073
	WT	-0.3337	0.0234	0.0806	0.5945
TW	**PWS**	-0.0414	0.7848	-0.2246	0.1335
	PWSn	-0.1493	0.3219	-0.2358	0.1147
	WT	-0.3559	0.0152	0.0423	0.7804

4. Discussion

4.1 Comparison with others modeling schemes

The comparisons of various risk factors to plaque progression between TW model and FSI model were reported in this paper. The relative errors of maximum PWS and PWSn using TW model data were slightly more than 10%, while the relation errors of the average PWS and PWSn were only about 5% in two cases. And it had much better performance in WT comparison. Those results indicated the TW model was better than 2D structure only model which had more than 30% relation error of PWS and PWSn [26].

TW models were in vivo IVUS based model, consequently, location specific pre-shrink ratio needed to be applied to for each slice. And there would be obvious differences between the results from TW model and those from FSI model in a small number of slices. As for TLS model based on the ex vivo data, due to the procedure without pre-shrink and the difference of model constructing the results of MPWS from TLS were closer to those from FSI model for each slice [7]. But overall statistical results, the accuracy of PWS from TW model was no less than it of TLS model.

4.2 TW models could be proper to be used for patient screening strategies.

Thin-Walled model involve the axial stretch to improve the accuracy and the time-saving characteristic is reserved comparing to 2D structure only model. What's more, TW model shorten the total time to several minutes, while there are still one or two hours needed to construct a TLS model. Since the TW model was expanded from 2D model, basically, the complexity to construct a TW model is almost same with that of a 2D model.

And it gives more possibility to finish the whole process automatically. From this demand, TW models might be practical to replace FSI models for the future clinical events.

4.3 Limitations for the models

In fact, the IVUS-VH cannot provide the data accurate enough, especially when it comes to the cap-thickness level. It is a common thing that the components cross over the lumen by segmented contour. Once this occurs, we keep the cap with thickness about 50 micron manually based the well accepted 65 micron threshold value for cap thickness. And this issue depends on the development of medical imaging. More accurate computational model could be constructed based on more accurate image. The other one is the lack of 3D vessel curvature data to make the 3D model more accurate.

Another limitation is that TW model is based on the single slice. This is why the TW model saves a lot of time and consequently it brings some disadvantages. Since it is current slice only, TW model cannot involve the impact from neighboring slices.

And the patient-specific and tissue-specific material properties were not considered in both models.

5. Conclusion

In this paper, a comparative analysis study between TW models and FSI models was performed. It was shown that the cost of TW model construction and data analysis was much less than that of FSI models. Errors of plaque stress and strain between TW and FSI models were less than 10%. However, correlation analysis results did not provide good agreement between the two models. Large scale studies are needed to investigate the feasibility of using TW models as approximation for FSI models.

Acknowledgements. This research was supported in part by NIH grant NIH/NIBIB R01 EB004759 and a Jiangsu Province Science and Technology Agency grant BE2016785.

References

[1] Fuster, V. (1999). The vulnerable atherosclerotic plaque: understanding, identification, and modification. *Cardiovascular Drugs and Therapy*, *13*(4), 363-363.

[2] Holzapfel, G. A., G. Sommer, and P. Regitnig. (2004). Anisotropic mechanical properties of tissue components in human atherosclerotic plaques. *Journal Of Biomechanical Engineering*,126(5):657 – 665.

[3] Holzapfel, G. A., Mulvihill, J. J., Cunnane, E. M., & Walsh, M. T. (2014). Computational approaches for analyzing the mechanics of atherosclerotic plaques: A review. *Journal Of Biomechanics*, *47*(4), 859-869.

[4] Huang, X. (2009). In Vivo MRI-based three-dimensional fluid-structure interaction models and mechanical image analysis for human carotid atherosclerotic plaques.

[5] Huang, X., Yang, C., Canton, G., Ferguson, M., Yuan, C., & Tang, D. (2012). Quantifying Effect of Intraplaque Hemorrhage on Critical Plaque Wall Stress in Human Atherosclerotic Plaques Using Three-Dimensional Fluid-Structure Interaction Models. *Journal Of Biomechanical Engineering-Transactions Of The Asme*, *134*(12),

[6] Huang, X., Yang, C., Zheng, J., Bach, R., Muccigrosso, D., & Woodard, P. K., et al. (2013). Abstract 9424: coronary artery disease-related death may be associated with higher critical plaque wall stress: a multi-patient 3d fluid-structure interaction study comparing plaque wall stress, flow shear stress, and plaque burden as risk indicators. *Circulation*(22), A9424.

[7] Huang, X., Yang, C., Zheng, J., Bach, R., Muccigrosso, D., Woodard, P. K., & Tang, D. (2016). 3D MRI-based multicomponent thin layer structure only plaque models for atherosclerotic plaques. *Journal Of Biomechanics*, *49*2726-2733.

[8] Huang, Y., Teng, Z., Sadat, U., Graves, M. J., Bennett, M. R., & Gillard, J. H. (2014). The influence of computational strategy on prediction of mechanical stress in carotid atherosclerotic plaques: Comparison of 2D structure-only, 3D structure-only, one-way and fully coupled fluid-structure interaction analyses. *Journal Of Biomechanics*, *47*1465-1471.

[9] Kural, M. H., Cai, M., Tang, D., Gwyther, T., Zheng, J., & Billiar, K. L. (2012). Planar biaxial characterization of diseased human coronary and carotid arteries for computational modeling. *Journal Of Biomechanics*, *45*(5), 790-798.

[10] Cai, M., Yang, C., Kural, MH., Bach, R., Muccigrosso, D., Yang, D,. Zheng, J., Billiar, K. L., Tang, D.(2011). Intravascular Ultrasound (IVUS)-Based Computational Modeling and Planar Biaxial Artery Material Properties for Human Coronary Plaque Vulnerability Assessment. *International Conference on Computational Engineering and Sciences (online Journal)*, ICCES, vol.19, no.4, pp.97-104.

[11] Mintz, G. S., Nissen, S. E., Anderson, W. D., Bailey, S. R., Erbel, R., Fitzgerald, P. J., & ... Winters, J. L. (2001). American College of Cardiology clinical expert consensus document on standards for acquisition, measurement and reporting of intravascular ultrasound studies (IVUS). *Journal Of The American College Of Cardiology*, *37*1478-1492.

[12] Naghavi, M., Libby, P., Falk, E., Casscells, S., Litovsky, S., Rumberger, J., & ... Willerson, J. (n.d). From vulnerable plaque to vulnerable patient - A call for new definitions and risk assessment strategies: Part I. *Circulation*, *108*(14), 1664-1672.

[13] Nair, A., Kuban, B. D., Tuzcu, E. M., Schoenhagen, P., Nissen, S. E., & Vince, D. G. (2002). Coronary plaque classification with intravascular ultrasound radiofrequency data analysis. *Circulation*, *106*(17), 2200.

[14] Nair, A., Margolis, M. P., Kuban, B. D., & Vince, D. G. (2007). Automated coronary plaque characterisation with intravascular ultrasound backscatter: ex vivo validation. *Eurointervention Journal of Europcr in Collaboration with the Working Group on Interventional Cardiology of the European Society of Cardiology, 3*(1), 113-20.

[15] Nasu, K., Tsuchikane, E., Katoh, O., Vince, D. G., Virmani, R., & Surmely, J. F., et al. (2006). Accuracy of in vivo coronary plaque morphology assessment: a validation study of in vivo virtual histology compared with in vitro histopathology. *Journal of the American College of Cardiology, 47*(12), 2405-2412.

[16] Rothwell, P., Gutnikov, S., & Warlow, C. (n.d). Reanalysis of the final results of the European Carotid Surgery Trial. *Stroke, 34*(2), 514-523.

[17] Tang, D., Yang, C., Kobayashi, S., & Ku, D. N. (2004). Effect of a lipid pool on stress/strain distributions in stenotic arteries: 3-d fluid-structure interactions (fsi) models. *Journal of Biomechanical Engineering, 126*(3), 363-370.

[18] Tang, D., Yang, C., Zheng, J., Woodard, P. K., Sicard, G. A., & Saffitz, J. E., et al. (2004). 3d mri-based multicomponent fsi models for atherosclerotic plaques. *Annals of Biomedical Engineering, 32*(7), 947.

[19] Tang, D. (2006). Flow in Healthy and Stenosed Arteries. Wiley Encyclopedia of Biomedical Engineering.

[20] Tang, D., Teng, Z., Canton, G., Hatsukami, T. S., Dong, L., & Huang, X., et al. (2009). Local critical stress correlates better than global maximum stress with plaque morphological features linked to atherosclerotic plaque vulnerability: an in vivo multi-patient study. *BioMedical Engineering OnLine, 8*(1), 15.

[21] Tang, D., Yang, C., Kobayashi, S., Zheng, J., Woodard, P. K., & Teng, Z., et al. (2009). 3d mri-based anisotropic fsi models with cyclic bending for human coronary atherosclerotic plaque mechanical analysis. *Journal of Biomechanical Engineering, 131*(6), 061010.

[22] Tang, D., Teng, Z., Canton, G., Yang, C., Ferguson, M., & Huang, X., et al. (2009). Sites of rupture in human atherosclerotic carotid plaques are associated with high structural stresses: an in vivo mri-based 3d fluid-structure interaction study. *Stroke, 40*(10), 3258.

[23] Tang, D., Kamm, R. D., Yang, C., Zheng, J., Canton, G., Bach, R., & ... Yuan, C. (2014). Image-based modeling for better understanding and assessment of atherosclerotic plaque progression and vulnerability: Data, modeling, validation, uncertainty and predictions. *Journal Of Biomechanics, 47*(SI: Plaque Mechanics), 834-846. doi:10.1016/j.jbiomech.2014.01.012

[24] Underhill, H. R., Hatsukami, T. S., Fayad, Z. A., Fuster, V., & Yuan, C. (2010). Mri of carotid atherosclerosis: clinical implications and future directions. *Nature Reviews Cardiology, 7*(3), 165-73.

[25] Wahle, A., Prause, P. M., Dejong, S. C., & Sonka, M. (1999). Geometrically correct 3-d reconstruction of intravascular ultrasound images by fusion with biplane angiography--methods and validation. *IEEE Transactions on Medical Imaging, 18*(8), 686-699.

[26] Wang, H., Wang, L., Zheng, J., Zhu, J., Maehara, A., & Yang, C., et al. (2015). Using 2d in vivo ivus-based models for human coronary plaque progression analysis and comparison with 3d fsi models. *Procedia Engineering, 126*(2), 451-455.

[27] Wang, L., Wu, Z., Yang, C., Zheng, J., Bach, R., & Muccigrosso, D., et al. (2015). Ivus-based fsi models for human coronary plaque progression study: components, correlation and predictive analysis. *Annals of Biomedical Engineering, 43*(1), 107-21.

[28] Yang, C., Bach, R. G., Zheng, J., & Ei Naqa, I. (2009). In vivo, ivus-based 3-d fluid–structure interaction models with cyclic bending and anisotropic vessel properties for human atherosclerotic coronary plaque mechanical analysis. *IEEE Transactions on Biomedical Engineering, 56*(10), 2420.

Effect of distal stenosis on the blood flow in right coronary arteries with serial stenoses

***†Biyue. Liu[1] and Dalin Tang[2]**

[1]Department of Mathematics, Monmouth University, USA.
[2] Department of Mathematical Sciences, Worcester Polytechnic Institute, USA

*Presenting and †Corresponding author: bliu@monmouth.edu

Abstract

Computer simulations of the blood flow in right coronary artery with two stenoses in the same arterial segment are carried out to investigate the interactions of serial stenoses, especially the effect of the distal stenosis. Various mathematical models are developed by varying the location of the distal stenosis. The numerical results show that the variation of the distal stenosis has significant impact on coronary hemodynamics, such as the pressure drop, the flow shifting, and the flow separation. Our simulations demonstrate that the distal stenosis has insignificant effect on the disturbed flow pattern in the upstream and across the proximal stenosis regions. In a curved artery segment with two moderate stenoses of the same size, the distal stenosis causes a larger pressure drop and a more disturbed flow field in post stenotic region than the proximal stenosis does. A distal stenosis located at a further downstream position causes a larger pressure drop and a stronger reverse flow.

Keywords: right coronary artery, serial stenoses, pressure drop, flow disturbance

1. Introduction

Atherosclerosis, a disease of large- and medium-size arteries, involves complex interactions between the artery wall and the blood flow. Hemodynamics plays an important role in the pathogenesis of atherosclerosis [1]-[5]. Vascular geometry is one of the risk factors strongly influencing the flow patterns and generating a pre-existing atherogenic hemodynamic environment [1][3]. Both clinical observations and experiment results show that there is a strong correlation between the sites of flow disturbance and preferential areas for the development of atherosclerotic diseases [2][4]. Atherosclerotic lesions originate preferably at the vessel branches, bifurcations and bends, where the blood vessel curvature creates dynamic environments of disturbed flow.

The obstruction presented by a moderate to severe stenosis can lead to a highly disturbed flow region at the downstream of the stenosis. The influence of an arterial stenosis in local blood flow is of considerable clinical importance. Many experimental and mathematical researches on hemodynamics in stenotic arteries have been reported to investigate the influence of the stenosis on the flow field downstream [6]-[9]. The presence of a moderate to severe stenosis significantly affects the downstream blood flow pattern, such as flow separation, secondary flow, low wall shear stress, and large pressure drop. These characteristic changes of the flow downstream may lead to the formation of a second stenosis [9]-[12]. In many clinical situations, the patient is found to have multiple stenoses in the same arterial segment.

The fluid dynamic interaction of multiple stenoses in coronary arteries is complex and cannot be adequately assessed by visual interpretation on the coronary angiogram [10]. The effect of serial stenoses on coronary hemodynamics has drawn much attention. Many researchers have

conducted the experimental and clinical investigations [13]-[16] and the numerical simulations [11][18]-[24].The existence of multiple stenoses can cause a decrease in blood flow at each stenosis, and then the pressure gradient at each stenosis can be lower than that at the stenosis if it were a single lesion in a vessel [17]. The study of De Bruyne *et al.* demonstrated that for multiple stenoses in the same vessel, the hemodynamic assessment of a stenosis and the potential benefit of angioplasty is significantly influenced by the presence of the second stenosis [25]. Talukder *et al.* showed that the total pressure drop across a series of mild stenoses increases linearly as the number of stenoses increases [16]. Using *in-vitro* experiments, D'Souza *et al.* found that the pressure drop coefficient of the upstream stenosis varies insignificantly in the presence of a downstream stenosis with a varying degree of severity [13]. Lee *et al.* performed numerical simulations of turbulent flow through tubes with double stenoses. They found that the maximum centerline velocity and disturbance intensity at the second stenosis are higher than those at the first stenosis when both stenoses have a 50% area reduction. The downstream stenosis usually does not have perceptible influences on the upstream flow fields [19][20].

In literature all of the numerical simulations examining the effect of multiple stenoses with varying size/location were performed using 2D models of axisymmetric straight tube. However, not much work has been reported on the effect of multiple stenoses on blood flow in curved arteries. No systematic study of the fluid dynamics of multiple stenoses in a segment of curved artery has been reported previously in the literature. It is expected that the curvature of the artery may intensify the complex of the blood flow in arteries with serial stenoses. In the present study, the blood flows in right coronary arteries with two stenoses are numerically examined. The coronary artery models are constructed based on the angiographic image of a patient with altered location of the stenoses.

2. Mathematical Models and Method

In the present study, the blood is assumed to be a laminar, incompressible, and non-Newtonian viscous fluid obeying the Carreau model [26][27]. The blood density ρ is assumed to be a constant at 1050 kg/m^3. The geometries of the models are constructed based on the inlet and outlet diameters and the center-line curve of the angiographic image of the coronary segment of a patient (see [28]), with varying locations of two moderate stenoses. The stenotic artery segments are 5.25cm in length with a diameter of 0.3256cm and 0.2954cm at the inlet and the outlet, respectively. In the computer simulation, to reduce the influence of the boundary conditions in the region of interest each coronary artery segment is extended at the inlet and the outlet by 0.6cm and 0.8cm, respectively. Including the extended inlet and outlet, the total length of coronary artery segment is 6.65cm in each model. At the inlet boundary, a time dependent pressure with waveform *Pin* (see Fig. 1) and a no-viscous-stress condition are applied. At the outlet boundary, a fully developed velocity profile with pulse waveform (see *Vout* in Fig. 1) is imposed as a normal outflow velocity. The pressure and the velocity waveforms in Fig.1 were extracted from the on-site blood pressure and flow velocity data of the patient [27][28]. A normalized time length $t/t_p = 1$ is used for simplicity. On the artery wall a non-slip boundary condition is imposed.

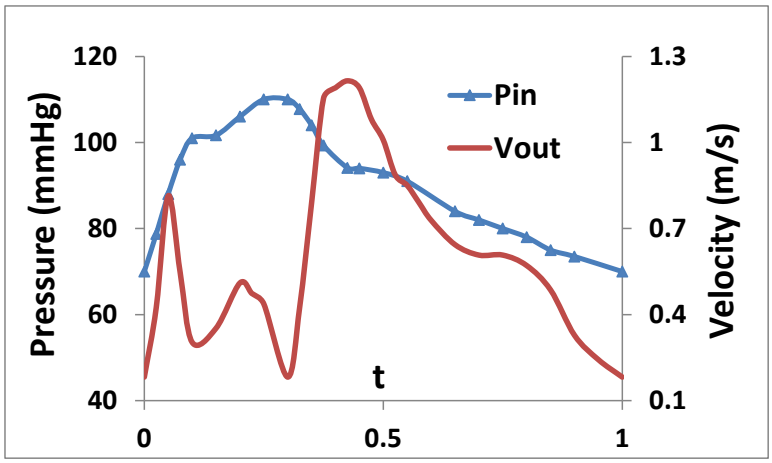


Figure 1. velocity and pressure waveforms

Unsteady Navier-Stokes equations are solved numerically with the finite element method over unstructured tetrahedral elements. The computations are performed using COMSOL 5.2a. To confirm that the numerical solutions are independent of the spatial mesh, the computations are repeated over the refined meshes till the maximum relative errors of the blood pressure over two consecutive refined meshes is less than 0.5%. Computations are performed over four consecutive cardiac cycles to ensure a truly periodic flow. The maximum relative error of the blood pressure between the third and the forth cycles is 0.53%. All numerical results presented in this paper are obtained over the meshes with about 2760000 elements.

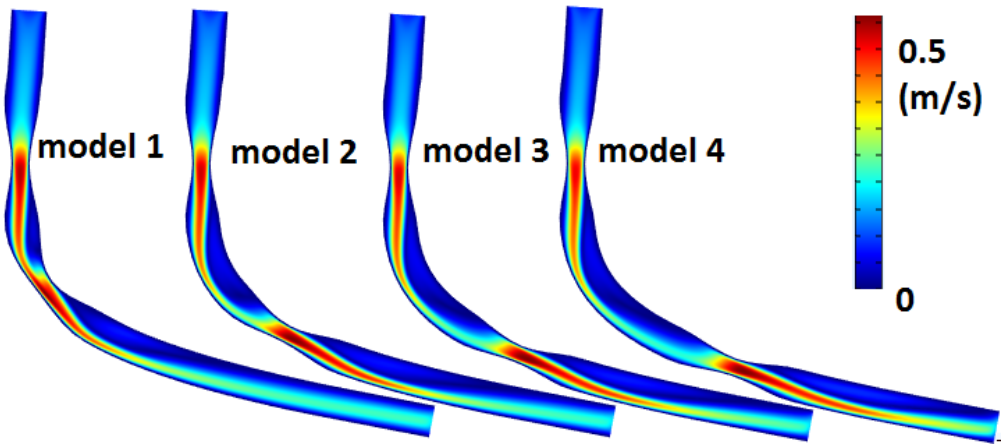


Figure 2. Magnitude of velocity at the center plane when t = 0.3 for four models

3. Numerical Results and Discussions

The shape of the stenotic coronary arteries and the location of stenoses are clearly shown in Fig. 2, including the extended inlet and outlet. The location of the proximal stenosis is the same in all four coronary models while the location of the distal stenosis is different in each model. The inter-stenotic distance between the necks of the proximal and the distal stenoses is 1.22*cm* (~3.9D), 1.96*cm* (~6.3D), 2.32*cm* (~7.5D), and 2.69*cm* (~8.7D) for model 1 through model 4, respectively. Here D is the average diameter of the artery. Both proximal and distal stenoses have the same length of 0.98*cm*, and a cross-section area reduction of 75% at the neck.

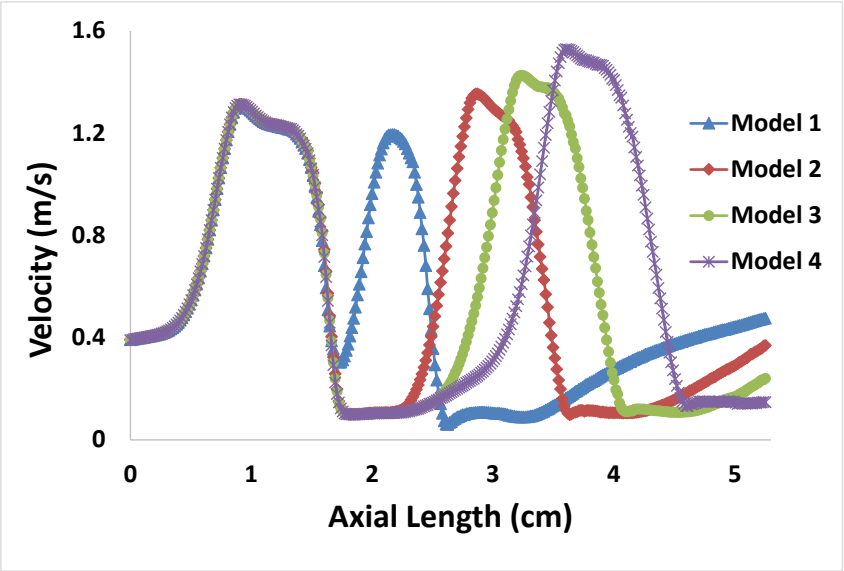


Figure 3. Temporal mean velocity magnitude along the center line of artery

Figure 2 plots the contour of velocity magnitude along the center plane for each coronary model at t = 0.3. Figure 3 plots the temporal mean velocity magnitude along the center line of the artery for each model. The horizontal axis is the axial length with z = 0*cm* and 5.25*cm* as the real inlet and outlet, respectively (The extended inlet and outlet are not included in Fig. 3). The neck of the proximal stenosis is at z = 0.86*cm* for all four models. The neck of the distal stenosis is at z = 2.08*cm*, 2.81*cm*, 3.18*cm* and 3.54*cm* for model 1, model 2, model 3 and model 4, respectively. Figure 4 includes the velocity field in the center plane of the artery at t = 0.8 for each model, including the extended inlet and outlet. Figure 5 is the plot of the slice-averaged PD_{mean} from the inlet to the outlet of the artery. Here PD_{mean} is the temporal mean pressure drop. A total of 128 points are picked evenly spaced on the lumen boundary of each cross-section. The PD_{mean} is averaged on the picked points of each slice.

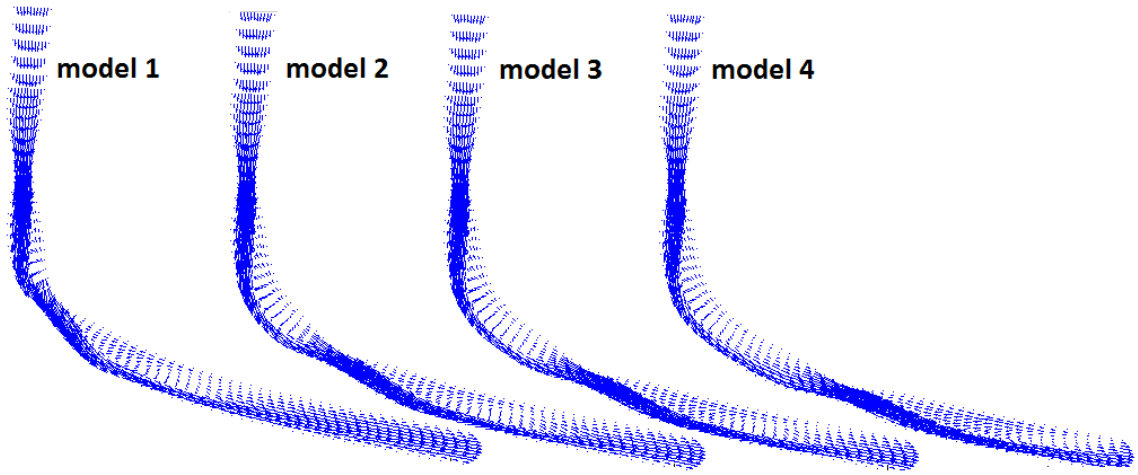


Figure 4. Plots of velocity field along the center plane at t = 0.8.

Flow Shifting

The contour plot of the velocity magnitude along the center plane for each model in Fig. 2 shows the pattern of flow shifting in right coronary artery with two moderate stenoses of the

same size. For models 2-4, the flow shifting in the stenotic area and post-stenotic area of the proximal stenosis is identical. This indicates that the location of the distal stenosis has no effect on the flow shifting pattern of the proximal stenosis when the inter-stenotic distance is six times the diameter of the artery or greater. In addition, we can also see that the flow shifting in the post stenotic region of the distal stenosis is stronger than that in the proximal stenosis even though both stenoses have the same severity. This observation can also be confirmed quantitatively by Fig. 3. The velocity magnitude along the center line is the same between z = 0*cm* (the inlet) and z = 2.32*cm* (the front edge of the distal stenosis of model 2) for models 2-4. The peak value at the neck of the distal stenosis is higher than that at the neck of the proximal stenosis for models 2-4. The further down the location of the distal stenosis, the higher of the velocity at the neck of the distal stenosis. The above discussed behaviors do not apply to model 1, where the flow shifting following the distal stenosis is weaker. This is probably due to the fact that the existence of the proximal stenosis will affect the flow disturbance of the stenosis in downstream when two stenosis are too close.

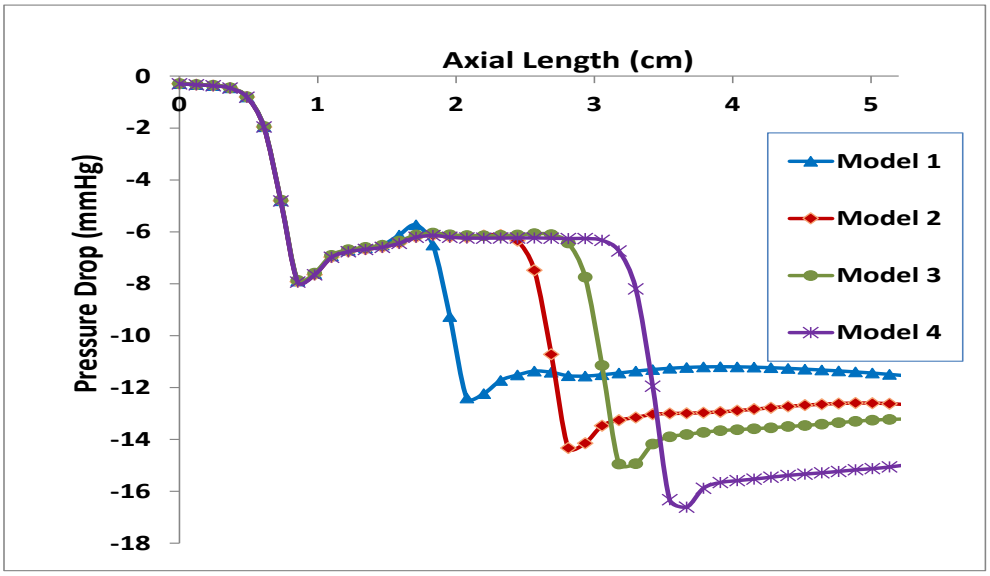


Figure 5. Slice averaged temporal mean pressure drop along the artery

Flow Separation

The formation and progression of a flow separation zone is one of the most important flow phenomena in the post stenosis region. Stenosis severity is strongly associated with the recirculation zone. The response of vascular endothelium to disturbed flow and its role in the pathogenesis of atherosclerosis have been studied extensively [29][30]. A net migration is directed away from the region of high shear gradient in a disturbed flow field. Following the reverse flow within the recirculation zone, the particle is exposed to low shear stress levels for a relative prolonged duration [30][31]. Figure 4 demonstrates the flow separation patterns in coronary arteries with two stenoses. In the post-stenosis area of each stenosis there is a region of flow separation near the inner wall of the bend. Comparing these two regions in each of the model plotted in Fig. 4, we can see that the distal stenosis results in a larger area of the flow separation and a stronger reverse flow near the inner wall. The further downstream the distal stenosis locates, the larger the area of the flow separation it results. It is evident that having the same area reduction the distal stenosis is more hemodynamically significant than the proximal stenosis in a right coronary artery segment with two stenoses.

Pressure Drop

The blood pressure difference along the coronary arterial length is one of the important initiating factors for atherosclerosis and intimal hyperplasia development [32]. The magnitude of the pressure drop has been clinically used to judge severity of the lesion. Several popularly used diagnostic parameters assessing the functionally significance of a stenosis are related to the pressure drop across the stenosis, such as the fractional flow reserve (*FFR*) and the pressure drop coefficient (*CDP*). Therefore it is of practical interest to examine the pressure drop pattern in coronary arteries with serial stenoses when investigating the effect of the distal stenosis. The curves of the slice averaged temporal mean pressure drop along the axial artery length plotted in Fig. 5 demonstrate that the distal stenosis does not affect the pressure drop across the proximal stenosis. On the other hand, the pressure drop across the distal stenosis could be influenced by the proximal stenosis if two stenoses are too close. When the inter-stenotic distance is six times the diameter of the artery or greater the pressure drop across the distal stenosis is larger than that across the proximal stenosis even though both stenoses have the same size of area reduction. A distal stenosis located in a further downstream position causes a larger pressure drop.

4. Conclusions

In this study coronary models are developed to investigate the hemodynamics of blood flows in right coronary artery segments with serial stenoses and to examine the effect of the distal stenosis. The center line of the artery segment and the location of the proximal stenosis are based on the angiographic image of the coronary segment of a patient, while the location of the distal stenosis varies in different models. Both stenoses are moderate with a 75% reduction in cross section area. Our simulations show that when the inter-stenotic distance is greater than or equal to six times the diameter of the artery the distal stenosis has insignificant effect on blood flow behavior from the inlet to the post-stenotic region of the proximal stenosis, such as the flow shifting pattern, the magnitude of velocity along the center line, the pressure drop, the area of flow recirculation and the strength of the reverse flow. In a right coronary artery segment with two moderate stenoses of the same size, the distal stenosis causes a more disturbed flow in the post stenotic region. Compared with those of the proximal stenosis, the flow shifting in the post stenotic region of the distal stenosis is stronger, the flow separation area downstream of the distal stenosis is larger, and the pressure drop across the distal stenosis is larger. The further downstream the distal stenosis locates, the more disturbed flow it creates in the post stenotic region. In summary, when two stenoses have the same size, the distal stenosis is more hemodynamically significant.

Acknowledgement. This work is partially supported by a grant from the Simons Foundation (#210082 to Biyue Liu), NIH grant NIH/NIBIB R01 EB004759, and a Jiangsu Province Science and Technology Agency grant BE2016785.

References

[1] Friedman, M.H., Hutchins, G.M., Bargeron, C.B., Deters, O.J., Mark, F.F. (1981) Correlation between intimal thickness and fluid shear in human arteries, *Atherosclerosis* **39**, 425–436.

[2] Karino, T. (1986) Microscopic structure of disturbed flows in the arterial and venous systems, and its implication in the localization of vascular diseases, *International Angiology: A Journal of the International Union of Angiology* **4**, 297–313.

[3] Myers, J.G., Moore, J.A., Ojha, M., Johnston, K.W., Ethier, C.R. (2001) Factors influencing blood flow patterns in the human right coronary artery, *Ann Biomed Eng* **29**, 109–20.

[4] Ku, D.N., Giddens, D.P., Zarins, C.K., Glagov, S. (1985) Pulsatile flow and atherosclerosis in the human carotid bifurcation: positive correlation between plaque location and low and oscillating stress, *Arteriosclerosis* **5**, 292–302.

[5] Prosi, M., Perktold, K., Ding, Z., Friedman, M.H. (2004) Influence of curvature dynamics on pulsatile coronary artery flow in a realistic bifurcation model, *J Biomech* **37**, 1767–75**.**

[6] Long, Q., Xu, X.Y., Ramnarine, K.V., Hoskins, P. (2001) Numerical investigation of physiologically realistic pulsatile flow through arterial stenosis, *J Biomech* **34**,1229–42.

[7] Stroud, J.S., Berger, S.A., Saloner, D. (2000) Influence of stenosis morphology on flow through severely stenotic vessels: implications for plaque rupture, *J Biomech* **33**, 443–55.

[8] Nosovitsky, V.A., Ilegbusi, O.J., Jiang, J., Stone, P.H., Feldman, C.L. (1997) Effects of curvature and stenosis-like narrowing on wall shear stress in a coronary artery model with phasic flow, *Comput Biomed Res* **30**, 61–82.

[9] Liu, B. (2007) The influences of stenosis on the downstream flow pattern in curved arteries. *Medical engineering & physics* **29(8)**, 868-876.

[10] Bernad, S.I., Bernad, E.S., Craina, M., Sargan, I., Totoran, A., Brisan, C. (2012) Particle Depositions and Related Hemodynamic Parameters in the Multiple Stenosed Right Coronary Artery, *J Clin Med Res* **4(3)**, 177-189.

[11] Bernad, S.I., Bernad, E.S., Totorean, A.F., Craina, M.L., Sargan, I. (2015) Clinical important hemodynamic characteristics for serial stenosed coronary artery, *Int. J. of Design & Nature and Ecodynamics.* **10**, 97–113.

[12] Dash, R.K., Jayaraman, G., Mehta, K.N. (1999) Flow in a catheterized curved artery with stenosis, *J Biomech* **32**, 49–61.

[13] D'Souza, G.A., Peelukhana, S.V., Banerjee, R.K. (2014) Diagnostic Uncertainties During Assessment of Serial Coronary Stenoses: An In Vitro Study, *Journal of Biomechanical Engineering* **136**, 021026-1-11**.**

[14] Pijls, N.H.J. De Bruyne, B., Bech, G.J.W., Liistro, F., Heyndrickx, G.R., M.Bonnier, H.J.R., Koolen, J.J. (2000) Coronary Pressure Measurement to Assess the Hemodynamic Significance of Serial Stenoses Within One Coronary Artery: Validation in Humans, *Circulation* **102**, 2371–2377.

[15] Sabbah, H.N., Stein, P.D. (1982) Hemodynamics of multiple versus single 50 percent coronary arterial stenoses, *Am. J. Cardiol* **50**, 276–280.

[16] Talukder, N., Karayannacos, P.E., Nerem, R.M., Vasko, J.S. (1977) An experimental study of the fluid dynamics of multiple noncritical stenoses, *J. Biomech. Eng.* **99**, 74–82

[17] Tanaka, K., Bezerra, H.G., Gaur, S., Attizzani, G.F., Botker, H.E., Costa, M.A., Rogers, C., Norgaard, B.L. (2015) Comparison Between Non-invasive (Coronary Computed Tomography Angiography Derived) and Invasive-Fractional Flow Reserve in Patients with Serial Stenoses Within One Coronary Artery: A NXT Trial substudy, *Annals of Biomedical Engineering*, **44**, 580–589.

[18] Bertolotti, C., Qin, Z., Lamontagne, B., Durand, L.G., Soulez, G., Cloutier, G. (2006) Influence of multiple stenoses on echo-Doppler functional diagnosis of peripheral arterial disease: a numerical and experimental study, *Ann. Biomed. Eng.* **34**, 564–574.

[19] Lee, T.S., Liao, W., Low, H.T. (2003) Numerical simulation of turbulent flow through series stenoses, *Int. J. Numer. Meth. Fluids* **42,** 717–740.

[20] Lee T.S., Liao W., Low, H.T. (2004) Numerical study of physiological turbulent flows through series arterial stenoses, *Int. J. Numer. Meth. Fluids* **46**, 315-344.

[21] Mustapha, N., Amin, N., Chakravarty, S., Mandal, P.K. (2009) Unsteady magnetohydrodynamic blood flow through irregular multi-stenosed arteries, *Comput. Biol. Med*. **39**, 896-906.

[22] Park, S.J., Ahn, J.-M., Pijls, N.H.J., et al, (2012) Validation of Functional State of coronary Tandem Lesions Using Computational Flow Dynamics," *Am. J. Cardiol*, **110**1578–1584.

[23] Jahangiri, M., Saghafian, M., Sadeghi, M.R. (2015) Numerical simulation of hemodynamic parameters of turbulent and pulsatile blood flow in flexible artery with single and double stenoses, *Journal of Mechanical Science and Technology* **29**, 3549~3560.

[24] Liu, B. Zheng, J., Bach R., Tang D. (2017) Influences of Flow Parameters on Pressure Drop in a Patient Specific Right Coronary Artery with Two Stenoses, *Computational Science and Its Application – ICCSA 2017*, Part I, LNCS 10404, Springer (to appear).

[25] De Bruyne, B., Pijls, N.H.J., Heyndrickx, G.R., Hodeige, D., Kirkeeide, R., Gould, K.L. (2000) Pressure-Derived Fractional Flow Reserve to Assess Serial Epicardial Stenoses Theoretical Basis and Animal Validation, *Circulation*. **101**, 1840-1847.

[26] Cho, Y.I., Kensey, K.R. (1991) Effects of the non-Newtonian viscosity of blood on flows in a diseased arterial vessel. Part 1: steady flows, *Biorheology* **28**, 241-262.

[27] Liu, B., Zheng, J., Bach, R., Tang D. (2015) Influence of model boundary conditions on blood flow patterns in a patient specific stenotic right coronary artery, *BioMedical Engineering OnLine* **14**(Suppl 1):S6

[28] Fan, R., Tang, D., Yang, C., Zheng, J., Bach, R., Wang, L., Muccigrosso, D., Billiar, K., Zhu, J., Ma, G., Maehara, A., Mintz, G.S. (2014): Human coronary plaque wall thickness correlated positively with flow

shear stress and negatively with plaque wall stress: an IVUS-based fluid-structure interaction multi-patient study, *BioMedical Engineering OnLine* 13:32.

[29] Traub, O., Berk, B.C. (1998) Laminar shear stress mechanics by which endothelial cells transducer an atheroprotective force, *Arterioscler Thromb Vasc Biol* **18**, 677-685.

[30] Tardy, Y., Resnick, N., Nagel, T, (1997) Gimbrone MA Jr, Dewey CF Jr. Shear stress gradients remodel endothelial monolayers in vitro via a cell proliferation-migration-loss cycle, *Arterioscler Thromb Vasc Biol* **17**, 3102-3106.

[31] Raz, S., EINAV, S., ALEMU, Y., BLUESTEIN, D. (2007) DPIV prediction of flow induced platelet activation – comparison to numerical predictions, *Annals of Biomedical Engineering* **35**, 493-504.

[32] Giannoglou, G.D., Soulis, J.V., Farmakis, T.M., Giannakoulas, G.A., Parcharidis, G.E., Louridas, G.E. (2005): Wall pressure gradient in normal left coronary artery tree. *Medical Engineering & Physics* 27, 455–464.

Simulation of sound transmission through thin elastic shell by the Coupled FEM/BEM

Z. Y. Yan

College of Aerospace Engineering, Nanjing University of Aeronautics & Astronautics, China

Email: jutsjtu@nuaa.edu.cn

Abstract

The coupled finite element/boundary element method is developed to simulate sound transmission through elastic shell. The vibration of the thin elastic shell is simulated using the finite element method and the acoustic fields both inside and outside the elastic shell are simulated using the boundary element method. To avoid the nonuniqueness problem occurring in the exterior boundary element method, the Burton and Miller formulation is applied. The algorithm is validated using the simulation of a fluid-filled submerged elastic spherical shell excited by a point sound source located at the center.

Keywords: FEM, BEM, Sound transmission

1. Introduction

Acoustic radiation and scattering from fluid-loaded elastic shells may be encountered in the aeronautical and naval industries as well as in underwater acoustics[1]. It is well known that the presence of fluid modifies considerably the resonance characteristics of the structure. In the mean time, the propagation of sound in fluids is altered by the presence of the elastic structure, which causes serious noise problems. The radiated noise from a vibrating structure is important for underwater-related applications. The scattering of acoustic waves from such structures contains information relating to the geometry and composition of the structure, which makes it possible to identify the structure by the remote sensing. Therefore, it is of considerable interest to predict the acoustic fields both radiated and scattered by a submerged vibrating structure.

Analytical approaches to such kind of fluid-structure interaction problems are almost invariably concerned with spherical or infinite circular cylindrical shells subjected to axisymmetric excitations for which the classical method of separation of variables is available. When analyzing the sound radiated or scattered by submerged elastic shells of more complicated shapes, it is almost indispensable to use numerical codes that can handle the complexity of the structure in question.

For a complex structure subjected to known applied forces, the finite element method (FEM) has become an accepted, well-proven, and highly successful analysis tool. Therefore, for the sound/structure interaction model, it is almost consistent to formulate the structural dynamic equations *via* FEM techniques. When applying to the case of interior problems where the fluid is inside the structure, FEM also gives satisfactory solutions. However, in the case of exterior problem in which fluid occupies an unbounded domain, the FEM is inefficient. When FEM has been applied to unbounded exterior problems, the domain has to be truncated and radiating boundary conditions have to be enforced. In addition, they are limited by computer memory and runtime considerations as well. To deal with the unbounded exterior acoustic field problems, the boundary element method based on the utilization of the Helmholtz integral equation has been the most popular numerical tool. The BEM method has several advantages over a FEM treatment of the acoustic problem, including a reduction of

dimensionality of the problem by one, and an automatic satisfaction of the radiation condition [2]. The elegance of this method is the mathematical simplicity of the resulting integral expressions.

There are many works on the coupled finite element/boundary element method. For example, Jeans and Mathews presented a unique coupled FEM/BEM method for the elastoacoustic analysis of fluid-filled thin shells [3]. They concluded that except for problems having a significant density difference between the internal and external acoustic fields, for example, air and water, their formulation was suitable.

2. The coupled FEM/BEM model

Our purpose is to investigate the acoustic transmission through thin elastic shells with different fluids on both the inside and outside. There are different fluids on the inside and outside of the thin elastic shell. A thin elastic shell is defined on the closed surface *S*. The shell submerged in an infinite fluid with density ρ^e in the exterior domain *E* and contains a fluid with density ρ^i in the interior domain *D*. The fluids on the inside and outside are assumed to be inviscid and compressible. The surface of the shell is assumed smooth. The normal vector at arbitrary point of the shell surfaces is uniquely determined. Its direction is defined to point into the exterior domain. In the following derivation, assuming there are a point sound source in the interior domain.

According to Hamilton's principle, the finite element governing equation for the dynamic fluid-structure interaction system is given by

$$[-\omega^2 M - i\omega C + K]\{U\} = \{F_I\} + \{F_A\} \tag{1}$$

where *M, K, C* are the mass, stiffness and damping matrices respectively. *U* denotes the displacement and ω is the circular frequency. F_A represents the known applied excitation forces and F_I represents the interaction forces generated by the acoustic fluid acting on the fluid-structure interaction surfaces. The vector of interaction force can be defined through the structure coupled matrix $L_s (L_s^i \text{ and } L_s^e)$ and the nodal acoustic pressures $\{\varphi^i\}$ and $\{\varphi^e\}$, that is

$$\{F_I\} = L_s^i\{\varphi^i\} - L_s^i\{\varphi^e\} \tag{2}$$

Where $\{\varphi^i\}$ and $\{\varphi^e\}$ are the interior and exterior surface acoustic pressures respectively. L_s^i and L_s^e are the structure coupled matrices on the interior surface and exterior surface. The coupled matrix of element is defined as

$$L = \iint_S [N]_f^T \{\vec{n}\}[N] dS \tag{3}$$

Where $[N]_f^T$ is the shape function matrix about displacement in the finite element method[5]. $[N]$ is the shape function matrix about acoustic pressure in the boundary element method and $\{\vec{n}\}$ is the vector of the direction cosines of normal vector.

In operator notations, the surface Helmholtz integral equations for exterior problem can be given as,

$$\left[-\frac{1}{2}I + M_k\right]\varphi^e = L_k \frac{\partial \varphi^e}{\partial n} \tag{4}$$

For interior problem

$$\left[\frac{1}{2}I + M_k\right]\varphi^i = L_k \frac{\partial \varphi^i}{\partial n} + I\varphi^I \tag{5}$$

where the integral operators M_k and L_k are defined as

$$M_k \mu = \iint_S \mu \frac{\partial G}{\partial n} dS \tag{6}$$

$$L_k \mu = \iint_S \mu G dS \tag{7}$$

where $G(p,q)$ is the free-space Green's function for the three-dimensional acoustic wave equation.

To form the coupled FEM/BEM model, the kinetic continuity boundary condition can be expressed as

$$\frac{\partial \varphi^e}{\partial n} = \omega^2 \rho^e Q^e \{U\} \tag{8}$$

$$\frac{\partial \varphi^i}{\partial n} = \omega^2 \rho^i Q^i \{U\} \tag{9}$$

where Q is defined as the kinetic coupled matrix.

Then the final coupled FEM/BEM formulation can be expressed as

$$[-\omega^2 M + K - i\omega C]\{U\} = L_s^i \{\varphi^i\} - L_s^e \{\varphi^e\} \tag{10}$$

$$\left[-\frac{1}{2}I + A^e\right]\{\varphi^e\} = \omega^2 \rho^e B^e Q^e \{U\} \tag{11}$$

$$\left[\frac{1}{2} + A^i\right]\{\varphi^i\} = \omega^2 \rho^i B^i Q^i \{U\} + \{\varphi^I\} \tag{12}$$

where A and B are defined the discretized coefficient matrices of the integral operators M_k and L_k respectively.

To avoid the nonuniqueness problem occurring in the exterior boundary element method, the Burton and Miller formulation [5] is applied.

3. Numerical example

In order to test the correctness of the coupled FEM and BEM method, a fluid-filled submerged elastic spherical shell excited by a point sound source located at the center has been analyzed. In this example, the fluid mediums inside the shell is air and outside the spherical shell is sea water.

The external radius of the spherical shell is 1.01 m and the internal radius is 1.0 m. Therefore, the relative thickness of the spherical shell is 1% . This shell is a thin shell. The density of the spherical shell is $\rho = 7.81\times10^{3}\,kg/m^{3}$. The elasticity modulus is $E = 2.07\times10^{11}\,Pa$ and Poisson's ratio is $\mu = 0.3$. The air density is $\rho^{a} = 1.21kg/m^{3}$ and sound velocity in air is $c^{a} = 346m/s$. While, the sea water density is $\rho^{w} = 1030kg/m^{3}$ and sound velocity in sea water is $c^{w} = 1500m/s$. The spherical shell is discretized using 96 surface elements.

In Fig. 1 and Fig. 2 the calculated frequency range is from 50 to 2400 Hz. In these figures, the numerical results agree quite well with the corresponding analytical solutions.

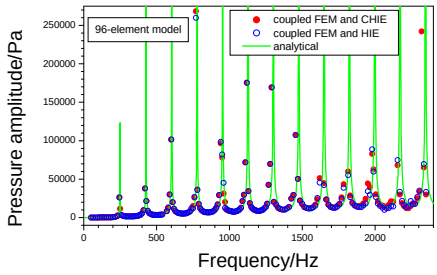


Figure 1. Frequency response of exterior acoustic surface pressure amplitude for a spherical shell

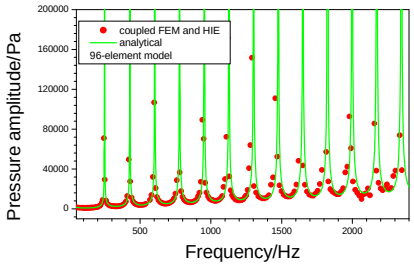


Figure 2. Frequency response of interior acoustic surface pressure amplitude for a spherical shell

4. Conclusions

A coupled finite element/boundary element method is developed for the simulation of sound transmission through thin elastic shell with different fluids on the inside and outside of the shell. This method and the corresponding in-house program is validated by the numerical simulation of sound transmission through a thin spherical shell with air inside and water outside of the shell.

Acknowledgements

This project was supported by NSFC (11172132) and NSFC (51405225); the Priority Academic Program Development of Jiangsu Higher Education Institutions.

References

[1] Junger, M.C. and Feit, D. (1986) Sound Structure and Their Interaction, 2nd ed., MIT, Cambridge, MA.
[2] Ciskowski, R.D. and Brebbia, C.A. (1991) Boundary element methods in acoustics, Souththampton Boston: Computational Mechanics Publications, 61-74; 109-129
[3] Jeans, R. A. and Mathews, I. C. (1990) Solution of fluid-structure interaction problems using a coupled finite element and variational boundary element technique, *J. Acoust. Soc. Am.*, **88**, 2459-2466

[4] Yan, Z. Y. (2000) Sound transmission through thin elastic shell with different fluids on both the inside and outside. Ph. D Thesis, Shanghai Jiao Tong University, P.R. China.
[5] Burton, A. J.; Miller, G. F. (1971) The application of integral equation methods to the numerical solution of some exterior boundary value problems. *Proc.R.Soc.London Ser.* **A323**, 201-210.

Numerical investigation on the gasoline mixture with side entering mechanical agitator

***X. X. Tang, Y. P. Qu, †S. Y. Chen, J. Y. Pan and D. D. Wan**

Key Laboratory of High-efficiency and Clean Mechanical Manufacture, School of Mechanical Engineering, Shandong University, China

*Presenting author:1245061506@qq.com
†Corresponding author: chensy66@sdu.edu.cn

Abstract

Employing the standard k-e turbulence model and sliding mesh technique, Fluent software was used to simulate the mixing flow peculiarity of the gasoline blending in the gasoline tank with installing side entering agitators. Simulation results show that high speed region is located in the vicinity of the Z=0.126 m plane paddle, where the speed can reach to 10.02 m/s. After continuous attenuation, the speed can reach to 0.35 m/s near the wall eventually, which meets the gasoline blending requirements according to the national standards. The stirring axis has a downward inclination with the horizontal plane in order to optimize the flow field in the middle and lower part of the tank further.

Keywords: Gasoline blending; Numerical simulation; Multiphase flow; Side entering agitator.

Introduction

As a main part of refinery production, gasoline from different refine processes should be blended together according with the national standards. The most widely used mixing equipment currently is the entering agitator installed in the gasoline tank. Especially in large storage tank, inputting a small power in the tank can get a relatively better effect [1]-[4]. The theoretical study of the flow field in the tank is very complicated and it is also very difficult to understand the distribution of the flow field in the tank by the test. In recent years, with the development of computational fluid dynamics (CFD) technology, numerical simulation has been widely used to study the characteristics of flow in gasoline tank, in which can shorten the research time and save funds. As a new means of scientific research, researchers pay more attention to the numerical investigation than before [5]-[6].

In the past 20 years, scholars have made further progress in the research of side entry mixing system. In 2004, Asghar [7] et al used the RNG k-e turbulence model to compute the flow field and mixing time of two different kinds of gasoline under the operation of single side entry agitator, of which the results are in agreement with the experimental results. In 2005, Janz [8] et al studied the agitating shaft's angle of the side entering agitator. In 2007, Saeed [9] et al simulated the power and flow parameters distribution in the pulping mixing tank with a single side entry agitator in

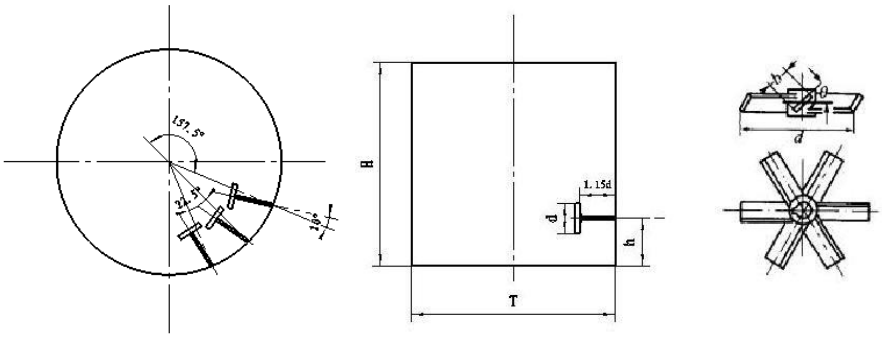
the flue gas desulfurization system of thermal power plant. In 2010, Gomez [10] et al carried out the numerical simulation on the side entry mixing tank under laminar flow condition and compared the results with PIV. The numerical simulation of the single-phase and multi-phase flow field in the side entry stirred reactor with different installation angle is carried out by Zhang [11] et al in 2013. In 2015, Liang [12] et al studied the liquid phase concentration distribution and the effect of particle diameter on the critical suspension speed of large double-layer side entry tank at different rotational speeds, and in the same year, Song [13] et al studied the stirring effect in the fermentation tank on the three-dimensional flow of solid-liquid (sludge and water) two-phase and obtained the velocity and concentration distribution in the tank.

In this paper, Fluent software was used to simulate the internal flow field of the mixing system so as to study the mixing peculiarity of multiphase gasoline ,which would have the guiding significance for the energy saving and emission reduction, the engineering application and the optimization design of the stirred reactor.

1 Calculation Model

1.1 *geometric model*

The engineering model adopted in this paper is based on the G408 gasoline blending tank which is owned by a Refining & Chemical Co., Ltd. Tank height is 14 m, the diameter is 22 m, nominal volume is 5000 m^3, safe height is 12.4 m, and the model is shown in Figure 1 (b). The stirrer is fixedly inserted into the angle-type lateral extension, whose arrangement is concentrated. Three agitators are installed as shown in Figure 1 (a). The angle of the axis of each stirrer is 22.5°, and the center line of the stirrer is deviated from the center line of the tank, whose angle is β. Three agitators are installed with a down angle. According to results of Liu [14], we takes the arrangement with the smallest power consumption, where the β is 10°. For the purpose of simplifying the model, we use the three-leaf propeller paddle type whose folding angle is 24°. In the case of ensuring the velocity of majority of the low-speed zones is not lower than the minimum mixing speed of 0.3 m/s, the similarity criterion of equal mixing power per unit volume is used for the side mixing model. Referring to agitator selection method from the "Mixing Equipment" [15], we determine the basic geometry of the model as follows: Inner diameter T is 1000 mm, height H is 1000 mm, and the arrangement of the blades is horizontal, lateral mounting height h is 126 mm, the diameter of the stirrer d is 100 mm, the diameter of the stirring shaft is 10 mm, the size of the hub is such as H15mm x 15mm, width of the blade is 15 mm, thickness is 2 mm, which is shown in Figure 1.

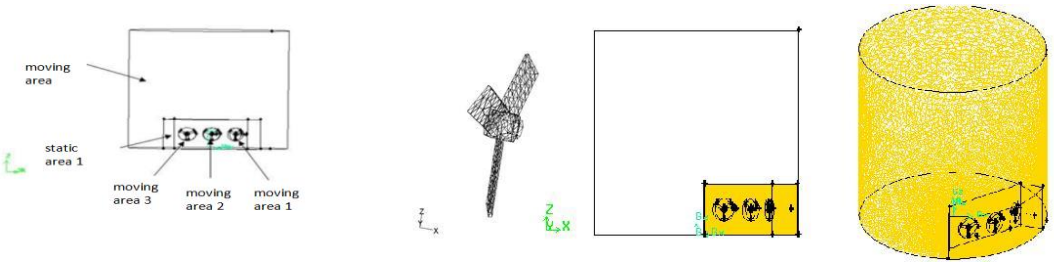


(a) (b) (c)

Figure 1. Geometric models: (a) Agitator arrangement orientation; (b) Blending tank model; (c) Schematic diagram of three leaf open turbine agitator

1.2 *Meshing and boundary conditions*

Gambit software is used to generate model grids [16]. As the agitator rotating, there are three moving and two static regions, which are shown in Figure 2 (a). Additionally, the three agitator positions are concentrated. Therefore, we block the model for meshing by using Tet/Hybrid scheme. The size of the boundary of the movement area is H 20 mm x 110 mm, the grids of motion region and the nearby static area are in local refinement, the total number of grid cells is 865817, as shown in Figure 2 (b).



（a） （b）

Figure 2. Grid division (a) Five regions of the mixing model; (b) Stirred tank area grid map

The boundary conditions for the side agitation model are set as follows, the top of the mixing tank is a free liquid surface, whose type is selected as Symmetry, and the relative motion of the stirring blade and the multiphase fluid in the tank is carried out by moving meshing method [17]. The rotational velocity of the three agitating blades is zero relative to the three corresponding moving regions, and the rotation axis remains the same. The model block interface is defined as the exchange wall Interface, and the remaining walls default to no slip standard boundary Wall.

1.3 *Simulation of working conditions and calculation strategies*

In the gasoline refining industry, a brand of gasoline is composed of the following four components [18], reforming gasoline, MTBE, catalytic gasoline and alkylated gasolines. So, there are four phases in the side agitation model, which is shown in Table 1. The density and viscosity of each in the table are measured at 20° C. In the

initial state, the four components are assumed to fill the entire blending tank in a layered form， and under the action of gravity, the first phase of the maximum density (reformed gasoline) is located at the tank bottom. Figure 3 shows the initial stratification distribution of each phase. Reynolds number of the fourth phase [19] is,

$$\mathrm{Re} = \frac{\rho n d^2}{\mu} \qquad (1)$$

Where ρ is density, kg/m^3； n is the speed of agitator, r/s； d is the blade diameter， m； μ is fluid dynamic viscosity, Pa•s.

Table 1. Five kinds of physical parameters and distribution region in stirred tank model

components	density （kg/m 3)	viscosity （Pa•s）	Z distribution region （m）
Reformed gasoline	860.7	0.0004	0-0.22
MTBE	741.3	0.00036	0.22-0.32
Catalytic gasoline	733.2	0.00035	0.32-0.9
Alkylation	699.3	0.00034	0.9-1.0

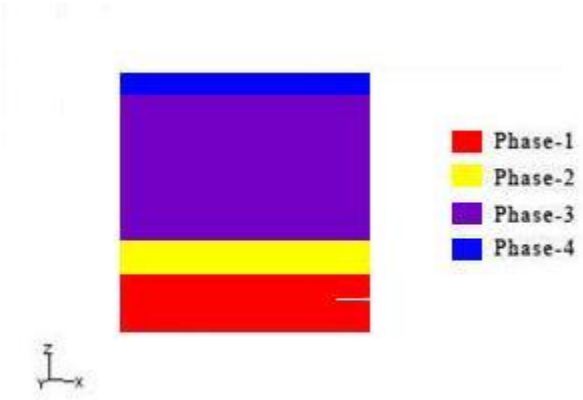


Figure 3. Initial distribution map of four phases

The Reynolds number of the region with the lowest density in the mixed flow field is 3.81x10^5. The mixing flow field is in turbulent state and the turbulence equation needs to be opened [20].

The steady-state implicit separation solver based on pressure is used for the side-mixing model. The pressure velocity coupling is solved by SIMPLE method [21], momentum, kinetic energy, dissipation rate is solved by Second Order Upwind scheme. The Eulerian mode is used in multiphase flow, the turbulence model is the

standard k-e model where the convergence of all variables is $1x10^{-5}$. After the velocity converges, the unsteady solver is used to calculate the volume distribution function. The function is discretized by QUICK format and the rest is selected by default. The working pressure is set on the cylindrical surface at the top of the mixing tank, and the pressure value is 101325Pa [22]. As we open the gravity acceleration option, the direction is -Z.

According to GB/T 4756-98 "Gasoline liquid manual sampling method" [16], in the tank model, three points of A(0.3,0,0.1), B(0.3,0,0.5) and C(0.3,0,0.9) were selected as the density monitoring points, and the mixing time is,

$$t_{95\%} = \text{time} \quad \text{for} \quad \left| \frac{d - \bar{d}}{\bar{d}} \right| \leq 0.003 \qquad (2)$$

Where $\bar{d}$ is the equilibrium concentration (when mixing is complete), d is the concentration measured at the monitoring point.

We need real-time monitoring of mixed phase density changes. When the relative density value of A, B, C in the mixed phase reaches 0.003 in the vertical direction, the mixed gasoline is uniform.

2. Simulation results and discussions

With the start of the stirrer, the flow velocity is increasing. After the iteration of 800 time steps, the agitator has a very good flow field range. The area with larger velocity of flow passes through the direction of middle diameter of blending tank, then returns back when the fluid hits the wall. After returning to the vicinity of the stirring blade, the velocity is further increased to complete a process cycle as shown in Figure 4.

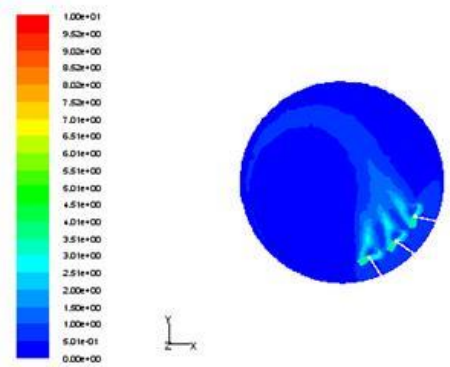


Figure 4. Velocity distribution of flow field at Z=0.126m

From the X-axis positive direction, the three agitators are in the clockwise rotation. From the Z-axis direction, the entire flow field will show a trend of counter-clockwise rotation. The velocity of the three agitators is superimposed on each other, and the liquid flows in the right direction of the opposite side. After the fluid hitting the tank wall, a vortex returned near the agitator, which is shown in Figure 5. As the centerline of the stirrer has a left angle to the centerline of the tank, the contact area between the

reflux field and the stirring paddle is larger [23], further increasing the flow velocity and forming a new cycle. The mounting angle of the three sides is intertwined with the blade mounting angle, which expands the flow of the agitator. From Figure 5, the radial flow effect of the stirring blade is very good, which directly causes the axial flow of the mixing tank and the multi-phase mixing in the vertical direction, which is the main source of power to break the initial phase and stratification.

From Figure 6, the axial flow of the side agitator is very significant. The high velocity region of the mixing flow is located in the vicinity of the stirring blade on the Z=0.126 m plane, where the velocity can reach to 10.02 m/s. After continuous attenuation, it can reach to 0.35 m/s at the wall in order to meet the petrol blending speed requirements. The location of the low velocity zone on the four horizontal sections is not the same, and the velocity distribution is similar and uniform in the whole. There is no increase in the range of the low velocity zone because of the distance from the plane of the stirring shaft. On the contrary, compared with the Z=0.126 m plane, the flow field of the three region is relatively mild, where the

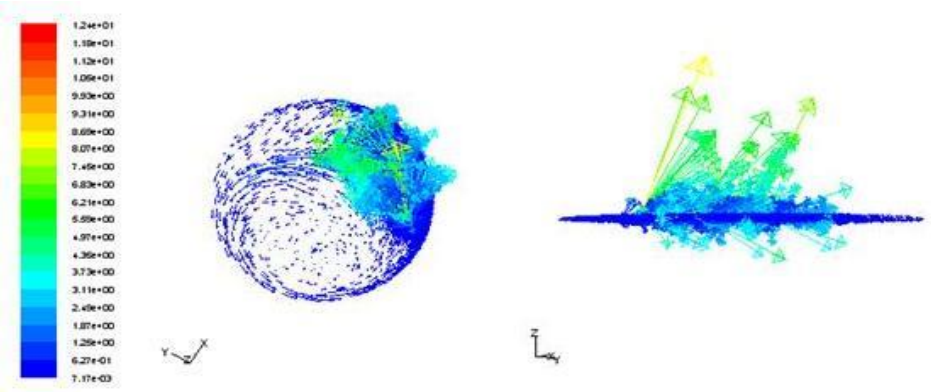


Figure 5. Velocity vector diagram of flow field at Z=0.126m

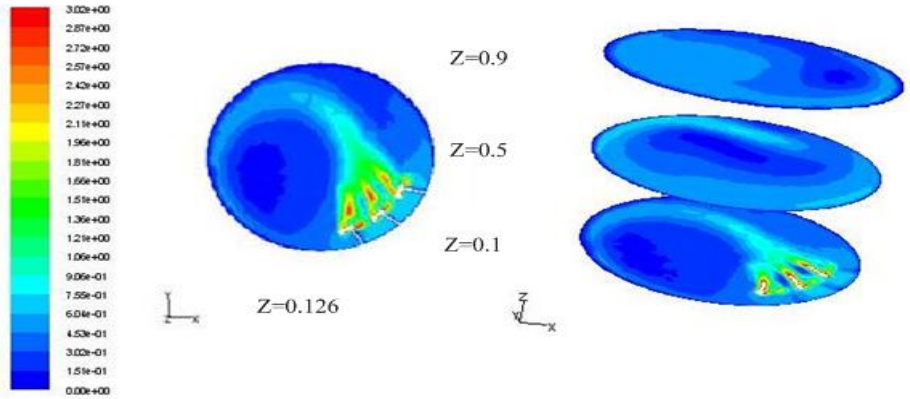


Figure 6. The velocity distribution on the four plane

velocity is large in the range of 0.30.6m/s, even which is smaller in a low velocity zone (deep blue), it is more conducive to the mix. It should be noted that a vortex will be formed, which determines that the low velocity region is not at the opposite tank wall but near the central area, as shown in Figure 7. This is very unfavorable for the

gasoline blending, and the flow structure determines the center of the low-speed area relatively stable and not easy to eliminate, which will certainly affect the mixing time and be also a disadvantage of mechanical mixing.

For the three vortices on Z=0.1 m, Z=0.5 m and Z=0.9 m planes, it can be seen from Figure 7 that the central regions of the three vortices are low velocity zone. The velocity is scattered to the surrounding area and different in each direction (which is caused by the rotation of the vortex). Additionally, there is a speed gradient in the same direction. The region with the largest velocity in the whole plane is located in the Z=0.1 m plane, and the velocity in the Z=0.9 m plane is the smallest one.

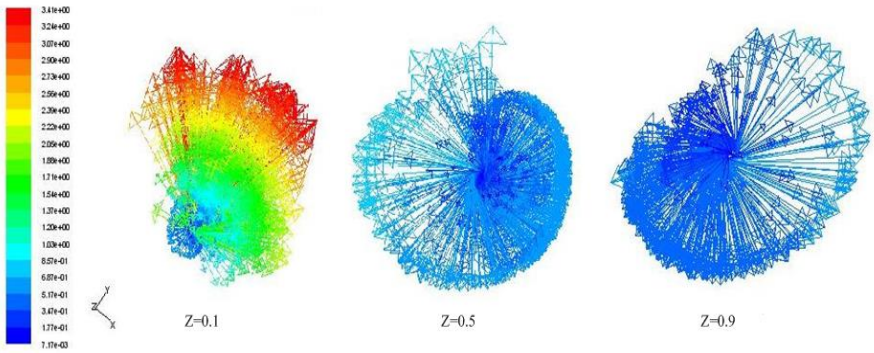


Figure 7. The velocity vector diagram on the three planes

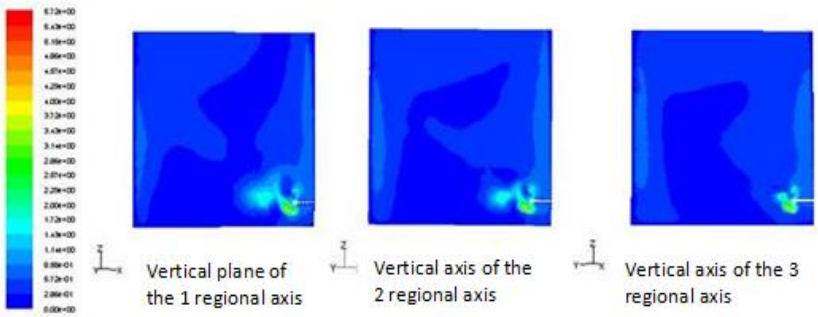


Figure 8. Velocity distribution in the vertical axis of three moving region

From Figure 8, the velocity field formed by three uniform impellers on the vertical plane of the axis is very similar and it is again verified that the vicinity of the central region of the transfer tank is a low velocity region. At the same time, the bottom of the tank is also a blind area for the mixing of the stirring blades. It is suggested that there should be a downward inclination angle between the stirring axis and the horizontal surface in the actual engineering installation, further optimizing the flow field in the lower and middle areas of the mixing tank [24].

The density values of three planes of Z=0.1 m, Z=0.5 m and Z=0.9 m obtained by Fluent are imported into Origin, and the density mixing time curve is obtained in Figure 9. The density of the three planes fluctuates violently between 10 s and 20 s after the stirring start, which is the result of axial flow in the flow field of the bleeding tank caused by radial flow. The upper layer of low density is brought into the lower region of the tank so that the density at the plane Z=0.1 m and Z=0.5 m decreases,

corresponding to the upper zone density. With the continuous action of the stirring blade, the density gradually reaches to the equilibrium density of 755.9kg/m^3 .On the Z=0.1 m plane, the mixed gasoline finally reaches the equilibrium density requirement, which is related to the effect of the low velocity zone near the stirring blade, resulting in a mixing time of 83.5 s.

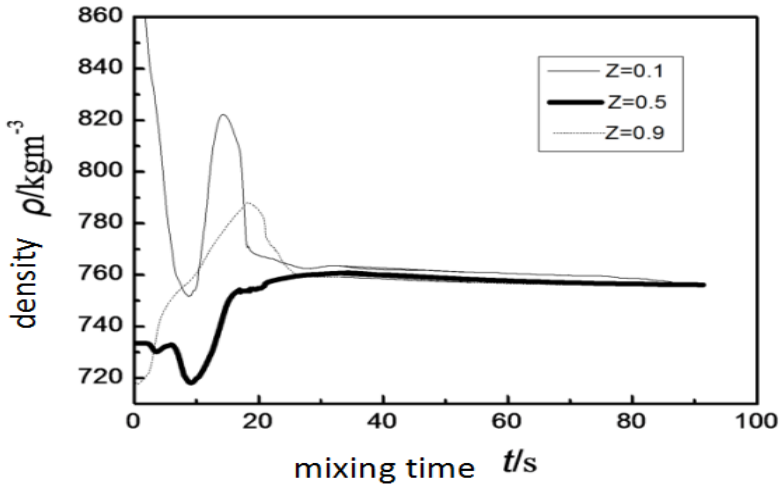


Figure 9. The mixed gasoline density profile and mixing time

3. Conclusion

In this paper, through the numerical simulation of the internal flow field of the side entering agitator system, the following conclusions are drawn:

(1) With the increase of stirring time, the flow velocity gradually increases, the radial flow affects the multi-phase mixing in axial and vertical direction, breaking the initial phase separation and then achieving gasoline blending. The velocity distribution of each section of the Z axis is similar, in the far zone, the stirring is relatively soft as the velocity range is 0.3~0.6 m/s. The range of the low velocity region is smaller, which is favorable for the mixing.

(2) The axial flow of the stirrer is very significant. The high velocity region of the mixed flow is located in the area near Z=0.126 m plane, where the speed can reach 10.02 m/s and then decay to 0.35 m/s at the wall to meet the speed requirements of gasoline blending.

(3)The bottom of the tank is a blind area for the mixing process. It is suggested that there should be a downward inclination angle between the stirring axis and the horizontal plane when the actual engineering installation is carried out. With the increase of stirring time, the mixing density gradually reach to the equilibrium density of 755.9kg /m^3, and the final mixing time is 83.5s.

Acknowledgement
This project is supported by Science and Technology Development Planning of Shandong Province (2014GGX108001), Key research and development program of

Shandong Province (2016GGX104018).

References

[1] Yu, Z. B. (2006) Comparison of Tank Mixing Ways in Petroleum Chemical Corporation. *Practical Scientific* **4**, 80-83.

[2] Cai, X. J. and Hu, L. Z. (2001) Research on the Structure of Side Inserted Agitator. *Practical Chemical Equipment* **30**, 45-46.

[3] Fang, J. and Sang, Z. F. (2009) Numerical Simulation of 3-D Flow Field in Side-entering Agitator. *China Petroleum Machinery* **37**, 30-34.

[4] Ma, Q.S. and Nie, Y.Q. (2003) Numerical Simulation of 3-D Flow Field in Side-entering Agitator. *Journal of Chemical Industry and Engineering* **54**, 612-618.

[5] Zhou, G. Z., Wang, Y. C. (2003) CFD Study of Mixing Process in Stirred Tank. *Journal of Chemical Industry and Engineering* **54**, 886-890.

[6] Guo, W. H. and Pan, J. Z. (2009) Application of Computational Fluid Dynamics to Flow Research and Structure Design of Stirred Tank. *Chemical Engineering* **37**, 20-23.

[7] Asghar, A. D. and Masoud, R. (2004) CFD Simulation of Homogenization in Large-scale Crude Gasoline Storage Tanks. *Journal of Petroleum Science and Engineering* **43**, 151-161.

[8] Janz, E. E., Fasano, J. and Myers, K. J. (2005) Different Solids Suspension Techniques in Flue Gas Desulfurization. *American Institute Of Chemical Engineers*, 4067-4068.

[9] Saeed, S., Ein MozaffariF. and Upreti, S. R. (2007) Using Computational Fluid Dynamics Modeling and Ultrasonic Doppler Velocimetry to Study-pulp Suspension Mixing. *Industrial and Engineering Chemistry Research* **46**, 2172-2179.

[10] GOMEZ, C., BENNINGTONC, P. J. and TAGHIPOUR, F. (2010) Investigation of the Flow Field in a Rectangular Vessel Equipped With a Side-entering Agitator. *Journal of Fluids Engineering* **132**, 1-13.

[11] Zhang, L. J. and Chen, G. G. (2013) Numerical Simulation of the Flow Field Characteristics of the Multi Phase Flow in a Stirred Tank. *Chinese Journal of Environmental Engineering* **4**, 1594-1600.

[12] Liang, J.Y. and Zhou, Y. J. (2015) Study on Solid Liquid Flow in Double Side Inlet Stirred Tank. *Light Industry Machinery* **33**, 35-38.

[13] Song, J. L. (2015) Numerical Simulation of Solid Liquid Two Phase Flow in a Fermentation Tank. *Energy Conservation* **5**, 22-25.

[14] Liu, G. Y. (2010) Numerical Simulation of Solid-liquid Suspension in Side Mixing. *Master's Degree Thesis of Shandong University*, 27-40.

[15] Wang, K. and Yu, J. (2003) Mixing equipment. *Chemical Industry Press*, 12-16.

[16] GB/T 4756-1998, (1998) Petroleum liquids-Manual sampling. *National standard press*, Beijing

[17] Xu, W. X. (2011) Analysis of the internal flow in a stirred tank bioreactor. *Hubei Agricultural Sciences* **13**, 2743-2745.

[18] Li, M. H. and Li, H. F. (2014) Study on Influencing Factors of Flow Field in Mixing Tank Bottom mixer. *Chemical Machinery* **6**:794-798.

[19] Cui, N. (2009) Simulation of Flow Field in Large Side Stirred Tank. *Shanghai Chemical Society* **2**.

[20] Chen, G. G., Zhang, L. J., Bo, Y. and Ye, X. C. (2012)Numerical Simulation of Influence of Blade Parameters on Flow Field and Power in Side-inlet Stirred Tank. *Journal of Beijing University of Chemical Technology* **3**, 29-34.

[21] Hou, H. G. (2010) Computational fluid dynamics simulation of gas-liquid flow in large stirred

fermentation tank. *Tianjin University*, China.

[22] Bi, J. W., Zhu, H. G., Shi, H. X., Li, Y. M., Rong, L. and Wang, T. (2010) CFD simulation and temperature field verification of mixing tank digester. *Journal of Agricultural Engineering* **10**, 283-289.

[23] Jiang, Y., Deng, L. K., Xu, G. H. and Lin, H. L. (2012) Simulation Study on Flow Field Characteristics of Side Stirred Fermentation Tank. *China Plant Engineering* **7**, 45-47.

[24] Ren, J. (2007) Simulation and Experimental Study on Flow Field and Dynamic Performance of Stirring Reactor. *Zhengzhou University, China*.

Patient-Specific Echo-Based Left Ventricle Models for Active Contraction and Relaxation Using Different Zero-Load Diastole and Systole Geometries

*** Longling Fan [1], Jing Yao[2], Chun Yang [3], Di Xu[2], † Dalin Tang [1,4]**

[1]School of Mathematics, Southeast University, Nanjing, 210096, China.
[2] Department of Cardiology, Nanjing Medical University, Nanjing 210029, China.
[3] China Information Tech. Designing & Consulting Institute Co., Ltd., Beijing, 100048, China.
[4] Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609.

*****Presenting author: fan_longling2008@126.com
†Corresponding authors. Dalin Tang, email: dtang@wpi.edu

Abstract

A new modeling approach using two different zero-load geometries (diastole and systole) was introduced to properly model active contraction and relaxation for more accurate stress/strain calculations. Ventricle diastole and systole material parameter values were also determined based on in vivo data. Echo image data were acquired from 10 healthy volunteers at the First Affiliated Hospital of Nanjing Medical University with consent obtained. Echo-based computational two-layer left ventricle (LV) models using one zero-load geometry (1G) and two zero-load geometries (2G) were constructed. Material parameter values in Mooney-Rivlin models were also adjusted during the cardiac cycle to match Echo volume data. Effective Young's moduli (YM) were calculated for ventricle materials for easy comparison.

Using the mean values of the 1G models as the baseline values, at begin-filling, the mean YM value for the fiber direction (YM_f) from the 2G model was 107% higher than that from the 1G model (723.57 kPa vs. 348.71 kPa). At begin-ejection, YM_f from 2G model was 47% lower than that from the 1G model (85.48 kPa vs. 162.77kPa). According to the total average values, begin-ejection stress and strain from the 2G model were 30% and 14.5% higher than that from the 1G model, respectively (345.16 kPa vs. 265.62 kPa; and 1.0489 vs. 0.9161). Begin-filling stress and strain from the 2G model were 11.5% and 55% higher than that from the 1G model, respectively (2.2613 kPa vs. 2.5543 kPa; and 0.0489 vs.0.1085). During a cardiac cycle, the 2G model begin-ejection YM_f, stress and strain were 19%, 495% and 29% higher than their end-filling value, end-ejection YM_f, stress and stain were 49%, 605% and 297% higher than their begin-filling value, respectively.

The 2G model took ventricle zero-load geometry difference between systole and diastole phases into consideration. This may lead to more accurate ventricle stress/strain calculations and material parameter value estimations.

Keywords: Ventricle modeling; active contraction; ventricle mechanics; Material properties.

Introduction

Myocardial contractility plays a central role in the cardiovascular circulatory system. A non-invasive method for estimating myocardial contractility would be a beneficial tool for cardiologists. As a first-order approximation, an approach using two zero-load geometries (2G) is proposed to model ventricle cardiac motion: one zero-load ventricle geometry is used to model the diastole phase where sarcomere has its relaxed zero-stress length, another zero-load ventricle

geometry is used to model the systole phase where sarcomere has its contracted zero-stress length (therefore the zero-load systole geometry is smaller than the zero-load diastole geometry). Essentially, we are using two models to model the cardiac cycle to handle the active contraction and relaxation which are caused by zero-stress sarcomere length changes.

Active contraction is caused by sarcomere shortening which leads to increased strain and stress (called active strain and stress). At the beginning of active contraction, the zero-stress sarcomere length is shortened in a very short time duration while ventricle volume has no change, so in vivo sarcomere length does not change, which leads to ventricle strain and stress increases, equivalent to the active tension in models in [4][8] for the stress part. McCulloch and Pfeiffer have made great contributions to passive and active ventricle modeling, such as the Physiome Project and the Continuity package [5][9][11]. Guccione et al. proposed the constitutive relations for active stress in cardiac muscle and developed three active tension models [2]-[3]. Liu et al. developed a dynamic cardiac elastography framework to assess the anisotropic viscoelastic passive properties and active contractility of myocardial tissues [6]. Pezzuto and Ambrosi focus on the contraction of the left ventricle in a finite elasticity framework, adopting the "prolate ellipsoid" geometry and the invariants-based strain energy proposed by Holzapfel and Ogden [10]. Our group introduced patient-specific cardiac magnetic resonance (CMR)-based right ventricle/left ventricle models with fluid-structure interactions with various surgical design and potential applications [12]-[15].

In this papers, a new modeling approach using different systole and diastole zero-load geometries was introduced to properly model active contraction and relaxation and obtain ventricle diastole and systole material parameter values, stress and strain conditions. New models were constructed for 10 patients and results were compared with our previously published one-geometry models [7].

Methods

Modeling active contraction and expansion by using different zero-load diastole and systole geometries

Since active LV contraction and relaxation are very complex and involve change of sarcomere zero-stress length which is hard to model using a single no-load LV geometry, some model simplifications are needed to obtain proper models to serve our purposes. Actual LV contraction and expansion involve two different RV zero-stress geometries (diastole and systole) and inter-connected changes of LV volume, pressure, stress/strain, and imposed active stress or active material properties. It is commonly accepted that a cardiac cycle may be divided into 4 phases, 2 in diastole (isovolumic relaxation followed by diastolic filling) and 2 in systole (isovolumic contraction followed by systolic emptying). To correctly model these 4 phases, two zero-stress geometries are needed. However, such two-geometry active contraction/expansion models are computationally expansive to construct. McCulloch et al. have introduced active tension in their sophisticated multiscale ventricle models with good success [2]-[3][5]. Tang et al. introduced LV/RV models with fluid-structure interactions using material stiffness variations to handle active contraction and relaxation [13]. Both active tension and stiffness variation approaches involved adding additional terms in tissue material strain energy functions.

Phase 1. Filling (diastole phase). The left ventricle starts with its minimum volume under minimum pressure with minimum stress and strain. One zero-load geometry (diastole geometry) is used for this phase, corresponding to diastole zero-stress sarcomere length (SL). It should be

noted that zero-stress status is a concept for stress/strain calculations. It is not observable in a living heart under in vivo conditions. At beginning-of-filling, mitral valve opens; LV volume increases, pressure increases, in vivo SL expands; strain and stress increases. Phase 1 ends when LV reaches its maximum volume under end-diastole pressure (denoted by Pdia) which is lower than the maximum pressure condition.

Phase 2. Isovolumic contraction: Both mitral (inlet) and aortic (outlet) valves are closed; LV volume has no change; zero-stress SL shortens (changing from diastole zero-stress length to systole zero-stress length); however, this sarcomere shortening is not physically observable. Roughly, average in vivo SL does not change much (small local SL changes are possible) since LV volume does not change. So zero-stress SL shortening leads strain and stress increase (This is similar to the active tension in other models, but our model have both strain and stress increase); increased stress pushes pressure to maximum. This phase is short. This phase involves dynamic change of zero-stress sarcomere length which is very difficult to implement. It was skipped in our model.

Phase 3. Ejection (systole phase): This phase starts from max volume, pressure, stress and strain. One zero-load geometry (for systole phase) is used for this phase, corresponding to systole zero-stress SL. At begin-ejection, aortic valve opens up and ejection starts; LV volume drops; in vivo SL shortens and strain decreases; pressure drops; stress drops. At end-systole (end-ejection), LV volume reaches its minimum, pressure drops to the end-systole pressure denoted as Psys, which is greater than minimum pressure. Pressure will continue to drop in Phase 4 when systole zero-stress SL changes to diastole zero-stress SL.

Phase 4. Isovolumic relaxation: Aortic valve closes (both valves closed); zero-stress SL relaxes from systole zero-stress length to diastole zero-stress length (non-contracted length); similar to the comments made in Phase 2, roughly, average in vivo SL does not change much since volume does not change; zero-stress SL relaxation leads to strain and stress decreases; pressure drops to minimum. This phase is short. It was also skipped in our model.

3D echo data acquisition

Patients were recruited to participate in this study with consent obtained (n=10, 7 males, mean age 54.9 years). Echo data acquisitions were performed at the First Affiliated Hospital of Nanjing Medical University, Nanjing, China. Standard echocardiograms were obtained using an ultrasound machine (E9, GE Mechanical Systems, Milwaukee, Wisconsin) with a 3V probe. Patients were examined in the left lateral decubitus position, and images were acquired at end expiration in order to minimize global cardiac movement. Details of the data acquisition procedures were previously described and are omitted here [7]. Figure 1 shows the echo images, zero-load and re-constructed 3D LV geometries.

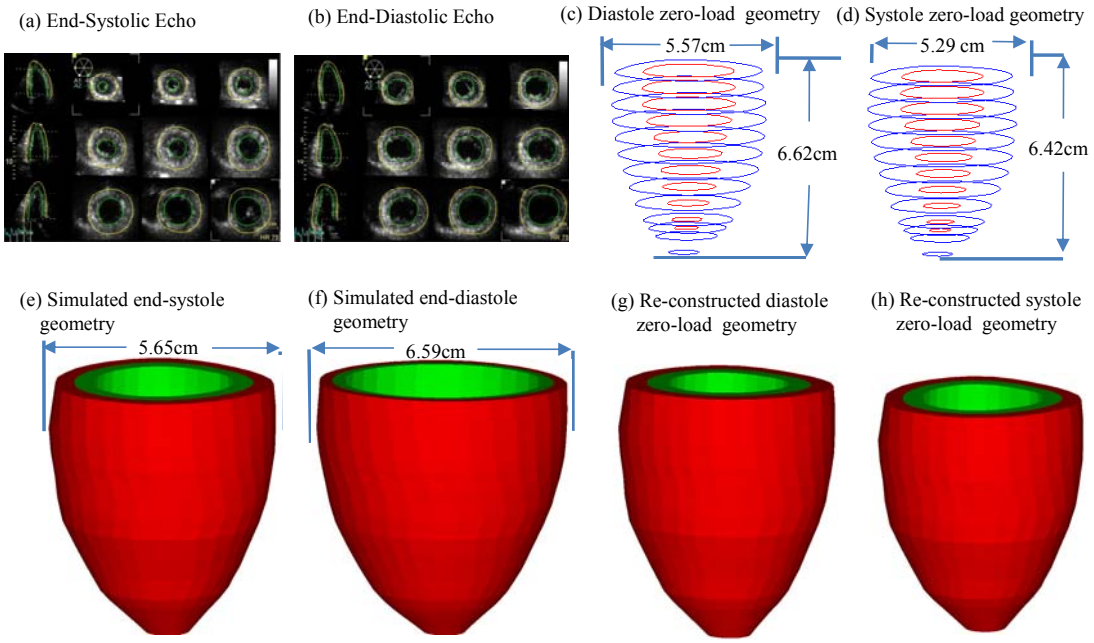


Figure 1. Echo image of a healthy volunteer, contours, zero-load diastole and systole geometries and re-constructed pressurized geometries.

Two-layer anisotropic LV moel construction with fiber orientations

Standard governing equations and boundary conditions for the LV model were given:

$$\rho v_{i,tt} = \sigma_{ij,j}\ ,\ i,\ j = 1,2,3;\ \text{sum over}\ \ j, \tag{1}$$

$$\varepsilon_{ij} = (v_{i,j} + v_{j,i} + v_{\alpha,i} v_{\alpha,j})/2,\ \ i,j,\alpha = 1,2,3, \tag{2}$$

where σ is the stress tensor, ε is the strain tensor, v is displacement, and ρ is material density. The normal stress was assumed to be zero on the outer (epicardial) LV surface and equal to the pressure conditions imposed on the inner (endocardial) LV surfaces.

The ventricle material was assumed to be hyperelastic, anisotropic, nearly-incompressible and homogeneous. The nonlinear Mooney-Rivlin model was used to describe the nonlinear anisotropic material properties. The strain energy function for the anisotropic modified Mooney-Rivlin model is given:

$$W = c_1(I_1-3) + c_2(I_2-3) + D_1[\exp(D_2(I_1-3)) - 1] + K_1/(2K_2)\exp[K_2(I_4-1)^2 - 1], \tag{3}$$

where I_1 and I_2 are the first and second strain invariants given by,

$$I_1 = \sum C_{ii},\ \ I_2 = \tfrac{1}{2}[I_1^2 - C_{ij}C_{ij}],\ \ I_4 = C_{ij}(\mathbf{n}_f)_i(\mathbf{n}_f)_j \tag{4}$$

$C = [C_{ij}] = X^T X$ is the right Cauchy-Green deformation tensor, $X = [X_{ij}] = [\partial x_i / \partial a_j]$, (x_i) is the current position, (a_i) is the original position, $\mathbf{n}_f$ is the fiber direction, c_i, D_i and K_i are material parameters chosen to match experimental measurements [1][13]. With parameters properly chosen, it was shown that stress-strain curves derived from Eq. (3) agreed very well with the stress-strain curves from the anisotropic (transversely isotropic) strain-energy function with respect to the local fiber direction given in McCulloch et al.[8]:

$$W = \frac{C}{2}(e^Q - 1), \tag{5}$$

$$Q=b_1E_{ff}^2+b_2(E_{cc}^2+E_{rr}^2+E_{cr}^2+E_{rc}^2)+b_3(E_{fc}^2+E_{cf}^2+E_{fr}^2+E_{rf}^2), \quad (6)$$

where E_{ff} is fiber strain, E_{cc} is cross-fiber in-plane strain, E_{rr} is radial strain, and E_{cr}, E_{fr} and E_{fc} are the shear components in their respective coordinate planes, C, b_1, b_2, and b_3 are parameters to be chosen to fit experimental data. For simplicity, we set b_1=0.8552, b_2=1.7005, b_3=0.7742 in Eq. (6) so that we can have a single parameter C for comparison. The least-squares method was used to find the equivalent Young's moduli (YM) for the material curves for easy comparison. Active contraction and expansion of myocardium were modeled by material stiffening and softening in our model. Material stiffening and softening were achieved by adjusting parameter values at each Echo-time step (28 Echo frames per cycle) to simulate active contraction and expansion and match LV volume data.

As patient-specific fiber orientation data was not available from these patients, we chose to construct a two-layer LV model and set fiber orientation angles using fiber angles given in Axel [1]. Fiber orientation angles were set at -60 degree and 80 degree for epicardium (outer layer) and endocardium (inner layer), respectively. Fiber orientation can be adjusted when patient-specific data becomes available [12].

A pre-shrink process and geometry-fitting technique for mesh generation

Under in vivo condition, ventricles were pressurized and the no-load ventricular geometries were unknown. In our model construction process, an iterative pre-shrink process was applied to the in vivo minimum volume ventricular geometry to obtain the two zero-load geometries so that when in vivo pressure was applied, the ventricle would regain its in vivo geometry. To get the zero-load diastole geometry, we start with a 4% shrinkage, construct the model, and apply the minimum pressure to see if the pressurized LV volume matches the Echo data. If not, we adjust the shrinkage, re-made the model, pressurize it and check again. The process is repeated until LV volume matches Echo volume with error < 0.5%. For the zero-load systole geometry, assuming a 10-15% sarcomere shortening, we start with a 14% shrinkage. Different shrinkage rates were used for LV inner and outer surfaces so that mass conservation law was enforced. The same process was repeated until the pressurized LV volume under end-systole pressure matched the Echo-measured end-systole volume data.

A geometry-fitting mesh generation technique was also used to generate mesh for our models [13]. Mesh analysis was performed by decreasing mesh size by 10% (in each dimension) until solution differences were less than 2%. The mesh was then chosen for our simulations.

Solution methods and Data collection for Statistical analysis

The anisotropic LV computational models were constructed for ten patients and the models were solved by ADINA (ADINA R&D, Watertown, MA, USA) using unstructured finite elements and the Newton-Raphson iteration method. Stress/strain distributions were obtained for analysis. Because stress and strain are tensors, for simplicity, maximum principal stress (Stress-P_1) and strain (Strain-P_1) were used and referred to as stress and strain in this paper. For each LV data set (11 slices. Slices are short-axis cross sections), we divided each slice into 4 quarters, each quarter with equal inner wall circumferential length. Ventricle wall thickness, circumferential curvature (C-curvature), longitudinal curvature (L-curvature) and stress/strain were calculated at all nodal points (100 points/slice, 25 points/quarter). The "quarter" values of those parameters were obtained by taking averages of those quantities over the 25 points for each quarter and

saved for analysis. The quarter values of those from the ten patients were compared to see if there are any statistically significant differences.

Results and Discussion

The purpose of this paper is to introduce the new model with 2 zero-load geometries (2G model), compare the results with our previous model which used 1 zero-load geometry (1G model). For the 1G model, results at begin-filling (BF) and begin-ejection (BE) corresponding to minimum and maximum pressure and LV volume were obtained for comparison. For the 2G model, results at begin-filling (BF), end-filling (EF), begin-ejection (BE), and end-ejection (EE) were obtained for comparison. Due to 1G model limitation, EF and BE from 1G model had the same geometry, pressure and material conditions. Therefore, 1G EF and BE stress/strain values were also the same.

Human ventricle tissue material properties are extremely hard to quantify noninvasively under in vivo conditions. Myocardium material stiffness changes in a cardiac cycle. Figure 2 gives the stress-stretch curves for 1G and 2G models from one patient to illustrate the material differences. Bar plots of mean YM value in the fiber direction (YM_f) of the 10 patients from 1G and 2G models were in Fig. 3 (a) showing clear comparisons. Using the mean values of 1G models as the baseline values, at BF, YM_f from 2G model was 107% stiffer than that from the 1G model (723.57 kPa vs. 348.71 kPa). At BE, YM_f from 2G model was 47% lower than that from the 1G model (85.48 kPa vs. 162.77kPa). At EF, YM_f from 2G model was 56% lower than that from the 1G model (71.683 kPa vs. 162.77kPa). At EE, YM_f from 2G model was 210% higher than that from the 1G model (1080.95 kPa vs. 348.71 kPa). This indicated the material parameter properties from 2G models were stiffer than that from 1G model at BF and EE corresponding to minimum LV volume, and softer than that from 1G model at BE and EF corresponding to maximum LV volume. From 2G models, BE YM_f was 19% higher than its EF value (85.48 kPa vs. 71.68 kPa), EE YM_f was 49% higher than its BF value (1080.95 kPa vs. 723.57 kPa).

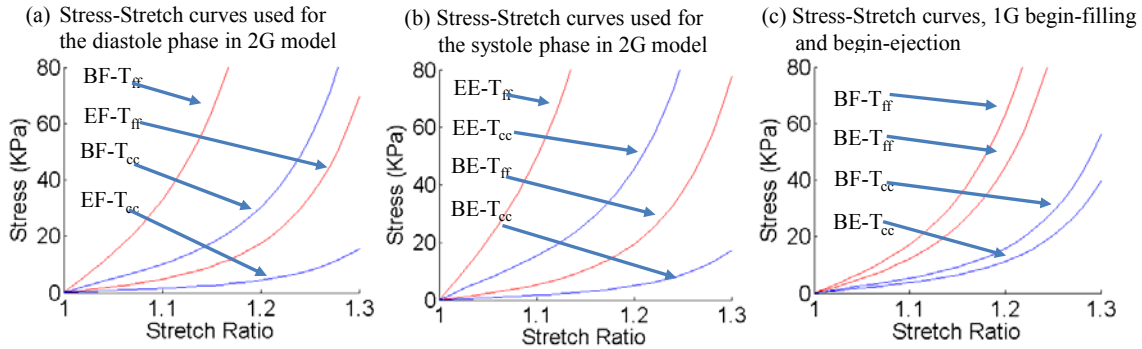


Figure 2. Material Stress-Stretch curves and YM_f comparison for 1G and 2G models. T_{ff}: stress in fiber direction; T_{cc}: stress in circumferential direction. YM_f: Young's Modulus in fiber direction. BF: Begin-Filling; EF: End-Filling. BE: Begin-Ejection; EE: End-Ejection.

Table 1. Material parameters from the 2G and 1G models.

	C(kPa)	YM_f(kPa)	C(kPa)	YM_f(kPa)	C(kPa)	YM_f(kPa)
	1G-BF		**2G-BF**		**2G-EF**	
P1	10.102	290.47	18.9420	544.62	2.3813	68.467
P2	11.185	321.59	25.9776	746.91	2.3452	67.429
P3	11.365	326.77	21.6480	622.43	3.1390	90.252
P4	10.283	295.65	19.1224	549.81	2.8503	81.953
P5	11.419	328.33	19.8440	570.56	2.1648	62.243
P6	14.432	414.95	28.8640	829.91	1.7860	51.350
P7	13.530	389.02	31.5700	907.70	1.5695	45.126
P8	10.102	290.47	21.6480	622.43	4.4198	127.08
P9	19.483	560.18	46.9040	1348.59	1.2628	36.308
P10	9.3808	269.72	17.1380	492.75	3.0127	86.621
Mean	**12.128**	**348.71**	**25.1658**	**723.57**	**2.4931**	**71.683**
	1G-BE		**2G-BE**		**2G-EE**	
P1	5.7367	164.94	2.4895	71.579	31.029	892.14
P2	5.2857	151.98	2.6519	76.247	38.425	1104.8
P3	8.0278	230.82	3.4998	100.625	32.652	938.82
P4	7.7211	102.23	3.7523	107.887	31.390	902.52
P5	6.0073	172.72	2.3272	66.911	33.735	969.95
P6	4.5100	129.67	2.0205	58.093	38.786	1115.2
P7	3.6621	105.29	1.8040	51.869	45.100	1296.7
P8	9.4710	272.31	6.4944	186.728	37.884	1089.2
P9	2.9586	85.065	1.3710	39.420	52.316	1504.2
P10	7.3964	212.66	3.3194	95.439	34.637	995.88
Mean	**6.0777**	**162.77**	**2.9730**	**85.480**	**37.595**	**1080.95**

Table 2 summarizes the average stress and strain values of the 10 patients from the 1G and 2G models. Bar plots of mean stress/strain values of the 10 patients from 1G and 2G models were in Fig. 3 (b) and (c) showing clear comparisons. According to the total average values in Table 3, BE stress values from the 2G model was 30% higher than that from the 1G model (345.16 kPa vs. 265.62 kPa). BF stress values from the 2G model was 14.5% lower than that from the 1G model (2.2613 kPa vs. 2.5543 kPa). BE strain values from the 2G model was 11.5% higher than that from the 1G model (1.0489 vs. 0.9161). BF strain values from the 2G model was 55% lower than that from the 1G model (0.0489 vs.0.1085). From 2G models, BE stress average value was 511% higher than its EF value (354.16 kPa vs. 57.96 kPa), EE stress average value was 605% higher than its BF value (15.94 kPa vs. 2.26 kPa). BE strain average value was 29% higher than its EF value (1.0489 vs. 0.8110), EE strain average value was 297% higher than its BF value (0.1942 vs. 0.0489). Bar plots of wall thickness and curvature results in Fig. 3 (d)-(f) shows the geometrical characteristics from the 1G and 2G models were about the same.

Table 2. Comparison of average stress and strain results from 2G and 1G models

Patient	1G-BF	1G-BE	2G-BF	2G-EF	2G-BE	2G-EE
Average stress						
P1	2.3290	224.68	1.9345	47.986	311.31	15.429
P2	2.8488	293.62	2.6241	58.033	372.43	17.785
P3	2.3849	185.83	2.0071	37.989	265.97	15.331
P4	2.4531	186.16	1.8084	30.841	255.31	14.907
P5	2.5214	290.74	2.3700	62.283	401.50	16.690
P6	2.2008	277.08	2.0724	64.571	353.67	13.015
P7	1.8442	292.55	1.7388	74.045	363.51	12.619
P8	3.4650	221.26	2.8673	41.191	261.86	18.875
P9	2.7979	408.15	2.6838	105.48	495.94	15.139
P10	2.6982	276.19	2.5062	57.221	370.09	19.644
Mean	**2.5543**	**265.62**	**2.2613**	**57.964**	**345.16**	**15.943**
Average strain						
P1	0.1121	0.8845	0.0508	0.7820	1.0297	0.2120
P2	0.1228	0.9206	0.0502	0.7972	1.0360	0.2021
P3	0.1103	0.8492	0.0497	0.7437	1.0135	0.2172
P4	0.1215	0.8438	0.0496	0.7150	0.9860	0.2145
P5	0.1110	0.9293	0.0613	0.8437	1.0948	0.2185
P6	0.0826	0.9910	0.0393	0.9049	1.1325	0.1700
P7	0.0716	0.9935	0.0291	0.9102	1.1110	0.1403
P8	0.1514	0.8061	0.0615	0.6632	0.8768	0.2042
P9	0.0726	1.0780	0.0289	0.9904	1.2036	0.1403
P10	0.1292	0.8654	0.0684	0.7598	1.0053	0.2227
Mean	**0.1085**	**0.9161**	**0.0489**	**0.8110**	**1.0489**	**0.1942**

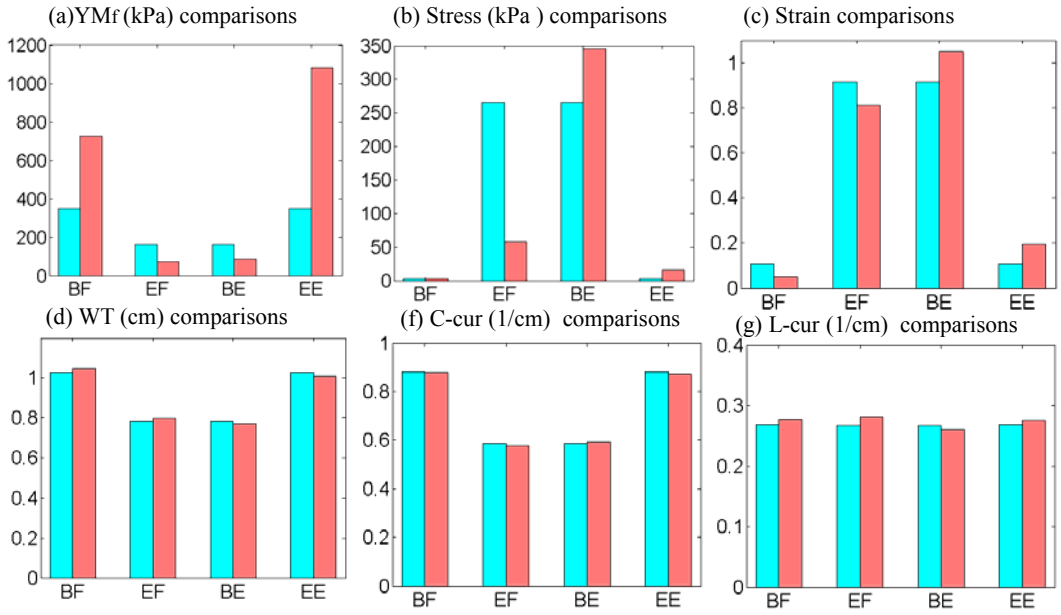


Figure 3. Bar plots for LV YM_f, Stres, Strain, wall thickness (WT), circumferential curvatures (C-cur) , longitudinal curvature (L-cur) comparisons from 1G and 2G models. Blue: 1G; Red: 2G.

Conclusions

Correct stress/strain calculation is of fundamental importance for many cardiovascular research where mechanical forces play a role. Ventricle remodeling, disease development, tissue regeneration, patient recovery after surgery and many other cell activities are closely associated with ventricle mechanical conditions. The 2G modeling approach is setting up the right stage for diastole and systole stress/strain calculations using proper zero-load geometries. 1G models do not use different reference geometries for systole and diastole phases, therefore have difficulties in giving right strain calculations. It should be noted that direct measurements of stress, strain, and zero-load sarcomere length are either extremely difficult or even impossible. Even by using tagging, the strain determined uses in vivo references and could not account for zero-stress SL changes. Actual ventricle contraction and relaxation are very complex. Our model is only a first-order approximation, an improvement over the 1G models. Lack of in vivo data and model construction cost are also considerations. Data from the literature or from ex vivo experiments have to be used to complete the computational models. We are in need of patient-specific data such as fiber orientation, sarcomere length contraction rate, regional material properties, etc.

Acknowledgement. This research was supported in part by National Sciences Foundation of China grants 11672001, 81571691. Longling Fan's research is supported in part by the Fundamental Research Funds for the Central Universities (KYLX15_0110) and the Scientific Research Foundation of Graduate School of Southeast University (YBJJ1617).

References

[1] Axel, L. (2002) Biomechanical dynamics of the heart with MRI, *Annu. Rev. Biomed. Eng.* **4**, 321-347.
[2] Guccione, J. M. and McCulloch, A. D. (1993a) Mechanics of active contraction in cardiac muscle: Part I–Constitutiverelations for fiber stress that describe deactivation, *J. Biomech. Eng.* **115**, 72–81.
[3] Guccione, J. M., Waldman, L. K. and McCulloch, A. D. (1993) Mechanics of active contraction in cardiac muscle: Part II-Cylindrical models of the systolic left ventricle, *J. Biomech. Eng.* **115**, 82–90.
[4] Hunter, P. J., Pullan, A. J., Smaill, B. H. (2003) Modeling total heart function. *Annu Rev Biomed Eng.* **5**, 147-77.
[5] Kerckhoffs, R. C. P., Healy, S. N., Usyk, T. P. and McCulloch, A. D. (2006) Computational methods for modeling cardiacelectromechanics, *Proc IEEE.* **94**,769-783.
[6] Liu, Y., Wang, G. and Sun, L. Z. (2006) Anisotropic elastography for local passive properties and active contractility of myocardium from dynamic heart imaging sequence, *Int J Biomed Imaging.* **2006**, 45957.
[7] Fan, L. L., Yao, J., Yang, C., Tang, D. and Xu, D. (2015) Infarcted Left Ventricles Have Stiffer Material Properties and Lower Stiffness Variation: 3D Echo-Based Modeling to Quantify In Vivo Ventricle Material Properties, *J.Biomech. Eng.* **137**, 081005.
[8] McCulloch, A. D., Waldman, L., Rogers, J. and Guccione, J. M. (1992) Large-scale finite element analysis of the beating heart, *Critical Rev. in Biomed. Eng.* **20**, 427-449.
[9] McCulloch, A. D., et al. (2007) Continuity 6, www.continuity.ucsd.edu.
[10] Pezzuto, S. and Ambrosi, D. (2014) Active contraction of the cardiac ventricle and distortion of the microstructural architecture, *Int J Numer Method Biomed Eng.* **30**, 1578-1596.
[11] Pfeiffer, E., Tangney, J. R., Omens, J. H. and McCulloch, A. D. (2014) Biomechanics of cardiac electromechanical coupling and mechanoelectric feedback, *J. Biomech. Eng.* **6**, 021007.
[12] Tang, D., Yang, C., Geva, T. and del Nido, P. J. (2008) Patient-specific MRI-based 3D FSI RV/LV/Patch models for pulmonary valve replacement surgery and patch optimization, *J. of Biomech. Eng.* **130**, 041010.
[13] Tang, D., Yang, C., Geva, T. and del Nido, P. J. (2010) Image-Based Patient-Specific Ventricle Models with Fluid-Structure Interaction for Cardiac Function Assessment and Surgical Design Optimization, *Progress in Pediatric Cardiology* **30**, 51-62.
[14] Tang, D., Yang, C., Geva, T., Gaudette, G. and del Nido, P. J. (2011) Multi-physics MRI-based two-layer fluid-structure interaction anisotropic models of human right and left ventricles with different patch materials: cardiac function assessment and mechanical stress analysis, *Computers & Structures* **89**, 1059-1068.
[15] Tang, D., del Nido, P. J., Yang, C., Zuo, H., Huang, X. Y., Rathod, R. H., Tang, A., Wu, Z. Y., Billiar, K. L. and Geva, T. (2016) Patient-Specific MRI-Based Right Ventricle Models Using Different Zero-Load Diastole and Systole Geometries for Better Cardiac Stress and Strain Calculations and Pulmonary Valve Replacement Surgical Outcome Predictions, *PLOS One* **11**, e0162986.

Multi-Field Coupled Simulation of Aqueous Humor in Rabbit Eye

***H.F. Song, L. Li, W.J. Wang and †Z.C. Liu**

School of Biomedical Engineering, Capital Medical University, China
Beijing Key Laboratory of Fundamental Research on Biomechanics in Clinical Application, Capital Medical University, Beijing, China

*Presenting author: songhf@ccmu.edu.cn
†Corresponding author: zcliu@ccmu.edu.cn

Abstract

Understanding the variation of the pressure difference between Anterior Chamber (AC) and Posterior Chamber (PC), as well as the state of aqueous flow, is the focus problem to know more about the progression and blinding mechanism of glaucoma. Aqueous humor plays an important role in intraocular pressure (IOP) and the increased IOP is one of the major risk factors for the primary angle closure glaucoma. Many studies did not consider the physiological variation of the Pressure Difference between PC and AC (PDPA), and the multi-field coupling effects. The aim of the present study is to investigate the aqueous humor flow based on the coupling effects of fluid-solid-heat by means of numerical simulation method.

Two models were constructed based on eye rabbit tissue sections. Model A is for multi-field coupling and Model B is for fluid-solid interaction under condition of normal IOP. ADINA (ADINA Inc.， USA) software was applied to mimic the aqueous humor flow and analyze the distribution of temperature, velocity, pressure and stress.

The contour of temperature demonstrates multiple peaks with larger value near the pupil. There is a temperature difference in the anterior segments, which plays an important role in the aqueous humor flow and affects the flow pattern and pressure distribution. The parameters demonstrate larger value when considering the temperature.

Considering the low temperature of 33^{o}C on the surface of cornea is of significance to exhibit the temperature distribution, flow pattern and pressure distribution of the aqueous humor flow. It is necessary to perform multi-field coupling simulation based on realistic PDPA measurement when studying the aqueous humor flow in eyes.

Keywords: Glaucoma; Fluid-Structure Interaction; Numerical Simulation; Aqueous Humor Flow; Ocular biomechanics

Introduction

Aqueous humor is developed in the Posterior Chamber (PC) at the ciliary body, then flows into the Anterior Chamber (AC) through the pupil, and is finally absorbed through trabecular meshwork at the angle of AC. Under steady flow condition, the ciliary body generates a steady aqueous flow rate of 2.5 μl/min[1].This flow process plays an important role in intraocular pressure (IOP) and the increased IOP is one of the major risk factors for the primary angle closure glaucoma. While the resistance of aqueous flow through the iris-lens channels increases, the Pressure Difference between PC and AC (PDPA) increases. Too high IOP will vary the aqueous humor flow patterns, the pressure distributions in the flow field and the deformations and stress distributions of it is and cornea. PDPA is the important force to maintain aqueous humor flow. Therefore, understanding the variation of PDPA, as well as the

state of aqueous flow, is the focus problem to know more about the progression and blinding mechanism of glaucoma.

The geometry dimensions of anterior segment are very small so that it is very difficult to measure its interior flow patterns via in vivo experiments. With the development of computer technology and finite element analysis, numerical simulation can be applied to mimic the flow field in the small domain of eyes.

As early as 1986, Ethier[2] calculated the flow impedance when aqueous humor flows out from juxtacanalicular. In 1988, Scott[3] performed numerical simulation of the aqueous humor flows in AC according to the temperature difference in AC. Based on a simplified 2D AC model according to the anatomical structure of eye. Canning[4] mimicked the aqueous humor flows in AC. In 2006 for the first time, Kumar [5] considered the porous property of trabecular meshwork, and constructed a simplified symmetric 3D AC model for aqueous humor flow of rabbit eye. Then, the IOP and aqueous humor flow were compared for different situations of temperature difference and pupil size. With regard to the pulsatile flow of aqueous humor from PC to AC, Abdulrazik [6] built a model to investigate the high risk factors of pupil impedance, and found the fluid dynamics difference between pulsatile flow and steady flow. Based on the establishment of a 2D model of aqueous humor flow in human AC, Ooi[7] studied the influence of aqueous humor flow on the distribution of intraocular temperature. However, this work lacks the support of experimental data, and the model was also an idealized simple model. Based on a 3D finite element model, Chai [8] mimicked influence of laser treatment on the aqueous flow and IOP, and the results might provide guidance for the treatment of decreasing IOP. A 3D symmetric model of human eye based on histological sections was built by Villamarina[9] in 2012, which considered the permeability of trabecular mesh and the temperature difference in AC, and then the flow pattern was simulated. The limitations of this work include no Fluid-Solid Interaction (FSI) was considered, the applied data was from hypothesis, and the reliability was discounted.

Although there are so many studies about aqueous humor flow, most of them are greatly simplified without considering the physiological variation of PDPA, and the coupling effects of fluid-solid-heat are rarely implemented.

The aim of the present study is to investigate the aqueous humor flow based on the coupling effects of fluid-solid-heat by means of numerical simulation method. Based on 24h continuous in vivo test of PDPA for normal IOP rabbit eyes, 3D rabbit eye models were reconstructed based on eye tissue sections, and finite element models for multi-field coupling analyses were established so as to mimic and analyze the variation rules of aqueous humor flow.

Methods

Geometry model construction of rabbit ocular segment

Histological slices of rabbit eye with HE dyeing (Figure 1) were applied to reconstruct the 3D geometry of ocular segment so as to obtain a more realistic model. Lens was not included in the slices. Thus, the lens was designed according to its realistic dimension when reconstructing the 3D model. The gap between the lens and iris at the pupil edge is H=10 μm for normal chamber angle. Model for normal rabbit eye was established according the anatomical characteristics of eye ball and the clinical characteristics of primary angle closure glaucoma (Figure 2).

Two models were constructed based on eye rabbit tissue sections. Model A is for multi-field coupling and Model B is for fluid-solid interaction.

Figure 1 Histological sections of rabbit eye

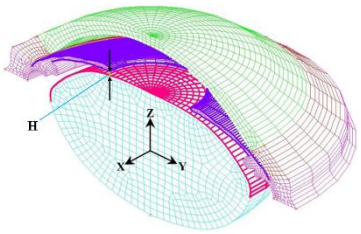


Figure 2 Reconstructed geometry model of rabbit eye

Model assumptions and parameters

Iris, cornea and lens are all of linear elastic property. The aqueous humor is incompressable laminar viscous Newtonian fluid. The upper and lower surfaces of iris and the inner suface of cornea are contacted with aqueous humor, and they are the interface of fluid-solid interaction where the no-slip condition is satisfied.

The temperature at the inner surface of cornea is 33 ^{o}C, and elsewhere 37 ^{o}C [10, 11]. The structure of trabecular meshwork at the chamber angle is porous with the porosity of α=0.5 and permeability of κ=1e^{-6} mm/s[5]. The rabbit eyes are in laying position.

Figure 3 shows the experimental data of PDPA which are the pressure boundary conditions for numerical simulation.

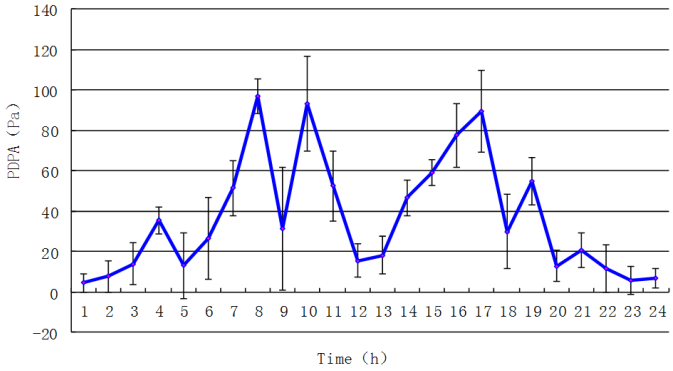


Figure 3 Experimental data of PDPA in rabbit eyes under normal IOP

The displacement of lens at y-z plane in x direction is 0 (Figure 4); The displacements at x-y plane in x, y, z directions are all 0; The displacement of trabecular meshwork at y-z plane in x direction is 0, and elsewhere are free; The displacement of cornea at y-z plane in x direction is 0, and elsewhere are free.

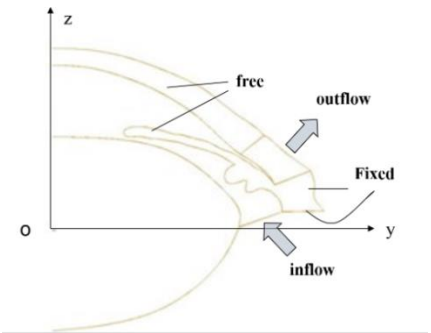


Figure 4 Boundary conditions

Material properties

According to references [7,12,13], the material properties of each part are listed in Table 1.

Table 1 Material properties

Parameter	E (kPa)	υ	c_p J(kg.K)	K (W/m.K)	ρ (kg/m^3)	μ (kg/ms)	β (1/K)
Cornea	1.19	0.49	3180	1.0042	-	-	-
Iris	6.1	0.45	4178	0.58	-	-	-
Lens	10^5	0.3	3000	0.4	-	-	-
Aqueous humor	-	-	4182	0.6	10^3	10^{-3}	$3*10^{-4}$

Note: E - Young's modulus, υ - Poison's ratio, c_p - Specific Heat Capacity, κ - coefficient of heat conductivity, ρ - density, μ - viscosity, β - coefficient of cubic expansion.

ADINA (ADINA Inc., USA) software was applied to mimic the aqueous humor flow under the multi-field coupling of fluid-solid-heat.
Five typical nodes in the model (N1~N5 in Figure 5) were selected to show the variation of flow velocity with time.

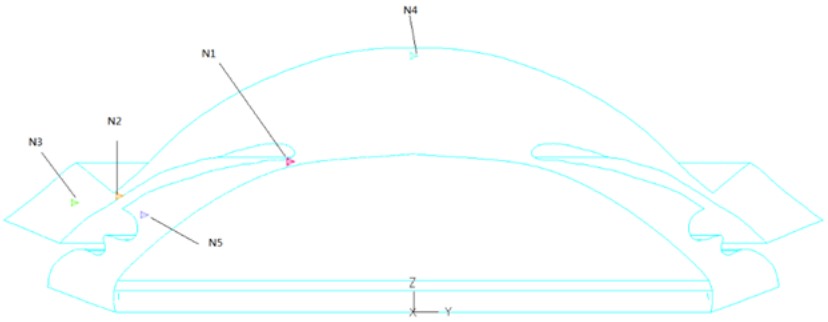


Figure 5 Nodes N1 to N5

Results

Figure 6 shows the distribution of temperature in Model A. The temperature distribution varies with time. The temperature is symmetric and decreases smoothly from iris to cornea. The contour of temperature demonstrates multiple peaks with larger value near the pupil. Temperature distribution at 8:00 AM (t=28800 s) corresponds to the moment with the largest PDPA.
Figure 7 and Figure 8 show the velocity distribution of Model A and Model B respectively. The velocity distribution is symmetric and there is a vortex in the left and right side respectively. The maximum velocity appears at the gap between the iris and pupil. The flow patterns in both models are similar while the velocity magnitude in model A is larger than that in model B.

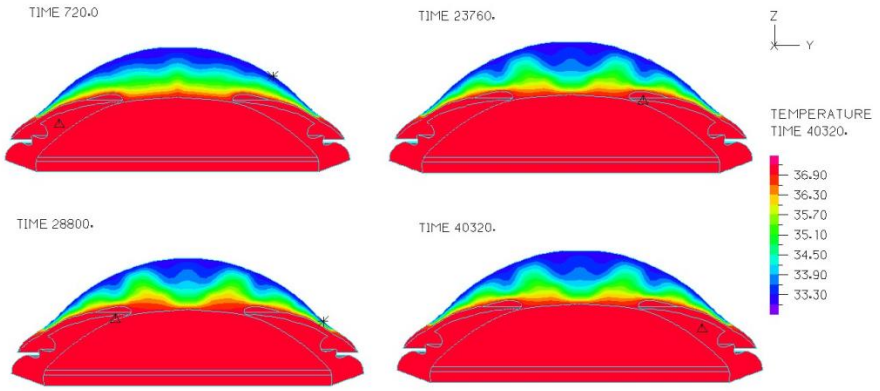


Figure 6 Distribution of temperature in model A

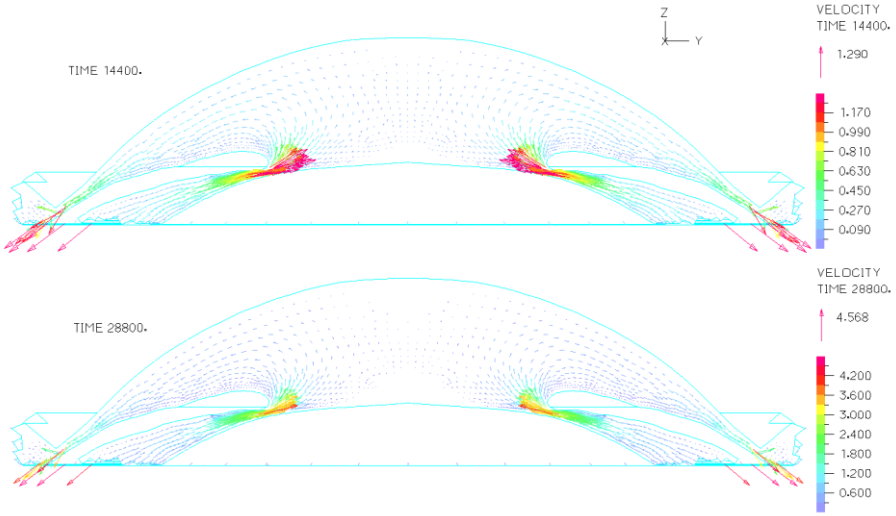


Figure 7 Velocity vector in Model A

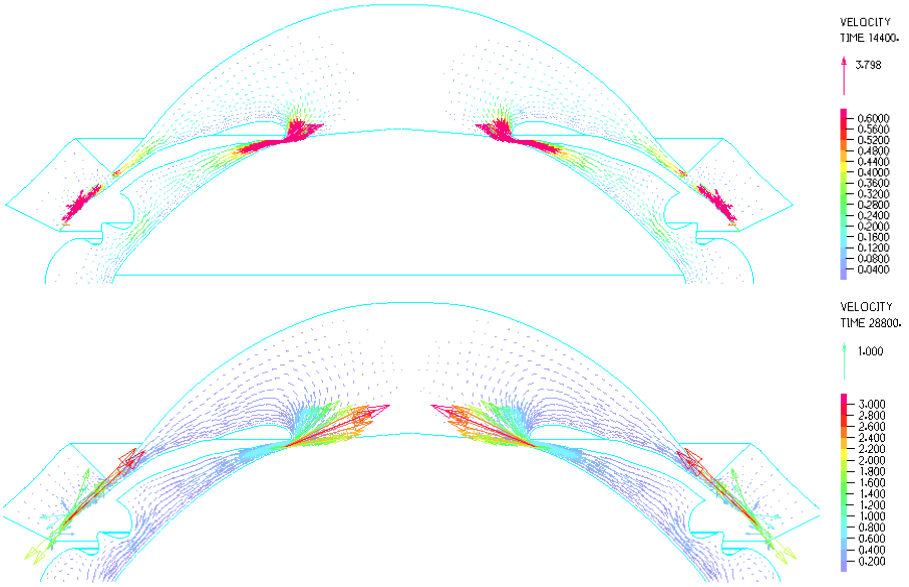


Figure 8 Velocity vector in Model B

Figure 9 shows the variation of velocities at the typical nodes in Model A. The variation rule of velocity is in agreement with that of PDPA, but the variation magnitudes are a little different. The maximum magnitudes are at N1 and N2 and that at N4 is largely constant. That is, the maximum variation of velocity locates at the gap between the iris and pupil and also at the chamber angle, while the velocity at the center of cornea rarely changes. The velocity at the gap between iris and pupil can reach 3.5 mm/s when the pressure difference is at the peak value, and the velocity at the center of cornea is nearly 0 which indicates that there is a stagnant point. The variation of velocities at the typical nodes in Model B is similar to that in Model A. Only the velocity at the AC is somewhat different. The velocity at N4 in Model A is 5.5e-3 mm/s while that in Model B is 1.49e-3 mm/s.

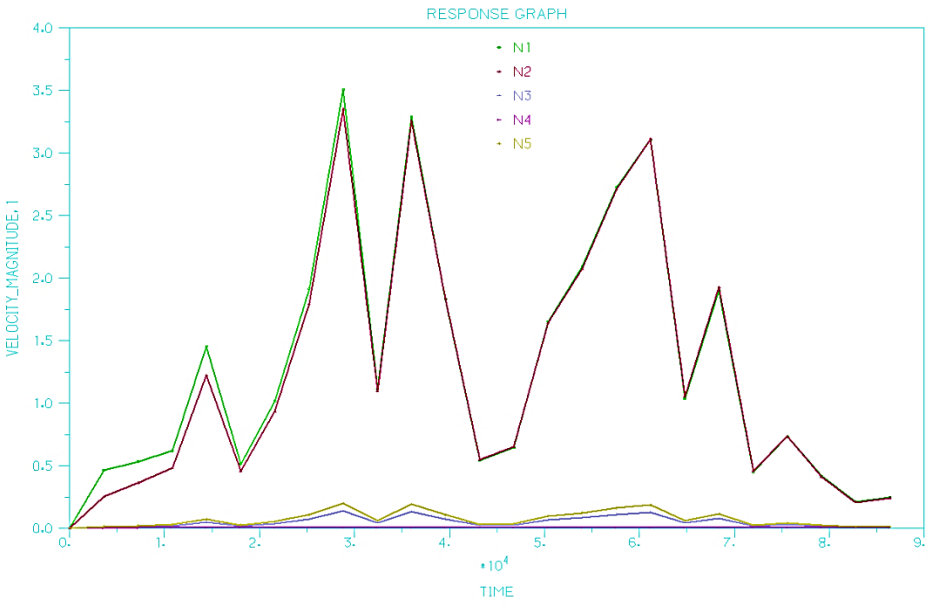


Figure 9 Variations of node velocity in Model A.

Figure 10 and Figure 11 show the contours of node pressure in Model A and Model B respectively. Pressure difference decreases gradually from PC to the cornea surface of AC. The pressure distributions in both models are similar while the pressure magnitude in model A is larger than that in model B.

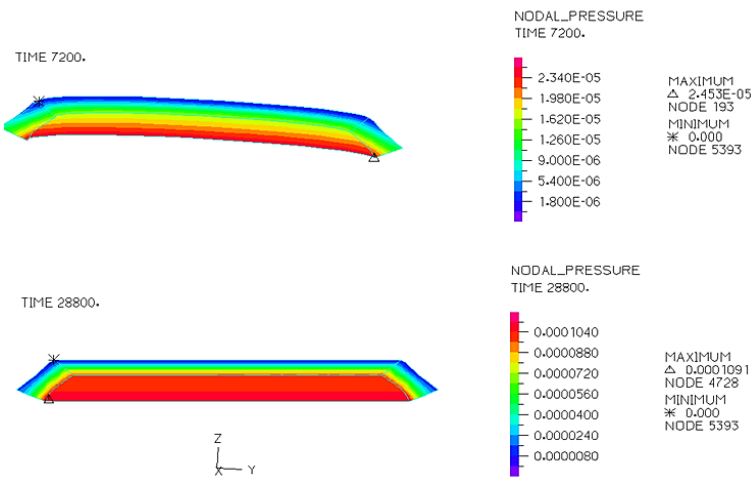


Figure 10 Contour of node pressure in Model A

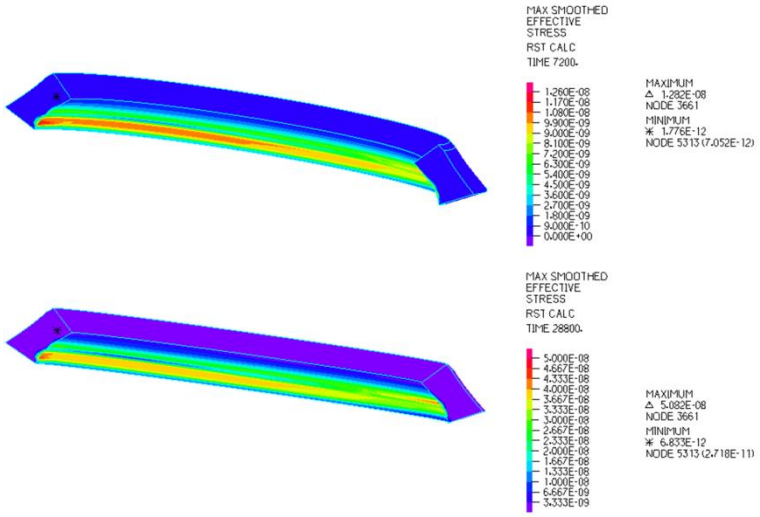


Figure 11 Contour of node pressure in Model B

Generally speaking, the parameters demonstrate larger value when considering the temperature, which indicates that the intraocular temperature plays significant rule in the aqueous humor flow, and it can influence the velocity, pressure and iris deformation.

Discussion

The distributions of temperature, velocity and pressure were investigated in this paper by means of numerical simulation. Multi-field coupling effect, which may influence the distributions of flow pattern, pressure and temperature, were considered in this study.

Results of temperature distribution in the present study are in accordance with those of Kumar[5], while quite different from those of Ooi E[7] who employed a 2D model where the geometry and boundary conditions are different.

The cornea is exposed to the air, and then some factors, such as the evaporation of tears, etc. can make the temperature of cornea outer surface equal to that of the surrounding environment, while the temperature of cornea inner surface is between 32-34^{o}C and the temperature of iris surface is the same as the body temperature of 37^{o}C. Thus, there is a temperature difference in the anterior segments, which plays an important role in the aqueous humor flow and affects the flow pattern and distribution.

The temperature difference in the AC induced buoyancy force drives the aqueous humor to flow through the iris and the cornea, then two symmetric vertices develop in the AC. This phenomenon is similar to the 2D model flow pattern in references[5,7]. The reversed flow pattern with the pressure in AC larger than that in PC can be confirmed by reference[14].

Conclusions

Considering the low temperature of 33^{o}C on the surface of cornea is of significance to exhibit the temperature distribution, flow pattern and pressure distribution of the aqueous humor flow. It is necessary to perform multi-field coupling simulation based on realistic PDPA measurement when studying the aqueous humor flow in eyes. The present work may provide reference value for understanding the flow patterns of aqueous humor flow in rabbit eye, and can help to investigate the mechanism of the development and blinding for primary angle closure glaucoma.

Acknowledgement: This work was supported by Beijing Natural Science Foundation (7152022) and National Natural Science Foundation of China (10802053).

References

[1] Mclaren, J.W., Trocme, S. D., Relf, S., and Brubaker R. F. (1990) Rate of flow of aqueous humor determined from measurements of aqueous flare. Invest. Opthalmol. *Vis. Sci*. **31**, 339-346.
[2] Ethier, C.R., Kamm, R.D., Palaszewski, B.A., Johnson, M.C., and Richardson, T.M. (1986) Calculations of flow resistance in the juxtacanalicular meshwork. *Invest Ophthalmol Vis Sci.***27**,1741-1750.
[3] Scott, J.A. (1988) A finite element model of heat transport in the human eye. *Phys Med Biol.* **33**, 227 -241.
[4] Canning, C.R, Greaney, M.J., Dewynne, J.N., and Fitt, A.D. (2002) Fluid flow in the anterior chamber of a human eye. *IMA J Math Appl Med Biol.***19**,31-60.
[5] Kumar, S., Acharya, S., Beuerman, R., and Palkama, A. (2006) Numerical solution of ocular fluid dynamics in a rabbit eye: parametric effects. *Ann Biomed Eng.* **34**,530-544.
[6] Abdulrazik, M. (2007) Model of pulsatile-flow of aqueous humor through the iris-lens canal. *Journal of glaucoma.***16**,129-136.
[7] Ooi, E., and Ng, E. (2008) Simulation of aqueous humor hydrodynamics in human eye heat transfer. *Computers in Biology and Medicine.* **38**, 252-262.
[8] Chai, D., Chaudhart, G., Mikula, E., Sun, H., and Juhasz, T. (2008) 3D finite element model of aqueous outflow to predict the effect of femtosecond laser created partial thickness drainage channels. *Lasers in surgery and medicine*, **40**,188-195.
[9] Villamarina, A., Roya, S., Hasballa, R., Vardoulis, O., Reymond, P., and Stergiopulos, N. (2012) 3D simulation of the aqueous flow in the human eye. *Med Eng Phys*, **34**,1462-1470.
[10] Mori, A., Oguchi, Y., Okusawa, Y., Ono, M., Fujishima, H., and Tsubota, K. (1997) Use of high-speed, high-resolution thermography to evaluate the tear film layer. *Am. J. Ophthalmol.***124**,729-735.
[11] Okuno, T. (1991) Thermal effect of infra-red radiation on the eye: A study based on a model. *Ann. Ocuup. Hyg.* **35**, 1-12.
[12] Yang, J., Bo, X., Qian, X., Wang, Q., Wang, J., Lu, J., Quan, H., and Liu, Z. (2010) Measuring iris elastic modulus based on holistic iris deformation experiment. *Journal of Medical Biomechanics*, **25**,182-185.
[13] Wu, L., Xie, Y., Fan, Y., Huang, X., and Luo, X. (2008) A new method used to measure corneal elastic modulus. *Journal of Sichuan University (Engineering Science edition)*, **40**, 80-85.
[14] Karickhoff, J.R. (1992) Pigmentary dispersion syndrome and pigmentary glaucoma: a new mechanism concept, a new treatment, and a new technique. *Ophthalmic Surg*. **23**, 269-277.

A dual wavelet shrinkage procedure for suppressing numerical oscillation for nonlinear hyperbolic equations

[*]Y. ZHAO[1], G. Y. ZHANG[2] , P. Y. YU[1] , S. J. SU[1], [†]T.L. WANG[1] , Z. Zong[2]

1.Transportation Equipment and Ocean Engineering College, Dalian Maritime University, Dalian 116026, P.R. China

2. School of Naval Architecture, Dalian University of Technology, Dalian 116024, P.R. China

[*]Presenting author: fluid@126.com

[†]Corresponding author: wangtianlin@dlmu.edu.cn

Abstract

In the numerical solution for nonlinear hyperbolic equations, numerical oscillation often appears and contaminates the real solution, and sometimes can even make the computation divergent. Using a signal processing approach, a dual wavelet shrinkage procedure is proposed, which allows one to extract the real solution hidden in the numerical solution with oscillation. The dual wavelet shrinkage procedure is introduced after applying the local differential quadrature (LDQ) method, which is a straightforward technique to calculate the spatial derivatives. Results free from numerical oscillation can be obtained, which can not only capture the position of shock and rarefaction waves, but also keep the sharp gradient structure within the shock wave. Three model problems, a one-dimensional dam break flow governed by shallow water equations, and the propagation of a one-dimensional and a two-dimensional shock wave controlled by the Euler equations, are used to confirm the validity of the proposed procedure.

Key Words: Wavelet shrinkage; numerical oscillation; shock wave; nonlinear hyperbolic equations; LDQ

1. Introduction

Due to the nonlinearity, most problems governed by hyperbolic PDEs in fluid dynamics have to be solved numerically. The main difficulty is that the solution of hyperbolic PDE are bound to develop discontinuities in finite time. A conventional numerical scheme, such as finite difference or finite volume, is directly applied to solve shock wave problem. Two serious problems may appear simultaneously or separately: (1) smearing out shock wave gradually; (2) polluting the shock wave by numerical oscillation. In the recent several decades, people have developed many numerical schemes to maintain the sharp gradient structure of shock wave, meanwhile avoiding the numerical oscillation. Godunov [1959] was credited with introducing the first Riemann solver for the Euler equations, by extending the previous Courant-Isaacson-Reeves (CIR) method to non-linear systems of hyperbolic conservation law. To increase computation's effectiveness, Roe[1981] proposed a second order ROE solver through taking average of square root of density on double sides of a cell. Subsequent works were HLL and HLLC schemes [Toro (1999)]. Efforts in this field led to the proposal of total variance diminishing (TVD) [Loubere *et al* (2014), Toro (1999)] and weighted essentially non-oscillating (WENO) schemes [Liu *et al*(1994), Abgrall (1994)]. Besides these sophisticated schemes, Shyy proposed a non-linear filtering algorithm to eliminate numerical oscillation from second-order central or upwind differencing in calculation of shock wave [Shyy *et al* (1992)]. The filter has proved to be very effective in

suppressing oscillation of short wavelength. However, the effect in removing the oscillation of long wavelength is not so promising. Kang introduced a multi-resolution analysis (MRA) for increasing computation's efficiency with preserving the high order numerical accuracy of a conventional solver [Kang *et al* (2014)]. Inspired by their work, we discovered that wavelet's application in suppressing numerical oscillation around the shock wave. Wavelet analysis is characterized by decomposing the signal to be analyzed into multi-scale coefficients, high frequency component is described by coefficients on small scale and low frequency component is described by coefficients on large scale [Gerolymos *et al* (2009), Mallat (1999)]. In this way, shock wave may be maintained and the numerical oscillation around the shock wave may be removed after some special treatment for wavelet coefficients.

In this paper, we propose a dual wavelet shrinkage procedure to suppress numerical oscillation from a straightforward numerical scheme, named localized differential quadrature (LDQ) method, to calculate shock wave problem. LDQ, proposed by Zong [2002], is a high order accurate numerical method The LDQ method was used to solve Riemann problem [Mahdavi *et al* (2012)]. However, high oscillation emerged. A dual wavelet transformation is then applied to process the highly oscillatory results. Three shock wave propagation problems governed by shallow water equations and Euler equations are presented for demonstration. The results are compared to their analytical solutions and very well agreement is achieved.

2. A brief introduction of LDQ

Hereinafter, only the main formulas of LDQ are introduced. Details of LDQ and their applications can be referred to Zong [2009]. The first step to LDQ method is to locate the neighborhood of a grid point of interest x_i. We use $r_{il}=|x_i-x_l|,i,l=1,\ldots,N$ to denote the distance between any two points in the solution domain. We find the permutation $s(1),s(2),\ldots,s(N)$ to satisfy $r_{is(1)} \le r_{is(2)} \ldots \le r_{is(N)}$.

It is clear that the points falling in the neighborhood of *i*-th point xi are the first m points. Denoting Si=(s(1),s(2),…,s(m)),i=1,…,N, and then Si defines the neighborhood of the grid point of interest. We may get the first and second spatial derivatives at point xi of function f(x) by its neighborhood Si as following:

$$\frac{df(x_i,t)}{dx} \approx \sum_{j\in S_i}^{N} a_{ij} f(x_j,t) \tag{1}$$

$$\frac{d^2 f(x_i,t)}{dx^2} \approx \sum_{j\in S_i}^{N} b_{ij} f(x_j,t) \tag{2}$$

where

$$a_{ij}(x) = \frac{1}{x_j - x_i} \sum_{\substack{k\in S_i \\ k\neq i,j}}^{N} \frac{x_i - x_k}{x_j - x_k},\ j \in S_i,\ j \neq i \tag{3}$$

$$a_{ii} = -\sum_{\substack{j\in S_i \\ j\neq i}}^{N} a_{ij} \tag{4}$$

$$b_{ij}(x) = 2[a_{ij}a_{ii} - \frac{a_{ij}}{x_i - x_j}], i, j = 1,...,m, i \neq j \tag{5}$$

$$b_{ii}(x) = -\sum_{j \neq i}^{N} b_{ij}(x), i = 1,...,m \tag{6}$$

Multi-dimensional formulas share the same form as in one dimensional because they are independent in each direction.

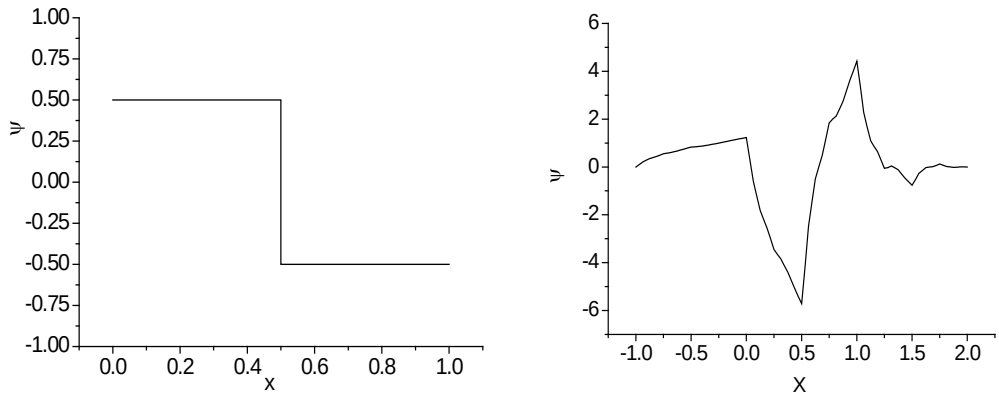
3. A dual wavelet shrinkage procedure

Wavelet analysis, introduced by Grossmann & Morlet in the early 1980s, has significantly impacted the signal and image processing [Mallat (1999)]and numerical solving in nonlinear partial differential equations (see Refs. Beylkin [1992], Zong *et al* [2010]，Vasilyev and Paolucci [1996], Vasilyev *et al* [1995])as well as its application on turbulence(see Refs. Farge *et al* [1999], Farge and Kaiser [2001], Goldstein and Vasilyev [2004], Schneider *et al* [1997], Zong et al [2010]). Wavelet provides compactly supported, orthogonal or bi-orthogonal basis functions with adjustable smoothness. Due to the compactly supported basis, wavelet coefficients contain local structure information, hence wavelet analysis is a proper mathematical tool to process signal with local structure. Further more, wavelet analysis enables us to obtain multi-scale information of the analyzed signal by introducing scale dilating.

Vanishing moment is an important property of wavelet analysis, which is directly related to wavelet basis functions' smoothness. Wavelet function is called M order vanishing moment if the following relation is satisfied:

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0, k = 0,1,2,\cdots,M-1 \tag{7}$$

Daubechies wavelet, one of the most important orthogonal wavelet categories, is widely used in many applications [Daubechies (1988)].The function's smoothness of Daubechies wavelet is adjustable by the vanishing moment order M. Let sup φ=[0,2M-1] be scaling function and sup ψ=[-M+1,M] be wavelet function's supported range. For example, the first or second order of Daubechies wavelet, denoted as DB1 or DB2, are shown in Fig.1.



(a) (b)

Fig.1 Daubechies wavelet functions of DB1(a) and DB2(b).

Wavelet bases provide us a multi-scale resolution to express function. We assume function f(x) can be totally approached on the finest scaling index J , i.e. $c_{J,k} = f(x_{J,k})$ and then it can be decomposed in terms of the sum of a series of wavelet bases functions

$$f^J(x) = \sum_k C_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j=j_0}^{J-1}\sum_k d_{j,k}\psi_{j,k}(x) \tag{8}$$

The coefficients are obtained by wavelet decomposition formula:

$$c_{j-1,k} = \sum_{l=0}^{L-1} h_l c_{j,l+2k} \tag{9}$$

$$d_{j-1,k} = \sum_{l=0}^{L-1} g_l c_{j,l+2k} \tag{10}$$

For function with most part is well smooth, most wavelet coefficients will be small. Consequently, we can retain a good approximation even discarding a large number of wavelets with small coefficients. More precisely, if we rewrite the approximation as a sum of two terms

$$f^J(x) = f^J_{|d|\geq\varepsilon}(x) + f^J_{|d|\leq\varepsilon}(x) \tag{11}$$

Where ε is a prescribed particular threshold and

$$f^J_{|d|\geq\varepsilon}(x) = \sum_k c_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j=j_0}^{J-1}\sum_{|d_{j,k}|\geq\varepsilon} d_{j,k}\psi_{j,k}(x) \tag{12}$$

$$f^J_{|d|<\varepsilon}(x) = \sum_{j=j_0}^{J-1}\sum_{|d_{j,k}|<\varepsilon} d_{j,k}\psi_{j,k}(x) \tag{13}$$

From Vasilyev and Paolucci [1996], the approximation error caused by the significant wavelets, whose coefficient amplitude is above threshold ε , is bounded by following restriction:

$$\left| f^J(x) - f^J_{|d|\geq\varepsilon}(x) \right| \leq C_1\varepsilon \tag{14}$$

and the number of significant wavelet coefficient *K* is bounded by ε and wavelet's vanishing moment *N* as

$$\mathrm{K} \leq C_2\varepsilon^{-1/2N} \tag{15}$$

C_1, C_2 in Eqs.(14) and (15) depend on wavelet vanishing moment and function *f(x)* .Threshold has two effects: making approximation adaptively and controlling the approximation error globally. The similar situation can be simple extended to multi-dimensional space by tensor product.

For the signal to be analyzed including unknown noise and unknown smoothness structures,

how to remove the noise and keep the unknown structure is a complicated problem. Donoho and Johnstone proposed a wavelet shrinkage procedure to extract the structure from noisy sampled data [Donoho and Johnstone (1995)]. We view numerical oscillation as the noise, discontinuity such as shock and rarefaction wave as the signal's structure. Based on this understanding, the wavelet shrinkage procedure is applied to suppress the highly numerical oscillation obtained from the LDQ Method. Indeed, wavelet shrinkage procedure can extract the unknown smoothness structures contaminated by heavy noise. Donoho and Johnstone also addressed that the reconstruction is essentially as smooth as the mother wavelet [Donoho and Johnstone (1995)].This indicates that the reconstructed signal's smoothness is closely related to the adopted wavelet basis functions' smoothness. The reconstructed signal is locally similar as the adopted wavelet basis function, so the optimal wavelet basis should be in the same order smoothness as the real physical signal. Based on the consideration, we propose a dual wavelet shrinkage procedure using DB1 and DB2 for suppressing numerical oscillation obtained from LDQ method. DB1 and DB2 are respectively suitable for shock and rarefaction wave's reconstruction, because DB1 is a sharp jump function and DB2 is a function with one-order smoothness. The dual wavelet shrinkage procedure is proposed as followings.

1) Decomposing the numerical result with highly oscillation obtained from LDQ method via the discrete wavelet transform using DB1 by Mallat algorithm [Goldstein and Vasilyev (2004)], then the wavelet coefficients $d_{j,k}$ is obtained, where $j_0 \leqslant j \leqslant J, 1 \leqslant k \leqslant 2^j$, j is scale index, j_0, J represent the largest and smallest scale, respectively.
2) Setting $t_j = \sigma_j \sqrt{2\ln(N_j)/N_j}$ be the threshold value at scale j, where σ_j is the standard deviation of coefficients $d_{j,k}$, and N_j is the number of wavelet coefficients $d_{j,k}$, i.e. $N_j = 2^j$. A threshold is assigned to each dyadic resolution level for threshold estimates, which is adaptive.
3) Thresholding wavelet coefficients $d_{j,k}$ to get revised coefficients $\bar{d}_{j,k}$ by following

$$\bar{d}_{j,k} = \begin{cases} \mathrm{sgn}(d_{j,k})(|d_{j,k}| - t_j), & if\ |d_{j,k}| \geq t_j \\ 0, & else \end{cases} \tag{16}$$

4) Reconstructing de-noisy data via the wavelet reconstruction transform using the revised wavelet coefficients $\bar{d}_{j,k}$.
5) Repeating the above procedure using DB1. At next time step, the dual wavelet shrinkage procedure is complemented again.

The additional of this computational effort of the overall procedure is only of order $N \bullet log(N)$ as a function of sample size N by Mallat algorithm, which is a fast transform similar as fast Fourier transform, brings little extra computation.

4. Numerical tests

4.1 One-dimensional dam break flow

One-dimensional shallow water equations (SWEs) is a typical Riemann problem [Stoker (1986)],

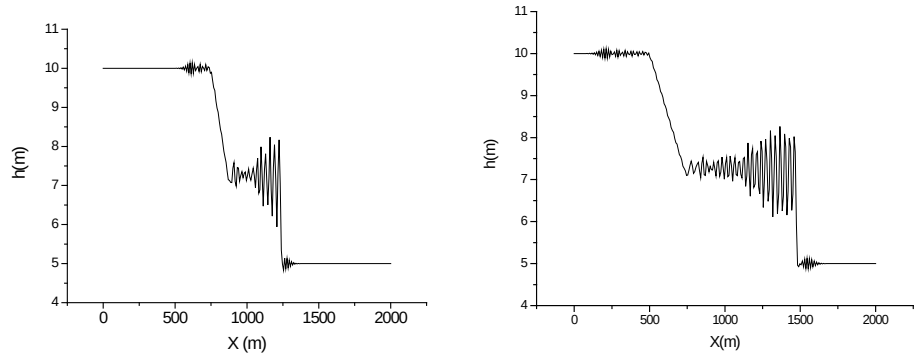
$$U_t + F(U)_x = 0 \tag{17}$$

$$U = \begin{bmatrix} h \\ hu \end{bmatrix}, F(U) = \begin{bmatrix} hu \\ hu^2 + gh^2/2 \end{bmatrix} \tag{18}$$

SWEs describe the flow at time $t \geq 0$ at point x, where $h(x,t)$ is the water depth, $u(x,t)$ is the average horizontal velocity and g the gravitational acceleration. A wide variety of physical phenomena are governed by the SWEs, such as tidal flows in coastal water regions, bore wave propagation, flood waves in rivers, surges, dam-break modeling and so on [Delis and Katsaounis (2003)]. Here we use SWEs to model dam-break problem in a channel of length L=2000 m and the initial condition is

$$\begin{aligned} u(x,0) &= 0 \\ h(x,0) &= \begin{cases} 10 & 0 \leq x \leq 1000 \\ 5 & 2000 \geq x > 1000 \end{cases} \end{aligned} \tag{19}$$

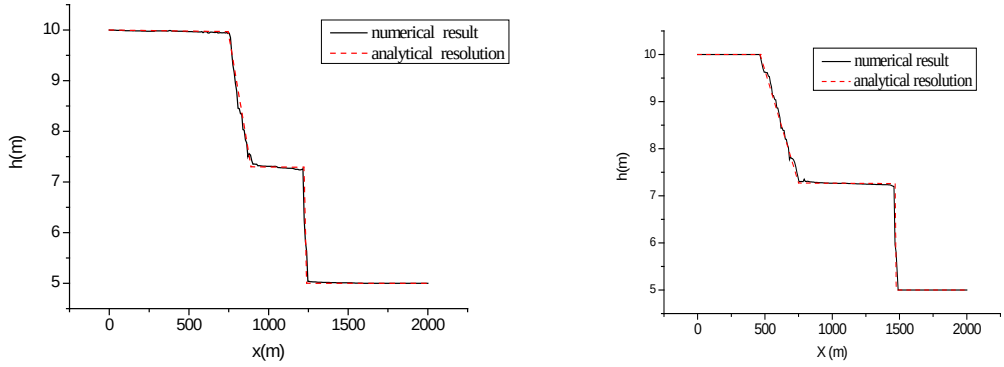
The dam collapses at t =0 and the resulting bow consists of a shock wave traveling downstream and a rarefaction wave traveling upstream. LDQ is applied to the calculation of shock wave tube. However, the viscosity is also included to avoid numerical oscillation in the article [Zhao *et al* (2011)]. In this research, the LDQ method is employed for spatial derivative calculation, where m=5. Uniformly spaced nodes N=256 are set in the channel length. Four-order Runge-Kutta method with time step Δt=0.05s is used for time integration. In order to check the dual wavelet shrinkage procedure's effect, results without wavelet shrinkage procedure at time t=25s and t=50s are shown in Fig.2. Highly numerical oscillatory results are obtained only by LDQ method, especially after the shock wave front. Further more, the oscillation develops along with the time. Real physical solutions are hidden in the highly oscillatory numerical results. The task of the wavelet shrinkage procedure is to reconstruct them by suppressing the numerical oscillation adaptively.



(a) (b)

Fig.2 Numerical result of water depth only by LDQ at *t=25s (a)* and *t=50s (b)*, respectively

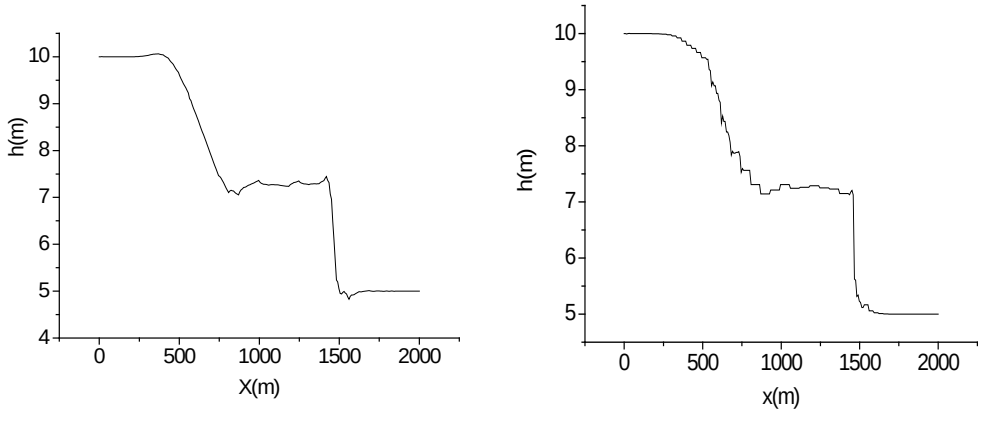
The dual wavelet shrinkage procedure is subsequently implemented after LDQ method to remove the numerical oscillation. Fig.3 shows the numerical result of water depth along the channel at time t=25s and t=50s, which is also compared to the analytical result referred to Stoker [1986]. The highly numerical oscillation is removed out by this dual procedure. Further more, it remains the sharp gradient structure near the shock wave front simultaneously. However, small disturbance is introduced in the domain of rarefaction wave, which could be ignored.



(a) (b)

Fig.3 Comparison of numerical result for water depth using the dual wavelet shrinkage procedure and analytical result at time *t=25*s (a) and *t=50*s (b), respectively

In order to examine the dual wavelet shrinkage's superiority over single wavelet shrinkage, we use only DB1 and DB2 in the procedure, respectively. Fig.4 is the result of water depth only using DB2 at time t=50s. As shown in Fig.4(a), DB2 can extract the rarefaction wave's structure well. However, DB2 introduces obvious disturbance with similar shape as DB2 function in the solution. Besides, DB2 makes the shock front excessively smooth, which does not comply with the physical truth. The result obtained from the procedure only using DB1 is shown in Fig.4(b). DB1 can reconstruct shock wave front with sharp gradient structure. But it also introduces many stepped shape error in rarefaction wave and before shock wave front.



(a) (b)

Fig.4 Water depth along the channel only with DB2 (a) and DB1(b) at *t=50s*, respectively

The dual procedure combines the respective advantages of DB1 and DB2. DB1 is a discontinuity function, which is the best candidate to capture the shock wave front. DB2, the second order vanishing moment function, is the optimal function to reconstruct the rarefaction wave. The dual wavelet procedure using DB1 and DB2 is a proper combination for processing highly numerical oscillatory results obtained from LDQ method in Riemann problem with shock wave.

4.2 One-dimensional shock tube problem

The Euler equations for one-dimensional unsteady ideal gas flow without heat conduction are

given in conservation form.

$$\frac{\partial \rho}{\partial t}+\frac{\partial(\rho u)}{\partial x}=0 \tag{20}$$

$$\frac{\partial(\rho u)}{\partial t}+\frac{\partial}{\partial x}(\rho u^2+p)=0 \tag{21}$$

$$\frac{\partial E}{\partial t}+\frac{\partial}{\partial x}(u(E+p))=0 \tag{22}$$

where ρ , u , E , p is gas' density, velocity, energy per unit volume and pressure, respectively. We need one more equation, i.e the state equation, to close the system.

$$p=(\gamma-1)(E-\frac{1}{2}\rho u^2) \tag{23}$$

Let ratio of specific heats γ=1.4 and spatial range $-15m \le x \le 15m$. Initial conditions are specified as followings:

$$(\rho,\rho u,E)^T=\begin{cases}(2,0,3), & -15 \le x \le 0,\\ (1,0,1.5), & 15 \ge x > 0\end{cases} \tag{24}$$

This is a shock tube problem. Setting uniformly spaced nodes with number N=512 in the tube length and m=5 in LDQ method for spatial derivatives' calculation. Fourth-order Runge-Kutta method is used for time integration with time step Δt=0.005s.

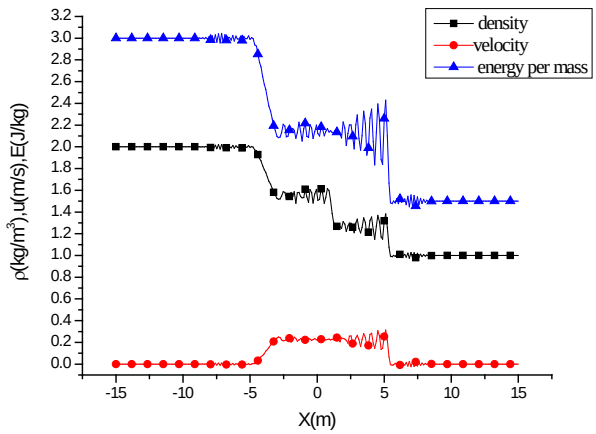


Fig. 5 Profiles of numerical density, velocity and energy of unit mass only using LDQ at time *t=5s*

In order to reveal the effect of wavelet shrinkage procedure, the numerical results obtained from pure LDQ method is shown in Fig. 5. Intense oscillation appears as in Fig.2.

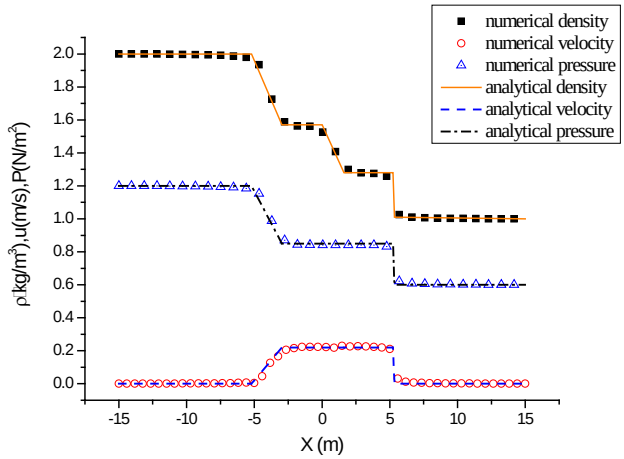


Fig.6 Comparison of numerical density, velocity and pressure de-noised by the dual wavelet shrinkage procedure and the corresponding analytical solutions at time *t=5s*

After applying the dual wavelet shrinkage procedure to remove the numerical oscillation, the gas's density, velocity and pressure at time t=5s are shown in Fig.6, also including the analytical solutions. As shown in Fig.6, numerical results are very close to their analytical counterparts as referred by Toro [1999].

4.3 Two-dimensional shock wave propagation

In order to confirm the procedure's effect in two-dimensional shock wave's computation, we consider two-dimensional Euler equations as following.

$$\begin{cases} \dfrac{\partial \rho}{\partial t}+\dfrac{\partial \rho u}{\partial x}+\dfrac{\partial \rho v}{\partial y}=0 \\ \dfrac{\partial \rho u}{\partial t}+\dfrac{\partial \rho uu}{\partial x}+\dfrac{\partial \rho uv}{\partial y}+\dfrac{\partial p}{\partial x}=0 \\ \dfrac{\partial \rho v}{\partial t}+\dfrac{\partial \rho vu}{\partial x}+\dfrac{\partial \rho vv}{\partial y}+\dfrac{\partial p}{\partial y}=0 \\ \dfrac{\partial E}{\partial t}+\dfrac{\partial (E+p)u}{\partial x}+\dfrac{\partial (E+p)v}{\partial y}=0 \\ p=(\gamma-1)(E-\dfrac{1}{2}\rho u^2-\dfrac{1}{2}\rho v^2) \end{cases} \tag{25}$$

The initial conditions are specified as below:

$$(\rho,\rho u,\rho v,E)^T=\begin{cases}(2,0,0,3), \sqrt{x^2+y^2}\le 10 \\ (1,0,0,1.5), 10<\sqrt{x^2+y^2}\le 20\end{cases} \tag{26}$$

Similar calculation parameters are set as in one-dimensional case, i.e. uniformly 512 ×512 nodes in x and y direction respectively; m=5 in LDQ method; and fourth-order Runge-Kutta method for time with time step Δt=0.005s.

A pressure contour at time t=3.75s is shown in Fig.7. Shock wave spreads outward with uniform speed and numerical results keep sharp interface. Pressure keeps the same at the two

sides of the contact discontinuity and varies slowly in rarefaction wave region whose position can be seen obviously in Fig.8. It can be clearly seen that the wave spreads uniformly and symmetrically. However, slightly distorted deformation is introduced in radial direction by the effects of the grid which cannot represent a circle by square grid.

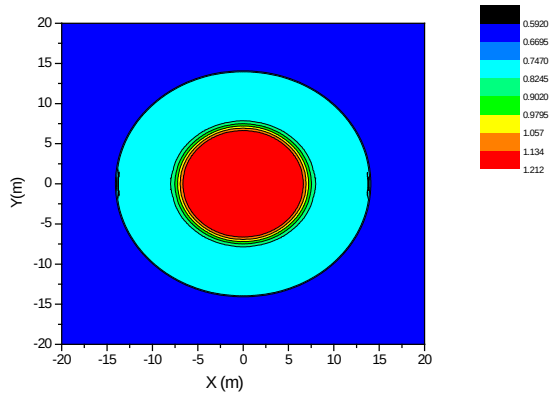


Fig.7 Pressure contour at *t=3.75*s

In order to get a clear version of density and pressure, we display their profiles at position of y=0 at time t=3.75s in Fig.8. Symmetry is maintained in the calculation and results with non-oscillation are obtained.

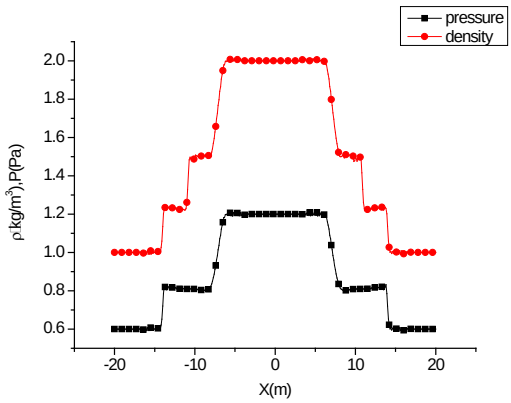


Fig.8 Profiles of numerical density, velocity components and pressure at position *y=0* at *t=3.75*s

This numerical model indicates that the dual wavelet shrinkage procedure can be extended to 2D Riemann problem with shock wave and the effect on numerical suppressing is as promising as in 1D case.

5. Conclusion

A dual wavelet shrinkage procedure is proposed to suppress the numerical oscillation for nonlinear hyperbolic equations, known as Riemann problem with shock wave in ideal fluid dynamics in this paper. The dual procedure combines the advantage of DB1 and DB2. DB1 is totally discontinuous, which is natural to capture the shock wave front. DB2, the second order vanishing moment function, is the optimal function to reconstruct the rarefaction wave. Based on these coincides, adaptive threshold value is applied to remove the numerical oscillation obtained from LDQ method.

Three numerical tests, *i.e.* one-dimensional water dam breaking problem and one/two-dimensional air shock wave propagating problems, are used to verify the procedure's performance. High quality results are obtained both in capturing discontinuity and suppressing the numerical oscillation. It's demonstrated that the dual wavelet procedure is a proper combination for processing highly numerical oscillation in Riemann problem with shock wave. Compared with the well known Riemann solvers and related complicated numerical schemes, LDQ method is qualified to solve shock wave problem with the aid of the dual wavelet shrinkage procedure.

The method proposed in this article could be applied into ship hydrodynamics numerical methods, such as fluid structure interaction (FSI). Due to most FSI analysis employ the iterative scheme that solvers the solid and fluid problems alternately at each time step, which could accumulate much numerical error and eventually could lead to convergence difficulties.

Acknowledgments

We thank these following funds i.e. National natural science foundation of China (51679021, 51609031); Liaoning province science and technology program (201602067); China postdoctoral science foundation (2016M601294); National key research and development project-the research and development of the key technology and equipment of the dish shaped underwater glider (2016YFC0301500); open research fund of state key laboratory of structural analysis for Industrial Equipment, Dalian University of Technology(15214) and open research fund of state key laboratory of ocean engineering, Shanghai Jiao Tong University (1402)
and fundamental research funds for the central universities (3132017016, 3132016339, 3132016346).

References

1. Abgrall, R. (1994). On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation, *Journal of Computational Physics*, **114**, 45-58.
2. Beylkin, G., Coifman, R., Daubechies, I. (1992) Wavelets and Their Application, Jones and Bartlett, Cambridge, MA.
3. Beylkin, G., Keiser, J.M. (1997). On the adaptive numerical solution of nonlinear partial differential equations in wavelet bases, *Journal of Computational Physics*, **132**, 233-259.
4. Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets, *Communication of Pure and Applied Mathematics*, **41**, 909-996.
5. Donoho, D.L., Johnstone, I.M. (1995). Adapting to unknown smoothness via wavelet shrinkage, *Journal of the American Statistical Association*, **90**, 1200-1224.
6. Delis, A.I. Katsaounis, T. (2003). Relaxation schemes for the shallow water equations, *International of Journal of Numerical Methods in Fluids*, **41**,695-719.
7. Farge, M., Kaiser, S. (2001). Coherent vortex simulation (CVS), a semi-deterministic turbulence model using wavelets, *Flow, Turbulence and Combustion* , **66**, 393-426.
8. Farge, M., Schneider, K., Kevlahan, N. K. R. (1999).Non-gaussianity and coherent vortex simulation for two-dimensional turbulence using an adaptive orthogonal wavelet basis, *Physics of Fluids*, **11**, 2187–2201.
9. Godunov, S.K. (1959). A finite difference method for computation of discontinuous solutions of the equations of fluid dynamics, *Matematicheski Sbornik* , **47**, 271-306.
10. Gerolymos, G.A., Senechal, D., Vallet, I. (2009). Very-high-order WENO schemes, *Journal of Computational Physics*, **228**, 8481-8524.
11. Goldstein, D. E., Vasilyev, O.V. (2004). Stochastic coherent adaptive large eddy simulation method, *Physics of Fluids*, **16**, 2497–2513.
12. Kang, H., Park, K., Kim, K., Lee, D. (2014). Multi-resolution analysis for high accuracy and efficiency of Euler computation, *International Journal of Numerical Methods in Fluids*, **74** 661-683.

13. Loubere, R., Dumbser, M., Diot, S. (2014). A new family of high order unstructured MOOD and ADER finite volume schemes for multidimensional systems of hyperbolic conservation laws, *Communications in computational physics*, **16**(3), 718-763.
14. Liu, X. D., Osher, S., Chan, T. (1994). Weighted essentially non-oscillatory schemes, *Journal of Computational Physics*, **115**,200-212.
15. Mallat, S. (1999). A wavelet tour of signal processing (2nd Edition), Academic Press, New York.
16. Mahdavi, A., Hashemi, M.R., Talebbeydokhti, N. (2012). A localized differential quadrature model for moving boundary shallow water flows, *Journal of Hydraulic Research*, **50**, 612-622.
17. Roe, P. L. (1981). Approximate Riemann solvers, parameter vectors and difference schemes, *Journal of Computational Physics*, **43**, 357-372.
18. Stoker, J.J. (1986). Water Waves, Interscience Publishers, Inc. New York.
19. Schneider, K., Kevlahan, N.K.R., Farge, M. (1997). Comparison of an adaptive wavelet method and nonlinearly filtered pseudo-spectral methods for two-dimensional turbulence, *Theory of Computational Fluid Dynamics*, **9**, 191–206.
20. Shyy, W., Chen, M.H., Mittal, R., Udaykumar H.S. (1992). On the suppression of numerical oscillations using a non-linear filter. *Journal of Computational Physics*, **102**, 49–62.
21. Toro, E. F. (1999). Riemann Solver and Numerical Methods for Fluid Dynamics (2nd edition), Springer, Berlin.
22. Vasilyev, O.V., Paolucci, S. (1996) A dynamically adaptive multilevel wavelet collocation method for solving partial differential equations in a finite domain, *Journal of Computational Physics*, **125**, 498-512.
23. Vasilyev, O.V., Paolucci, S., Sen, M. (1995). A multilevel wavelet collocation method for solving partial differential equations in a finite domain, *Journal of Computational Physics*, **120**, 33-47.
24. Zong, Z., Lam, K.Y. (2002). A localized differential quadrature method and its application to the 2D wave equation, *Computational Mechanics*, **29**, 382-391.
25. Zong, Z., Zhang, Y. Y. (2009). Advanced Differential Quadrature Methods, Chapman & Hall/CRC.
26. Zhao, Y., Zong, Z., Li, Z.R. (2011). Shock calculation based on second viscosity using local differential quadrature Method. *Applied Mathematics and Mechanics(English edition)*, **32**,349-360.
27. Zhao, Y., Zong, Z., Zou, W.N. (2011). Numerical simulation of vortex evolution based on adaptive wavelet method, *Applied Mathematics and Mechanics(English edition)*, **32**, 33-44.
28. Zong, Z., Zhao, Y., Zou, W.N. (2010). Numerical solution for differential evolutional equation using adaptive interpolation wavelet method, *Chinese Journal of Computational Physics*,**27**, 65-69.

Hybrid approximation spaces for solving the compressible Navier-Stokes equations with high Reynolds number

*Fan Zhang, Jian Cheng, †Tiegang Liu
School of Mathematics and Systems Science,
Beihang University, Beijing 100191, PR China.
*Presenting author: nmsdzf90@163.com
†Corresponding author: liutg@buaa.edu.cn

Abstract

Boundary layer will be produced if the Reynolds number of the compressible Navier-Stokes equations is sufficiently high. As there has a steep gradient of flow variables inside the boundary layer, it is very difficult to compute the Navier-Stokes equations stably and accurately. There are mainly two approaches to solve this problem. One is to use a very fine mesh whose grid size is approximate to the thickness of boundary layer. However, for the high dimensional problem, this may lead to an unbearable computing cost. The other is to use a high order numerical method, such as the Runge-Kutta discontinuous Galerkin (RKDG) method. However, as the thin thickness and large slope of the boundary layer, the traditional high order RKDG methods based on piecewise polynomial basis functions may not provide the best approximation to the solution and normal derivative inside the boundary layer unless a very fine spatial grid is divided.
In this paper, we consider the one-dimensional compressible Navier-Stokes equations which has an analytic solution under some assumptions, and according to its analytic solution, we provide an exponential approximation space for solving the computing domain inside the boundary layer. And for the computing domain outside the boundary layer, we still use the traditional polynomial space to approximate it. From numerical experiments we can see that compared to the traditional RKDG method based on polynomial approximation space, the RKDG method based on the hybrid approximation space provided in this paper can yield better results of the flow field values and gradient values over the same computing grid.

Keywords: Compressible Navier-Stokes equations, boundary layer theory, discontinuous Galerkin method, approximation space, exponential basis functions

Introduction

Flows of fluids with low viscosity values and thus very high Reynolds numbers occur in many technical applications. When a viscous fluid flows along a fixed impermeable wall, or past the rigid surface of an immersed body, an essential condition is that the velocity at any point on the wall or other fixed surface is zero. The extend to which this condition modifies the general character of the flow depends upon the value of the viscosity. If the body is of streamlined shape and if the viscosity is small without being negligible, which leads to a large value of Reynolds number, the modifying effect appears to be confined within narrow regions adjacent to the solid surfaces which are called as boundary layers [1]. Within such layers the fluid velocity changes rapidly from zero to its main-stream value, and this may imply a steep gradient of shearing stress. As a consequence, not all the viscous terms in the equation of motion will be negligible, even though the viscosity, which they contain as a factor, is itself very small.
The concept and theory of boundary layers have been developed by Ludwig Prandtl and presented in a historic paper in 1905 [2]. After then, physicists and engineers have written hundreds of articles and books about various aspects of boundary layer theory. And the most classic and best-known is Hermann Schlichting's boundary layer theory [3, 4]. The main characteristic of

the boundary layer theory is that the solution for the in-viscid outer flow and the solution for the boundary layer are being determined separately and matched properly, that is the flow region can be divided into two parts:

- Away from the surface of the object, viscous effects can be considered negligible, and potential flow can be assumed.
- In a thin region near the surface of the object, viscous effects cannot be neglected, and are as important as inertia.

Prandtl's boundary layer theory had a tremendous effect on the development of fluid mechanics and had attracted the attention of many researchers in finding high order numerical method to solve the the complete equations of motion of a viscous fluid - Navier-Stokes equations.
As there exist a large velocity gradient normal to the boundary in a very thin layer, it is considerably difficult to resolve it. One approach to solve this problem is mesh fitting. Because the accuracy of the numerical solutions are determined to some extent by the quality of the computing mesh, in order to obtain accurate solutions it would need to take a very fine mesh. Many researchers have worked in the field of performance and generation of boundary layer elements for CFD simulations. Karman presented a linear-elastic smoothing scheme to push bulk mesh and generate a new, unstructured viscous layer of elements [5–9]. The main idea of all these methods is to take a directional grid refinement procedure for accurate solution for the boundary layer and wake flow regions [10]. However, for high dimensional flow problems, this may lead to an unbearable computational cost and a low computing efficiency. Taking the two dimensional flat plate problem as an example, approximately 75% of the grid points are inside the boundary layer, which decrease the efficiency of numerical computing severely.
Another approach to obtain an accurate solution inside the boundary layer is to take a high order method, such as the finite difference method, finite volume method and finite element method, see review paper [11]. As the higher order the numerical method is, the larger number of solution unknowns or degrees of freedom are needed, this approach also brings a tremendous amount of computing. For the simulation of high dimensional turbulent flow problems, systems of several million degrees of freedom are common. Unfortunately, grid convergence, and hence reliable accuracy, is not always attained. What's more, high order methods applied to non-linear problems tend to become unstable when the approximating apace is inadequate to resolve the main features of the true solution. Unresolved boundary layers produce Gibbs oscillations which, in the presence of non-linearly, often lead to solution blow up [12].
Since the Navier-Stokes equations consist of the Euler equations plus shear-stress and heat flux terms, one of the major differences that occurs when solving the Navier-Stokes equations, as compared to the Euler equations, is the need to use fine meshes or high order scheme in order to properly resolve viscous layers. As the property of easily handling complex geometries and boundary conditions and achieving high order accuracy, the finite element method, especially the discontinuous Galerkin (DG) methods have received much attention during the last decade due to their ability to produce stable and high order accurate discretizations of conservation laws on fully unstructured meshes [13]. The DG method is a finite element method using a completely discontinuous piecewise polynomial space for the numerical solution and the test functions [14, 15]. The major development of DG methods was carried out by Cockburn and Shu in a series of papers [16–20]. They constructed high order Runge-Kutta discontinuous Galerkin (RKDG) method for the scaled conservation laws. And then this method was extended to one-dimensional and multi-dimensional systems.
When DG methods are used to solve partial differential equations (PDEs), the piecewise polynomial space is the commonly chosen finite element approximation space. However, for some

PDEs and initial/boundary conditions, piecewise polynomials may not provide the best approximation to the solution if the mesh is coarse, such as the Navier-Stokes equations with high Reynolds number [21]. And for the DG method based on non-classical piecewise polynomial basis functions, it has been studied in several literatures. In [22], Li and Shu proposed the use of locally divergence-free polynomial space in the DG method to solve the Maxwell equations and they achieved better results compared to the DG method based on the classical piecewise polynomial space P^k. Subsequently, the locally divergence-free polynomial space for approximating MHD equations, Hamilton-Jacobi equations and Laplace equation were discussed in [23–25]. However, these locally divergence-free polynomial space is still based on polynomials. Then in [26], Yuan and Shu developed discontinuous Galerkin methods based on non-polynomial approximation space, such as exponential approximation space and trigonometric approximation space, etc., for numerically solving time-dependent hyperbolic and parabolic and steady state hyperbolic and elliptic partial differential equations. However, this paper only discussed the scalar equation and what's more, the non-polynomial approximation space proposed in this paper was constructed without taking the analytic solution of the equation into account.

For the steady one-dimensional compressible Navier-Stokes equations, it can be solved analytically under the assumptions that the Prandtl number is taken as $\frac{3}{4}$ and the enthalpy is uniform on the boundaries [27]. As the analytic behaviour of the solution is available, the approximation of the solution can be improved by taking this information into account. Thus, in this paper, our main purpose is to propose a suitable approximation space which can approximate the Navier-Stokes equations accurately without needing to take a very fine mesh. Based on the traditional direct DG (DDG) method [28–30], we introduce a new DDG method based on a hybrid approximation space which is taken as an exponential approximation space inside the boundary layer and a polynomial approximation space outside the boundary layer in order to obtain a good approximation to the solutions over a coarse grid.

One-dimensional Compressible Navier-Stokes Equations

1. Non-dimensional compressible Navier-Stokes equations

The one-dimensional flow of viscous, heat-conducting, compressible fluid is described by the following hydrodynamic equations in conservative form

$$\frac{\partial Q}{\partial t}+\frac{\partial F}{\partial x}=\frac{\partial F_v}{\partial x}, \qquad x\in[x_L,x_R], \tag{1}$$

where $[x_L, x_R]$ is the solution domain, the conservation variable Q, the non-viscous flux F and the viscous flux F_v are defined as

$$Q=\begin{bmatrix}\rho\\ \rho u\\ E\end{bmatrix}, \quad F=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho u H\end{bmatrix}, \quad F_v=\begin{bmatrix}0\\ \tau_{xx}\\ u\tau_{xx}-q_x\end{bmatrix},$$

In order to make Eq.(1) closed, we also need an equation of state (EOS) and is assumed as

$$p=\rho RT. \tag{2}$$

Here,

$$\tau_{xx}=(2\mu+\lambda)\frac{\partial u}{\partial x}, \quad q_x=-k\frac{\partial T}{\partial x},$$

ρ is mass density, p is pressure, u is $x-$component of velocity. And in this paper, we assume the flow is from left to right, parallel to x-axis which means the velocity vector is positive. E is total energy per unit mass, H is enthalpy which is defined as

$$H = (E + p)/\rho = \frac{1}{2}u^2 + C_p T,$$

T is absolute temperature, μ is coefficient of viscosity and for simplicity we take it as a constant $\mu = \mu_r$. By using Stokes' hypothesis $\lambda = -\frac{2}{3}\mu$ the viscous stress τ_{xx} can be simplified to

$$\tau_{xx} = \frac{4}{3}\mu\frac{\partial u}{\partial x}.$$

k is the coefficient of thermal conductivity and is defined by

$$k = \frac{\mu C_p}{P_r},$$

in which P_r is the Prandtl number, $C_p = \dfrac{\gamma}{\gamma - 1}R$ is the specific heat at constant pressure, $\gamma = \frac{C_p}{C_v}$ is the ratio of specific heat, R is a gas constant which depends on the particular gas under consideration.
The boundary conditions of Eq.(1) are given as follows

$$\begin{cases} \rho(x_L) = \rho_L, \quad u(x_L) = u_L, \quad T(x_L) = T_L, \\ \rho(x_R) = \rho_R, \quad u(x_R) = u_R, \quad T(x_R) = T_R. \end{cases} \tag{3}$$

Next, we define the following dimensionless variables

$$\begin{aligned} & x = \frac{x^*}{L}, \quad u = \frac{u^*}{u_r}, \quad t = \frac{t^*}{L/u_r}, \\ & \rho = \frac{\rho^*}{\rho_r}, \quad p = \frac{p^*}{\rho_r u_r^2}, \quad T = \frac{T^*}{T_r}, \\ & R = \frac{R^*}{R_r}, \quad \mu = \frac{\mu^*}{\mu_r}, \\ & C_v = \frac{C_v^*}{U_r^2/T_r} = \frac{C_v^*}{\gamma \cdot R_r \cdot M_a^2} = \frac{1}{\gamma(\gamma-1)M_a^2}, \\ & C_p = \frac{C_p^*}{U_r^2/T_r} = \frac{C_p^*}{\gamma \cdot R_r \cdot M_a^2} = \frac{1}{(\gamma-1)M_a^2}, \\ & C_r = \sqrt{\gamma R_r T_r}, \quad u_r = M_a \cdot C_r, \end{aligned} \tag{4}$$

where the superscript * denotes the dimensional variables, subscript r denotes dimensional reference quantities and any non-marked variable denotes the non-dimensional variable. Then the non-dimensional form of (1) can be written as

$$\frac{\partial Q}{\partial t} + \frac{\partial F}{\partial x} = \frac{1}{Re}\frac{\partial F_v}{\partial x}. \tag{5}$$

Here, the forms of conservation variable Q, non-viscous flux F and viscous flux F_v are same

with those in Eq.(1). The Reynolds number which is defined as

$$R_e = \frac{\rho_r u_r L}{\mu_r},$$

significantly, corresponds to a non-dimensional reference length that directly controls the steepness of the non-dimensional gradients. Increasing R_e, as expected, thus leads to solutions that may be difficult to compute accurately.
The non-dimensional form of EOS (2) is

$$p = \rho T/(\gamma M_a^2) = \frac{\gamma - 1}{\gamma} C_p \rho T. \tag{6}$$

2. Analytic solution of the steady one-dimensional Navier-Stokes equations with $P_r = \frac{3}{4}$
For the non-dimensional Navier-Stokes equations (5), if the computing time is large enough, that is $t \to \infty$, it can be got that $\frac{\partial Q}{\partial t} \to 0$. Then we will have the steady one-dimensional compressible Navier-Stokes equations

$$\begin{cases} \dfrac{\partial \rho u}{\partial x} = 0, \\ \dfrac{\partial(\rho u^2 + p)}{\partial x} = \dfrac{1}{Re}\dfrac{\partial \tau_{xx}}{\partial x}, \\ \dfrac{\partial \rho u H}{\partial x} = \dfrac{1}{Re}\dfrac{\partial(u\tau_{xx} + \frac{\mu C_p}{P_r} T_x)}{\partial x}. \end{cases} \tag{7}$$

Integrating Eq.(7) once with respect to x in domain $[x_L, x]$, where x is any point in the whole computing domain $[x_L, x_R]$, leads to the following system

$$\begin{cases} \rho u = F_1, \\ \rho u^2 + p - \varepsilon\mu \frac{du}{dx} = F_2, \\ \rho u H - \varepsilon\mu(u\frac{du}{dx} + \frac{C_p}{\frac{4}{3}P_r}\frac{dT}{dx}) = F_3. \end{cases} \tag{8}$$

Here, $\varepsilon = \frac{4}{3}/R_e$ is a constant. And when the Reynolds number satisfies $R_e \to \infty$, we have $\varepsilon \to 0$. The constants $F_i\,(i = 1, 2, 3)$ are defined as

$$\begin{cases} F_1 = (\rho u)|_{x_L}, \\ F_2 = (\rho u^2 + p)|_{x_L}, \\ F_3 = (\rho u H)|_{x_L}. \end{cases} \tag{9}$$

According to [27], if we take the assumption that the Prandtl number is taken as $P_r = \frac{3}{4}$ and the coefficient of viscosity μ is taken as a constant, then the integrated Navier-Stokes equations (8) can be decoupled and simplified into a non-linear ordinary differential equation of velocity

$$\varepsilon\mu u\frac{du}{dx} - \frac{\gamma + 1}{2\gamma}F_1 u^2 + F_2 u - \frac{\gamma - 1}{\gamma}F_3 = \frac{\gamma - 1}{\gamma}F_1(H_R - H_L)e^{-F_1\frac{x_R - x}{\varepsilon\mu}}. \tag{10}$$

If the boundary conditions (3) satisfy $H_L = H_R$, then the equation of velocity (10) can be further reduced to

$$\varepsilon\mu u\frac{du}{dx} - \frac{\gamma + 1}{2\gamma}F_1 u^2 + F_2 u - \frac{\gamma - 1}{\gamma}F_3 = 0. \tag{11}$$

Eq.(11) can be solved analytically which leads to the following lemma

Lemma 1 *The solution of velocity of the steady one-dimensional compressible Navier-Stokes equations (7) under the assumptions that (i) $P_r = \frac{3}{4}$; (ii) $H_L = H_R$ is*

$$(u - u_0)(u - u_1)^{k_1} = (u_R - u_0)(u_R - u_1)^{k_1} \cdot e^{-k\xi}. \tag{12}$$

Here, $\xi = \frac{x_R - x}{\varepsilon\mu}$, u_0 and u_1 are the solutions which satisfy the non-viscous Rankine-Hugnoit relations

$$u_0 = u_L, \quad u_1 = (\frac{\gamma - 1}{\gamma + 1} + \frac{2}{\gamma + 1}\frac{1}{M_a^2})u_0. \tag{13}$$

$k = \frac{\frac{\gamma+1}{\gamma}F_1 u_0 - F_2}{u_0} > 0$ *and* $k_1 = -\frac{u_1}{u_0}$.

After obtaining the solution of velocity, by using the relationship between ρ and u

$$\rho u = F_1, \tag{14}$$

and the relationship between T and u

$$\frac{1}{2}u^2 + C_p T = H_L, \tag{15}$$

obtained from Eq.(8), we can obtain the complete solution of the steady one-dimensional Navier-Stokes equations (7).

From Eq.(12) we can see that when the Reynolds tends to infinity, that is the parameter ε tends to zero, the velocity in fact varies in the form of an exponential function in a thin region near the wall x_R. This region is called as boundary layer [2–4]. Outside the boundary layer, due to the exponential function $e^{-k\xi}$ on the right hand of the solution (12) closes to zero, the velocity remains as the constant u_L. Thus, the velocity varies rapidly inside the boundary layer from u_L to u_R which leads to an extremely large gradient value at the wall. And this usually cannot be calculated accurately by the traditional DG method based on polynomial approximation space unless a very fine mesh is used in numerical computing. However, for high dimensional problems, the cost of a very fine mesh is unbearable. Next, our main work is to construct a suitable non-polynomial approximation space for solving the one-dimensional compressible Navier-Stokes equations (5) with high Reynolds number inside the boundary layer in order to get a more accurate solution without needing to refine the computing mesh.

DDG Method Based on Hybrid Approximation Space

1. Hybrid approximation space for one-dimensional compressible viscous flow

The computational domain $[x_L, x_R]$ is divided into N cells with cell interfaces $x_L = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = x_R$, and we denote the center of cell $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ by x_j, and $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$ is the cell size of I_j. By dimensional analysis of the one-dimensional compressible Navier-Stokes equations (5), we can roughly estimate the thickness of the boundary layer δ has the same order of magnitude with $\frac{1}{R_e}$ [27]. Then according to the thickness of boundary layer δ and the cell size Δx, we can determine which computing cell is inside the boundary layer.

For the computing cells inside the boundary layer, from the solution (12), it can be seen that the solution of velocity is a linear combination of the following exponential functions

$$E_u = \{u : u \in span\{1, e^{-k\xi}, e^{-2k\xi}, e^{-3k\xi}, e^{-4k\xi}, \cdots\}, \xi = \frac{x_R - x}{\varepsilon\mu}\}. \tag{16}$$

Here, k is the same as defined in Lemma 1. Then according to the relationship between ρ and u (14) and the relationship between T and u (15), it can be seen that we can take the same

non-polynomial basis functions for approximating ρ and T with that for u (16).
Next, we consider the approximation space for conservative variables: ρ, ρu, E. For the energy variable E, we can take the approximation space of T (16) for solving it. For the momentum variable ρu, however, it should be noticed that, according to Eq.(8), as

$$\rho u = F_1 = const,$$

hence, we could just take the local orthogonal Legendre polynomial approximation space

$$P_{\rho u} = \{\rho u : \rho u \in span\{1,\, x - x_j,\, (x - x_j)^2 - \frac{1}{12}\Delta x^2,\, (x - x_j)^3 - \frac{3}{20}\Delta x^2 (x - x_j), \cdots\}\}$$

for solving it.
Thus the third order hybrid polynomial and exponential approximation space which is denoted as H^2 for solving the conservative variables inside the boundary layer are given as follows

$$H_0^{(j)}(x) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad H_1^{(j)}(x) = \begin{bmatrix} e^{-k\frac{x_R - x}{\varepsilon\mu}} \\ x - x_j \\ e^{-k\frac{x_R - x}{\varepsilon\mu}} \end{bmatrix}, \quad H_2^{(j)}(x) = \begin{bmatrix} e^{-2k\frac{x_R - x}{\varepsilon\mu}} \\ (x - x_j)^2 - \frac{1}{12}\Delta x^2 \\ e^{-2k\frac{x_R - x}{\varepsilon\mu}} \end{bmatrix}. \quad (17)$$

And for the computing cells outside the boundary layer, we still use the third order Legendre polynomial approximation space P^2

$$P_0^{(j)}(x) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad P_1^{(j)}(x) = \begin{bmatrix} x - x_j \\ x - x_j \\ x - x_j \end{bmatrix}, \quad P_2^{(j)}(x) = \begin{bmatrix} (x - x_j)^2 - \frac{1}{12}\Delta x^2 \\ (x - x_j)^2 - \frac{1}{12}\Delta x^2 \\ (x - x_j)^2 - \frac{1}{12}\Delta x^2 \end{bmatrix}. \quad (18)$$

Next, we will briefly conclude the steps for implementing the hybrid DDG method.

2. DDG method based on hybrid approximation space

The hybrid DDG method for solving the one-dimensional compressible Navier-Stokes equations (5) is set up as follows

Step 1 For each computing cell I_j ($j = 1, 2, \cdots, N$), determining whether it is inside the boundary layer region or not.

Step 2 If the computing cell is inside the boundary layer, multiplying Eq.(1) by the arbitrary test functions $H_l^{(j)}(x)$ ($l = 0, 1, 2$) in (17), integrating over the interval I_j and then integrating by parts, we will obtain the weak formulation of this equation

$$\int_{I_j} Q_t H_l^{(j)} dx - \int_{I_j} V(Q) \frac{dH_l^{(j)}}{dx} dx + \Delta_+ [V(Q_{j-\frac{1}{2}}) H_l^{(j)}(x_{j-\frac{1}{2}})] = 0, \quad (19)$$

where $\Delta_+ w_j = w_{j+1} - w_j$, $V(Q) = F(Q) - F_v(Q)$ and $Q_{j+\frac{1}{2}} = Q(x_{j+\frac{1}{2}})$. And if the computing cell is outside the boundary layer, the test functions will be changed to the second order Legendre polynomial functions $P_l^{(j)}(l = 0, 1, 2)$ in (18).

Step 3 For the computing cells inside the boundary layer, defining the degrees of freedom as

$$Q_j^{(l)} = Q_j^{(l)}(t) = \int_{I_j} Q(x, t) H_l^{(j)}(x) dx, \quad l = 0, 1, 2, \quad (20)$$

and

$$Q^h(x,t) = \sum_{l=0}^{2} A_l Q_j^{(l)}(t) H_l^{(j)}(x) \quad for \quad x \in I_j, \tag{21}$$

in which $A_l = \dfrac{1}{\int_{I_j}(H_l^{(j)}(x))^2 dx}$ and $Q^h(x,t)$ is the approximation of the solution $Q(x,t)$ in H^2.

Then replacing the solution Q by Q^h and taking it into Eq.(19), we will arrive at an ODE which are the degrees of freedom $Q_j^{(l)}$ $(l = 0, 1, 2)$ must satisfy

$$\frac{d}{dt}Q_j^{(l)} + \Delta_+[\widehat{V}_{j-\frac{1}{2}} H_l^{(j)}(x_{j-\frac{1}{2}})] - \int_{I_j} V(Q^h(x,t))\frac{d}{dx}H_l^{(j)}(x)dx = 0, \quad l = 0, 1, 2, \tag{22}$$

where the numerical flux $\hat{V}$ is defined as $\hat{V} = V(Q^h, Q_x^h)$.

Similarly, if the computing cell is outside the boundary layer, the test functions will be changed to the Legendre polynomial functions $P_l^{(j)}(l = 0, 1, 2)$ (18) instead.

Step 4 For the numerical flux $\widehat{V}_{j+\frac{1}{2}} = \widehat{F}_{j+\frac{1}{2}} - \frac{1}{R_e} \cdot \widehat{F_v}_{j+\frac{1}{2}}$, we use the Lax-Friedrichs scheme

$$\begin{aligned} \widehat{F}_{j+\frac{1}{2}} &= F(Q^+_{j+\frac{1}{2}}, Q^-_{j+\frac{1}{2}}) \\ &= \frac{1}{2}[F(Q^+_{j+\frac{1}{2}}) + F(Q^-_{j+\frac{1}{2}}) - \alpha(Q^+_{j+\frac{1}{2}} - Q^-_{j+\frac{1}{2}})], \end{aligned}$$

where $\alpha = max|F'(Q)|$, to achieve total variation stability. And for the derivative term Q_x in $\hat{F}_v$, we use the piecewise linear approximation proposed in DDG method [28–30] to calculate it

$$\widehat{Q_x} = \beta_0 \frac{[Q]}{\Delta x} + \overline{Q_x}. \tag{23}$$

Here, $[Q] = Q^+ - Q^-$, $\overline{Q} = \frac{1}{2}(Q^+ + Q^-)$ and β_0 is a constant which is commonly taken as $\frac{1}{2}$.

Step 5 For the term of time, we use the third order Runge-Kutta scheme [13]

$$\begin{cases} u^{(1)} = u^n + \Delta t L(u^n), \\ u^{(2)} = \dfrac{3}{4}u^n + \dfrac{1}{4}u^{(1)} + \dfrac{1}{4}\Delta t L(u^{(1)}), \\ u^{n+1} = \dfrac{1}{3}u^n + \dfrac{2}{3}u^{(2)} + \dfrac{2}{3}\Delta t L(u^{(2)}). \end{cases} \tag{24}$$

Numerical results

Example 1. We solve the non-dimensional compressible Navier-Stokes equations (5) in $[0, 1]$ with the Reynolds number equals to 1.0×10^4. And the initial and boundary conditions are given as follows.

- Boundary condition

$$\begin{cases} u_L = 1.0, & u_R = 1.3, \\ \rho_L = 1.0, & \rho_R = (\rho_L \cdot u_L)/u_R, \\ T_L = 1.0, & T_R = (\frac{1}{2}u_L^2 + C_p T_L - \frac{1}{2}u_R^2)/C_p. \end{cases} \tag{25}$$

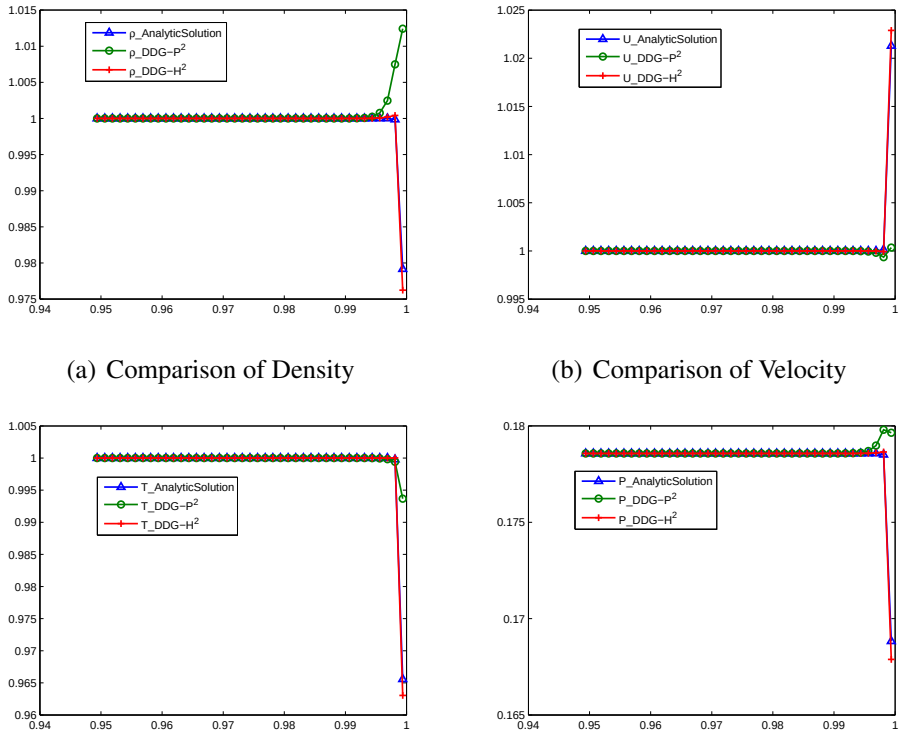
- Initial condition

$$\rho(x,0) = \begin{cases} \rho_L, & x \le x_B \\ \rho_L + \dfrac{x - x_L}{x_R - x_L}(\rho_R - \rho_L), & x > x_B, \end{cases}$$

$$u(x,0) = \begin{cases} u_L, & x \le x_B, \\ u_L + \dfrac{x - x_L}{x_R - x_L}(u_R - u_L), & x > x_B, \end{cases}$$

$$T(x,0) = \begin{cases} T_L, & x \le x_B, \\ T_L + \dfrac{x - x_L}{x_R - x_L}(T_R - T_L), & x > x_B. \end{cases}$$

Dividing the computing domain $[0,1]$ into $N = 800$ cells and taking $x_B = x_{N-3}$, then we use the DDG method based on standard P^2 polynomial basis functions ($DDG - P^2$), and DDG method based on hybrid basis functions ($DDG - H^2$) which is taken as P^2 basis functions (18) in the cells $i \le N-1$ and H^2 basis functions (17) in the cell $i = N$ to solve this equation, respectively, then compare the numerical results with the analytic solution (12). The results are shown in the following figure and table.



(a) Comparison of Density

(b) Comparison of Velocity

(c) Comparison of Temperature

(d) Computing error of Velocity

Figure 1: Example 1: Comparison of $DDG - P^2$ & $DDG - H^2$ with analytic solution.

Table 1: Example 1: Computing errors of flow variables (L_∞).

Error / Method	ρ	U	T	P
$DDG-P^2$	0.0333	0.0209	0.0281	0.0108
$DDG-H^2$	0.0029	0.0016	0.0025	0.0009

From Fig. 1 and Tab. 1 we can see that as the grid size we take is not small enough compared to the thickness of boundary layer, as expected, the traditional $DDG-P^2$ gives a poor flow variables profile. The flow field values outside the boundary layer might be correct but inside the boundary layer, they are obviously inaccurate and even have numerical oscillations. However, from Fig. 1 and Tab. 1 it can be seen that although we only change the basis functions in one computing cell $i=N$, the numerical results computed by the $DDG-H^2$ method proposed in this paper are very close to the analytic solution everywhere.

As mentioned above, because the Reynolds number taken in this example is very high, there have a steep gradients of flow variables near the wall, and these gradient values have an important effect on computing the coefficient of skin friction and other physical quantities in high dimensional problems. Thus, next, we will compare the numerical results of gradient values at the wall x_R.

Table 2: Example 1: Percentage errors of gradient values at the wall x_R.

Error(%) / Method	ρ_x	u_x	T_x	P_x
$DDG-P^2$	26.1745	41.8165	45.2866	38.7611
$DDG-H^2$	1.1076	2.0544	3.3219	2.6517

It is obvious to see from Tab. 2 that, under the same computing grid, although it seems that the numerical results computed by $DDG-P^2$ method are acceptable from the point of L_∞ computing error, but it gives poor approximation for the gradient values of flow variables at the wall x_R. However, the $DDG-H^2$ method is far more effective than the $DDG-P^2$ method in computing the gradient values and it can provide a very accurate approximation results.

Example 2. We also solve Eq.(5) with the same boundary and initial conditions with those of Example 1. We increase the number of computing grids N to 1600, 2400 and 3200, respectively, and compute the gradients of flow variables at the wall x_R by the $DDG-P^2$ method again to see the variations of percentage errors with the number of grids. The results are shown in the following table.

Table 3: Example 2: Percentage errors of gradient values at the wall x_R ($DDG - P^2$).

Error(%) / N	ρ_x	u_x	T_x	P_x
800	26.1745	41.8165	45.2866	38.7611
1600	7.4528	17.4923	16.9579	13.9999
2400	2.5276	8.7675	7.2083	5.8779
3200	0.9195	5.0338	3.4916	2.8179

From Tab. 3 we can see that the percentage errors of gradient values are decreasing slowly with the grid size becoming smaller. We can also that compared to the $DDG - H^2$ method, the $DDG - P^2$ method should take nearly four times the number of grids more than that of the $DDG - H^2$ method does in order to obtain the same order of percentage errors. And this, from the other perspective, demonstrate that the DDG method based on hybrid basis functions (17) is effective in decreasing the number of grids needed in solving the Navier-Stokes equations with high Reynolds number and can obtain a very accurate result of gradient value on a coarse mesh. It can save nearly 75% of the grids number.

Example 3. Here, we use higher order accuracy methods: $DDG - P^3$ and $DDG - P^4$ methods to solve Example 1 on a coarse mesh ($N = 800$), then compare the numerical results with that of $DDG - H^2$ method, the comparisons are shown in the following table.

Table 4: Example 3: Percentage errors of gradient values at the wall x_R.

Error(%) / Method	ρ_x	u_x	T_x	P_x
$DDG - P^2$	26.1745	41.8165	45.2866	38.7611
$DDG - P^3$	4.6980	12.7386	11.0475	8.9381
$DDG - P^4$	0.3093	1.7141	2.1390	1.6020
$DDG - H^2$	1.1076	2.0544	3.3219	2.6517

It is obviously to see from Table 4 that, better results are obtained when using a high order DDG method based on polynomial basis functions (18). And it nearly needs to take fifth order accuracy DDG method based on the polynomial basis functions in order to obtain the same order of percentage errors with those of the $DDG - H^2$ method proposed in our paper.

Conclusions

The main objective of this paper is to propose a suitable non-polynomial approximation basis function for solving the compressible Navier-Stokes equation with high Reynolds number which has a thin boundary layer near the wall of object. According to the analytic solution of the one-dimensional steady compressible Navier-Stokes equations under the assumptions that the Prandtl number is taken as $\frac{3}{4}$ and the coefficient of viscosity is taken as a constant, we proposed a DDG method based on a hybrid exponential and polynomial approximation space for solving the Navier-Stokes equations. Numerical tests have shown compared to the DDG method based on the standard polynomial approximation space, the DDG method based on the hybrid approximation space proposed in this paper can obtain a more accuracy results, especially for

the gradients values of flow variables on the wall, with less number of degrees of freedom over a coarse grid. Further research focusing on high dimensional problems is on going.

References

1. I. MARUSIC, B. J. MCKEON, P. A. MONKEWITZ, H. M. NAGIB, A. J. SMITS, *Wall-bounded Turbulent Flows at High Reynolds Numbers: Recent Advances and Key Issues*, Physics of Fluids, 22(2010), pp. 065103.

2. L. PRANDTL, *Verhandlung des III Internationalen Mathematiker-Kongresses (heidelberg, 1904)*, A. Krazer, ed., Teubner, Leipzig, Germany 1905, pp. 484–491.

3. H. SCHLICHTING, *Boundary-layer Theory*, 7th ed., McGrawHill, New York, 1979.

4. TUNCER CEBECI, JEAN COUSTEIX, *Modeling and Computation of Boundary-layer Flows: Laminar, Turbulent and Transitional Boundary Layers in Incompressible and Compressible Flows (second revised and extended edition)*, Horizon Publishing, 2005.

5. KARMAN S. L., *Unstructured Viscous Layer Insertion Using Linear-elastic Smoothing*, AIAA Journal, 45(2007), pp. 168–180.

6. GUILLAUME V., FORNIER Y., AND BOUBEKEUR T, *Hybrid Viscous Layer Insertion in a Tetrahedral Mesh*, In IMR (Research Note), 2012.

7. ITO Y., NAKAHASHI K., *Unstructured Mesh Generation for Viscous Flow Computations*, IMR, 2002, pp. 367–377.

8. BAHRAINIAN S. S., MEHRDOOST Z., *An Automatic Unstructured Grid Generation Method for Viscous Flow Simulations*, Mathematics and Computers in Simulation, 2012.

9. GARIMELLA R. V., SHEPHARD M. S., *Boundary Layer Mesh Generation for Viscous Flow Simulations*, International Journal for Numerical Methods in Engineering, 49(200), pp. 193–218.

10. RAJEEV J., TIMOTHY J. T., *PostBL: Post-Mesh Boundary Layer Mesh Generation Tool*, In: Sarrate J., Staten M. (eds) Proceedings of the 22nd International Meshing Roundtable, Springer, Cham, 2014.

11. Z.J. WANG, *High-order Methods for the Euler and NavierCStokes Equations on Unstructured Grids*, Progress in Aerospace Sciences, 43(2007), pp. 1–41.

12. , *RANS Solutions Using High Order Discontinuous Galerkin Methods*, Aiaa Aerospace Science Meeting & Exhibit Reno for Discontinuous Galerkin, 2007.

13. B. COCKBURN, C. W. SHU,*Runge-Kutta Discontinuous Galerkin Methods for Convection-Dominated Problems*, Journal of Scientific Computing, 16 (2001) , pp. 173–261.

14. BERNARDO COCKBURN , GEORGE E. KARNIADAKIS AND CHI-WANG SHU,*The Development of Discontinuous Galerkin Methods*, 1999.

15. W. H. REED AND T. R. HILL, *Triangular Mesh Methods for the Neutron Transport Equation*, Tech.report LA-UR-73-479, Los Alamos Scientific Laboratory, Loa Alamos, NM, 1973.

16. B. COCKBURN, C. W. SHU, *TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Scalar Conservation Laws II: General framework*, Math. Comp., 52 (1989), pp. 411–435.

17. B. COCKBURN, C. W. SHU, *TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Scalar Conservation Laws III: One dimensional systems*, J. Comput. Phys., 84 (1989), pp. 90–113.

18. B. COCKBURN, C. W. SHU, *TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Scalar Conservation Laws IV: The multidimensional case*, Math. Comp., 54 (1990), pp. 545–581.

19. B. COCKBURN, C. W. SHU, *The Runge-Kutta Local P^1 Discontinuous Galerkin Finite Element Method for Scalar Conservation Laws*, RAIRO Modél. Math. Comp., 52 (1989), pp. 411–435.

20. B. COCKBURN, G. E. KARNIADAKIS AND C.-W. SHU, *The Local Discontinuous Galerkin Method for Time-dependent Convection-diffusion Systems*, SIAM J. Numer. Anal., 35 (1998), pp. 2440–2463.

21. YUAN LING, C. W. SHU, *Discontinuous Galerkin Method Based on Non-polynomial Approximation Spaces*, J. Comput. Phys., 218 (2006), pp. 295–323.

22. B. COCKBURN, F. LI AND C.-W. SHU, *Locally Divergence-free Discontinuous Galerkin Methods for the Maxwell Equations*, Journal of Computational Physics, 194(2004), pp. 588–610.

23. F. LI, C.-W. SHU, *Locally Divergence-free Discontinuous Galerkin Methods for MHD Equations*, Journal of Scientific Computing, 22-23(2005), pp. 413-442.

24. F. LI, C.-W. SHU, *Reinterpretation and Simplified Implementiation of a Discontinuous Galerkin Method for Hamilton-Jacobi Equations*, Applied Mathematics Letters, 18(2005), pp. 1204–1209.

25. F. LI, C.-W. SHU, *A Local-structure-preserving Local Discontinuous Galerkin Method for the Laplace Equation*, Methods and Applications of Analysis, 2(2006), pp. 215-C234

26. L. YUAN, C.-W. SHU, *Discontinuous Galerkin Method Based on Non-polynomial Approximation Spaces*, Journal of Computational Physics, 218(2006), pp. 295–323.

27. FAN ZHANG, TIEGANG LIU, *Solutions of Steady One-dimensional Compressible Navierstokes Equations under Uniform Enthalpy with $P_r = \frac{3}{4}$*. (To appear)

28. HAILIANG LIU, JUE YAN, *The Direct Discontinuous Galerkin (DDG) Methods for Diffusion Problems*, SIAM J.Numer. Anal, 47 (2009), pp. 675–698.

29. JIAN CHENG, XIAOQUAN YANG, XIAODONG LIU, TEIGANG LIU AND HONG LUO, *A Direct Discontinuous Galerkin Method for the Compressible NavierCStokes Equations on Arbitrary Grids*, Journal of Computational Physics, 327(2016), pp. 484–502.

30. JIAN CHENG, XIAODONG LIU, TIEGANG LIU AND HONG LUO, *A Parallel, High-Order Direct Discontinuous Galerkin Method for the Navier-Stokes Equations on 3D Hybrid Grids*, Communications in Computational Physics, 21(2017), pp. 1231–1257.

A novel method to improve the multiple-scales solution of the forced strongly nonlinear oscillators

***Hai-En Du, Guo-Kang Er, and Vai Pan Iu**

Department of Civil and Environmental Engineering, University of Macau, Macau SAR, China.

*Presenting author and Corresponding author: haiendu@outlook.com

Abstract

We propose a novel procedure to improve the solution obtained by perturbation methods for analyzing the solutions of strongly nonlinear systems. The multiple-scales method, one of the perturbation method, is widely used in many areas. However, multiple-scales method fails in analyzing the solutions of oscillators if the oscillator nonlinearity is strong. We apply the proposed procedure to improve the multiple-scales method to obtain the optimum solution of the forced oscillator with strongly nonlinear restoring and inertial forces. The solutions obtained from multiple-scales method and the proposed method are examined by the numerical solution obtained from 4^{th}-order Runge-Kutta method. The results show that the proposed method is effective for the oscillators with nonlinear restoring force as well as nonlinear inertial force even if the nonlinearities are strong. Numerical results and comparison show that the proposed method can improve the solution a lot in comparison to the solution obtained by conventional multiple-scales method.

Keywords: Perturbation method; Strong nonlinearity; Nonlinear restoring force; Nonlinear inertial force; Forced Vibration.

Introduction

Strongly nonlinear systems can be found in many structural applications, such as the vibrations of mass-spring system, cable, cantilever with large deflections, etc [1]. Though numerical methods have been widely applied for numerical solutions of nonlinear vibration problems, the solutions of some strongly nonlinear oscillators can still not be completely obtained. Therefore, the studies on the methods for approximate analytical solutions of strongly nonlinear oscillators are attractive. Perturbation methods have been used to obtain the approximate analytical solution for a long time. However, the assumption for perturbation methods limits their applications. Perturbation methods are invalid if the nonlinearity within the system is strong because of the assumption of small perturbation parameter in the system. In order to analyze the vibration of oscillator with strong nonlinearity, some methods have been developed and studied in recent years. They can be categorized as (1) harmonic balance method, (2) variational iteration method, (3) linearized perturbation method, (4) parameter expansion perturbation method, (5) various modified Lindstedt-Poincaré methods and (6) homotopy analysis method. Each of these methods can be applied for obtaining the approximate solutions of a wide class of nonlinear oscillators without introducing a small perturbation parameter as classical perturbation methods do [2]. Wu and Lim proposed a method by combining the linearization of equation of motion and harmonic balance method to analyze the free vibration of an ordinary differential equation with odd nonlinear restoring force [3]. Cheung and Iu applied the harmonic balance method to analyze the forced vibration of a dynamical system with quadratic and cubic nonlinearities [4]. Hu applied a modified iteration procedure to a quadratic nonlinear oscillator (QNO) and obtained an improved solution in comparison to those obtained by the first-order harmonic balance method [5].

Shakeri and Dehghan adopted the variational iteration method to solve the Klein-Gordon equation and it shows that the solution converges fast [6]. Marinca and Herisanu proposed a perturbation technique by combining the iteration methods and the solution obtained by this new method agrees well with exact solution [7]. The linearized perturbation technique is applied to a Duffing oscillator with 5th-order nonlinearity [8]. Xu applied He's parameter-expanding method (PEM) to determine the limit cycles of the strongly nonlinear oscillators [9]. With this method, a strongly nonlinear oscillator with large perturbation parameter is transformed into an oscillator with small parameter. Chen et al. proposed a modified Lindstedt-Poincaré method for the analytical approximate solution of limit cycles in three-dimensional nonlinear autonomous dynamical systems [10]. In 2009, Pakdemirli proposed a method named multiple-scales Lindstedt-Poincaré (MSLP) method by combining the multiple-scales (MS) method and Lindstedt-Poincaré (LP) method. This method has been applied to analyze the free vibration of three oscillators which are the damped linear oscillator, undamped Duffing oscillator and damped Duffing oscillator [11]. Later, the MSLP method was extended to analyze the forced vibration of strongly nonlinear Duffing oscillator [12]. Liao proposed an optimal homotopy analysis method by introducing a two-parameter family equation to find the fastest convergence solution to the Blasius boundary-layer flows problem in fluid mechanics [13]. Razzak and Molla combined the general Struble's technique and homotopy perturbation method to analyze damped and driven strongly nonlinearDuffing oscillator and strongly nonlinear van der Pol oscillator with damping [14].

Since the validity condition for perturbation method to give a valid solution is that the ratio of the amplitude of $O(\varepsilon^1)$ solution and that of $O(\varepsilon^0)$ solution is much less than unity [15], the method proposed in this paper is based on the objective that the ratio of the amplitude of $O(\varepsilon^1)$ solution and that of $O(\varepsilon^0)$ solution is minimized. To do so, an equivalent oscillator is formulated by splitting the parameters in nonlinear restoring force and nonlinear inertial force. The introduced unknown nonlinearity parameters can be determined with the objective that the ratio of the amplitude of $O(\varepsilon^1)$ solution and that of $O(\varepsilon^0)$ solution is minimized. An oscillator is analyzed by multiple-scales (MS) method and the proposed method which is named parameter-split-multiple-scales (PSMS) method. The solutions obtained by these methods are compared to the numerical solutions obtained by the 4th-order Runge-Kutta method. The accuracy and the effectiveness of PSMS method are examined by numerical analysis.

Procedures for optimizing the solution obtained by the multiple-scales method

Considering the following nonlinear oscillator

$$\ddot{y} + c\varepsilon^2 \dot{y} + \omega_0^2 y + \varepsilon g\left(y, \dot{y}, \ddot{y}\right) = F\varepsilon^2 \cos\left(\Omega t\right) \tag{1}$$

where y is displacement, t is time, c is damping coefficient, ω_0 is the natural frequency of the oscillator, ε is perturbation parameter, F is excitation amplitude, Ω is excitation frequency and $g\left(y, \dot{y}, \ddot{y}\right)$ is a nonlinear function and given as

$$g\left(y, \dot{y}, \ddot{y}\right) = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{l} \eta^{(i,j,k)} y^i \dot{y}^j \ddot{y}^k \tag{2}$$

where $\eta^{(i,j,k)}$ are nonlinear parameters which reflect the degrees of nonlinearity and $\sum i + j + k \geq 2$. The nonlinear parameters $\eta^{(i,j,k)}$ are split and expressed by two terms as follows.

$$\eta^{(i,j,k)} = \eta_1^{(i,j,k)} + \eta_2^{(i,j,k)}\varepsilon \tag{3}$$

Then, Eq. (1) is written as

$$\ddot{y} + c\varepsilon^2\dot{y} + \omega_0^2 y + \varepsilon g_1\left(y,\dot{y},\ddot{y}\right) + \varepsilon^2 g_2\left(y,\dot{y},\ddot{y}\right) = F\varepsilon^2\cos\left(\Omega t\right) \tag{4}$$

where $g_1\left(y,\dot{y},\ddot{y}\right) = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{l}\eta_1^{(i,j,k)} y^i \dot{y}^j \ddot{y}^k$ $\quad g_2\left(y,\dot{y},\ddot{y}\right) = \sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{l}\eta_2^{(i,j,k)} y^i \dot{y}^j \ddot{y}^k$ (5)

In the analysis with perturbation method, the response of the oscillator is assumed to be

$$y = y_0 + \varepsilon y_1 + \varepsilon^2 y_2 + O\left(\varepsilon^3\right) \tag{6}$$

Substituting Eq. (7) into Eq. (4) leads to

$$\begin{aligned}
&\ddot{y}_0 + \varepsilon\ddot{y}_1 + \varepsilon^2\ddot{y}_2 + c\varepsilon^2\dot{y}_0 + \omega_0^2\left(y_0 + \varepsilon y_1 + \varepsilon^2 y_2\right) \\
&+\varepsilon\sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{l}\eta_1^{(i,j,k)}\left(y_0 + \varepsilon y_1 + \varepsilon^2 y_2\right)^i\left(\dot{y}_0 + \varepsilon\dot{y}_1 + \varepsilon^2\dot{y}_2\right)^j\left(\ddot{y}_0 + \varepsilon\ddot{y}_1 + \varepsilon^2\ddot{y}_2\right)^k \\
&+\varepsilon^2\sum_{i=0}^{n}\sum_{j=0}^{m}\sum_{k=0}^{l}\eta_2^{(i,j,k)}\left(y_0 + \varepsilon y_1 + \varepsilon^2 y_2\right)^i\left(\dot{y}_0 + \varepsilon\dot{y}_1 + \varepsilon^2\dot{y}_2\right)^j\left(\ddot{y}_0 + \varepsilon\ddot{y}_1 + \varepsilon^2\ddot{y}_2\right)^k \\
&= F\varepsilon^2\cos\left(\Omega t\right)
\end{aligned} \tag{7}$$

Equating the coefficients of ε^s (s=0, 1, 2) to zero and eliminating the secular terms one can obtain an approximate steady-state response to the oscillator in the form of

$$y = A\cos\left(\Omega t - \gamma\right) + Y_1\cos\left[3\left(\Omega t - \gamma\right)\right] + Y_2\cos\left[5\left(\Omega t - \gamma\right)\right] \tag{8}$$

in which A is the steady-state response amplitude, γ is the steady-state phase angle, Y_1 and Y_2 are the amplitudes of $O(\varepsilon^1)$ solution and $O(\varepsilon^2)$ solution, respectively. They are also the functions of $\eta_1^{(1,1,0)},\ldots,\eta_2^{(n,m,l)}$. Then the values of $\eta_1^{(i,j,k)}$ and $\eta_2^{(i,j,k)}$ are determined by a numerical iteration procedure.

Damped and driven Duffing equation with nonlinear inertial force

Consider a damped and driven Duffing equation with nonlinear inertial force as follows.

$$\ddot{q} + 2u\varepsilon^2\dot{q} + \omega_0^2 q + \alpha\varepsilon q\dot{q}^2 + \alpha\varepsilon q^2\ddot{q} + \beta\varepsilon q^3 = F\varepsilon^2\cos\left(\Omega t\right) \tag{9}$$

which can find its applications in the nonlinear vibrations of cantilever with large deflection. The nonlinear parameter α and β are split by

$$\alpha = \alpha_1 + \alpha_2\varepsilon \tag{10}$$

$$\beta = \beta_1 + \beta_2\varepsilon \tag{11}$$

The oscillator response is expressed as

$$q = q_0\left(T_0,T_1,T_2\right) + \varepsilon q_1\left(T_0,T_1,T_2\right) + \varepsilon^2 q_2\left(T_0,T_1,T_2\right) + O\left(\varepsilon^3\right) \tag{12}$$

where T_0, T_1 and T_2 are different time scales with multiple-scales method which are given by

$$T_0 = t, T_1 = \varepsilon t, T_2 = \varepsilon t^2. \tag{13}$$

By chain rule, the operators of time derivatives are

$$\frac{d}{dt} = D_0 + \varepsilon D_1 + \varepsilon^2 D_2 + \dots, \tag{14}$$

$$\frac{d^2}{dt^2} = D_0^2 + 2\varepsilon D_0 D_1 + \varepsilon^2 \left(D_1^2 + 2D_0 D_2\right) + \dots, \tag{15}$$

where $D_n = \partial / \partial T_n$ and $D_n^2 = \partial^2 / \partial T_n^2$. Substituting Eq. (10)-(15) into Eq. (9) and setting the coefficients of $\varepsilon^m \left(m = 0, 1, 2\right)$ to zero lead to the following equations.

$$O\left(\varepsilon^0\right): \quad D_0^2\left(q_0\right) + \omega_0^2 q_0 = 0, \tag{16}$$

$$O\left(\varepsilon^1\right): \quad D_0^2\left(q_1\right) + \omega_0^2 q_1 = -2D_0 D_1\left(q_0\right) - \alpha_1 q_0 \left[D_0\left(q_0\right)\right]^2 - \alpha_1 q_0^2 D_0^2\left(q_0\right) - \beta_1 q_0^3 \tag{17}$$

$$\begin{aligned} O\left(\varepsilon^2\right): \quad & D_0^2\left(q_2\right) + \omega_0^2 q_2 = -2D_0 D_1\left(q_1\right) - D_1^2\left(q_0\right) - \beta_2 q_0^3 - 2D_0 D_2\left(q_0\right) \\ & -2uD_0\left(q_0\right) - 3\beta_1 q_0^2 q_1 - \alpha_1 q_1 \left[D_0\left(q_0\right)\right]^2 - \alpha_2 q_0 \left[D_0\left(q_0\right)\right]^2 - \alpha_2 q_0^2 D_0^2\left(q_0\right) \\ & -2\alpha_1 q_0 D_0\left(q_0\right) D_0\left(q_1\right) - 2\alpha_1 q_0 D_0\left(q_0\right) D_1\left(q_0\right) - \alpha_1 q_0^2 D_0^2\left(q_1\right) \\ & -2\alpha_1 q_0^2 D_0 D_1\left(q_0\right) - 2\alpha_1 q_0 q_1 D_0^2\left(q_0\right) + F\cos\left(\Omega t\right). \end{aligned} \tag{18}$$

The solution of the $O\left(\varepsilon^0\right)$ equation is

$$q_0 = C\left(T_1, T_2\right) e^{i\omega_0 T_0} + \overline{C}\left(T_1, T_2\right) e^{-i\omega_0 T_0} \tag{19}$$

where C is a function of time scales T_1 and T_2 which can be determined by omitting the secular terms in the $O\left(\varepsilon^1\right)$ equation in the following. Substituting Eq. (19) into the righthand side of the $O\left(\varepsilon^1\right)$ equation and eliminating the secular terms yield

$$2i\omega_0 D_1\left(C\right) + 3\beta_1 C^2 \overline{C} - 2\alpha_1 \omega_0^2 C^2 \overline{C} = 0 \tag{20}$$

and

$$q_1 = \Lambda e^{3i\omega_0 T_0} + \overline{\Lambda} e^{-3i\omega_0 T_0}, \tag{21}$$

in which

$$\Lambda = \frac{\beta_1 C^3}{8\omega_0^2} - \frac{\alpha_1 C^3}{4}. \tag{22}$$

Substituting the expressions of q_0 and q_1 into the $O\left(\varepsilon^2\right)$ equation, eliminating the secular terms, and using the expression $\Omega = \omega_0 + \varepsilon^2 \sigma$ where σ is a detuning parameter that can be determined if Ω is given, it gives

$$q_2 = \Gamma_1 e^{3i\omega_0 T_0} + \Gamma_2 e^{5i\omega_0 T_0} + \overline{\Gamma}_1 e^{-3i\omega_0 T_0} + \overline{\Gamma}_2 e^{-5i\omega_0 T_0}, \tag{23}$$

in which

$$\Gamma_1 = \frac{9\alpha_1 C^4 \overline{C}}{16} - \frac{\alpha_2 C^3}{4} + \frac{\beta_2 C^3}{8\omega_0^2} - \frac{\alpha_1\beta_1 C^4 \overline{C}}{8\omega_0^2} - \frac{21\beta_1^2 C^4 \overline{C}}{64\omega_0^4}, \tag{24}$$

$$\Gamma_2 = \frac{3\alpha_1^2 C^5}{16} - \frac{\alpha_1\beta_1 C^5}{8\omega_0^2} + \frac{\beta_1^2 C^5}{64\omega_0^4}. \tag{25}$$

$D_2\left(C\right)$ is selected to eliminate the secular terms and expressed as

$$D_2\left(C\right) = \frac{Fe^{i\sigma T_2}}{4i\omega_0} - uC - \frac{9\alpha_1^2\omega_0 C^3 \overline{C}^2}{4i} + \frac{9\alpha_1\beta_1 C^3 \overline{C}^2}{4i\omega_0} + \frac{15\beta_2^2 C^3 \overline{C}^2}{16i\omega_0^3} \tag{26}$$

The time derivative of C can be expressed as

$$\frac{dC}{dt} = \varepsilon D_1\left(C\right) + \varepsilon^2 D_2\left(C\right) + O\left(\varepsilon^3\right). \tag{27}$$

The polar form of C is assumed to be

$$C = \frac{1}{2} Ae^{ib}, \tag{28}$$

where A is the response amplitude and b is the phase of oscillator response. Substituting Eqs. (20), (26) and (28) into Eq. (27) and separating the real and imaginary parts yield

$$\dot{A} = \frac{F\varepsilon^2}{2\omega_0} \sin\gamma - uA\varepsilon^2 \tag{29}$$

and

$$\dot{\gamma} = \varepsilon^2\sigma + \varepsilon\left(\frac{A^2\alpha_1\omega_0}{4} - \frac{3A^2\beta_1}{8\omega_0}\right) + \varepsilon^2\left(\begin{array}{l}\frac{F\cos\gamma}{2A\omega_0} - \frac{9A^4\alpha_1^2\omega_0}{64} + \frac{9A^4\alpha_1\beta_1}{64\omega_0} \\ + \frac{15A^4\beta_1^2}{256\omega_0^3} + \frac{\alpha_2 A^2\omega_0}{4} - \frac{3\beta_2 A^2}{8\omega_0}\end{array}\right), \tag{30}$$

where $\gamma = \sigma T_2 - b$.

For steady state, $\dot{A}$ and $\dot{\gamma}$ are equal to zero. Then the frequency response curve can be obtained by eliminating γ and σ in Eq. (30). The relation between the excitation frequency and the response amplitude at steady state can be obtained to be

$$\begin{aligned}\Omega = \omega_0 &- \frac{A^2\alpha_1\omega_0\varepsilon}{4} + \frac{3A^2\beta_1\varepsilon}{8\omega_0} + \frac{9A^4\alpha_1^2\omega_0\varepsilon^2}{64} - \frac{9A^4\alpha_1\beta_1\varepsilon^2}{64\omega_0} - \frac{15A^4\beta_1^2\varepsilon^2}{256\omega_0^3} \\ &- \frac{\alpha_2 A^2\omega_0\varepsilon^2}{4} + \frac{3\beta_2 A^2\varepsilon^2}{8\omega_0} \mp \frac{F\varepsilon^2}{2A\omega_0}\sqrt{1 - \frac{4A^2u^2\omega_0^2}{F^2}}.\end{aligned} \tag{31}$$

The approximate response of the oscillator can be expressed as

$$q = A\left\{\cos\left(\Omega t - \gamma\right) + X_1 \cos\left[3\left(\Omega t - \gamma\right)\right] + X_2 \cos\left[5\left(\Omega t - \gamma\right)\right]\right\}, \tag{32}$$

in which

$$X_1 = \frac{9A^4\alpha_1^2\varepsilon^2}{256} - \frac{A^2\alpha_1\varepsilon}{16} - \frac{21A^4\beta_1^2\varepsilon^2}{1024\omega_0^4} + \frac{A^2\beta_1\varepsilon}{32\omega_0^2} - \frac{\alpha_2 A^2\varepsilon^2}{16} + \frac{\beta_2 A^2\varepsilon^2}{32\omega_0^2} - \frac{A^4\alpha_1\beta_1\varepsilon^2}{128\omega_0^2} \tag{33}$$

and

$$X_2 = \frac{3A^4\alpha_1^2\varepsilon^2}{256} + \frac{A^4\beta_1^2\varepsilon^2}{1024\omega_0^4} - \frac{A^4\alpha_1\beta_1\varepsilon^2}{128\omega_0^2}. \tag{34}$$

Due to the relations given by Eqs. (10) and (11), two of the parameters α_1, α_2, β_1 and β_2 are independent if the values of α and β are given. Consider the parameters α_1 and β_1 as independent parameters. In order get the optimum solution, the values of α_1 and β_1 are determined such that the absolute value of X_1 is minimized.

Case 1: $\alpha = 0$

When $\alpha = 0$, the considered oscillator can be regarded as a damped and driven Duffing oscillator which can be found in many applications such as the forced vibrations of pendulum, isolator, electrical circuit [1].

The frequency response curves obtained by the proposed method and the multiple-scales method are compared to the frequency response curve obtained by the fourth-order Runge-Kutta method to examine for the effectiveness of the methods.

The parameters of nonlinear oscillators are listed in Table 1.

Table 1. Oscillator parameters

Oscillator	ε	ω_0	u	α	β	F
1	0.1	1	2	0	10	30
2	0.1	1	2	0	0.1	30
3	0.1	1	2	0	10	30

The frequency response curves of oscillators 1, 2 and 3, obtained by the proposed method, the multiple-scales method and the numerical simulation are presented in Figs. 1-3, respectively.

Case 2: $\alpha \neq 0$

When $\alpha \neq 0$, the considered oscillator can be regarded as a damped and driven Duffing oscillator with nonlinear inertial forces ($q^2\ddot{q}$ and $q\dot{q}^2$) which can be found in the forced vibrations of beams [16].

The frequency response curves obtained by the proposed method and the multiple-scales method are compared to the frequency response curve obtained by the fourth-order Runge-Kutta method to examine the effectiveness of the methods.

The parameters of the nonlinear oscillators are listed in Table 2.

Table 2. Oscillator parameters

Oscillator	ε	ω_0	u	α	β	F
4	0.1	1	2	0.1	10	30
5	0.1	1	2	2	0.1	30
6	0.1	1	2	2	10	30

The frequency response curves of oscillators 4, 5 and 6, obtained by the proposed method, the multiple-scales method and the numerical simulation are presented in Figs. 4-6, respectively.

Conclusions

A novel method named parameter-split-multiple-scales method is proposed to improve the solution obtained by perturbation method based on the objective that the ratio of the amplitude of $O(\varepsilon^1)$ solution and that of $O(\varepsilon^0)$ solution is minimized. The forced vibration of an oscillator with strongly nonlinear restoring and inertial forces is analyzed by the proposed PSMS method, MS method and 4th-order Runge-Kutta method. We have first studied the case that $\alpha = 0$ to examine the validity of the proposed method when nonlinear restoring force is large. After that, we have studied the oscillator with nonlinear restoring and inertial forces ($\alpha \neq 0$). The results show that the proposed method works for the oscillators with strongly nonlinear restoring force and/or strongly nonlinear inertial force. It can improve make the solutions improved a lot compared to the conventional multiple scales method.

Acknowledgement

The results presented in this paper were obtained under the supports of the Research Committee of University of Macau (Grant No. MYRG2014-00084-FST) and the Science and Technology Development Fund of Macau (Grant No. 043/2013/A).

References

[1] Kovacic I., Brennan M.J. (2011) *The Duffing equation: nonlinear oscillators and their behaviour*, John Wiley & Sons, New York.
[2] He J.H. (2000) A review on some new recently developed nonlinear analytical techniques, *International Journal of Nonlinear Sciences and Numerical Simulation*, **1(1)**, 51--70.
[3] Wu B.S., Lim C. W. (2004) Large amplitude non-linear oscillations of a general conservative system, *International Journal of Non-Linear Mechanics*, **39**,859--870.
[4] Cheung Y.K. and Iu V.P. (1987) An implicit implementation of harmonic balance method for non-linear dynamic systems, *Engineering Computation*, **5(2)**, 134—140.
[5] Hu H. (2006) Solutions of a quadratic nonlinear oscillator: Iteration procedure, *Journal of Sound and Vibration*, **298(4)**, 1159--1165.
[6] Shakeri F. and Dehghan M. (2008) Numerical solution of the Klein-Gordon equation via he's variational iteration method, *Nonlinear Dynamics*, **51(1-2)**, 89-97.
[7] Marinca V. and Herisanu N. (2006) A modified iteration perturbation method for some nonlinear oscillation problems, *Acta Mechanica*, **184(1-4)**, 231-242.
[8] He J.H. (2003) Linearized perturbation technique and its applications to strongly nonlinear oscillators, *Computational Mathematics and Computation*, **45(1)**, 1--8.
[9] Xu L. (2007) Determination of limit cycle by He's parameter-expanding method for strongly nonlinear oscillators, *Journal of Sound and Vibration*, **302(1)**, 178--184.
[10] Chen S.H., Shen J.H., Sze K.Y. (2009) A new perturbation procedure for limit cycle analysis in three-dimensional nonlinear autonomous dynamical systems, *Nonlinear Dynamics*, **56(3)**, 255--268.
[11] Pakdemirli M., Karahan M. and Boyacı H. (2009) A new perturbation algorithm with better convergence properties: Multiple scales lindstedt poincaré method, *Mathematical and Computational Applications*, **14(1)**, 31-44.
[12] Pakdemirli M., Karahan M. and Boyacı H. (2011) Forced vibrations of strongly nonlinear systems with multiple scales lindstedt poincaré method, *Mathematical and Computational Applications*, **16(4)**, 879-889.

[13] Liao S.J. (2010) An optimal homotopy-analysis approach for strongly nonlinear differential equations, *Communication in Nonlinear Science and Numerical Simulation*, **15**, 2003--2016.
[14] Razzak M.A. and Molla M.H.U. (2015) A new analytical technique for strongly nonlinear damped forced systems, *Ain Shams Engineering Journal*, **6(4)**, 1225—1232.
[15] Nayfeh A.H. (2008) *Perturbation methods*, Wiley-VCH Verlag GmbH and Co. KGaA, Weinheim.
[16] Malatkar P. (2003) Nonlinear vibrations of cantilever beams and plates. Doctoral thesis in Virginia Polytechnic Institute and State University.

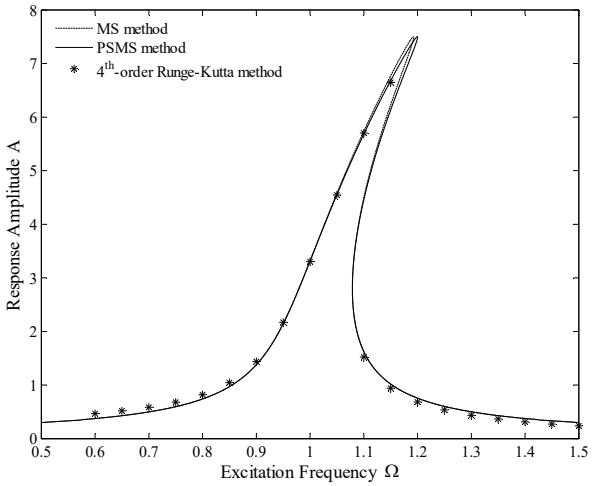


Figure 1. FRCs of oscillator 1 by the proposed method, the MS method and numerical simulation.

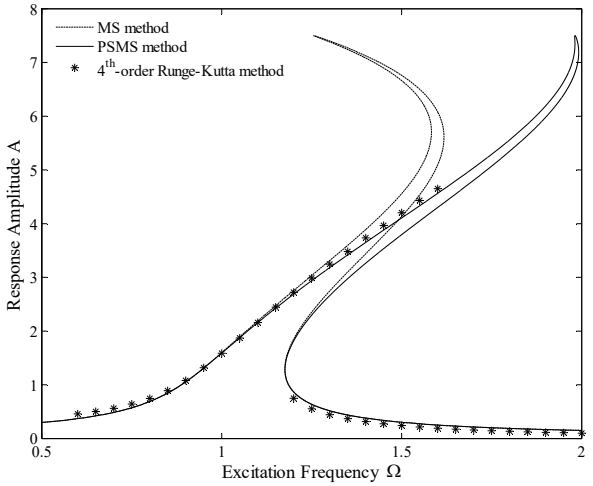


Figure 2. FRCs of oscillator 2 by the proposed method, the MS method and numerical simulation.

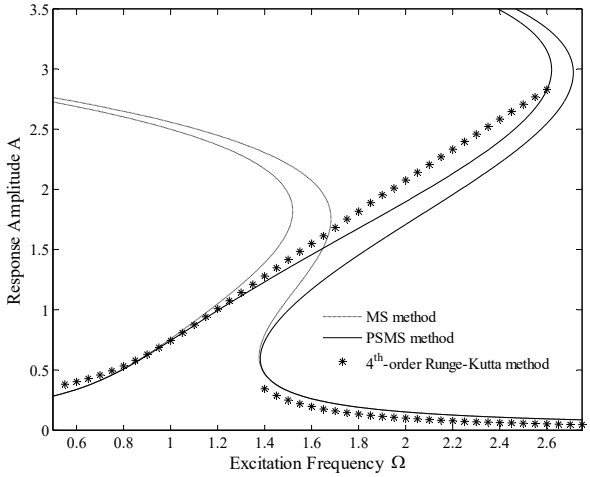


Figure 3. FRCs of oscillator 3 by the proposed method, the MS method and numerical simulation.

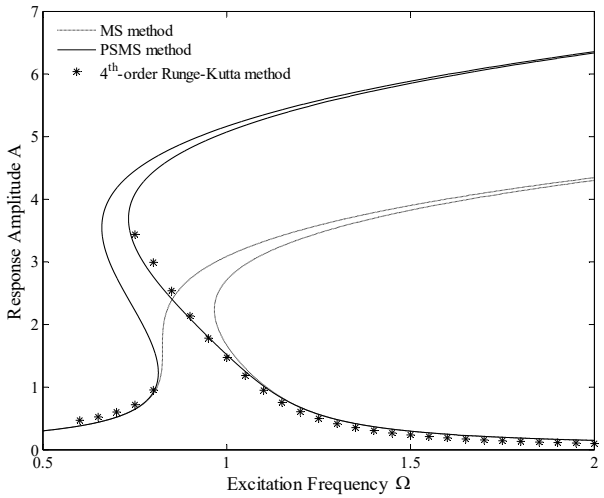


Figure 4. FRCs of oscillator 4 by the proposed method, the MS method and numerical simulation.

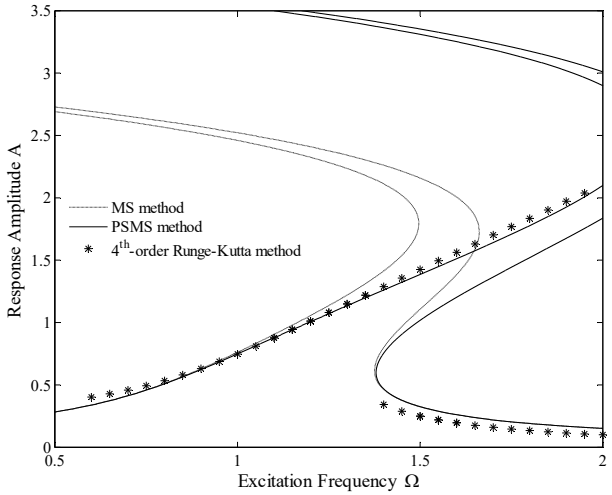


Figure 5. FRCs of oscillator 5 by the proposed method, the MS method and numerical simulation.

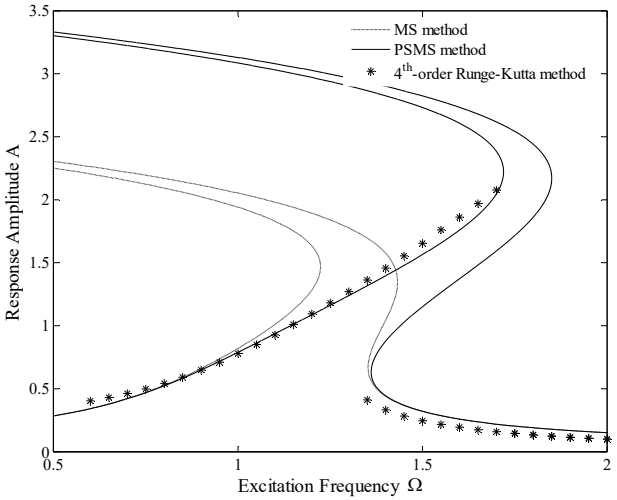


Figure 6. FRCs of oscillator 6 by the proposed method, the MS method and numerical simulation.

On boundary-value problems of elasticity theory with mixed boundary conditions

M. Kovalenko and †* I. Menshova

Laboratory of Geodynamics, Institute of Earthquake Prediction Theory and Mathematical Geophysics RAS, Russia

*Presenting author: menshovairina@yandex.ru
†Corresponding author: menshovairina@yandex.ru

Abstract

We consider a class of mixed boundary value problems of elasticity theory for the junction of two rectangular horizontal semi-strips of the same width with different boundary conditions on their long sides. On the junction of semi-strips, the continuity conditions of solutions or the discontinuity of displacements and stresses can be known.

Keywords: Mixed boundary value problems of elasticity theory; junction of two semi-strips

Introduction

In the general case, the solutions of such problems are represented as series in Papkovich–Fadle eigenfunctions, in particular as series in trigonometric functions. The unknown expansion coefficients are determined from the conditions on the junction of semi-strips. However, since two complete and minimal systems of functions (for example, trigonometric ones) take part in the solutions, their union on the junction will not be minimal. Therefore, it is impossible to construct a system of functions that is biorthogonal to this union. Hence, it is impossible to find a closed form solution. The main idea is to form a minimal system of functions, then to construct biorthogonal systems of functions and determine the unknown expansion coefficients with their help.

Statement of the problem and its solving

Let us consider the horizontal strip $\{\Pi : |y| \le 1, |x| < \infty\}$ with the following boundary conditions:

$$\begin{aligned} u(x,\pm 1) = 0,\ \sigma_y(x,\pm 1) = p^+(x),\ (x>0); \\ v(x,\pm 1) = \tau_{xy}(x,\pm 1) = 0,\ (x<0), \end{aligned} \tag{1}$$

where u and v are displacements along the x- and y-axes respectively.

Suppose that $p(x) \in L_2(-\infty,\infty)$ is a certain continuation of $p^+(x)$ to the whole real axis. We will assume that the null boundary functions are continued by zero. The solutions in the left semi-strip $\{\Pi^- : x \le 0, |y| \le 1\}$ and in the right semi-strip $\{\Pi^+ : x \ge 0, |y| \le 1\}$ can be represented in the form of Fourier series and integrals [1][2]. The corresponding formulas of the displacements and the stresses in the semi-strips $\Pi^{\mp}$ are as follows:

$$U^-(x,y) = \frac{1-\nu}{2} A_0 x - \frac{1+\nu}{2} \sum_{k=1}^{\infty} A_k \left(q_k^2 + B_k q_k + B_k q_k^2 x \right) e^{q_k x} q_k \cos(q_k y),$$

$$V^-(x,y) = \sum_{k=1}^{\infty} \left\{ \left(\frac{1+\nu}{2} A_k q_k^2 + 2 B_k q_k \right) + \frac{1+\nu}{2} B_k q_k^2 x \right\} e^{q_k x} \sin(q_k y),$$

$$\sigma_x^-(x,y)=A_0-\sum_{k=1}^{\infty}\left\{\left((1+\nu)A_kq_k^3+2B_kq_k^2\right)+(1+\nu)B_kq_k^3x\right\}e^{q_kx}\cos(q_ky),\tag{2}$$

$$\sigma_y^-(x,y)=\nu A_0+\sum_{k=1}^{\infty}\left\{\left((1+\nu)A_kq_k^3+2(2+\nu)B_kq_k^2\right)+(1+\nu)B_kq_k^3x\right\}e^{q_kx}\cos(q_ky),$$

$$\tau_{xy}^-(x,y)=\sum_{k=1}^{\infty}\left\{\left((1+\nu)A_kq_k^3+(3+\nu)B_kq_k^2\right)+(1+\nu)B_kq_k^3x\right\}e^{q_kx}\sin(q_ky)$$

and

$$U^+(x,y)=-\frac{1+\nu}{2}\sum_{k=1}^{\infty}(a_kp_k^2+b_kp_k+b_kp_k^2x)e^{p_kx}\cos(p_ky)+U^\circ(x,y),$$

$$V^+(x,y)=\sum_{k=1}^{\infty}\left(\frac{1+\nu}{2}a_kp_k^2+2b_kp_k+\frac{1+\nu}{2}b_kp_k^2x\right)e^{p_kx}\sin(p_ky)+V^\circ(x,y),$$

$$\sigma_x^+(x,y)=-\sum_{k=1}^{\infty}\left\{(1+\nu)a_kp_k^3+2b_kp_k^2+(1+\nu)b_kp_k^3x\right\}e^{p_kx}\cos(p_ky)+\sigma_x^\circ(x,y),\tag{3}$$

$$\sigma_y^+(x,y)=\sum_{k=1}^{\infty}\left\{(1+\nu)a_kp_k^3+2(2+\nu)b_kp_k^2+(1+\nu)b_kp_k^3x\right\}e^{p_kx}\cos(p_ky)+\sigma_y^\circ(x,y),$$

$$\tau_{xy}^+(x,y)=\sum_{k=1}^{\infty}\left\{(1+\nu)a_kp_k^3+(3+\nu)b_kp_k^2+(1+\nu)b_kp_k^3x\right\}e^{p_kx}\sin(p_ky)+\tau_{xy}^\circ(x,y).$$

In formulas (2) and (3) the following notations are introduced: G is the shear modulus, ν is the Poisson ratio, $q_k=k\pi$, $p_k=-(2k-1)\pi/2$, $U(x,y)=Gu(x,y)$ and $V(x,y)=Gv(x,y)$. The superscript $\circ$ indicates the quantities corresponding to the solution in terms of Fourier integrals for the infinite strip under the boundary conditions $U(x,\pm1)=0$, $\sigma_y(x,\pm1)=p(x)$. This solution can be found easily, for example:

$$U^\circ(x,y)=\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{1+\nu}{4}(\sin t\cos ty-y\cos t\sin ty)\frac{e^{tx}}{\cos^2 t}\Im[p](t)dt.\tag{4}$$

$\Im[p](t)$ is the Fourier transform of the function $p(x)$.

Let $p^+(x)=p=const$ and $p(x)=p$. In this case the solution written in terms of Fourier integrals has the form

$$U^\circ(x,y)=0,\ V^\circ(x,y)=\frac{1}{2}(1-\nu)py,\ \sigma_x^\circ(x.y)=p\nu,\ \tau_{xy}^\circ(x,y)=0.\tag{5}$$

The required coefficients A_k,B_k and a_k,b_k must be found from the continuity conditions of the displacements $U^\pm(0,y)$ and $V^\pm(0,y)$ and the stresses $\sigma_x^\pm(0,y)$ and $\tau_{xy}^\pm(0,y)$ on the junction of the semi-strips. As a result, we obtain four functional equations containing the four complete minimal systems of functions $\{1,\cos q_ky\}\cup\{\cos p_ky\}$ and $\{\sin q_ky\}\cup\{\sin p_ky\}$ $(k\ge1)$. However, these unions are not minimal. It follows that the unknown coefficients cannot exactly be found from these equations. We must first eliminate the unnecessary functions from these equations. In order to do that, we consider the two functions analytic in the semi-strips $\Pi^\pm$:

$$\Phi^{\pm}(x,y)=\frac{i}{2}\left(2(1+\nu)\frac{dV^{\pm}(x,y)}{dy}-(1-\nu)\sigma_x^{\pm}(x,y)\right)-\tau_{xy}^{\pm}(x,y),$$
$$\Psi^{\pm}(x,y)=(1+\nu)\frac{dU^{\pm}(x,y)}{dy}-\frac{3+\nu}{2}\tau_x^{\pm}(x,y)+i\left(\nu\sigma_x^{\pm}(x,y)+2(1+\nu)\frac{dV^{\pm}(x,y)}{dy}\right). \quad (6)$$

By expanding the equations $\Phi^{+}(0,y)=\Phi^{-}(0,y)$ and $\Psi^{+}(0,y)=\Psi^{-}(0,y)$, we obtain the system of two functional equations

$$-(1-\nu)\frac{A_0}{2}+\sum_{k=1}^{\infty}[(1+\nu)A_k q_k^3+(3+\nu)B_k q_k^2]e^{iq_k y}-$$
$$-\sum_{k=1}^{\infty}[(1+\nu)a_k p_k^3+(3+\nu)]b_k p_k^2 e^{iq_k y}=\frac{\Phi^{\circ}(y)}{i},$$
$$mA_0+\sum_{k=1}^{\infty}[(1+\nu)A_k q_k^3+(4+2\nu)B_k q_k^2]e^{iq_k y}- \quad (7)$$
$$-\sum_{k=1}^{\infty}[(1+\nu)a_k p_k^3+(4+2\nu)b_k p_k^2]e^{ip_k y}=\frac{\Psi^{\circ}(y)}{i}.$$

The functions $\Phi^{\circ}(y)$ and $\Psi^{\circ}(y)$ are determined according to (6) for the variables indicated by the degree superscript. In accordance with (5), we have

$$\Phi^{\circ}(y)=(1-\nu)pi/2,\ \Psi^{\circ}(y)=pi.$$

Now we introduce the new notations:

$$\omega_1=q_1,\ \omega_2=p_1,\omega_3=q_2,\quad \omega_4=p_2,...,$$
$$D_1=A_1,D_2=-a_1,D_3=A_3,D_4=-a_2,..., \quad (8)$$
$$C_1=B_1,C_2=-b_1,C_3=B_3,C_4=-b_2,....$$

Theorem 1. The function system $\{e^{i\omega_k y}\}_{k=1}^{\infty}$ is complete and minimal in $L_2(-\infty,\infty)$.

Theorem 2. There exist a unique function system $\{\psi_k(y)\}_{k=1}^{\infty}$ biorthogonal to $\{e^{i\omega_k y}\}_{k=1}^{\infty}$. These functions are given by the formulas

$$\psi_k(y)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{L(\omega)\exp(-i\omega y)d\omega}{(\omega-\omega_k)L'(\omega_k)}\quad(|y|\le 1) \quad (9)$$

We multiply Eqs. (7) by $\psi_k(y)$ and integrate the result over $[-1,1]$ for each $k\ge 1$, thus obtaining the system of two algebraic equations for the unknowns C_k and D_k. Solving it and writing the result in the original notations (8), we have

$$A_k=\frac{(1-\nu)p}{(1+\nu)p_k^4L'(q_k)},\ B_k=\frac{(1-\nu)p}{2q_k^3L'(q_k)},$$
$$a_k=\frac{-(1-\nu)p}{(1+\nu)p_k^4L'(p_k)},\ b_k=\frac{(1-\nu)p}{2p_k^3L'(p_k)}. \quad (10)$$

By introducing the function

$$\varphi(n)=\frac{(-1)^{n-1}(n-1)!}{\sqrt{\pi}\Gamma\left(n+\frac{1}{2}\right)},$$

we can write out the final formulas for the stresses:

$$\begin{aligned}
\frac{\sigma_x^-(x,y)}{p}&=\nu+\frac{x}{2}(1-\nu^2)\sum_{n=1}^{\infty}\frac{\exp(n\pi x)}{\varphi(n)}\cos(n\pi),\\
\frac{\sigma_y^-(x,y)}{p}&=\nu^2-(1-\nu^2)\sum_{n=1}^{\infty}\left\{\frac{1}{n\pi}+\frac{x}{2}\right\}\frac{\exp(n\pi x)}{\varphi(n)}\cos(n\pi y),\\
\frac{\tau_{xy}^-(x,y)}{p}&=-\frac{(1-\nu^2)}{2}\sum_{n=1}^{\infty}\left\{\frac{1}{n\pi}+x\right\}\frac{\exp(n\pi x)}{\varphi(n)}\sin(n\pi y),
\end{aligned}\tag{11}$$

$$\begin{aligned}
\frac{\sigma_x^+(x,y)}{p}&=\nu+\frac{x}{2}(1-\nu^2)\sum_{n=1}^{\infty}\frac{1}{\varphi(n)}\exp\left(-\frac{2n-1}{2}\pi x\right)\cos\frac{2n-1}{2}\pi y,\\
\frac{\sigma_y^+(x,y)}{p}&=1+(1-\nu^2)\sum_{n=1}^{\infty}\left\{\frac{2}{(2n-1)\pi}-\frac{x}{2}\right\}\frac{1}{\varphi(n)}\exp\left(-\frac{2n-1}{2}\pi x\right)\cos\frac{2n-1}{2}\pi y\\
\frac{\tau_{xy}^+(x,y)}{p}&=-\frac{(1-\nu^2)}{2}\sum_{n=1}^{\infty}\left\{\frac{2}{(2n-1)\pi}-x\right\}\frac{1}{\varphi(n)}\exp\left(-\frac{2n-1}{2}\pi x\right)\sin\frac{2n-1}{2}\pi y.
\end{aligned}\tag{12}$$

As an illustration, Fig. 1 shows the distribution of the stresses $\sigma_x^{\pm}(\pm 0.05, y)$.

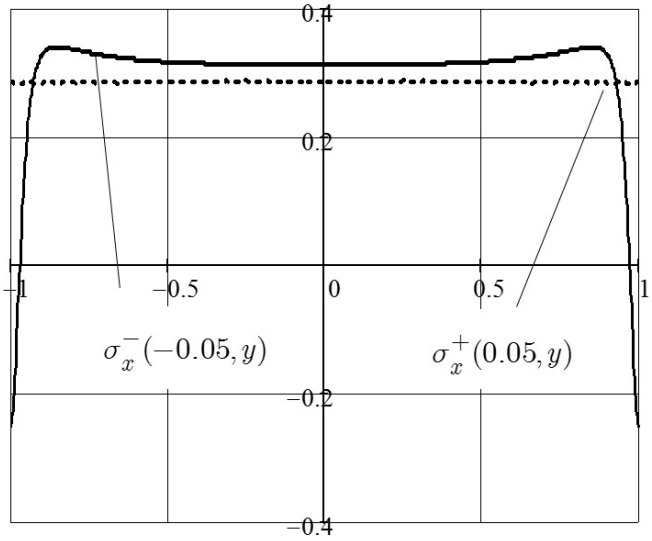


Figure 1. Distribution of the stresses $\sigma_x^{\pm}(\pm 0.05, y)$

Conclusions

It can be shown that the biorthogonal functions have a singularity of the type $\left(1\pm y\right)^{-1/2}$, therefore, the stresses at these points will also have the same singularity.

Acknowledgment

This research was supported by the Russian Foundation for Basic Research, research project no. 15-41-02644 r_povolzhye_a.

References

[1] Kovalenko, M. D., Menshova, I. V. and Kerzhaev, A. P. (2015) Method of initial functions and the integral Fourier transform in the boundary value problem of elasticity theory for an infinite strip (I), *Mechanics of Composite Materials and Structures* **21**(1), 47–64 [in Russian].

[2] Kovalenko, M. D., Menshova, I. V. and Kerzhaev, A. P. (2015) Method of initial functions and the integral Fourier transform in the boundary value problem of elasticity theory for an infinite strip (II), *Mechanics of Composite Materials and Structures* **21**(2), 237–250 [in Russian].

Generalized-strain – An Efficient Local Meshfree Method in Linear Elasticity

***Tiago Oliveira[1], Wilber Vélez[1], †Artur Portela[1]**

[1]Department of Civil and Environmental Engineering, University of Brasilia, Brazil

∗Presenting author: tiago_antuney@hotmail.com

†Corresponding author: aportela@unb.br

Abstract

This paper is concerned with the overall performance of the Generalized-strain Meshfree formulation, a new local meshless method, when compared to other meshless methods, for solving two-dimensional linear elastic problems.

Four methods are compared in this study, namely, the Generalized-Strain Mesh-free (GSMF) formulation, also known as the weak-form collocation meshless formulation; the Rigid-body Displacement Mesh-free (RBDMF) formulation, the Element-free Galerkin (EFG) and the Meshless Local Petrov-Galerkin Finite Volume Method (MLPG FVM). While the RBDMF, EFG and MLPG FVM rely on integration and quadrature process to obtain the stiffness matrix, the GSMF is completely integration-free, working as a weighted-residual weak-form collocation. This weak-form collocation readily overcomes the well-known difficulties of the strong-form collocation, such as low accuracy and instability of the solution.

A numerical example was analyzed with these methods, in order to assess the accuracy and the computational effort. The results obtained are in agreement with those of the available analytical solution. The numerical results show that the GSMF is superior not only regarding the computational efficiency, but also regarding the accuracy, when compared to the other methods.

Keywords: Local Meshless, Generalized-strain, Weak-form collocation, Element-free Galerkin, Meshless Local Petrov-Galerkin.

Introduction

The meshless methods or meshfree methods have intrinsic advantages over the element-based approaches, mostly due to the elimination of the mesh and the high-order continuity of the trial functions.

The main feature of these methods is that only a set of scattered nodes in the physical domain is required to approximate the solutions, and the nodes do not need to be connected to form closed polygons. In contrast with the finite element method, the meshless methods can save the preprocessing cost of mesh generation, as no element is required for the whole model [1]. In general, their formulation is based in the weighted-residual method [2].

Some meshless methods are based on a weighted-residual weak-form formulation. After discretization, the weak form is used to derive a system of algebraic equations through a process of numerical integration using sets of background cells, globally or locally constructed in the domain of the problem. Research on meshfree methods, based on a weighted-residual weak-form formulation, significantly increased after the publication of the Diffuse Element Method (DEM), introduced by [3]. The Reproducing Kernel Particle Method (RKPM), presented by [4], and the Element-free Galerkin (EFG) method, presented by [5], were the first weak-form meshless methods applied in solid mechanics.

All these weak-form meshless methods rely on background cells for the integration of the weighted-residual weak form over the global domain, in the process of the generation of the system of algebraic equations and therefore, they are not truly meshless methods.

To avoid the general background mesh generation, a class of meshfree methods based on local weighted-residual weak forms, such as the Meshless Local Petrov–Galerkin (MLPG) method [6, 7], the Meshless Local Boundary Integral Equation (MLBIE) method [8], the Local Point Interpolation Method (LPIM) [9] and the Local Radial Point Interpolation Method (LRPIM) [10], have been developed. The most popular of these methods is the MLPG, based on a moving least-squares (MLS) approximation. The main difference of the MLPG method to other global meshless methods, such as EFG or RKPM, is that local weak forms are used for integration on overlapping regular-shaped local subdomains, instead of global weak forms and consequently the method does not require the use of a background global mesh, but only a background local grid, which usually has a simple shape.

An implementation of the meshless Finite Volume Method (FVM) through the MLPG mixed approach was presented in [11] for solving elasto-static problems. In this approach, both the strains and displacements are independently interpolated, at randomly distributed points in the domain, through a local meshless interpolation schemes, in this case the MLS. Then, the nodal values of strains are expressed in terms of the interpolated nodal values of displacements, by simply enforcing the strain-displacement relationships directly by collocation at the nodal points. This formulation eliminates the expensive process of directly differentiating the MLS interpolations for displacements in the entire domain to compute the strains, leading to a high computational efficiency.

In order to further improve the computational efficiency, two formulations were presented by [12], the Rigid-body Displacement Mesh-free (RBDMF) formulation and the Generalized-Strain Mesh-free (GSMF) formulation. In the first formulation, the local work theorem leads to a weak form that is a regular local boundary integral equation. In the second formulation, the local work theorem generates a weak form that is completely integration free, working as a weighted-residual weak-form collocation.

In the present paper a numerical comparison between the Generalized-strain Mesh-Free (GSMF) formulation and three other meshless methods: the RBDMF, the EFG and the MLPG FVM; is performed for the solution of two-dimensional problems in linear elasticity. The results obtained in this study shows that the GSMF performs better than the other meshless methods regarding both computational efficiency and accuracy, as can be seen in the numerical results. It is expected that the GSMF framework will be implemented in a variety of problems, including large deformations and fracture mechanics, in the very near future.

MLS Approximation

Let Ω be the domain of a body with boundary Γ and let $N = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N\} \in \Omega$ be a set of scattered nodal points that represents a meshless discretization, in which some of them are located on the boundary Γ, where Ω_s, represented as Ω_P, Ω_Q and Ω_R, is the local compact support of a

node $\mathbf{x}_i$, represented as $\mathbf{x}_P$, $\mathbf{x}_Q$ and $\mathbf{x}_R$; $\Omega_{\mathbf{x}}$ is the domain of definition of a sampling point $\mathbf{x}$ and Ω_q is the local weak-form domain or quadrature domain of a node $\mathbf{x}_i$, as represented in Fig. 1.

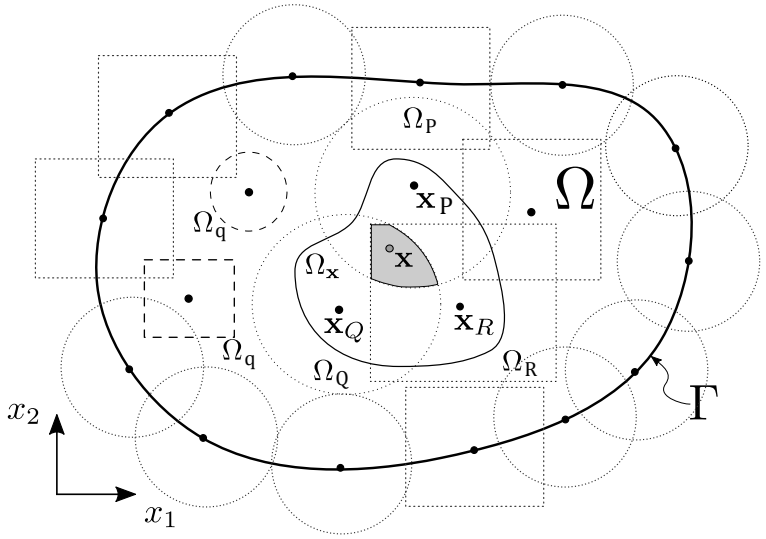


Figure 1. Representation of a global domain Ω and boundary Γ in a meshless discretization, with $\mathbf{x}_i$ nodes distributed within the body.

Circular or rectangular local supports, centered at each nodal point, can be used. In a neighborhood of a sampling point $\mathbf{x}$, the domain of definition of MLS approximation is the subdomain $\Omega_{\mathbf{x}}$, where the approximation is defined.

Shape Functions

Let $\Omega_{\mathbf{x}}$ be the domain of definition of the MLS approximation, in a neighbourhood of a sampling point $\mathbf{x}$. To approximate the displacement $u(\mathbf{x}) \in \Omega_{\mathbf{x}}$, over a number of scattered nodes $\mathbf{x}_i \in \Omega$, $i = 1, 2, \ldots, n$, where the nodal parameters $\hat{u}_i$ are defined, the MLS approximation is given by

$$u^h(\mathbf{x}) = \mathbf{p}^T(\mathbf{x})\mathbf{a}(\mathbf{x}), \tag{1}$$

for $\mathbf{x} \in \Omega_{\mathbf{x}}$, in which

$$\mathbf{p}^T(\mathbf{x}) = \left[p_1(\mathbf{x}), p_2(\mathbf{x}), \ldots, p_m(\mathbf{x})\right], \tag{2}$$

is a vector of the complete monomial basis of order m and $\mathbf{a}(\mathbf{x})$ is the vector of unknown coefficients $a_j(\mathbf{x})$, $j = 1, 2, \ldots, m$ that are functions of the space coordinates $\mathbf{x} = [x_1, x_2]^T$, for 2-D problems.

The coefficient vector $\mathbf{a}(\mathbf{x})$ is determined by minimizing the weighted discrete L_2 norm

$$J(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{n} w_i(\mathbf{x})\left[u^h(\mathbf{x_i}) - \hat{u}_i\right]^2 = \frac{1}{2}\sum_{i=1}^{n} w_i(\mathbf{x})\left[\mathbf{p}^T(\mathbf{x}_i)\mathbf{a}(\mathbf{x}) - \hat{u}_i\right]^2, \tag{3}$$

with respect to each term of $\mathbf{a}(\mathbf{x})$, in which $w_i(\mathbf{x})$ is the weight function associated with the node $\mathbf{x}_i$, with compact support that is $w_i(\mathbf{x}) > 0$, for all $\mathbf{x}$ in the support of $w_i(\mathbf{x})$. Figure 1 represents schematically the compact support of the MLS weight functions associated with a few nodes.

Finding the extremum of $J(\mathbf{x})$ with respect to each term of $\mathbf{a}(\mathbf{x})$, leads to

$$\mathbf{A}(\mathbf{x})\mathbf{a}(\mathbf{x}) = \mathbf{B}(\mathbf{x})\hat{\mathbf{u}}, \tag{4}$$

in which

$$\mathbf{A}(\mathbf{x}) = \sum_{i=1}^{n} w_i(\mathbf{x})\mathbf{p}(\mathbf{x}_i)\mathbf{p}^T(\mathbf{x}_i), \tag{5}$$

$$\mathbf{B}(\mathbf{x}) = [w_1(\mathbf{x})\mathbf{p}(\mathbf{x}_1), w_2(\mathbf{x})\mathbf{p}(\mathbf{x}_2), \ldots, w_n(\mathbf{x})\mathbf{p}(\mathbf{x}_n)] \tag{6}$$

and

$$\hat{\mathbf{u}} = [\hat{u}_1, \hat{u}_2, \ldots, \hat{u}_n]\,. \tag{7}$$

Solving Eq. (4) for $\mathbf{a}(\mathbf{x})$ yields

$$\mathbf{a}(\mathbf{x}) = \mathbf{A}^{-1}(\mathbf{x})\mathbf{B}(\mathbf{x})\hat{\mathbf{u}}, \tag{8}$$

provided $n \geq m$, for each sampling point $\mathbf{x}$, as a necessary condition for a well-defined MLS approximation. In the end, substituting for $\mathbf{a}(\mathbf{x})$ into Eq. (1) results in the MLS approximation

$$u^h(\mathbf{x}) = \sum_{i=1}^{n} \phi_i(\mathbf{x})\hat{u}_i, \tag{9}$$

in which

$$\phi_i(\mathbf{x}) = \sum_{j=1}^{m} p_j(\mathbf{x}) \left[\mathbf{A}^{-1}(\mathbf{x})\mathbf{B}(\mathbf{x})\right]_{ji} \tag{10}$$

is the shape function of the MLS approximation corresponding to the node $\mathbf{x}_i$, schematically represented in Fig. 2. The MLS shape functions are not nodal interpolants that is $\phi_i(\mathbf{x}_j) \neq \delta_{ij}$. The

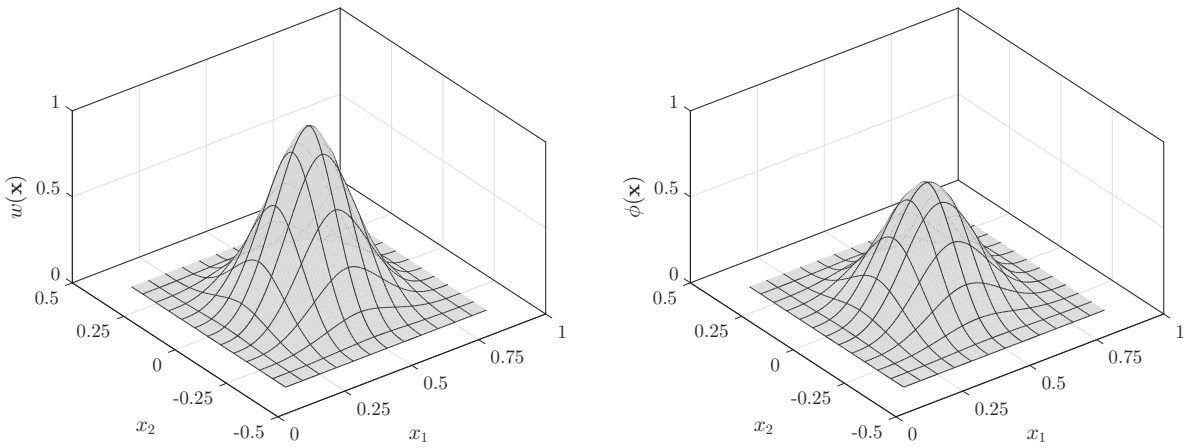


Figure 2. Respectively the typical weight function and shape function of the MLS approximation.

local character of the MLS approximation is preserved, since $\phi_i(\mathbf{x})$ vanishes for $\mathbf{x}$ not in the local domain of the node $\mathbf{x}_i$. The nodal shape function is complete up to the order of the basis. Also,

the smoothness of the nodal shape function is determined by the smoothness of the basis and of the weight function. The spatial derivatives of the shape function $\phi_i(\mathbf{x})$ are given by

$$\phi_{i,k} = \sum_{j=1}^{m} \left[p_{j,k}(\mathbf{A}^{-1}\mathbf{B})_{ji} + p_j(\mathbf{A}^{-1}\mathbf{B}_{,k} - \mathbf{A}^{-1}\mathbf{A}_{,k}\,\mathbf{A}^{-1}\mathbf{B})_{ji} \right], \tag{11}$$

in which $()_{,k} = \partial()/\partial x_k$.

Weight Functions

Weight functions $w_i(\mathbf{x})$, schematically represented in Fig. 2, firstly introduced in Eq. (3) for each node $\mathbf{x}_i$, have a compact support which defines the subdomain where $w_i(\mathbf{x}) > 0$, for all sampling point $\mathbf{x}$. For the sake of simplicity, this paper considers rectangular compact supports with weight functions defined as

$$w_i(\mathbf{x}) = w_{i_x}(\mathbf{x})\, w_{i_y}(\mathbf{x}) \tag{12}$$

with the weight function given by the quartic spline function

$$w_{i_x}(\mathbf{x}) = \begin{cases} 1 - 6\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^2 + 8\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^3 - 3\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^4 & \text{for } 0 \le d_{i_x} \le r_{i_x} \\ 0 & \text{for } d_{i_x} > r_{i_x} \end{cases} \tag{13}$$

and

$$w_{i_y}(\mathbf{x}) = \begin{cases} 1 - 6\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^2 + 8\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^3 - 3\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^4 & \text{for } 0 \le d_{i_y} \le r_{i_y} \\ 0 & \text{for } d_{i_y} > r_{i_y}, \end{cases} \tag{14}$$

in which $d_{i_x} = \|x - x_i\|$ and $d_{i_y} = \|y - y_i\|$. The parameters r_{i_x} and r_{i_y} represent the size of the support for the node i, respectively in the x and y directions.

Elastic Field

The elastic field is now approximated at a sampling point $\mathbf{x}$. Considering Eq. (9), displacement and strain components are respectively approximated as

$$\mathbf{u} = \begin{bmatrix} u^h(\mathbf{x}) \\ v^h(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \phi_1(\mathbf{x}) & 0 & \dots & \phi_n(\mathbf{x}) & 0 \\ 0 & \phi_1(\mathbf{x}) & \dots & 0 & \phi_n(\mathbf{x}) \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{v}_1 \\ \vdots \\ \hat{u}_n \\ \hat{v}_n \end{bmatrix} = \Phi\, \hat{\mathbf{u}} \tag{15}$$

and

$$\boldsymbol{\varepsilon} = \mathbf{L}\,\mathbf{u} = \mathbf{L}\,\Phi\,\hat{\mathbf{u}} = \mathbf{B}\,\hat{\mathbf{u}}, \tag{16}$$

in which geometrical linearity is assumed in the differential operator $\mathbf{L}$ and thus,

$$\mathbf{B} = \begin{bmatrix} \phi_{1,1} & 0 & \dots & \phi_{n,1} & 0 \\ 0 & \phi_{1,2} & \dots & 0 & \phi_{n,2} \\ \phi_{1,2} & \phi_{1,1} & \dots & \phi_{n,2} & \phi_{n,1} \end{bmatrix}. \tag{17}$$

Stress and traction components are respectively approximated as

$$\boldsymbol{\sigma} = \mathbf{D}\,\boldsymbol{\varepsilon} = \mathbf{D}\,\mathbf{B}\,\hat{\mathbf{u}} \tag{18}$$

and

$$\mathbf{t} = \mathbf{n}\,\boldsymbol{\sigma} = \mathbf{n}\,\mathbf{D}\,\mathbf{B}\,\hat{\mathbf{u}}, \tag{19}$$

in which $\mathbf{D}$ is the matrix of the elastic constants and $\mathbf{n}$ is the matrix of the components of the unit outward normal, defined as

$$\mathbf{n} = \begin{bmatrix} n_1 & 0 & n_2 \\ 0 & n_2 & n_1 \end{bmatrix}. \tag{20}$$

Equations (15) to (19) show that, at a sampling point $\mathbf{x} \in \Omega_{\mathbf{x}}$, the variables of the elastic field are defined in terms of the nodal unknowns $\hat{\mathbf{u}}$.

Local Form of the Work Theorem

This section present the development of the local form of the work theorem, first introduced in [12].

Let Ω be the domain of a body and Γ its boundary, subdivided in Γ_u and Γ_t that is $\Gamma = \Gamma_u \cup \Gamma_t$; nodal points P, Q and R have corresponding local domains Ω_P, Ω_Q and Ω_R, as represented in Fig. 3. The mixed fundamental boundary value problem of linear elastostatics aims to determine the distribution of stresses $\boldsymbol{\sigma}$, strains $\boldsymbol{\varepsilon}$ and displacements $\mathbf{u}$ throughout the body, when it has constrained displacements $\overline{\mathbf{u}}$ defined on Γ_u and is loaded by an external system of distributed surface and body forces with densities denoted by $\overline{\mathbf{t}}$ on Γ_t and $\mathbf{b}$ in Ω, respectively.

A totally admissible elastic field is the solution of the posed problem that simultaneously satisfies the kinematic admissibility and the static admissibility. If this solution exists, it can be shown that it is unique, provided linearity and stability of the material are admitted [13, 14].

The general work theorem establishes an energy relationship between any statically-admissible stress field and any kinematically-admissible strain field that can be defined in the body. Derived as a weighted residual statement, the work theorem serves as a unifying basis for the formulation of numerical models in Continuum Mechanics [15].

In the domain of the body, consider a statically-admissible stress field that is

$$\mathbf{L}^T\boldsymbol{\sigma} + \mathbf{b} = \mathbf{0}, \tag{21}$$

in the domain Ω, with boundary conditions

$$\mathbf{t} = \mathbf{n}\,\boldsymbol{\sigma} = \overline{\mathbf{t}}, \tag{22}$$

on the static boundary Γ_t, in which the vector $\boldsymbol{\sigma}$ represents the stress components; $\mathbf{L}$ is a matrix differential operator; the vector $\mathbf{t}$ represent the traction components; $\overline{\mathbf{t}}$ represent prescribed values of tractions and $\mathbf{n}$ represents the outward unit normal components to the boundary.

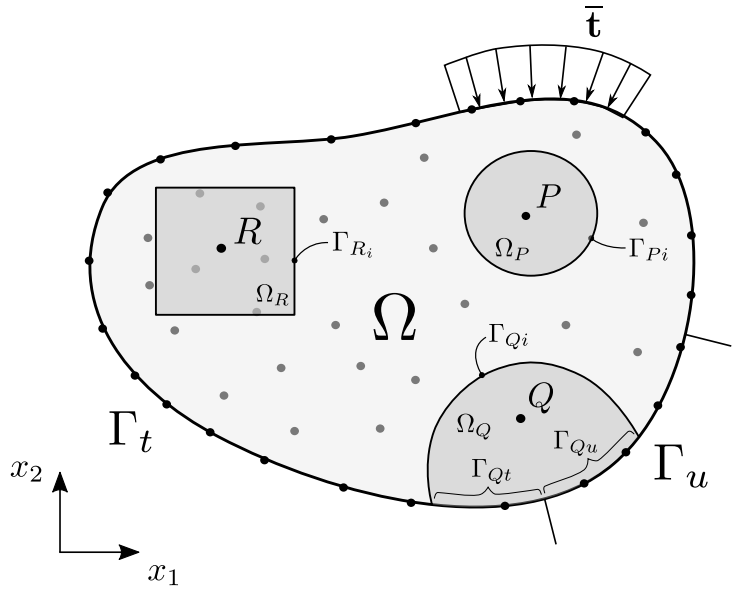


Figure 3. Meshless discretization of the global domain Ω and the local domains Ω_P, Ω_Q and Ω_R, with boundary $\Gamma = \Gamma_u \cup \Gamma_t$ represented.

In the global domain Ω, consider an arbitrary local subdomain Ω_Q, centered at the point Q, with boundary $\Gamma_Q = \Gamma_{Qi} \cup \Gamma_{Qt} \cup \Gamma_{Qu}$, in which Γ_{Qi} is the interior local boundary, while Γ_{Qt} and Γ_{Qu} are local boundaries that respectively share a global boundary, as represented in Fig. 3. Due to its arbitrariness, this local domain can be overlapping with other similar subdomains. For the local domain Ω_Q, the strong form of the weighted-residual equation is written as

$$\int_{\Omega_Q} \left(\mathbf{L}^T\boldsymbol{\sigma} + \mathbf{b}\right)^T \mathbf{W}_\Omega \, \mathrm{d}\Omega + \int_{\Gamma_{Qt}} \left(\mathbf{t} - \overline{\mathbf{t}}\right)^T \mathbf{W}_\Gamma \, \mathrm{d}\Gamma = \mathbf{0}, \tag{23}$$

in which $\mathbf{W}_\Omega$ and $\mathbf{W}_\Gamma$ are arbitrary weighting functions defined, respectively in Ω and on Γ. When the domain term of Eq. (23) is integrated by parts, the following local weak form of the weighted residual equation is obtained

$$\int_{\Gamma_Q} (\mathbf{n}\boldsymbol{\sigma})^T \mathbf{W}_\Omega \, \mathrm{d}\Gamma - \int_{\Omega_Q} \left(\boldsymbol{\sigma}^T \mathbf{L}\mathbf{W}_\Omega - \mathbf{b}^T\mathbf{W}_\Omega\right) \, \mathrm{d}\Omega + \int_{\Gamma_{Qt}} \left(\mathbf{t} - \overline{\mathbf{t}}\right)^T \mathbf{W}_\Gamma \, \mathrm{d}\Gamma = \mathbf{0} \tag{24}$$

which now requires continuity of $\mathbf{W}_\Omega$, as an admissibility condition for integrability. For the sake of convenience, the arbitrary weighting function $\mathbf{W}_\Gamma$ is chosen as

$$\mathbf{W}_\Gamma = -\mathbf{W}_\Omega, \tag{25}$$

on the boundary Γ_{Qt}. Thus, Eq. (24) leads to

$$\int_{\Gamma_Q - \Gamma_{Qt}} \mathbf{t}^T \mathbf{W}_\Omega \, \mathrm{d}\Gamma + \int_{\Gamma_{Qt}} \overline{\mathbf{t}}^T \mathbf{W}_\Omega \, \mathrm{d}\Gamma - \int_{\Omega_Q} \left(\boldsymbol{\sigma}^T \mathbf{L}\mathbf{W}_\Omega - \mathbf{b}^T\mathbf{W}_\Omega\right) \, \mathrm{d}\Omega = \mathbf{0}. \tag{26}$$

Consider further an arbitrary kinematically-admissible strain field $\boldsymbol{\varepsilon}^*$, with continuous displacements $\mathbf{u}^*$ and small derivatives, in order to assume geometrical linearity, defined in the global domain that is

$$\boldsymbol{\varepsilon}^* = \mathbf{L}\,\mathbf{u}^*, \tag{27}$$

in the domain Ω, with boundary conditions

$$\mathbf{u}^* = \overline{\mathbf{u}}, \tag{28}$$

on the kinematic boundary Γ_u.

When the continuous arbitrary weighting function $\mathbf{W}_\Omega$, is defined as

$$\mathbf{W}_\Omega = \mathbf{u}^*, \tag{29}$$

the weak form (26), of the weighted residual equation, becomes

$$\int\limits_{\Gamma_Q-\Gamma_{Qt}-\Gamma_{Qu}} \mathbf{t}^T\mathbf{u}^*\,\mathrm{d}\Gamma + \int\limits_{\Gamma_{Qu}} \mathbf{t}^T\overline{\mathbf{u}}^*\,\mathrm{d}\Gamma + \int\limits_{\Gamma_{Qt}} \overline{\mathbf{t}}^T\mathbf{u}^*\,\mathrm{d}\Gamma - \int\limits_{\Omega_Q} \left(\boldsymbol{\sigma}^T\,\mathbf{L}\mathbf{u}^* - \mathbf{b}^T\mathbf{u}^*\right)\,\mathrm{d}\Omega = \mathbf{0} \tag{30}$$

which can be written in a compact form as

$$\int\limits_{\Gamma_Q} \mathbf{t}^T\mathbf{u}^*\,\mathrm{d}\Gamma + \int\limits_{\Omega_Q} \mathbf{b}^T\mathbf{u}^*\,\mathrm{d}\Omega = \int\limits_{\Omega_Q} \boldsymbol{\sigma}^T\boldsymbol{\varepsilon}^*\,\mathrm{d}\Omega. \tag{31}$$

This equation is the starting point of the kinematically admissible formulations of the local meshfree methods presented in this paper. Equation (31) which expresses the static-kinematic duality, is the local form of the well-known work theorem, the fundamental identity of solid mechanics [16].

It is important to notice that the stress field $\boldsymbol{\sigma}$, is any one that satisfies equilibrium with the applied external forces $\mathbf{b}$ and $\mathbf{t}$, which is not necessarily the stress field that actually settles in the body. Also, the strain field $\boldsymbol{\varepsilon}^*$, is any one that is compatible with the constraints $\mathbf{u}^* = \overline{\mathbf{u}}$, which is not necessarily the strain field that actually settles in the body. This two fields are not connected by any constitutive relationship; indeed, as a consequence of the arbitrariness of the weighting function $\mathbf{W}_\Omega$ they are completely independent. For that reason Eq. (31) can be used under the only assumption of geometrical linearity.

It is the independence of the two admissible fields of the Eq. (31) that allows the generation of different meshfree methods, when the strain field is locally defined through different options, as carried out in this paper.

A final important remark, worth of mentioning, is that the local domain Ω_Q, is any arbitrary subdomain of the global domain Ω, of the body.

Modeling Strategy

Different formulations of local meshfree methods can be derived when the arbitrary kinematically-admissible field $\boldsymbol{\varepsilon}^*$, is locally defined in the work theorem, Eq. (31). In the following section, simple kinematically-admissible local fields will be used to derive the meshless formulation presented in this paper, the Generalized-Strain Mesh-Free (GSMF) formulation.

On the other hand, the statically-admissible local field $\boldsymbol{\sigma}$, will be always assumed as the elastic field that actually settles in the body. Not only satisfying static admissibility, through Eq. (21) and (22), but also satisfying kinematic admissibility in this elastic field defined as

$$\boldsymbol{\varepsilon} = \mathbf{L}\,\mathbf{u}, \tag{32}$$

in the domain Ω, with boundary conditions

$$\mathbf{u} = \overline{\mathbf{u}}, \tag{33}$$

on the kinematic boundary Γ_u; in which the displacements $\mathbf{u}$, are assumed continuous with small derivatives, in order to allow for geometrical linearity of the strain field $\boldsymbol{\varepsilon}$. Therefore, Eq. (33) must be enforced in the numerical model, in order to provide a unique solution of the posed problem.

For a meshless discretization of the body, the local weak-form domain or quadrature domain Ω_Q, centered at a node Q, can be defined in this paper as a rectangular or circular subdomain, as represented in Fig. 3.

Generalized-Strain Formulation

This section briefly discuss the development of the Generalized-Strain Mesh-free (GSMF) formulation. For the complete and detailed development see [12].

In the local form of the work theorem, Eq. (31), the kinematically-admissible displacement field $\mathbf{u}^*$, was assumed as a continuous function leading to a regular integrable function that is the kinematically-admissible strain field $\boldsymbol{\varepsilon}^*$. However, this continuity assumption on $\mathbf{u}^*$, enforced in the local form of the work theorem, is not absolutely required but can be relaxed by convenience, provided $\boldsymbol{\varepsilon}^*$ can be useful as a generalized function, in the sense of the theory of distributions [17]. Hence, this formulation considers that the kinematically-admissible displacement field is a piecewise continuous function, defined in terms of the Heaviside step function and therefore the corresponding kinematically-admissible strain field is a generalized function, defined in terms of the Dirac delta function.

For the sake of the simplicity, in dealing with Heaviside and Dirac delta functions in a two-dimensional coordinate space, consider a scalar function d, defined as

$$d = \| \mathbf{x} - \mathbf{x}_Q \| \quad \text{that is} \quad \begin{cases} d = 0 & \text{if } \mathbf{x} \equiv \mathbf{x}_Q \\ d > 0 & \text{if } \mathbf{x} \neq \mathbf{x}_Q, \end{cases} \tag{34}$$

which represents the absolute-value function of the distance between a field point $\mathbf{x}$ and a particular reference point $\mathbf{x}_Q$, in the local domain $\Omega_Q \cup \Gamma_Q$ assigned to the field node Q. Therefore, this definition always assumes $d = d(\mathbf{x}, \mathbf{x}_Q) \geq 0$, as a positive or null value, in this case whenever $\mathbf{x}$ and $\mathbf{x}_Q$ are coincident points. It is important to remark that, in Eq. (34), neither the field point $\mathbf{x}$ nor the reference point $\mathbf{x}_Q$ is necessarily a nodal point of the local domain.

For a scalar coordinate $d \supset d(\mathbf{x}, \mathbf{x}_Q)$, the Heaviside step function can be defined as

$$H(d) = \begin{cases} 1 & \text{if } d \leq 0 \ (d = 0 \text{ for } \mathbf{x} \equiv \mathbf{x}_Q), \\ 0 & \text{if } d > 0 \text{ that is } \mathbf{x} \neq \mathbf{x}_Q, \end{cases} \tag{35}$$

in which the discontinuity is assumed at $\mathbf{x}_Q$ and consequently, the Dirac delta function is defined with the following properties

$$\delta(d) = H'(d) = \begin{cases} \infty & \text{if } d = 0 \text{ that is } \mathbf{x} \equiv \mathbf{x}_Q, \\ 0 & \text{if } d \neq 0 \ (d > 0 \text{ for } \mathbf{x} \neq \mathbf{x}_Q) \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(d)\, \mathrm{d}d = 1, \tag{36}$$

in which $H'(d)$ represents the distributional derivative of $H(d)$. Note that the derivative of $H(d)$, with respect to the coordinate x_i, can be defined as

$$H(d)_{,i} = H'(d)\ d_{,i} = \delta(d)\ d_{,i} = \delta(d)\ n_i. \tag{37}$$

Since the result of this equation is not affected by any particular value of the constant n_i, this constant will be conveniently redefined later on.

Now Consider that d_l, d_j and d_k represent the distance function d, defined in Eq. (34), for corresponding collocation points $\mathbf{x}_l$, $\mathbf{x}_j$ and $\mathbf{x}_k$. Then, when Eq. (34) to (36) are considered, the displacement field $\mathbf{u}^*(\mathbf{x})$, can be conveniently defined as

$$\mathbf{u}^*(\mathbf{x}) = \left[\frac{L_i}{n_i} \sum_{l=1}^{n_i} H(d_l) + \frac{L_t}{n_t} \sum_{j=1}^{n_t} H(d_j) + \frac{S}{n_\Omega} \sum_{k=1}^{n_\Omega} H(d_k) \right] \mathbf{e}, \tag{38}$$

in which $\mathbf{e} = [1 \ \ 1]^T$ represents the metric of the orthogonal directions and n_i, n_t and n_Ω represent the number of collocation points, respectively on the local interior boundary $\Gamma_{Qi} = \Gamma_Q - \Gamma_{Qt} - \Gamma_{Qu}$ with length L_i, on the local static boundary Γ_{Qt} with length L_t and in the local domain Ω_Q with area S. This assumed displacement field $\mathbf{u}^*(\mathbf{x})$, a discrete rigid-body unit displacement defined at collocation points, is schematically represented in Fig. 4.

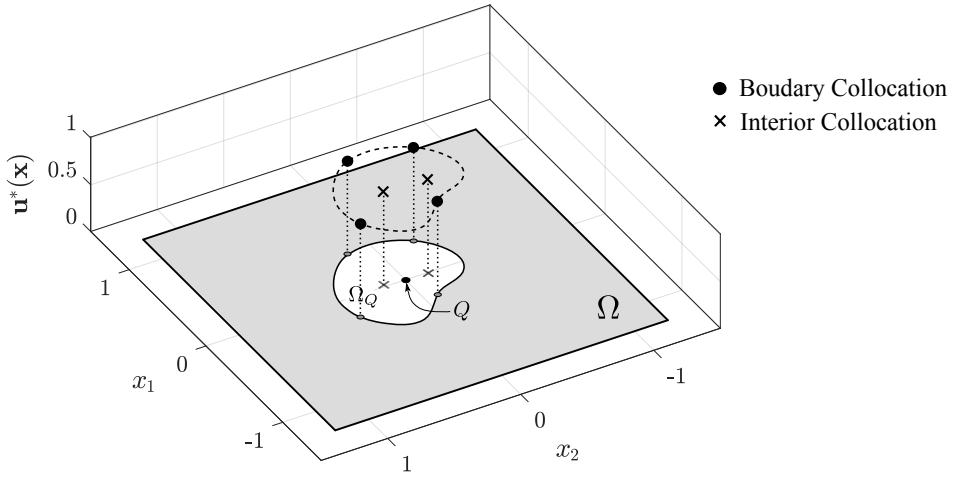


Figure 4. Schematic representation of the displacement $\mathbf{u}^*(\mathbf{x})$ of Eq. (38), a discrete rigid-body unit displacement defined at collocation points, of the Generalized-Strain Mesh-free formulation, for a local domain associated with a field node Q.

Therefore, when Eq. (37) are taken into account, the strain field $\boldsymbol{\varepsilon}^*(\mathbf{x})$, is given by

$$\begin{aligned}\boldsymbol{\varepsilon}^*(\mathbf{x}) = \mathbf{L}\,\mathbf{u}^*(\mathbf{x}) = \left[\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{L}\,H(d_l) + \frac{L_t}{n_t}\sum_{j=1}^{n_t}\mathbf{L}\,H(d_j) + \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{L}\,H(d_k)\right]\mathbf{e} = \\ = \left[\frac{L_i}{n_i}\sum_{l=1}^{n_i}\delta(d_l)\,\mathbf{n}^T + \frac{L_t}{n_t}\sum_{j=1}^{n_t}\delta(d_j)\,\mathbf{n}^T + \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\delta(d_k)\,\mathbf{n}^T\right]\mathbf{e},\end{aligned} \tag{39}$$

in which $\mathbf{n}$ is given by Eq. (20), with arbitrary components n_i that will be defined later on.

Having defined the displacement and the strain components of the kinematically-admissible field, respectively with Eq. (38) and (39), the local work theorem, Eq. (31), can be written as

$$\int_{\Gamma_Q-\Gamma_{Qt}}\mathbf{t}^T\mathbf{u}^*\,\mathrm{d}\Gamma + \int_{\Gamma_{Qt}}\overline{\mathbf{t}}^T\mathbf{u}^*\,\mathrm{d}\Gamma + \int_{\Omega_Q}\mathbf{b}^T\mathbf{u}^*\,\mathrm{d}\Omega = \int_{\Omega_Q}\boldsymbol{\sigma}^T\boldsymbol{\varepsilon}^*\,\mathrm{d}\Omega \tag{40}$$

that is

$$\begin{aligned}\frac{L_i}{n_i}\sum_{l=1}^{n_i}\int_{\Gamma_Q-\Gamma_{Qt}}\mathbf{t}^T H(d_l)\mathbf{e}\,\mathrm{d}\Gamma + \frac{L_t}{n_t}\sum_{j=1}^{n_t}\int_{\Gamma_{Qt}}\overline{\mathbf{t}}^T H(d_j)\mathbf{e}\,\mathrm{d}\Gamma + \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\int_{\Omega_Q}\mathbf{b}^T H(d_k)\mathbf{e}\,\mathrm{d}\Omega = \\ = \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\int_{\Omega_Q}\boldsymbol{\sigma}^T\delta(d_k)\,\mathbf{n}^T\mathbf{e}\,\mathrm{d}\Omega.\end{aligned} \tag{41}$$

Taking into account the properties of the Heaviside step function, defined in Eq. (35), Eq. (41) simply leads to

$$\mathbf{e}^T\left[\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{t}_{\mathbf{x}_l} + \frac{L_t}{n_t}\sum_{j=1}^{n_t}\overline{\mathbf{t}}_{\mathbf{x}_j} + \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathbf{x}_k} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{n}\int_{\Omega_Q}\delta(d_k)\boldsymbol{\sigma}\,\mathrm{d}\Omega\right] = \mathbf{0} \tag{42}$$

which, after considering the selective properties of Dirac delta function, leads to

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{t}_{\mathbf{x}_l} - \frac{S}{n_\Omega}\mathbf{n}\sum_{k=1}^{n_\Omega}\boldsymbol{\sigma}_{\mathbf{x}_k} = -\frac{L_t}{n_t}\sum_{j=1}^{n_t}\overline{\mathbf{t}}_{\mathbf{x}_j} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathbf{x}_k}. \tag{43}$$

Finally, when the variable $\mathbf{n}$, given by Eq. (20), is arbitrarily defined with identically null components $n_i = 0$, as allowed by Eq. (37), the Eq. (43) leads to

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{t}_{\mathbf{x}_l} = -\frac{L_t}{n_t}\sum_{j=1}^{n_t}\overline{\mathbf{t}}_{\mathbf{x}_j} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathbf{x}_k}. \tag{44}$$

Equation (44) states the equilibrium of tractions and body forces, pointwisely defined at collocation points, as schematically represented in Fig. 5. It can be seen that this is the pointwise version of the Euler - Cauchy stress principle. This is the equation used in the Generalized-Strain Mesh-free

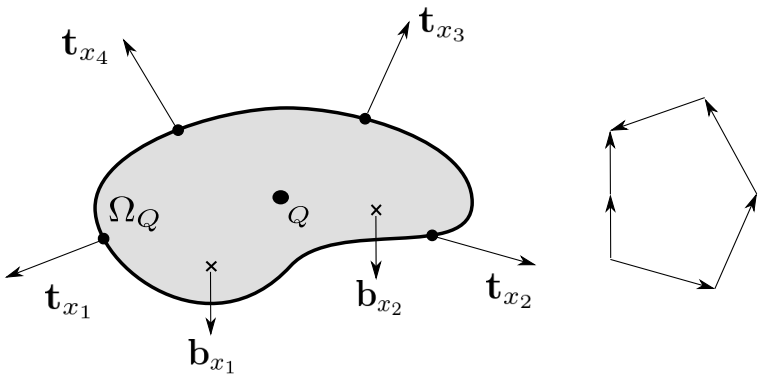


Figure 5. Schematic representation of the equilibrium of tractions and body forces of Eq. (44), pointwisely defined at collocation points of a local domain associated with a field node Q, of the Generalized-Strain Mesh-free formulation.

(GSMF) formulation which, therefore, is free of integration. Since the work theorem is a weighted-residual weak form, it can be easily seen that this integration-free formulation is nothing else other than a weighted-residual weak-form collocation.

Equations (44), of the Generalized-Strain Mesh-free formulation, can be derived from another kinematically-admissible displacement field, defined as a linear combination of Kronecker delta function evaluations at an arbitrary number of collocation points, conveniently arranged in the local domain $\Omega_Q \cup \Gamma_Q$ of the field node Q, as see in [12].

Discretization of Eq. (44) is carried out with the MLS approximation, Eq. (15) to (19), for the local domain Ω_Q, in terms of the nodal unknowns $\hat{\mathbf{u}}$, thus leading to the system of two linear algebraic equations

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{n}_{\mathbf{x}_l}\mathbf{D}\mathbf{B}_{\mathbf{x}_l}\hat{\mathbf{u}} = -\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{\mathbf{x}_j} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathbf{x}_k} \tag{45}$$

that can be written as

$$\mathbf{K}_Q\,\hat{\mathbf{u}} = \mathbf{F}_Q, \tag{46}$$

in which $\mathbf{K}_Q$, the nodal stiffness matrix associated with the local domain Ω_Q, is a $2 \times 2n$ matrix given by

$$\mathbf{K}_Q = \frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{n}_{\mathbf{x}_l}\mathbf{D}\mathbf{B}_{\mathbf{x}_l} \tag{47}$$

and $\mathbf{F}_Q$ is the respective force vector given by

$$\mathbf{F}_Q = -\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{\mathbf{x}_j} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathbf{x}_k} \tag{48}$$

Consider that the problem has a total of N field nodes Q, each one associated with the respective local region Ω_Q. Assembling Eq. (46), for all M interior and static–boundary field nodes leads to the global system of $2M \times 2N$ equations

$$\mathbf{K}\,\hat{\mathbf{u}} = \mathbf{F}. \tag{49}$$

Finally, the remaining equations are obtained from the $N - M$ boundary field nodes on the kinematic boundary. For a field node on the kinematic boundary, a direct interpolation method is used to impose the kinematic boundary condition as

$$u_k^h(\mathbf{x}_j) = \sum_{i=1}^{n} \phi_i(\mathbf{x}_j)\hat{u}_{ik} = \overline{\mathbf{u}}_k, \tag{50}$$

or, in matrix form as

$$\mathbf{u}_k = \Phi_k\, \hat{\mathbf{u}} = \overline{\mathbf{u}}_k, \tag{51}$$

with $k = 1, 2$, where $\overline{\mathbf{u}}_k$ is the specified nodal displacement component. Equations (50) are directly assembled into the global system of equations (49).

Numerical Results

This section presents some numerical results comparing the Generalized-Strain Mesh-free (GSMF) formulation with the Rigid-Body Displacement Mesh-free (RBDMF) formulation, the Element-free Galerkin (EFG) and the Meshless Local Petrov–Galerkin Finite Volume Method (MLPG FVM). Also, the best values for α_s and α_q are investigated in order to obtain the best results possible with GSMF.

For a generic node i, the size of the local support Ω_s and the local domain of integration Ω_q are respectively given by

$$r_{\Omega_s} = \alpha_s\, c_i, \tag{52}$$

and

$$r_{\Omega_q} = \alpha_q\, c_i, \tag{53}$$

in which c_i represents the distance of the node i, to the nearest neighboring node; for the applications presented in this paper, $\alpha_s = 3.0 \sim 4.5$ and $\alpha_q = 0.5 \sim 0.6$ were used. Only local meshless methods like the RBDMF, the GSMF and the MLPG FVM use local domains of integration; the EFG use background cells for integration purpose.

Displacement and energy norms can be used for error estimation. These norms can be computed, respectively as

$$\|\mathbf{u}\| = \left[\int_{\Omega} \mathbf{u}^T\mathbf{u}\, \mathrm{d}\Omega\right]^{1/2} \tag{54}$$

and

$$\|\boldsymbol{\varepsilon}\| = \left[\frac{1}{2}\int_{\Omega} \boldsymbol{\varepsilon}^T\mathbf{D}\,\boldsymbol{\varepsilon}\, \mathrm{d}\Omega\right]^{1/2}. \tag{55}$$

The relative error for $\|\mathbf{u}\|$ and $\|\boldsymbol{\varepsilon}\|$ is given, respectively by

$$r_u = \frac{\|\mathbf{u}_{num} - \mathbf{u}_{exact}\|}{\|\mathbf{u}_{exact}\|} \tag{56}$$

and

$$r_\varepsilon = \frac{\|\boldsymbol{\varepsilon}_{num} - \boldsymbol{\varepsilon}_{exact}\|}{\|\boldsymbol{\varepsilon}_{exact}\|}. \tag{57}$$

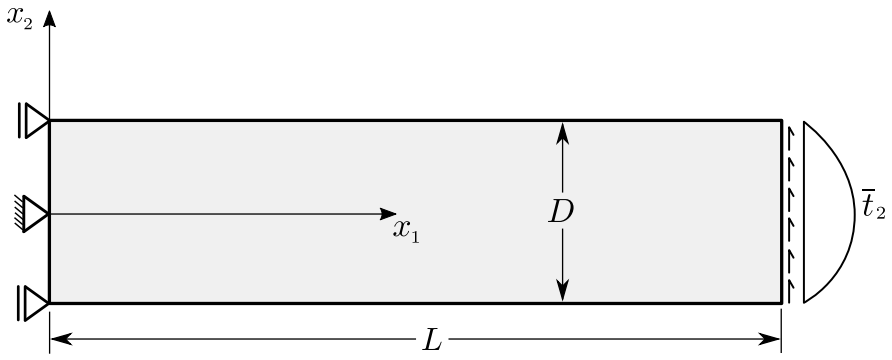


Figure 6. Timoshenko cantilever beam problem.

Now consider a beam of dimensions $L \times D$ and of unit depth, subjected to a parabolic traction at the free end as shown in Fig. 6. The beam is assumed in a plane stress state and the parabolic traction is given by

$$\bar{t}_2(x_2) = -\frac{P}{2I}\left(\frac{D^2}{4} - x_2^2\right), \tag{58}$$

where $I = D^3/12$ is the moment of inertia. The exact displacement components for this problem are given by

$$u_1(x_1, x_2) = -\frac{Px_2}{6EI}\left[(6L - 3x_1)x_1 + (2+\nu)\left(x_2^2 - \frac{D^2}{4}\right)\right] \tag{59}$$

and

$$u_2(x_1, x_2) = \frac{P}{6EI}\left[3\nu x_2^2(L - x_1) + (4+5\nu)\frac{D^2 x_1}{4} + (3L - x_1)x_1^2\right] \tag{60}$$

and the exact stress components are given by

$$\sigma_{11}(x_1, x_2) = -\frac{P(L - x_1)x_2}{I}, \quad \sigma_{22}(x_1, x_2) = 0. \tag{61}$$

and

$$\sigma_{12}(x_1, x_2) = -\frac{P}{2I}\left(\frac{D^2}{4} - x_2^2\right) \tag{62}$$

Material properties are taken as Young′s modulus $E = 3.0 \times 10^7$ and the Poisson′s ratio $\nu = 0.3$ and the beam dimensions are $D = 12$ and $L = 48$. The shear force is $P = 1000$.

Effects of the local support domain size on GSMF

The local support domain size is a very important meshless parameter, related to both accuracy and computational efficiency. Usually, the parameter α_s is greater than 1.0, to make sure that there are enough points to support the nodes on the global boundary. For a small size, the algorithm of MLS approximation may be singular and the shape function cannot be constructed, because there is not enough nodes for interpolating.

Three regular nodal distributions were considered with a discretization of $13 \times 4 = 52$ nodes, $33 \times 5 = 165$ nodes and $65 \times 9 = 585$ nodes. In the present study, 7 ratios are used for first-order polynomial basis in MLS approximation.

Figure 7 displays the variation of the energy error as a function of the size of the local support

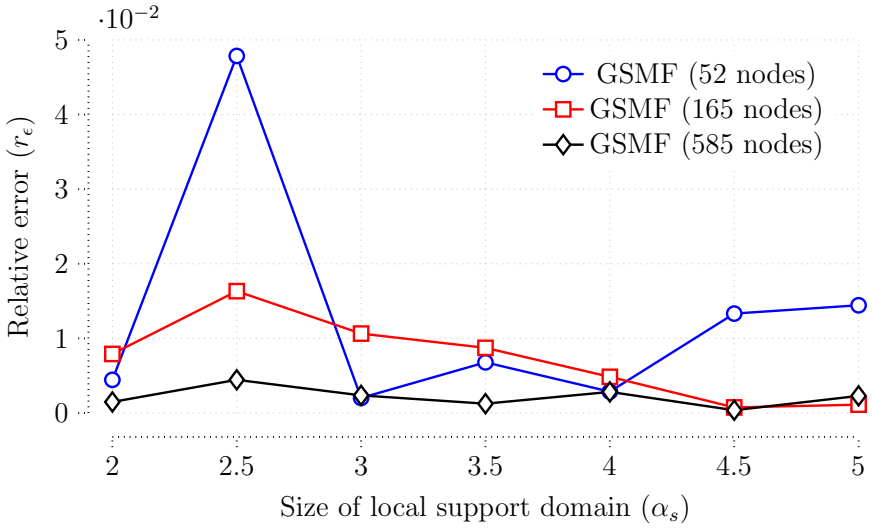


Figure 7. Influence of the local support domain size (α_s) for different nodal distributions.

domain, for fixed value of α_q, in this case, $\alpha_q = 0.5$. For the regular nodal distribution of 52 nodes, the best results are obtained when $3 \leq \alpha_s \leq 4$ and the most accurate result is obtained with $\alpha_s = 3$, leading to $r_\varepsilon = 2.02 \times 10^{-3}$. Now, for the nodal distribution of 165 and 585 nodes, the most accurate results are obtained with $\alpha_s = 4.5$, leading respectively to $r_\varepsilon = 7.51 \times 10^{-4}$ and $r_\varepsilon = 3.77 \times 10^{-4}$.

In general, for greater nodal distributions and first-order polynomial basis, $2.5 \leq \alpha_s \leq 5$ can be selected and the method is yet convergent.

Effects of the local quadrature/collocation domain size on GSMF

The weak-form domain or local quadrature/collocation domain is one of the key concepts for local meshless methods in general that is also related to both accuracy and computational efficiency. The parameter α_q is chosen to be less than 1.0 in the present study to ensure that the local sub-domains of the internal nodes are entirely within the solution domain, without being intersected by the global boundary.

The same three regular nodal distributions were considered (52, 165 and 585 nodes). In the present study, 5 ratios are used for first-order polynomial basis in MLS approximation.

Figure 8 displays the variation of the energy error as a function of the size of the local quadrature/collocation domain, for fixed value of α_s. For all nodal distributions the best results are obtained when $0.5 \leq \alpha_s \leq 0.7$ and the most accurate result is obtained with $\alpha_q = 0.5$, leading to $r_\varepsilon = 3.77 \times 10^{-4}$.

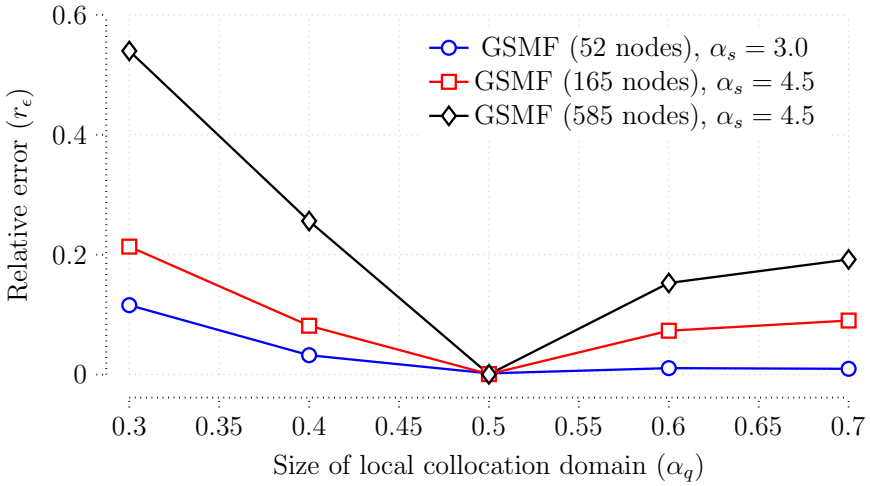


Figure 8. Influence of the local quadrature/collocation domain size (α_q) for different nodal distributions.

Displacement comparison

A initial regular nodal distribution was considered to solve the problem, with a discretization of $33 \times 5 = 165$ nodes, represented in Fig. 9.

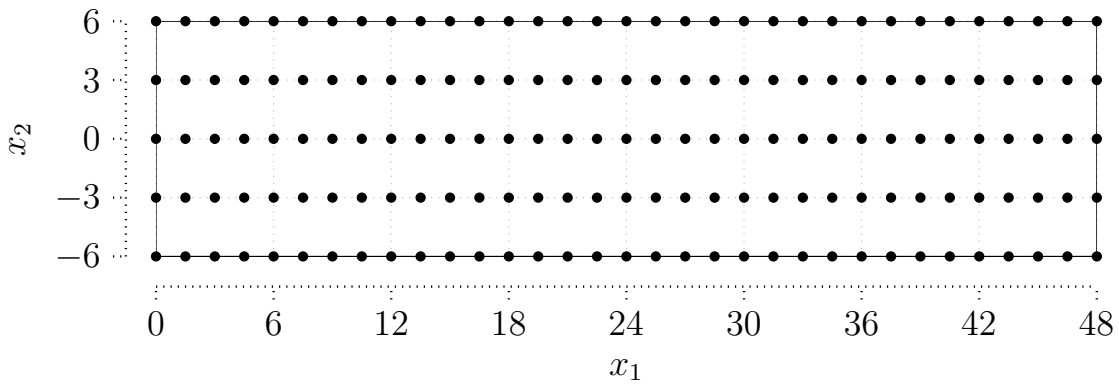


Figure 9. The regular nodal distribution of $33 \times 5 = 165$ nodes.

Rectangular local domains were considered for the local kinematic formulations, with 1 collocation point to compute the weak form of GSMF and 10 Gauss-quadrature points to integrate the weak-form of RBDMF, placed on each boundary of the local domain. The EFG considered 10 Gauss-quadrature points on each background cell and the MLPG FVM considered 10 Gauss-quadrature points distributed on the local domain. A first-order polynomial basis was considered in MLS approximation.

The displacements obtained with the four methods at x_1/L, represented in Fig. 10, show very good agreement with the results of the exact solution. The best results are obtained with GSMF, RBDMF and EFG, while the MLPG FVM got the least accurate results among them.

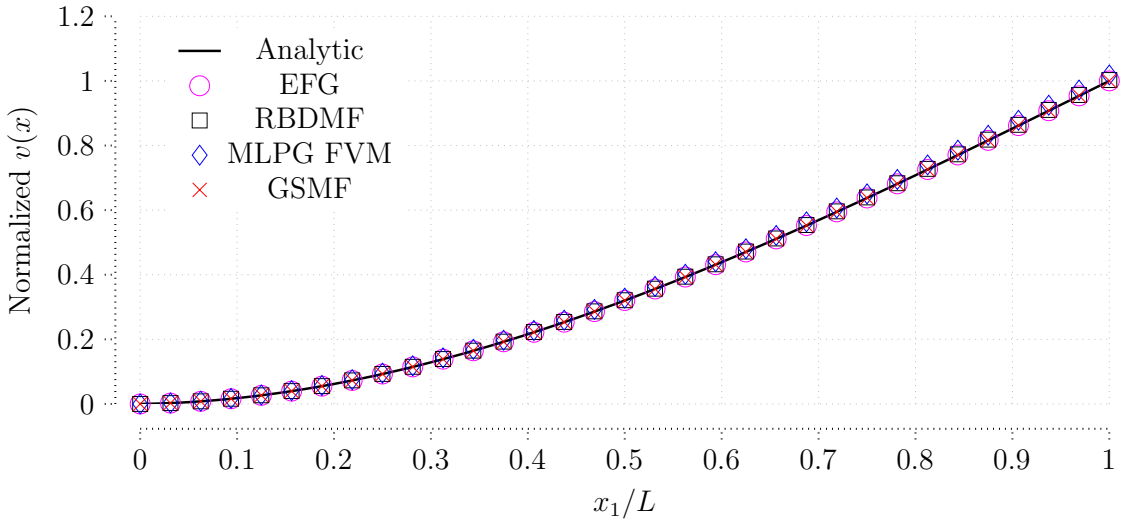


Figure 10. Normalized displacements of the cantilever-beam discretization with 165 nodes.

Computational efficiency comparison

The weak-form collocation of GSMF represents a clear reduction of the computational effort when compared to other meshless methods. The GSMF require only 1 collocation point, placed on each boundary of the local domain, to obtain the most accurate results, see [12]; while the other methods require at least 10 Gauss-quadrature points in order to obtain a good accuracy. This important feature is measure through CPU time consumption and convergence rates.

In order to further the study of the computational efficiency of the presented method, three additional regular nodal distributions with $65 \times 9 = 585$, $97 \times 13 = 1261$ and $129 \times 17 = 2193$ nodes of the cantilever-beam were considered. Only the major computational cost that is the cost of generating the global stiffness matrix and solving the system of algebraic equations, was measured. All the routines were compared when using MATLAB 2015a on an Intel Core I7-4700MQ computer with CPU of 2.4GHz and 16 GB of RAM.

The results obtained are presented in Fig. 11, where it can be seen that CPU time of GSMF is always much lower than CPU time of the other methods, when the same parameters are considered. The CPU time consumption of the GSMF is 3.62 times faster than the second best value that is the one obtained with the MLPG FVM. This important result clearly evidences the high computational efficiency of GSMF.

Accuracy and convergence comparison

Another test was performed to assess the accuracy and the convergence rate of the analyzed methods, using the relative energy norm. Since the MLPG FVM obtained the least accurate result among all methods, it was not compared in this test. The same three regular discretizations of the cantilever-beam were considered. Figure 12 presents the results obtained for the accuracy and convergence rates. As expected, the best results are obtained with the nodal distribution of 441 nodes, with a relative error of $r_\varepsilon = 6.38 \times 10^{-5}$ for the GSMF, $r_\varepsilon = 2.94 \times 10^{-4}$ for the RBDMF and $r_\varepsilon = 2.80 \times 10^{-3}$ for the EFG. A stable convergence rate is obtained with all tested methods.

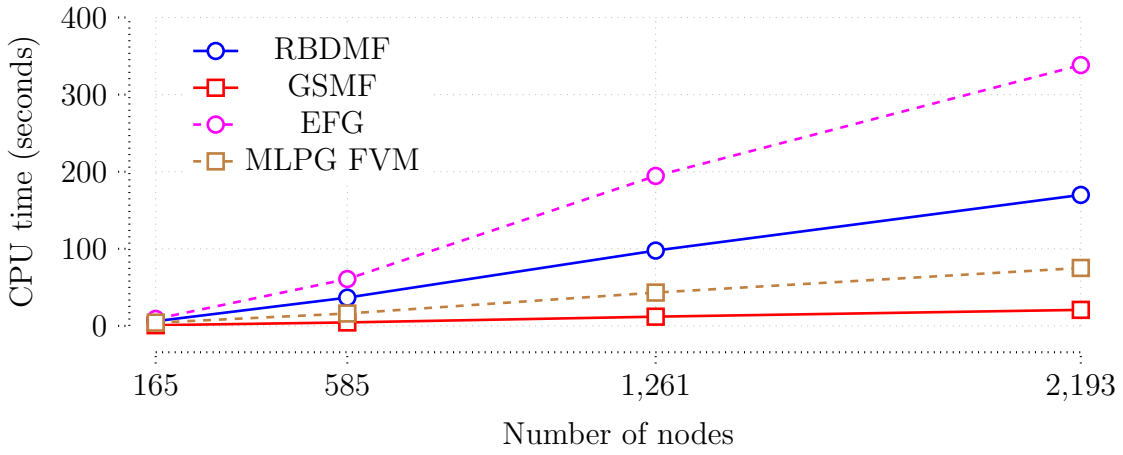
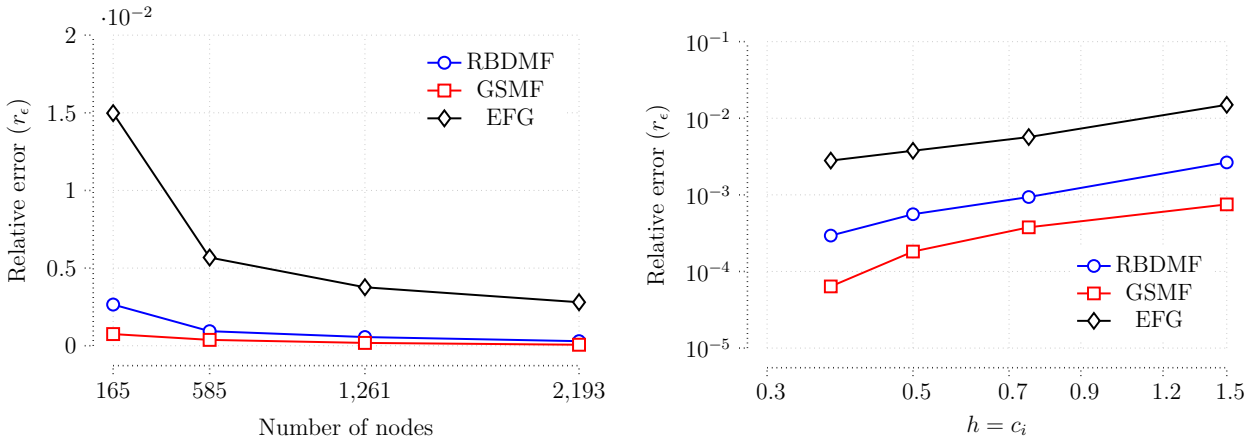


Figure 11. CPU time consumption of a cantilever-beam with 165, 585, 1261 and 2193.



(a) Accuracy.

(b) Convergence rates.

Figure 12. Accuracy and convergence rates for the cantilever-beam discretization with 165, 585, 1261 and 2193 nodes; c_i is the distance of a generic node i, to the nearest neighboring node, as defined in Eq. (52) and (53).

The results show that the GSMF is more accurate than the RBDMF and the EFG, with better convergence rates when compared to both of them.

Conclusions

A numerical comparison of the overall performance and efficiency of the Generalized-strain mesh-free formulation and three other meshless methods is performed, for for solving two-dimensional linear elastic problems.

While the Rigid-body Displacement Mesh-free (RBDMF) formulation, the Element-free Galerkin (EFG) and the Meshless Local Petrov-Galerkin Finite Volume Method (MLPG FVM) rely on integration and quadrature process to obtain the stiffness matrix of the posed problem; the Generalized-Strain Mesh-free (GSMF) formulation is completely integration free, working as a weighted-residual weak-form collocation.

A numerical example was analyzed with these methods, in order to compare the accuracy and the computational effort under the same parameters. The results obtained with all methods are in agreement with those of the available analytical solution. The MLPG FVM led to very fast computations, which are compromised by the low accuracy obtained. The EFG and the RBDMF obtained very accurate results with good convergence rates, although are computationally more expensive than the other methods. Among all methods, the GSMF obtained the most accurate results with the fastest computation.

All the numerical results obtained clearly demonstrate that this weighted-residual weak-form collocation readily overcomes the well-known difficulties posed by the weighted-residual strong-form collocation, regarding accuracy and stability of the solution. The results obtained using only 1 collocation point led to accurate results with incredible fast computations, surpassing all the other analyzed methods. This features make the GSMF superior when compared to the other meshless methods presented in this paper, making it an efficient local meshfree method for solving two-dimensional problems in linear elasticity.

Finally, it is expected that the GSMF framework will be implemented in a variety of problems in the very near future, specially for solving large deformations and fracture mechanics, where it is known that there are challenges in developing computationally efficient algorithms, with high accuracy that can overcome the issue of the computational cost.

Acknowledgments

The first author acknowledges the program *PECC – Pós-Graduação em Estruturas e Construção Civil*, Department of Civil and Environmental Engineering, Faculty of Technology, University of Brasília and *CNPq – Brazilian National Counsel of Technological and Scientific Development* for his PhD scholarship.

References

[1] Liu, G. R. (2009). Meshfree methods moving beyond the finite element method. CRC Press, 2nd edition edition.

[2] Finalyson, B. A. (1972). The Method of Weighted Residuals and Variational Principles. Academic Press.

[3] Nayroles, B., Touzot, G., and Villon, P. (1992). "Generalizing the Finite Element Method: Diffuse Approximation and Diffuse Elements." Computational Mechanics, 10:307–318.

[4] Liu, W. K., Jun, S., and Zhang, Y. F. (1995). "Reproducing Kernel Particle Methods." International Journal for Numerical Methods in Engineering, 20:1081–1106.

[5] Belytschko., T., Lu, Y. Y., and Gu, L. (1994). "Element-free Galerkin methods." International Journal for Numerical Methods in Engineering, 37(2):229–256. ISSN 1097-0207.

[6] Atluri, S. N. and Zhu, T. (1998). "A new Meshless Local Petrov-Galerkin (MLPG) approach in computational mechanics." Computational Mechanics, 22(2):117–127.

[7] Atluri, S. N. and Shen, S. (2002). "The Meshless Local Petrov-Galerkin (MLPG) Method: A Simple and Less-costly Alternative to the Finite Element and Boundary Element Methods." CMES: Computer Modeling in Engineering and Sciences, 3(1):11–51.

[8] Zhu, T., Zhang, J., and Atluri, S. N. (1998). "A Local Boundary Integral Equation (LBIE) Method in Computational Mechanics and a Meshless Discretization Approach." Computational Mechanics, 21:223–235.

[9] Liu, G. R. and Gu, Y. T. (2001). "A Local Point Interpolation Method for Stress Analysis of Two-Dimensional Solids." Structural Engineering and Mechanics, 11(2):221–236.

[10] Liu, G. R., Yan, L., Wang, J. G., and Gu, Y. T. (2002). "Point Interpolation Method Based on Local Residual Formulation Using Radial Basis Functions." Structural Engineering and Mechanics, 14:713–732.

[11] Atluri, S. N., Han, Z. D., and Rajendran, A. M. (2004). "A New Implementation of the Meshless Finite Volume Method Through the MLPG Mixed Approach." CMES: Computer Modeling in Engineering and Sciences, 6:491–513.

[12] Oliveira, T. and Portela, A. (2016). "Weak-Form Collocation – a Local Meshless Method in Linear Elasticity." Engineering Analysis with Boundary Elements, 73:144–160.

[13] Fredholm, I. (1906). "Solution d'un probléme fondamental de la théorie de l'élasticité." Arkiv för Matematik Astronomi och Fysk 2, 28(1):1–8.

[14] Fichera, G. (2006). Linear Elliptic Differential Systems and Eigenvalue Problems. Springer.

[15] Brebbia, C. A. and Tottenham, H., editors. Variational Basis of Approximate Models in Continuum Mechanics. Southampton and Springer Verlag.

[16] Sokolnikoff, I. S. (1956). Mathematical theory of elasticity. McGraw-Hill New York.

[17] Gelfand, I. M. and Shilov, G. E. (1964). Generalized Functions, volume 1. Academic Press.

Extrapolation Methods for Computing Supersingular Integral on a Circle

Xiaoping Zhang[1], and †Jin Li[1]*

[1]School of Science, Shandong Jianzhu University, Jinan 250101, P. R. China ;

*Presenting author: lijin@lsec.cc.ac.cn.
†Corresponding author: lijin@lsec.cc.ac.cn.

Abstract

The trapezoidal rule for the computation of supersingular integrals on circle is discussed, and the asymptotic expansion of error function is obtained. A series to approach the singular point is constructed. The extrapolation algorithm is presented and the convergence rate is proved. Some numerical results are also reported to confirm the theoretical results and show the efficiency of the algorithms.

Keywords: Supersingular integral; Extrapolation method; Composite trapezoidal rule; Posteriori estimate

1.Introduction

Accurate calculation of boundary element methods(BEM) arising in boundary integral equations has been a subject of intensive research in recent years. The formulation of certain classes of boundary value problems in terms of supersingular integral equations:

$$\oint_a^b \frac{f(x)}{(x-s)^{p+1}}dx = g(s) \quad s \in (a,b), p = 1,2 \tag{1}$$

have drawn lots of interests. In the literature different definitions of singular integrals are found which can be shown to be the same. We mention the following one

$$\oint_a^b \frac{f(x)}{(x-s)^{p+1}}dx = \lim_{\varepsilon\to 0}\left\{\left(\int_a^{s-\varepsilon}+\int_{s+\varepsilon}^b\right)\frac{f(x)}{(x-s)^{p+1}}dx - \frac{2f^{(p-1)}(s)}{\varepsilon}\right\},$$
$$s \in (a,b), \quad p = 1,2. \tag{2}$$

where $\oint_a^b$ denotes a supersingualr integral and s the singular point. Here the supersingular integral is one order higher singularity than hypersingular integral.

One of the major problems arising from boundary element methods is how to evaluate such supersingular integral efficiently. For the case $p = 1$, numerous work has been devoted to developing efficient quadrature formulas for hypersingular integral such as the Gaussian method [7,8], the Newton-Cotes method [15,22-25], the transformation method [3,5] and some other methods [4,10,19]Because of the high-order singularity of the kernels, the rules for Hadamard finite-part integrals (including hypersingular and supersingular integrals) are less accurate than their counterparts for Riemann integrals. Newton-Cotes rules for evaluating hypersingular integrals were firstly suggested by Linz [15]. The superconvergence phenomenon of trapezoidal rule and Simpson's rule for hypersingular integrals was found in [22,24], which showed that one order higher superconvergence rate than that in general case if the singular point is located at some a prior known point. Then, the superconvergence for arbitrary degree Newton-Cotes rules of hypersingular integrals were studied in [22,23] and the superconvergence rate was $O(h^{k+1})$. Recently, Newton-Cotes rules and the corresponding superconvergence for evaluating hypersingular integrals on a circle were discussed in [29,30]. Integrals with kernels beyond hypersingular have not been extensively studied with $p = 2$. Du [4] studied the composite Simpson's rule and showed the optimal global convergence rate is $O(h)$. Then, Wu & Sun [22] studied the superconvergence of trapezoidal rule, the $O(h^2)$ superconvergence rate was obtained when the singular point is located at the middle point of each subinterval far away from two endpoints. Recently, Zhang et.al [28] discussed the superconvergence phenomenon of the composite Simpson's rule and also the $O(h^2)$ rate was obtained for those superconvergence points far away from the endpoints.

In this paper, we consider the supersingular integral defined in a circle which have been paid less attention to it.

$$\begin{aligned} I(f,s) : &= \oint_c^{c+2\pi} \frac{\cos\frac{x-s}{2} f(x)}{\sin^3\frac{x-s}{2}} dx \\ &= \lim_{\epsilon\to 0}\left\{ \int_c^{s-\epsilon} \frac{f(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}} dx + \int_{s+\epsilon}^{c+2\pi} \frac{f(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}} dx - \frac{2\epsilon f'(s)}{\sin^2\frac{\epsilon}{2}} \right\}. \end{aligned} \tag{3}$$

To our knowledge, maybe the reference [12-14] have entire on the subject, where the superconvegence rate of Simpson rule and trapezoidal rule to compute the supersingular integral have been considered. When the singular point coincides with some priori known point, the convergence rate of the trapezoidal rule is higher than the global one which is considered as the superconvergence phenomenon. Then the error functional of density function is derived and the superconvergence phenomenon of composite trapezoidal rule occurs at certain local coordinate of each subinterval which is different from the case supersingular defined on the interval.

Extrapolation methods as an accelerating convergence technique has been applied to many fields in computational mathematics [16,21]. The most famous one is Richardson extrapolation with the error function as
$T(h) - a_0 = a_1 h^2 + a_2 h^4 + a_3 h^6 + \cdots$,
where $T(0) = a_0$ and a_j are constant independent of *h*. Then in the paper of Li [11] et.al, the trapezoidal rule for computation hypersingular integral on interval by extrapolation methods was given.

Before presented the extrapolation methods to compute the supersingular integral in a circle, we firstly give notation as below. Let

$$\psi_{ik}(t) = \begin{cases} \oint_{-1}^{1} \dfrac{\mathcal{M}_{ik}(\tau,t)\cos\frac{\tau-t}{2}}{(k-i)!\sin^3\frac{\tau-t}{2}} d\tau, & |t|<1, \\ \displaystyle\int_{-1}^{1} \frac{\mathcal{M}_{ik}(\tau,t)\cos\frac{\tau-t}{2}}{(k-i)!\sin^3\frac{\tau-t}{2}} d\tau, & |t|>1, \end{cases} \tag{4}$$

where

$$\mathcal{M}_{ik}(\tau,t) = (\tau^2-1)(\tau-t)^{k-i}[(\tau+1)^i - (\tau-1)^i]. \tag{5}$$

Let $J := (-\infty,-1)\bigcup(-1,1)\bigcup(1,\infty)$ and the operator $W : C(J) \to C(-1,1)$ be defined by

$$T_{ik}(\tau) := \psi_{ik}(\tau) + \sum_{j=1}^{\infty} \psi_{ik}(2j+\tau) + \sum_{j=1}^{\infty} \psi_{ik}(-2j+\tau), \quad k = i, i+1. \tag{6}$$

In this paper, based on the asymptotic error expansion of the composite trapezoidal rule for the computation of supersingular integrals. We firstly obtain the asymptotic error expansion as follow

$$E_n(f,s) = \sum_{i=1}^{l-2} \frac{h^{i-1}}{2^{i-1}} f^{(i+1)}(s) a_i(\tau) + O(h^{l-2}), \tag{7}$$

where $a_i(\tau)$ are functions independent of h, τ the local coordinate of the singular point. Then an extrapolation algorithm is presented to compute the supersingular integral. For a given τ a series of s_j is selected to approximate the singular point *s* with local coordinate equal the superconvergence point accompanied by the refinement of the meshes. Moreover, by means of the extrapolation technique, we not only obtain an approximation with higher order accuracy, but also get a posteriori error estimate.

The rest of this paper is organized as follows. In Sect.2, after introducing some basic formulas of the general (composite) trapezoidal rule, some notations and preliminaries, we present our

main result. In Sect.3 the proof is completed. In Sect.4, extrapolation algorithm is presented and the convergence rate is proved. Finally, several numerical examples are given to validate our analysis.

2.Main result

Let $c = x_0 < x_1 < \cdots < x_{n-1} < x_n = c + 2\pi$ be a uniform partition of the interval $[c, c + 2\pi]$ with mesh size $h = 2\pi/n$ and $f_L(x)$ the piecewise linear interpolant for $f(x)$:

$$f_L(x) = \frac{x - x_{j-1}}{h} f(x_j) + \frac{x_j - x}{h} f(x_{j-1}), \quad x \in [x_{j-1}, x_j], \quad 1 \le j \le n, \tag{8}$$

and a linear transformation

$$\mathrm{x}=_j(\tau) := (\tau + 1)(x_j - x_{j-1})/2 + x_{j-1}, \quad \tau \in [-1, 1], \tag{9}$$

maps the reference element $[-1, 1]$ onto the subinterval $[x_{j-1}, x_j]$. Replacing $f(x)$ in (3) with $f_L(x)$ gives the composite trapezoidal rule:

$$I_n(f, s) := \fint_c^{c+2\pi} \frac{\cos\frac{x-s}{2} f_L(x)}{\sin^3\frac{x-s}{2}} dx = \sum_{j=0}^{n} \omega_j(s) f(x_j) = I(f, s) - E_n(f, s), \tag{10}$$

where $\omega_j(s)$ denotes the Cotes coefficients, see [13] and $E_n(f, s)$ the error functional.

We define

$$F_i(\tau) = (\tau - 1)(\tau + 1)[(\tau + 1)^i - (\tau - 1)^i] \tag{11}$$

and

$$\phi_{i,i+2}(t) = \begin{cases} -\frac{1}{2} \fint_{-1}^{1} \frac{F_i(\tau)}{\tau - t} d\tau, & |t| < 1, \\ -\frac{1}{2} \int_{-1}^{1} \frac{F_i(\tau)}{\tau - t} d\tau, & |t| > 1. \end{cases} \tag{12}$$

Suppose $F_i(\tau)$ is replaced by the Legendre polynomial, then $\phi_{i,i+2}(t)$ defines the Legendre function of the second kind \cite{candrews1992}. Let

$$\phi_{i,i+1}(t) = \phi'_{i,i+2}(x), \tag{13}$$

$$\phi_{ii}(t) = 2\phi'_{i,i+1}(t) = 2\phi''_{i,i+2}(t) \tag{14}$$

and

$$\phi_{ik}(t) = -\frac{1}{2} \int_{-1}^{1} \frac{F_i(\tau)(\tau - t)^{k-i-3}}{(k-i)!} d\tau, \quad k > i + 2. \tag{15}$$

Before presenting the main results, we firstly define $K_s(x)$ as follows

$$K_s(x) = \begin{cases} \frac{(x-s)^3 \cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}} & x \ne s, \\ 8, & x = s. \end{cases} \tag{16}$$

We present our main results below.

Theorem 1 Assume $f(x) \in C^l[a, b], \quad l \ge 3$. For the trapezoidal rule $I_n(f; s)$ defined in (10), there exists a positive constant *C*, independent of *h* and *s* and functions $a_i(\tau)$, independent of *h*, such that

$$E_n(f, s) = \sum_{i=1}^{l-2} \frac{h^{i-1}}{2^{i-1}} f^{(i+1)}(s) a_i(\tau) + R_n(s), \tag{17}$$

where $s = x_{m-1} + (1 + \tau)h/2, \quad m = 1, 2, \cdots, n$ and

$$|R_n(s)| \le C \max_{x \in (c, c+2\pi)} \{K_s(x)\}(|\ln h| + \gamma^{-2}(\tau))h^{l-2} \tag{18}$$

and

$$\gamma(\tau) = \min_{0\le j\le n} \frac{|s - x_j|}{h} = \frac{1 - |\tau|}{2} \tag{19}$$

3.Proof of the Theorem 1

In the following analysis, *C* will denote a generic positive constant which is independent of *h* and *s*. Let P_l and Q_l denote the Legendre polynomial [1] of degree *l* and the associated Legendre function of the second kind, respectively.

Define

$$M_{ik}^j(x,s) \quad = F_i^j(x)(x-s)^{k-i}, \quad k \ge i, \tag{20}$$

where

$$F_i^j(x) = (x - x_{j-1})(x - x_j)[(x - x_{j-1})^i - (x - x_j)^i]. \tag{21}$$

By (9), we have

$$M_{ik}^j(x,s) \quad = \frac{h^{k+3}}{2^{k+3}} F_i(\tau)(\tau - c_j)^{k-i}, \tag{22}$$

where $F_i(\tau)$ is defined by (21) and

$$M_{ik}(\tau, c_j) = (\tau^2 - 1)(\tau - c_j)^{k-i}[(\tau + 1)^i - (\tau - 1)^i] \tag{23}$$

and

$$c_j = \frac{2}{h}(s - x_{j-1}) - 1. \tag{24}$$

Lemma 3.1 Let $K_s(x)$ be defined in (16). For $x \in (x_{j-1}, x_j)$, by linear transformation (9), we have

$$K_s(x) = K_{c_j}(\tau), \quad \tau \in (-1, 1) \tag{25}$$

where

$$K_{c_j}(\tau) = 8 + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j - 2ln)^3} + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j + 2ln)^3} \tag{26}$$

And c_j is defined as (24).

Proof:By the identity in [1]

$$\frac{\pi^3 \cos \pi t}{\sin^3 \pi t} = \sum_{l=-\infty}^{\infty} \frac{1}{(t + l)^3}; \tag{27}$$

then we get

$$\frac{\cos \frac{x-s}{2}}{\sin^3 \frac{x-s}{2}} = \frac{8}{(x - s)^3} + 8\sum_{l=1}^{\infty} \frac{1}{(x - s - 2l\pi)^3} + 8\sum_{l=1}^{\infty} \frac{1}{(x - s + 2l\pi)^3} \tag{28}$$

and

$$\begin{aligned} K_s(x) \quad &= \quad \frac{(x-s)^3 \cos \frac{x-s}{2}}{\sin^3 \frac{x-s}{2}} = 8 + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j - 4l\pi/h)^3} + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j + 4l\pi/h)^3} \\ &= \quad 8 + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j - 2ln)^3} + 8\sum_{l=1}^{\infty} \frac{(\tau - c_j)^3}{(\tau - c_j + 2ln)^3} \\ &= \quad K_{c_j}(\tau), \end{aligned}$$

which completes the proof.

Lemma 3.2 Assume $s \in (x_{m-1}, x_m)$ for some *m* and let $c_j (1 \le j \le n)$ be given by (24). Then we have

$$\phi_{ik}(c_j)=\begin{cases} -\dfrac{2^{k-1}}{h^{k-1}}\oint_{x_{m-1}}^{x_m}\dfrac{\cos\frac{x-s}{2}M_{ik}^m(x,s)}{\sin^3\frac{x-s}{2}}dx, & j=m,\\ -\dfrac{2^{k-1}}{h^{k-1}}\displaystyle\int_{x_{j-1}}^{x_j}\dfrac{\cos\frac{x-s}{2}M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx, & j\neq m, \end{cases} \tag{29}$$

Proof: If $j=m$, by the definition of (3) and noting $k=i$, we have

$$\begin{aligned}\oint_{x_{j-1}}^{x_j}\frac{\cos\frac{x-s}{2}M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx &= \oint_{x_{j-1}}^{x_j}\frac{M_{ik}^j(x,s)K_s(x)}{(x-s)^3}dx\\ &= \oint_{x_{j-1}}^{x_j}\frac{F_i^j(x)K_s(x)}{(x-s)^3}dx\\ &= \lim_{\epsilon\to0}\left\{\left(\int_{x_{j-1}}^{s-\epsilon}+\int_{s+\epsilon}^{x_j}\right)\frac{F_i^j(x)K_s(x)}{(x-s)^3}dx-\frac{2F_i^j(s)K_s(s)}{\epsilon}\right\}\\ &= \frac{h^{k-1}}{2^{k-1}}\lim_{\epsilon\to0}\left\{\left(\int_{-1}^{c_m-\frac{2\epsilon}{h}}+\int_{c_m+\frac{2\epsilon}{h}}^{1}\right)\frac{F_i(\tau)K_s(x)}{(\tau-c_m)^3}d\tau-\frac{hF_i(c_m)K_s(c_m)}{\epsilon}\right\}\\ &= \frac{h^{k-1}}{2^{k-1}}\oint_{-1}^{1}\frac{F_i(\tau)K_s(c_m)}{(\tau-c_m)^3}d\tau\\ &= -\frac{h^{k-1}}{2^{k-1}}\psi_{ii}(c_m).\end{aligned}$$

The case $j\neq m$ can be proved by applying the same approach to the correspondent Riemann integral.

Lemma 3.3 Under the assumption of Lemma3.2, for $k=i+1$, there holds that

$$\phi_{ik}(c_j)=\begin{cases} -\dfrac{2^{k-1}}{h^{k}}\oint_{x_{m-1}}^{x_m}\dfrac{\cos\frac{x-s}{2}M_{ik}^m(x,s)}{\sin^3\frac{x-s}{2}}dx, & j=m,\\ -\dfrac{2^{k-1}}{h^{k}}\displaystyle\int_{x_{j-1}}^{x_j}\dfrac{\cos\frac{x-s}{2}M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx, & j\neq m, \end{cases} \tag{30}$$

Proof:By the definition of (3), we have:

$$\begin{aligned}\oint_{x_{j-1}}^{x_j}\frac{\cos\frac{x-s}{2}M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx &= \oint_{x_{j-1}}^{x_j}\frac{M_{ik}^j(x,s)K_s(x)}{(x-s)^3}dx\\ &= \oint_{x_{j-1}}^{x_j}\frac{F_i^j(x)K_s(x)}{(x-s)^2}dx\\ &= \lim_{\varepsilon\to0}\left\{\int_{x_{j-1}}^{s-\varepsilon}\frac{F_i^j(x)K_s(x)}{(x-s)^2}dx+\int_{s+\varepsilon}^{x_j}\frac{F_i^j(x)K_s(x)}{(x-s)^2}dx-\frac{2F_i^j(s)K_s(s)}{\varepsilon}\right\}\\ &= \frac{h^{k}}{2^{k}}\lim_{\varepsilon\to0}\left\{\left(\int_{-1}^{c_m-\frac{2\varepsilon}{h}}+\int_{c_m+\frac{2\varepsilon}{h}}^{1}\right)\frac{F_i(\tau)K_s(\tau)}{(\tau-c_m)^2}d\tau-\frac{hF_i(c_m)K_s(c_m)}{\varepsilon}\right\}\\ &= \frac{h^{k}}{2^{k}}\oint_{-1}^{1}\frac{F_i(\tau)K_s(\tau)}{(\tau-c_m)^2}d\tau\\ &= -\frac{h^{k}}{2^{k-1}}\psi_{i,i+1}(c_m).\end{aligned}$$

The first identity in (30) is then verified. The second identity can be obtained by applying the approach to the correspondent Riemann integral. The proof is completed.

Lemma 3.4 Under the assumption of Lemma3.3, for $k=i+2$, there holds that

$$\phi_{ik}(c_j)=\begin{cases} -\dfrac{2^{k}}{h^{k+1}}\oint_{x_{m-1}}^{x_m}\dfrac{\cos\frac{x-s}{2}M_{ik}^m(x,s)}{\sin^3\frac{x-s}{2}}dx, & j=m,\\ -\dfrac{2^{k}}{h^{k+1}}\displaystyle\int_{x_{j-1}}^{x_j}\dfrac{\cos\frac{x-s}{2}M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx, & j\neq m, \end{cases} \tag{31}$$

Proof:By the definition of (3), we have:

$$\begin{aligned}
\int\!\!\!\!\!\!-_{x_{j-1}}^{x_j} \frac{\cos\frac{x-s}{2} M_{ik}^j(x,s)}{\sin^3\frac{x-s}{2}}dx &= \int\!\!\!\!\!\!-_{x_{j-1}}^{x_j} \frac{M_{ik}^j(x,s)K_s(x)}{(x-s)^3}dx \\
&= \int\!\!\!\!\!\!-_{x_{j-1}}^{x_j} \frac{F_i^j(x)K_s(x)}{x-s}dx \\
&= \lim_{\varepsilon\to 0}\left\{\int_{x_{j-1}}^{s-\varepsilon}\frac{F_i^j(x)K_s(x)}{x-s}dx + \int_{s+\varepsilon}^{x_j}\frac{F_i^j(x)K_s(x)}{x-s}dx\right\} \\
&= \frac{h^{k+1}}{2^{k+1}}\lim_{\varepsilon\to 0}\left\{\left(\int_{-1}^{c_m-\varepsilon}+\int_{c_m+\varepsilon}^{1}\right)\frac{F_i(\tau)K_s(\tau)}{\tau-c_m}d\tau\right\} \\
&= \frac{h^{k+1}}{2^{k+1}}\int\!\!\!\!\!\!-_{-1}^{1}\frac{F_i(\tau)K_s(\tau)}{\tau-c_m}d\tau \\
&= -\frac{h^{k+1}}{2^k}\psi_{i,i+2}(c_m)
\end{aligned}$$

The first identity in (31) is then verified. The second identity can be obtained by applying the approach to the correspondent Riemann integral. The proof is completed.

Lemma 3.5 Under the assumption of Lemma3.3 and for $k > i+2$, there holds that

$$\phi_{ik}(c_j) = -\frac{2^{k+1}}{h^{k+2}}\int_{x_{j-1}}^{x_j}\frac{M_{ik}^j(x,s)}{(k-i)!}dx, \tag{32}$$

The proof of this lemma can be obtained in a way similarly to that of Lemma 3.3 or Lemma 3.4.

Lemma 3.6 Suppose $f(x)\in C^l[c,c+2\pi], l\geq 3$. If $s\neq x_j$, for any $j=1,2,\cdots,n$, then there holds

$$\begin{aligned}
f(x)-f_L(x) &= \sum_{i=1}^{l-1}\sum_{k=i}^{l-1}\frac{(-1)^{i+1}f^{(k+1)}(s)}{h(i+1)!}\frac{M_{ik}^j(x,s)}{(k-i)!} \\
&\quad + \sum_{i=1}^{l-2}\frac{(-1)^{i+1}}{h(i+1)!}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}M_{i,l-1}^j(x,s) \\
&\quad + \frac{(-1)^l}{hl!}\tilde{M}_l^j(x), \xi_{ij}\in(x_{j-1},x_j),
\end{aligned} \tag{33}$$

where

$$\begin{aligned}
\tilde{M}_l^j(x) &= (x-x_{j-1})(x-t_j)\left[f^{(l)}(\eta_j)(x-x_{j-1})^{l-1} - f^{(l)}(\zeta_j)(x-x_j)^{l-1}\right] \\
&\quad - f^{(l)}(s)M_{l-1,l-1}^j(x,s), \eta_j,\zeta_j\in(x_{j-1},x_j).
\end{aligned}$$

Define

$$\mathcal{H}_m(x) = f(x)-f_L(x)-\sum_{i=1}^{l-1}\sum_{k=i}^{l-1}\frac{(-1)^{i+1}f^{(k+1)}(s)}{h(i+1)!}\frac{M_{ik}^m(x,s)}{(k-i)!}, \quad x\in(x_{m-1},x_m). \tag{34}$$

Lemma 3.7 Under the same assumptions of Theorem 1, for $\mathcal{H}_m(x)$ in (34), there holds that

$$\left|\int\!\!\!\!\!\!-_{x_{m-1}}^{x_m}\frac{\mathcal{H}_m(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}dx\right| \leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\gamma^{-2}(\tau)h^2. \tag{35}$$

where $\gamma(\tau)$ is defined in (19).

Proof. By the definition of $\mathcal{H}_m(x)$, we have

$$|\mathcal{H}_m^{(i)}(x)| \leq Ch^{l-i}, i=0,1,2. \tag{36}$$

As we known

$$\oint_{x_{m-1}}^{x_m} \frac{\mathcal{H}_m(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}dx = 8\oint_{x_{m-1}}^{x_m} \frac{\mathcal{H}_m(x)}{(x-s)^3}dx + \oint_{x_{m-1}}^{x_m} \frac{\mathcal{H}_m(x)[\kappa_s(x)-8]}{(x-s)^3}dx. \quad (36)$$

From the identity

$$\begin{aligned}\oint_a^b \frac{f(x)}{(x-s)^3}dx = \; & \tfrac{f(s)}{2}\left[\tfrac{1}{(a-s)^2} - \tfrac{1}{(b-s)^2}\right] \\ & -\frac{(b-a)f'(s)}{(b-s)(s-a)} + \frac{f''(s)}{2}\ln\frac{b-s}{s-a} \\ & + \int_a^b \frac{f(x)-f(s)-f'(s)(x-s)-f''(s)(x-s)^2/2}{(x-s)^3}dx,\end{aligned} \quad (37)$$

we have

$$\begin{aligned}\oint_{x_{m-1}}^{x_m} \frac{\mathcal{H}_m(x)}{(x-s)^3}dx = \; & \frac{\mathcal{H}_m(s)}{2}\left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right] \\ & -\tfrac{h\mathcal{H}'_m(s)}{(x_m-s)(s-x_{m-1})} \\ & +\frac{\mathcal{H}''_m(s)}{2}\ln\frac{x_m-s}{s-x_{m-1}} + \int_{x_{m-1}}^{x_m}\frac{\mathcal{H}'''_m(\theta(x))}{6}dx\end{aligned} \quad (38)$$

where $\theta(x) \in (x_{m-1}, x_m)$. Since

$$\begin{aligned}& \left|\frac{\mathcal{H}_m(s)}{2}\left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right| \\ &= \left|\frac{\mathcal{H}_m(s)-\mathcal{H}_m(x_{m-1})}{2}\left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right| \\ &= \left|\frac{\mathcal{H}'_m(\xi_{m-1})(s-x_{m-1})}{2}\left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right| \\ &\le C\gamma^{-1}(\tau)h^{l-2},\end{aligned} \quad (39)$$

where $\xi_{m-1} \in (x_{m-1}, x_m)$ and we have used $\mathcal{H}_m(x_{m-1}) = 0$.

$$\left|\frac{h\mathcal{H}'_m(s)}{(x_m-s)(s-x_{m-1})}\right| \le C\gamma^{-1}(\tau)h^{l-2},$$

$$\left|\frac{\mathcal{H}''_m(s)}{2}\ln\frac{x_m-s}{s-x_{m-1}}\right| \le C[|\ln\gamma(\tau)| + |\ln h|]h^{l-2}$$

and

$$\left|\int_{x_{m-1}}^{x_m}\frac{\mathcal{H}'''_m(\theta(x))}{6}dx\right| \le Ch^{l-2}.$$

As for the second term,

$$\begin{aligned}& \left|\oint_{x_{m-1}}^{x_m}\frac{\mathcal{H}_m(x)[\kappa_s(x)-8]}{(x-s)^3}dx\right| \\ &= \max_{x\in[x_{m-1},x_m]}|\mathcal{H}_m(x)|\int_{x_{m-1}}^{x_m}\frac{\kappa_s(x)-8}{(x-s)^3}dx \\ &= \max_{x\in[x_{m-1},x_m]}|\mathcal{H}_m(x)|\left\{\oint_{x_{m-1}}^{x_m}\frac{\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}dx - \oint_{x_{m-1}}^{x_m}\frac{8}{(x-s)^3}dx\right\} \\ &= \max_{x\in[x_{m-1},x_m]}|\mathcal{H}_m(x)|\left\{\frac{1}{\sin^2\frac{s-x_{m-1}}{2}} - \frac{1}{\sin^2\frac{s-x_m}{2}} + \left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right\} \\ &= \max_{x\in[x_{m-1},x_m]}|\mathcal{H}_m(x)-\mathcal{H}_m(x_m)|\left\{\frac{1}{\sin^2\frac{s-x_{m-1}}{2}} - \frac{1}{\sin^2\frac{s-x_m}{2}} + \left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right\} \\ &= \max_{x\in[x_{m-1},x_m]}|\mathcal{H}'_m(\xi_m)(s-x_m)|\left\{\frac{1}{\sin^2\frac{s-x_{m-1}}{2}} - \frac{1}{\sin^2\frac{s-x_m}{2}} + \left[\frac{1}{(x_{m-1}-s)^2} - \frac{1}{(x_m-s)^2}\right]\right\} \\ &\le C\gamma^{-2}(\tau)h^{l-2}.\end{aligned} \quad (40)$$

(35) can be obtained by putting together from (39) to (40) which completes the proof.

Lemma 3.8 Under the assumption of Theorem 2.1, we have

$$\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^l}{hl!}\int_{x_{j-1}}^{x_j}\frac{\tilde{M}_l^j(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx\right|\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\gamma^{-2}(\tau)\frac{h^{l-2}}{l!} \tag{41}$$

and

$$\begin{aligned}&\left|\sum_{i=1}^{l-2}\frac{(-1)^{i+1}}{h(i+1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}\frac{M_{i,l-1}^j(x,s)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx\right|\\ &\leq\begin{cases}C\max\limits_{x\in(x_{j-1},x_j)}\{K_s(x)\}\gamma^{-1}(\tau)\dfrac{h^{l-2}}{(l-1)!}, & i=l-2,\\ C\max\limits_{x\in(x_{j-1},x_j)}\{K_s(x)\}\dfrac{h^{l-2}}{(l-2)!}(|\ln\gamma(\tau)|+|\ln h|), & i=l-3,\\ C\dfrac{h^{l-2}}{(l-i-1)!}, & i<l-3.\end{cases}\end{aligned} \tag{42}$$

Proof: By (33), we see that $|\tilde{\mathcal{M}}_l^j(x)|\leq Ch^{l+1}$, and thus

$$\begin{aligned}&\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^l}{hl!}\int_{x_{j-1}}^{x_j}\frac{\tilde{M}_l^j(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx\right|\\ &=\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^l}{hl!}\int_{x_{j-1}}^{x_j}\frac{\tilde{\mathcal{M}}_l^j(x)K_s(x)}{(x-s)^3}dx\right|\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\frac{h^l}{l!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{1}{(x-s)^3}dx\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\gamma^{-2}(\tau)\frac{h^{l-2}}{l!}.\end{aligned} \tag{43}$$

Now, we estimate (42) and get

$$\begin{aligned}&\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^{i+1}}{h(i+1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}\frac{M_{i,l-1}^j(x,s)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx\right|\\ &=\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^{l-1}}{h(l-1)!}\int_{x_{j-1}}^{x_j}(f^{(l)}(\xi_{il})-f^{(l)}(s))\frac{\mathcal{M}_{l-2,l-1}^j(x,s)K_s(x)}{(x-s)^3}dx\right|\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\frac{h^{l-1}}{(l-1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\left|\frac{1}{(x-s)^2}\right|dt\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\gamma^{-1}(\tau)\frac{h^{l-2}}{(l-1)!}.\end{aligned} \tag{44}$$

and

$$\begin{aligned}&\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^{i+1}}{h(i+1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}\frac{M_{i,l-1}^j(x,s)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx\right|\\ &=\left|\sum_{j=1,j\neq m}^{n}\frac{(-1)^{l-1}}{h(l-1)!}\int_{x_{j-1}}^{x_j}(f^{(l)}(\xi_{il})-f^{(l)}(s))\frac{\mathcal{M}_{l-2,l-1}^j(x,s)K_s(x)}{(x-s)^3}dx\right|\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}\frac{h^{l-1}}{(l-1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\left|\frac{1}{x-s}\right|dt\\ &\leq C\max_{x\in(x_{j-1},x_j)}\{K_s(x)\}(|\ln\gamma(\tau)|+|\ln h|)\frac{h^{l-2}}{(l-1)!}.\end{aligned} \tag{45}$$

The proof is completed.

Lemma 3.9Under the same assumption of Theorem 1 with $k=i,i+1,i+2$, there holds that

$$T_{ik}(\tau) = 8 \sum_{j=-\infty}^{\infty} \psi_{ik}(2j+\tau). \tag{46}$$

Proof: By (23), we have

$$\begin{aligned} \oint_{-1}^{1} \frac{\mathcal{M}_{ik}(\tau,t)\cos\frac{\tau-t}{2}}{\sin^3\frac{\tau-t}{2}}d\tau = \; & 8\oint_{-1}^{1}\frac{\mathcal{M}_{ik}(\tau,t)}{(\tau-t)^3}d\tau + 8\sum_{l=1}^{\infty}\int_{-1}^{1}\frac{\mathcal{M}_{ik}(\tau,t)}{(\tau-t-2l\pi)^3}d\tau \\ & +8\sum_{l=1}^{\infty}\int_{-1}^{1}\frac{\mathcal{M}_{ik}(\tau,t)}{(\tau-t+2l\pi)^3}d\tau, \end{aligned} \tag{47}$$

which means

$$\begin{aligned} T_{ik}(\tau) &= \sum_{i=1}^{n}\psi_{ik}(t) \\ &= 8\sum_{i=1}^{n}\phi_{ik}(2(m-i)+\tau) + 8\sum_{i=1}^{n}\sum_{l=1}^{\infty}\phi_{ik}(2(m-i-nl)+\tau) \\ &\quad + 8\sum_{i=1}^{n}\sum_{l=1}^{\infty}\phi_{ik}(2(m-i+nl)+\tau) \\ &= 8\sum_{l=-\infty}^{\infty}\sum_{i=1}^{n}\phi_{ik}(2(m-i+nl)+\tau) \\ &= 8[\phi_{ik}(\tau) + \sum_{l=1}^{\infty}\phi_{ik}(2l+\tau) + \sum_{l=1}^{\infty}\phi_{ik}(-2l+\tau)]. \end{aligned} \tag{48}$$

where

$$\phi_{ik}(t) = \begin{cases} \oint_{-1}^{1}\dfrac{\mathcal{M}_{ik}(\tau,t)}{(\tau-t)^3}d\tau, & |t|<1, \\ \displaystyle\int_{-1}^{1}\frac{\mathcal{M}_{ik}(\tau,t)}{(\tau-t)^3}d\tau, & |t|>1, \end{cases} \tag{49}$$

which is related with second kind of legendre function. Since

$$Q_0(t) = \frac{1}{2}\log\left|\frac{1+t}{1-t}\right|, \tag{50}$$

then we have

$$\sum_{j=0}^{\infty}Q_0(2j+\tau) + \sum_{j=1}^{\infty}Q_0(-2j+\tau) = \frac{1}{2}\log\frac{2(n-m)-1+\tau}{2m+1-\tau} = 0. \tag{51}$$

By Lemma 3.1 and Lemma 3.2 in [11], we can easily show that $T_{ik}(\tau)$ converges to certain function.

Proof of Theorem 1:By Lemma 3.6, we have

$$\begin{aligned} & \left(\int_{c}^{x_{m-1}} + \int_{x_m}^{c+2\pi}\right)\frac{\cos\frac{x-s}{2}[f(x)-f_L(x)]}{\sin^3\frac{x-s}{2}}dx \\ = \; & \sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{\cos\frac{x-s}{2}[f(x)-f_L(x)]}{\sin^3\frac{x-s}{2}}dx \\ = \; & \sum_{i=1}^{l-1}\sum_{k=i}^{l-1}\frac{(-1)^{i+1}f^{(k+1)}(s)}{h(i+1)!(k-i)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{\cos\frac{x-s}{2}M_{ik}^{j}(x,s)}{\sin^3\frac{x-s}{2}}dx \\ & + \sum_{i=1}^{l-2}\frac{(-1)^{i+1}}{h(i+1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}\frac{\cos\frac{x-s}{2}M_{i,l-1}^{j}(x,s)}{\sin^3\frac{x-s}{2}}dx \\ & + \sum_{j=1,j\neq m}^{n}\frac{(-1)^n}{hl!}\int_{x_{j-1}}^{x_j}\frac{\cos\frac{x-s}{2}\tilde{M}_{l}^{j}(x)}{\sin^3\frac{x-s}{2}}dx. \end{aligned} \tag{52}$$

By the definition of $\mathcal{H}_m(x)$ in (34), we have

$$\begin{aligned}
&⨍_{x_{m-1}}^{x_m} \frac{\cos\frac{x-s}{2}[f(x)-f_L(x)]}{\sin^3\frac{x-s}{2}}\,dx \\
&= ⨍_{x_{m-1}}^{x_m} \frac{\cos\frac{x-s}{2}\mathcal{H}_m(x)}{\sin^3\frac{x-s}{2}}\,dx \\
&+ \sum_{i=1}^{l-1}\sum_{k=i}^{l-1} \frac{(-1)^{i+1}f^{(k+1)}(s)}{h(i+1)!(k-i)!} ⨍_{x_{m-1}}^{x_m} \frac{\cos\frac{x-s}{2}M_{ik}^m(x,s)}{\sin^3\frac{x-s}{2}}\,dx
\end{aligned} \tag{53}$$

Putting (52) and (53) together yields

$$⨍_c^{c+2\pi} \frac{\cos\frac{x-s}{2}[f(x)-f_L(x)]}{\sin^3\frac{x-s}{2}}dx = \sum_{i=1}^{l-1}\sum_{k=i}^{l-1} \frac{(-1)^i h^{k-1}f^{(k+1)}(s)}{(i+1)!2^{k-1}} \sum_{j=1}^{n}\phi_{ik}(c_j) + R_f(s),$$

where

$$R_f(s) = \mathcal{R}^1(s) + \mathcal{R}^2(s)$$

$$\mathcal{R}^1(s) = ⨍_{x_{m-1}}^{x_m} \frac{\mathcal{H}_m(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}dx,$$

$$\begin{aligned}
\mathcal{R}^2(s) = &\sum_{i=1,i\neq m}^{n}\int_{x_{i-1}}^{x_i} \frac{\tilde{M}_l^j(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx \\
&+\sum_{i=1}^{l-2}\frac{(-1)^{i+1}}{h(i+1)!}\sum_{j=1,j\neq m}^{n}\int_{x_{j-1}}^{x_j}\frac{f^{(l)}(\xi_{ij})-f^{(l)}(s)}{(l-i-1)!}\frac{M_{i,l-1}^j(t,s)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}\,dx.
\end{aligned}$$

By Lemma 3.7 and Lemma 3.8, we have

$$|R_n(s)| \le C \max_{x\in(c,c+2\pi)}\{K_s(x)\}(|\ln h| + \gamma^{-2}(\tau))h^{l-2}.$$

The proof is complete.

Based on the Theorem 1, assume $f(x)\in C^4[c,c+2\pi]$, we present the modify trapezoidal rule

$$\tilde{I}_n(f;s) = I_n(f;s) - 4\pi f''(s)\tan\frac{\tau\pi}{2},$$

and

$$\tilde{E}_n(f;s) = ⨍_c^{c+2\pi} \frac{f(x)\cos\frac{x-s}{2}}{\sin^3\frac{x-s}{2}}dx - \tilde{I}_n^2(f;s)$$

then we have

$$\tilde{E}_n(f;s) \le C \max_{x\in(c,c+2\pi)}\{K_s(x)\}[|\ln h| + \gamma^{-2}(\tau)]h^2$$

where $\gamma(\tau)$ is defined as (19).

4.Extrapolation method

In the above sections, we have proved that the error functional of the trapezoidal rule has the following asymptotic expansion of (17). We present our extrapolation algorithms as follow: Assume there exists positive integer n_0 such that

$$m_0 := \frac{n_0(s-c)}{2\pi}$$

is a positive number. Firstly $[c,c+2\pi]$ is partitioned into n_0 equal subinterval denoted by Π_1 with mesh size $h_1 = 2\pi/n_0$. Then we refine Π_1 to get mesh Π_2 with mesh size $h_2 = h_1/2$. In

this way, a series of meshes $\{\Pi_j\}(j=1,2,\cdots)$ in which Π_j is refined from Π_{j-1} with mesh size denoted by Π_{j-1}. Define

$$s_j = s + \frac{h_j}{2}, j = 1, 2, \cdots . \tag{54}$$

and

$$T(h_j) = I_{2^{j-1}n_0}(f, s_j).$$

We present the following extrapolation algorithm: first compute

$$T_1^{(j)} = T(h_j), j = 1, \cdots, m.$$

Second compute

$$T_i^{(j)} = T_{i-1}^{(j+1)} + \frac{T_{i-1}^{(j+1)} - T_{i-1}^{(j)}}{2^{i-1}-1}, \quad i = 2, \cdots, m \quad j = 1, \cdots, m-i.$$

Thoerem 2 Under the asymptotic expansion of theorem 2.1, for $\tau = 0$ and the series of meshes defined by (54), we have

$$\left| I(f,s) - T_i^{(j)} \right| \leq Ch^i$$

and a posteriori asymptotic error estimate is given by

$$\left| \frac{T_{i-1}^{(j+1)} - T_{i-1}^{(j)}}{2^{i-1}-1} \right| \leq Ch^{i-1}.$$

5.Numerical example

In this section, computational results are reported to confirm our theoretical analysis.

Example 1 Consider the supersingular integral

$$\oint_c^{c+2\pi} \frac{\cos\frac{x-s}{2} f(x)}{\sin^3\frac{x-s}{2}} dx = g(s), s \in (c, c+2\pi)$$

with $f(x) = 1 + \sin(3x) + \cos(2x)$ and the exact analysis is $4\pi(-9cos(3s) + 4sin(2s))$

Table1 Errors of the trapezoidal rule rules $s = x_{[n/4]} + (1+\tau)h/2$

n	$\tau = 0$	$\tau = 2/3$	$\tau = -2/3$	$\tau = 1/2$
32	8.1300e-01	8.7751e+01	-1.0333e+02	5.3026e+01
64	1.0380e-01	1.0320e+02	-1.0747e+02	6.0147e+01
128	1.3044e-02	1.0739e+02	-1.0849e+02	6.2136e+01
256	1.6327e-03	1.0846e+02	-1.0874e+02	6.2655e+01
512	2.0416e-04	1.0874e+02	-1.0881e+02	6.2787e+01
1024	2.5555e-05	1.0881e+02	-1.0882e+02	6.2821e+01
h^α	2.9915	-	-	-

Table2 Errors of the mod-trapezoidal rule $s = x_{[n/4]} + (1+\tau)h/2$

n	$\tau = 0$	$\tau = 2/3$	$\tau = -2/3$	$\tau = 1/2$
32	8.1300e-01	-2.5666e+00	4.7362e+00	-1.1117e+00
64	1.0380e-01	-9.2221e-01	1.1658e+00	-4.8090e-01
128	1.3044e-02	-2.5972e-01	2.8810e-01	-1.4320e-01
256	1.6327e-03	-6.8134e-02	7.1544e-02	-3.8476e-02
512	2.0416e-04	-1.7405e-02	1.7822e-02	-9.9397e-03
1024	2.5555e-05	-4.3961e-03	4.4477e-03	-2.5241e-03
h^α	2.9915	1.8379	2.0113	1.7566

Table3 Errors of the trapezoidal rule $s = x_0 + (1+\tau)h/2$

n	$\tau=0$	$\tau=2/3$	$\tau=-2/3$	$\tau=1/2$
32	3.2188e+00	5.2573e+00	-6.3854e+01	3.6217e+00
64	8.2217e-01	3.9187e+01	-7.6374e+01	2.5780e+01
128	2.0567e-01	6.2972e+01	-8.1977e+01	3.7852e+01
256	5.1301e-02	7.5017e+01	-8.4588e+01	4.4033e+01
512	1.2802e-02	8.1044e+01	-8.5843e+01	4.7145e+01
1024	3.1973e-03	8.4055e+01	-8.6457e+01	4.8705e+01
h^α	1.9951	-	-	-

Table4 Errors of the mod-trapezoidal rule $s=x_0+(1+\tau)h/2$

n	$\tau=0$	$\tau=2/3$	$\tau=-2/3$	$\tau=1/2$
32	3.2188e+00	4.6428e+00	3.8210e+00	3.8759e+00
64	8.2217e-01	8.8449e-01	1.0294e+00	8.3812e-01
128	2.0567e-01	1.7963e-01	2.6605e-01	1.8823e-01
256	5.1301e-02	3.9364e-02	6.7560e-02	4.4115e-02
512	1.2802e-02	9.1278e-03	1.7019e-02	1.0645e-02
1024	3.1973e-03	2.1907e-03	4.2717e-03	2.6124e-03
h^α	1.9951	2.2099	1.9610	2.1070

For the case of $s=x_{[n/4]}+(1+\tau)h/2$,Table1 show that when the local coordinate of singular point $\tau=0$, the quadrature reach the convergence rate of $O(h^2)$ as for the non-supersingular point the the convergence rate which agree with our theorematically analysis. From the Table 2 shows the modify trapezoidal rule have the convergence rate of $O(h^2)$ at both the superconvergence point and non-superconvergence point which coincide with our results.For the case of $s=x_0+(1+\tau)h/2$because of no influence of the boundary condition, from table 3 and table 4, we get the superconvergence point the same as $s=x_{[n/4]}+(1+\tau)h/2$ and the superconvergence rate as following which coincide with our theoretically analysis.

Example 2 We still consider the supersingular integral as example 1 with $f(x)=1+\sin(3x)+\cos(2x)$ and the exact analysis is $4\pi(-9cos(3s)+4sin(2s))$

Table 5 Errors of the linear trapezoidal rule $s=-\pi/2$

	$\tau=0$	h^2-extrapolation	h^3-extrapolation	h^4-extrapolation
32	8.1836e+00			
64	3.7710e+00	-6.4164e-01		
128	1.7978e+00	-1.7536e-01	-1.9940e-02	
256	8.7613e-01	-4.5544e-02	-2.2699e-03	2.5439e-04
512	4.3227e-01	-1.1587e-02	-2.6881e-04	1.7053e-05
1024	2.1467e-01	-2.9214e-03	-3.2661e-05	1.0752e-06

Table 6 Errors of the linear trapezoidal rule $s=-\pi/2$

	$\tau=0$	h^2-extrapolation	h^3-extrapolation	h^4-extrapolation
32				
64	4.4126e+00			
128	1.9732e+00	-1.5542e-01		
256	9.2167e-01	-4.3274e-02	-2.5243e-03	
512	4.4386e-01	-1.1319e-02	-2.8587e-04	1.5822e-05
1024	2.1760e-01	-2.8887e-03	-3.3736e-05	1.0652e-06

Now we consider $s=-\pi/2$ and choose the approximation series $s_j=-\frac{\pi}{2}+\frac{\tau+1}{2}h_j$with starting meshes $n_0=32$, then *s* can be located at the mesh point. We list the error and a posteriori estimate

in table 5 and table 6 with the extrapolation rate is h, h^2, h^3 and h^4, respectively. From the table we can see that the numerical results agree with the theoretical analysis very well.

Acknowledgment

The work of Jin Li was supported by National Natural Science Foundation of China (Grant No. 11471195 and Grant No. 11101247), China Postdoctoral Science Foundation (Grant No. 2013M540541 and 2015T80703).

References

[1] Andrews,L.C., Special Functions of Mathematics for Engineers, seconded., McGraw-Hill, Inc, 1992
[2] Ainsworth,M. and Guo,B. An additive Schwarz preconditioner for p-version boundary element approximation of the hypersingular operator in three dimensions. Numer. Math. 85 (2000) 343-366
[3] Choi,U.J. Kim,S.W.and Yun,B.I. Improvement of the asymptotic behaviour of the Euler-Maclaurin formula for Cauchy principal value and Hadamard finite-part integrals. Int. J. Numer. Methods Eng. 61 (2004) 496-513
[4] Du,Q.K. Evaluations of certain hypersingular integrals on interval.Int. J. Numer. Methods Eng. 51 (2001) 1195-1210
[5] Elliott,D. and Venturino,E. Sigmoidal transformations and the Euler-Maclaurin expansion for evaluating certain Hadamard inite-part integrals. Numer. Math. 77 (1997) 453-465
[6] Frangi,A. and Bonnet,M. A direct approach for boundary integral equations with high-order singularities.Int. J. Numer. Meth. Engng.49 (2000)871-898
[7] Hasegawa,T. Uniform approximations to finite Hilbert transform and its derivative. J. Comput. Appl. Math. 163 (2004) 127-138
[8] Hui,C.W.and Shia,D. Evaluations of hypersingular integrals using Gaussian quadrature. Int. J. Numer. Methods Eng. 44 (1999) 205-214
[9] Ioakimidis,N.I. On the uniform convergence of Gaussian quadrature rules for Cauchy principal value integrals and their derivatives. Math. Comp. 44 (1985) 191-198
[10] Kim,P. and Jin,U.C. Two trigonometric quadrature formulae for evaluating hypersingular integrals. Inter. J. Numer. Methods Eng. 56(2003) 469-486
[11] Li,J. Wu,J.M. and Yu,D.H. Generalized Extrapolation for Computation of Hypersingular Integrals in Boundary Element Methods. CMES:Computer Modeling in Engineering & Sciences,42 (2009) 151-175
[12] Li,J. ,Rui,H.X.and Yu,D.H.: Composite Simpsons rule for computing supersingular integral on circle. CMES 97(6), 463-481 (2014)
[13] Li,J. ,Rui,H.X.and Yu,D.H.:Trapezoidal Rule for Computing Supersingular Integral on a Circle.J Sci Comput, Doi 10.1007/s10915-015-0042-3
[14] Yang, C.X.: A unified approach with spectral convergence for the evaluation of hypersingular and supersingular integrals with a periodic kernel. J. Comput. Appl. Math. 239, 322-332 2013
[15] Linz,P. On the approximate computation of certain strongly singular integrals. Computing 35 (1985) 345-353
[16] Liem,C.B. Tao,L. Shih,T.M. The Splitting Extrapolation Method, World Scientific, Singapore, 1995
[17] Lu,T. and Huang,Y. A generalization of discrete Gronwall inequality and its application to weakly singular Volterra integral equality of the second kind, J. Math. Anal. Appl. 282 56-62 (2003)
[18] Lu,T. and Huang,Y., Extrapolation method for solving weakly singular nonlinear Volterra integral equations of the second kind, J. Math. Anal. Appl. 324 225-237 (2006)
[19] Moore,M.N.J. Gray,L.J. and Kaplan,T. Evaluation of supersingular integrals: Second-order boundary derivatives. Int. J. Numer. Meth. Engng 69 (2007) 1930-1947
[20] Monegato,G. Numerical evaluation of hypersingular integrals. J. Comput. Appl. Math. 50 (1994) 9-31
[21] Sidi,A. Practical Extrapolation Methods Theory and Applications. Cambridge University Press, 2003
[22] Wu,J.M.and Sun,W.W. The superconvergence of the composite trapezoidalrule for Hadamard finite part integrals. Numer. Math. 102 (2005)343-363
[23] Wu,J.M.and Sun,W.W.The superconvergence of Newton-Cotes rules for the Hadamard finite-part integral on an interval. Numer. Math. 109 (2008) 143-165
[24] Wu,J.M.and Lu,Y. A superconvergence result for the second order Newton-Cotes formula for certain finite part integrals.IMA Journal of Numerical Analysis. 25 (2005) 253-263
[25] Wu,J.M. and Yu,D.H. The approximate computation of hypersingular integrals on interval, Chinese J. Numer. Math. Appl. 21 (1999) 25-33
[26] Yu,D.H. Natural Boundary Integrals Method and its Applications, Kluwer Academic Publishers, 2002
[27] Yu,D.H. The approximate computation of hypersingular integrals on interval, Numer. Math. J. Chinese Univ. (English Ser.) 1 (1992) 114-127

[28] Zhang,X.P., Wu, J.M. and Yu, D.H. Superconvergence of the composite Simpsons rule for a certain finite-part integral and its applications, Int J Comput Math, 223 (2009) 598-613.
[29] Zhang,X.P., Wu, J.M. and Yu, D.H. , The superconvergence of the composite trapezoidal rule for Hadamard finite-part integral on a circle and its application, Int J Comput Math, 87: (2010) 855-876.
[30] Zhang,X.P., Yu, D.H. and Wu, J.M. , The superconvergence of the composite Newton-Cotes rules for Hadamard finite-part integral on a circle, Computing, 85 (2009) 219-244.

A method to personalize the lumped parameter model of coronary artery

***Bao Li[1] , Wenxin Wang[1] , Xi Zhao[1] , Boyan Mao[1] , †Youjun Liu[1] , Zhou Zhao[2]**

[1]College of Life Science and Bio-engineering, Beijing University of Technology, No. 100 Pingleyuan, Chaoyang District, Beijing 100124, P. R. China.

[2]Peking University People's Hospital, Beijing 100124, P. R. China.

*fengchileizhen@emails.bjut.edu.cn

†lyjlma@bjut.edu.cn

Abstract

Lumped parameter model is a 0 dimensional (0D) model which uses circuit to simulate the cardiovascular circulatory system based on the similarity of circuit and blood vessel. Due to the invasive measurement of physiological indicators in the coronary heart disease surgery, lumped parameter model was used to compute the pressure and flow rate in coronary artery before the surgery. It's a conventional and effective simulation method for appropriate surgical planning. However, due to the specificity of blood vessel structure, thickness and blood viscosity and other physiological parameters of each patient, differences of hemodynamic simulation results ought to exist. If the same parameters of every circuit element in the model were used to simulate cardiovascular system, what is difficult to convincing is that results couldn't reflect the actual situation of each one. In this study, a method which can personalize the lumped parameter model of cardiovascular system based on the non-invasive physiological parameters has been developed. The parameters of cardiovascular system were determined by different physiological parameters. The heart module was determined by systolic blood pressure (SBP), diastolic blood pressure (DBP) and heart rate; the systemic circulation module was determined by cardiac output and Cardio-ankle vascular index (CAVI), while the CAVI was determined by age, height, SBP and DBP; the coronary module, which is most important, was determined by the target waveforms of coronary flow rate predicted from cardiac output. Based on the physiological parameters of a patient provided by Beijing Anzhen Hospital, genetic algorithm was used to optimize the parameters of each module. The considerable results after optimization proved that this method can be applied to each patient. According to the actual physiological parameters of patient, the corresponding changes in the structure and the optimization of parameters of model were necessary procedures to determine the patient-specific treatment.

Keywords: personalize, coronary artery, lumped parameter model, physiological parameters

Introduction

For patients with coronary heart disease, the internal hemodynamic parameters of coronary artery such as coronary flow, pressure, wall shear stress (WSS) and oscillatory shear index (OSI) is not easy to be measured before and during the surgery due to technical and ethical limitations. But these indicators are of great reference significance for preoperative surgical planning and decision-making. Although hemodynamics in the human body can be simulated by animal experiments, the details of the blood flow still will miss due to the limitations of the experimental conditions [1, 2]. While with the development of computer science, computational numerical simulation of fluid flow has been applied to many studies [3]. Therefore, establishing the coronary artery model for patients with coronary heart disease by the technology of computational numerical modeling, and executing detailed simulation focus on the intracoronary hemodynamics of coronary artery are very effective and necessary means.

Lumped parameter model is a common 0D numerical model, which uses circuit elements to simulate the cardiovascular circulation system. The physiological parameters and characteristics of human body can be reflected through it [4]. In the model, the resistance is used to simulate the viscous resistance of blood flow and the resistance of vessel wall; the capacitance is used to simulate the compliance of the vessel wall; the inductance is used to simulate the flow inertia of blood; the diode is used to simulate the valve, which makes the unidirectional flow in some particular structure. The model is based on the Kirchhoff's law, and

the differential-algebraic equations were used to describe the circuit [5]. Frank [6] compared the computation results of lumped parameter model with the blood flow characteristics of blood circulation system, and many similarities were found to prove the effectiveness of the model. Manor [7] has an initially discussion on the lumped parameter model of coronary artery, which provided a theoretical basis for the construction of the model. In the study of treatment direction for coronary artery disease, Taylor [8] used a multi-scale method coupled by lumped parameter model and 3D model. On the basis of this method, Kim [9] predicted the flow rate and pressure inside the coronary arteries for personalized patient with coronary heart disease, but the parameters of lumped parameter model used were not personalized. Burattini and Van Huis [10, 11], respectively, estimated the proximal and distal resistance and compliance of coronary artery through the animal experiment, the parameters used provided a reference value in this study. Taylor and Draney had more deep research in the application of the description of blood flow by lumped parameter model [12].

Based on the CTA image of patients with coronary heart disease, the patient-specific 3D model of coronary artery can be obtained by the technology of three-dimensional reconstruction. Because of the actual characteristics of intracoronary blood flow, and in consideration of the interaction between the heart and the coronary artery, the lumped parameter model of heart should be coupled at the inlet of 3D model; and because of the resistant and compliance exist in the coronary artery, microcirculation and veins, the lumped parameter vascular bed model should be coupled at the outlet of the coronary artery of 3D model [9]. Zhao [13] developed the coupling algorithm. By this algorithm, the 3D model and the lumped parameter model were coupled to obtain the patient's geometric multi-scale model, which was used to simulate the circulation system focus on the coronary artery. The geometric multi-scale model is shown in Fig. 1.

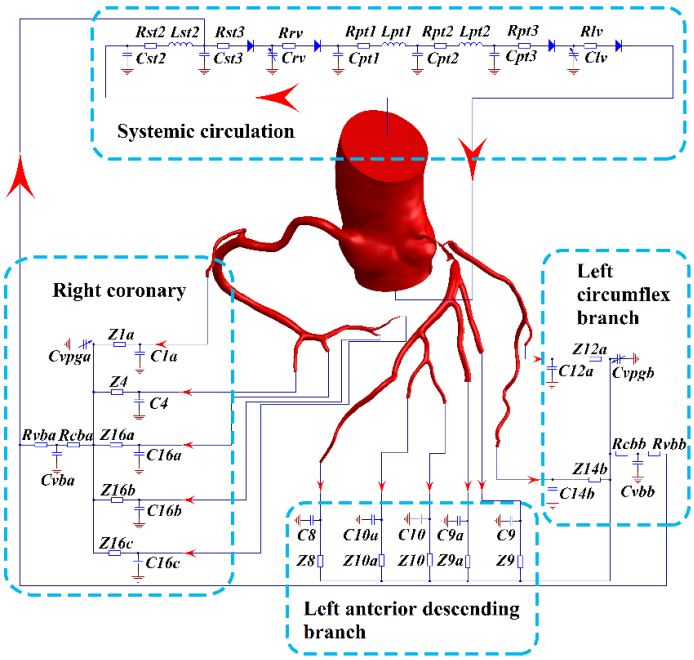


Figure 1 The geometric multi-scale model of coronary artery

In the geometric multi-scale model, the lumped parameter model is used to provide boundary condition of inlet and outlet for the coronary artery of 3D model. For patients with coronary artery stenosis, a similar geometric multi-scale model can be established, and the simulation of different bypass surgery based on this model can be executed. After the hemodynamic simulation calculations, the flow rate, pressure, WSS and OSI inside coronary artery can be observed. Optimal values of these indicators under different bypass surgery, resulting in the best surgical approach for each patient. However, it is precisely because of the specificity in structure, vascular thickness and vascular length of each patient, the parameters of lumped parameter model coupled to the inlet and outlet of coronary artery should be adjusted based on the physiological parameters appropriately, thus to ensure that the simulation results are personalized.

In this study, according to a patient's physiological data provided by Beijing Anzhen Hospital, including age, height, blood pressure, cardiac output and coronary artery CTA images, based on the personalized CTA images and the technology of 3D reconstruction, personalized geometric multi-scale model of coronary artery was built after the coupling by 3D model and lumped parameter model. The parameters of lumped parameter model were adjusted to promise the specificity of patient according to the measured non-invasive physiological parameters. We adjusted the heart module coupled to the inlet of coronary artery according to the patient's aortic pressure (including systolic and diastolic blood pressure), adjusted the systemic module coupled to the outlet of circulatory artery based on its cardiac output, aortic pressure, age and height, and finally genetic algorithm was used to optimize the parameters of coronary module coupled to the outlet of coronary artery. Once the simulation waveforms were similar to the target waveforms, the personalized parameters suitable for the patient would be got.

Methods

Building the personalized geometric multi-scale model

According to the coronary artery CTA images of 512 * 512 pixels of 200 patients provided by Anzhen Hospital, 3D model of coronary artery was reconstructed by three-dimensional reconstruction software such as Mimics. According to the results of 3D reconstruction, the lumped parameter model representing heart was coupled to the inlet of 3D model, the lumped parameter model representing systemic circulation was coupled to the outlet of systemic circulation artery, and the lumped parameter coronary vascular bed model was coupled at the outlet of coronary artery, then the personalized geometric multi-scale model of patient was got, as shown in Fig. 2.

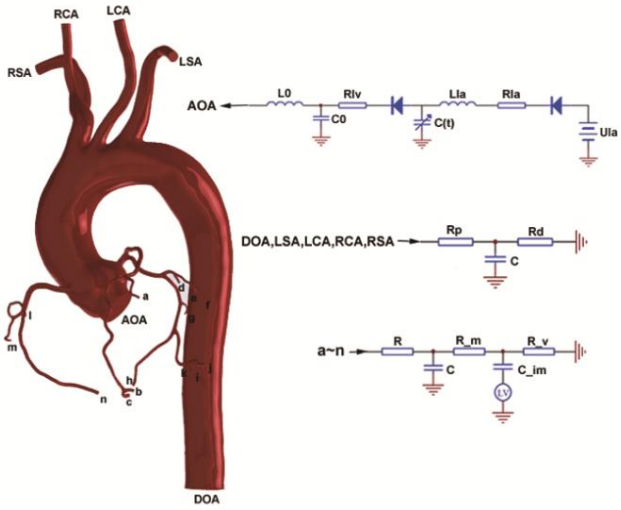


Figure 2 The personalized geometric multi-scale model of patient

In the lumped parameter model of heart module, the function to simulate the ventricular systole was applied to the variable capacitance *C(t)*. Suga and Sagawa [14] established the ventricular pressure-volume relationship by animal experiments, which can be expressed by the time-varying function *E(t)* as Eq. (1):

$$E(t) = \frac{P_{SV}(t)}{V_{SV}(t) - V_0} \tag{1}$$

Where *E(t)* is time-varying function (mmHg/ml), $P_{sv}(t)$ is time function of ventricular pressure

(mmHg), $V_{sv}(t)$ is time function of ventricular volume (ml), V_0 is ventricular reference volume (ml), and that is theoretical volume relative to the "zero ventricular pressure". Based on Eq. (1), Boston [15] sorted out the function which can express ventricular systole by mathematical fitting:

$$E(t) = (E_{\max} - E_{\min}) \cdot E_n(t_n) + E_{\min} \tag{2}$$

Where the E_{max} is related to ventricular pressure-volume relation during end-systole, while the E_{min} is related to ventricular pressure-volume relation during end-diastolic, and the $E_n(t_n)$ is double hill function [16]:

$$E_n(t_n) = 1.55\left[\frac{\left(\frac{t_n}{0.7}\right)^{1.9}}{1+\left(\frac{t_n}{0.7}\right)^{1.9}}\right]\left[\frac{1}{1+\left(\frac{t_n}{1.17}\right)^{21.9}}\right] \tag{3}$$

Where t_n is t/T_{max}, while the T_{max} can be calculated by the cardiac cycle t_c:

$$T_{max} = 0.2 + 0.15t_c \tag{4}$$

$E_n(t_n)$ was called the normalization function of ventricular elasticity, which was applied to the variable capacitance $C(t)$ in the Fig. 1. Then a pressure waveform that simulates ventricular systole generated on $C(t)$ as shown in Fig. 3. After the computation by lumped parameter model, the time-varying pressure was converted to a waveform of flow rate at AOA in Fig. 2 to provide boundary condition of flow rate for the inlet of 3D model.

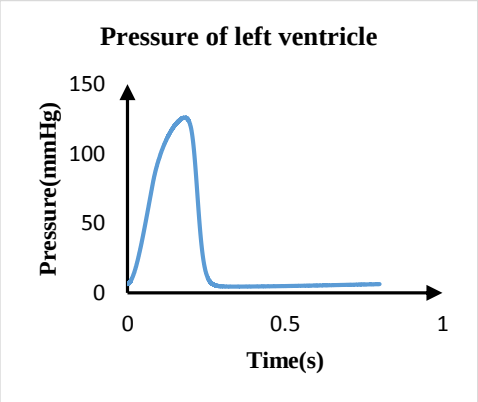


Figure 3 The pressure waveform of left ventricular

While, in the lumped parameter model of coronary module, R represents the resistance of coronary artery, R_m represents the resistance of coronary microcirculation, R_v represents the resistance of coronary vein, C represents the compliance of coronary artery, C_im represents the compliance of coronary microcirculation [9]. Different values of resistance and capacitance produce different waveforms of coronary flow rate at the outlet of coronary artery. Therefore, for personalized patients, the value of each parameter should be personalized. Similarly, in the lumped parameter model coupled to the outlet of systemic circulation arteries, R and C represent the resistance and compliance of the systemic circulation vessels, respectively. There is an impedance at each outlet of 3D model, which provides pressure boundary conditions for the

outlet of it. The next step is personalizing the parameters of the lumped parameter model at the inlet and outlet of 3D model.

The personalization of parameters in lumped parameter model of each module

Due to the calculation of geometric multi-scale model is very time-consuming, it is impractical to try many different parameters to execute the calculation. Therefore, the pure lumped parameter model was used to replace the multi-scale model. In this case, the trunks of coronary artery in the 3D model were replaced with the classical three-element windkessel model [17], and the lumped parameter vascular bed model was connected to the end of each branch, just as in the multi-scale model. Based on the personalized coronary structure of the patient, the final structure of pure lumped parameter model is shown in Fig. 4.

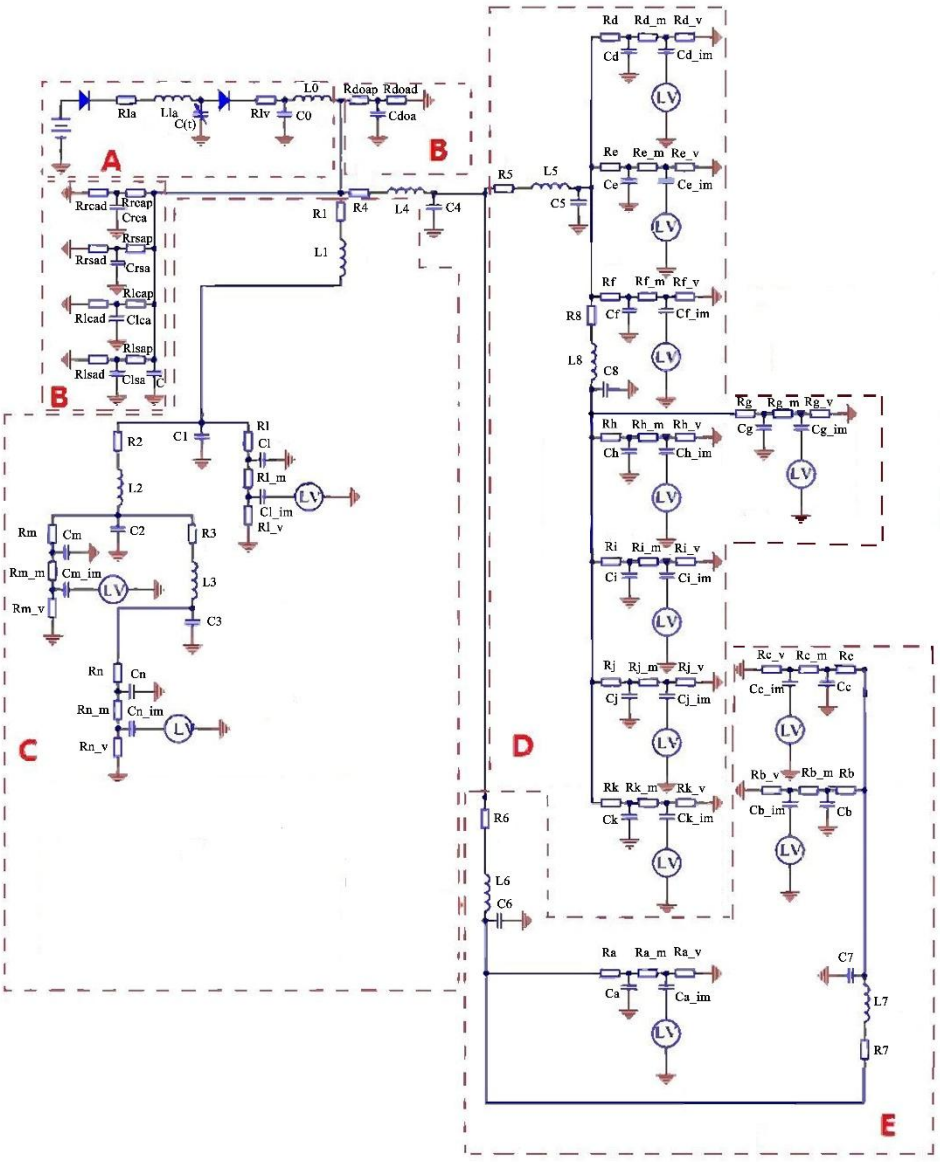


Figure 4 Lumped parameter model of blood circulation system, A represents the heart, B represents the systemic circulation artery, C represents right coronary, D represents left anterior descending artery, and E represents left circumflex branch

Based on the final structure, the parameters of blood circulation system could be personalized. The system could be divided into three modules: the heart module, the systemic circulation module, and the coronary module. The personalization of each module was based on different physiological parameters. Firstly, the parameters of heart module were personalized. The function of this module is providing the aortic pressure (including systolic and diastolic blood pressure), cardiac output and cardiac cycle (waveform cycle) to the circulatory system. It can be obtained from Eq. (2) that different values of E_{max} and E_{min} will affect the output of ventricular systole function, resulting in different aortic pressure and cardiac output waveforms. Therefore, the SBP, DBP and cardiac output of patient should be considered when the E_{max} and E_{min} were adjusted. While, the aortic pressure and cardiac output were not determined by E_{max} and E_{min} entirely, because there exists impedance of systemic circulation and coronary at the outlet of heart, and the magnitude of impedance also affects cardiac output and aortic pressure. Secondly, the personalization of parameters in systemic circulation module was executed. Because of the effects on flow rate and pressure of coronary artery by different resistance and compliance of systemic circulation, the parameters of this module were personalized. The resistance of this module, such as *Rdoap*, *Rdoad* in Fig. 4, were determined by the cardiac output of patient. The specific adjustment method of E_{max}, E_{min} and resistance of this module will be described later. The capacitance values of systemic circulation module, such as *Cdoa* and *C*, could be adjusted according to the heart-ankle index (CAVI) and the aortic pressure. CAVI is an index which can evaluate the compliance of aorta, thoracic aorta, abdominal aorta, femoral artery to the ankle artery. The relationship between CAVI and the pulse wave velocity is shown in Eq. (5) [18].

$$CAVI = \frac{2\rho}{\Delta P} ln(\frac{P_s}{P_d}) PWV^2 \tag{5}$$

Where ρ is blood density, ΔP is the differential pressure between systolic and diastolic blood pressure, P_s is systolic blood pressure, P_d is diastolic blood pressure, PWV is pulse wave velocity, and according to the famous Moens-Korteweg Eq. (6) [19]:

$$PWV = \sqrt{\frac{Eh}{2\rho R}} \tag{6}$$

Where E is the elastic modulus of blood vessel, h is the average arterial thickness, and R is the average radius of artery. Combining with Eq. (5) and Eq. (6), the relationship of elastic modulus and CAVI could be summarized as Eq. (7).

$$E = \frac{CAVI(P_s - P_d)R}{ln\left(\frac{P_s}{P_d}\right)h} \tag{7}$$

While in the establishment of lumped parameter model, the vascular compliance equation is:

$$C = \frac{3\pi R^3 L}{2Eh} \tag{8}$$

Where C is the compliance of blood vessels, L is the total length of blood vessels, and h is the average thickness of blood vessels. Combining with Eq. (7), the equation of CAVI about the compliance of lumped parameter model would be got.

$$C = \frac{3\pi r^2 L \ln(\frac{Ps}{Pd})h}{2CAVI(Ps-Pd)} \quad (9)$$

Where L is the distance from heart to ankle, about the height of human body multiplied by 0.8, h is average arterial thickness from heart to ankle, usually is 0.64mm, r is the average arterial radius, Usually is 2.1mm. By the above equation, the personalized average arterial compliance could be estimated through the patient-specific SBP, DBP, height and CAVI. But the measurement of CAVI is more cumbersome, so this study carried out a research about the relationship between CAVI and the non-invasive physiological parameters. Through the physical examination data of 195 teachers provided by Beijing University of Technology Hospital, such as height, body weight, Body Mass Index (BMI), SBP, DBP, CAVI, ankle brachial index (ABI), Heart rate and blood viscosity, by the method of multiple linear regression analysis, conclusion could be obtained. The critical threshold of CAVI is recognized as 9.0. In the analysis, IBM SPSS Statistics 19 was used to gather statistics, *t* test was used for the comparison between groups, and the x^2 test was used for technical data.

Finally, the personalization of lumped parameter coronary vascular bed model coupled to the outlet of 3D model was executed. In this study, the genetic algorithm [20] was used to optimize the parameters of this module according to the patient-specific waveforms of coronary flow rate. Genetic algorithm is a computational model of the evolutionary process of natural selection and genetic mechanism of Darwinian biological evolution. It is an optimization method to search the optimal solution by simulating the natural evolutionary process. The first step of genetic algorithm was establishing the target of optimization. The waveforms of patient-specific aortic pressure, cardiac output and coronary flow rate of each branch were treated as target waveforms. The SBP, DBP and cardiac output of patient could be measured, and the waveforms of aortic pressure and cardiac output could be fitted based on the measured data. While, the flow rate of coronary blood is difficult to be measured in clinical, but it could be predicted by cardiac output. According to the study of patient-specific flow rate of coronary by Kim [9], under the rest state, the total flow rate of coronary is 4% of the cardiac output, the average flow rate of left anterior descending branch is 1.3 ml/s, and the average flow rate of left circumflex branch is 1.5 ml/s, and the average flow rate of right coronary is 0.6 ml/s. Of course, the above data will not be applicable to each patient, but could be used as a general reference range, because of the little difference of coronary flow rate between each patient. In this study, the personalized coronary flow rate of patient was predicted according to the theory of Hagen - Poiseuille hemodynamics. The WSS is similar in the vascular of same level, and by the Hagen - Poiseuille shear stress equation:

$$WSS = \frac{4\mu}{\pi d^3} Q \quad (10)$$

Where μ is the blood viscosity, d is the diameter of blood vessel, and Q is the flow rate of blood vessel. Based on Eq. (10), in the ideal model, the WSS in the vascular of same level is approximately equal, so the flow ratio is third power of the diameter ratio at the bifurcation of vessel. Based on the above theory, the flow rate of left and right coronary was predicted in this study. The diameters of each branch of coronary arteries were measured in the SolidWorks based on the result of 3D reconstruction of coronary CTA images. And according to Eq. (10),

the flow rate of each branch was allocated based on the diameters measured and the total flow rate of coronary. For example, for the vessel in Fig. 5, the principle of blood flow allocation is shown in Eq. (11).

$$\frac{Q_1}{Q_2} = \left(\frac{D_1}{D_2}\right)^3 \tag{11}$$

Figure 5 The bifurcation of vessel

Where Q is the flow rate of vessel and D is the diameter of vessel. After obtaining the average flow rate of each branch, the waveform of each flow can be fitted according to the above data. Due to the distal branch of left anterior descending and left circumflex are mainly attached to the left ventricle, they will be squeezed by the ventricular systole, so the blood of left coronary is mainly supplied in the systole; and due to the distal branch of right coronary is mainly attached to the right ventricle, in which the systolic pressure is much smaller compared with the left ventricle, so the blood of right coronary is mainly supplied in the systole [9]. In addition, the peak value of left coronary flow rate in the systole should be 40-50% of it in the diastole flow [21]. The waveforms of left anterior descending branch, left circumflex branch, and right coronary flow rate were shown in Fig. 6.

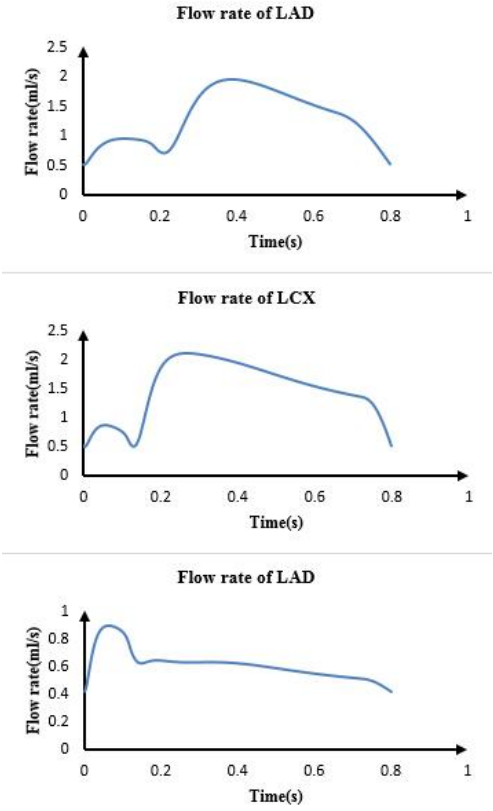


Figure 6 The waveforms of LAD, LCX, and right coronary flow rate

Through the above method, the personalized waveforms of aortic pressure, cardiac output and

flow rate of each branch could be obtained and treated as the target waveforms. In the process of optimization, the genetic algorithm invoked the simulation function of cardiovascular system, and calculated the root mean square error of simulation waveforms and standard waveforms to be the target function, as shown in Eq. (12). And these errors are the fitness values of the various groups in the process of heredity, but also the ultimate goal of optimization.

$$f = \sqrt{\frac{1}{K}\sum_{n=1}^{K}|y_n - y_n'|^2} \tag{12}$$

Where y_n is the data of target waveforms, y_n' is the data of simulation waveforms. After determining the optimization goals, the search range of parameters value in the lumped parameter model needs to be given, including E_{max}, E_{min}, *Rdoap*, *Rdoad*, and the other parameters in cardiovascular system. We have initial value of each parameter based on the experience of adjusting them manually. Of course, these values are not applicable to each patient, but we can set the search range according to them. In the algorithm, all parameters of the simulation circuit were regarded as each individual of population, and each population generation was generated from the search range. In this study, the simulation waveforms and the target waveforms were compared innovatively, and the program would adjust the search range automatically based on comparison results, which made the optimization better and faster. Finally, in the main part of the algorithm, the parameters of lumped parameter model were optimized based on the root mean square error between the standard waveforms and the simulation waveforms. If the error was less than 5%, the obtained parameters could be considered as patient's personalized parameters of lumped parameter model. The flow chart of the optimization algorithm is shown in Fig.7.

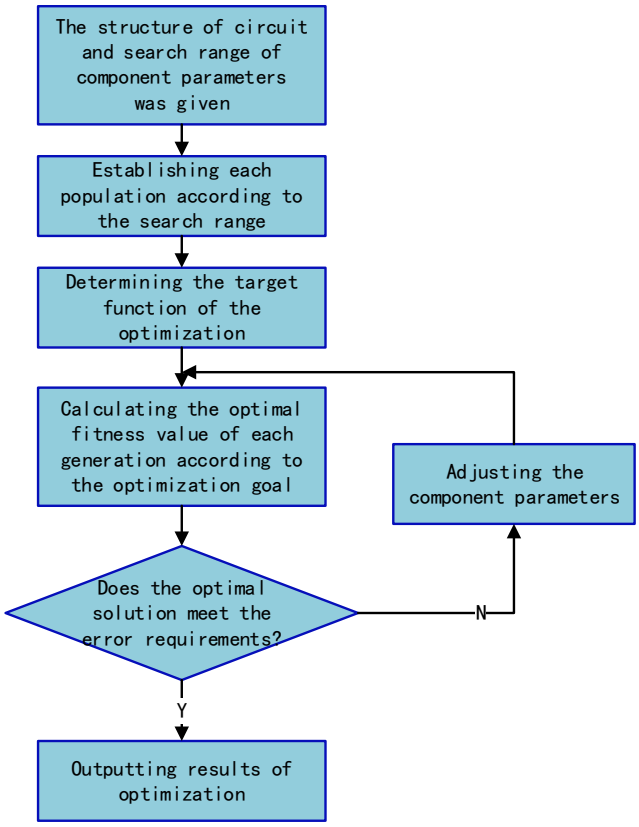


Figure 7 The flow chart of optimization algorithm

Simulation details

In this study, the genetic algorithm was used to optimize the parameters of circulatory system. Each parameter variable was the chromosome of each individual, and each chromosome was binary coded. For each chromosome, the range of values was given according to the different parameter types. The maximum range was the coronary microcirculation resistance R_m, the range was [10, 300], and the smallest range was the coronary artery capacitance C, the range was [0, 1]. Because of the values of capacitance were small, the required accuracy was 4 decimal places, which would require the range of chromosome values x_i was $(b_i - a_i) \times 10^5$ at least. According to Eq. (13), the coding string length m_i, and the gene length would be determined:

$$2^{m_i-1} < (b_i - a_i) \times 10^5 < 2^{m_i} \tag{13}$$

From the above equation, it could be obtained that m_i is 25, so the length of chromosome gene was 25. In this study, the number of coronary branches was 14, and five parameters required to be optimized for each branch. Including E_{max}, E_{min} and *Rdoap* et al., the total number of parameters required to be optimized was 102, so the number of chromosomes was 102. The genetic crossover rate was 0.7, the mutation rate was 0.02, and the genetic generations were 400. Through programming, CPUs were fully invoked by multi-threaded technology for simulation. The size of population in the algorithm was 64, in order to adapt to the core number of CPU, and achieve multi-threaded computation. As a result, computation time reduced and optimization efficiency was improved.

Results and conclusions

Correlation between CAVI and Physiological Parameters

There were 195 samples, and the comparison between normal group and risk group about the correlation between CAVI and physiological parameters is shown in Table 1.

Table 1 The comparison between normal group and risk group

Physiological parameters	CAVI≤9.0(n=143)		CAVI＞9.0(n=52)		t/x^2	P
	Average	Standard deviation	Average	Standard deviation		
Age(year)	52.04	10.44	67.12	9.05	-9.22	0.001
Height(cm)	164.34	7.06	165.77	6.63	-1.273	0.212
Weight(kg)	64.88	10.51	67.21	10.04	-1.38	0.178
BMI(kg/m^2)	23.89	2.90	24.37	3.14	-0.98	0.321
SBP(mmHg)	125.72	12.72	143.33	15.26	-8.09	0.003
DBP(mmHg)	78.61	8.58	85.08	8.68	-4.64	0.003
Heart rate(times)	69.15	10.65	69.21	9.77	-0.04	0.976
ABI	1.08	0.09	1.11	0.08	-2.78	0.013
Blood viscosity (mPa·s)	5.39	0.42	5.24	0.34	2.31	0.025

Pearson correlation analysis of CAVI and physiological parameters showed that CAVI was linearly related to age, systolic and diastolic blood pressure. The correlation coefficient between CAVI and age is $r = 0.760$, $P < 0.01$; correlation coefficient between CAVI and SBP is $r = 0.540$, $P < 0.001$; correlation coefficient between CAVI and DBP is $r = 0.365$, $P < 0.01$. The results of their multiple linear regression analysis are shown in Table 2 and Eq. (14).

Table 2 The results of multiple linear regression analysis of influencing factors of CAVI

Physiological parameters	Regression coefficients	Standard error	Standardized regression coefficient	T	P
Age(year)	0.071	0.639	0.599	10.336	0.001
SBP(mmHg)	0.023	0.008	0.253	3.031	0.003
DBP(mmHg)	-0.005	0.012	-0.034	-0.645	0.003

$$CAVI = 0.071 * Age + 0.023 * SBP - 0.005 * DBP + 1.393248 \quad (14)$$

By the above equation, the value of CAVI can be estimated by age, systolic and diastolic blood pressure, and according to Eq. (9), the personalized average arterial compliance of patient can be estimated by CAVI and height, which could be used to determine the capacitance of systemic circulation module.

Optimization results of parameters

In this study, coronary CTA images and physiological parameters of patient with coronary heart disease were collected, such as aortic pressure (systolic and diastolic blood pressure), cardiac output, height, age, heart rate and so on. The patient was 50 years old, the height was 1.62 meters, SBP was 140 mmHg, DBP was 80 mmHg, cardiac output was 5.5 L/min, heart rate was 85 beats/min. According to the prediction method of coronary flow described above, the total flow rate of coronary was 4% of cardiac output, and that was 3.667 ml/s. The dominant type of coronary artery was left dominant type, and the flow ratio of left and right coronary was 14:3. The flow rate of left coronary was 3.020 ml/s and the flow rate of right coronary was 0.647 ml/s. The flow rate of LAD was 1.786 ml/s, and the flow rate of LCX was 1.881ml/s. The diameter measured by Solidworks and flow rate predicted following the Hagen - Poiseuille law of each branch are shown in Table 3.

Table 3 The diameter and predicted flow rate of each branch of coronary artery

	Branch number	Diameter(mm)	Flow rate(ml/s)
LAD	a	1.730	0.684
	b	2.196	0.623
	c	2.007	0.479
LCX	d	1.324	0.309
	e	1.302	0.210
	f	1.209	0.168
	g	1.064	0.148
	h	2.183	0.536

	i	1.298	0.074
	j	1.077	0.055
	k	2.058	0.381
Right coronary	l	1.773	0.136
	m	1.732	0.126
	n	2.282	0.385

Based on the flow rate and standard waveform of each branch, the personalized waveforms of coronary flow rate could be fitted. The waveforms of aortic pressure, cardiac output and each coronary flow were treated as target waveforms. The results of optimization by genetic algorithm is shown in Fig. 8.

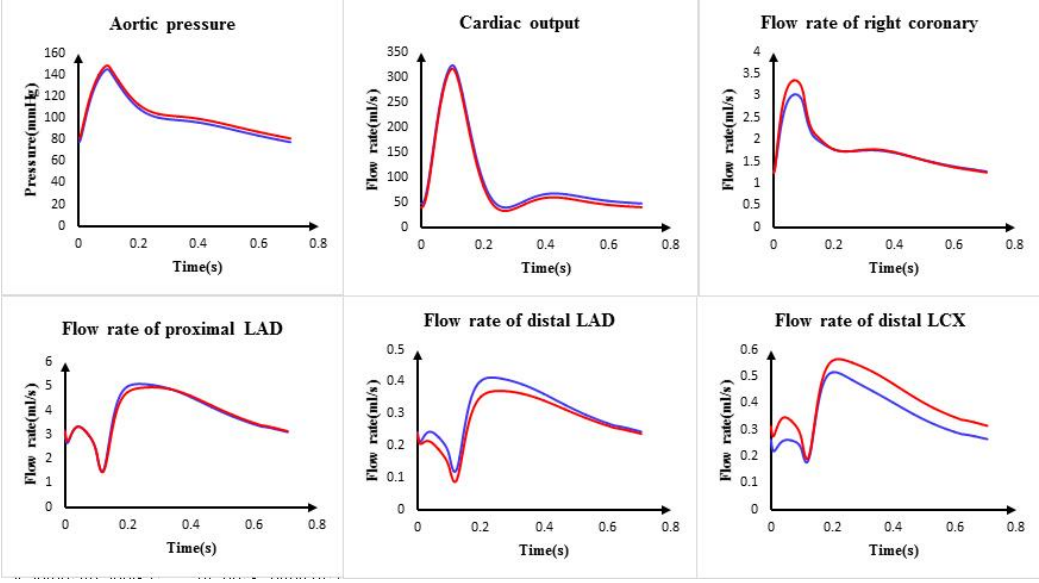


Figure 8 Waveforms of optimization results

As it was shown that the target waveforms were similar to the optimized simulation waveform, and the optimization effect was good, which proves the practicability of this method. The values of parameters of lumped parameter model are shown in Table 4.

Table 4 The parameter values of LPM

Units：Resistance (mmHg s/ml), Inductance (mmHg s^2/ml), Capacitance (ml/mmHg)

Heart module		***Rla***	***Rlv***	***Lla***	***L0***	***C0***
		0.0037	0.0010	0.0028	0.0045	0.9500
		Rp	***Rd***	***C***		
Systemic circulation module	***DOA***	0.2830	1.6046	0.5241		
	RCA	3.3875	23.3258	0.0328		
	RSA	1.5228	10.8912	0.0330		
	LCA	2.0988	2.8714	0.0361		
	LSA	2.4877	16.1566	0.0287		
		R	***R_m***	***R_v***	***C***	***Cim***
Coronary module	***a***	74.5897	118.1439	59.7817	0.00067	0.0045
	b	66.0806	121.7862	63.1510	0.00066	0.0066

c	150.0712	282.3035	85.3090	0.00077	0.0069
d	108.8882	150.1549	79.29208	0.00049	0.0044
e	114.9609	295.7461	118.7321	0.00074	0.0071
f	132.6122	234.5194	110.2857	0.00059	0.0049
g	134.1361	170.0063	104.3336	0.00051	0.0081
h	117.0208	141.4654	122.2045	0.00058	0.0072
i	77.79236	156.5535	125.8896	0.00038	0.0049
j	108.8599	103.9168	120.3035	0.00030	0.0061
k	133.5083	179.2889	139.7658	0.00059	0.0066
l	59.1354	238.6547	80.0720	0.00051	0.0031
m	55.3238	169.4415	132.2594	0.00025	0.0023
n	70.4592	104.8977	118.2161	0.00023	0.0035

While the E_{max} was 1.0, E_{min} was 0.0305, and the value of C was 0.0031. In Fig. 4, the parameters such as $R1$, $C1$ and $L1$, which were used to simulate the trunk of coronary artery, had little effect on the flow rate of each branch. So the values of these parameters were small. In order to save the calculation time, it's unnecessary to optimize them. In this study, R was 3.5, C was 0.00005, L was 0.01.

summary

In this study, by the CTA images of coronary artery of a patient provided by Beijing Anzhen Hospital, the personalized lumped parameter model structure of coronary artery was established; by SBP, DBP and age of patient, the personalized CAVI value was estimated, and the average arterial compliance was estimated further; by cardiac output, the personalized flow rate waveforms of each branch of coronary artery were predicted, and the genetic algorithm was used to optimize the parameters of personalized lumped parameter model of coronary artery, which was applied to the patient. All the programming was finished by C# based on the visual studio 2013 development environment. When a simulation of the personalized cardiovascular system was executed, these parameters were significant basis, which will be great significance to clinical preoperative surgical planning.

Limitation and expectation

In this study, we propose a series of methods to determine the parameters of lumped parameter model of cardiovascular system. The genetic algorithm was used to optimize the parameters to achieve the simulation waveforms is similar to the target waveforms. However, since the target waveform of coronary flow is not a real clinical measure, the parameter values determined in the present study can be used to calculate the fractional flow reserve (FFR) of patients with coronary artery stenosis in subsequent studies. Comparing with the clinical FFR, so as to verify and improve the method. In addition, although the technology of multi-threading was used to optimize the parameters of lumped parameter model to improve the efficiency, but the reconstruction process of 3D model for coronary is time-consuming, preoperative surgical planning still need some time to prepare, there are some limitations. In the future work, we are supposed to find a quick method to understand the coronary structure and predict the flow rate of each branch, which will achieve the rapid personalization of lumped parameter model, and provide effective reference value for the rapid clinical preoperative surgical planning of patients.

References

[1] Bertolotti C, Zhao Q, Lamontagne B, et al. (2006) Influence of Multiple Stenoses on Echo-Doppler Functional Diagnosis of Peripheral Arterial Disease: A Numerical and Experimental Study, *Annals of Biomedical Engineering* **34**, 564-574.

[2] Pekkan K, Kitajima H D, De Z D, et al. (2005) Total cavopulmonary connection flow with functional left pulmonary artery stenosis: angioplasty and fenestration in vitro, *Circulation* **112**, 3264-71.

[3] Wilding NB. (2001) Computer simulation of fluid phase transitions, *American Journal of Physics* **69**, 1147-1155.

[4] Liang FY, Senzaki H, Yin ZF. (2013) Transient hemodynamic changes upon changing a BCPA into a TCPC in staged Fontan operation: A computational model study, *Scientific World Journal* **2013**, 77-87.

[5] Pietrabissa R, Mantero S, Marotta T, et al. (1996) A lumped parameter model to evaluate the fluid dynamics of different coronary bypasses, *Medical Engineering & Physics* **18**, 477-484.

[6] Frank O. (1899) Die Grundform des arteriellen Pulses: Erste Abhandlung mathematische Analyse, *Z Biol* **37**, 483-526.

[7] Manor D, Beyar R, Sideman S. (1994) Pressure-flow characteristics of the coronary collaterals: a model study, *American Journal of Physiology* **266**, H310-8.

[8] Taylor CA, Fonte TA, Min JK. (2013) Computational fluid dynamics applied to cardiac computed tomography for noninvasive quantification of fractional flow reserve: scientific basis, *Journal of the American College of Cardiology* **61**, 2233-41.

[9] Kim HJ, Vignon-Clementel IE, Coogan JS, et al. (2010) Patient-Specific Modeling of Blood Flow and Pressure in Human Coronary Arteries, *Annals of Biomedical Engineering* **38**, 3195-3209.

[10] Burattini R, Sipkema P, Van Huis GA, et al. (1985) Identification of canine coronary resistance and intramyocardial compliance on the basis of the waterfall model, *Annals of Biomedical Engineering* **13**, 385-404.

[11] Van Huis GA, Sipkema P, Westerhof N. (1987) Coronary input impedance during cardiac cycle as determined by impulse response method, *American Journal of Physiology* **253**, H317-24.

[12] Taylor CA, Draney MT. (2004) Experimental and computational methods in cardiovascular fluid mechanics, *Annual Review of Fluid Mechanics* **36**, 197-231.

[13] Zhao X, Liu YJ, Ma LC. (2015) Hemodynamic comparison between normal graft and Y-type graft in coronary artery bypass grafting: a numerical study using 0D/3D coupling method, *Journal Of Mechanics In Medicine And Biology* **15**, 1-20.

[14] Suga H, Sagawa K. (1974) Instantaneous Pressure-Volume relationships and heir ratio in the excised, supported canine left ventricle, *Circulation Research* **35**, 117-26.

[15] Goldberger AL, Amaral LA, Glass L, et al. (2000) PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, *Circulation* **101**, 215-20.

[16] Stergiopulos N, Meister J J, Westerhof N. (1996) Determinants of stroke volume and systolic and diastolic aortic pressure, *American Physiological Society* **270**, 2050-9.

[17] Formaggia L, Quarteroni A , Veneziani A. (2009) *Cardiovascular Mathematics*, Springer.

[18] Shirai K, Utino J, Otsuka K, et al. (2006) A novel blood pressure-independent arterial wall stiffness parameter; cardio-ankle vascular index (CAVI), *Journal of Atherosclerosis & Thrombosis* **13**, 101-7.

[19] Newman DL, Greenwald SE. (1978) *Validity of the Moens-Korteweg Equation, The Arterial System*, Springer Berlin Heidelberg.

[20] GOLDBERG DE. (1989) *Genetic algorithms in search optimization and machine learning*, Reading Menlo Park: Addison-wesley,.

[21] Spaan J A, Breuls N P, Laird J D. (1981) Diastolic-systolic coronary flow differences are caused by intramyocardial pump action in the anesthetized dog, *Circulation Research* **49**, 584-93.

An adaptive control dynamic-grids generation method for numerical simulation of moving and deforming boundary flow field

***Zeyu GUO[1], † Zuogang CHEN [1,2]**

1 State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;
2 Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration(CISSE), Shanghai 200240, China

*Presenting author: sjguozeyu@sjtu.edu.cn
†Corresponding author: zgchen@sjtu.edu.cn

Abstract

A novel adaptive control dynamic-grids generation method based on two-dimensional body fitted grids is developed, which offers a solution to the application of high-precision structured grids in the complicated moving and deforming boundary flow fields. There have been some dynamic-grids methods proposed by researchers. In these methods, advantages of structured grids in the accuracy and rate of generation, and the applicability of turbulence models, have not been made full use of. The adaptive control method manages to generate updated high-precision body fitted structured grids in each time step according to the movement and deformation of flow field boundaries. Researches of flow fields with moving and deforming boundaries are supposed to benefit from this new dynamic-grids method.

Keywords: dynamic-grids; grid generation; body fitted structured grids; deforming boundary flow field; fish-like undulation

Introduction

In the field of simulation and quantification of biological locomotion in fluids, a series of challenges are presented for the development of suitable numerical methods in front of researchers. The movement and deformation of the boundary in an unsteady flow field remains a challenge for numerous CFD researches. The dynamic-grids technology is a commonly used method to tackle with such situation. Then more dynamic-grids methods are proposed and developed by researchers.

The dynamic-grids methods are generally coupled with unstructured triangular grids owing to its brilliant solution to the complex geometric configurations. Cavallon et al. [1] developed an edge-based unstructured flow solver for flow fields with moving and/or deforming boundaries. In the work of Batina [2], an improved algorithms for the solution of the time-dependent Euler equations are presented for unsteady aerodynamic analysis involving unstructured dynamic

meshes. Blom [3] presents an investigation on the spring analogy, which serves for deformation in a moving boundary problem. Many researchers have used and improved the spring analogy coupled with unstructured grids in a variety of flow fields. Bottasso et al. [4] complement a network of edge springs with an additional set of linear springs that oppose element collapsing to achieve better robustness. Pérez et al. [5] applied the 2D dynamic mesh based on spring-smoothing dynamic mesh model for deposit shape prediction in boiler banks. However, it is still unable to handle the large deformation using only spring-smoothing method for dynamic-grids. Therefor this method is often coupled with globe or regional remeshing methods.

But the inherent defect of unstructured triangular grids is inextricable. The structured grid has more advantages compared with the unstructured grid in the computer memory saving and the computational efficiency. To employ the structured grids on the dynamic-grids methods, the first requirement is to generate the body-fitted grids for different boundaries. In 1999, the generation of various forms of grids has been systematically introduced [6]. It is a challenge for researchers to maintain the grids quality and orthogonality during the movement and deformation of boundaries. The commercial CFD software, FLUENT, is adopted a dynamic-grids method for structured grids, which is called diffusion-smoothing method. But this method is still of strong limitation. At present, few researchers stimulate the flow fields with moving and deforming boundaries with structured grids, which is a treasure trove for CFD.

With the development of CFD technology, many other methods are introduced into the problem. The wake structure of a single swimmer is simulated by Mattia et al. [7] using a vortex particle method coupled with a penalization technique. Yigang Xu and Decheng Wan [8] approach the problem with multi-block and overset grid method.

We present a novel adaptive control dynamic-grids generation method based on two-dimensional body fitted grids in the paper. The method, which is named as adaptive control dynamic-grids method, aims to solve the application of high-precision structured grids in the complicated flow field of moving and deforming boundaries. It can control the accuracy of structured grids of the entire flow field readily, especially in the near-wall and wake region. The computational expense is less than most present dynamic-grids methods.

Methodology

Structured body fitted grids generation method

The kernel of body fitted structured grids generation is to build up the mapping relation between the computational domain and the physical domain. Many pioneering researchers have proposed various schemes such as the TTM [9] method based on solving a set of elliptic partial differential equations. However, considering that the grids of entire flow field should update in each time steps or certain time steps when using dynamic-grids method, the computational speed and burden are an important factor. As a result, we construct the new method by the algebraic method to achieve the high generating speed. The quality of body fitted grids can be of enough precise. To mapping the physical domain into the computational domain, the Eq. (1) needs to be figured out.

$$\begin{aligned} x &= x(\xi,\eta) \quad \xi = \xi(x,y) \\ y &= y(\xi,\eta)' \eta = \eta(x,y) \end{aligned} \tag{1}$$

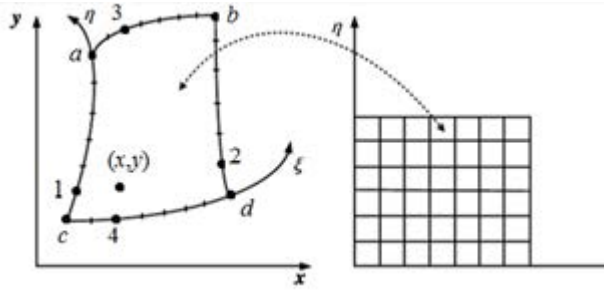
where Domain ***xy*** refers to the physical domain and Domain ***ξη*** refers to the computational domain. In the application of CFD, the boundary condition is the input parameter, which means that the one-to-one correspondence can be formed between physical domain boundaries and computational domain boundaries. The relation is presented as Eq. (2):

$$\begin{aligned} &(x_i, y_{\min}) \Leftrightarrow (\xi_{i,0}, \eta_{i,0}) \qquad (i = 0,1,\ldots,M_x) \\ &(x_i, y_{\max}) \Leftrightarrow (\xi_{i,M_y}, \eta_{i,M_y}) \qquad (i = 0,1,\ldots,M_x) \\ &(x_{\min}, y_j) \Leftrightarrow (\xi_{0,j}, \eta_{0,j}) \qquad (i = 0,1,\ldots,M_y) \\ &(x_{\max}, y_j) \Leftrightarrow (\xi_{M_x,j}, \eta_{M_x,j}) \qquad (i = 0,1,\ldots,M_y) \end{aligned} \tag{2}$$

where M_x and M_y are the number of nodes in x and y dimension separately. In order to implement the coordinate conversion between two different coordinate systems, the Jacobian matrix is introduced. The Jacobian matrix can be deduced as

$$J = \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} = \xi_x\eta_y - \eta_x\xi_y = \begin{vmatrix} x_\xi & x_\eta \\ y_\xi & y_\eta \end{vmatrix}^{-1} \tag{3}$$

where J is the Jacobian matrix. Subscripts represent partial differentiations with respect to the referred variables. It should be noted that these partial differentiations are actually discrete form because the grids nodes are discrete. The computational process is illustrated as Fig. 1 below.



$$\begin{aligned} \xi_x &= \frac{1}{x_2 - x_1}, \eta_x = \frac{(y_1 - y_c)/(y_a - y_c) - (y_2 - y_d)/(y_b - y_d)}{(x_2 - x_1)} \\ \eta_y &= \frac{1}{y_3 - y_4}, \xi_y = \frac{(x_3 - x_a)/(x_b - x_a) - (x_4 - x_c)/(x_d - x_c)}{(y_3 - y_4)} \end{aligned} \tag{4}$$

Figure 1 Algebraic interpolation principle

In Eq. (4), subscripts of ξ and η represent the partial differentiations while the subscripts of x and y are the indications of different points as the figure illustrated. The robustness of the computational process is important to deal with various flow field boundaries. There are divisions performed to obtain J. Therefore an extra procedure is applied to complete the coordinate conversion if the denominators happen to be 0 or close to 0.

A subroutine is programmed into the computational process. When the absolute value of the denominator is less than 10^{-6}, the coordinate system is rotated by the subroutine. By rotation of the coordinate of a particular angle, the value of each node is converted and the difference is enlarged to avoid morbid matrix. After the coordinate values of all nodes in the new coordinate system are achieved, let the coordinate just rotate backward by the same angle. By completing the steps above, the grids generation is of generality for different flow field.

Adaptive control dynamic-grids method

Commercial software (FLUENT) has several built-in methods, which can help users to solve dynamic-grids problem readily. But yet, these methods are not able to satisfy researches of various flow fields universally.

The adaptive control dynamic-grids method is coupled with the User Define Function (UDF) based on FLUENT. The UDF is written by the users themselves in C code. Actually, users are able to interfere in any steps of the entire process and redefine them by UDF. However, most of the dynamic-grids applications based on UDF are confined to the moving rigid boundary or unstructured grids. We present the new dynamic-grids method based on structured grids to solve the moving and deforming boundary flow field problem, which is a breakthrough for structured dynamic-grids.

The key part of dynamic-grids is how to control the movement of boundaries and the update of grids. For instance, when a flexible object is locomote in a flow, the movement and deformation occurs simultaneously, which is a challenge for the quality of grids. The software has provided several functional interfaces to settle the dynamic boundaries, while there is hardly any one of them that has been used to control the nodes system of the entire flow field. Fig. 2 presents the process diagram of the adaptive control method.

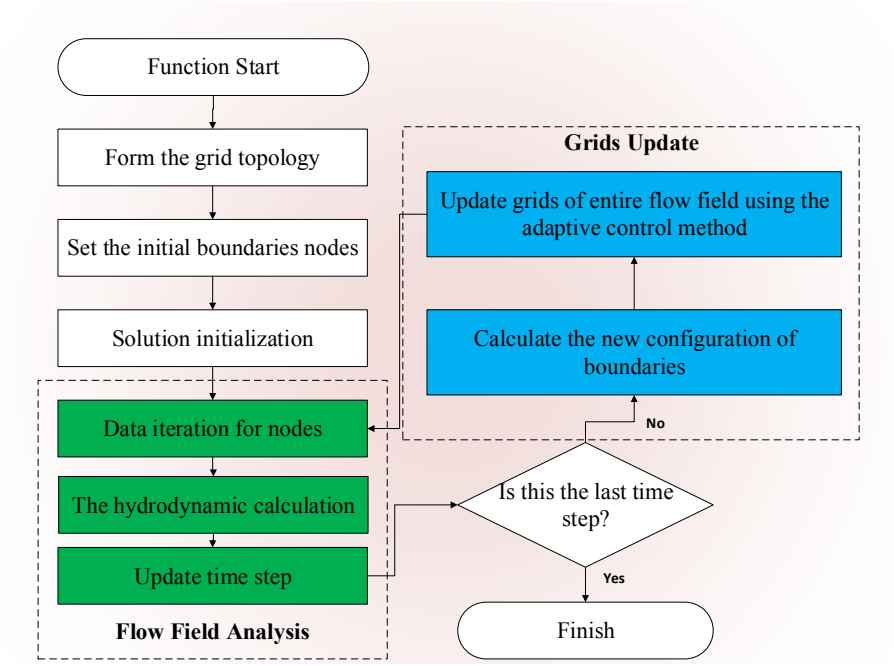


Figure 2: the process diagram of the adaptive control dynamic-grids method

In the Grids Update step, some functions and function functional interfaces are adopted by UDF. The adaptive control method manages to employ the DEFINE_GRID_MOTION functional interface in the entire flow field instead of only boundaries. Firstly, function NODE_X and function NODE_Y traverse all nodes to get their coordinate information. And then, the location and configuration of boundaries are defined according to motion equations. Afterward, the new grids of present time step can be generated by the algebraic interpolation mentioned above.

During the update process, the sequence of nodes is rigid when functions traverse and coordinates update, which needs extra attention. The match-up of the original nodes and update nodes is important for the date iteration. To arrange the node sequence properly, we number all the nodes and the sequence number is used to iterate back the update coordinates and other discrete flow field data.

Due to the complete control of grid nodes, the quality of grids does not depend on smooth models of FLUENT. The quality of both globe and regional grids can be assured by adjusting the UDF. In general, the adaptive control dynamic-grids method updates grid nodes in each time step rapidly and control the spacing of boundaries to assure the grid quality and orthogonality.

Application and Validation

In hydrodynamic researches concerning fish swimming, the hydrodynamic characters of flexible boundaries remain a hot spot. The high propulsive efficiency and energy saving locomotion style of fish undulation are attractive for scientists to study and simulate. Research of flexible body locomotion is subject to the difficulties of dealing with moving and deforming boundaries. Methods provided by most CFD softwares are still not able to settle the problem perfectly. To validate the advantages of the adaptive dynamic-grids method, we compared it with several other methods in a fish-like undulation flow field.

The moving and deforming boundary problem of fish-like undulation

The simulated body undulates actively in unbounded oncoming flow. Related variables are normalized by the body length L and oncoming velocity U, as well as the time is normalized by L/U. The movement of the undulating body in y direction is given as

$$y(x,t) = ax^n \sin[2\pi b(x - Sp \cdot t)] \tag{5}$$

where a is amplitude, n=1.1, $2\pi b$ is wave number, Sp is phase velocity, t is time, where x and y are the stream-wise and the lateral coordinates whose origin locates at the leading edge of the body. $0 \leqq x \leqq Xend(t)$ where $Xend(t)$ is x-coordinate of the trailing edge calculated as

$$\int_0^{Xend(t)} \sqrt{1+\left(\frac{\partial y}{\partial x}\right)^2}\, dx = 1.0 \tag{6}$$

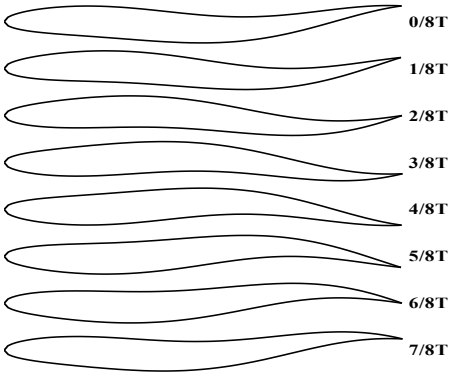


Figure 3: Configurations of an undulating NACA0010 during one period (a=0.051, b=1.0)

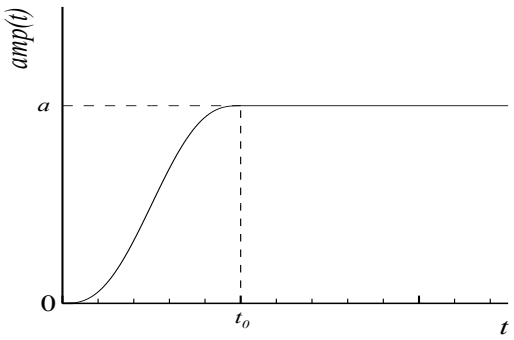


Figure 4: Employed acceleration modes for undulating amplitude

Equation (2.2) keeps the body length constant during the undulation. The obtained smooth deforming surface of NACA0010 can be seen in Figure 3.

In the present study, in the flow field acceleration stage, which is within the non-dimensional time, the undulating amplitude increase from zero to a final constant. As shown in Fig.4, when $t \geq t_0$, it keeps constant. The acceleration model satisfies Eq. (7):

$$amp = \begin{cases} at^3(6t^2 - 15t + 10) & 0 \leq t \leq 1 \\ a & t > 1 \end{cases} \tag{7}$$

It is noted that the boundary of the fish body is flexible. When the fish is swimming, the movement and deformation of the boundary both occur, which lead to the enormous problems of dynamic-grids.

Application of present dynamic-grids methods

In the simulation of fish undulation flow field, on the one hand, the movement and deformation of boundaries are of significance. On the other hand, a finer grid scheme in the boundary layer region should be maintained consistently. As different turbulence models have different requirements for the grids quality in the boundary layer region, the initial spacing of boundary grids must be carefully maintained to meet various requirements during the dynamic computation

The spring-based smoothing method regards the edges between any two grid nodes as a network of interconnected springs. When the grids deteriorated because of the boundary displacement, a remeshing method will be enabled to update the new grids to avoid convergence problems.

In Fig. 5, different remeshing parameters are chose for the triangular grids. The Maximum Length Scale (MLS) parameters, which specifies the upper limit of cell size above which the cells are marked for remeshing, are 0.01 for Fig. 5(a) and 0.001 for Fig. 5(b). When the MLS is 0.01 which is relatively a large number, the alteration of the globe grids is slight. While if MLS becomes 0.001, which is relatively a small number, clusters of dense grids appear stochastically. To obtain acceptable grids, many researchers have made lots of improvements for the choices of relevant parameters based on their particular projects. But the disadvantages of unstructured triangular grids on the study of the boundary layer are inevitable.

The diffusion-smoothing method is proposed based on structured grids. The grids motion is governed by a diffusion equation, which is a Laplace equation. The equation describes how the prescribed boundary motion diffuses into the interior grids. Fig. 6(*a*) and Fig. (*b*) show the grids at t=1 and t=6 separately. When the fish just finished the acceleration at t=1, the orthogonality of grids in the near-wall region is no longer as good as the initial grids. When computation proceed to t=6, the skewness of structured quadrilateral grid in some regions, especially in the intense motion region, is so deteriorated that the convergence problem may happen. Compared with spring-smoothing method, the diffusion-smoothing method is computationally more expensive but is able to achieve better grid quality.

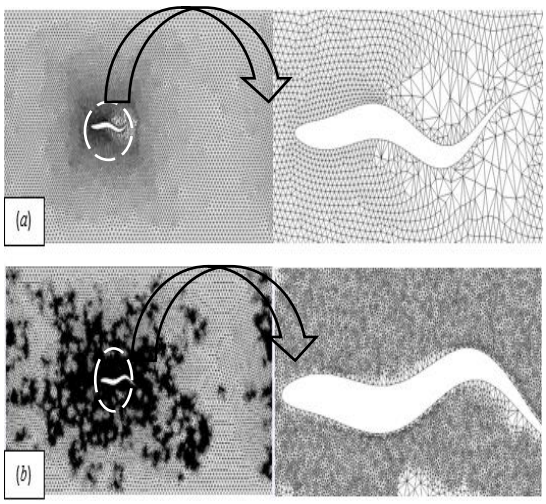


Figure 5: Grids based on spring smoothing and remeshing method at different Maximum Length Scale

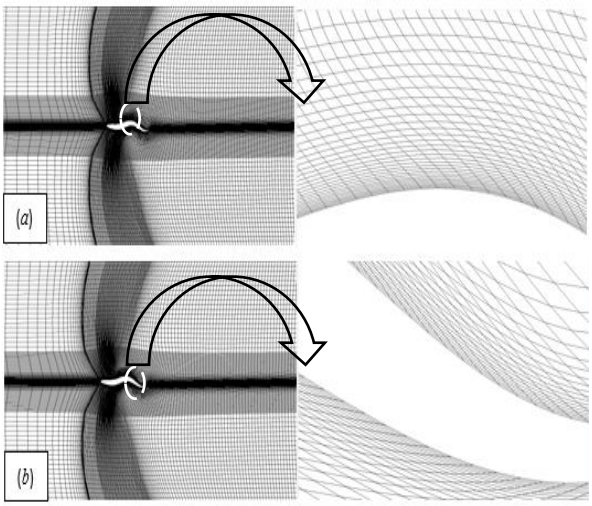


Figure 6: Domain grid and grid at boundary layer based on diffusion-smoothing method at *t*=1 (*a*) and *t*=6 (*b*)

In the overset grid method, the background grid and the embedded grid are generated independently. When the fish swims, only the embedded grid need to move and deform without having to remesh the background grid. The connection between background and the embedded grid is the interfaces, which interpolate cell data in overlapping regions. To avoid the interpolation errors, the embedded region must overlap sufficiently, which is expensive computationally.

Application of Adaptive control dynamic-grids method

The initial grid topology of the adaptive control dynamic-grids method is similar with the diffusion-smoothing method besides the enhancement of boundary and wake grids. The computational domain is $-3.0 \leqq x \leqq 5.0$, $-2.0 \leqq y \leqq 2.0$, where the coordinates are normalized by fish length L. The grid consists of $[140 \times 195]$ points in x, y directions where minimum normal spacing is 2×10^{-5} to meet the requirement of SST turbulence model. The obtained grids movement and deformation can be seen in Fig. 7.

During one complete period of a fish undulation (Fig. 7), the density of grid nodes maintains consistently. Although the motion of boundary is intense, the orthogonality of grid in the near-wall region is well preserved. Considering the wake flow of the fish undulation, grids in the wake region are diffused gradually. Moreover, the edge curvature of grid cells are correlated with the camber line of the fish body. From the generation method aforementioned, it can be noticed that the update grids in new time step are independent of the grids in last time step, which means that the quality of grids will not deteriorate time-dependently. When the movement and deformation of boundaries are cyclical, the motion of grids can be accordingly cyclical by using this new method, which ensure the quality of grids and the accuracy of computation.

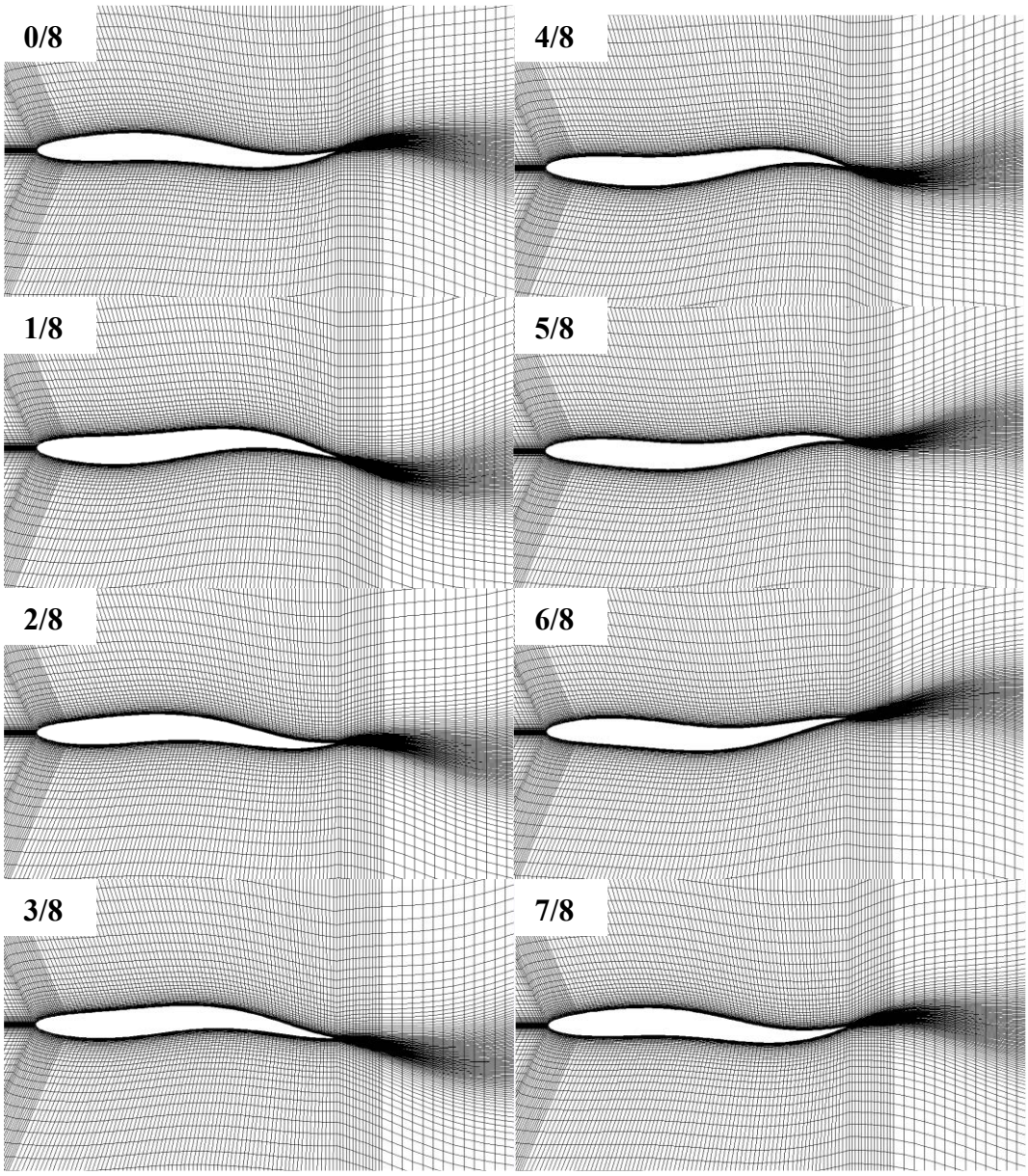


Figure 7: The dynamic movement and deformation of structured grids based on the adaptive control method during one period (*a*=0.051, *b*=1.0)

Validation on hydrodynamic force coefficients and

The total force acting on the body varies during the undulating motion. Therefore, in the present study, self-propulsion is defined as the condition when the time-averaged total force becomes zero. The hydrodynamic force coefficients are defined as follows[10]:

$$C_{Fx} = \frac{F_{xf}}{\frac{1}{2}\rho {U_0}^2 S}, C_{Px} = \frac{F_{xp}}{\frac{1}{2}\rho {U_0}^2 S} \tag{8}$$

$$\overline{C}_{Fx} = \frac{\int_0^T F_{xf}\mathrm{dt} \Big/ T}{\frac{1}{2}\rho {U_0}^2 S}, \overline{C}_{Px} = \frac{\int_0^T F_{xf}\mathrm{dt} \Big/ T}{\frac{1}{2}\rho {U_0}^2 S}, \overline{C}_{Tx} = \overline{C}_{Fx} + \overline{C}_{Px} \tag{9}$$

where F_{xf} is x-component of frictional force exerted on the body, F_{xp} is x-component of pressure

exerted on the body, T is the period of the undulation, ρ is the density of the fluid, U_0 is the oncoming velocity, S is the area of the body. The negative symbol of pressure component F_{xp} (against to x-axis) represents thrust and the positive symbol of frictional force component F_{xf} (same to x-axis) represents resistance. When the self-propulsion condition is achieved, the time-averaged total force exerted on the body becomes zero. The hydrodynamic force coefficients computed by these methods mentioned above are compared. C_{Fx}, C_{Px} and C_{Tx} are the primary focus.

Fig. 8 presents the simulated time history of hydrodynamic force coefficient C_{Fx} of overset grid method. The computed curves of C_{Fx} oscillate by time with large noise as shown in Fig. 8. The noise comes from the interpolation errors resulted from the data exchange. The noise can be decreased by densifying background grids, which causes the computational expense increased significantly. Nevertheless, encryption can only attenuate the noise, but it is unable to remove the noise completely.

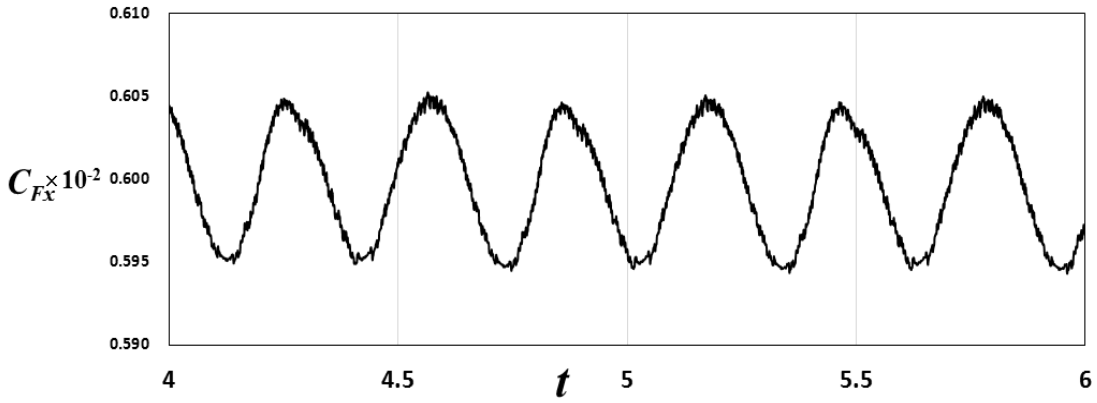


Figure 8: Simulated time history of hydrodynamic force coefficients at (Re=10^7, a=0.051, b=1.5, Sp=1.17）based on overset grids method

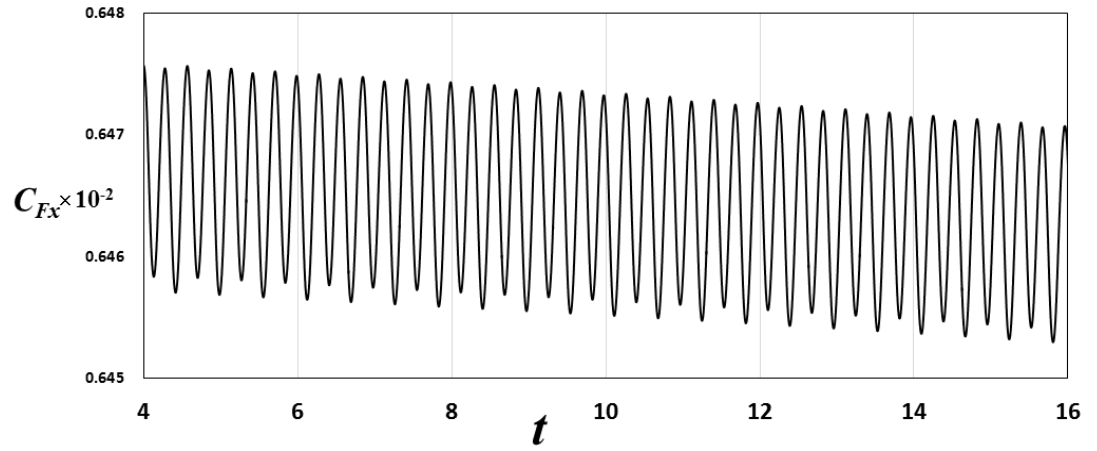


Figure 9: Simulated time history of hydrodynamic force coefficients at （Re=10^7, a=0.051, b=1.5, Sp=1.17）based on diffusion-smoothing method

Fig. 9 is the computed results of diffusion-smoothing method. The time-averaged value of C_{Fx} gradually decrease with the periodical locomotion of fish, which means that the result is unauthentic. If finer grids with smaller height value of the first boundary layer grids are adopted to meet the requirement of advanced turbulence models, the deterioration of grids quality will lead to computational divergence.

The results of the adaptive control method are presented in Fig.10. Simulated time history of indicates that the self-propulsion condition is achieved. All the hydrodynamic force coefficients are continuous and smooth, and undulate periodically. The results will not be deteriorative or even divergent as computation goes on.

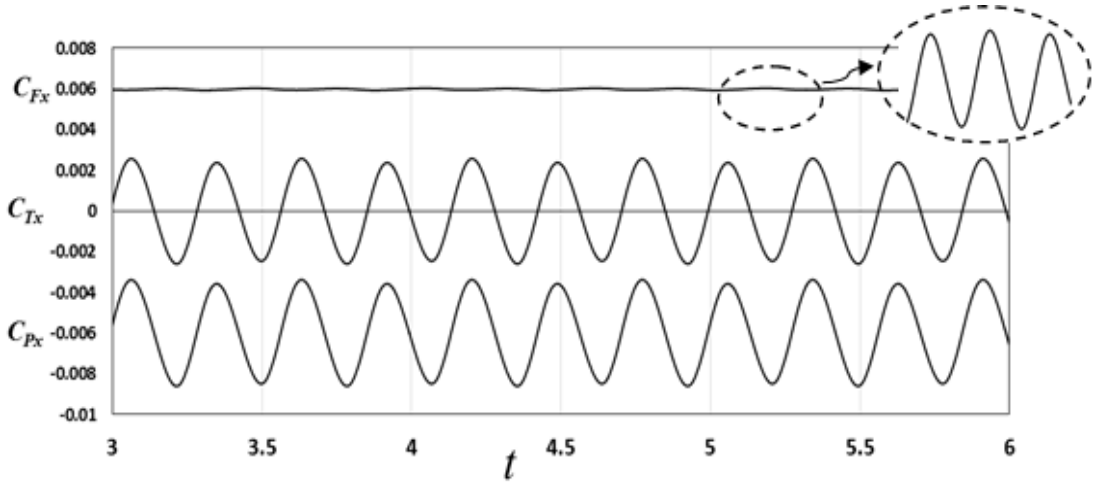


Figure 10: Simulated time history of hydrodynamic force coefficients at （Re=10^7，a=0.051，b=1.5，Sp=1.17）based on adaptive control method

The structure and characteristics of the flow field are also part and parcel of the computation. The globe and regional velocity contours based on the adaptive control method and the spring smoothing method are compared in Fig. 11. The globe velocity contours are similar, while in the wake region, the obvious differences can be noticed. The adaptive control method better simulates the process of shedding vortices, which is expected to be smooth without noise according to the previous works.

If more detailed observation is taken on the boundary layer, significant errors will be noticed in the boundary layer calculated by the spring-smoothing method. Because of the limitation of unstructured grids on the simulation of the flow field in near-wall region, the accurate flow field can hardly be achieved. The shape of grid cells has evident influences on the contours, which lead to the uncertainty of computation.

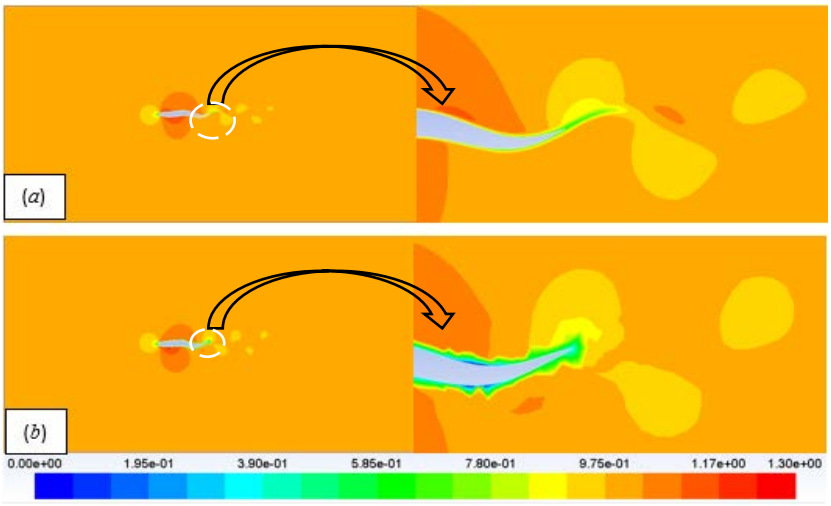


Figure 11: Velocity contours based on adaptive control method (*a*) and spring-smoothing method (*b*) method

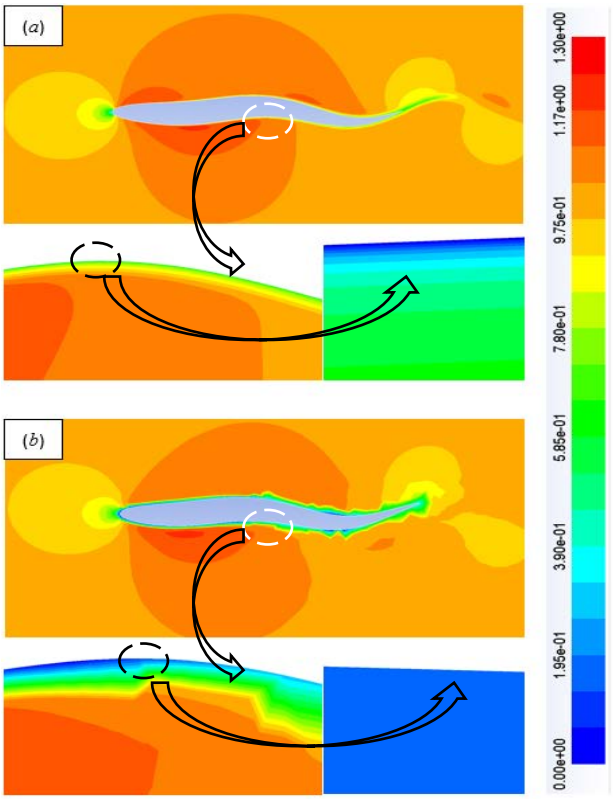


Figure 12: Boundary layer based on adaptive control method (*a*) and spring-smoothing method

The velocity contours illustrate that a row of low velocity zones appear in the wake of fish undulating. With the help of accuracy computational results based on the adaptive control method, the complete process of the shedding vortices can be simulated and researched. Fig. 13 illustrates the morphology of the wake and the flow adjacent to the fish body. The results are also compared with previous experimental and numerical research. The vorticity distributions are correspondent to the experimental flow field of a swimming eel visualized by

PIV[11]. The wake structure resembles a reverse von Karman vortex street. The numerical simulation of Chen [10] and Mattia[7] with different methods came to the similar conclusion.

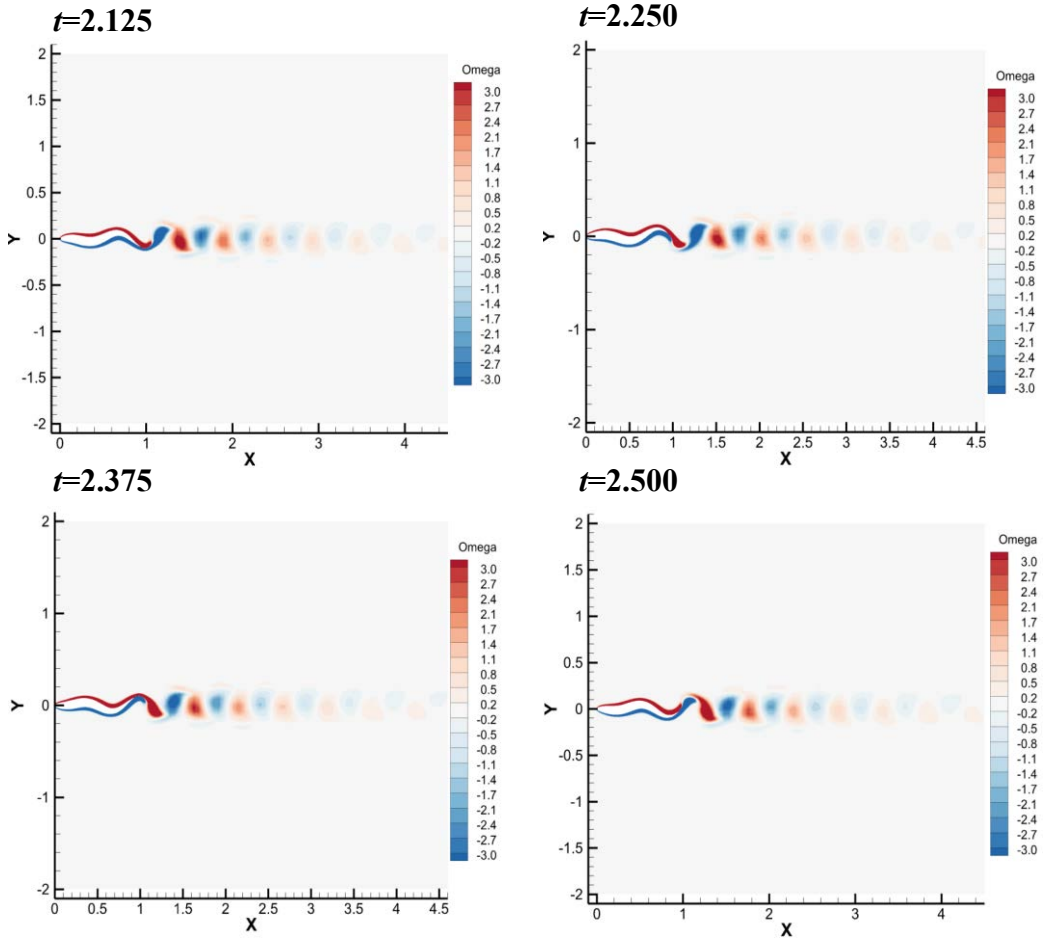


Figure 13: Vorticity contour of the wake flow field by adaptive control dynamic grids method

Concluding Remarks

Numerous dynamic-grids methods have been proposed by previous researchers for the flow fields with moving and deforming boundary. But these methods are not able to solve this problem perfectly. In the paper, the adaptive control dynamic-grids method is compared with several methods by being adopted in a fish-like undulation flow field

The method is based on structured grids, which can achieve higher grid quality in the near-wall region and the wake. The deterioration of grid quality can be avoided with the solution processing. There is no interpolation error, which is inevitable when using overset grid method. Some frequent problems occurred in previous dynamic-grids methods, including grids intersection and quality deterioration, are solved efficiently. Especially in the boundary layer region, the quality of near-wall structured grids can be guaranteed, which is vital to match up with advanced turbulence models. The comparisons have demonstrated advantages of the new methods in grids quality, the computational cost and the simulation accuracy. These advantages make it an efficient alternative dynamic-grids method to assist with related research.

References

[1]. Cavallo, P., Hosangadi, A., Lee, R., Dash, S., Cavallo, P., Hosangadi, A. & Dash, S. (1997). Dynamic unstructured grid methodology with application to aero/propulsive flowfields. In 15th Applied Aerodynamics Conference (p. 2310).

[2]. Batina, J.T., Implicit flux-split Euler schemes for unsteady aerodynamic analysis involving unstructured dynamic meshes. Aiaa Journal, 2012.

[3]. Blom, F.J., Considerations on the spring analogy. International Journal for Numerical Methods in..., 2015.

[4]. Bottasso, C.L., D. Detomi and R. Serra, The ball-vertex method: a new simple spring analogy method for unstructured dynamic meshes. Computer Methods in Applied Mechanics and Engineering, 2005. 194(39-41): p. 4244-4264.

[5]. Pérez, M.G., E. Vakkilainen and T.H. Nen, 2D dynamic mesh model for deposit shape prediction in boiler banks of recovery boilers with different tube spacing arrangements. Fuel, 2015.

[6]. Thompson, J. F., Soni, B. K., & Weatherill, N. P. (Eds.). (1998). Handbook of grid generation. CRC press.

[7]. Gazzola, M., et al., Simulations of single and multiple swimmers with non-divergence free deforming geometries. Journal of Computational Physics, 2011.

[8]. XU, Y. and D. WAN, Numerical simulation of fish swimming with rigid pectoral fins. Journal of Hydrodynamics, Ser. B, 2012. 24(2): p. 263-272.

[9]. Thompson, Joe F., Frank C. Thames, and C. Wayne Mastin. "Automatic numerical generation of body-fitted curvilinear coordinate system for field containing any number of arbitrary two-dimensional bodies." Journal of computational physics 15.3 (1974): 299-319..

[10]. Chen Z, Doi Y. Numerical Study on Propulsion by Undulating Motion in Laminar-Turbulent Flow [C]. 24th Symposium on Naval Hydrodynamics, Fukuoka, Japan, July 8-13, 2002..

[11] Müller U K, Smit J, Stamhuis E J, et al. How the body contributes to the wake in undulatory fish swimming[J]. Journal of Experimental Biology, 2001, 204(16): 2751-2762.

Optimizing the Geometric Parameters of Cutting Edge for Finishing Machining of 30Cr2Ni4MoV Alloy Steel

***F. Jiang[1], T.K. Liao[1], H. Xie[3], F.Z. Wang[1] and L. Yan[2]**

[1] Institute of Manufacturing Engineering, Huaqiao University, China.
[2] Machine Tool Factory, Dongfang Turbine Co., Ltd., China.
[3] College of Mechanical Engineering and Automation, Huaqiao University, China.

* jiangfeng@hqu.edu.cn

Abstract

In order to optimize the geometric parameters of cutting edge for finishing machining of 30Cr2Ni4MoV (30Cr) alloy steel, a 2D finite element (FE) model of orthogonal cutting has been built with FE software AdvantEdge. The optimized methodology of cutting edge geometric parameters has been proposed based on simulated results. Then, the geometric parameters of cutting edge have been optimized based on a comprehensive criterion combining chip deformation coefficient and tool stress. The chip deformation coefficient indirectly determines the surface roughness and tool stress determines tool wear, which affect the dimensional precision of parts. The range of rake angle is selected from 12° to 20° and the range of cutting edge radius is selected from 12μm to 20μm in the optimization process. The optimal rake angle for finishing machining 30Cr alloy steel is 16° and the optimal cutting edge radius is 14μm with given relief angle of 7°.

Keywords: Numerical simulation, Finishing machining, Chip deformation coefficient, Tool stress, Geometric parameters of cutting edge

Introduction

The deformation of workpiece material mainly depends on the cutting edge geometries, including rake angle, relief angle and edge radius which have great effect on flow stress, chip morphology and machined surface quality. Metal machining process is generally divided into rough and finishing machining. Rough machining pays more attention to the tool life while finishing machining pays more attention to the machining surface quality and tool wear.

Theoretical calculation or cutting experiments is difficult to determine the optimal cutting tool geometry quickly and effectively due to the complexity and instability of the cutting process. The machining process is thermo-mechanical coupled process. The finite element method (FEM) has been proven to be a useful tool to analyze the metal cutting process [1, 2] and optimize the process parameters [3, 4]. Many researchers focus on the optimization of cutting edge geometric parameters based on the FEM. In the previous research of our team, X. Cheng et al. [5] built the FE model to optimize the geometric parameters of cutting edge for rough machining of Fe-Cr-Ni stainless steel. The optimized methodology of cutting edge geometric parameters has been proposed, and the optimized methodology is based on the stress, and the cutting parameters are designed based on the equal material removal rate. Keyvan and E.Ng [6] used a combined empirical-numerical (FE) approach for predicting the tool life. This approach is based on the similarities found among the worn cutting edge geometries which have been obtained from the orthogonal tool life tests at different cutting speeds.

However, few studies have used the numerical simulation technology to optimize the parameters of cutting processes or cutting tools for machined surface quality. The main reason

is that, the mesh size and remesh process or inputting unstable factors such as material failure of FEM model affect the formation process of machined surface in the cutting simulation. But the indirect indexes, chip deformation coefficient, obtained from numerical simulation analysis could be used in the optimization of machined surface quality. Related research pointed out that the chip morphology indirectly reflects the machined surface quality. The smaller chip deformation coefficient is, the better machined surface quality is. Su. G et al. [7] have found the correlations between chip morphology and machined surface micro-topography at different chip serration stages encountered in high speed cutting, and get a conclusion that the principal factor influencing surface roughness is the thickness of the sawed segment (tooth) of saw-tooth chip. Schultheiss et al. [8] have present a new model for predicting the surface roughness during turning operations, while the influence of the minimum chip thickness on the obtained surface roughness have been analyzed.

In this paper, a 2D FE model of orthogonal cutting process of 30Cr alloy steel has been built with commercial software Third Wave AdvantEdge. The temperature and stress and chip thickness under different cutting edge geometric parameters have been obtained. The influence of rake angle and cutting edge radius on the stress and chip deformation coefficient has been analyzed. A comprehensive criterion has been used in this paper, which is combining the chip deformation coefficient and stress. The chip deformation coefficient could reflect the machined surface quality while the tool stress could reflect the tool wear. Through this criterion, the rake angle and cutting edge radius of finishing machining 30Cr alloy steel have been optimized.

Nomenclature

f	feed rate	γ_0	rake angle
m	the strain rate hardening coefficient	ε_0	reference strain
n	the strain hardening coefficient	ε_s	actual stress
T	current deformation temperature	ρ_s	the density of specimen
v	cutting speed	σ_0	yield stress at reference strain
σ	tool stress	σ_s	flow stress
Λ	the comprehensive coefficient	ζ_a	the deformation coefficient
a_c	undeformed thickness	ξ_f	the factor of feed rate determine minimum element size
a_{ch}	chip thickness	ξ_r	the factor of cutting edge radius determine minimum element size
a_p	depth of cut	ΔT	the adiabatic temperature rise
c_0-c_3	the coefficient of heat softening	$\bar{\sigma}$	average of tool stress
C_1 C_2	the coefficient of normalization	S'_f	feed rate to determine minimum element size
C_p	the specific heat capacity of workpiece	S'_{min}	preset minimum element size
r_e	cutting edge radius	S'_r	cutting edge radius to determine minimum element size
S_{max}	maximum element size	$\dot{\varepsilon}_0$	reference strain rate
S_{min}	actual minimum element size	$\dot{\varepsilon}_s$	actual stress rate
T_{ini}	the initial deformation temperature	$\bar{\zeta}_a$	average of deformation coefficient
α_0	relief angle		

Model of orthogonal cutting 30Cr alloy steel

30Cr alloy steel used in the turbine rotor of the steam turbine and other key industrial areas. After hardening and tempering methods or high-frequency surface hardening, the comprehensive mechanical properties have been improved. The hardness is about 270HV.

The chemical compositions of 30Cr alloy steel are listed in Tab. 1. The microstructure of the material is shown in Fig. 1.

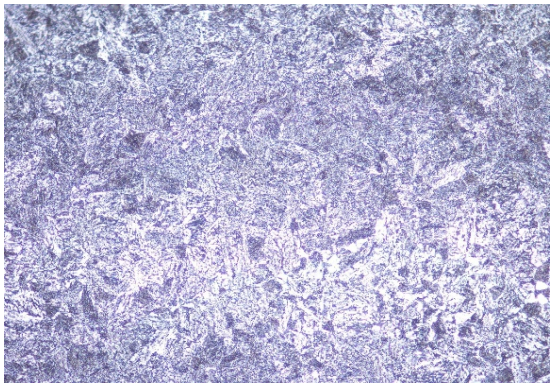


Figure 1. Microstructure of 30Cr alloy steel (500×)

Table 1. Chemical compositions of 30Cr alloy steel (wt.%)

Composition	C	Si	Mn	P	S	Ni	Cr	Mo	V	Fe
Content (wt.%)	≤0.35	≤0.03	≤0.05	≤0.004	≤0.002	1.6	3.5	0.3	0.08	Bal.

Cutting parameters of finishing machining

In the practical machining process, the cutting speed, feed rate and depth of cut could be varied in a reasonable range. It is essential to optimize the rake angle and cutting edge radius considering different cutting parameters with limited range. Lower cutting depth and feed and higher cutting speed have been generally used in finishing machining process. In order to obtain the optimal value of rake angle and cutting edge radius with a good adaptability, the process parameters are set in a reasonable range. The machining parameters are listed in the Tab. 2.

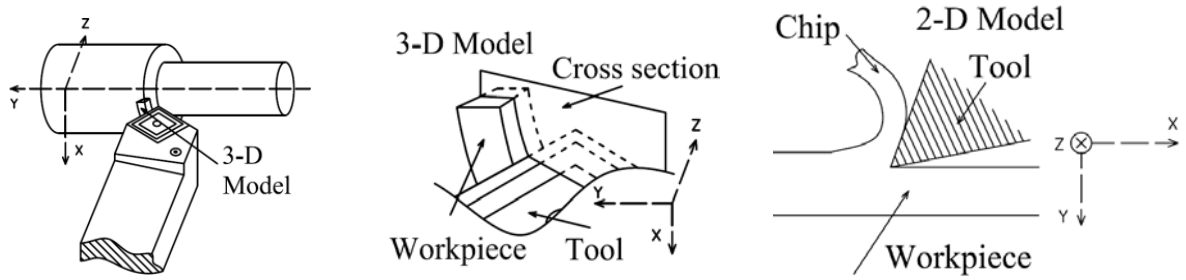
Table 2. Finishing machining cutting parameters

a_p (mm)	f (mm/rev)	v (m/min)	Number (No.)
1	0.1	100	1
		140	2
		180	3
	0.15	100	4
		140	5
		180	6
	0.2	100	7
		140	8
		180	9

Cutting speed is set at 3 levels and feed rate is also set at 3 levels. The depth of cut is set as 1mm in all simulations. In this paper, the single-factor design of experiment has been used, so there are 9 combinations of finishing machining cutting parameters.

The simplified model and meshing control

Simulations have been performed with AdvantEdge which integrates advanced FE models appropriate for machining operations. Assuming the workpiece material is isotropic, the material is removed in same state at all times. Besides, each point of cutting edge is equivalent in the direction of cutting depth.



(a) turning process (3D) (b) 3D model (c) 2D plane strain problem

Figure 2. A complex 3D problem transforms into a simple 2D plane strain problem

Therefore, a complex three-dimensional (3D) problem, as shown in Fig.2 (a) and (b), transforms into a simple two-dimensional (2D) plane strain problem as shown in Fig.2 (c). In this model, cutting tool is regarded as rigid and applied by fully constrained in a non-contact surface. In the 2D plane strain model, the workpiece moves in X direction with the cutting speed while the tool is fixed, and the feed direction translate into Y direction, as shown in Fig.2 (c).

The deformation of the material will cause the mesh distortion during the process of FE simulation. The adaptive FE mesh is the important features of the AdvantEdge software, therefore, it not only solve the above problems, but also improve the accuracy for the edge region after refinement processing and ensures the calculation efficiency for the region away from the cutting edge after coarsening processing.

Mesh grading (0.1~1.0) determine the nature of the transition from fine elements near the cutting edge to coarse elements away from the cutting edge. Increasing this parameter will result in a more coarsen mesh. In this study, the mesh grading is set to 0.4.

The workpiece meshing parameters, the mesh coarsening factor and the mesh refinement factor are set to 6 and 2 respectively; S_{max} is set to 0.1mm, and S'_{min} is set to 0.02mm; S'_f defines the smallest element length through the chip thickness, thus, the factor ξ_f setting in the study is use as default 0.1; S'_r defines the element length in the vicinity of the cutting edge. In order to keep minimum element size of all models be built by the same standard, the factor ξ_r setting in the study is 0.4. In this way, the minimum element size of all models is determined by the cutting edge radius. The FE model with mesh is shown in Fig. 3.

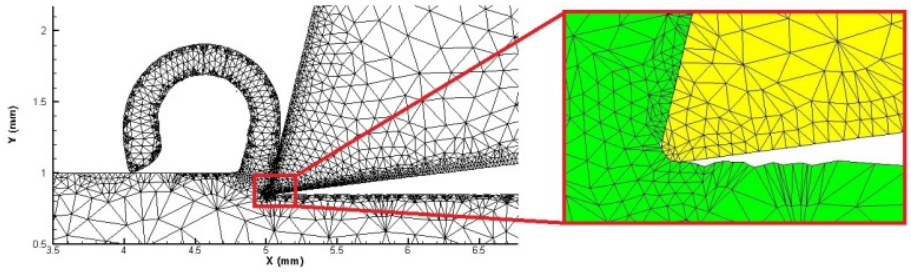


Figure 3. Finite element model with mesh

The method of how to ensure the minimum element size of all models with the same standard is as follows: for example, there are 3 kinds of feed rate (f_1, f_2, f_3) and 6 kinds of cutting edge radius (r_{e1}, r_{e2}, r_{e3}, r_{e4}, r_{e5}, r_{e6}) in this study. In order to keep the minimum element size of all models be determined by the cutting edge radius, it should be pay more attention to adjusting the value of ξ_r to make the S_{min} equal to S'_r.

$$S'_r=\xi_r \cdot max(r_{e1}, r_{e2}, r_{e3}, r_{e4}, r_{e5}, r_{e6}) \quad (1)$$

$$S'_f=\xi_f \cdot min(f_1, f_2, f_3) \quad (2)$$

$$S_{min}=min(S'_r, S'_f, S'_{min}) \quad (3)$$

Material constitutive model and friction mode

The numerical simulation of metal cutting involves many models, such as friction model [9], material constitutive model [10], heat transfer model and thermal physical property model [11]. Material constitutive model has been modified continuously to form the thermal-mechanical coupling phenomenon and the iterative process of the algorithm. When the thermal-mechanical coupling reaches the equilibrium and the cutting simulation process reaches the steady state, the system output the final results which include the cutting stress, chip deformation and other physical quantities. It could be seen that the key models for the numerical simulation of metal cutting is the material constitutive model.

Material constitutive model is a set of equations that reflect the stress-strain relationship of the deformable body material [12]. At present, the spilt Hopkinson pressure bar (SHPB) technique is usually used to determine the flow stress and strain data in a certain range of strain rate and temperature [13]. According to these data and the corresponding empirical formula, the constitutive equation is established. There are many types of constitutive equations of materials [14], power-law (P-L) constitutive model has been introduced the strain hardening, strain rate hardening and thermal softening parameters, which is a comprehensive reflection of large strain, high strain rate and high temperature load [15]. At the same time, it is simple in its own form and applied to a variety of computer encoding, so it is a practical model for analysis and calculation of. The P-L constitutive model is expressed as [16]:

$$\sigma_s=\sigma_0\left(1+\frac{\varepsilon_s}{\varepsilon_0}\right)^{1/n}\times\left(1+\frac{\dot{\varepsilon}_s}{\dot{\varepsilon}_0}\right)^{1/m}\times\left(c_0+c_1T+c_2T^2+c_3T^3\right) \tag{4}$$

$$T=T_{ini}+\Delta T \tag{5}$$

$$\Delta T=\frac{1}{\rho_s C_p}\int\sigma_s \mathrm{d}\varepsilon_s \tag{6}$$

The SHPB technique has been used to study the dynamic deformation behavior of 30Cr alloy steel. The true stress-true strain curves of workpiece material have been obtained. The P-L model constitutive model is given as the following [17]:

$$\sigma_s=547.92\left(1+\frac{\varepsilon_s}{0.01}\right)^{1/8.9047}\times\left(1+\frac{\dot{\varepsilon}_s}{100}\right)^{1/6.9979}\times(1.0308-1.8124e^{-3}T+2.1826e^{-6}T^2-1.2745e^{-9}T^3 \tag{7}$$

σ_s (MPa), $\dot{\varepsilon}_s$ (s^{-1}), T (℃).

Comprehensive criterion combining chip deformation coefficient and tool stress

AdvantEdge utilizes Tecplot software to display and assist in analyzing simulation results. The average stress in the tool can be computed and obtained easily. The stress history is average over the highest 10% of the elements in the cutting tool, as show in the Fig. 4. Then, the average of stress history which has achieved steady state is calculated as a quantitative data.

However, the quality of machined surface is difficult to be observed directly from the 2D FE model, and the surface quality has great relationship with chip deformation coefficient in the finishing machining process. Generally small chip deformation coefficient means small material deformation in the cutting process, which result in better surface quality. Thus, the chip deformation coefficient is proposed as an indirect index to evaluate the machined surface quality. The deformation coefficient of workpiece can be calculated as follows:

$$\zeta_a=\frac{a_{ch}}{a_c} \tag{9}$$

The ac is the undeformed thickness, which is equal to the feed rate in the 2D FE model. Chip thickness is obtained by measuring the radius of a circle which the center at the inner side of the chip, and the circle are tangent to the outer side of the chip. In order to ensure the accuracy of the measured chip thickness, the chip thickness is measured 6 times and the average value is obtained, as show in the Fig. 5.

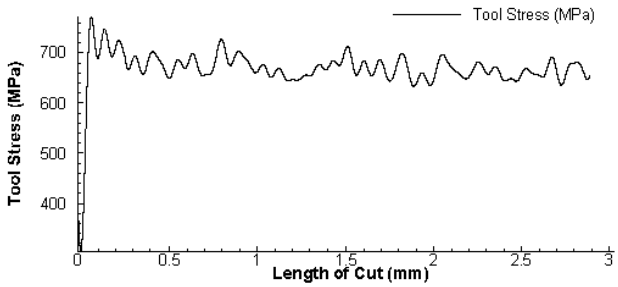


Figure 4. Average over the highest 10% of the elements in the cutting tool

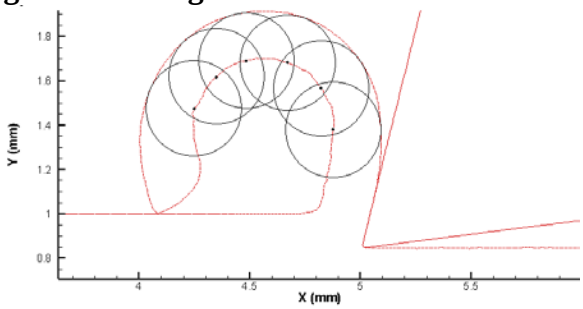


Figure 5. Schematic illustration of measuring chip thickness

The comprehensive criterion combining deformation coefficient and stress is expressed as:

$$\Lambda = C_1 \frac{\zeta_a}{\bar{\zeta}_a} + C_2 \frac{\sigma}{\bar{\sigma}} \tag{10}$$

The determination of C_1, C_2, $\bar{\zeta}_a$ and $\bar{\sigma}$ is as follow:

Assuming the simulation results of different cutting edge radius with one rake angle are show in the Tab. 3. The method of comprehensive criterion to optimize the cutting edge radius as follows:

The values of $\Lambda_1 \sim \Lambda_5$ are compared, and the minimum value of the corresponding cutting edge radius is the optimal under this rake angle. The method of comprehensive criterion to optimize the rake angle is the same as mentioned above.

Table 3. Simulation results of different cutting edge radius with one rake angle

cutting edge radius(μm)	12	14	16	18	20
stress(MPa)	σ_1	σ_2	σ_3	σ_4	σ_5
deformation coefficient	ζ_{a1}	ζ_{a2}	ζ_{a3}	ζ_{a4}	ζ_{a5}

$$\bar{\sigma} = \frac{\sigma_1+\sigma_2+\sigma_3+\sigma_4+\sigma_5}{5} \tag{11}$$

$$\bar{\zeta}_a = \frac{\zeta_{a1}+\zeta_{a2}+\zeta_{a3}+\zeta_{a4}+\zeta_{a5}}{5} \tag{12}$$

$$\Delta\sigma = \max\left(\frac{\sigma_n}{\bar{\sigma}}\right) - \min\left(\frac{\sigma_n}{\bar{\sigma}}\right) \tag{13}$$

$$\Delta\zeta_a = \max\left(\frac{\zeta_{an}}{\bar{\zeta}_a}\right) - \min\left(\frac{\zeta_{an}}{\bar{\zeta}_a}\right) \tag{14}$$

$$C_1 = \frac{\Delta\sigma}{\Delta\sigma+\Delta\zeta_a} \tag{15}$$

$$C_2 = \frac{\Delta\zeta_a}{\Delta\sigma+\Delta\zeta_a} \tag{16}$$

$$\Lambda_n = c_1 \frac{\zeta_{an}}{\bar{\zeta}_a} + c_2 \frac{\sigma_n}{\bar{\sigma}} \qquad n=1,2,3,4,5 \tag{17}$$

Optimization procedures

According to the characteristics of geometric parameters of cutting tool, the optimization procedures are proposed. Relief angle (α_0) impacts the contact length between the machined surface and flank face of tool. A reasonable relief angle for finishing machining is 5°~8° [18]. According to the carbide indexable inserts design standards [19], relief angle of 7° is select. In the process of finishing, the cutting tool wear and machined surface quality are more important. Larger rake angle (γ_0) can get a better machined surface quality, while excessively rake angle not only weakens the cutting edge strength, but also reduce the cutting edge of the volume and lead to the increase of cutting tool wear, which lower the dimensional precision. At the same time, the deformation coefficient decreases with the increase of the rake angle. Therefore, it is an optimal rake angle to balance the tool stress and the deformation coefficient. Larger cutting edge radius (r_e) gets a smaller tool stress but increases the chip deformation coefficient, while smaller cutting edge radius gets smaller chip deformation coefficient but increases the tool stress. Thus, it is an optimal cutting edge radius to balance the tool stress and the deformation coefficient for the specific process parameters and specific materials.

According to reference [17], 12°~20° is the reasonable range for rake angle and 12μm~20μm is the reasonable range for cutting edge radius in the finishing machining process. The optimization steps are shown in ANNEX A, and the numbers of cutting parameter are shown in Tab. 2.

Step 1: The optimization process is divided into two parts, one is the optimization of the rake angle, and the other is the optimization of the cutting edge radius. The two are respectively independent and not related to each other.

Step 2: 9 sets of cutting parameters are mentioned above. In each set of cutting parameters, all of the rake angle and cutting edge radius should be carried crossover. The rake angle is changed from 12° to 20° with interval of 2° and a total of 5 values (12°, 14°, 16°, 18°, 20°) are selected; the cutting edge radius is changed from 12μm to 20μm with interval of 2μm and a total of 5 values (12μm, 14μm, 16μm, 18μm, 20μm) are selected. Therefore, there are 225 sets of simulations to be performed.

Step 3: 2 optimal rake angles are obtained according to the comprehensive criterion of 225 sets of simulations, while 2 optimal cutting edge radiuses are also obtained. Finally, the 4 sets of combination with 2 optimal rake angles and 2 optimal cutting edge radiuses are obtained.

Step 4: The optimal combination of cutting edge geometric parameters is obtained through the comprehensive criterion which based on the above 4 sets of combination.

The simulated results

Summary of optimal cutting edge geometric parameters

A total of 45 optimal cutting edge radiuses are obtained in 5 different rake angles while based on 9 sets of cutting parameters in the Tab. 4. The value of each column is obtained by the fixed rake angle and changed cutting edge radius with 9 sets of cutting parameter based on the

comprehensive criterion. Then the total number obtained by summing the number of each cutting edge radius, and the total number of the first two will be elected to alternative area.

Table 4. Summary of optimal cutting edge radius

Number as optimal value		Rake angle 12°	14°	16°	18°	20°	Total(45)
Cutting edge radius	12μm	4	5	4	4	3	20
	14μm	3	2	2	3	3	13
	16μm	1	0	2	1	1	5
	18μm	1	0	0	0	0	1
	20μm	0	2	1	1	2	6

A total of 45 optimal rake angles are obtained in 5 different cutting edge radiuses while based on 9 sets of cutting parameters in the Tab. 5. The value of each line is obtained by the fixed cutting edge radius and changed rake angle with 9 sets of cutting parameter based on the comprehensive criterion. Then the total number obtained by summing the number of each rake angle, and the total number of the first two will be elected to alternative area.

Table 5. Summary of optimal rake angle

Number as optimal value		Rake angle 12°	14°	16°	18°	20°
Cutting edge radius	12μm	0	4	5	0	0
	14μm	0	4	5	0	0
	16μm	0	1	4	3	1
	18μm	1	3	4	1	0
	20μm	0	0	7	0	2
Total (45)		1	12	25	4	3

It can be seen obviously in Tab. 4. and Tab. 5. In a series of cutting edge radius, the number of 12μm and 14μm are higher, and the number of the 14° and 16° in a series of rake angles are higher, too. Therefore, 4 sets of optimization combination are obtained by the cross combination of 2 rake angle and 2 cutting edge radius. These 4 combinations are as follows: 14°-12μm, 14°-14μm, 16°-12μm, 16°-14μm (fixed relief angle is 7°)

The optimal combination of cutting edge geometric parameters was determined

Table 6. New addition of machining cutting parameters

a_p (mm)	f (mm/rev)	v (m/min)	Number (No.)
1	0.08	120	10
		160	11
		200	12
	0.13	120	13
		160	14
		200	15
	0.18	120	16
		160	17
		200	18

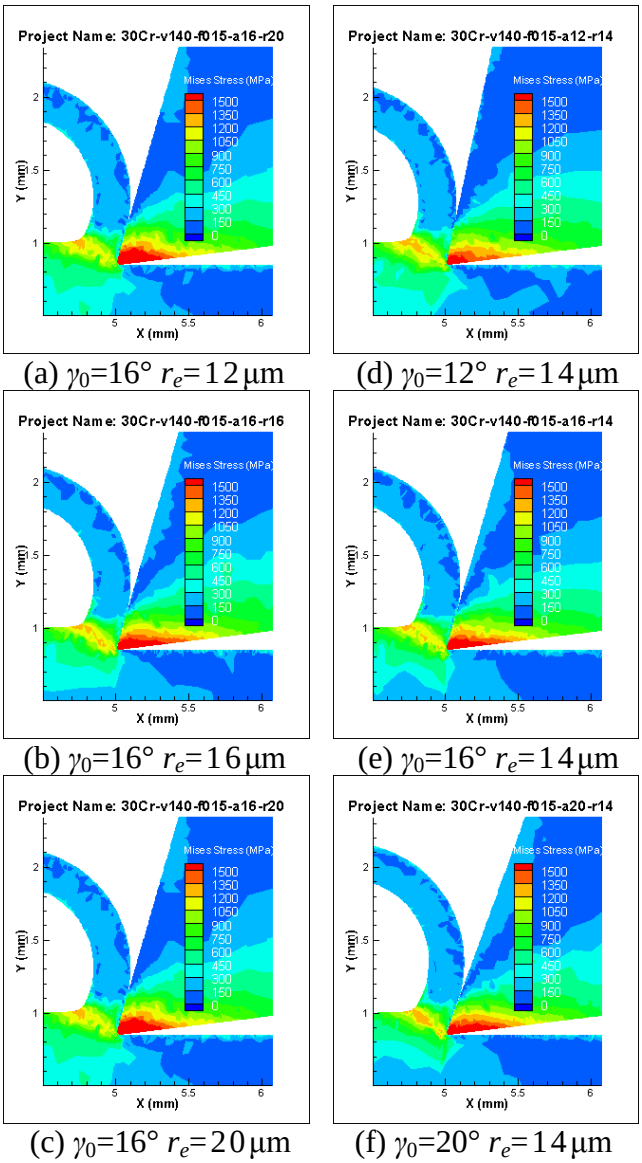
In order to ensure the correctness of the optimization results, 3 sets of feed rate (0.08mm/rev, 0.13mm/rev, 0.18mm/rev) and 3 sets of cutting speed (120m/min, 160m/min, 200m/min) are added to carry out the additional simulation experiment for these 4 combinations. Thus, a total of 36 new simulation experiments are given. The new machining cutting parameters are list in the Tab. 6.

Then, a total of 4 combinations of rake angle and cutting edge radius are obtained under 18 sets of cutting parameters respectively and the total number of the first will be elected to be the optimal combination of cutting edge geometric parameters. The 4 combinations of rake angle and cutting edge radius are listed in the Tab. 7. It can be seen obviously that the third combination, 16°-14μm, is the optimal.

Analysis on contours of temperature and stress

Several combinations of cutting edge geometric parameters were taken as examples to analyze the contours of temperature and stress.

As shown in Fig.6, it can be seen obvious change in the stress distribution of the tool with the increase of rake angle and little change with the increase of cutting edge radius. Increasing rake angle results in the highest values of stress expand to the flank face of insert. But the stress of the primary deformation is increasing with the increase of cutting edge radius.



(a) γ_0=16° r_e=12μm (d) γ_0=12° r_e=14μm

(b) γ_0=16° r_e=16μm (e) γ_0=16° r_e=14μm

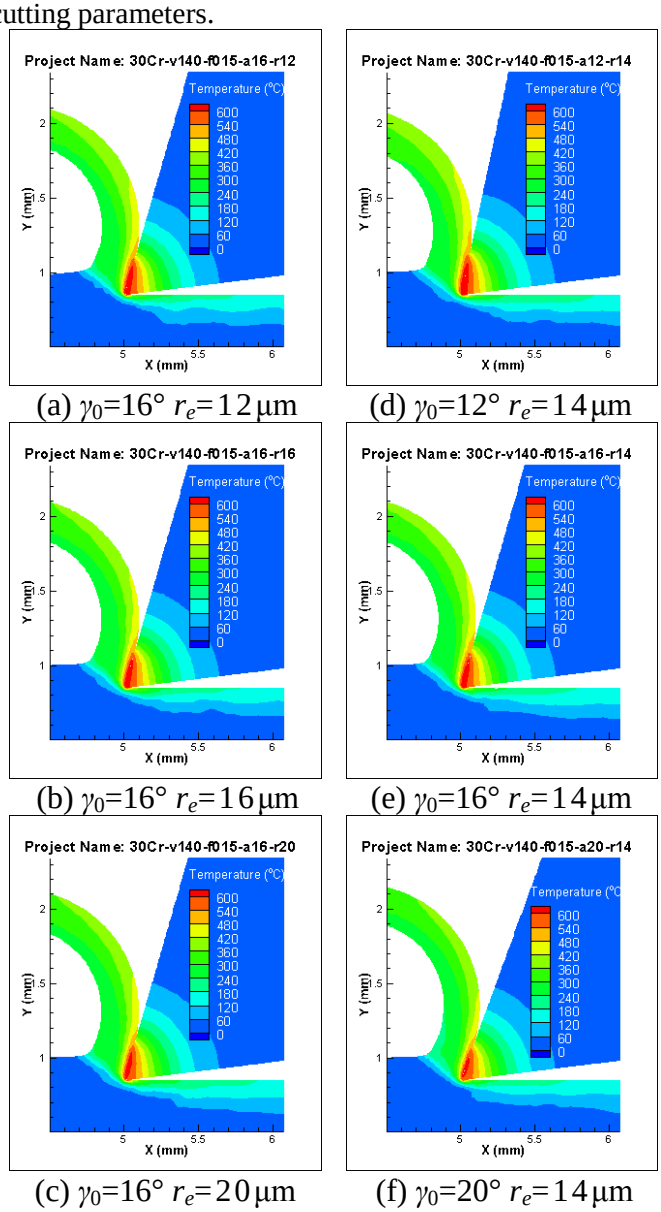
(c) γ_0=16° r_e=20μm (f) γ_0=20° r_e=14μm

Figure 6. The contours of stress in different combination (v=140m/min f=0.15mm/rev a_p=1mm α_0= 7°)

Table 7. Optimal combination of rake angle and cutting edge radius

4 combinations	The number of cutting parameters																		Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
14°-12μm				*										*				*	2
14°-14μm																			0
16°-12μm					*	*	*						*			*	*	*	7
16°-14μm	*	*	*					*	*	*	*	*			*				9

The meaning of "*" in the Tab.8 is that the combination of this cutting edge radius and rake angle is optimal for this set of cutting parameters.



(a) γ_0=16° r_e=12μm (d) γ_0=12° r_e=14μm

(b) γ_0=16° r_e=16μm (e) γ_0=16° r_e=14μm

(c) γ_0=16° r_e=20μm (f) γ_0=20° r_e=14μm

Figure 7. The contours of temperature in different combination (v=140m/min f=0.15mm/rev a_p=1mm α_0= 7°)

As shown in Fig. 7 there is no obvious change in the temperature distribution of the tool and primary deformation zone with the change of the rake angle and cutting edge radius.

Conclusions

In this paper, a 2D FE model of orthogonal cutting 30Cr alloy steel has been built to analyze the deformation coefficient and stress under different cutting edge geometric parameters. The

rake angle and cutting edge radius for finishing machining Fe-Cr-Ni alloy steel have been optimized. The conclusions are as follows:

In the finishing machining 30Cr alloy steel, the cutting edge geometric parameters have greater influence on stress than temperature.

The optimized methodology of cutting edge geometric parameters has been proposed which is a comprehensive criterion combining chip deformation coefficient and tool stress.

In the finishing machining 30Cr alloy steel, the optimal rake angle 16°, and the optimal cutting edge radius is 14μm when the relief angle is fixed 7°. The correctness of the optimized results can be guaranteed by a uniform comprehensive criterion and multiple sets of validation experiments.

Acknowledgments

This research is sponsored by the National Natural Science Foundation of China (No.51475173), the Major Project on the Integration of Industry and Research of Fujian Province (No.2014H6018), the Research and Innovation Ability of Graduate Students in Huaqiao University (No.1511403008).

References

[1] Atlati, S., B. Haddag, M. Nouari, and M. Zenasni. (2014). Thermomechanical modelling of the tool-workmaterial interface in machining and its implementation using the ABAQUS VUINTER subroutine. *INT J MECH SCI* **87:**102-117.
[2] Ben Moussa, N., H. Sidhom, and C. Braham. (2012). Numerical and experimental analysis of residual stress and plastic strain distributions in machined stainless steel. *INT J MECH SCI* **64:**82-93.
[3] Bharilya, R. K., R. Malgaya, L. Patidar, R. K. Gurjar, and A. K. Jha. (2015). Study of Optimised Process Parameters in Turning Operation Through Force Dynamometer on CNC Machine. *Materials Today: Proceedings* **2:**2300-2305.
[4] Krol, T. A., C. Seidel, and M. F. Zaeh. (2013). Prioritization of Process Parameters for an Efficient Optimisation of Additive Manufacturing by Means of a Finite Element Method. *Procedia CIRP* **12:**169-174.
[5] Cheng, X., X. Zha, and F. Jiang. (2016). Optimizing the geometric parameters of cutting edge for rough machining Fe-Cr-Ni stainless steel. *The International Journal of Advanced Manufacturing Technology* **85:** 683-693.
[6] Hosseinkhani, K., and E. Ng. (2015). A Combined Empirical and Numerical Approach for Tool Wear Prediction in Machining. *Procedia CIRP* **31:**304-309.
[7] Su, G., Z. Liu, L. Li, and B. Wang. (2015). Influences of chip serration on micro-topography of machined surface in high-speed cutting. *International Journal of Machine Tools and Manufacture* **89:**202-207.
[8] Schultheiss, F., S. R. H Gglund, V. Bushlya, J. Zhou, and J. St Hl. (2014). Influence of the Minimum Chip Thickness on the Obtained Surface Roughness During Turning Operations. *Procedia CIRP* **13:**67-71.
[9] Taureza, M., X. Song, and S. Castagne. (2014). Depth-dependent stress-strain relation for friction prediction. *INT J MECH SCI* **86:**46-53.
[10] Jiang, F., L. Yan, and Y. Rong. (2013). Orthogonal cutting of hardened AISI D2 steel with TiAlN-coated inserts simulations and experiments. *The International Journal of Advanced Manufacturing Technology* **64:**1555-1563.
[11] Jiang, F., T. Zhang, and L. Yan. (2016). Estimation of temperature-dependent heat transfer coefficients in near-dry cutting. *The International Journal of Advanced Manufacturing Technology.* **86:**1207-1218.
[12] Hu, W., Y. Lin, S. Yuan, and Z. He. (2015). Constitutive models for regression of various experimental stress-strain relations. *INT J MECH SCI* **101-102:**1-9.
[13] Malakizadi, A., S. Cedergren, I. Sadik, and L. Nyborg. (2016). Inverse identification of flow stress in metal cutting process using Response Surface Methodology. *SIMUL MODEL PRACT TH* **60:**40-53.
[14] Buchkremer, S., B. Wu, D. Lung, S. Münstermann, F. Klocke, and W. Bleck. (2014). FE-simulation of machining processes with a new material model. *J MATER PROCESS TECH* **214:**599-611.
[15] Jiang, F., J. Li, J. Sun, S. Zhang, Z. Wang, and L. Yan. (2010). Al7050-T7451 turning simulation based on the modified power-law material model. *The International Journal of Advanced Manufacturing Technology* **48:**881.
[16] Yu, J., F. Jiang, Y. Rong, H. Xie, and T. Suo. (2014). Numerical study the flow stress in the machining

process. *The International Journal of Advanced Manufacturing Technology* **74:**509-517.
[17] Xie, H., Yu, J., Wang, G., Rong, Y., Ye, J. and Liu, Y. (2016). Constitutive model for dynamic mechanical response of high strength steel. *Journal of Sichuan University (Nature Science Edition)* **53:**382-388.
[18] Trent, E. M., and P. K. Wright (2000). *Metal Cutting (Fourth Edition) in Book Metal Cutting (Fourth Edition) Butterworth-Heinemann Ltd; 4th Revised edition.*
[19] GB/T2076-2007 (ISO6987/1-1993). The Carbide Indexable Inserts Design Standards. *China National Standardization Management Committee*.

3D Simulation of the Defect Generation by Hydrogen at $Si - SiO_2$ Interface

***Zhaocan Ma**[1], **Jingjie Xu**[1], **Hongliang Li**[2], **Song Yu**[2], **Linbo Zhang**[1], and **†Benzhuo Lu**[1]

[1] *State Key Laboratory of Scientific and Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China*
[2] *Microsystem & Terahertz Research Center, China Academy of Engineering Physics, Chengdu, China*
* *Presenting author: zhaocanma@lsec.cc.ac.cn*
† *Corresponding author: bzlu@lsec.cc.ac.cn*

Abstract

In order to study defect generation and radiation-induced interface in semiconductor influenced by dose rate response and H_2, we provide a 3D finite element model based on Poisson-Nersnt-Planck equations to simulate the electro-diffusion process in numerical experiment. Multi-scale method is used in discretization and the restricted additive Schwarz preconditioner is applied to solve the linear system in simulation. The algorithm we establish in this work is solved by our parallel finite element software based on Parallel Hierarchical Grid with high efficiency. It is shown that the numerical results from our simulation agree well with experimental data of the devices affected by enhanced low dose-rate sensitivity in different H_2 environments.

Keywords: defect generation, $Si - SiO_2$ interface, enhanced low dose-rate sensitivity (EDLRS), multi-scale method.

Introduction

Total ionizing dose (TID) effect is well known as the results of generating charges in semiconductor devices designing for the space environment. As reported in previous work, [7] degradation of bipolar device characteristics increases at low dose rates irradiation, and this phenomenon is referred as enhanced low dose-rate sensitivity (ELDRS). ELDRS has great effect on modern silicon devices, and it is observed to be strongly related to hydrogen. For hermetically sealed integrated circuits (ICs) used in satellites and other space-based systems in radiation environments, the presence of high concentrations of H_2 in the package can enhance radiation-induced degradation, which greatly reduces their operation reliability. Bulk oxide (like SiO_2) and interface defects play an important role as charge trapping sites. When radiation induced, the SiO_2 layers of the device is charged and the interface traps N_{it} at the Si/SiO_2 interface increase (see Figure 1). Once this situation occurs, subsequent physical effects, such as the changing of threshold voltage, will disturb the working state of devices. After plenty of experiments, several models have been proposed to describe the phenomenon. [5] [11] [12] However, these models are

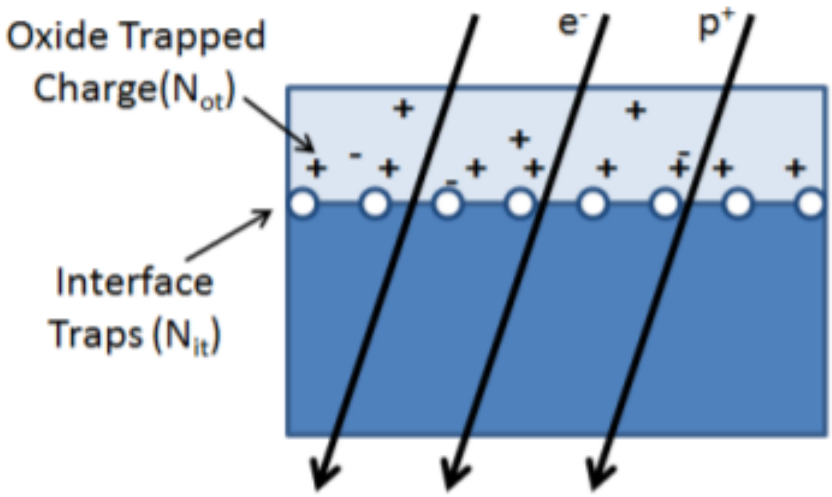


Figure 1: An illustration showing ionizing dose in SiO_2-on-Si structures.

limited in 1D or 2D cases with low efficiency [9].

In this paper, we recalculate the same Poisson-Nernst-Planck (PNP) model studied by Nicole L. Rowsey [12] with a new 3D numerical simulation, which is more complex and reasonable in describing the coupling of potential field and diffusion processes. We solve the PNP equations by finite element method, just as the numerical simulations of biochemical systems in previous works. [16][17][14] However, various simulated species along with the electrostatic voltage are considered in the model, some of which are extremely sensitive to time scale, so that the multi-scale method is an optional choice in our simulation. In addition, restricted additive Schwarz preconditioner[2] is used to solve the linear systems generated in diffusion-convection equations with high efficiency. As computations increase rapidly in 3D model compared to previous cases, we develop a parallel software to reduce the cost of simulation time. The parallel efficiency is presented in "Parallel Efficiency" section.

Drift-Diffusion Modeling

Trapping Species

Oxygen vacancies are the dominant defects in SiO_2 [1]. The neutral oxygen vacancy includes one $Si-Si$ bond instead of two $Si-O$ bonds. Because of the distinct energies, There are two detect species, detects $V_{o\delta}^+$ and $V_{o\gamma}^+$, and two neutral precursors, $V_{o\delta}$ and $V_{o\gamma}$. It is shown by observing the associated energies that $V_{o\delta}^+$ is a shallow hole trap, while $V_{o\gamma}^+$ is much deeper. These two kinds of defects can be hydrogenated or doubly hydrogenated to form $V_{o\gamma}H$, $V_{o\gamma}H^+$, $V_{o\gamma}H_2$, $V_{o\gamma}H_2^+$, $V_{o\delta}H$, $V_{o\delta}H^+$ and $V_{o\delta}H_2$, $V_{o\delta}H_2^+$, respectively.

Chemical Reactions

Here, we briefly summarize the influences of radiation-induced electrons and holes on metal-oxide-semiconductor (MOS) structures.

- Radiation emerges Electron-hole pairs (EHPs).

- The electric field pushes holes escaping initial recombination towards the interface, while electrons are towards the metal gate.
- Neutral oxygen vacancies become positively charged when capturing holes and then are neutralized by electrons.
- Protons and hydrogenated defects are created when molecular hydrogen recombines the positively charged defects.
- Protons can also be cracked by positively-charged hydrogenated defects directly.
- Interface traps are formed by the react of protons and the $Si-H$ bonds on the SiO_2/Si interface.

$$H^+ + Si - H \Leftrightarrow N_{it} + H_2 \tag{1}$$

Poisson-Nernst-Planck Equations

In this model, every kind of species participates in chemical reactions and therefore we use generation and recombination terms in Nernst-Planck equation to simulate the ion transport process. The electrostatic field is determined by applied voltage and charged species, including electron-hole pairs (EHPs) and positively charged defects, which are simulated by Poisson equation shown as follows:

$$\epsilon \nabla^2 \phi = -Q \tag{2}$$

$$Q_{SiO_2} = q(p + H^+ + V_{o\delta}^+ + V_{o\delta}H^+ + V_{o\delta}H_2^+ + +V_{o\gamma}^+ + V_{o\gamma}H^+ + V_{o\gamma}H_2^+ - n) \tag{3}$$

Electrons, holes and H^+ are charged and mobile, so the diffusion-convection process is considered. The $U_{radiation}$ is the EHP generation term [10]. During the radiation, recombination and generation of spices also arise by the chemical reactions, and they must also be accounted in the continuity equations, which we will describe later.

$$\frac{\partial n}{\partial t} = \nabla \cdot (e\mu_n nE + D_n \nabla n) + U_{radiation} + G_n - R_n \tag{4}$$

$$\frac{\partial p}{\partial t} = -\nabla \cdot (e\mu_p pE - D_p \nabla p) + U_{radiation} + G_p - R_p \tag{5}$$

$$\frac{\partial H^+}{\partial t} = -\nabla \cdot (e\mu_{H^+} H^+ E - D_{H^+} \nabla H^+) + G_{H^+} - R_{H^+} \tag{6}$$

Hydrogen is mobile, but not charged.

$$\frac{\partial H_2}{\partial t} = D_{H_2} \nabla H_2 + G_{H_2} - R_{H_2} \tag{7}$$

The defects including $V_{o\gamma}, V_{o\gamma}^+, V_{o\delta}, V_{o\delta}^+, V_{o\gamma}H, V_{o\gamma}H^+, V_{o\delta}H, V_{o\delta}H^+, V_{o\gamma}H_2, V_{o\delta}H_2, V_{o\gamma}H_2^+, V_{o\delta}H_2^+$ are not mobile with no drift nor diffusion, but they still have recombination and generation terms.

$$\frac{dT_i}{dt} = G_i - Ri \tag{8}$$

Recombination and Generation Terms

As the chemical reactions are complicate, we use recombination and generation terms to stand for chemical reactions in PNP equations, and these terms are nonzero after irradiation. Consider the bulk reactions:

$$A + B \Leftrightarrow C \qquad (Reaction \quad 0)$$
$$A + M \Leftrightarrow C + N \qquad (Reaction \quad 1)$$

For non-equilibrium conditions, each specie has recombination and generation terms, we take spice A for example, whose rates have the following form:

- the generation term of A can be described as:

$$G_A = k_{r_0} \cdot [C] + k_{r_1} \cdot [C][N]$$

- the recombination term of A can be described as:

$$R_A = k_{f_0} \cdot [A] \cdot [B] + k_{f_1} \cdot [A] \cdot [M]$$

where k_f and k_r are reaction rates; the bracket represents concentration of the species in $/cm^3$. The recombination and generation terms of other spices can be defined as the same.

Then we can add the recombination and generation terms to each continuity equation. The reaction rates k_f, k_r are formulated in 3 different cases, which can be found in Ref.[13].

Boundary Conditions

Here we describe the boundary conditions of mobile particles. On the gate contact and SiO_2/Si interface, Electron, holes and protons are allowed to move freely. For H_2 case, there are two possible choices depending on whether the edge is exposed or shielded.

1 Exposed case: H_2 concentration on the gate is constant and equals to the ambient concentration but reflection boundary condition on the interface.

2 Non-exposed case: reflection boundary condition on both edges.

For the surrounding edges of the SiO_2 bulk, all moving particles have reflection boundaries.
H^+ has to be emphasized on Si/SiO_2 interface because of the reaction:

$$H^+ + Si - H \Leftrightarrow N_{it} + H_2$$

The reverse reaction is negligible at room temperature and we can give the continuity equation of interface traps:

$$\frac{dN_{it}}{dt} = k_{int} \cdot [H^+]_{int}[SiH]_{int} \tag{9}$$

Since the $[H^+]_{int}, [SiH]_{int}$ is 2D with the unit of$/cm^2$ on the interface, we present the equation in another form as in Ref. [8].

$$\frac{dN_{it}}{dt} = \sigma_{int}\vec{J}_{H^+} \cdot \vec{n}([SiH] - [Nit]) \tag{10}$$

σ_{int} is the transformation coefficient with the unit of cm^2

Multi-scale Methods

Finite Element Method is used to spatially discrete the PNPs with generation and recombination terms and backward Euler method is applied to discrete the time step. Actually, the reaction process is an ODE problem, since some particles are sensitive about time.

Take $V_{o\gamma}H_2^+$ for example:

$$\begin{aligned}\frac{dV_{o\gamma}H_2^+}{dt} =&(1.03e-13)[V_{o\gamma}H_2][h^+]+(1.03e-19)[V_{o\gamma}][H^+]+(4.02e-21)[V_{o\gamma}^+][H_2]\\ &+(3.21e-138)[V_{o\gamma}H_2]\\ &-(4.16e+3)[V_{o\gamma}H_2^+]-(3.81e+5)[V_{o\gamma}H_2^+]-(1.90e+5)[V_{o\gamma}H_2^+]\\ &-(2.06e-07)[V_{o\gamma}H_2^+][e^-]\end{aligned} \tag{11}$$

Some parameter can be regarded as 0. And we can simplify it as:

$$\frac{\partial V_{o\gamma}H_2^+}{\partial t} = -O(10^5)V_{o\gamma}H_2^+ + C \tag{12}$$

To get this equation solved, the time step should be less than $10^{-5}s$. During the simulation, if we choose a time step that small, considering the amount of irradiation, it will take too much time at this time scale.

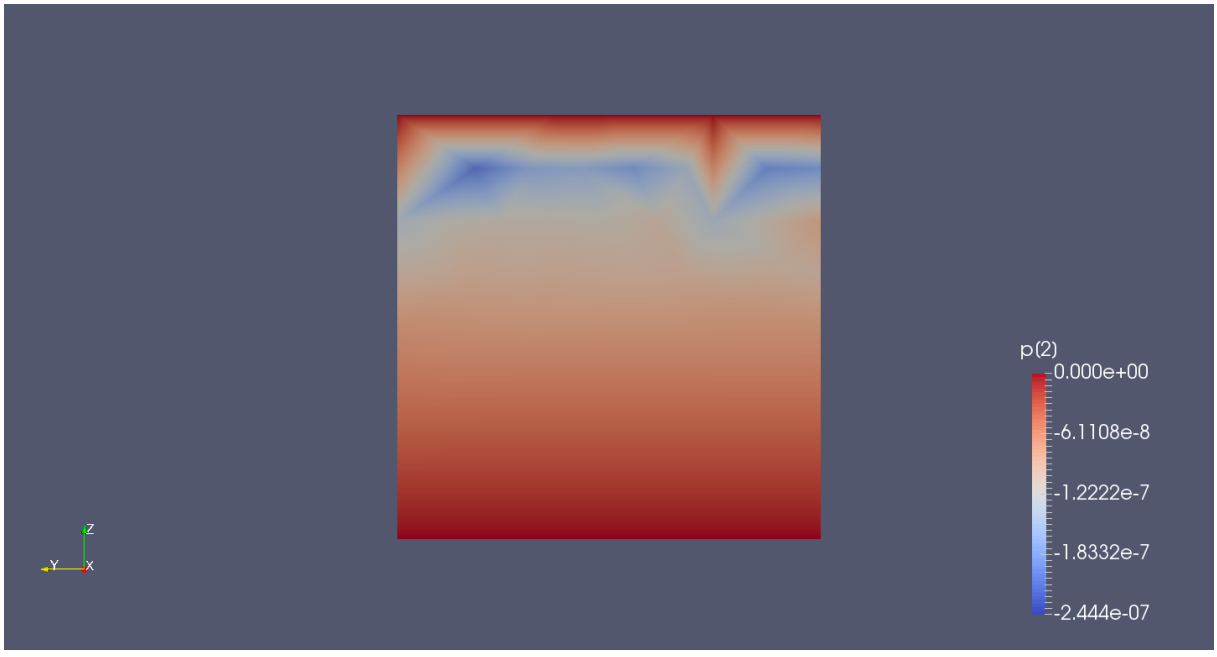


Figure 2: Distribution of H^+ without multi-scale method

So we explore a kind of multi-scale method to overcome the problem. We firstly set drift and diffusion terms as the macro-variable reaction parameters, like e^-, h^+, H^+, H_2. The detects, that can't move, treat as the micro-variable, like $V_{o\gamma}, V_{o\delta}$, then we can give the algorithm:

- Step 1. Given the current state of the reaction parameters, initialize the macro-variable of e^-, h^+ using the macro-solver with generating terms and reaction terms, and take their results as current state;

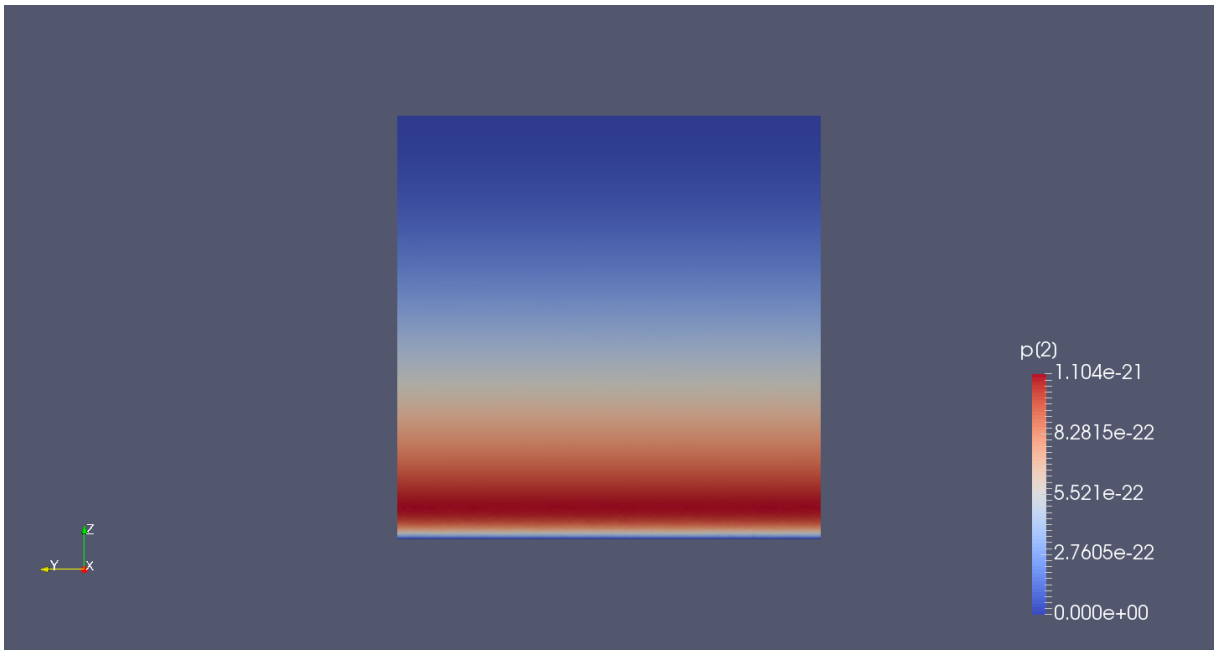


Figure 3: Distribution of H^+ with multi-scale method

- Step 2. Evolve the micro-variables like $V_{o\gamma}, V_{o_\delta}$ for M micro time steps with iteration method like Euler method using the current step of e^-, h^+ to ensure the convergence;

$$\frac{c_i^{n+1} - c_i^n}{\delta t} = G_i^n - R_i^n \quad (n = 0, ..., M-1) \tag{13}$$

- Step 3. Estimate the generating terms and reaction terms of other macro variables and evolve the macro-variable of H^+, H_2 for one macro time step using the macro-solver;

- Step 4. Set the current state of the macro-variables and repeat

We choose restricted additive Swcharz precondioner in Ref.[2] and GEMRS method as macro-solver to get the solutions, which will improve the parallel efficiency comparing with regular additive Schwarz precondioner.

As suggested above, in order to accelerate the simulation, we implement the multi-scale algorithm to calculate the time-dependent Nersnt-Planck equation. The difference of result between the two algorithms, with and without multi-scale method, is shown in Figure 2 and Figure 3.

Numerical Results

H_2 *Trend*

For different kind of semiconductor, the doping situation changes. And the difficulties of simulation are different, since the linear system of high doping case is much more singular to solve. We present two different doping parameters in Table 1. Considering the time cost is very expensive even with the parallel method, we have to try the total dose of 0.1krad, which is less than the experiment. Every point we get need nearly 3000s, but it

is still much faster than other simulation time [9]. The mesh we use to calculate below has 168986 vertices and 977221 tetrahedra. Figure 4 shows the trend of N_{it} data in different H_2 concentrations, which matches the data taken by the experiments [5]. The high doping case has a swift change despite the similar trend, which could be explained by the cracking of double-hydrogenated defects.

Table 1: Semiconductor with different doping [15] [6]

parameters	low doping	high doping
H_2	$10^{11} - 10^{18} cm^{-3}$	$10^{11} - 10^{18} cm^{-3}$
$V_{o\gamma}$	$1.0e14 cm^{-3}$	$1.0e15 cm^{-3}$
$V_{o\delta}$	$1.0e18 cm^{-3}$	$1.0e18 cm^{-3}$
$V_{o\gamma}H$	$1.0e14 cm^{-3}$	$1.0e14 cm^{-3}$
$V_{o\delta}H$	$1.0e14 cm^{-3}$	$1.0e14 cm^{-3}$
$V_{o\gamma}H_2$	$1.0e14 cm^{-3}$	$1.0e16 cm^{-3}$
$V_{o\delta}H_2$	$1.0e13 cm^{-3}$	$1.0e16 cm^{-3}$
$Si - H$	$1.0e13 cm^{-2}$	$1.0e13 cm^{-2}$

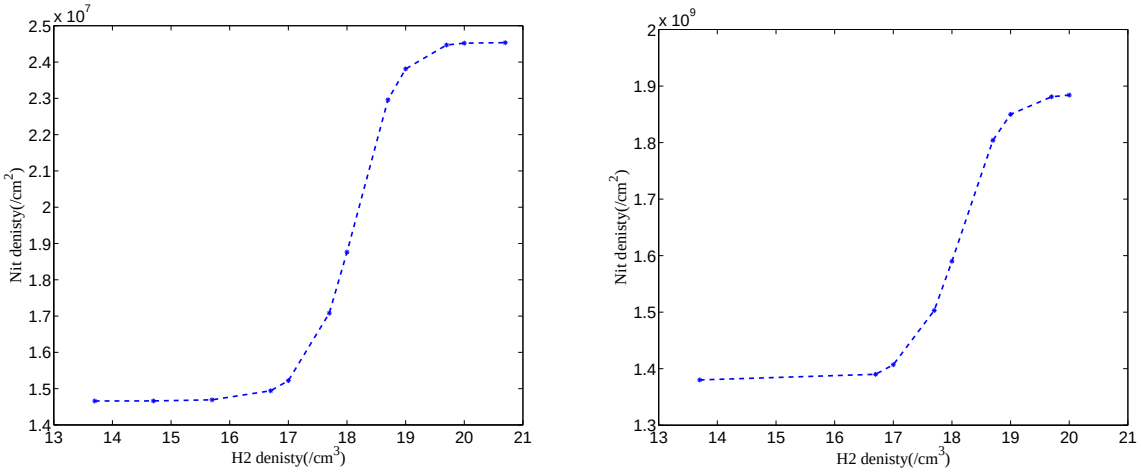


Fig.a Fig.b

Figure 4: Total dose of 0.1krad with 10rad/s in low doping (Fig.a) and high doping(Fig.b)

EDLRS Effect

We also simulate the EDLRS response in Figure 5 and show the general expected trend which matchs the trend of experiment data [11] and the simulation result of 1D case [4]. The quantities of interface traps reduce by the increase of dose rate and we predict to raise the amount of total dose and present better characterization in our later work.

Parallel Efficiency

Here we present parallel efficiency results. The implementation of the algorithm is based on the package PHG [18], which is a parallel toolbox for writing adaptive finite element programs. In order to assess the parallel scalability and efficiency with 1024 processes,

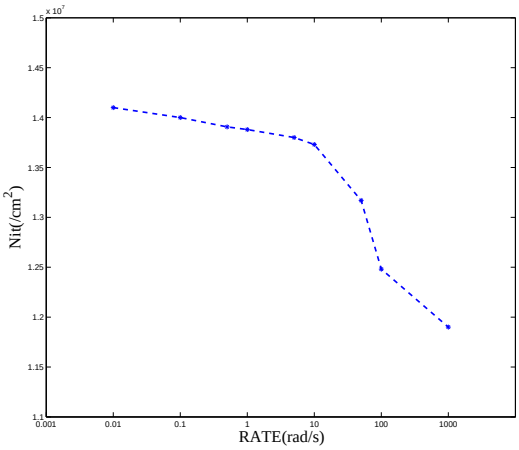


Figure 5: Total dose of 0.1krad and the H2 density of $3e14cm^{-3}$

we simulate a much larger system with a mesh containing 1418778 vertices and 8637254 tetrahedra, as MPI communication costs too much time in small system. We use GMRES method with restricted additive Schwarz preconditioner as our linear system solver. Table 4 gives the wall-clock time and parallel efficiency for different amount of MPI processes. Considering the large memory requirement, our tests start with 32 processes, whose parallel efficiency is regarded as 100%, and the parallel efficiency for p processes can be defined as:

$$E_p = \frac{32T_{32}}{pT_p}$$

where T_p denotes the time to irradiate 0.1krad with the rate of 10 rad/s for solving the PNPEs using p processes. In most cases, it proves that our method has good parallel scalability.

Table 2: parallel extendiability

Num of procs	Times	Efficiency
32	182000s	100%
64	75888s	120%
128	30444s	149%
256	16566s	137%
512	14570s	78%

Conclusion

We have developed a parallel finite element code to give a glimpse at the TID effects. The parallel solution of time-dependent PNP equations with reaction functions have been explored. During the simulation, multi-scale method is presented to dispose the time scale sensitive parameters and restricted additive Schwarz preconditioner is used to overcome the singularity of linear system. Our result shows a good match with the influence of H_2 concentration and dose rate. Further study will focus on the error estimate and reducing calculation time.

Acknowledgement

All the authors are supported by Science Challenge Program (SCP JCKY2016212A503) and the China NSF (NSFC 21573274).

References

[1] Peter E Blöchl. First-principles calculations of defects in oxygen-deficient silica exposed to hydrogen. *Physical Review B*, 62(10):6158, 2000.

[2] Xiao Chuan Cai and Marcus Sarkis. A restricted additive schwarz preconditioner for general sparse linear systems. *Siam Journal on Scientific Computing*, 21(2):792–797, 1999.

[3] Tony F Chan and Tarek P Mathew. Domain decomposition algorithms. *Acta numerica*, 3:61–143, 1994.

[4] X. J Chen, H. J Barnaby, P Adell, and R. L Pease. Modeling the dose rate response and the effects of hydrogen in bipolar technologies. *IEEE Transactions on Nuclear Science*, 56(6):3196–3202, 2010.

[5] X. J. Chen, H. J. Barnaby, B. Vermeire, and K. Holbert. Mechanisms of enhanced radiation-induced degradation due to excess molecular hydrogen in bipolar oxides. *IEEE Transactions on Nuclear Science*, 54(6):1913–1919, 2007.

[6] RAB Devine, D Mathiot, WL Warren, DM Fleetwood, and B Aspar. Point defect generation during high temperature annealing of the si-sio2 interface. *Applied physics letters*, 63(21):2926–2928, 1993.

[7] E. W. Enlow, R. L. Pease, W. Combs, and R. D. Schrimpf. Response of advanced bipolar processes to ionizing radiation. *IEEE Transactions on Nuclear Science*, 38(6):1342–1351, 1991.

[8] Ivan Sanchez Esqueda. *Modeling of total ionizing dose effects in advanced complementary metal-oxide-semiconductor technologies*. PhD thesis, Arizona State University, 2011.

[9] Harold P. Hjalmarson, Ronald L. Pease, and Roderick A. B. Devine. Calculations of radiation dose-rate sensitivity of bipolar transistors. *IEEE Transactions on Nuclear Science*, 55(6):3009–3015, 2008.

[10] Timothy R Oldham. *Ionizing radiation effects in MOS oxides*. World Scientific, 2000.

[11] Ronald L. Pease, Philippe Claude Adell, Bernard G. Rax, and Xiao Jie Chen. The effects of hydrogen on the enhanced low dose rate sensitivity (eldrs) of bipolar linear circuits. *IEEE Transactions on Nuclear Science*, 55(6):3169–3173, 2008.

[12] N. L Rowsey, M. E Law, R. D Schrimpf, and D. M Fleetwood. A quantitative model for eldrs and degradation effects in irradiated oxides based on first principles calculations. *Nuclear Science IEEE Transactions on*, 58(6):2937–2944, 2011.

[13] Nicole L Rowsey. *Quantitative modeling of total ionizing dose reliability effects in device silicon dioxide layers*. 2012.

[14] Bin Tu, Minxin Chen, Yan Xie, Linbo Zhang, Bob Eisenberg, and Benzhuo Lu. A parallel finite element simulator for ion transport through three-dimensional ion channel systems. *Journal of computational chemistry*, 34(24):2065–2078, 2013.

[15] WL Warren, DM Fleetwood, MR Shaneyfelt, JR Schwank, PS Winokur, RAB Devine, and D Mathiot. Links between oxide, interface, and border traps in high-temperature annealed si/sio2 systems. *Applied physics letters*, 64(25):3452–3454, 1994.

[16] Yan Xie, Jie Cheng, Benzhuo Lu, and Linbo Zhang. Parallel adaptive finite element algorithms for solving the coupled electro-diffusion equations. *Molecular Based Mathematical Biology*, 1:90–108, 2013.

[17] Jingjie Xu, Yan Xie, Benzhuo Lu, and Linbo Zhang. Charged substrate and product together contribute like a nonreactive species to the overall electrostatic steering in diffusion-reaction processes. *The Journal of Physical Chemistry B*, 120(33):8147–8153, 2016.

[18] Lin-Bo Zhang. A parallel algorithm for adaptive local refinement of tetrahedral meshes using bisection. *Numer. Math.: Theory, Methods and Applications*, 2:65–89, 2009.

Simplified and fast modeling of automotive body frame

Chungyang Gui[1,2] and *†Wenjie Zuo[1,2]

[1] State Key Laboratory of Automotive Simulation and Control, Jilin University, China
[2] School of Mechanical Science and Engineering, Jilin University, China

*Presenting author: zuowenjie@jlu.edu.cn
†Corresponding author: zuowenjie@jlu.edu.cn

Abstract

At conceptual design stage, beam element is extensively used to create the frame structure of automotive body, which can not only archive the accurate stiffness but also reduce much design period. However, so far there is no perfect method to apply the beam element to create the automotive frame composed of the plate element. This paper presents a solution to this problem in order to help engineers to fast carry out the vehicle body problem at conceptual design stage. Firstly, formulations of geometric properties of complex section are reviewed. Secondly, the method of establishing the cross beam with reference to the midpoint deflection and mass of the plate is presented to simplify the plate with a higher precision. Thirdly, regarding the joint elements of vehicle body, the spatial semi-rigid beam element and its stiffness matrix are expressed. Lastly, a numerical example of car frame proves that the proposed method can analyze the stiffness of the body more fast and accurately.

Keywords: Conceptual design, automotive frame, cross beam, plate element, semi-rigid beam element

Introduction

The design of automobile body can be divided into conceptual design stage and detailed design stage. Conceptual design is requisite during the whole design process and can reduce the design period and manufacturing risk for detailed design [1][2][3]. Especially, many automobile manufacturers make great efforts to shorten production cycles and to broaden the spectra of vehicles, so the demand for conceptual design will continue to increase [4][5].

Since the body-in-white (BIW) structure occupies about one third of the total weight of a passenger vehicle, many researches have been concentrating on this area. Some methods have been put forward to promote the development of conceptual design, in which it is well solved by using a simplified frame consisting of thin-walled beams (TWBs) [6][7]. Cross-sectional shapes are determined to describe the simplified frame. Therefore, sufficient CAD geometry data of TWBs is necessary to design automotive body [8][9][10]. Also, much effort has been devoted to establish simplified model, for example, the first order analysis (FOA) was originally proposed for graphic interfaces using Microsoft Excel to achieve the product oriented analysis, and open, single-cell and double-cell sections were applied to the frame [11][12]. Nishigaki and Kikuchi [13] focused on the crashworthiness of FOA, and predicted the collapse behavior of the beam members. Moreover, crashworthiness design and optimization for TWB with complex cross-sectional shapes under axial impact load was conducted by using genetic algorithm [14]. Then, BIW frame with semi-rigid joints was created to improve the accuracy of stiffness evaluation [15]. Besides, component sensitivity analysis was proposed to modify and optimize the BIW frame with rectangular tubes [16][17][18]. Recently, the torsional moment of inertia of the three-cell section was formulated [19]. However, above studies did not clearly present the mathematical expressions of moments of inertia, product of inertia, and torsional

moment of inertia of the cross section.

Meanwhile, the joint structures are important parts of the BIW frame [20][21]. Mostly, the simplified joint was regarded as spring elements, whose properties were usually from the reduced joint model of plate finite element (FE) or experimental test of trial-manufactured joint [22][23]. Actually, detailed FE joint or trial-manufactured joint is unavailable at the conceptual design stage [24]. Therefore, the properties of spring elements should be approximately calculated by TWB, which is a only feasible method at that stage. Among them, properties of the entire structure were evaluated by the FE analyses of a model made of beam elements frames and torsional spring elements joints, created from the selected joints and joined frames [25]. Also, plate structures, such as ceiling, floor and firewall, are all importantly load-bearing structures, but the DoFs between plate element and beam element is inconsistent, so plate structures are usually omitted in the BIW frame [26][27][28].

Therefore, This paper focuses on the formulations of the torsional moment of inertia of open, single-cell, double-cell and three-cell sections, simplification of plates. Additionally, performance evaluation of the refined BIW frame is conducted and compared with the benchmarking BIW structure.

Formulations of properties of complex section

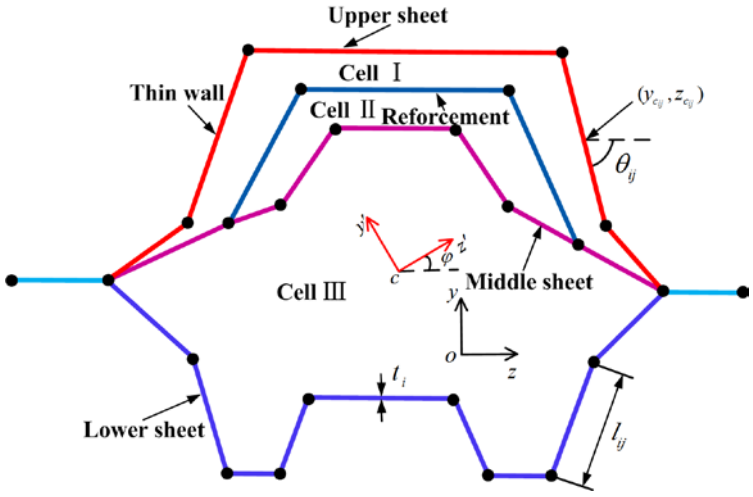


Figure 1. A typical cross section

A typical cross section is shown in Figure 1. Engineers design the cross-sectional shape in the yoz coordinates. Point C is the centroid of cross section. Each sheet can be viewed as a folded line consisting of rectangle segments. So the cross-sectional area can be written as

$$A = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{m} l_{ij} t_i \tag{1}$$

where n is the number of sheets; m is the number of segment of the i-th sheet; l_{ij} and A_{ij} are the length and area of the j-th segment of i-th sheet, respectively; t_i is the thickness of the i-th sheet. The cross-sectional centroid can be expressed as

$$y_c = \frac{1}{A}\sum_{i}^{n}\sum_{j}^{m} A_{ij} y_{c_{ij}} \quad \text{and} \quad z_c = \frac{1}{A}\sum_{i}^{n}\sum_{j}^{m} A_{ij} z_{c_{ij}} \tag{2}$$

where $(y_{C_{ij}}, z_{C_{ij}})$ indicates the coordinate of the center of the *j*-th segment of *i*-th sheet, as shown in Figure 1. Besides, moments of inertia I_y, I_z and product of inertia I_{yz} with respect to the centroid can be, respectively, derived as

$$I_y = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\left(\frac{l_{ij}t_i^3}{12}\right)\sin^2\theta_{ij} + \left(\frac{l_{ij}^3 t_i}{12}\right)\cos^2\theta_{ij} + l_{ij}t_i y_{C_{ij}}^2\right] \tag{3}$$

$$I_z = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\left(\frac{l_{ij}t_i^3}{12}\right)\cos^2\theta_{ij} + \left(\frac{l_{ij}^3 t_i}{12}\right)\sin^2\theta_{ij} + l_{ij}t_i z_{C_{ij}}^2\right] \tag{4}$$

$$I_{yz} = \sum_{i=1}^{n}\sum_{j=1}^{m}\left[\left(\frac{l_{ij}^3 t_i - l_{ij}t_i^3}{24}\right)\sin 2\theta_{ij} + l_{ij}t_i z_{C_{ij}} y_{C_{ij}}\right] \tag{5}$$

where θ_{ij} is the angle between the positive *z* axis and the *j*-th segment of *i*-th sheet. From I_y, I_z and I_{yz}, the principal moment of inertia are obtained by

$$I_{\max} = \frac{1}{2}\left(I_y + I_z\right) + \sqrt{\frac{1}{2}\left(I_y - I_z\right)^2 + I_{yz}^2} \tag{6}$$

$$I_{\min} = \frac{1}{2}\left(I_y + I_z\right) - \sqrt{\frac{1}{2}\left(I_y - I_z\right)^2 + I_{yz}^2} \tag{7}$$

The angle φ of principle direction of inertia with respect to the reference *z*' axis is

$$\varphi = \frac{1}{2}\tan^{-1}\left(\frac{-2I_{yz}}{I_y - I_z}\right) \tag{8}$$

The y'cz' coordinate axis is called principle coordinate axis of inertia. The procedure for calculating the torsional moment of inertia depends on the types of the cross-sectional shape, as shown in Figure 1. The torsional moment of inertia of an open section is calculated as

$$J^{\mathrm{o}} = \sum_{i=1}^{o}\sum_{j=1}^{m} l_{ij}t_i^3 \tag{9}$$

where *o* denotes the number of sheets for open section. The torsional moment of inertia of single-cell section J_1^c, double-cell section J_2^c and three-cell section J_3^c can, respectively, be expressed as

$$J_1^{\mathrm{c}} = \frac{4F_1^2}{L_l/t_l + L_u/t_u} \tag{10}$$

$$J_2^{\mathrm{c}} = \frac{4\left\{F_1^2\left(\frac{L_l}{t_l} + \frac{L_r}{t_r} + \frac{L_u - L_u'}{t_u}\right) - F_2^2\left(\frac{L_r}{t_r} + \frac{L_u'}{t_u}\right)\right\}}{\left(\frac{L_r}{t_r} + \frac{L_u'}{t_u}\right)\left(\frac{L_l}{t_l} + \frac{L_r}{t_r} + \frac{L_u - L_u'}{t_u}\right) - \frac{L_r^2}{t_r^2}} \tag{11}$$

$$J_3^c = 4\left(q_1 F_1 + q_2 F_2 + q_3 F_3\right) \tag{12}$$

where q_1, q_2 and q_3 are solved by equation (13)

$$\begin{bmatrix} \frac{L_u}{t_u}+\frac{L_r}{t_r}+\frac{L_m-L_m'}{t_m} & -\frac{L_r}{t_r} & -\frac{L_m-L_m'}{t_m} \\ -\frac{L_r}{t_r} & \frac{L_r}{t_r}+\frac{L_m'}{t_m} & -\frac{L_m'}{t_m} \\ -\frac{L_m-L_m'}{t_m} & -\frac{L_m'}{t_m} & \frac{L_l}{t_l}+\frac{L_m}{t_m} \end{bmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix} \tag{13}$$

where F_1, F_2 and F_3 are the enclosed area of Cell Ⅰ, Cell Ⅱ, and Cell Ⅲ; L_u, L_l, L_m and L_r are the length of Upper Sheet, Lower Sheet, Middle Sheet and Reforcement; L_u' is the length of Upper Sheet' which is the part of Upper Sheet as shown in Figure 2 (c); L_m' is the length of the shared part of Cell Ⅱ and Cell Ⅲ as shown in Figure 2 (d); t_u, t_l, t_m and t_r are the thickness of Upper Sheet, Lower Sheet, Middle Sheet and Reforcement, respectively.

When a more complex section consists of open and close sections, the torsional moment of inertia can be expressed as

$$J = J^o + J_k^c \quad k = 1, 2 \text{ and } 3 \tag{14}$$

where k represents the number of close cells on the section.

In summary, the formulations of moments of inertia I_y and I_z are the same for the four types of cross sections. However, the formulations of the torsional moment of inertia J for the four types of cross sections are different, as shown in Figure 2

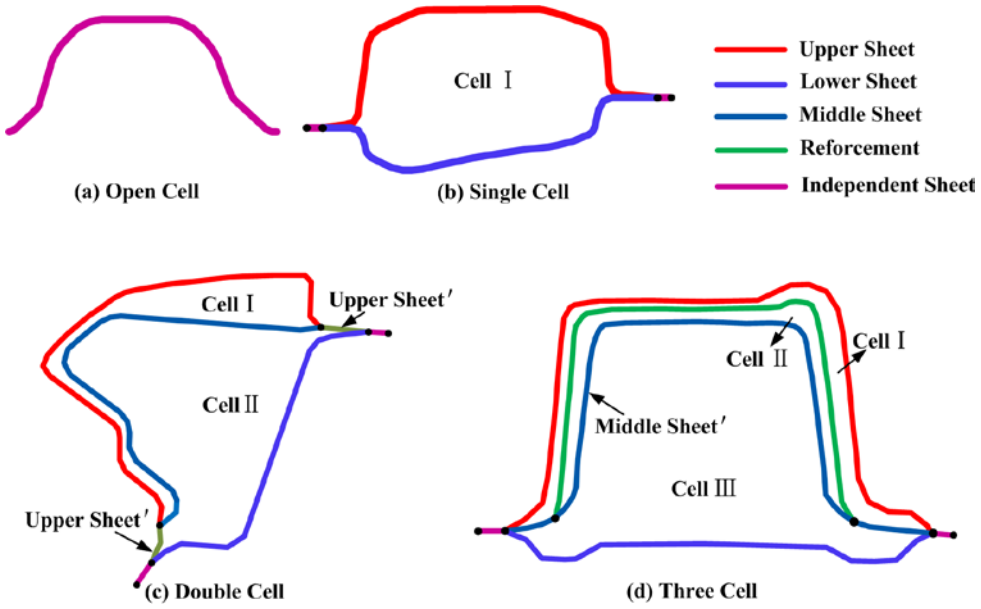


Figure 2. Four types of cross section

Simplification of plate structure

The plates such as the ceiling, floor and firewall not only contributes to the mass of vehicle, but also the stiffness. Moreover, the DoFs between plate element and beam element are inconsistent, therefore, the cross beams are introduced to simplify the plate structure as shown in Figure 3

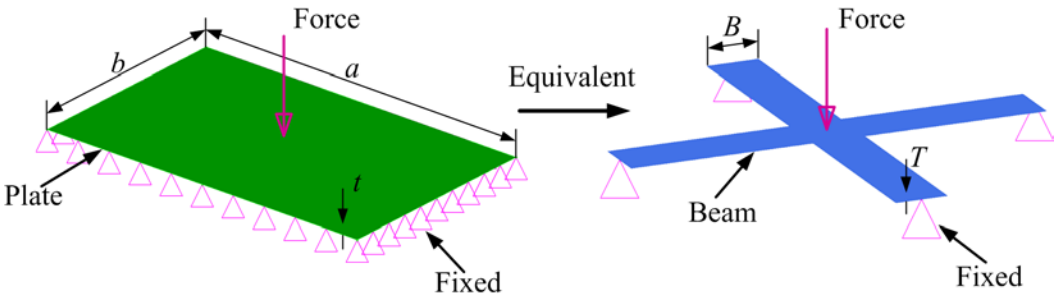


Figure 3 Equivalence from plate to cross beams

The mass of the rectangular thin plate and cross beams are, respectively, calculated as

$$m = \rho abt \tag{15}$$

$$M = \rho LBT \tag{16}$$

where ρ, a, b and t are the density, length, breadth and thickness of plate, respectively; L is the sum of the length of the two diagonal beams; B and T are the breadth and thickness of each beam among cross beams, respectively. The central deflection of the rectangular plate and cross beams can be, respectively, obtained by

$$w = \beta Fab/D \tag{17}$$

$$W == FL^3/256\,EBT^3 \tag{18}$$

where coefficient $\beta = 0.0056$ when the boundary of the plate is fixed, and stiffness of the plate $D = Et^3/12\left(1-\mu^2\right)$.

The respective equality of mass and deflection between plate and cross beams is necessary to the respective equivalence of them. Therefore, let equations (15) and (17) be equal to equations (16) and (18), respectively. Then, the width and thickness of cross beams can be calculated as

$$B = \frac{32a^2b^2\sqrt{3\beta\left(1-\mu^2\right)}}{L^3} \tag{19}$$

$$T = \frac{L^2t}{32ab\sqrt{3\beta\left(1-\mu^2\right)}} \tag{20}$$

Modeling and evaluation of BIW frame

The proposed methods are used to simplify the Toyota Yaris BIW, which contains 232 components, 495000 plate elements and 1510000 DoFs, as shown in Figure 4. The simplified frame, as shown in Figure 5 contains 470 semi-rigid beam elements, 50 sections, and 5600 DOFs.

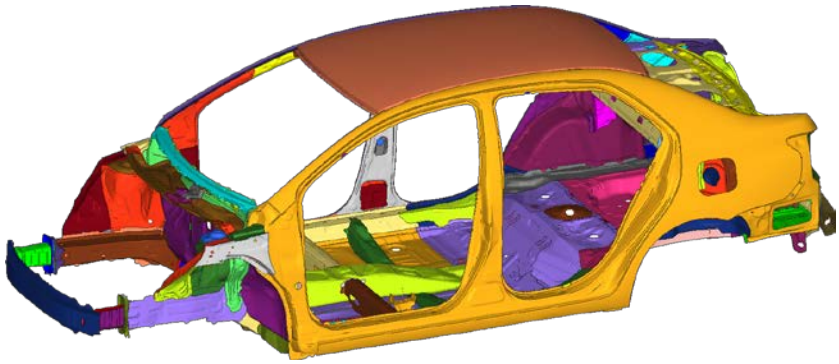


Figure 4 A detailed FE model of Yaris BIW

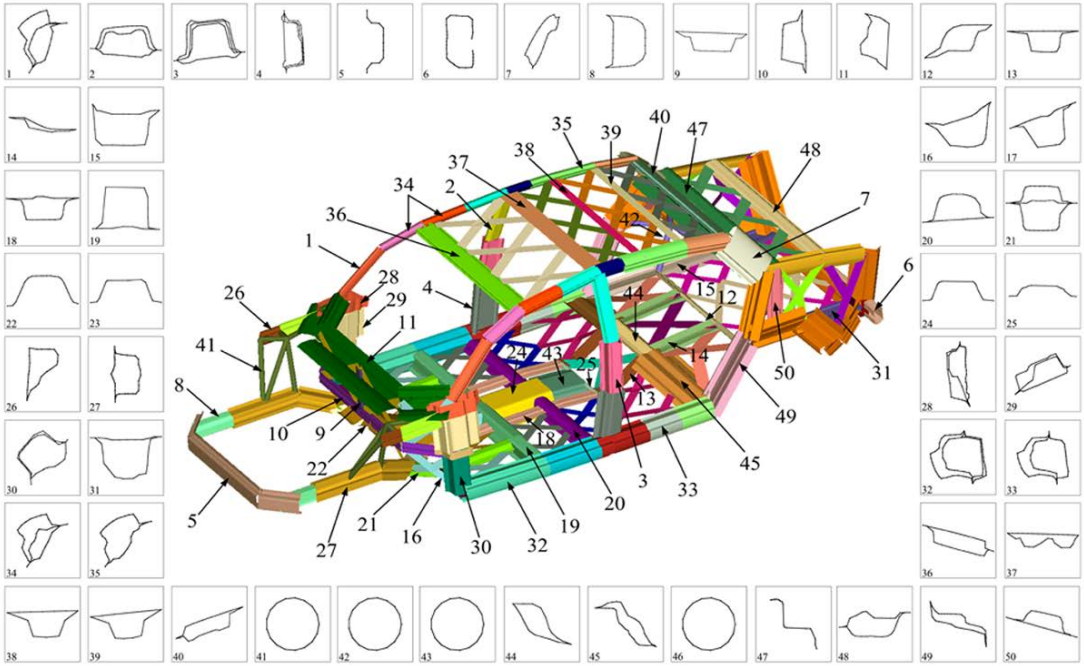


Figure 5 Simplified BIW frame

At the torsional loadstep, the DoFs of the rear suspension are all constrained. Moment of couple is exerted to the front suspension, whose size of force is 1980 N. At the bending loadstep, xyz-translational DoFs of the front suspension and the *z*-translational DoFs of the rear suspension are constrained. At the fixed places of the seats, five forces which each of them is 1670 N are exerted to replace the weight of passengers, respectively.

Detailed BIW, solved by Optistruct software, is regarded as a benchmarking example. Two simplified BIW frames are solved by CarFrame CAE software to compare with the Detailed BIW. One BIW frame uses the rigid connection, the other uses the semi-rigid connection. All the results are listed in Table 1. The modeling cost for the simplified BIW frame is about 2 days, which is less than the 3 months of detailed BIW. Referring to the benchmarking detailed BIW, these two types of BIW frame almost acquire the same mass and centroid coordinates. However, the simplified frame with semi-rigid joints obtains the more accurate torsional stiffness, bending stiffness and frequencies than the simplified frame with rigid joints. Especially, the errors of those evaluation indexes of simplified frame with semi-rigid joints are all limited within 10%, compared to the detailed model, which can be accepted at the conceptual design stage.

Table 1 Comparison of torsional stiffness, bending stiffness and frequency

Evaluation index		BIW				
		Detailed	Simplified frame with rigid joints	Error (%)	Simplified frame with semi-rigid joints	Error (%)
Mass (Kg)		263.7	263.7	0.00%	263.7	0.00%
Centroid coordinates (mm)	x	-2.223.9	-2224.0	0.00%	-2224.0	0.00%
	y	3.8	-4.5E-03	0.00%	-4.5E-03	0.00%
	z	619.6	619.7	0.01%	619.7	0.01%
Torsional stiffness ($N \cdot m / ^o$)		7418	19876	167.94%	7583	2.22%
Bending stiffness (N/m)		17996	39567	119.87%	16442	8.64%
Frequency (Hz)	1st	28.6	50.6	76.92%	25.9	9.44%
	2nd	35.5	65.6	84.79%	39.0	9.86%
	3rd	52.0	69.8	34.23%	48.3	7.11%

Conclusion

The aim of this study is to propose a fast and simplified modeling method of BIW frame at the conceptual design stage. TWBs with complex section and cross beams can be together used to fast create BIW frame, which can be readily designed and modified for the development of new automobile body. Numerical example proves that simplified BIW frame with semi-rigid joints obtains the more accurate torsional stiffness, bending stiffness and frequencies than the simplified frame with rigid joints. Especially, the errors of those evaluation indexes are all limited within 10%, which can be accepted at the conceptual design stage.

References

[1] Zuo W. (2015) Bi-level optimization for the cross-sectional shape of a thin-walled car body frame with static stiffness and dynamic frequency stiffness constraints, *Proceedings of the Institution of Mechanical Engineers Part D Journal of Automobile Engineering* **229**, 1046-1059.

[2] Zuo W., Bai J., and Yu J. (2015) Sensitivity reanalysis of static displacement using Taylor series expansion and combined approximate method, *Structural & Multidisciplinary Optimization* **53**, 953-959.

[3] Xue L., Chen B., Zhao Z., Dang Z., and Zou M. (2016) Experimental Study of Torque Using a Small Scoop on the Lunar Surface, *International Journal of Aerospace Engineering* **2016**, 1-8.

[4] Hou W., Shan C., and Zhang H. (2015) *Multi-level optimization method for vehicle body in conceptual design*, ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Boston, Massachusetts, USA.

[5] Zuo W., Xu T., Zhang H., and Xu T. (2011) Fast structural optimization with frequency constraints by genetic algorithm using adaptive eigenvalue reanalysis methods, *Structural and Multidisciplinary Optimization* **43**, 799-810.

[6] Zou M., Wei C. G., Li J. Q., Xu S. C., and Zhang X. (2015) The energy absorption of bamboo under dynamic axial loading, *Thin-Walled Structures* **95**, 255-261.

[7] Zou M., Xu S., Wei C., Wang H., and Liu Z. (2016) A bionic method for the crashworthiness design of thin-walled structures inspired by bamboo, *Thin-Walled Structures* **101**, 222-230.

[8] Huang G., Wang H., Gao G., and Li G. (2011) Modified combined approximations and the application in reanalysis of vehicle frame stiffness, *Journal of Mechanical Engineering* **47**, 86-92.

[9] Huang G., Wang H., and Li G. (2015) Some modifications for ES-FEM and its application for vehicle design, *Engineering Computations* **32**, 1432-1459.

[10] Wang H., Li G., and Li E. (2010) Time-based metamodeling technique for vehicle crashworthiness optimization, *Computer Methods in Applied Mechanics and Engineering* **199**, 2497-2509.

[11] Nishigaki H., Amago T., Sugiura H., Kojima Y., Nishiwaki S., and Kikuchi N. (2004) *First order analysis for automotive body structure design - part 1: overview and applications*, SAE 2004 World Congress & Exhibition, Detroit, Michigan, USA.

[12] Nishigaki H., Nishiwaki S., Amago T., Kojima Y., and Kikuchi N. (2001) *First order analysis - New CAE tools for automotive body designers*, SAE 2001 World Congress, Detroit, Michigan, USA.
[13] Nishigaki H. and Kikuchi N. (2004) *First order analysis for automotive body structure design - part 3: crashworthiness analysis using beam elements*, SAE 2004 World Congress & Exhibition, Detroit, Michigan, USA.
[14] Bai J., Li Y., and Zuo W. (2017) Cross-sectional shape optimisation for thin-walled beam crashworthiness with stamping constraints using genetic algorithm, *International Journal of Vehicle Design* **73**, 76-95.
[15] Zuo W., Li W., Xu T., Xuan S., and Na J. (2012) A complete development process of finite element software for body-in-white structure with semi-rigid beams in. NET framework, *Advances in Engineering Software* **45**, 261-271.
[16] Chen W. and Zuo W. (2014) Component sensitivity analysis of conceptual vehicle body for lightweight design under static and dynamic stiffness demands, *International Journal of Vehicle Design* **66**, 107-123.
[17] Zuo W., Yu J., and Saitou K. (2016) Stress sensitivity analysis and optimization of automobile body frame consisting of rectangular tubes, *International Journal of Automotive Technology* **17**, 843-851.
[18] Zuo W. and Saitou K. (2017) Multi-material topology optimization using ordered SIMP interpolation, *Structural and Multidisciplinary Optimization* **55**, 477–491.
[19] Zuo W. and Bai J. (2016) Cross-sectional shape design and optimization of automotive body with stamping constraints, *International Journal of Automotive Technology* **17**, 1003-1011.
[20] Hou W., Zhang H., Zhang W., and Hu P. (2011) Rapid structural property evaluation system for car body advanced design, *International journal of vehicle design* **57**, 242-253.
[21] Shahhosseini A. M., Prater G., Osborne G. M., Kuo E. Y., and Mehta P. R. (2010) Major compliance joint modelling survey for automotive body structures, *International Journal of Vehicle Systems Modelling and Testing* **5**, 1-17.
[22] Nikolaidis E. and Lee K. (1992) *A 3-D joint model for automotive structures*, International Conference On Vehicle Structural Mechanics & Cae, Traverse City, Michigan, USA.
[23] Mundo D., Donders S., Stigliano G., and Auweraer H. V. D. (2011) Concept design of vehicle bodies using reduced models of beams, joints and panels, *International Journal of Vehicle Design* **57**, 71-83.
[24] Zou M., Fan S., Shi R., Yang Y., and Li J. (2015) Effect of gravity on the mechanical properties of lunar regolith tested using a low gravity simulation device, *Journal of Terramechanics* **60**, 11–22.
[25] Lyu N., Lee B., and Saitou K. (2006) Optimal subassembly partitioning of space frame structures for in-process dimensional adjustability and stiffness, *Journal of Mechanical Design* **128**, 527-535.
[26] Hamza K. and Saitou K. (2004) Design optimization of vehicle structures for crashworthiness using equivalent mechanism approximations, *Journal of Mechanical Design* **127**, 485-492.
[27] Mohsin M. E. and Sadek E. A. (1980) On the dynamics of plates using a beam-analog, *Computers & Structures* **12**, 267-272.
[28] Xue L., Zhang R. R., Zong W., Song J. F., and Zou M. (2016) Bionic Design for Mars Sampling Scoop Inspired by Himalayan Marmot Claw, *Applied Bionics & Biomechanics* **2016**, doi: 10.1155/2016/5713683.

A truss section size optimization design method based on constraint variation principle

*** LIU Yu-Bin ¦, ZHENG Xin ¦, ZHANG Sheng-Jun 2, SHEN Feng 2,**

and †FU Xiang-Rong [1]

1 College of Water Conservancy and Civil Engineering, China Agricultural University, China.
2 State Nuclear Power Research Institute, China.
*Presenting author: liuyubin@cau.edu.cn
†Corresponding author: fuxr@cau.edu.cn

Abstract

In the actual structure, rational optimization of truss structure section size, can make each bar bear maximum load and play a role. By this way, we can not only make the truss material give full play to its own performance, but also reduce weight of truss, and cost savings. At present, there are two methods for section size optimization: Mechanical criterion method and mathematical programming method. This paper combines the two methods and analyze a truss-braced structure. In this paper, the truss node displacement is used as the constraint condition. It takes the cross section area of the bar in the truss as the independent variable and takes the steel weight of the truss as objective function. It uses constrained variation principle and force method equation to create a hybrid numerical method. Through concrete examples, the results of this algorithm are compared with the simulation results, and the reliability and correctness of truss section size optimization are verified.

Keywords: Truss; Section size; Optimization; Constrained variation principle; Software simulation;

1. Introduction

In the building structure, supporting truss is one of the important structure form which consists of a number of bar through some articulated points. Truss structure is widely used in homes, factories and other buildings because of its great Wholeness, large stiffness, great seismic ability. But in the actual structure, the force of the bar in supporting truss is not the same in affected by external force, so that it leads to the different bar stress. Rational design of truss structure, can meet the requirements of the actual working conditions, so that the bar can bear the maximum load and play a role. In this way, we can not only make the truss material give full play to its own performance, but also achieved the goal that reduce weight of truss, materials and save cost.

In the theoretical calculation stage, this paper firstly analyze the static-truss structure, discusses the case of single-independent and multivariable ones, deduces the iterative criterion. And then it explains the application of the optimization method by an example of cantilever statically indeterminate structure. At last it is compared with the software simulation results in order to obtain a more secure form of structure, so that the truss structure tends to rationalize the force.

2. Theoretical analysis

2.1 Statically determinate truss single variable optimization theoretical derivation

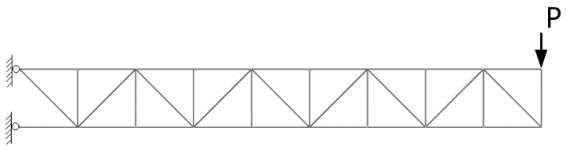


Figure 3-1.statically determinate truss single variable optimization

As is shown in the picture, the distance between two nodes of a truss is 0.5m.The truss is made in NO.45 steel：Density ρ=7800kg/m3，E=210GPa，μ=0.3。It has the same cross-sectional area A in chord member and web member. There is each 70000N static-state load which is downward as we can see in the picture.
Nodal displacement calculation formula:

$$\Delta = \sum \int \frac{\bar{F}_N F_P}{EA} ds \tag{1}$$

Truss quality calculation formula:

$$M(A) = k_1 \sum_i \rho l A + k_2 \sum_j \rho l A \tag{2}$$

Make $\Delta < 0.1m$, $g(A) = \Delta - 0.1$, The above questions can be expressed as:

$$\begin{array}{ll} \min & A \\ s.t. & A \le g \end{array} \tag{3}$$

The problem is an extreme one including an unknown value. It takes the cross section area of the bar in the truss as the independent variable, takes the points of force application as constraint condition and takes the steel weight of the truss as objective function by finite element software. We make $\Delta < 0.1m$, determine the optimum case when the independent variable is approaching.

2.2 Statically determinate truss double variables optimization theoretical derivation

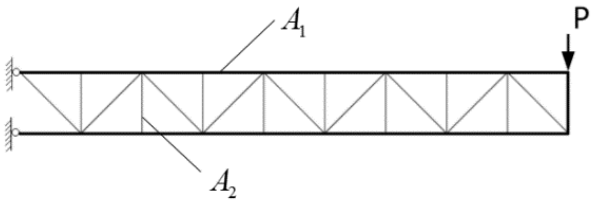


Figure 3-2.statically determinate truss two variables optimization

The parameter information is the same as we can see in the first example. The cross-sectional area in chord member is A_1 and the cross-sectional in web member is A_2 .

Nodal displacement calculation formula:

$$\Delta = \sum_i \int \frac{\bar{F}_N F_P}{EA_i} ds \tag{4}$$

Truss quality calculation formula:

$$M(A_1, A_2) = k_1 \sum_1 \rho l A_1 + k_2 \sum_2 \rho l A_2 \tag{5}$$

Make $\Delta < 0.1m$, $g(A) = \Delta - 0.1$.The above questions can be expressed as:

$$\begin{aligned} &\min && \mathrm{M}(A_1, A_2) \\ &s.t. && \mathrm{g}(A_1, A_2) \le 0 \end{aligned} \tag{6}$$

It makes A_1, A_2 as design variables.

Vector form:

$$\boldsymbol{A} = \left[A_1, A_2 \right]^{\mathrm{T}} \tag{7}$$

First we construct an internal punishment function:

$$\phi(A_1, A_2, r) = M(A_1, A_2) - r^k \frac{1}{\mathrm{g}(A_1, A_2)} \tag{8}$$

Use analytical method to seek extreme value.

$$\begin{cases} \dfrac{\partial \phi}{\partial A_1} = 0 \\ \dfrac{\partial \phi}{\partial A_2} = 0 \end{cases} \tag{9}$$

Solve it simultaneously

$$\begin{cases} A_1(\mathrm{r}^k) \\ A_2(\mathrm{r}^k) \end{cases}, \begin{cases} A_1'(\mathrm{r}^k) \\ A_2'(\mathrm{r}^k) \end{cases} \tag{10}$$

Meet

$$\begin{cases} A_1(\mathrm{r}^k) > 0 \\ A_2(\mathrm{r}^k) > 0 \\ \mathrm{g}(A_1, A_2) \le 0 \end{cases} \begin{cases} A_1(\mathrm{r}^k) > 0 \\ A_2(\mathrm{r}^k) > 0 \\ \mathrm{g}(A_1, A_2) \le 0 \end{cases} \tag{11}$$

Ignore solutions that don't meet the requirement. $r^{(k)}$ is the penalty factor in the formula.it is descending positive sequence:

$r^{(0)} > r^{(1)} > r^{(2)} > \cdots > r^{(k)} > r^{(k+1)} > \cdots > 0, r^{(k-1)} \cdot c = r^{(k)}$ $r^{(k-1)} \cdot c = r^{(k)}$

Reduction coefficient c:

$$0 < c < 1, \lim_{k \to \infty} r^{(k)} = 0, \phi(A_1, A_2, r) \to M(A_1, A_2) \tag{12}$$

Flow diagram:

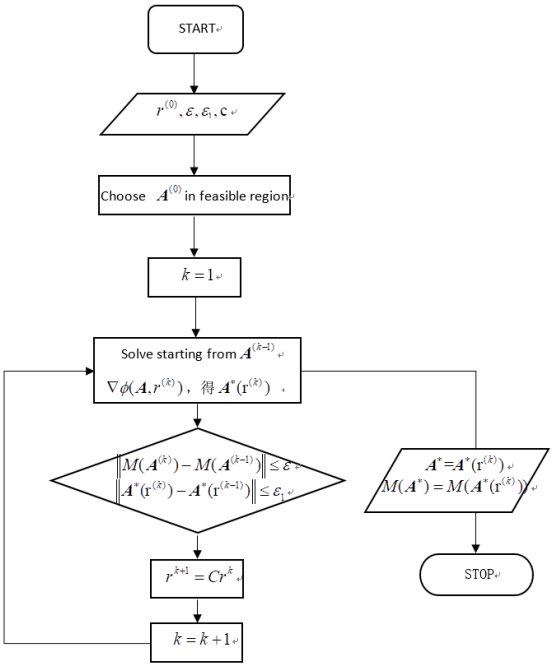


Figure 3-3.Flow diagram

2.3 Cantilever statically indeterminate structure single variable optimization theoretical derivation.

A project used support truss structure. The distance between two nodes of a truss is 0.5m.The truss is made in NO.45 steel：Density ρ=7800kg/m3，E=210GPa，μ=0.3。It has the same cross-sectional area A in chord member and web member. Each cross-sectional area is 350mm2 before optimized. The mechanical model is shown below:

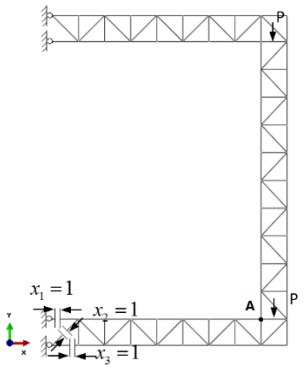


Figure 3-3.statically indeterminate structure single variable optimization

The truss itself is subjected to gravity, and there is each 70000N static-state load which is downward as we can see in the picture.

Establish force method equation:

$$\begin{cases} \delta_{11}x_1+\delta_{12}x_2+\delta_{13}x_3+\Delta_{1p}=0 \\ \delta_{21}x_1+\delta_{22}x_2+\delta_{23}x_3+\Delta_{2p}=0 \\ \delta_{31}x_1+\delta_{32}x_2+\delta_{33}x_3+\Delta_{3p}=0 \end{cases} \tag{13}$$

Get unknown coefficients and free term coefficients in force method equation by graphic multiplication method, then put it into formula to solve out $[x_1, x_2, x_3]^T$, which is the fundamental unknown force vector. Put the vector into formula to get internal force each in the statically indeterminate structure.

Transform mechanical problem into optimization problem:

$$\begin{aligned} &\min && \mathrm{M}(A) \\ &s.t. && \mathrm{g}(A)\le 0 \end{aligned} \tag{14}$$

Determine the optimum case by finite element software when the independent variable is approaching.

3.4 Cantilever statically indeterminate structure double variables optimization theoretical derivation.

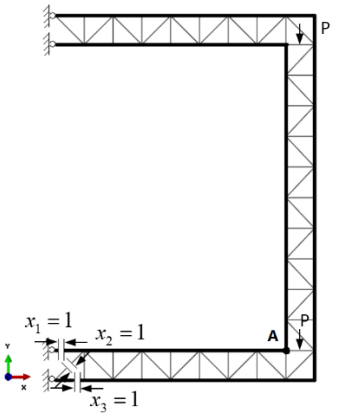


Figure 3-4.statically indeterminate structure double variables optimization

The parameter information is the same as we can see in the third example. The calculation method is the same as we can see in the second example. Then we get a series of solutions by finite element software.

Table 1. Comparison between optimization results

example		Ⅰ	Ⅱ	Ⅲ	Ⅳ
before optimization	A1/m2	1e-3	1e-3	4e-4	4e-4
	A1/m2	1e-3	1e-3	4e-4	4e-4
	M/kg	158.83	158.83	191.25	191.25
algorithm optimization	A1/m2	8.59e-4	1.02e-3	3.43e-4	4.29e-4
	A1/m2	8.59e-4	2.2e-4	3.43e-4	2.26e-4
	M/kg	136.46	94.17	164.00	152.3
simulation optimization	A1/m2	8.62e-4	1.04e-3	3.25e-4	4.34e-4
	A1/m2	8.62e-4	2.0e-4	3.25e-4	1.95e-4
	M/kg	136.46	94.012	155.39	145.4

Conclusions

Through the comparison and research of the derivation and numerical simulation of the truss structure, the following conclusions can be drawn:

(1) After the cross-sectional size optimization, the cross-sectional area of the structural bar is reduced, but it has made full use of the material properties, that is, the amount of steel is reduced, the cost is less, and the material performance can be fully utilized. It is suitable for the scientific development concept.

(2) Based on the previous research, the cross section area of the structural chord and the web are optimized as the optimization variables. The optimized effect is better than the effect of using the cross-sectional area as a single optimization variable. As the cross-sectional area of the chord increases, the total amount of steel is reduced and the strength of the material is brought into full play.

(3) Compared with the numerical results and the results of the algorithm, the results show that the numerical simulation results are slightly smaller than the algorithm optimization results, and there exist about 5% errors.

Acknowledgements:

Project supported by the National Natural Science Foundation of China (Nos. 11672362), the National Science and Technology Major Project(Nos. 2015ZX06004004-003)

References

[1] JIANG, D. J., WANG, D. X. (2004) Intelligent layout optimization design of truss, *Engineering mechanics* **26**, 160-165.

[2] Li, D. Z., Yu, D. Y., Ma, X. R. (2009) Truss topology optimization with uncertain loading scenarios, *Journal of Beijing university of aeronautics and astronautics* **35**, 1170-1173, 1179.

[3] Zhou, Q. C., Wu, Q. L., Xiong, X. L., et al. (2016) Optimal design of topology and section size of truss structures, *journal of Xi'an Jiao tong university* **50**, 2-9.

[4] Zhang, B. H., Hou, C. (1998) Optimization design of civil engineering structure, *Tong Ji university press*, Shanghai.

[5] Zhuang, Z., Y, X. C., Liao, J. H., et al. (2009) FEM analysis and application based on ABAQUS, *Tsinghua university press*, Beijing.

Flutter frequency based on bending - torsion coupling theory

***ZHENG Xin [1], LIU Yu-Bin [1],CHEN Pu[2], SHEN Feng [3],ZHANG Sheng-Jun [3],and †FU Xiang-Rong [1]**

[1] College of Water Conservancy and Civil Engineering, China Agricultural University, China.

[2] College of Engineering, Peking University, China.

[3] State Nuclear Power Research Institude, China.

*Presenting author: 476238231@qq.com

†Corresponding author: fuxr@cau.edu.cn

Abstract

This paper establishes a method which can forecast the general value range of flutter frequency by the following two steps, namely 1)Based on the theory of frequency superposition, this paper chooses the cantilever panel structure as a main object of the analysis, and controls the chord length and the root chord length as single variable separately. By the establishing and simulating of flutter procedures, this paper studies the correlation between the flutter mode and the bending - torsional coupling modes of the wing model under different geometric parameters. 2) Considering the flexible characteristics of the elastic support in the selection of node positions together with its agility to descript an intricate mechanical state of a structure, subjoined spring supports to the original model to simulated flutter statement. This approximation model simulates the flow-solid coupling state caused by the additional aerodynamic force in the flutter problem, so that provide a reference for the study of the internal mechanism of the problem.

Keywords: Bending-torsion coupling theory; Flutter frequency; Equivalent spring model;

1.Introduction

In the series of aero-elastic problems caused by the load of the aircraft, the avoidance of the flutter problem has become an extremely important part of the aircraft design due to the abruptness and destructiveness of the accident caused by it. The aero-elastic flutter damage is caused by the interaction among inertial, elastic and aerodynamic forces, thus the vibration of the structure becomes considerable significance when the mechanism of flutter is studied [1] . In the process of flutter analysis of the wing, the uncertainty and ineffectiveness of the simulation results is often caused by the geometric and physical parameters of the model, the selection of the structural and the aerodynamic mesh, the match of the Mach number and the flutter frequency.

In the process of flutter analysis, the following methods are used to predict the flutter parameters: the time-frequency domain of the flutter determinant can be used to obtain the judgment of the flutter coupling mode; or the V-g method, p-k method and other solution methods can also get the relevant paramount, such as flutter frequency, damping and flutter critical velocity; or through the frequency coupling trend and the damping zero-crossing branch to predict the

flutter mode; the analysis of the matrix eigenvector and the calculation of the contribution coefficient of the flutter mode can also achieve the prediction of the flutter mode[4].

The study is based on a simplified frequency coincidence theory proposed that the critical state of the flutter concerning coupling between bending and torsional vibrations. Firstly, analyzing rectangular cantilever models with chord length as its single variable, a convenient and effective formula for calculating the flutter frequency is obtained. Secondly, by selecting the wing tip length individually, the calculation error is compared and the reliability of the formula is also verified. At last, finding out a spring stiffness value for the model that makes its first-order vibration mode frequency corresponds to the original model flutter value. The research content provides a relatively effective forecasting method for the initial selection of the flutter frequency before analysis, and also provides some reference value for the simplification of the internal mechanism of the flutter problem.

2. Theoretical development

2.1 Frequency coincidence theory

The earliest aerodynamic problem is the bending and torsional coupling vibration of the wing. The bending caused by the wing flight leads to the angle of attack to change, and causes the disturbance of the lift and the aerodynamic moment, and again causes the bending which feedback to the bending and torsional vibration[2]. It is the initial explanation of flutter. A large number of flutter analysis of the case also shows that the increase in the flow rate will cause the two branches of the frequency changes in the speed continues to increase in the process, the two branch frequencies close to each other, the coupling is strengthened, then the coupling vibration will draw energy from the airflow and reach the critical point of flutter[5]. It should be noticed that this theory is a highly simplified analysis of the mechanism of flutter, which is not an exact solution concept compared to the flutter theory defined by the excitation and damping forces. The matrix form of the simplified equation of motion for a typical binary wing is given as follows

$$\begin{bmatrix} m\lambda^2 + k_a & S_a\lambda^2 + qS\dfrac{\partial C_L}{\partial a} \\ S_a\lambda^2 & I_a\lambda^2 + k_a - qSe\dfrac{\partial C_L}{\partial a} \end{bmatrix} \begin{bmatrix} f_0 \\ a_0 \end{bmatrix} = 0 \qquad (1)$$

Where $\partial C_L / \partial a$ is the slope of the wing lift line, m is the mass of the wing, k_a is the stiffness of the torsion spring, S is the reference area of the wing, S_a is the static moment of stiffness center ($S_a = m\sigma$), I_a is the inertia of stiffness center ($I_a = I_0 + m\sigma^2$), e is the distance of aerodynamic center to stiffness center, backward for positive; q is the air flow pressure($q = \rho V^2 / 2$), ρ is the air density, f_0 is the deflection of stiffness center, down for positive, and a_0 is the rotation angle of stiffness center, counterclockwise for positive;
The characteristic equation is expressed as:

$$
\begin{aligned}
&A\lambda^4 + B\lambda^2 + C = 0 \\
&A = mI_a - S_a^2 \\
&B = mk_a - I_a k_a (me + S_a) qS \frac{\partial C_L}{\partial a} \\
&C = k_a (k_a - qSe \frac{\partial C_L}{\partial a})
\end{aligned}
\tag{2}
$$

The latent root can be expressed as $\lambda^2 = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$.

When $q = 0, B > 0, B^2 - 4AC > 0$, it is a state of no air flow, the two natural frequencies in the vacuum are obtained, that is, the free vibration frequency; when there is air flow, $q > 0, B^2 - 4AC > 0$, there are still two vibration frequencies change with the value of q; When q increased to the point that $B^2 - 4AC < 0$, λ^2 will become a complex plural. If the real part of λ is positive, the amplitude will continue to expand, when the movement is unstable, that is flutter.

2.2 p-k method flutter equation

In this paper, the flutter frequency of the structure is calculated by the p-k method. In p-k method, the solution for the flutter problem is found by solving the eigenvalue problem.
One key advantage of using the p-k method for determining the flutter characteristics is that it allows flutter analysis to be carried out, based on any given velocity [3]. Here is the simplified equation of plate motion

$$
\begin{cases}
f = f_0 e^{(\gamma + i)\omega t} = f_0 e^{pt} \\
a = a_0 e^{(\gamma + i)\omega t} = a_0 e^{pt}
\end{cases}
\tag{3}
$$

Where f is the displacement of stiffness center, a is the corner of stiffness center, ω is the reduced frequency, then by Eq. (3) the decay rate coefficient γ can be written as

$$
\gamma = \frac{1}{2\pi} \ln \left[\frac{f(t + 2\pi/\omega)}{f(t)} \right] \tag{4}
$$

The *p-k* flutter solution of two-dimensional rectangular structures can be written as

$$
[\frac{V^2}{b^2} \boldsymbol{M} p^2 + \boldsymbol{K} + \frac{1}{2} \rho V^2 \boldsymbol{A}(K)] \begin{pmatrix} f/b \\ a \end{pmatrix} = 0 \tag{5}
$$

Where, $\boldsymbol{M}$ is the mass matrix, $\boldsymbol{K}$ is the stiffness matrix, $\boldsymbol{A}$ is the aerodynamic force matrix, ρ is the free stream air density, V is the velocity, b is a reference semi-chord of the lifting surface, and p refers to the complex response frequency and eigenvalue. The complex response frequency and eigenvalue, p can be expressed as

$$
p = \gamma\omega + i\omega, \; i = \sqrt{-1} \tag{6}
$$

Using the Eq.(5), the flutter phenomenon of two-dimensional plate wing can be found, when γ=0. So, the flutter speed and flutter frequency can also be found, when the sign of damping value changes from negative to positive.

3. Model analysis

3.1 Flutter characteristics of rectangular cantilever wing

Selecting the aluminum alloy as the material of the model, the structure and mechanical properties are tabulated in Table 3-1. In the premise of keeping the half-wingspan l = 1m unchanged, select the reference semi-chord ***b*** as a single variable.

Table 3-1. Rectangular cantilever wing properties

Half-wingspan, l	Thickness, t	Young's Modulus, E	Poisson Ratio, μ	Density, ρ
1.0 m	0.0018 m	70.0GPa	0.3	2700 kg/m^3

When the reference semi-chord ***b*** changes, there is no effect on the bending frequency, but will have a direct effect on the torsional frequency, so that the torsion mode will "jump" between different modes, due to changes in the wing's structural geometry. The main purpose of this paper is to investigate the relationship between frequency coupling modal and the flutter modal. Based on this prerequisite, we find out seven models whose modal frequency of the first-order torsional mode is close to that of the different order bending modes separately. The corresponding values of the reference semi-chord and flutter parameters summarized as shown in Table 3-2.

Table 3-2 Corresponding parameters for 7 models

Label	1	2	3	4	5	6	7
Semi-chord, b(m)	0.3279	0.1069	0.0578	0.0535	0.0316	0.0224	0.0201
Flutter velocity , V_F(m/s)	24.3	33.2	44.4	46.0	58.0	66.0	68.5
The flutter order	2nd	3rd	4th	4th	5th	6th	6th
Coupling order	2nd &3rd	3rd &4th	4th &5th	5th &6th	6th &7th	7th &8th	8th &9th
Coupling vibration mode	2nd-bend& 1st-torsion	3th-bend& 1st-torsion	4th-bend& 1st-torsion	5th-bend& 1st-torsion	6th-bend& 1st-torsion	7th-bend& 1st-torsion	8th-bend& 1st-torsion

Conclusion: It can be seen from Table 3-2 that the flutter order is not exactly the same as the flexural and torsional coupling order, and the coupling order is hysteresis. In addition, with the

increase of the reference semi-chord *b*, the flutter order increases(i.e., the frequency of participation in the coupling decreases)and the flutter velocities assume a decreasing trend and tends to be stable.

3.2 Flutter frequency prediction of rectangular cantilever wing

The first five-order bending frequency, the first-order torsional frequency and the flutter frequency of the six models are plotted in Fig.3-1. Significant trend consistency in natural frequencies can be observed from the flutter curve and the first-order torsional curve, which also meet the frequency coincidence theory. At the same time, the trend also proves that the torsional frequency is the dominant frequency of the flutter coupling in this model. In the results of some studies on the contribution coefficient of the flutter mode, it can be seen that the participation ratio of the natural frequencies is relatively disparity when the flutter occurs.

Therefore, based on the theory of frequency coincidence, using the first-order torsional modal and the first-order bending modal as the main coupling modals of the prediction formula. The main inherent frequencies, the flutter frequencies, the predicted frequencies, and the resulting error analysis of the seven modes are summarized in Table 3-3. The comparison of models' original flutter frequencies and estimated frequencies have been shown in Fig.3-2.

$$F_0 = f_{n1} + f_{t1} / 2 \tag{7}$$

Where, F_0 the estimated flutter frequency, f_{n1} is the first bending natural frequency, f_{t1} the first torsional natural frequency.

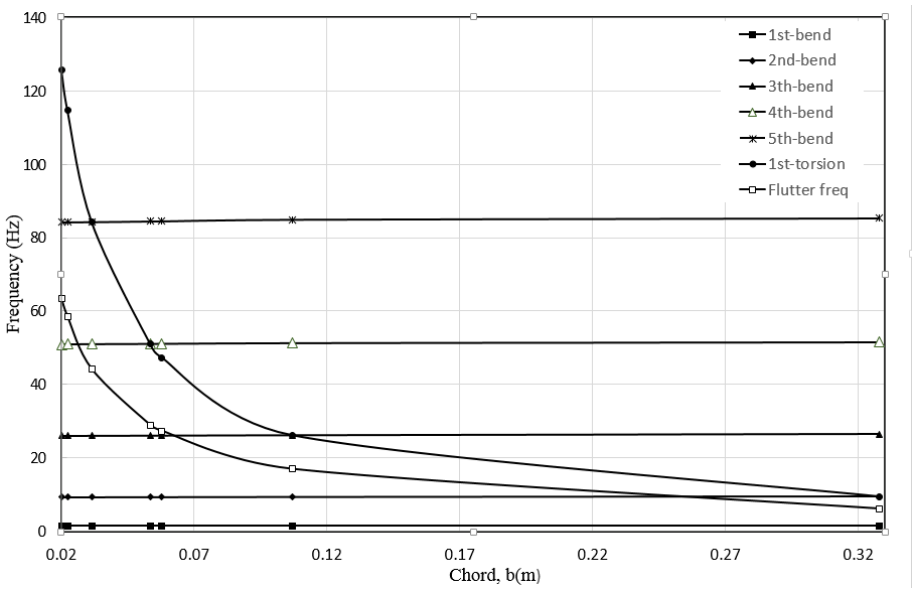


Figure 3-1 the flutter frequency and natural frequencies of seven models

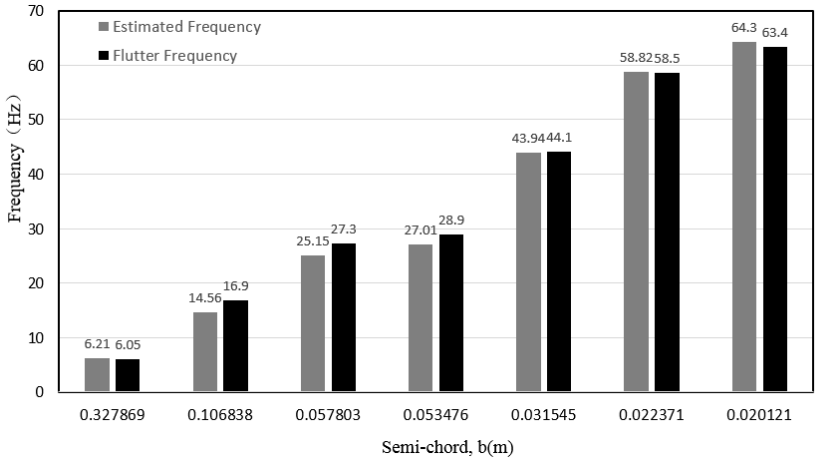


Figure 3-2 Flutter frequencies and estimated frequencies of seven models

Table 3-3 Frequencies and relative error rates of seven models

Label	1	2	3	4	5	6	7
Semi-chord, *b*(m)	0.3279	0.1069	0.0578	0.0535	0.0316	0.0224	0.0201
1st torsional freq (Hz)	1.5084	1.4911	1.4863	1.4859	1.4837	1.4829	1.4827
1st bending freq (Hz)	9.4067	26.140	47.336	51.046	84.128	114.67	125.63
Flutter frequency, *F*(Hz)	6.05	16.9	27.3	28.9	44.1	58.5	63.4
Estimated freq, F_0(Hz)	6.212	14.561	25.154	27.009	43.938	58.818	64.298
Relative error rate (%)	2.68	13.8	7.86	6.53	0.367	0.543	1.41

From Table 3-3, it can be observed that；

1) The frequency coincidence theory can be proved through the data, and there is synergy between the data of flutter frequencies and the 1st-order torsional frequencies.

2) The prediction frequencies based on the 1st-order torsional modal and the 1st-order bending modal are showing a consistency compared with the actual flutter frequencies. In addition to the first model, the relative error rate of the formula is also shrinking with the semi-chord decreases from 0.328 to 0.020

3.3 Regular validation of flutter frequency by trapezoidal cantilever wing

On the basis of the above research, the remaining parameters are kept relatively unchanged, and the symmetrical trapezoidal cantilever model is used to further verify the eq.(7), the tip chord length is chosen as a single variable. Under standard atmospheric pressure, keeping the wing root length b_1 = 0.3m unchanged and selecting the wing tip length b_2 of 0.30, 0.25, 0.20, 0.15, 0.10, 0.05, 0.00 (unit: m) respectively. A right-angle trapezoidal model is selected (R0.1) as a control model when b_2=0.1m. The grid density of the structure model is 15×50 and is 5×10 for the aerodynamic model. The structural meshing and the aerodynamic grid partition are shown in Fig. 3-3. The first four modes for T0.2, R0.1 and T0.0 are visualized in Fig. 3-4.

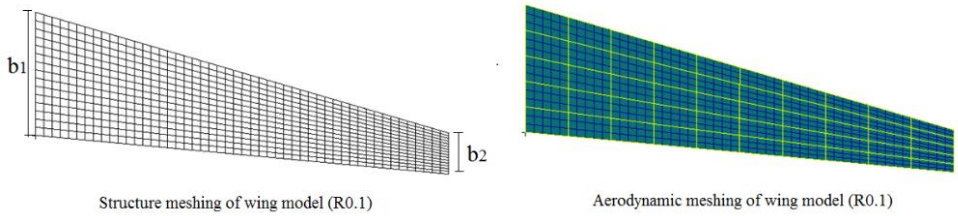


Figure 3-3 The structural meshing and the aerodynamic meshing

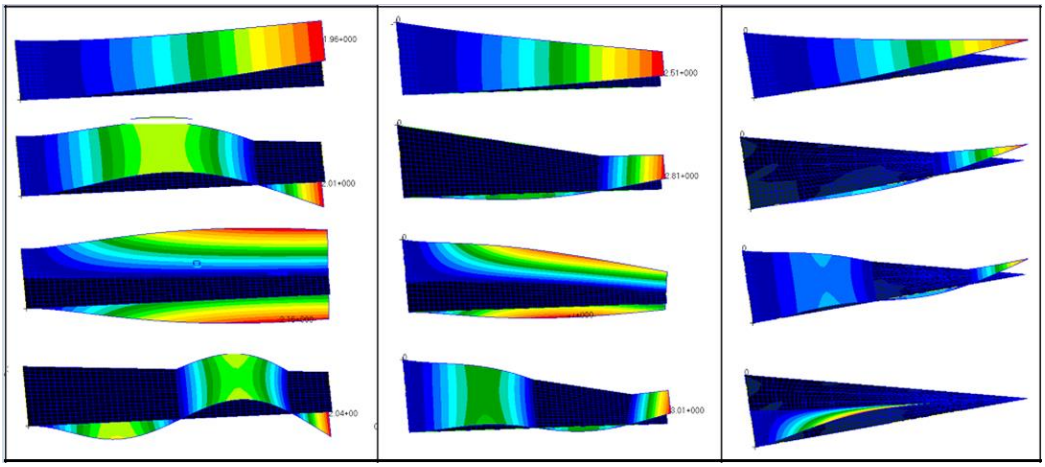


Figure 3-4 The first four - order vibration modes of T0.2、R0.1 and T0.0

According to the forecast frequency of Eq.(7), the estimated frequency can be obtained through the 1st-order torsional frequency and the 1st-order bending frequency. The necessary frequency data and the resulting error analysis of the eight models are organized in Table 3-3. And the comparison between original flutter frequencies and estimated frequencies of seven models is visualized in Fig. 3-5.

Table 3-4 Frequencies and relative error rates of eight models

Label	T0.3	T0.25	T0.2	T0.15	T0.1	T0.05	T0.0	R0.1
Wing tip length, b_2(m)	0.3	0.25	0.2	0.15	0.1	0.05	0.0	0.1
The 1st torsional frequency (Hz)	1.5067	1.5905	1.6976	1.8411	2.0473	2.3792	3.0409	2.031
The 1st bending frequency (Hz)	10.201	11.966	14.311	17.479	21.728	27.110	33.191	21.717
Flutter frequency, F(Hz)	6.5	7.3	8.1	12.8	15.4	18.8	23.1	15.3
Estimated frequency, F_0(Hz)	6.607	7.574	8.853	10.581	12.911	15.934	19.637	12.89
Error rate (%)	1.65	3.75	9.29	17.34	16.16	15.38	14.95	15.75

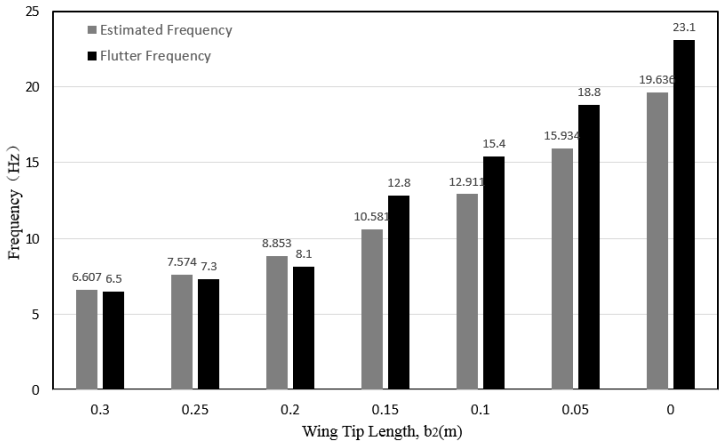


Figure 3-5 Flutter frequencies and estimated frequencies of models

Under the premise of the presumption formula, the relationship between flutter and frequency coincidence theory is analyzed as follows:

1) Through the comparison between original flutter frequencies and estimated frequencies of eight models, it can be observed that the absolute error values are controlled within 3Hz, except the triangle model (T0.0) which has a relatively obvious deviation.

2) The regularity of the Eq.(7) can be proved from Fig. 3-5 that the trend of the data obtained by the bending-torsion coupling law is basically the same as that of the flutter frequency.

3) It can be noticed that the approximation phenomenon of the previous group is different from of the second group, due to the error value does not show obvious regularity in the trapezoidal model.

4. Equivalent spring model simulation

Using the flexibility of the spring elastic support to describe a complex mechanical state of a model, it is placed in the two corresponding corner positions of the end of the cantilever model[6]. By changing the spring stiffness so that its first order vibration equal to the flutter frequency of the original model, and the modal changes after the application of the spring are compared with the flow-solid coupling states caused by the additional aerodynamic forces under the same spring model. Keep the rest of the conditions unchanged, and then add the spring at the midpoint of the end, analysis of the various modes. The inherent frequencies of two spring models are tabulated in Table 4-1.

Table 4-1 The inherent frequencies of two spring models

The flutter order	1st	2nd	3th	4th	5th
Natural frequency of 2 spring bearing model (Hz)	6.5860 (1st-bend)	20.937 (1st-torsion)	21.170 (2nd-bend)	42.724 (3th-bend)	44.817 (2nd-torsion)
Natural frequency of 3 spring bearing model (Hz)	6.6065 (1st-bend)	20.937 (1st-torsion)	21.505 (2nd-bend)	44.818 (2nd-torsion)	45.123 (3th-bend)

It can be concluded from Table 4-1 that in the 2-spring bearing model, keeping the first-order frequency equal to the flutter frequency, the first-order is still a bending model, the final spring stiffness $K = 5\times10^6$. With the same spring stiffness, the torsional frequency of the 3-spring bearing model does not change, but each bending frequency increases. The result is consistent with the flexural and torsional coupling simplified model setting.

5. Conclusion and Further work

From the results presented in this paper, it can concluded that:
1）Rectangular cantilever models with chord length as single variable: after the flutter result analysis of the cantilever structure models, consider the participation difference of the natural modals for the flutter motion, an estimated formula of the chatter frequency is deduced by the first-order torsional modal and the first-order bending modal. The relative errors between the estimated frequency and the actual flutter frequency are controlled within a relatively ideal range. In addition to the first group of values, with the reference semi-chord length decreases, the absolute error rate of the formula is also shrinking.

2) Trapezoidal cantilever models with tip chord length as single variable: analysis the relationship between the presumed formula and the flutter frequency, it can be observed that the absolute error approximation phenomenon of the first group of models does not occur in the second group of models. The main reason is that the wing tip chord changes will have an obvious disturbance to the torsional frequency, so that the value of the error does not show a clear regularity, but the frequency values remains still the same as the first group.

3) When applying the same structural analysis under the elastic support simulation, the application of the spring at the midpoint of the end of the cantilever beam does not affect the torsional frequency which conforms to the theory of plate wing analysis under the assumption of bending and the frequency coincidence theory.

4) Prospects: The speculative formula of this paper provides a convenient way to estimate the approximate range of the flutter frequency, and further confirms the simplicity of the frequency coincidence theory in dealing with some flutter cases. The model types studied in this paper are limited and the theoretical estimates are also relatively rough, thus we can refine and correct natural frequencies weights of the wing model in order to realize the more ideal prediction results in future research.

Acknowledgements:

Project supported by the National Natural Science Foundation of China (Nos. 11672362), the National Science and Technology Major Project(Nos. 2015ZX06004004-003).

References

[1] Li, Z. W., Cao, P. K. and Zhou, T. X. (2008)Numerical simulation of wing flutter based on modal coupled method, *Aeronautical Computing Technique Research Institute* **38**, 43–46.
[2] Yang, Y. X., Ge Y. J. and Xiang, H. F. (2006) Research on the coupled bending-torsional flutter mechanism

for thin plate sections, *State Key Laboratory for Disaster Reduction in Civil Engineering* **23,** 1–8.
[3] Pratik Shrestha, Min-Soo Jeong and In Lee. (2013) Flutter characteristics of a morphing flight vehicle with varying inboard and outboard folding angles, *International Journal of Aeronautical and Space Sciences* **14**, 133–139.
[4] S. Irani, S. Sazesh. (2013) A new flutter speed analysis method using stochastic approach, *Journal of Fluids and Structures* **40,** 105–114.
[5] Shreyas Mandre, L. Mahadevan. (2009) A generalized theory of viscous and inviscid flutter, *Proceedings of the Royal Society A* **466**, 141–156.
[6] Wang, J. R., Zuo, S. G. and Lei, L.(2009) Simulation of static and dynamic stiffness characteristics of convex coil spring based on MSC. Marc, *College of Automotive Engineering of Tongji University* **27,** 807-810.

Examples of applications of two-stage method in calculations of statically indeterminate trusses

***Janusz Rębielak[1]**

[1]Laboratory of Building Structures, Faculty of Architecture, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland.

*Presenting author: j.rebielak@wp.pl

Abstract

The two-stage method belongs to the group of approximate methods of calculation of statically indeterminate systems because differences of the stiffness of component parts, joined in the same node, are not considered in this method. The paper presents a comparison between the values of forces calculated in the members of a selected type of statically indeterminate plane truss under load by forces applied in a symmetrical way and in an asymmetrical way. The comparison is drawn between values of forces calculated for the same trusses by the application of the two-stage method and by the application of suitable computer software. The values of the forces determined by the computer software are considered as exact results because in this case the stiffness differences of the component parts are taken into consideration. The point of the two-stage method is to remove, from the area of the statically indeterminate truss, a certain number of members, which number equals the statically indeterminacy of the basic truss. Appropriate statically determinate truss is calculated in each stage what implies, that is one of very simple methods, like for instance Cremona's method, can be used for this purpose. The trusses calculated in each stage are of the same clear span and they have identical construction depth like the basic indeterminate truss, however they are loaded by forces of half values compared to forces applied to the basic one. The final forces in the members of the statically indeterminate truss are calculated as resultants of forces defined in each stage for members having appropriate positions in the area of the truss system.

Keywords: Truss system, Calculus of vectors, Cremona's method, Superposition method, Statically indeterminate system, Approximate solution.

Introduction

Statically indeterminate systems are calculated by the application of suitable methods like for instance, the force method, the displacement method, the iterations methods, the method of successive approximations, and the finite elements method etc., which nowadays are adapted to the requirements of numerical technology and are applied in various types of computer software [1-4]. Values of forces determined by the means of these methods are considered as exact because in their calculation procedures are taken into consideration, among numerous others, the stiffness differences between members connected to the same nodes. It implies that these methods are of a very complex computational structure what further implies, that procedures developed for suitable computer calculation software have to be very complex. The two-stage method of approximate calculation of statically indeterminate trusses has been developed during initial static analysis of certain types of tension-strut structures. More comprehensive analyses of calculated forces in members of the basic geometry of the truss system are presented in papers [5,6]. In the paper are calculated trusses of similar geometry, however devoid of vertical members. Values of forces determined by means of the two-stage method are compared with the results obtained for the same structural conditions by application of suitable computer software.

Subject of static calculations

Schemes of the initial static calculations have been carried out, according to the principles of the two-stage method, for the truss systems shown in Fig. 1a-c. The basic truss system, see Fig. 1a, is a plane statically indeterminate truss with vertical members. It is loaded by concentrated loads of equal values (F) applied to each node of its upper chord. Figure 1b presents the geometry of the statically determined truss recommended for the first stage of this method together with the way of its loading. The number of members excluded from the geometry of the basic truss equals the degree of statically indeterminacy of the basic system. The scheme of the truss and the way of the loading assumed in the second stage of the proposed method is shown in Fig. 1c. Analyses of results obtained for this type of the truss system are presented in paper [7].

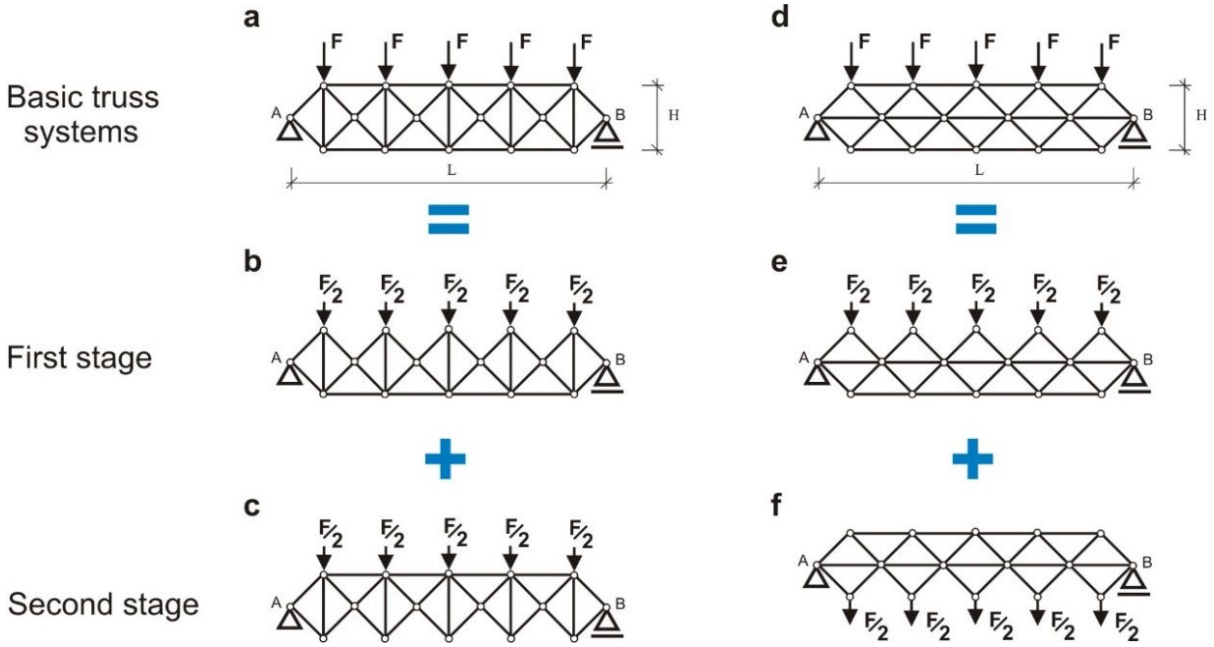


Figure 1. General schemes of calculation procedures applied for, a-c) truss system with vertical members, d-f) truss system not having vertical members

In order to reveal more characteristic features of the two-stage method a series of static calculations have been carried out for the truss system shown in Fig. 1d. The calculated truss has no vertical members but most other geometrical parameters are almost the same like of the truss shown in Fig. 1a, except horizontal members located in the middle chord. Schemes of the appropriate trusses considered in the first stage, Fig. 1e, and in the second stage, Fig. 2f, are defined according to the basic rules of the two-stage method. It was assumed that the number of nodes is defined by symbol "w", while symbol "p" defines number of members. The condition for the inner statically determinacy of the plane truss is defined as follows:

$$p = 2 \cdot w - 3 \tag{1}$$

The considered truss system shown in Fig.1d is built by a number of nodes, w = 16, what implies that the statically determinate truss created by means of this number of nodes has to be built by the help of the following number of members:

$$29 = 2 \cdot 16 - 3 \tag{2}$$

Because the truss shown in Fig.1d is built by the number of members, p = 33, it indicates that the considered structure is the fourfold the statically indeterminate system. It implies that in order to make the basic truss a statically determined system one should exclude 4 suitable members. Proceeding according to rules of the two-stage method in its first sage one should exclude 4 members of upper chord, see Fig 1e, and to apply half values of the unit forces to the same nodes like in the basic truss, see Fig. 1d. In the second stage it is necessary to exclude also

4 members from the lower chord, see Fig. 1f, but in this case suitable load forces have to be applied also to nodes of the lower layer. In each stage the statically determinate truss is calculated, due to which one can apply one of the simple methods like e.g. Cremona's method, to determine values of the forces in its members. Final values of the forces in the basic system will be resultants of forces calculated in each stage for members of appropriate positions. All the assumptions are congruent with rules of calculus of vectors, principle of superposition and with the three fundamental conditions of equilibrium presented below:

$$\sum_{i=1}^{n} F_{ix} = 0 \tag{3}$$

$$\sum_{i=1}^{n} F_{iy} = 0 \tag{4}$$

$$\sum_{i=1}^{n} M_{i} = 0 \tag{5}$$

The statically indeterminate truss system, of geometry of members shown in Fig. 1d, is subject of the static calculations. It was assumed that the basic truss is subjected to two types of loading. In the first one the unit load forces F of value equal to 1,00 kN, are symmetrically applied to all nodes of upper chord. In the second type of loading these unit forces F are applied in an asymmetric way only along a part of its clear span. In this case it means that these loads are applied only to two successive nodes of the upper chord, which are adjacent e.g. to support node B. Basic structures are of 5,00 m clear span and their construction depth is equal to 1,00 m.

Values of forces calculated for symmetrically loaded truss

The first static calculations have been made for symmetrical loading of the basic truss. In the first stage the concentrated forces, each of value equal to 0,50 kN, are uniformly distributed to all nodes of the upper chord of a suitably assumed scheme of the statically determinate truss. Values of forces defined in this stage for all members of the truss, together with appropriate Cremona's polygon of forces, are shown in Fig. 2. In the horizontal member located, for example, between nodes 12 and 13 acts as a tension force of value equal to +3,00 kN. Because in this stage the truss member is not present between node number 3 and node number 4, that is why for this area the force value is estimated as equal to zero. In this stage the value of the compression force calculated for cross brace placed between nodes 3 and 8 equals -0,353 kN. The horizontal member placed between e.g. node 8 and node 9 acts compression force of value equal to -2,250 kN.

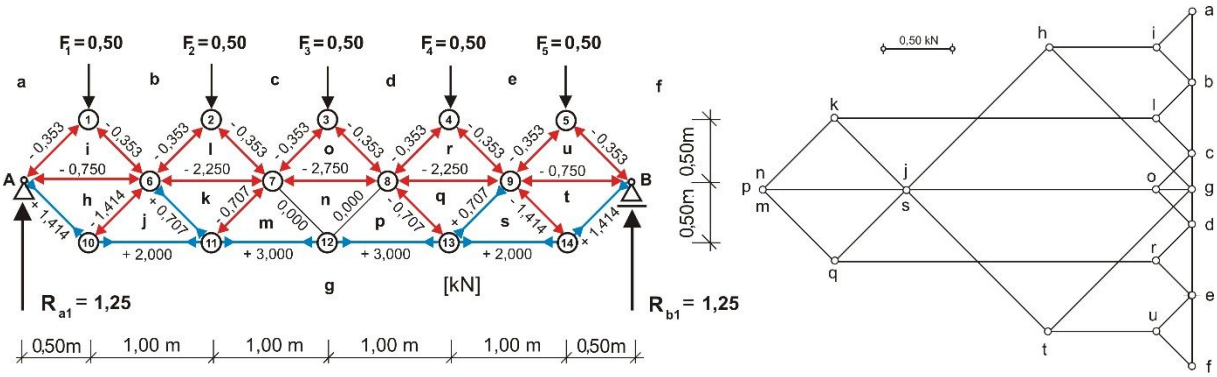


Figure 2. Values of forces determined in the first stage of calculations for symmetrically loaded truss together with Cremona's polygon of forces

Static scheme of the investigated truss considered in the second stage of calculations, together with their results and Cremona's polygon of forces, are shown in Fig. 3.

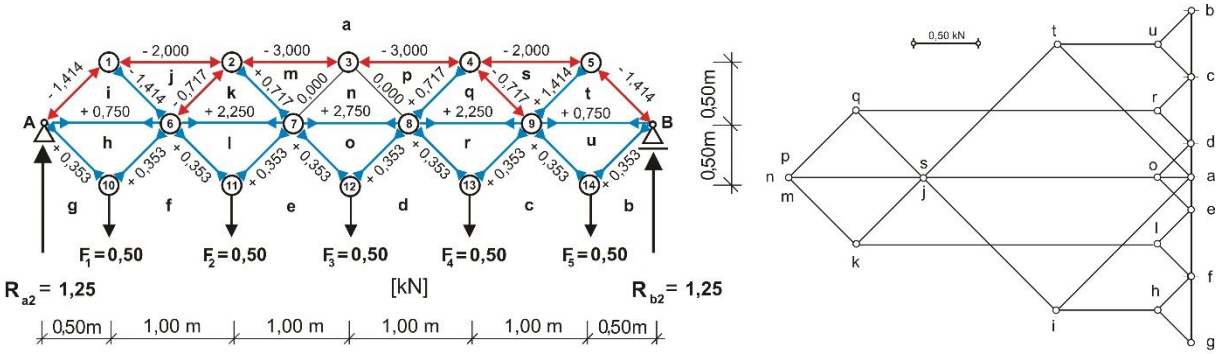


Figure 3. Values of forces calculated in the second stage of calculations for symmetrically loaded truss together with Cremona's polygon of forces

The values of the forces calculated in the second stage for some counterpart members are following. In the cross brace located between nodes 3 and 8 acts the force of zero value, a horizontal member placed between nodes 8 and 9 is subjected to an act of tension force having the value +2,250 kN, see Fig. 3. Moreover in the member of the upper chord located between nodes 3 and 4 acts compression force of the value equal to -3,000 kN. The final values of the forces calculated by the means of the two-stage method in members of the basic form of statically indeterminate truss are presented in Fig. 4a. For instance the final value of force acting in the cross brace placed between nodes 3 and 8 is defined as -0,35 kN and it is the resultant of the force value calculated for the same member in the first stage, (-0,353 kN) and the force value calculated in the second stage (0,000 kN). Similar procedures have been done for all members of the calculated truss.

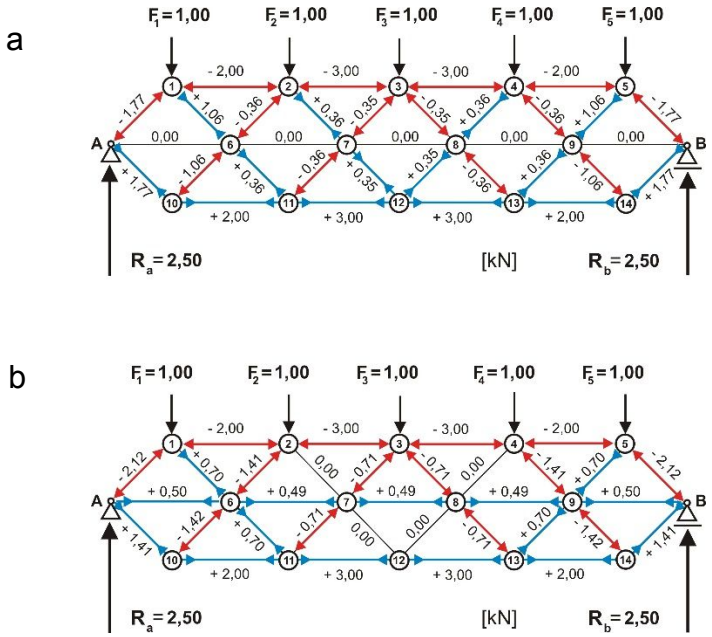


Figure 4. Values of forces in members of symmetrically loaded truss calculated, a) in the two-stage method, b) by means of suitable computer software

The same basic truss has been subjected to static calculation carried out by the application of the Autodesk Robot Structural Analysis Professional 2016, for which the software is designed for the precise calculation of force values acting in members of the statically indeterminate systems. Static calculations were made by assumption that the truss consists of steel tubular members having diameter of 30.00 mm, the thickness of the section equals to 4.00 mm and the steel material has the Young's modulus equal to 210 GPa. Results achieved in this way are presented in Fig. 4b. From the analysis of results obtained from both compared methods for the

same members follows, that the biggest forces calculated for the same members are of the same values and of the same sense. This remark refers to the forces acting in the members of the outer chords of the truss. One can notice bigger differences in values of the forces calculated in the cross braces. For instance the value of the compression force defined in the two-stage method in members located, for example, between node 5 and support node B is equal to -1,77 kN, while by application of computer software the compression force acting in the same cross brace is estimated as -2,12 kN, which constitutes about 16,5 % of the biggest value. Much bigger relative differences are noticeable in other members, however the absolute values of these forces are rather very small.

Values of forces calculated for asymmetrically loaded truss

The second group of calculations have been made for the same basic form of the statically indeterminate truss but this time loaded in an asymmetric way. In this case the load forces are applied only to two nodes of upper chord, which are located in the close vicinity to the support node B. In the first stage of the two-stage method there are removed four members of the upper chord and then to suitable nodes of this chord are applied load forces of half unit, which are of value equal to 0,50 kN, see Fig. 5.

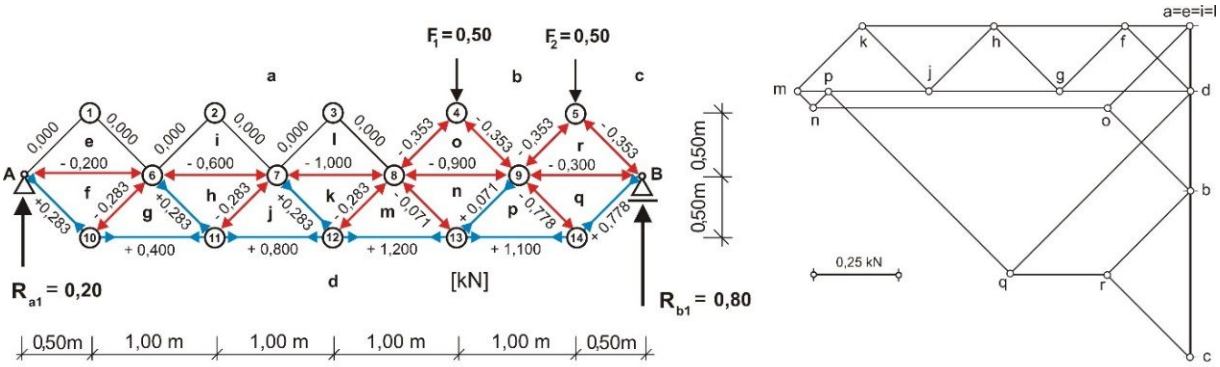


Figure 5. Values of forces determined in the first stage of calculations for asymmetrically loaded truss together with Cremona's polygon of forces

According to the rules of the two-stage method in its second stage it is considered a statically determinate truss, the pattern of which is defined by removing four members from the bottom chord, see Fig. 6. Then to appropriate nodes of this chord are applied load forces having a value of half of the unit load. In this case values of the load forces are also equal to 0,50 kN. The final values of the forces calculated by the means of the two-stage method in members of the investigated truss are sown in Fig. 7a.

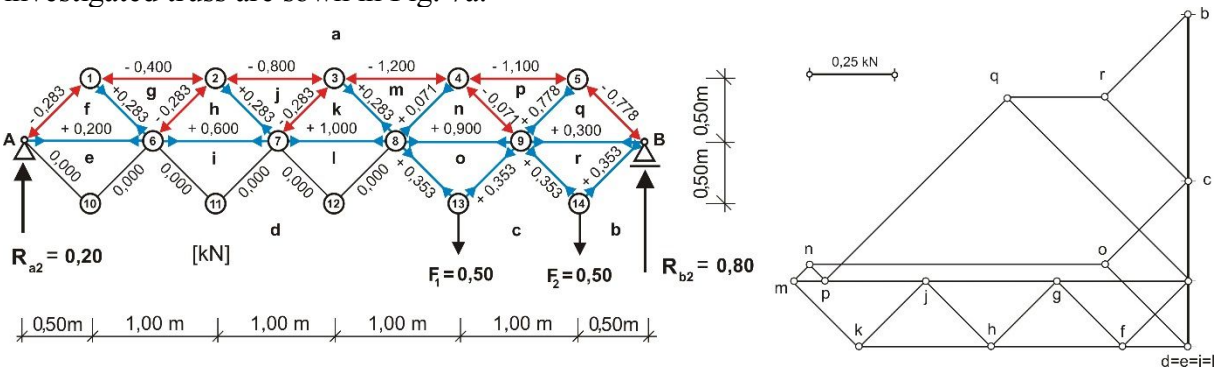


Figure 6. Values of forces defined in the second stage of calculations for asymmetrically loaded truss together with Cremona's polygon of forces

A statically indeterminate truss of the same static scheme was calculated also by the application of the Autodesk Robot Structural Analysis Professional 2016 for identical material parameters

of component parts as it was previously listed. The results of these calculations are presented in Fig. 7b. The values of the forces acting in the members of the outer chords calculated in both methods are identical. Differentiation of force values defined for other members of the basic statically indeterminate structure is bigger. For instance the value of the tension force calculated in the two-stage method for the cross brace placed between node 8 and 13 equals +0,28 kN, see Fig. 7a, while by application of computer software in the same member is defined a compression force of value equals -0,07 kN, see Fig. 7b. In this case forces calculated for the same member in two compared methods are of different values as well as of various senses. In spite of this relatively big differences one should notice that they appear only in members subjected to act of the smallest values of forces.

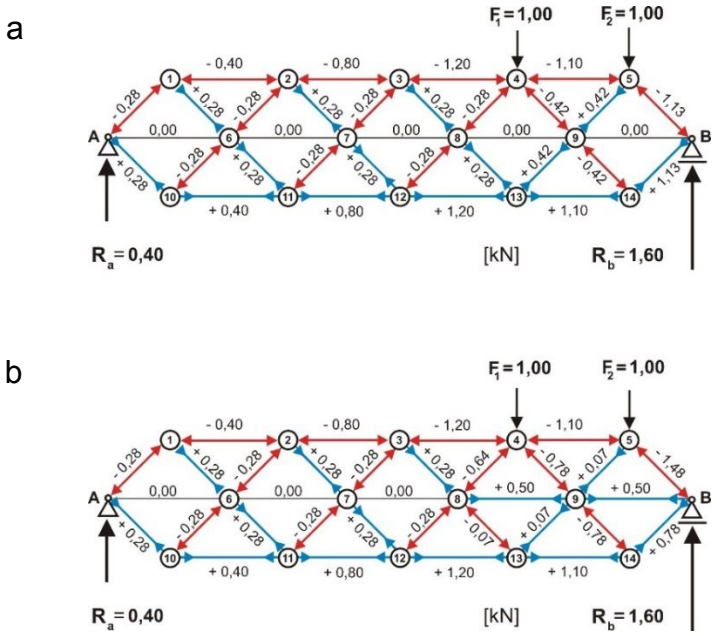


Figure 7. Values of forces in members of asymmetrically loaded truss calculated, a) in the two-stage method, b) by means of suitable computer software

Conclusions

The two-stage method for the calculation of the statically indeterminate trusses easy produces results which closely approximate the values of forces determined by the application of the exact methods of calculations. Accuracy of the results obtained by the two-stage method can be significantly enhanced by the application of suitable sets of coefficients determined individually for each node, due to which the stiffness differences between members connected in a particular node can be taken into consideration during definition of the final values in the calculation processes.

References

[1] Timoshenko, S.P. (1966) *History of strength of materials*, Arkady, Warszawa, - in Polish
[2] Makowski, Z.S. (1981) *Analysis, design and construction of double-layer grids*, Applied Science Publishers, London.
[3] Zienkiewicz, O.C. and Taylor R.L. (2000) *The finite element method*, Oxford Press, UK.
[4] Allen E., Zalewski W. and Boston Structures Group (2010) *Form and forces. Designing efficient, expressive structures*, John Wiley & Sons, Hoboken, New Jersey.
[5] Rębielak, J. (2014) A two-stage method for an approximate calculation of statically indeterminate trusses, *Journal of Civil Engineering and Architecture* **78**, 567-572.
[6] Rębielak, J. (2015) Examples of application of principle of superposition in the design of structural systems and in static analyses, *Journal of Mathematics and System Science* **5**, 150-155.
[7] Rębielak, J. (2016) Simple method of approximate calculation of statically indeterminate trusses, *Proceedings of The 7th International Conference on Computational Methods (ICCM2016)*, August, 1-4, 2016, University of California, Berkeley, USA, 748-753.

Structural concept of system of combined foundation designed for buildings located in earthquake areas

***Janusz Rębielak[1]**

[1]Laboratory of Building Structures, Faculty of Architecture, Cracow University of Technology, ul. Warszawska 24, 31-155 Kraków, Poland.

*Presenting author: j.rebielak@wp.pl

Abstract

The paper presents initial structural and static analyses of an innovative type of foundation structure, which can obtain theoretically unlimited horizontal surface. This implies that the proposed system of combined foundation can be applied for construction of heavily loaded buildings located on subsoil of very small load carrying ability. Basic rules of theory of structures, like e.g. the principle of superposition, are respected during process of design of the proposed structural system. Important inspiration in this process have played patterns of stress trajectories in the free-end-beam as well as the concept of Michell's beam. The inner build of the system of combined foundation makes possible the very uniform distribution of the load forces onto its horizontal surface. Moreover this system has inherent structural features of a damper, due to which it is able to absorb safely a significant part of energy of vibrations evoked by earthquakes. Suitable application of the proposed system makes possible to design safe structures of high-rise building located in earthquake areas.

Keywords: Foundation, Building structure, Subsoil, Load capacity, Earthquake, Tall building.

Introduction

Structural system of a building, together with its foundation, has to be designed according basic rules of theory of structures and application of modern theories [1,2,3]. Safety problems of structural systems of buildings located on ground of small loading ability or in earthquake areas belong to the constantly actual research tasks in engineering [4-7], likewise aspects of damping of the energy vibrations caused by the earthquake and accidental land's slides or casual displacement of a ground beneath the foundation structure [8-11]. In case of tall buildings their bearing systems should provide them with suitable rigidity and stability defined by buildings codes mostly under the wind load. Tall buildings, as well as all types of buildings, are time to time subjected to act of huge values of dynamic loads. Bearing systems of such objects have to at the same time of two contradictory structural features; on the one hand they have to be very stiff but on the other hand they have to be flexible to some degree [12-14]. Some similar remarks refer also to the foundation systems, while structural problems to be solved are very complex. There are numerous structural systems of foundation applied in the practice for a long time, which are permanently modified and gradually improved. They may be subdivided generally into two main groups of the shallow and the deep foundation systems. Systems of the shallow foundations are mostly applied when the passive earth pressure beneath the bottom surface of the foundation is also within the range of load carrying ability of the ground and they are usually not complex because of simplicity of their structural forms. Complexity of the deep foundations makes them often difficult to construct that is why they are the very expensive technical solutions.

Definition of expected structural features of foundation system

Systems of deep foundations are complex, they need a long time to be constructed, they are expensive technical solutions and their applications may have serious impact onto the

environment. Thus it seems to be more convenient to apply one of shallow types of the foundation systems but application of this concept is in certain cases also not easy. The requested foundation has to be of considerable big horizontal surface or length (L1), see Fig. 1a, while the clear span of cantilever part of the foundation structure (Ks) is strictly limited by its mutual relationship to the construction depth (Fd), what indirectly is determined by permissible deflection of the cantilever structure. Surface of the foundation may be significantly enlarged in simple way by increasing of its horizontal dimensions along two orthogonal directions, due to which value of passive earth pressure could be significantly smaller than in the typical configuration of the shallow foundation placed beneath the ground floor of a building. Expected shape of the widely spread foundation should be characterized by big stability under acting of complex set of big loads, including vertical load (G), sometimes huge horizontal load caused by wind (W), see Fig. 1a. and also under acting of forces caused by rapid dislocation of subsoil along its parameter (Gd) or e.g. evoked by displacement of ground beneath even a vast part of foundation area (Gd2) of the foundation. New structural system should provide the foundation with very large horizontal surface, while its structural depth has to be relatively very small. Structural task defined in this way is very difficult to solve because the searched technical solution has to fulfil at the same time some contradictory structural requirements, which have been shortly discussed above. In processes of solving such complex tasks, there are mostly very helpful conclusions following from analyses of structures existing in the nature as well as the suitable application of basic rules and theorems of theory of structures referring to stress trajectories in a free-ends beam or to structural concept of Michell's beam.

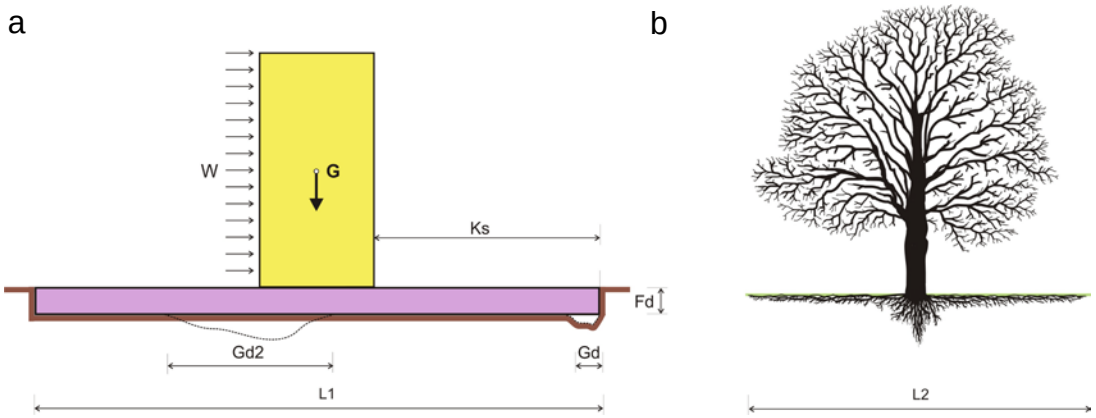


Figure 1. Analytic scheme of vertical cross section of building having very large form of shallow foundation and similar scheme of tree structure together with its root system

Shape of the root system of certain kind of a deciduous tree, see Fig. 1b, was one of inspirations in the design process of the proposed structural system of foundation, which could obtain the above defined features. The own weight of a tree is transmitted through its trunk to extensive root system composed of numerous small components distributed quite uniformly and shallowly in the subsoil.

Structural concept of proposed foundation

Heavily loaded object can be mostly located safely on subsoil of very small load capacity if its shallow foundation will have big enough horizontal surface. It can be represented by an example and simple form of foundation structure, see Fig. 2, which consists of two beams (Bm) parallel to each other, located on a common horizontal slab (SLb), joined also by help of distanced elements (De) put in selected places of upper parts of beams, see Fig. 2d, [15]. The length of these beams should be considerably big (Ld). In this case the concentrated force (Fk) can be spread onto a suitable large surface. The point of the proposed technical solution is to arrange, in a narrow space between two main beams, an intermediate system (Int) of straight structural members connected together by means of structural nodes (Sn) being theoretically the articulated joints. Structural nodes (Sn) have to be not connected to the main beams. Outer concentrated force (Fk) is applied to upper node (A1) of a short vertical member

situated between boundary nodes A1 and B1, which is put in vertical guides and having possibly of displacement only along the vertical direction.

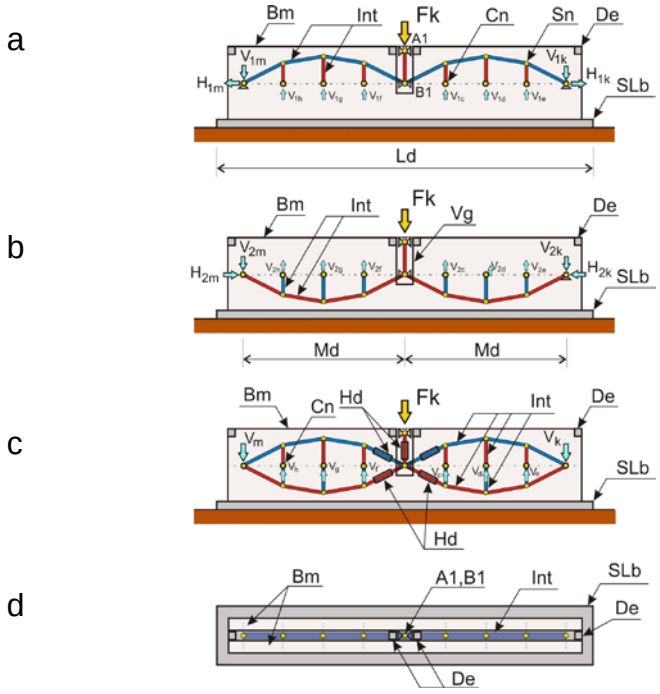


Figure 2. General schemes of basic structural configurations of the system of combined foundation

The intermediate system is built of two subsystems, while each of them is connected to lower node (B1) of the short vertical member. In this particular case the subsystems are composed of two parts arranged symmetrically towards vertical axis of central member put between nodes A1 and B1. The first one of these subsystems is called funicular system, see Fig. 2a, and it is built of struts placed along broken curves of convexity directed up. Structural nodes (Sn) of this subsystem are the upper nodes of short vertical members, lower nodes of which are the central nodes (Cn). In the boundary central nodes act, except vertical (V_{1m}, V_{1k}), also the horizontal components of reactions (H_{1m}, H_{1k}) and senses of the last vectors are directed outside the foundation. The central nodes are uniformly arranged along the neutral horizontal axis of the main beams, where deflections of bending beams are theoretically the smallest. Second subsystem is called an arch system, see Fig. 2b, its structure is similar to the first one but its pattern is symmetrical towards the neutral axis of horizontal beams. The main members are located along a broken curve of the convexity directed down. As previously structural nodes are connected by means of short vertical members to the central nodes. In this case in the boundary nodes act horizontal components of reactions (H_{2m}, H_{2k}) of the same values like reactions H_{1m} and H_{1k} but they are directed oppositely, it means towards center of the foundation. When these both subsystems are combined together, see Fig. 2c, then in the boundary central nodes act only vertical components of reactions (V_m, V_k), which in this case are directed down. The distance between node B1 and the boundary central nodes can be considered as structural module (Md) of system of the combined foundation. The final form of the intermediate system takes a lenticular shape, due to which it has some inherent features of damping of energy of vibrations evoked by dynamic load. This ability can be significantly increased e.g. by arrangement of computer controlled hydraulic jacks (Hd) in selected members of the intermediate system as well as in the main vertical member situated between nodes A1 and B1. The hydraulic jacks will be integral parts of structures of suitable members. The main beams and horizontal slab will be made as concrete or reinforced concrete structures. Component parts of the intermediate system can be made as steel members, as well as the reinforced concrete structures or certain types of composite materials of very high strength. Proposed structural system of foundation has been patented [16].

The proposed system of combined foundation can be shaped in various ways, see Fig. 3. Members of the intermediate system can be arranged on both sides of a single beam, see Fig. 3c, what implies that appropriate pairs of central nodes have to be joined by bolts interpenetrating matter of the beam (Bm), due to which the reaction forces are applied theoretically in the middle points inside space of the beam. Single segments of the intermediate system, having modular length (Md), can be repeated many times along the horizontal direction. Adjacent segments have to be connected by help of special type of central nodes B1 or Cx, which are also not connected to matter of the beam and have freedom of dislocation along vertical plane. Number of replications of single segments of intermediate system is optional what implies, that foundation structure designed in the proposed way may obtain extremely large length or surface, which theoretically can be unlimited.

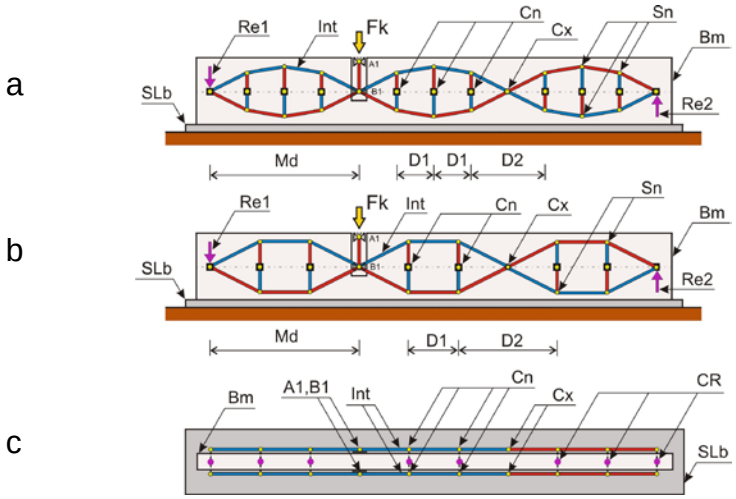


Figure 3. Simplified schemes of arrangement of components of intermediate system around single shape of the main beam

Vectors of reactions acting in boundary central nodes (Re1, Re2), depending on mutual structural configuration and number of replication of these modular units, can be of various senses. The outer concentrated force (Fk) can be apply at any part of the foundation but always according to structural requirements described previously. Regular geometric subdivision of single modular unit does not ensure uniform distribution of vertical components of reactions appearing in the central nodes. However distances between these nodes within single modules are of the same length (D1) but distances between two central nodes in two adjacent modules (D2) are twice bigger. The main structural goal of proposed system of the combined foundation is to distribute the concentrated load force (Fk) to numerous central nodes (Cn) uniformly arranged along neutral axes of horizontal beams, where vertical reactions can be of smaller values and will be applied to points (CR) located in the middle plane of a beam, see Fig. 3c. When number of central nodes located within a single module (Md) is big, then values of these reactions can be significantly smaller. But on the other hand the shape of the foundation structure is complex what causes, that useful advantages of its inner space are considerably restricted. Therefore it is proposed to reduce the number of central nodes only to two, which are situated within the structural module, see Fig. 3b, as well as to slightly modify geometry of components of the intermediate system. The point of the modification is to unify sizes distances (Dn) between central nodes, see Fig. 4a, what is presented on example of foundation, see Fig. 4b, built by means of single beam (Bm) put on slab (SLb). Members of the intermediate system are connected together in suitable nodes at the right angle or at angle of 45 degrees. Due to this recommended geometry the necessary technical or technological openings inside the main beams can be relatively big. Alternately directed senses of vectors of the vertical reaction forces acting in some adjacent pairs of central nodes cause, that the main beams (Bm) have to be of appropriate big stiffness because they are subjected to big values of bending moments and shear forces acting along the vertical surface. Construction depth of beams will be equal to one or to a set of typical floors.

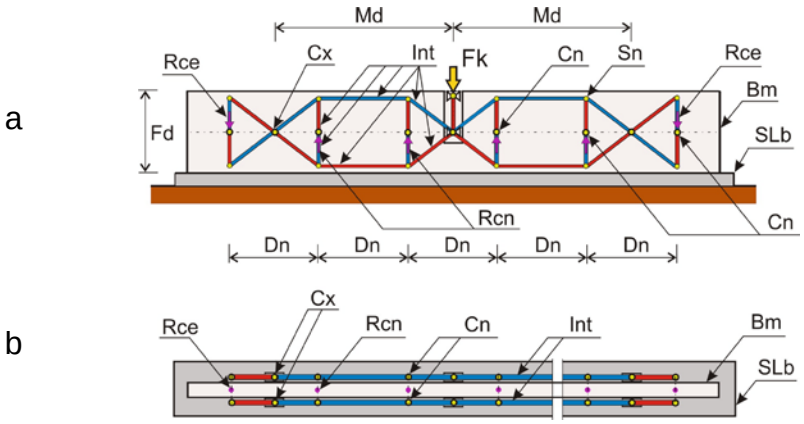


Figure 4. Recommended geometry of components inside space of proposed structural system of foundation

One should strive to distribute the main load in form of concentrated force (Fk) along at least two various horizontal directions, what ensures significant decreasing the values of reactions applied in the central nodes of the large horizontal surface of the combined foundation shaped like a specific type of foundation framework. Figure 5 shows scheme of structural module of such foundation formed around selected part of single beam having modular length (Md) and construction depth (Fd). This modular part of beam has suitable cut-outs located in the middle of vertical edges, where in the free spaces are located boundary nodes of single segments of the intermediate system (Int) arranged symmetrically around both vertical sides of the main beam. When the structural module, consisting of a beam and intermediate system, will be situated in perimeter zone of the foundation structure then its outer side have to be protected from the adjacent ground by help of additional border wall put vertically in relatively short distance to the main matter of beam.

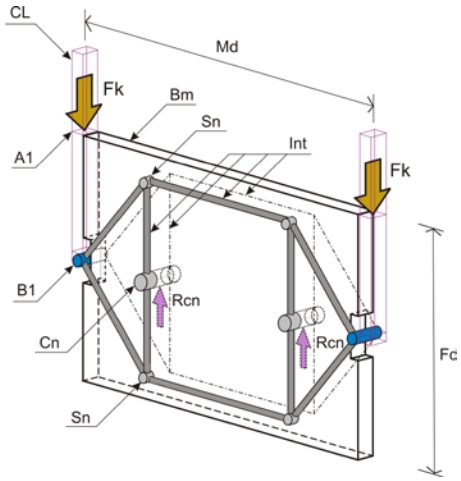


Figure 5. Example of shape of structural unit of proposed type of foundation

Such structural units can be component parts of the foundation frameworks having patterns of rectangular, triangular or other type of a grid. Figure 6a shows an exemplary shape of set of four structural modules creating spatial module of square grid of foundation framework. Modular segments of the main beams (Bm) should be connected together by means of e.g. rigid joints (Rjt), which can be made in various ways. Lower parts of shorts struts placed between nodes A1 and B1 are connected to the middle parts (Mp) of nodes of the type B1 placed in spaces of suitable cut-outs (Bx). Geometric dimensions of this spatial foundation module, measured along horizontal directions, can be in the real structure quite big as well as the value of its construction depth (Fd) what implies, that its inner space has to be supplemented by some additional structural members. Figure 6b shows simplified scheme of such foundation module containing certain number of vertical columns (Col) located inside the space of a structural box, which columns are intermediate supports for slabs (FL) of the

inner floors. In each vertical side wall (Bm) can be arranged appropriate and large technical openings (Op), presence of which makes possible to design the inner spaces of the foundation for numerous useful purposes of the building.

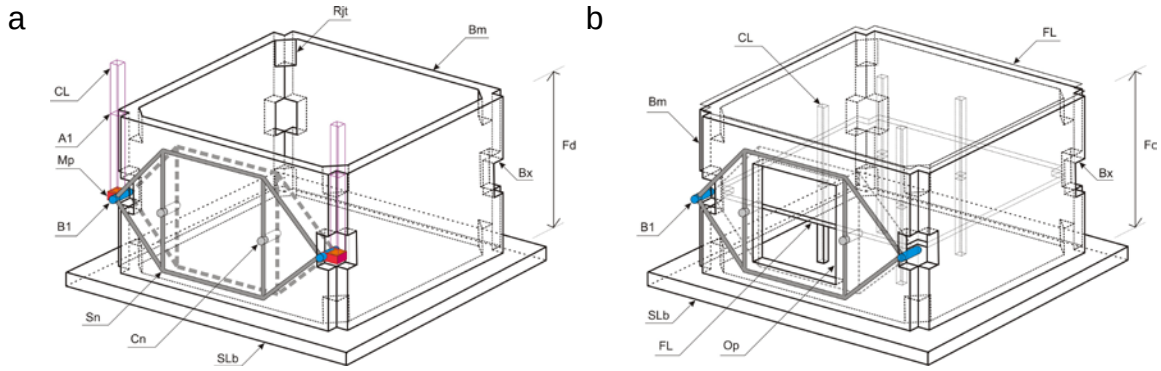


Figure 6. Views of simple forms of spatial structural module of system of combined foundation

Structural system of the combined foundation makes possible to spread the matter of shallow foundation on the large horizontal surface by keeping the basic rules of theory of structure. It can be also a stabile support for objects even after large displacement of subsoil beneath the foundation structure. Therefore, this shape of technical solution can be consider as satisfactory solution of the research problem defined previously. However, the structural system can be considered as somewhat complex but it enables to construct the safe buildings in areas of very difficult or dangerous environmental conditions.

Proposal of application

System of combined foundation can take various forms and it fulfils the main structural conditions required from foundation of buildings located in the earthquake areas defined at the beginning of this research. The main body of this foundation can be of very big length or large surface, Fig. 7a, and it can be horizontally separated (Hrs) from the background.

Proposed type of combined foundation, due to its very large surface, can ensure stability to the whole object even after dynamic dislocations of big parts of the ground. In selected members of the intermediate system (Int) one can put the electronic controlled hydraulic jacks (Hd), what will considerably increase the inherent ability of damping vibrations, which is characteristics for this structural form. Because to the boundary nodes (Ce) are applied vertical reactions directed down therefore, it is proposed to locate below these nodes additional foundation plies (Fp) in order to stabilize this part of structure.

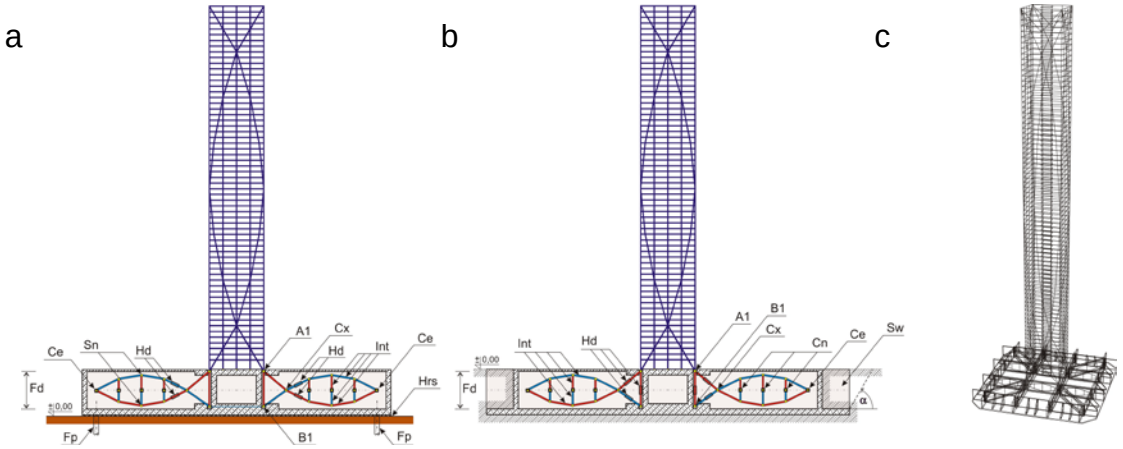


Figure 7. Schemes of combined structural system of tall building located on proposed type of combined foundation

Another way of stabilization of these parts is to shape there a type of big scoop, see Fig. 7b, where the weight of ground located in the soil wedge (Sw), determined by angle of slide (α), will play similar role like the foundation plie. When the multi-storey building is supported on the proposed type of foundation and inside its aboveground structure is applied suitable bracing in shape of the lenticular girder, then the whole bearing structure is called the combined structural system of the tall building. Figure 7c shows simplified scheme of structural system of tall building shaped in similar way, being a suitable visualization of numerical model of the system defined in programming language Formian [17]. Initial static analyses of this structural system have been done on basis of this numerical model by application of computer software Autodesk Robot Structural Analysis Professional 2016.

Conclusions

The proposed structural system of combined foundation fulfils requirements defined at the beginning of the research task. Its horizontal surface is theoretically unlimited and it can be the safe and reliable foundation structure for the heavily loaded buildings, which have to be placed even on ground of very small load capacity or located in seismic area. Application of this structural system does not need to make deep foundation trenches, that is why its presence should not destroy the natural underground water system and it can be relatively inexpensive technical solution. However, it is somewhat complex structure but it will make possible the safe foundation of objects in areas, which nowadays are considered as difficult or almost impossible to use for building purposes. It is expected that specific form of system of combined foundation can be applied to straighten the previously inclined objects. The proposed structural solution of combined foundation has to be subjected to numerous and complex structural, static and dynamic analyses and also to research tests in order to verify all the assumptions and expectations.

References

[1] Allen E., Zalewski W. and Boston Structures Group (2010) *Form and forces. Designing efficient, expressive structures*, John Wiley & Sons, Hoboken, New Jersey.

[2] Timoshenko, S.P. (1966) *History of strength of materials*, Arkady, Waszawa, - in Polish

[3] Zienkiewicz, O.C. and Taylor R.L. (2000) *The finite element method*, Oxford Press, UK.

[4] Kowalczyk, R.M. and Sim, R. and Kilmister, M.B. eds. (1993) *Structural systems for tall buildings*, Council on Tall Buildings and Urban Habitat, McGraw-Hill, New York.

[5] Allen, E. and Iano, J. (2014) *Fundamentals of building construction. Materials and methods*, John Wiley & Sons, Hoboken.

[6] Coduto, D.P. (2001) *Foundation design. Principles and practices*, Prentice Hall, NJ.

[7] Das, B.M. (2014) *Principles of foundation engineering*. SI, Global Engineering, Stamford, CT.

[8] Hamburger, R.O. and Krawinkler, H. and Malley, J.O. and Adan, S.C. (2009) Seismic design of steel special moment frames. A guide for practicing engineers, *NEHRP Seismic Design Technical Brief No. 2*, National Institute of Standards and Technology, USA.

[9] Hamburger, R.O. (2009) *Earthquakes and seismic design*, Facts for Steel Buildings, American Institute of Steel Construction, USA.

[10] Hansbo, S. *Foundation engineering* (1994) Elsevier, Amsterdam, London, New York, Tokyo.

[11] Greenfield, S.J. (1992) *Foundation in problem soil. A guide to lightly loaded foundation construction for challenging soil and site conditions*, Prentice Hall, NJ.

[12] Ali, M.M. and Moon, K.S. (2007) Structural developments in tall buildings. Current trends and future prospects, *Architectural Science Review*, **50** (3), 205–223.

[13] Chew, M.Y.L. (2012) *Construction technology for tall Buildings*, World Scientific Publishing Co. Pte. Ltd, Singapore.

[14] Matso, K. (1995) Lessons from Kobe, *Civil Engineering*, **4**, 42-47.

[15] Rębielak, J. (2012) System of combined foundation for tall buildings, *Journal of Civil Engineering and Architecture*, **61**, 1627-1634.

[16] Rębielak, J. (2016) *Systemowy fundament zespolony, (System of Combined Foundation - in Polish)*, Patent No 221971, Patent application No P.394745, Patent Office of Republic of Poland.

[17] Nooshin, H. and Disney, P. (2000) Formex configuration processing 1, International Journal of Space Structures, **17**(1), 1-52.

A Molecular Dynamic Study on Nonlinear Vibration Behaviors of Fe Nanowires

***Zhuoqun Zheng[1], Eric Li[2], Nan Ding[1], †Xu Xu[1]**

[1]College of Mathematics, Jilin University, 2699 Qianjin Street, Changchun, 130012, China
[2] Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong, China

*Presenting author: zzq20050305@126.com

†Corresponding author: xuxu@jlu.edu.cn

Abstract

In this paper, nonlinear behaviors in the vibration of Fe nanowires are investigated by using the large scale Molecular Dynamics (MD) simulations. It is observed that the vibration frequency of nanowires rises slightly and nonlinearly with the increase of initial actuation amplitude. Based on the atomic arrangement, a discrete spring-mass model is developed. Its geometric nonlinearity is used to explain this phenomenon. In addition, Fe nanowires with different sizes show different vibration properties in this work. The ratio between the length (L) and the height (h) of nanowires has a significant influence on vibration behaviors. The vibration frequency changes linearly with h/L^2 when the ratio is relatively large, while it changes nonlinearly when the ratio is relatively small.

Keywords: nonlinear vibrations, Fe nanowires, molecular dynamics, initial actuation, size, spring-mass model.

Introduction

Nanowires have been widely investigated with the rapid development of nano science and technology. Due to their unique electronic, thermal, mechanical and optical properties, nanowires have been widely applied as active components of nanoelectromechanical systems (NEMS), such as force and pressure sensing [1, 2], nanowire–nanopore sensors [3], field effect transistor [4], lithium battery anodes [5] and other devices [6-8]. These NEMS utilize the nanowire as a resonating beam, in which the nanowire vibrates continuously at or near its resonant frequency. In addition, the changes in local environment including force, pressure or mass can be detected by the corresponding changes in the resonance frequency of the nanowire [9]. Therefore, it is of great significance to study the nanowire's mechanical properties under vibration.

In the past few years, there have been lots of studies on the nanowire's vibrational behaviors and properties, including the experimental studies [10-13], theoretical modeling [14-18] and computational simulations [19-26]. Copper, silver, gold and silicon nanowires have been studied on their vibrational properties respectively by several researchers [19, 22-26]. Two-dimensional vibration has been studied by Conley et al. theoretically [14] and used for mass sensing and stiffness spectroscopy by Gil-Santos [6]. In the nanowire's two-dimensional vibration, beat phenomenon is a special physical characteristic, which is reported and investigated by Zhan [19, 20]. Unfortunately, little attention has been paid to nonlinear vibration behaviors of nanowires, especially the Fe nanowires. Additionally, initial actuation and nanowires' sizes are rarely considered in the study on nonlinear vibration behaviors of nanowires. Therefore, the mechanism behind nonlinear vibration behaviors is still not clear, which is quite important to study the properties of nanowire vibrations.

Therefore, a fundamental numerical and theoretical study of nonlinear vibration behaviors of Fe nanowires is presented in this work. We investigate the nonlinear vibration behaviors of Fe nanowires in two aspects: initial actuation amplitude and nanowires' size. Based on the atomic structure, a valid discrete spring-mass model has been proposed to study the vibrations of the nanowire. With the results of MD simulations and the spring-mass model, the vibration frequency of nanowires is found to increase slightly and nonlinearly with the rise of initial actuation amplitude. On the other hand, based on the MD simulations performed on a number of Fe nanowires with different sizes, it is found that the ratio between the length and the height of nanowires has a significant influence on vibration behaviors. We introduce the slender beam and thick beam depend on this ratio *L/h*. The results from MD simulations signify that the slender beams' vibration frequencies have a linear correlation with h/L^2. In contrast, thick beams have a nonlinear correlation between vibration frequency and h/L^2.

Models and Methods

Atomistic Simulation.

In this work, double clamped Fe nanowires are investigated for their vibration properties. We use the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [27] to perform these simulations. Fig. 1 shows the model of nanowires used in MD simulations, which is created with iron atoms in positions corresponding to a perfect BCC crystal lattice. The shape of cross-section is chosen to be rectangle. The Fe lattice constant *a* is chosen as 0.287 nm. For the Fe nanowires, the length *L* ranges from 20*a* to 100*a*, while the width *b* and height *h* ranges from 4*a* to 30*a*. Atoms at both ends of the nanowires are fixed in their positions, which are denoted by "A" in Fig. 1. The rest of the nanowires are free to move. We assume that the vibrations of these Fe nanowires occur in a vacuum. Thus no damping is applied in these simulations. The periodic boundary conditions are not imposed in any directions.

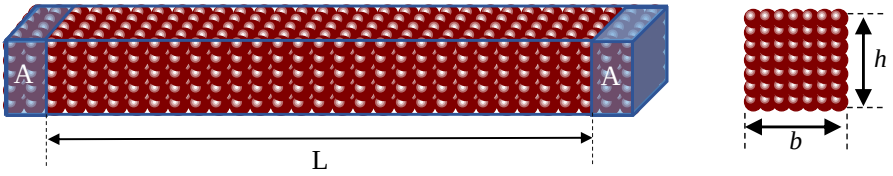


Figure 1. A schematic visualization of a fixed-fixed Fe nanowire used in simulation. The atoms in the areas 'A' are fixed in all directions and the residual atoms can move freely.

The embedded-atom-method (EAM) potential developed by Mendelev et al [28] is utilized to describe the atomic interactions between Fe atoms in these simulations, which is a semi-empirical function fitted to a group of parameters, including elastic constants, equilibrium lattice constant, cohesive energy, unrelaxed vacancy formation energy and others. In this model of atomic interaction, the total energy E_{tot} of a system of N atoms is a sum of two terms, which are classical pair potential and many-body embedding energy [29].

$$E_{tot} = \sum_{i=1}^{N} F_i(\rho_i) + \sum_{i=1}^{N}\sum_{j=1}^{N} \phi_{ij}(R_{ij}) \tag{1}$$

Here, F, ρ, ϕ are the embedded energy, pair potential and electron cloud density. R_{ij} is the distance between atom *i* and *j*.

At the beginning of each simulation, the nanowires are relaxed to the initial equilibrium configuration using the conjugate [30]gradient energy minimization. Then, the Nose-Hoover thermostat [30, 31] is employed to equilibrate the nanowires at 0.2 K. Finally, an initial velocity excitation $\mathbf{v}(z)$ is imposed on the nanowires along the *z*-axis.

$$\mathbf{v}(z) = \lambda \sin(\frac{\pi z}{L}) \tag{2}$$

Here, λ is actuation amplitude and L is the effective length of the nanowires that exclude

the two fixed edges.

Theoretical Models.

As mentioned before, nanowires are usually utilized as resonating beams in NEMS. According to the classical Euler-Bernoulli beam theory, the governing partial differential equation for the beam is

$$EI\frac{\partial^4 w}{\partial x^2}+\rho A\frac{\partial^2 w}{\partial t^2}=0 \tag{3}$$

where E is Young's modulus, I is moment of inertia, ρ is density and A is the cross-sectional area of the nanowire. $w(x,t)$ is the nanowire's transverse displacement. By using the clamped-clamped boundary conditions, the vibration frequency can be determined by

$$\omega=\pi^2\sqrt{\frac{EI}{\rho AL^4}} \tag{4}$$

where L is the length of the nanowire. With $I=bh^3/12$ and $A=bh$ (b and h are the width and the height of the nanowire's cross-section), the vibration frequency equals

$$\omega=\pi^2\sqrt{\frac{Eh^2}{12\rho L^4}}=\sqrt{\frac{E}{12\rho}}\times\frac{\pi^2 h}{L^2} \tag{5}$$

For the theoretical calculation of the vibration frequency for Fe nanowire, density ρ equals 7.87 g/cm^3.

To explore the relation of vibrations in different directions from the view of lattice structure, we develop a simplified discrete spring-mass model. According to the atomic arrangement of the Fe nanowire, each internal lattice is surrounded by other six lattices. However, the two lattices along the *z*-axis have little effect on the displacements of atoms in the intermediate lattice because of the fixed end in the *z*-axis. Thus, we neglect the lattices along the *z*-axis. Besides, lattices along the *x*-axis and *y*-axis have the same influence on the internal lattice. Fig. 2a shows the structure.

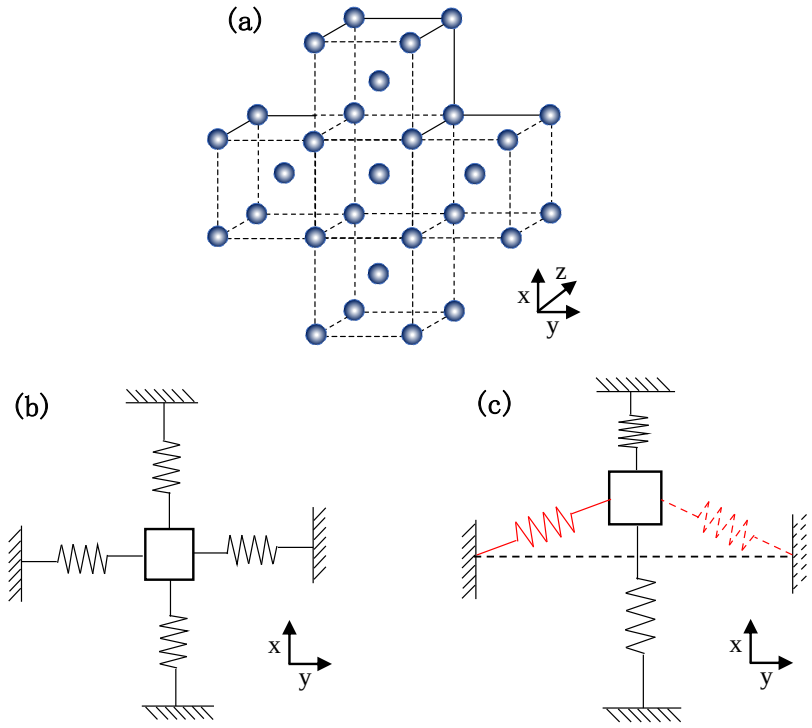


Figure 2. (a) The typical atomic arrangement of Fe nanowires with several lattices. (b)

A schematic of simplified discrete spring-mass model according to the atomic arrangement of Fe nanowires. (c) A schematic of this spring-mass model when the mass move along the *x*-axis. The red springs result in geometric nonlinearity of this model.

For simplicity, we interpret this structure as a spring-mass system as illustrated in Fig. 2b. This spring-mass system consists of linear springs, which are defined as

$$m\ddot{\mathbf{x}} = -k\Delta\mathbf{x} \tag{6}$$

where m is mass of the object, k is elastic constant of the spring. Due to the symmetric cross-sectional plane and the same influence along the *x*-axis and *y*-axis, the springs are chosen to be the same. The dynamic equations of this spring-mass system are described as

$$\begin{cases} m\ddot{y} = -4ky - kl_0\left(\dfrac{l_0 - y}{\sqrt{x^2 + (y - l_0)^2}} + \dfrac{-l_0 - y}{\sqrt{x^2 + (y + l_0)^2}} + \dfrac{-y}{\sqrt{(x + l_0)^2 + y^2}} + \dfrac{-y}{\sqrt{(x - l_0)^2 + y^2}}\right) \\ m\ddot{x} = -4kx - kl_0\left(\dfrac{-x}{\sqrt{x^2 + (y - l_0)^2}} + \dfrac{-x}{\sqrt{x^2 + (y + l_0)^2}} + \dfrac{-l_0 - x}{\sqrt{(x + l_0)^2 + y^2}} + \dfrac{l_0 - x}{\sqrt{(x - l_0)^2 + y^2}}\right) \end{cases} \tag{7}$$

where m is mass of the object, k and l_0 are elastic constant and initial length of the spring.

In this work, we focus on the one-dimensional vibration. Thus, we can let *y*=0 for all the time. Then, we can get the simplified dynamic equation as

$$m\ddot{x} = -4kx + 2kx\frac{l_0}{\sqrt{x^2 + l_0^2}} \tag{8}$$

In Eq. 8, it is obvious that the second item on the right is a nonlinear item for this equation. As shown in Fig. 2c, when this system vibrate along the *x*-axis, because of the spring-mass model's structure, the red springs produce the nonlinear forces, which correspond to the nonlinear item in Eq. 8. This is what we call the geometric nonlinearity.

Results and Discussion

Nanowires in Vibration

We begin this paper with MD simulations and corresponding analysis on the vibrational Fe nanowires. External energy is defined as the difference of the potential energy before and after the transverse velocity actuation is applied to the nanowire [19]. We can obtain the energy data directly from the simulations. By analyzing the external energy time history with the Fast Fourier transform (FFT), the vibration frequency and modes of the nanowire can be obtained. In this section, MD simulations are carried out on the Fe nanowire with the size of $20a \times 10a \times 100a$. Fig. 3a depicts the time history of the external energy during the free vibration after the initial velocity actuation. Fig. 3b presents part of the periodogram regarding the power of the discrete Fourier transformation versus frequency. The whole simulation last about 4000 picoseconds at a time step of 1 femtoseconds. The amplitude frequency curve is depend on these data. As is seen in Fig. 3b, one main frequency component is identified, which is about 53.41 GHz. This result indicates that the Fe nanowire is under a vibration with the frequency of 26.70 GHz, because the frequency of the external energy is twice that of the actual vibration.

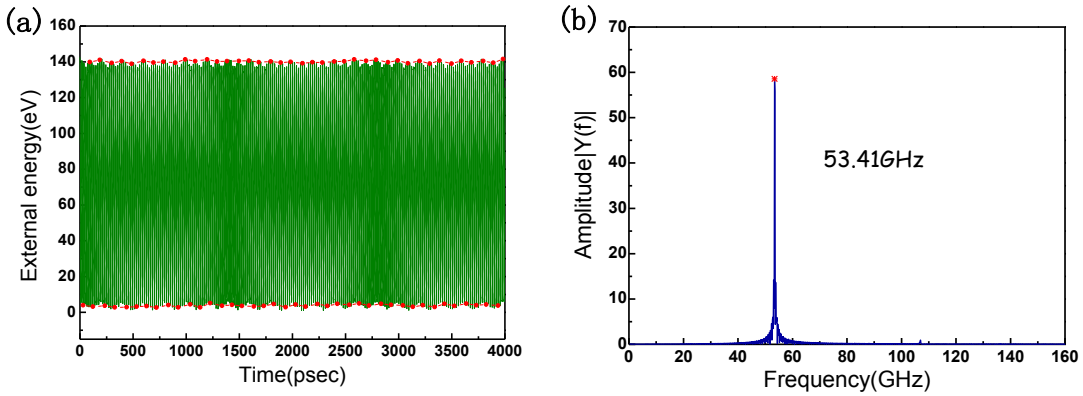


Figure 3. (a) External energy time history of the Fe nanowire with the size of $20a\times10a\times100a$ for a free vibration MD simulation at 0.2K. The simulation time is from 0 to 4000 picosecond. Circle markers highlight the maximum and minimum of external energy during each vibration circle. (b) The frequency spectrum from FFT analysis on the previous simulation (from 0 to 160 GHz). The star marker highlights the maximum of the main frequency component.

Initial Actuation Amplitude

Then, the relation between initial actuation amplitudes and vibration frequencies is studied. We perform MD simulations on the Fe nanowires with the size of $4a\times4a\times20a$, $10a\times10a\times50a$, $5a\times5a\times50a$ and $10a\times10a\times100a$. Initial velocity excitations as Eq. 2 are imposed on these nanowires. In Eq. 2, several different amplitudes λ ranging from 0.2 Angstrom/psec to 1.4 Angstrom/psec are chosen for tests. Figs. 4 describe the frequencies received from the free vibration MD simulations of Fe nanowires under different initial actuation amplitudes. As observed in Figs. 4, it is found that the vibration frequency from the MD simulations rises slightly and nonlinearly with the increase of initial actuation amplitude.

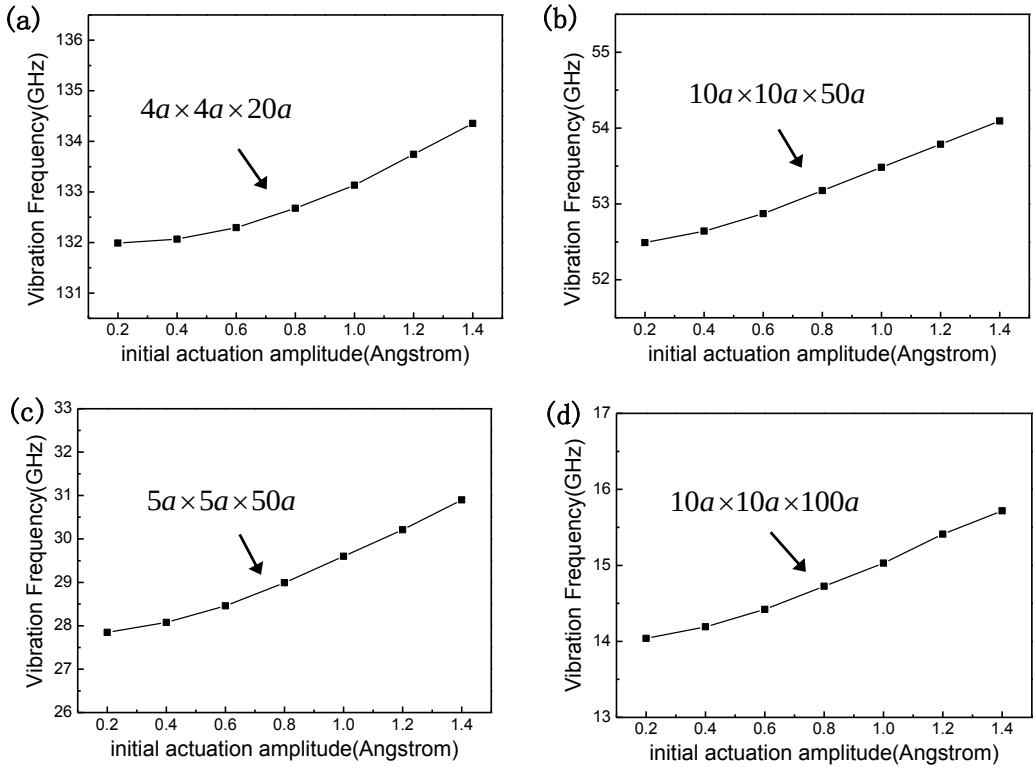


Figure 4. These lines represent the vibration frequencies vs initial actuation amplitude.

In this situation, we use the frequencies obtained from FFT analysis when we apply different initial actuation amplitude on different Fe nanowires. They are nanowires with (a) the size of $4a\times4a\times20a$. (b) the size of $10a\times10a\times50a$. (c) the size of $5a\times5a\times50a$. (d) the size of $10a\times10a\times100a$.

In addition, we test initial boundary conditions of different actuation amplitudes on the spring-mass model. The results in Fig. 5 show that with the increase of initial actuation amplitude, the vibration frequency rises slowly, which is similar to the results from the MD simulations.

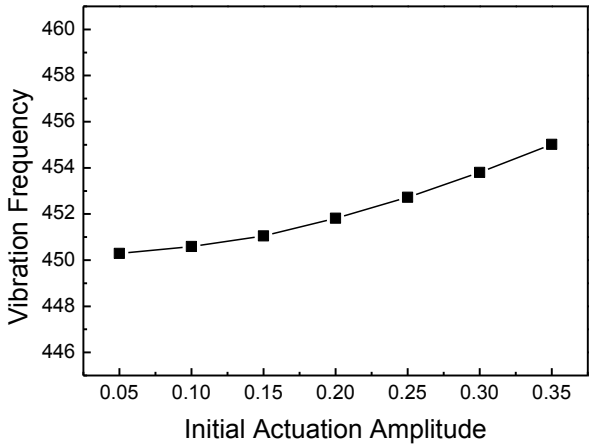


Figure 5. The vibration frequencies vs initial actuation amplitude. In this situation, we apply different initial conditions' value on the spring-mass model. Then, the frequencies can be obtained from FFT analysis on the numerical results.

According to the classical theory (Eq. 5), the Euler-Bernoulli beams have fixed natural frequencies, which is irrelevant to initial actuation amplitude. However, the results from MD simulations signify that the vibration frequency has a nonlinear correlation with initial actuation amplitude. We suggest that, in the nano scale, the lattice structure has a big influence on the properties of the nanowires' vibrations. As mentioned above, according to the atomic arrangement, the spring-mass model has geometric nonlinearity, which may result in the nonlinear correlation between vibration frequency and initial actuation amplitude. However, the classical Euler-Bernoulli model is a macro-continuity model, which may be not suitable for the nanowire in this scale. Therefore, in this case, it is found that the vibration frequency increase slightly with the growth of initial actuation amplitude.

Size Effect

Further interest is laid on the relation between nanowires size and vibration frequencies. During the simulations in this section, the initial actuation amplitude λ is fixed at 1.5 Angstrom/psec. Figs. 6a and 6b shows the vibration frequency height curve obtained from the simulations with the fixed *L* and *b*. As observed in Figs. 6a and 6b, it is obvious that the vibration frequency rises with the increase of height. According to the classical Euler-Bernoulli beam theory (Eq. 5), the vibration frequency has a linear correlation with the height of the nanowire. In the results from MD simulations, we can also find that when the value of height is relatively small, the correlation between vibration frequency and height is linear. However, the vibration frequency rises nonlinearly with the increase of height when the value of height becomes relatively large. Since *L* is a fixed value, the ratio between length and height (*L/h*) declines with the increase of height. We define the slender beams as the nanowires with relatively large value of *L/h* (usually larger than 5), while the thick beams are defined as the nanowires with relatively large value of *L/h* (usually smaller than 5). Thus, the observations signify that, in the nano size, the correlation between vibration frequency and height is linear for the slender beams, while the correlation is nonlinear for the thick beams.

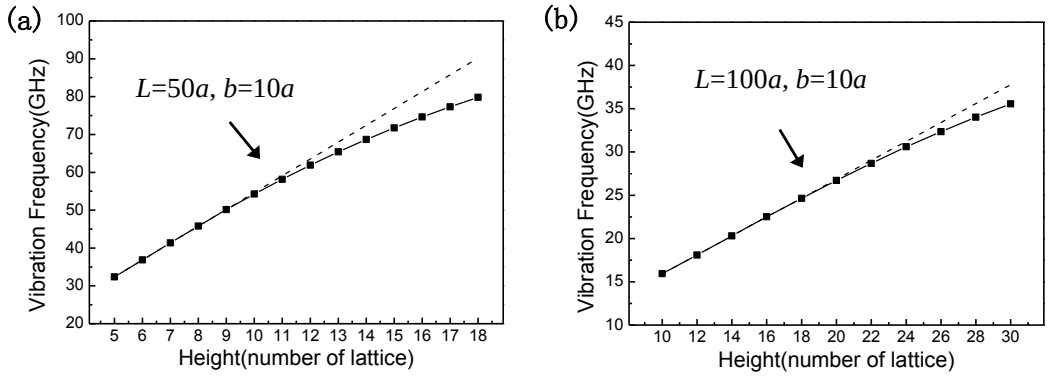


Figure 6. The solid lines in these figures represent the vibration frequencies in different heights. The dotted lines are the fitted straight lines based on vibration frequency values in the first five heights. In this situation, we use the frequencies obtained from FFT analysis on the MD simulations carried on different Fe nanowires. They are nanowires with the fixed *L* and *b* ((a) *L*=50*a*, *b*=10*a*.(b) *L*=100*a*, *b*=10*a*.).

We also perform MD simulations on the nanowires with the fixed *h* and *b*. Fig. 7a depicts vibration frequency length curve, which received from several MD simulations on Fe nanowires in different lengths. From Fig. 7a, it is found that, in the fixed *h* and *b*, the vibration frequency declines with the increase of the nanowires' length. According to the classical beam theory (Eq. 5), there is a linear correlation between vibration frequency and $1/L^2$. Thus, Fig. 7b describes the vibration frequency with $1/L^2$. It is found that the left half curve is almost linear, while the right half curve shows the nonlinear trend. Since *h* is fixed at 50a, it is suggested that the correlation between vibration frequency and $1/L^2$ is linear for the slender beams, while the correlation is nonlinear for the thick beams.

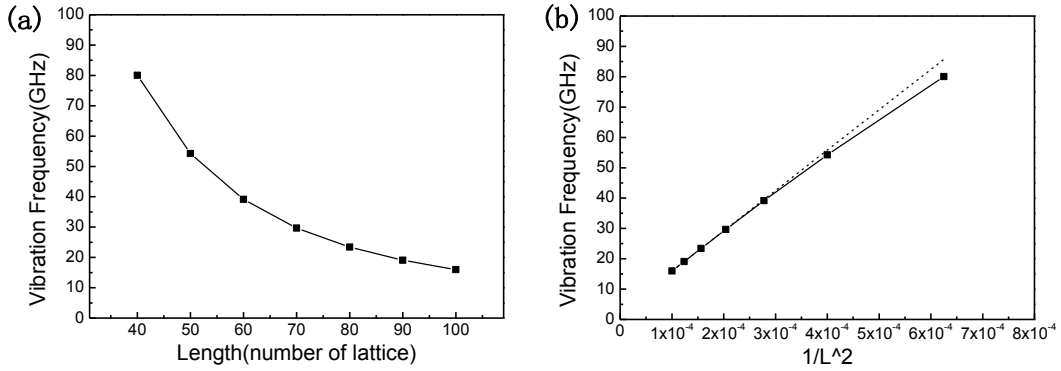


Figure 7. (a) The vibration frequencies vs nanowires' length. In this situation, we use the frequencies obtained from FFT analysis on the MD simulations carried on different Fe nanowires. They are nanowires with the fixed *h*=10*a* and *b*=10*a*. (b) The solid line shows the vibration frequencies vs $1/L^2$, while the dotted lines are the fitted straight lines based on vibration frequency values in the first three values of $1/L^2$.

Conclusion

In summary, based on large scale molecular dynamics simulations, we investigated nonlinear behaviors in the vibration of Fe nanowires. It is found that with the increase of initial actuation amplitude, the vibration frequency of nanowires grows slowly and nonlinearly. A discrete spring-mass model abstracted from the atomic arrangement is developed in this work. The geometric nonlinearity of this model is applied to explore this nonlinear correlation between initial actuation amplitude and vibration frequency, which cannot be explained by the classical Euler-Bernoulli model. Furthermore, simulations on Fe nanowires with different sizes are

performed to show the vibration properties. Fe nanowires with the change of L and h are considered respectively. Both the results shows that the vibration frequency has a linear correlation with h/L^2 for the slender beams, while the correlation becomes nonlinear for thick beams.

Acknowledgements

This work was supported in part by The National Natural Science Foundation of China (NSFC) under grant number 11072086, 11372117, J1310022, by the 985 program, by the 211 project of Jilin University, and by Basic funds for science and research in Jilin University. This work was also supported by the Jilin Province Computing Centre through the use of its high-performance computing facilities.

References

[1] D. Rugar, R. Budakian, H.J. Mamin, B.W. Chui, Single spin detection by magnetic resonance force microscopy, Nature, 430 (2004) 329-332.
[2] T. Stowe, K. Yasumura, T. Kenny, D. Botkin, K. Wago, D. Rugar, Attonewton force detection using ultrathin silicon cantilevers, Applied Physics Letters, 71 (1997) 288-290.
[3] P. Xie, Q. Xiong, Y. Fang, Q. Qing, C.M. Lieber, Local electrical potential detection of DNA by nanowire-nanopore sensors, Nature nanotechnology, 7 (2012) 119-125.
[4] R.-M. Ma, L. Dai, H.-B. Huo, W.-J. Xu, G. Qin, High-performance logic circuits constructed on single CdS nanowires, Nano letters, 7 (2007) 3300-3304.
[5] C.K. Chan, H.L. Peng, G. Liu, K. McIlwrath, X.F. Zhang, R.A. Huggins, Y. Cui, High-performance lithium battery anodes using silicon nanowires, Nature Nanotechnology, 3 (2008) 31-35.
[6] E. Gil-Santos, D. Ramos, J. Martínez, M. Fernández-Regúlez, R. García, Á. San Paulo, M. Calleja, J. Tamayo, Nanomechanical mass sensing and stiffness spectrometry based on two-dimensional vibrations of resonant nanowires, Nature nanotechnology, 5 (2010) 641-645.
[7] M. Li, T.S. Mayer, J.A. Sioss, C.D. Keating, R.B. Bhiladvala, Template-grown metal nanowires as resonators: Performance and characterization of dissipative and elastic properties, Nano Letters, 7 (2007) 3281-3284.
[8] X.-Q. Zeng, Y.-L. Wang, H. Deng, M.L. Latimer, Z.-L. Xiao, J. Pearson, T. Xu, H.-H. Wang, U. Welp, G.W. Crabtree, Networks of ultrasmall Pd/Cr nanowires as high performance hydrogen sensors, Acs Nano, 5 (2011) 7443-7452.
[9] S.Y. Kim, H.S. Park, Utilizing mechanical strain to mitigate the intrinsic loss mechanisms in oscillating metal nanowires, Physical review letters, 101 (2008) 215502.
[10] X. Feng, R. He, P. Yang, M. Roukes, Very high frequency silicon nanowire electromechanical resonators, Nano Letters, 7 (2007) 1953-1959.
[11] H.S. Park, W. Cai, H.D. Espinosa, H. Huang, Mechanics of crystalline nanowires, MRS bulletin, 34 (2009) 178-183.
[12] S.S. Verbridge, J.M. Parpia, R.B. Reichenbach, L.M. Bellan, H. Craighead, High quality factor resonance at room temperature with nanostrings under high tensile stress, J Appl Phys, 99 (2006) 124304.
[13] Y. Yang, C. Callegari, X. Feng, M. Roukes, Surface adsorbate fluctuations and noise in nanoelectromechanical systems, Nano letters, 11 (2011) 1753-1759.
[14] W.G. Conley, A. Raman, C.M. Krousgrill, S. Mohammadi, Nonlinear and nonplanar dynamics of suspended nanotube and nanowire resonators, Nano letters, 8 (2008) 1590-1595.
[15] J. He, C.M. Lilley, Surface stress effect on bending resonance of nanowires with different boundary conditions, Applied physics letters, 93 (2008) 263108.
[16] H.S. Park, Surface stress effects on the resonant properties of silicon nanowires, J Appl Phys, 103 (2008) 123504.
[17] G.F. Wang, X.Q. Feng, Timoshenko beam model for buckling and vibration of nanowires with surface effects, J Phys D Appl Phys, 42 (2009).
[18] Z. Yan, L.Y. Jiang, The vibrational and buckling behaviors of piezoelectric nanobeams with surface effects, Nanotechnology, 22 (2011).
[19] H.F. Zhan, Y.T. Gu, A fundamental numerical and theoretical study for the vibrational properties of nanowires, J Appl Phys, 111 (2012).
[20] H.F. Zhan, Y.T. Gu, H.S. Park, Beat phenomena in metal nanowires, and their implications for resonance-based elastic property measurements, Nanoscale, 4 (2012) 6779-6785.
[21] J. Zhang, C. Wang, Beat vibration of hybrid boron nitride-carbon nanotubes–A new avenue to atomic-scale mass sensing, Computational Materials Science, 127 (2017) 270-276.
[22] P.A.T. Olsson, Transverse resonant properties of strained gold nanowires, J Appl Phys, 108 (2010).
[23] S.H. Park, J.S. Kim, J.H. Park, J.S. Lee, Y.K. Choi, O.M. Kwon, Molecular dynamics study on size-dependent

elastic properties of silicon nanocantilevers, Thin Solid Films, 492 (2005) 285-289.
[24] L.F. Wang, H.Y. Hu, Size Effects on Effective Young's Modulus of Nano Crystal Copper Wires, Int J Comp Meth-Sing, 2 (2005) 315-326.
[25] H. Yu, W.W. Zhang, S.Y. Lei, L.B. Lu, C. Sun, Q.A. Huang, Study on Vibration Behavior of Doubly Clamped Silicon Nanowires by Molecular Dynamics, J Nanomater, (2012).
[26] Y.G. Zheng, Y.T. Zhao, H.F. Ye, H.W. Zhang, Size-dependent elastic moduli and vibrational properties of fivefold twinned copper nanowires, Nanotechnology, 25 (2014).
[27] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, Journal of computational physics, 117 (1995) 1-19.
[28] M.I. Mendelev, S. Han, D.J. Srolovitz, G.J. Ackland, D.Y. Sun, M. Asta, Development of new interatomic potentials appropriate for crystalline and liquid iron, Philos Mag, 83 (2003) 3977-3994.
[29] M.T. McDowell, A.M. Leach, K. Gall, Bending and tensile deformation of metallic nanowires, Model Simul Mater Sc, 16 (2008).
[30] W.G. Hoover, Canonical dynamics: equilibrium phase-space distributions, Physical review A, 31 (1985) 1695.
[31] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, The Journal of chemical physics, 81 (1984) 511-519.

Modelling flow-diverting stent as porous medium with different permeabilities in the treatment of intracranial aneurysms: a comparison of a successfully treated case and an unsuccessful one

†*Y. Li[1,2], M. Zhang[1,2], D. I. Verrelli[2], W. Yang[3], W. Chong[4], M. Ohta[5] and Y. Qian[2]

[1] Graduate School of Engineering, Tohoku University, Japan
[2] Faculty of Medicine and Health Sciences, Macquarie University, Australia
[3] Mineral Resources, CSIRO, Australia
[4] Neuroradiology Department, Monash Medical Centre, Australia
[5] Institute of Fluid Science, Tohoku University, Japan

* Presenting author: yujie.li@biofluid.ifs.tohoku.ac.jp
† Corresponding author: yujie.li@biofluid.ifs.tohoku.ac.jp

Abstract

Using a porous medium as a computational model for the real flow-diverting (FD) stent in computational fluid dynamics (CFD) improves the simulation efficiency. Adjustment of permeability level of the applied porous medium can result in various flow-resistance effects, which impact on the flow dynamics around and inside the aneurysm dome. Moreover, diversity in patient-specific aneurysm geometry also contributes to the difference in both the resistance force induced by the FD device and the aneurysmal haemodynamics. However, few studies have discussed the relationship between the setting of permeability and the intra-aneurysmal haemodynamics with different aneurysms.

In this study, in order to distinguish FD stents with different porosity, we simulated the porous medium stent with a range of permeabilities, respectively in a successfully treated aneurysm and an unsuccessfully treated one observed clinically. Haemodynamic parameters of intra-aneurysm mass flow rate (MFR) and energy loss (EL) were calculated to investigate their response to the alteration of permeability, as well as to the aneurysm morphology.

In comparison between the two patient-specific aneurysms, we found marked changes (70 and 40 % in MFR, 40 and 35 % in EL, respectively with successful and unsuccessful cases) in the aneurysmal haemodynamics as the porosity level of the implanted FD stents was increased by a factor of 25. The simulation results showed considerable differences in the relative flow-diversion between the clinically observed successful and unsuccessful case (up to 30 % in MFR and 45 % in EL). This study will help to provide future FD modellers with information about suitable selection of permeability level for different aneurysm cases.

Keywords: Permeability; Porous Medium; Computational Fluid Dynamics; Flow Diverting Stent; Cerebral Aneurysm

Introduction

Intracranial aneurysm (IA) is a vascular disease, observed by digital subtraction angiography (DSA) as a bulge that dilates out of the cerebrovascular wall. Untreated IAs may rupture, leading to subarachnoid haemorrhage, a severe condition that threatens the patient's life [1-3]. As an endovascular therapy, the flow-diverting (FD) stent is commonly used by neuroradiologists to keep most of the blood flow within the parent artery, thereby inducing blood clotting inside an aneurysm [4-5].

The flow-diversion efficacy of an implanted FD device is thought to be closely associated with the treatment outcome. To quantitatively analyse the flow-diversion efficacy, computational fluid dynamics (CFD) is frequently used to resolve the aneurysmal haemodynamics, as well as to quantify the haemodynamic alterations caused by different FD devices. Recently, an emerging technique that models an FD stent as a porous medium markedly decreased the computational time in CFD simulations [6]. By adjusting the permeability of the porous medium on the basis of the diversity in FD stent structures and porosity, the flow resistance induced by the porous medium stent model can be varied to represent different FD stents [6-8].

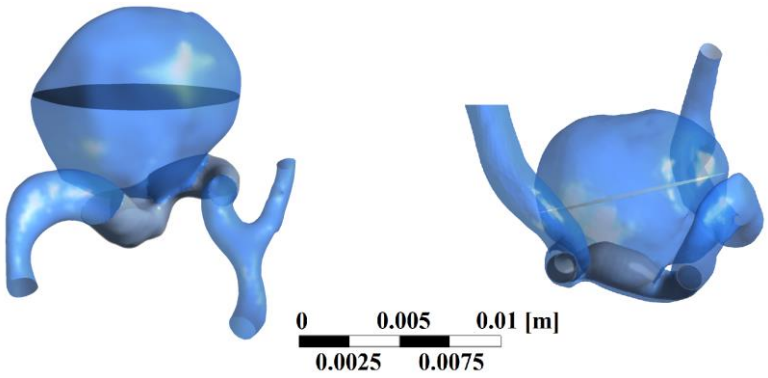
A problematic issue in FD treatment is that the outcomes may vary with respect to patients, *i.e.* a number of patients may still suffer from incomplete aneurysm occlusion even after two or three FDs were implanted. Previous studies suggest that the porosity of an FD stent can greatly affect the flow-diversion efficacy, thereby determining if an aneurysm can be completely occluded. However, the effects of device porosity on flow-diversion efficacy have not yet been quantitatively studied.

In this study, we aim to compare the flow-diversion efficacy of FD stents with different porosities, using permeability as a surrogate for porosity, and investigate the treatment outcomes of such devices in different patients, by contrasting the haemodynamic alterations in a successfully treated aneurysm and an unsuccessfully treated one.

Methods

Patient-specific Aneurysms

Two patient-specific intracranial aneurysm model geometries were studied, after we obtained the institutional ethics approvals. One of the aneurysms ('successful') was confirmed fully occluded 6 months after treatment, whereas the DSA of the other ('unsuccessful') revealed the existence of a residual aneurysm. A centre plane across the aneurysm lumen was selected to demonstrate and summarise the haemodynamic differences after FD treatments. The aneurysm geometries and the positions of the two planes are as shown in Figure 1.



(a) **(b)**

Figure 1. Geometries of two patient-specific aneurysms and the positions of the centre planes: (a) the successful aneurysm and (b) the unsuccessful aneurysm.

Flow-diverting Stent Modelling

The FD stent geometry was modelled as a fitted tube, with varied diameters, running through the parent artery that covers the aneurysm neck (see Fig. 1). Instead of the conventional stent geometry with individual FD wires, the FD stent was defined as a homogeneous porous medium tube in the simulation, with constant thickness of 100 μm. By defining the permeability, the

resistance force induced by the porous medium can be adjusted, which is expressed as a momentum source term added to the standard fluid flow equation:

$$S_i = -\left(\frac{\mu}{\alpha} v_i + C_2 \frac{1}{2} \rho |v| v_i\right), \quad (1)$$

where α is the permeability, C_2 is the quadratic loss coefficient, i represents the x, y, or z coordinate, v is velocity, μ is viscosity, and ρ is density. The equation describing the correlation between the momentum source term and pressure drop is simplified as

$$\Delta p = -S_i \Delta e, \quad (2)$$

where Δe is the thickness of the porous medium [6].

Initial values of both permeability and inertial resistance factor were obtained based on a test model in numerical simulation that has a similar geometry to the *Silk* stent [6]. The permeability of an FD device is directly associated with the FD stent porosity. We adjusted the permeability by a series of factors — from 20% to 500% of the initial value (referred to as porosity level 20 to 500), to measure the effects of device porosity on the post-stenting haemodynamic alterations.

CFD simulation

The blood was assumed to be a Newtonian, incompressible fluid undergoing steady laminar flow. To match the properties of physiological blood flow, density and viscosity were specified as 1050 kg/m^3 and 0.0035 Pa·s, respectively. Mass flow rates of 250 mL/min at the inlet boundary were studied, while a 0 Pa static pressure was set at the outlets for all cases. A mesh dependency test was carried out before the simulation to ensure mesh quality and numerical accuracy. The fluid zone was discretised into 0.7 to 2.8 million elements with tetrahedral mesh elements, while denser mesh was added to the artery wall and stent surface. We used a commercial finite-volume-method based solver (CFX, Ansys, U.S.A.) to perform the CFD simulation.

Results

Figure 2 demonstrates the intra-aneurysmal haemodynamic result for the untreated condition and treated cases with the porosity level 20, 100, and 500. Figure 3 presents the alterations of flow-diversion efficacy, measured as intra-aneurysmal mass flow rate (MFR) and the Energy Loss (EL), at different levels of porosity.

We found that the velocity magnitude on the centre plane increases, when the device porosity increases. This trend is observed in both successful and unsuccessful aneurysms (Fig. 2).

By comparing the streamlines before and after FD treatments, an inflow reduction can be clearly observed in both aneurysms. With the successful aneurysm, the increasing porosity levels result in a decreasing blocking effect of aneurysmal inflow, while in the unsuccessful aneurysm, the flow-diversion effects remain similar from porosity level 20 to 500 (Fig. 2).

Visualisation of velocity iso-surfaces in both cases shows that a strong inflow jet exists in the unsuccessful case even under the condition of the lowest porosity level, whereas the inflow jet can be markedly reduced in the successful case when a FD stent is implanted (Fig. 2).

Quantitative results also show distinctions in the response to different porosity levels between the two aneurysms. For the successful case, the MFR drastically decreased from 80% to 16% when the porosity level was reduced from 500 to 20. For the unsuccessful case, however, the MFR only decreased to 40% when the porosity experienced the same reduction. Moreover, EL in the unsuccessful case reveals a 25–55% higher tendency than that of the successful case (Fig. 3).

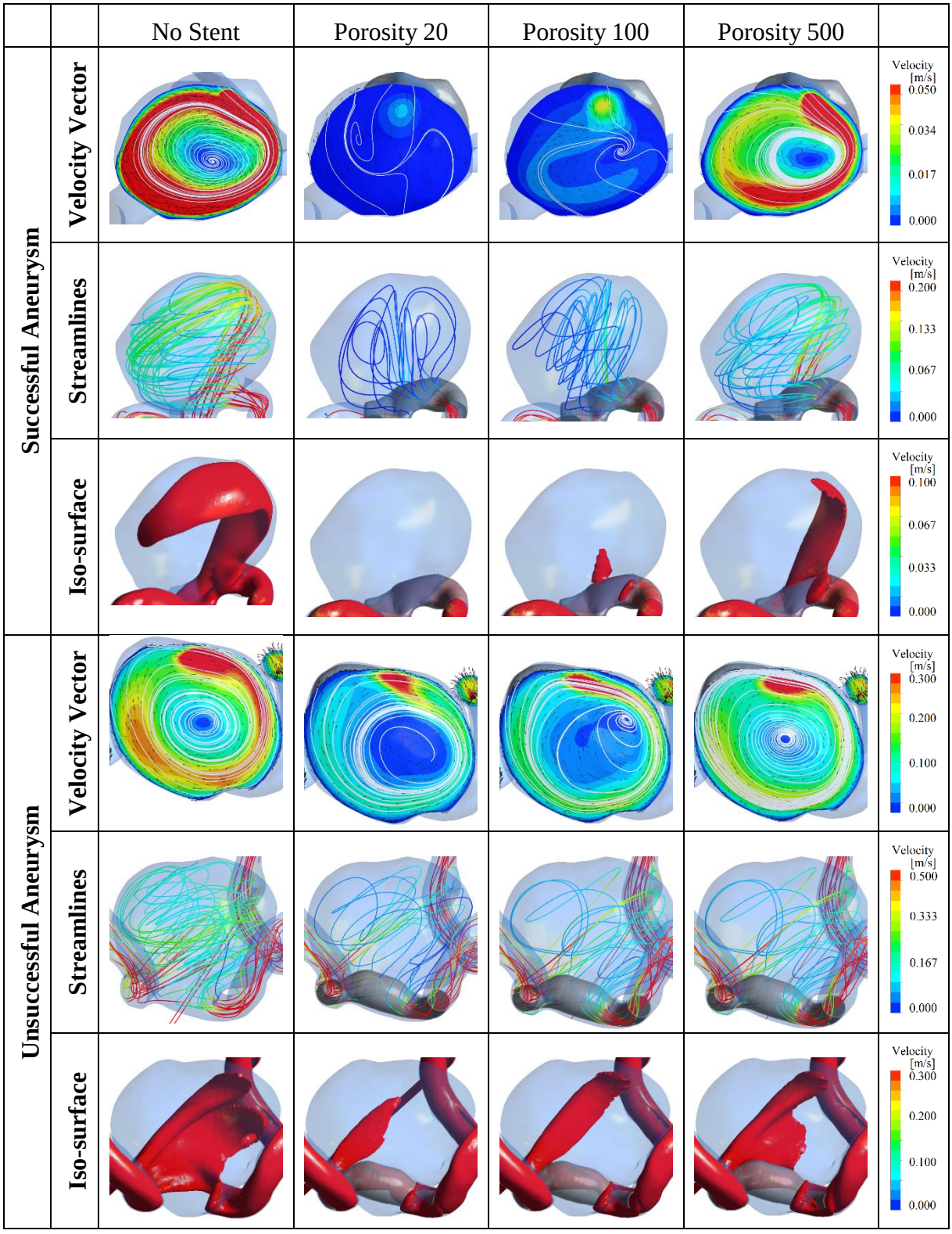
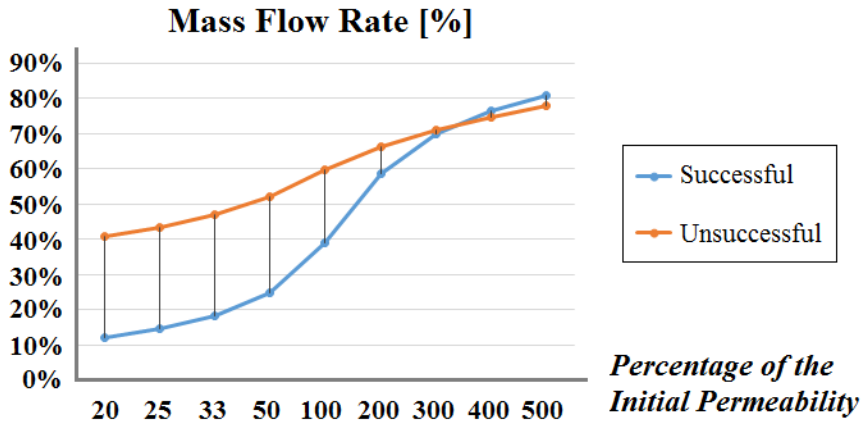
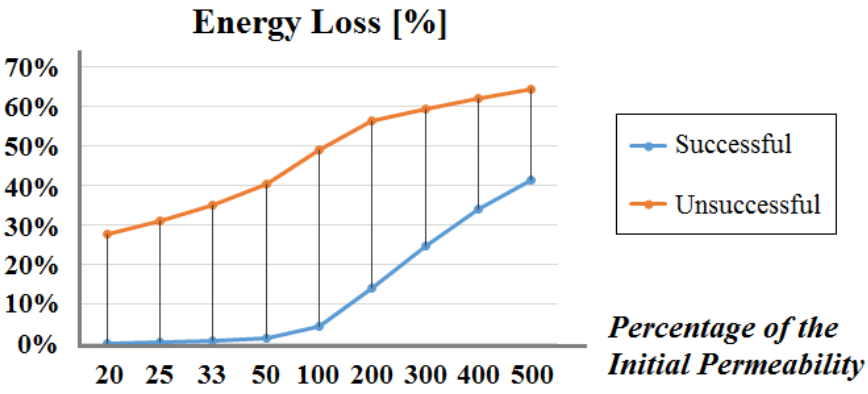


Figure 2. The velocity distributions on the centre plane, the streamlines and the velocity iso-surfaces for the successful and unsuccessful aneurysms, with a porous medium FD stent of various permeability levels, and the untreated cases.



(a)



(b)

Figure 3. Comparison of aneurysmal haemodynamics between successful and unsuccessful aneurysms treated with a porous medium FD stent at a range of permeability levels: (a) Mass flow rate and (b) Energy loss.
(*Haemodynamic results are given as percentages of the respective untreated cases.*)

Discussion and conclusions

Techniques, such as stent compaction or multi-stent implantation [9-10] that decrease the porosity of the FD wires across the aneurysm ostium, can effectively improve the flow-diversion efficiency post-treatment. However, the highest level of flow diversion that can be achieved by decreasing the permeability level may differ across different aneurysm geometries, e.g. the successful case (16%) and the unsuccessful one (40%) in this study.

Furthermore, as can be seen in the unsuccessful case, a gap exists between the stent layer and the vascular wall, through which a strong inflow jet enters the aneurysm lumen. The DSA taken in the follow-up of the unsuccessful case also revealed the existence of the strong inflow jet. This suggests that not only the porosity level of an FD stent, but also the selection of device size in combination with parent artery morphology should be taken into consideration by neuroradiologists in the design of an FD stent treatment.

To sum up, we investigated the effect of device permeability levels on a successfully treated aneurysm and an unsuccessful one with an implanted FD stent. We found that decreasing the permeability of an FD stent can improve the flow-diversion efficacy post-treatment, but how much improvement can be achieved depends on the morphological characteristics of the aneurysm and parent artery, as well as the selection of device size.

Acknowledgements

This study was supported by the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), the Macquarie University Research Excellence Scholarship (iMQRES, No. 2015256), and Australian Research Council Linkage Projects (Grant ID: LP130100423).

References

[1] J. Cebral, E. Ollikainen, B. J. Chung, F. Mut, V. Sippola, B. R. Jahromi, R. Tulamo, J. Hernesniemi, M. Niemela, A. Robertson and J. Frosen. Flow conditions in the intracranial aneurysm lumen are associated with inflammation and degenerative changes of the aneurysm wall, *Am. J. Neuroradiol.*, vol. 38, no. 1, pp. 119–126, Jan. 2017.

[2] S. E. Francis, J. Tu, Y. Qian and A. P. Avolio. A combination of genetic, molecular and haemodynamic risk factors contributes to the formation, enlargement and rupture of brain aneurysms, *J. Clin. Neurosci.*, vol. 20, no. 7, pp. 912–918, Jul. 2013.

[3] M. D. Ford, H. N. Nikolov, J. S. Milner, S. P. Lownie, E. M. DeMont, W. K. F. Loth, D. W. Holdsworth and D. A. Steinman. PIV-measured versus CFD-predicted flow dynamics in anatomically realistic cerebral aneurysm models, *J. Biomech. Eng.*, vol. 130, no. 2, p. 21015, 2008.

[4] F. Mut, M. Raschi, E. Scrivano, C. Bleise, J. Chudyk, R. Ceratto, P. Lylyk and J. R. Cebral. Association between hemodynamic conditions and occlusion times after flow diversion in cerebral aneurysms, *J. NeuroInterventional Surg.*, vol. 7, no. 4, pp. 286–290, Apr. 2015.

[5] S. Fischer, M. A. Perez, W. Kurre, G. Albes, H. Bäzner and H. Henkes. Pipeline Embolization Device for the treatment of intra- and extracranial fusiform and dissecting aneurysms: initial experience and long-term follow-up, *Neurosurgery*, vol. 75, no. 4, pp. 364–374, Oct. 2014.

[6] L. Augsburger, P. Reymond, D. A. Rufenacht and N. Stergiopulos, Intracranial stents being modeled as a porous medium: flow simulation in stented cerebral aneurysms, *Ann. Biomed. Eng.*, vol. 39, no. 2, pp. 850–863, Feb. 2011.

[7] M. Ohta, H. Anzai, Y. Miura and T. Nakayama, Parametric study of porous media as substitutes for flow-diverter stent, *Biomater. Biomed. Eng.*, vol. 2, no. 2, pp. 111–125, Jun. 2015.

[8] M. Kim, D. B. Taulbee, M. Tremmel and H. Meng, Comparison of two stents in modifying cerebral aneurysm hemodynamics, *Ann. Biomed. Eng.*, vol. 36, no. 5, pp. 726–741, May 2008.

[9] J. Xiang, D. Ma, K. V. Snyder, E. I. Levy, A. H. Siddiqui and H. Meng. Increasing flow diversion for cerebral aneurysm treatment using a single flow diverter, *Neurosurgery*, vol. 75, no. 3, pp. 286–294, Sep. 2014.

[10] D. Ma, J. Xiang, H. Choi, T. M. Dumont, S. K. Natarajan, A. H. Siddiqui and H. Meng. Enhanced aneurysmal flow diversion using a dynamic push-pull technique: an experimental and modeling study, *Am. J. Neuroradiol.*, vol. 35, no. 9, pp. 1779–1785, Sep. 2014.

A treatment planning of radiofrequency ablation for spinal tumor

Zhen Tian[1], Tong Dong[1], Yanyan Cheng[1], Jian Hu[1], * Qun Nan[1]

[1] College of Life Science and Bio-Engineering, Beijing University of Technology, China.

*Corresponding author: nanqun@bjut.edu.cn

Abstract

Background: To establish a treatment planning to predict therapeutic parameters accurately and avoid other healthy tissue damage by radiofrequency ablation (RFA) in spinal tumor.

Materials and Methods: According different therapeutic parameters (voltages 10-30V and heating-up time 120-1200s) and structures of electrode, the ablation results were achieved by finite element method to build a database. For the purpose of predicting therapeutic parameters in the treatment planning through the database, we chose three modes to compare the best one, including back propagation (BP) neural network, general regression neural network (GRNN) and support vector machine (SVM).

Results: Small tumor (maximum diameter<30.0mm) could be ablated by single needle electrode, and large tumor (maximum diameter≥30.0mm) should be applied by multi needle electrode. Besides, acquired 383 sets of data were regarded as database by the simulation. Compared with the other two models, SVM model is the best one to be applied in the treatment planning, in which error rates(0 and 0.33%) were the smallest and correlation coefficients(1.000 and 0.999) were the closest to 1. To verify whether the ablation results are correct, we have compared them to the clinical treatment results and previous studies.

Conclusions: This treatment planning by SVM is accurate to predict voltage and heating-up time, and it has higher correlation coefficient and few errors.

Keywords: Spinal tumor; radiofrequency ablation (RFA); SVM (support vector machine); treatment planning

Abbreviations: RFA, radiofrequency ablation; BP, back propagation; GRNN, general regression neural network; SVM, support vector machine; MWA, microwave ablation

Introduction

Spinal tumor is an uncommon disease in the primary tumors, while, it has been a kind of serious metastatic spinal tumors in clinic. It has been divided into three types, including primary benign, primary malignant and metastasis spinal tumors [1]. As reported in USA, there were 11712 patients suffered from spinal tumors including 2576 cases of malignant primary spinal tumors and 9136 cases of primary benign spinal tumors in 2004-2007 [2]. So applying a highly efficient

method to treat spinal tumors is desiderated in clinic. Despite surgical resection is still the main effective method to treat various tumors [3], thermal ablation is under development for spinal tumors currently [4]-[5], on the basis of its clinical utility and safety for treating liver, lung and bone tumors [6]-[8]. Raising tissue temperature in the target volume to induce coagulative necrosis is the goal of thermal ablation, which contains RFA, MWA and so on. The energy of radiofrequency is less than microwave energy, which could avoid damage of nervous tissue availably by small tissue of coagulation necrosis [9]. So the method of RFA is widely used in treating spinal tumors. In addition, the complexity of the spine anatomical structures in the surgical treatment of spinal tumors is difficult, in recent years, some scholars explore imaging guided RFA technology was applied to part of the treatment of spinal tumor, has obtained the good effect [10]-[12], for the minimally invasive treatment of spinal tumor provides a new way. In some studies [13]-[14], osteoid osteoma whose diameters are generally between 2 and 14mm has been treated by RFA, and it has verify that applying RFA is both feasible and safe for the treatment of painful primary benign tumors. According to the previous studies [4][9][15][16], using RFA to treat metastasis spinal tumors has obtained satisfying effect in clinical. Hillen et al [15] applied CT to achieve the aim of monitor and navigation to guide the RF electrode and had better effect (the maximum ablation area about 30mm long and width 20mm) in clinical based on 26 patients. And it's also illustrated the feasibility and safety by RFA in spinal tumor. Furthermore, Wallace et al [16] researched 72 patients of spinal metastases by RFA from 2012 to 2014, and explained the availability of this method at the average 8min32s ±4min49s heating-up time. However, due to the complex environment of spinal tumors which could be located close to the spinal cord or neural, studies have demonstrated that the aim is to both get large target volumes and prevent the risk of neural or spinal cord damage [9][17]. There are still many challenges in clinical hasn't been solved yet, such as the other tissues get thermal injured caused by different heterogeneity, how to real-time monitoring melting process, avoid the interference of the vascular structures and neural structure as well as electrode of the RFA probe insertion ways. So some researchers [8][13][14][17][18][19] have developed the planning system to assist surgeon or use ultrasound assistance in the treatment of spinal and other tumors, based on the imaging technology, such as CT, MR or X ray. Among that, studying possibility of subject specific precise offline microwave thermal ablation treatment planning is based on the CT and MR [18]. However, nerve injuries have been reported more frequently in series of heat-based ablation of spinal tumors in close proximity to the spinal cord or sacral nerve roots [20]-[25]. Many neural structures are not well depicted with the imaging modalities most commonly used to guide thermal ablation, namely ultrasound and CT. Pre-procedural imaging with MRI is frequently valuable to evaluate the proximity of neural structures to the target tumor. Likewise, CT angiography may be helpful to depict vulnerable vascular structures supplying the central neuro axis, namely the radicular and vertebral arteries. Alternatively, tumor ablation may be performed under MRI guidance in some medical centers to improve visualization of these neural structures intraprocedurally. Immediate pre-ablation myelography may facilitate recognition of the proximity of the spinal cord or nerve roots when performing ablation in the spine and sacrum [26]. And according to the open source image-guide surgery toolkit, presenting three image-guided systems for needle-based interventional radiology procedures is significant [19]. Therefore, establishing a feasible treatment planning is essential to assist physicians, which is focused on the accuracy ablation to avoid spinal cord damage.

Empirically, it's believed that data of vitro experiment and simulation is the basis of clinical treatments [27]. In the previous studies we have based on a true database set up personalized microwave thermal therapy of spinal tumor surgery planning ,which includes according to the patient's medical slice images 3d reconstruction of the real geometry of tumor and its surrounding tissues, imposed by the in vitro experiment measured SAR boundary conditions, using the finite element method for temperature field simulation [28].Nevertheless, containing all kinds of spinal tumors ablation results completely in the database is impossible. According to the shape and size of spinal tumor including maximum axial length, maximum transverse width and maximum vertical depth, it's essential to calculate and predict precisely the therapeutic parameters containing voltage and heating-up time. In addition, the great target volume is sensitive to the heating-up time and voltage in the radiofrequency ablation.

Owing to nonlinearity between therapeutic parameters (voltage and heating-up time) and ablation results, artificial neural network is employed to achieve the aim of treatment planning in the view of limited samples [29]. Currently, SVM [30], BP neural network and GRNN [31]-[32] are benefit to regression analysis prediction model for nonlinear multi-input multi-output.

This prospective study was to build a database about the relationship between accurate ablation results and therapeutic parameters. And the model of SVM was applied in the treatment planning to assist surgeons to estimate the therapeutic parameters including setting voltage and heating-up time. A treatment planning in spinal tumors is an essential tool to determine the therapeutic parameters rapidly and accurately, and it also can assist surgeons in the radiofrequency ablation of spinal tumors.

Methods

Simulation of spinal tumors in RFA

Because of multi-physics field solving comprehensive model exactly, this study applied ANSYS workbench as simulation soft to establish electromagnetic field and thermal field coupling. Among that, the frequency of electromagnetic field is at the lower frequency (460KHz), so using the Maxwell formulation to calculate the field is following by equation1-4 [27][33]-[35]. In the simulation, the blood perfusion rate could be ignored for ideal model. So the power generated by the RF was converted into the heat energy to make tumor temperature rise. Meanwhile, Pennes equation [36] was applied in simulation as the heat field in equation5.

$$\nabla \cdot (\sigma \nabla V) = 0 \tag{1}$$

$$E = -\nabla V \tag{2}$$

$$J = \sigma \cdot E \tag{3}$$

$$Q_r = J \cdot E \tag{4}$$

where V is the voltage (V), σ is the electrical conductivity (S/m), J is the current density (A/m^2) and Q_r is the volumetric heat generation rate (W/m^3) about electromagnetic heating.

$$\rho c\frac{\partial T}{\partial \tau}=k\nabla^2 T+\omega_b c_b(T_b-T)+Q_m+Q_r \tag{5}$$

where ρ is the tissue density (kg/m^3), c is tissue specific heat capacity(J/(kg·℃)), k is the thermal conductivity of tissue (W/(m·℃)), T is the temperature of tissue(℃), τ is the time (s), c_b is the specific heat capacity of blood (J/(kg·℃)), ω_b is the blood perfusion rate ($kg/(m^3 \cdot s)$), Ta is the blood temperature in the heating zone(℃), Q_m is the heat energy of biological tissue ($J/(m^3 \cdot s)$), Q_r is the energy of the external heat quantity.

Due to different sizes of spinal tumors, we built two kinds of the 3D ideal models to descript the comprehensiveness of ablation results in this study, including single needle electrode and multi needle electrode, which were regarded as common electrodes in clinic. But is questionable, the multi-needles electrode of the RFA probe may not even be opened as some spinal tumors are hard in consistency. In spite of this, the STAR Tumor Ablation System (DFINE, San Jose, CA) has designed can be curved up 90 degrees of bipolar needle electrode used in bone tumors [37], prove its effectiveness and safety , declare that there has further work can be carried out. The structures (Fig.1) of electrode is from previous studies [9][38]-[40].

(a) **(b)**

Figure1. (a) The structure of single needle electrode is divided into two cylinders including electrode tip and insulating shaft, meanwhile the diameter of cylinder is 1.8mm. The length of electrode tip and insulating shaft are 20mm and 60mm respectively.
(b) The structure of multi needles electrode [30] is divided into two cylinders and four needles. Among that, the length of two cylinders including electrode tip and insulating shaft are 10mm and 90mm respectively, meanwhile the diameter of cylinder is 1.8288mm. The distance is about 15mm between the bottom of electrode tip and each needle. Moreover, the diameter of four needles is 0.5334mm.

The boundary of electrode tip and insulating shaft are regarded as Dirichlet boundary and

Neumann boundary respectively [9][41]. Therefore, constant voltage is applied on the active part of the electrode and the ground is acted as other boundary. Material parameters in the numerical simulation are shown in Table 1.

Table 1. Thermal parameters of each material

Materials	k (W/m ·℃)	c (J/kg ·℃)	ρ (kg/m^3)	F(s/m)
Spinal tumor	0.42	4100	1000	0.4
Electrode tip	71	21500	132	4e6
Insulating shaft	0.026	1045	70	1e5

As reported, the time of solving pre-processing and post-processing could account for 95% by finite element simulation, besides, the time of analysis and solution spends less time which occupied 5% approximately [42]. So solving pre-processing is great significant to simulate by finite element method, especially mesh generation to determine whether or not the scientific and rational calculation results. We applied three kinds of grids namely coarse, medium and fine to generate, which could create mesh refinement. Among that, the mesh elements are respectively 16891, 55265 and 57937, as well as the element qualities are 0.93, 0.97 and 0.90 sequentially. According to the standard of meshing quality [43], the second case has high meshing quality by comparison, as enough mesh elements. Besides, in the second case, a majority of element quality is almost more than 0.95 by calculation, which could achieve the aim of great precise and logical ablation results.

Owing to the therapeutic parameters including voltage and heating-up time, we choose effective parameters in clinic. Based on previous studies [9][27][33][44][45], voltage supplied from transducer was allowed from 10V to 30V, moreover, the heating-up time was set from 120s to 1200s at intervals of 60 seconds. Meanwhile, to avoid high temperature (>100℃) to cause carbonization and hurt healthy tissue, keeping the highest time below 100℃ is necessary.

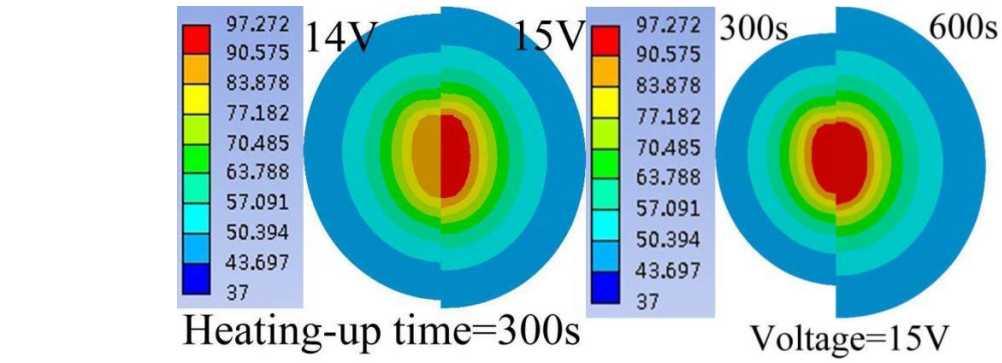
The treatment planning

To predict the therapeutic parameters (voltage and heating-up time) for providing the treatment data of surgical planning, it is essential to apply regressive prediction model to explore the relationship between therapeutic parameters and ablation results, based on several uncorrelated factors to affect ablation results. In the light of current studies about prediction models [29]-[32], three kinds of models have been used popularly, and they fit our simulation results characteristics, including BP neural network, GRNN and SVM. To verify the accuracy of prediction data and evaluate the superiority-inferiority in three models, randomly selecting 20 sets of data from simulation data (383 sets), including therapeutic parameters and ablation results, is regarded as test data. Among that, the voltage and heating-up time were predicted by the different models on the basis of characteristic between therapeutic and ablation results in database except the selected 20sets of data. Besides, the prediction data must be contracted with the actual data, which could prove the scientific nature of the evaluation model.

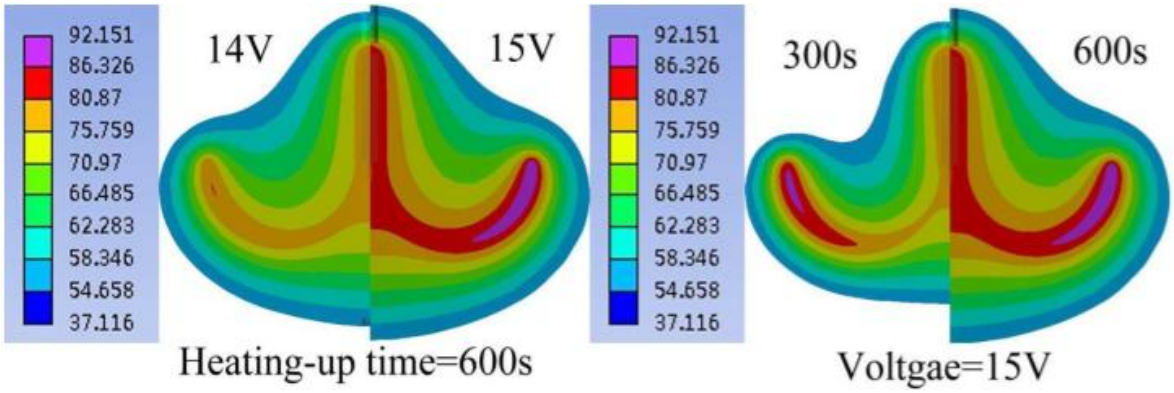
Results

Simulation of spinal tumors in RFA

Fig.2(a) shows the temperature distribution of various therapeutic parameters after simulation by single needle electrode. It was observed that the temperature distribution was symmetric and similar to rotundity. Whereas, the temperature distribution of multi needle electrode presented a mushroom shape in Fig.2(b), and with increasing the voltages or heating-up time, the temperature distribution had a tendency to expand a mushroom shape, particularly the concave changing, which was conformed the clinical research status compared with other studies[38][40].



(a)

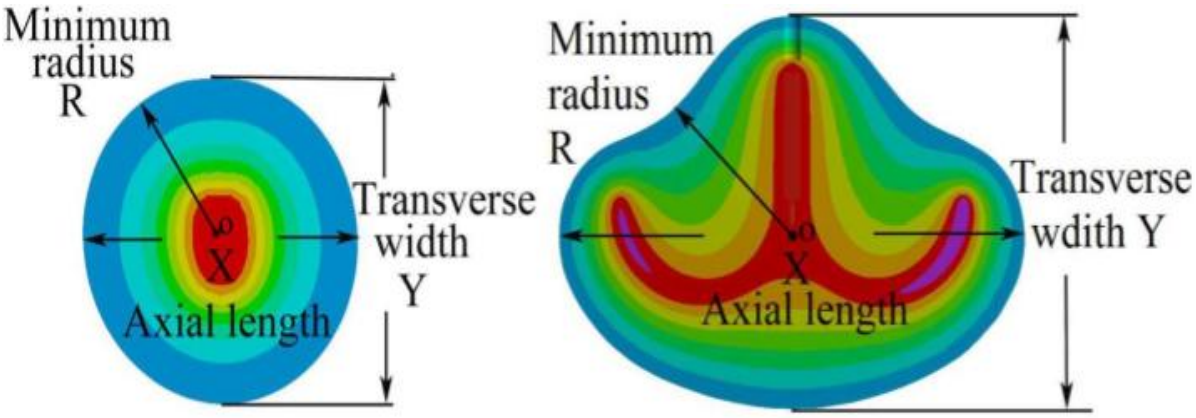


(b)

Figure2. (a) With increasing heating-up time and voltage, the temperature distribution(54℃) by single needle electrode is expanding as spherical shape.
(b) The temperature distribution by multi needles electrode is a mushroom shape at different heating-up times and voltages.

To illustrate the influence of therapeutic parameters (voltage and heating-up time) on ablation results including axial length(X), transverse width(Y) and vertical depth, we have measured the parameters of lesion to indicate the changing ablation results (see in Fig.3) and chosen some simulation results to summarize the variation in Table2 and Table3. The simulation results by single needle electrode were that the volume range of ablation zone was 70-17549mm^3, and the transverse diameter range of ablation zone was 5.4-31.8mm. So the small diameter (<30mm) spinal tumor could be treated by the single needle electrode in clinic. Furthermore, the volume range of simulation results by multi needles electrode was 81-130179mm^3. And the ranges of transverse diameter and minimum radius were 31.4-69.4mm and 3.9-30.1mm. The big diameter

(≥30mm) spinal tumor could be treated by the multi needles electrode in clinic. Compared with previous studies [13]-[16] about spinal tumors in clinical, our simulation results contain different sizes ranged between 2 and 40mm spinal tumors including primary benign, primary malignant and metastasis spinal tumors.



(a) By single needle electrode **(b) By multi needles electrode**

Figure3. The diagram vertical profile about temperature distribution (54℃) of the parametersmeasured. Among that, it shows axial length(X) and transverse width(Y) at the origin(o). The minimum radius(R) is to illustrate the peculiarity of temperature distribution, especially by multi needles electrode.

Table 2. Ablation results at different therapeutic parameters by single needle electrode

		14V				15V			
	Heating up time	X (mm)	Y (mm)	R (mm)	Volume (mm^3)	X (mm)	Y (mm)	R (mm)	Volume (mm^3)
Single needle electrode	300s	13.39	13.37	6.68	954.56	13.41	13.37	6.69	1175.13
	600s	16.36	15.94	7.97	1576.54	17.33	16.49	8.25	2032.27
expansion rate(%)		22.16	19.22	19.22	65.15	29.23	23.33	23.31	72.94

Table 3. Ablation results at different therapeutic parameters by multi needle electrode

		14V				15V			
	Heating up time	X (mm)	Y (mm)	R (mm)	Volume (mm^3)	X (mm)	Y (mm)	R (mm)	Volume (mm^3)
Multi needle electrode	300s	40.94	28.55	8.38	9089.2	42.34	29.52	9.65	10315.00
	600s	45.94	34.58	10.41	19650.3	48.93	36.88	12.95	23637.10
expansion rate(%)		12.21	21.12	24.22	116.19	15.56	24.93	34.19	129.15

As shown in Table2, with the increase the heating-up time, the expansion rate of axial length was 22.16%, meanwhile the expansion rate of transverse width was 19.22% at the voltage of 14V. Moreover, the expansion rate of axial length and transverse width were 29.23% and 23.33% at 15V. Furthermore, the volume was obvious to increase at different therapeutic parameters in Table2. After 300s, the volume was 954.56 mm^3 at 14V, and when the heating-up time was 600s, the volume was 1576.54mm^3, which was extended with a rate up to 65.15%. At 15V, the volumes were 1175.13mm^3 and 2032.27mm^3 respectively while the heating-up time were 300s and 600s. As the voltage increasing from 14V to 15V, it was evident that the expansion rate of volume also grew accordingly ranging 65.15% to 72.94%.

In addition, the change of ablation results had the similar rule to expand by multi needle electrode in Table3. By the multi needle electrode, the rates of enlarging axial length and transverse width were 15.56% and 24.93% from 300s to 600s at 15V. Moreover, the expansion rate of volume at 14V was larger than that at 15V.

Through the simulation, acquired 383 sets of data were regarded as database. On the basis of the change regularity of ablation results, the expand rates of the axial length, transverse width and vertical depth are difficult to seek a linear rule. Thus, the relationship between therapeutic parameters and ablation results is the characteristic of nonlinearity. It's essential to generate logical method to illustrate the influence of therapeutic parameters on ablation results.

The treatment planning

In terms of above method, Fig.4 shows the predicting data and actual data for different prediction models of SVM, BP neural network and GRNN.

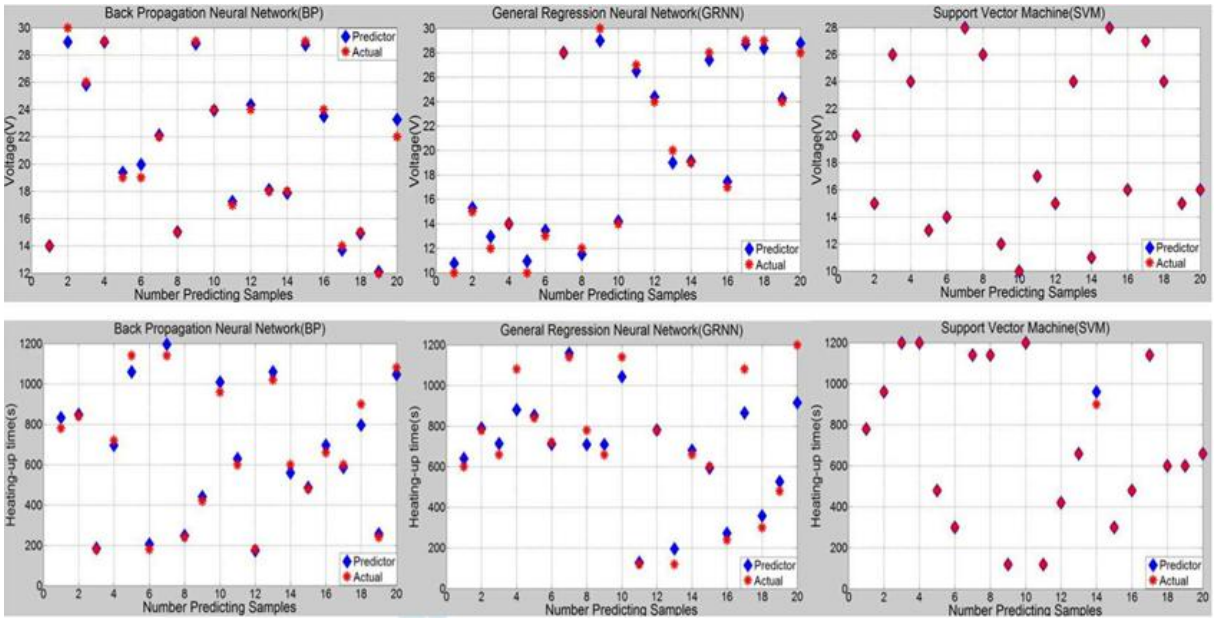


Figure4. The prediction results is to illustrate the accuracy by three models in selecting 20 sets of data. The icon of red asterisk represents actual data, which is about voltage and heating-up time by simulation. Meanwhile, the icon of blue rhombus stands for predictor data of voltage and heating-up time by three model. If the predictor data is

closer to actual data, the distance between blue rhombus and red asterisk is less. Beside, the complete covering of two kinds icons signifies that the predictor and actual data is the same, which means the high accuracy of prediction.

Although it has realized the goal to predict voltage and heating-up time by three models, the errors of prediction were different in Table4. By the statistical analysis between predicting data and actual data, we have calculated the error rate and correlation coefficients to evaluate the accuracy. In Table4, the maximum error rate of heating-up time prediction was 11.28% in GRNN model, meanwhile, the error rate of voltage prediction is also maximum in GRNN model. In the model of BP neural network, the prediction data was more accuracy than that in GRNN model; the error rates of voltage and heating-up time were 1.50% and 5.09%; the correlation coefficient were 0.996 and 0.991 respectively. However, in model SVM, the error rates were the smallest and the correlation coefficients were the closest to 1, no matter what prediction data, compared with that in other two models. Therefore, the model of SVM is the best one to predict in this database whether the error rate or the correlation coefficient.

Table 4. The statistical analysis in three models

	BP neural network		GRNN		SVM	
	Error rate(%)	Correlation coefficient	Error rate(%)	Correlation coefficient	Error rate(%)	Correlation coefficient
Voltage	1.50	0.996	3.22	0.992	0	1.000
Heating up time	5.09	0.991	11.28	0.964	0.33	0.999

In clinic, the voltage and heating-up time are estimated by the shape and size spinal tumor. Fig.5 shows the treatment planning to display the interface and function. In Fig.5, after writing the input value, the prediction data including voltage, heating-up time and volume could was shown in the interface. To illustrate the scientifically and precision of prediction, it was calculated the accuracy and correlation coefficient. Beyond that, the coverage was calculated to indicate whether ablating completely or not. Furthermore, the minimum radius was to descript the minimum radius by the multi needle electrode.

Figure5. The main program interface of software has three parts including input area, output area and evaluation area. According to input value about spinal tumor of axial length, transverse width, vertical depth and volume in input area, it's convenient to choosing different model containing single needle electrode model and multi needles electrode model. Moreover, the function is to predict correspondingly precise voltage, heating-up time and volume. The evaluation value is to verify the accuracy of prediction, based on the accuracy rate and correlation coefficient to show the accuracy, the coverage and minimum radius to explain whether complete cover.

Discussions and conclusions

This study describes the results of spinal tumor by radiofrequency ablation and the treatment planning to assist surgeons in surgical treatment.

Based on the clinical experience of ablation results, to make sure the ablation results scientifically and accurately is essential by numeral simulation analysis which can avoid the outside influential factor. In this study, the temperature distributions in Fig.2 are produced by different voltages and heating-up times, on the basis of the good meshing quality to improve calculating speed in addition to ensure the accuracy of simulation data. Compared with the previous study [5], the temperature distribution of simulation by RFA in this study is similar to that by MWA in previous study. It is also found that the highest temperature of ablation results is nearly located at the center of temperature distribution, and with increasing therapeutic parameters, the temperature distribution is increasing, and the center temperature is higher and higher.

It's believed that the spinal tumor ablation results by FEM in this study and other acquired results in clinical or ex-vivo studies could be combined together to verify the accuracy of data.

In the previous ex-vivo study [44], when the heating-up time was 600s and average energy was 1.52±0.76Kcal, the average long-axis diameter along the electrode was 3.05±0.46cm, meanwhile the average largest short-axis perpendicular diameters were 1.85±0.59cm and 1.26±0.48cm, moreover the average volume was 4.19±2.95cm^3. At the same conductions, the average energy 1.52±0.76Kcal is equal to the range from 14V-28V because of the conversion formula and mature research's conclusion [45]. When the voltage and heating-up time were 14V and 600s respectively, the axial length(X) was 1.64cm, the transverse width(Y) was 1.59cm, and the ablation volume was 1.57cm^3. Furthermore, 2.81cm axial length(X), 2.73cm transverse width(Y) and 7.29cm^3 volume were generated by 28V voltage and 600s heating-up time. Among that, the range of ablation volume is between 1.57cm^3 and 7.29cm^3, and other range of axial length(X) and transverse width(Y) is 1.64-2.81cm and 1.59-2.73cm. Compared with previous ablation results in ex-vivo spine, our simulation ablation results show that it has errors (0.33cm and 0.15cm) about ablation volume. Combined our ablation results with previous ablation at the same situations, there are some unavoidable elements to cause the significant errors, including different structures of electrode, different thermal parameters causing different heat transfer and the idealized numerical simulation model ignored some physiological parameters irrefutably. In spite of this, the ablation results are still of certain value for reference, which has the relatively accurate data of ablation results and the correct temperature distribution. Besides, the trendy of change in simulation of spinal tumor by RFA is the same to the situation in clinical. In consequence, the ablation results by single needle electrode are believed as the certain practical and significant reference and foundation in clinical.

Compared with other study [38] by multi needle electrode, the shape of our temperature distribution is resemblance. Due to the different tissue, we have changed the materials to illustrate the validity of simulation. The study has been reported that the diameter and the depth of the lesion were approximately 42mm and 36mm, and the lesion volume was approximately 19.3cm^3. At the same heating-up time, the simulation ablation results were 45mm(diameter), 35mm(depth) and 21.0cm^3 respectively, but relatively less error can be ignored caused by the other boundary condition, including 7.14%, 2.78% and 8.81%. Therefore, it's believed that the ablation results are regarded as basic database of treatment planning by numeral simulation.

In the terms of simulation data, we can find that at the same voltage, the expansion rate of axial length is bigger than that of transverse width with increasing heating-up time by single needle electrode. Moreover, as increasing the voltage, the expansion rates are all rising from 22.16% to 29.23% about the axial length. Actually the extending volume is more obvious than others in Table2 no matter which variable in therapeutic parameters. So the important factors are still voltage and heating-up time by the single needle electrode, and the axial length varies much more than transverse width by different heating-up time. Whereas, the expansion rate of axial length is smaller than that of transverse width with increasing heating-up time by multi needle electrode. Moreover, the expansion rate of R varies much more than the other two length, which ranges between 24.22% and 34.19%. Consequently, small tumor (maximum diameter<30.0mm) could be ablated by the single needle electrode and the large tumor (maximum diameter≥30.0mm) should be applied by the multi needle electrode.

Generally, in order to overcome the problem of the irregular shape of the spinal tumor, a treatment planning is provided to control the ablation region both completely and accurately. Due to the nonlinear relationship between therapeutic parameters and ablation results, it's advisable to select SVM method to predict data in treatment planning. It illustrates that the prediction data is approximately equal to actual data in the pictures in Fig.4.

Zheng [46] have built a prediction model by BP neural network model to illustrate feasibility of method in 0.988 correlation coefficient. Compared with the results by BP neural network model, the correlation coefficients are 0.999 and 0.996 respectively in prediction voltage and heating-up time. According to the study of Zheng, our study could be deemed to have accuracy to predict in BP neural network model. In fact, the correlation coefficients are 1 and 0.999 by SVM model, which is closer to 1 than that by BP neural network model. Besides, it has high accuracy for prediction, and the errors of prediction are respective 0 and 0.33% by SVM model. According to the prediction results, higher correlation coefficient and less error indicate that the prediction model by SVM is suitable for ablation in spinal tumor. And predicting the testing data describes the relationship between actual data and prediction data in Fig.4. Because the actual data is almost same to the prediction data, it explains the treatment planning is estimated successfully by SVM model in consequence. However, the reasons of prediction inaccuracy about other models attributes to the database characteristic and volume. Especially, the BP neural network model always depends on the rules of data, equivalently large data volume. Therefore, SVM model should be used in the system to predict surgical treatment of spinal tumors.

According to the SVM model and small sample by simulation, we have built the treatment planning, which could calculated the therapeutic parameters through inputting spinal tumor axial length, transverse width, vertical depth and volume. Meanwhile, it was calculated the accuracy and correlation coefficient. Beyond that, the coverage was calculated to indicate whether ablating completely or not. Therefore, this treatment planning have assisted surgeon to make precise surgical decision quickly and comprehensively.

Although the treatment planning based on numerical simulation is scientific and accurate, some functions are still generated imperfectly. Measuring axial length, transverse width and vertical depth of patient-specific spinal tumor should be provided in the interface of software. And because of the idealized numerical simulation model ignored some physiological parameters irrefutably, we will consider some medically important parameters in the model to acquire more accurate ablation results and consummate the database for reference in clinical. Besides, based on the complex environment of spinal tumor, we will study multi-material such as spinal cord, intervertebral disk, and spine and so on. In addition, it's essential to combine the imaging technology (CT, MR or X ray) with this treatment planning. And designing the path of electrode to insert spinal tumor seemly should be developed.

Acknowledgment

This research is supported by National Natural Science Foundation of China (No.31070754),

Education Project Scientific and Technological Program of Beijing Municipal Commission (KM201410005028).

References

[1] Stitzlein RN, Abdullah KG, Mroz TE. (2011) The Management of Upper Thoracic Spine Tumors. *Operative Techniques in Orthopaedics* **21**,225-234.

[2] Duong LM, Mccarthy BJ, Mclendon RE, Dolecek TA, Carol Kruchko BA, Ctr LL D, et al.(2012) Descriptive epidemiology of malignant and nonmalignant primary spinal cord, spinal meninges and cauda equina tumors, United States, 2004-2007. *Cancer* **118**, 4220–4227.

[3] Omeis IA, Mashaal D, Sciubba DM, Gottfried ON, Mcgirt MJ, Attenello FJ. (2011) Postoperative surgical site infections in patients undergoing spinal tumor surgery: incidence and risk factors. *Spine* **36**, 1410-1419.

[4] Saumet L, Deschamps F, Marec-Berard P, Gaspar N, Corradini N, Petit P.(2015) Radiofrequency Ablation of Metastases from Osteosarcoma in Patients Under 25 Years: The SCFE Experience. *Pediatric Hematology-Oncology* **32**, 41-49.

[5] Ding J, Zhang H, Nan Q, Liu Y, Du J, and Lu Y. (2013) Surgical Planning of Microwave Thermal Ablation for a Patient-specific Spinal Tumor. *Journal of Beijing University of Technology* **39**, 1264-1268.

[6] Tal A O, Vermehren J, Friedrichrust M, Bojunga J, Sarrazin C, Zeuzem S. (2014) Intraductal endoscopic radiofrequency ablation for the treatment of hilar non-resettable malignant bile duct obstruction. *World Journal of Gastrointestinal Endoscop*y **6**, 13-9.

[7] Lanuti M, Sharma A, Willers H, Digumarthy SR, Mathisen DJ, Shepard JO. (2012) Radiofrequency ablation for stage I non-small cell lung cancer: management of locoregional recurrence. *The Annals of thoracic surgery* **93**, 921-928.

[8] Abdullah B, Yeong C, Goh K, et al. (2015) Robotic-assisted thermal ablation of liver tumours. *European Radiology* **25**, 246-257.

[9] Gazis AN, Beuing O, Franke J, Jöllenbeck B, Skalej M.(2014) Bipolar radiofrequency ablation of spinal tumors: predictability, safety and outcome. *The Spine Journal* **14**, 604-608.

[10] Vanderschueren GM，Obermann WR，Dijkstra SP，et al. (2009) Radiofrequency ablation of spinal osteoid osteoma：clinical outcome[J].*Spine* **34**, 901-904.

[11] Nakatsuka A，Yamakado K，Takaki H，et al. (2009) Percutaneous radiofrequency ablation of painful spinal tumors adjacent to the spinal cord with real-time monitoring of spinal canal temperature:a prospective study[J]. *Cardiovascular & Interventional Radiology* **32**, 70-75.

[12] Neeman Z, Patti JW, Wood BJ. (2002) Percutaneous Radiofrequency Ablation of Chordoma [J]. *American Journal of Roentgen ology* **179**, 1330-2.

[13] Elian MMM, Sadek AF. (2015) What makes CT guided radiofrequency ablation for osteoid osteoma superior to open surgery in terms of pain control and patient's quality of life? *Egyptian Journal of Radiology & Nuclear Medicine* **186**, 949-955.

[14] Martel J, Ángel Bueno, Nieto-Morales ML, Ortiz EJ. (2009) Osteoid osteoma of the spine: CT-guided monopolar radiofrequency ablation. *European Journal of Radiology* **71**, 564-569.

[15] Hillen TJ, Anchala P, Friedman MV, Jennings JW. (2014) Treatment of metastatic posterior vertebral body osseous tumors by using a targeted bipolar radiofrequency ablation device: technical note. *Radiology* **273**, 261-267.

[16] Wallace AN, Greenwood TJ, Jennings JW. (2015) Radiofrequency ablation and vertebral augmentation for palliation of painful spinal metastases. *Journal of Neuro-Oncology* **124**, 1-8.

[17] Zhou H, Miller D, Schulte DM, Benes L, Bozinov O, Sure U, et al. (2011) Intraoperative ultrasound assistance in treatment of intradural spinal tumors. *Clinical neurology and neurosurgery* **113**, 531-537.

[18] Ali M F, Ray S. (2014) Offline RF thermal ablation planning using CT/MRI scan data. *Egyptian Journal of Radiology & Nuclear Medicine* 46, 141–150.

[19] Yaniv Z, Cheng P, Wilson E, Popa T, Lindisch D, CamposNanez E, et al. (2009) Needle-based interventions with the image-guided surgery toolkit (IGSTK): from phantoms to clinical trials. *IEEE transactions on biomedical engineering* **57**, 922-933.

[20] Glaiberman C B, Brown D B. (2004) Reversible neuropathy caused by overuse following radiofrequency ablation of metastatic pelvic lesions [J]. *Journal of Vascular & Interventional Radiology* **15**, 1307-1310.

[21] Goetz MP, Callstrom MR, Charboneau JW, et al. (2004) Percutaneous image-guided radiofrequency ablation of painful metastases involving bone: a multicenter study. *Journal of Clinical Oncology* **22**, 300–306.

[22] Kojima H, Tanigawa N, Kariya S, Komemushi A, Shomura Y, Sawada S.(2006) Clinical assessment of percutaneous radiofrequency ablation for painful metastatic bone tumors. *Cardiovascular & Interventional Radiology* **29**, 1022–1026.

[23] Nakatsuka A, Yamakado K, Maeda M, et al. (2004) Radiofrequency ablation combined with bone cement injection for the treatment of bone malignancies. *Journal of Vascular Interventional Radiology* **15**, 707–712.

[24] Nakatsuka A, Yamakado K, Takaki H, et al. (2009) Percutaneous radiofrequency ablation of painful spinal tumors adjacent to the spinal cord with real-time monitoring of spinal canal temperature: a prospective study. *Cardiovascular & Interventional Radiology* 32, 70–75.

[25] Philip A, Gupta S, Ahrar K, Tam AL. (2013) A spectrum of nerve injury after thermal ablation: a report of four cases and review of the literature. *Cardiovascular & Interventional Radiology* **36**, 1427–1435.

[26] Kurup A N, Schmit G D, Morris J M, et al. (2016) Avoiding Complications in Bone and Soft Tissue Ablation[J]. *Cardiovascular & Interventional Radiology* **40**, 166.

[27] Zorbas G, Samaras T. (2013) Parametric study of radiofrequency ablation in the clinical practice with the use of two-compartment numerical models. *Electromagnetic Biology & Medicine* **32**, 236-243.

[28] Ding J L, Zhang H J, Nan Q, et al. (2013) Personalized spine microwave thermal therapy of tumor surgery planning [J]. *Journal of Beijing University of Technology* **8**, 1264-1268.

[29] Shaban K, El-Hag A, Matveev A. (2009) A cascade of artificial neural networks to predict transformers oil parameters. *Dielectrics & Electrical Insulation IEEE Transactions on* **16**, 516 - 523.

[30] Doran G, Ray S. (2014) A theoretical and empirical analysis of support vector machine methods for multiple-instance classification. *Machine Learning* **97**, 79-102.

[31] Stamenkovi LJ, Antanasijevi DZ, Risti MĐ, Perić-Grujić AA, Pocajt VV. (2014) Modeling of methane emissions using artificial neural network approach. *Journal of the Serbian Chemical Society* **80**, 212-231.

[32] Lu W, Chu H, Zhang Z. (2015) Application of generalized regression neural network and support vector regression for monthly rainfall forecasting in western Jilin Province, China. *Iwa Publishing* **64**, 95.

[33] Ng E Y K, Jamil M. (2014) Parametric sensitivity analysis of radiofrequency ablation with efficient experimental design. *International Journal of Thermal Sciences* **80**, 41-47.

[34] Berjano EJ. (2006) Theoretical modeling for radiofrequency ablation: state-of-the-art and challenges for the future. *Biomedical Engineering Online* **5**, 24.

[35] Schutt D, Berjano EJ, Haemmerich D. (2009) Effect of electrode thermal conductivity in cardiac radiofrequency catheter ablation: A computational modeling study. *International Journal of Hyperthermia* **25**, 99–107

[36] Pennes HH. (1998) Analysis of tissue and arterial blood temperatures in the resting human forearm. Journal of Applied Physiology **85**, 5-34.

[37] Ma Y, Wallace A N, Madaelil T P, et al. (2016) Treatment of osseous metastases using the Spinal Tumor

Ablation with Radiofrequency (STAR) system[J]. *Expert Rev Med Devices* **13**, 1137.

[38] Tungjitkusolmun S, Staelin T, Haemmerich D, Jang-Zern T, Cao H, Webster JG, et al. (2002) Three-Dimensional finite-element analyses for radio-frequency hepatic tumor ablation. *IEEE transactions on bio-medical engineering* **49**, 3-9.

[39] Mulier S, Jiang Y, Jamart J, et al. (2015) Bipolar radiofrequency ablation with 2×2 electrodes as a building block for matrix radiofrequency ablation: Ex vivo liver experiments and finite element method modeling. *International Journal of Hyperthermia* **31**, 1-17.

[40] Haemmerich D, Tungjitkusolmun S, Staelin ST, Lee FT, Mahvi DM, Webster JG. (2002) Finite-element analysis of hepatic multiple probe radio-frequency ablation. *IEEE transactions on bio-medical engineering* **49**, 836-842.

[41] Zhang B, Moser M, Zhang E, Luo Y, Zhang W. (2015) Numerical analysis of the relationship between the area of target tissue necrosis and the size of target tissue in liver tumors with pulsed radiofrequency ablation. *International Journal of Hyperthermia* **31**, 715-725.

[42] Wang R, Chen H, Wang G. (2002) Analysis of ANSYS finite element mesh dividing. *Journal of Tianjin Polytechnic University* **21**, 8-11.

[43] Zhong Z, Guo X, Wang W, et al. (2013) Particle-based anisotropic surface meshing. *Acm Transactions on Graphics* **32**, 96.

[44] You NK, Lee HY, Shin DA, et al. (2013) Radiofrequency ablation of spine: an experimental study in an ex vivo bovine and in vivo swine model for feasibility in spine tumor. *Spine* **38**, E1121- E1127.

[45] Foppen S. (2010) Experimental and Numerical Analysis of Lesion Growth during Cardiac Radiofrequency Ablation[J]. *Journal of the Austrian Society of Agricultural Economics.*

[46] Zheng D, Zhu M, Liu W, Yao X, Pan K. (2014) Research on coagulation zone prediction induced by cooled-tip radiofrequency ablation base on the BP neural network. *Laser Journal* **35**, 95-97.

High Order Discontinuous Galerkin Method for the Euler Equations Using Curved Elements

†SU Penghui[1], ZHANG Liang[1]

[1]China Academy of Aerospace Aerodynamics, Beijing, China.

†Corresponding author: penghui.su@gmail.com

Abstract

In this study, a high order discontinuous Galerkin method for the two dimensional Euler equations is presented, the physical domain is divided into triangular elements which form an unstructured mesh. High order 6-node triangular elements with curved edges are introduced for curved physical boundaries. Polynomial functions up to fourth order are used as basis functions in each computational element. Fluxes between elements are calculated using HLLC approximate Riemann solver. An explicit third order Runge-Kutta time integration method is employed to solve the discretized systems. A number of test cases are presented to demonstrate the accuracy of this method. The results show that when curved elements are utilized, this method could archive its designed accuracy on domains with curved geometry boundaries.

Keywords: Discontinuous Galerkin Method, Euler equations, unstructured mesh.

Introduction

The discontinuous Galerkin method (DGM) was first proposed by Reed and Hill[1] for solving neutron transportation problems, since then, the DGM are extensively used in many areas, which include fluid simulations, MHD simulations, shallow water simulations and many others. In the field of computational fluid dynamics, finite difference method (FDM) and finite volume method (FVM) have long history of applications. FDM is suitable for building high order numerical schemes, but it has many difficulties when dealing with complex geometries and unstructured meshes. FVM could be implemented on unstructured meshes and complex geometries easily, but it is hard to construct high order compact FVM schemes. The DGM has both the advantages of FDM and FVM. The discrete unknown variables of DGM are linear combinations of element-wise polynomial basis functions, high order DGM schemes could be built on arbitrary meshes with compact stencils.

High order DGM is much more sensitive to the numerical boundary conditions than other methods[2,3]. When the physical boundaries are curved, boundary elements with straight edges may not meet the demand of DGM and lead to unphysical solutions. In these cases, accurate representations of curved boundaries are crucial for performing high order DGM simulations.

The present authors have developed a 2D DGM for Euler equations on unstructured meshes. High order elements with curved edges are utilized at physical boundaries. Numerical tests show that when curved elements are utilized, DGM could archive its designed accuracy on domains with curved physical boundaries.

Governing Equations

The conservative forms of Euler equations are

$$\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}=0 \tag{1}$$

Where *U*, *F* and *G* refer to conservative state vector, *x*-direction inviscid flux and *y*-direction inviscid flux respectively.

$$U=\begin{bmatrix}\rho\\ \rho u\\ \rho v\\ \rho E\end{bmatrix};\ F=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ u\left(\rho E+p\right)\end{bmatrix};\ G=\begin{bmatrix}\rho v\\ \rho uv\\ \rho v^2+p\\ v\left(\rho E+p\right)\end{bmatrix} \tag{2}$$

To enclose the equation system, the equation of state is introduced.

$$p=\rho RT \tag{3}$$

Discontinuous Galerkin Method

The physical domain is divided into non-overlapping elements *K*. The unknowns U_h on each element are expressed as linear combinations of basis functions.

$$U_h=U_i\phi_i \tag{4}$$

In each element, the weak form of governing equations is introduced.

$$\int_{K_j}\left(\frac{\partial U}{\partial t}+\frac{\partial F}{\partial x}+\frac{\partial G}{\partial y}\right)\phi_i\mathrm{d}\Omega=0 \tag{5}$$

With some manipulations, the equations have the following form.

$$\int_{K_j}\frac{\partial U}{\partial t}\phi_i\mathrm{d}\Omega-\int_{K_j}\left(F\frac{\partial\phi_i}{\partial x}+G\frac{\partial\phi_i}{\partial y}\right)\mathrm{d}\Omega+\int_{\partial K_j}\phi_i\ \ Fn_x+Gn_y\ \ \mathrm{d}S=0 \tag{6}$$

The solution of DGM has multiple values on element boundaries. In order to determine a unique value of inter-cell fluxes, numerical flux functions are introduced, in this study, the HLLC Riemann solver[4] is adopted.

At each time step, a system of ordinary differential equation is formed

$$M_{K_j}\frac{dU}{dt}=R \tag{7}$$

Where *R* is residual term, *M* refers to the mass matrix on element K_j. A third order explicit Runge-Kutta scheme is utilized to solve this ordinary equation system.

Numerical Results

The 2D subsonic flow around a cylinder[5] at Mach number 0.38 is chosen to demonstrate the performance of DGM with the existence of curved boundaries. The radius of cylinder is 0.5, computational domain is bounded by a circle of r=20. four successively refined meshes are generated, which contains 16×4, 32×8, 64×16 and 128×32 points. Details about these meshes could be found in [6], the meshes are shown in Fig.1.

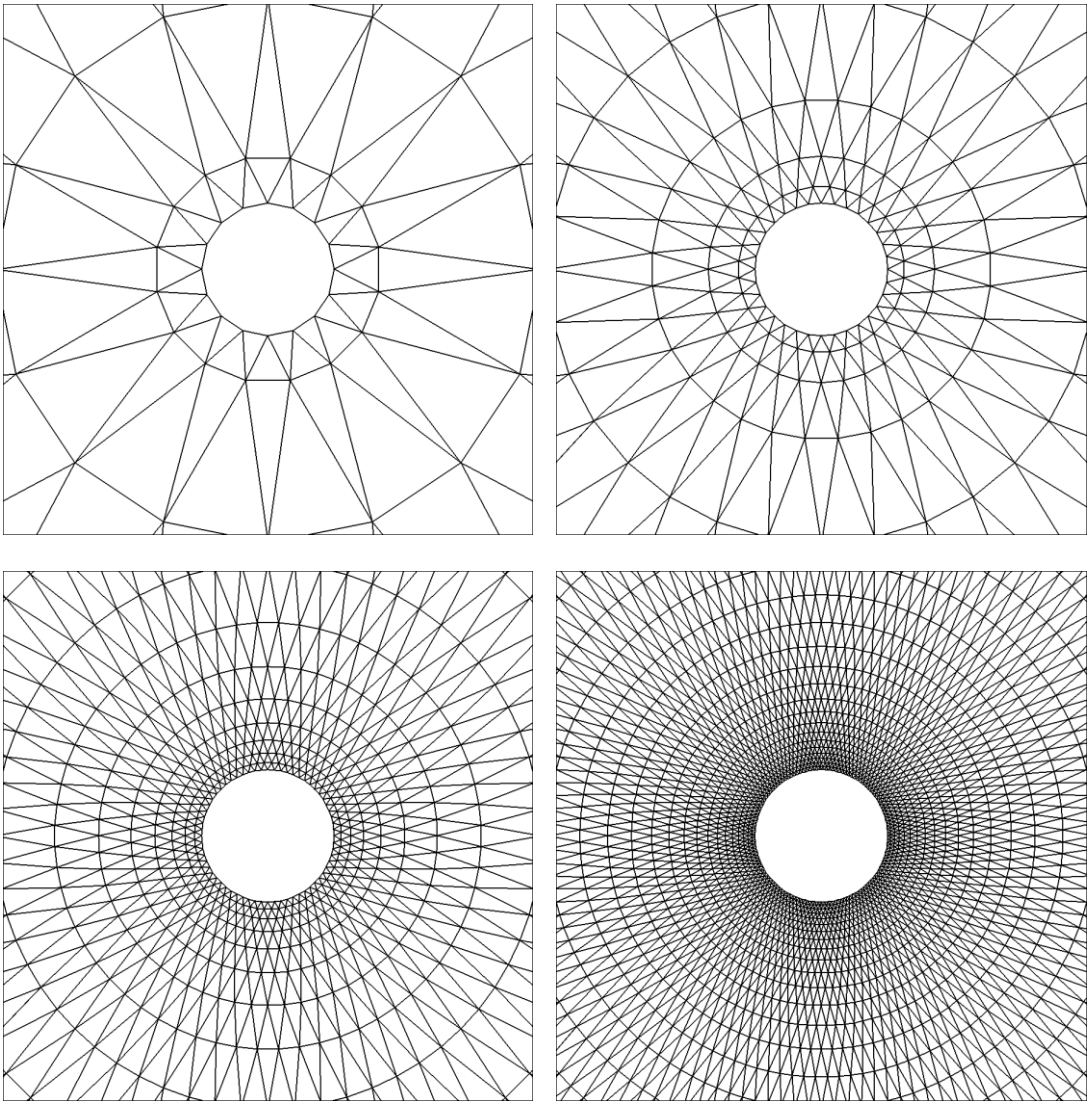
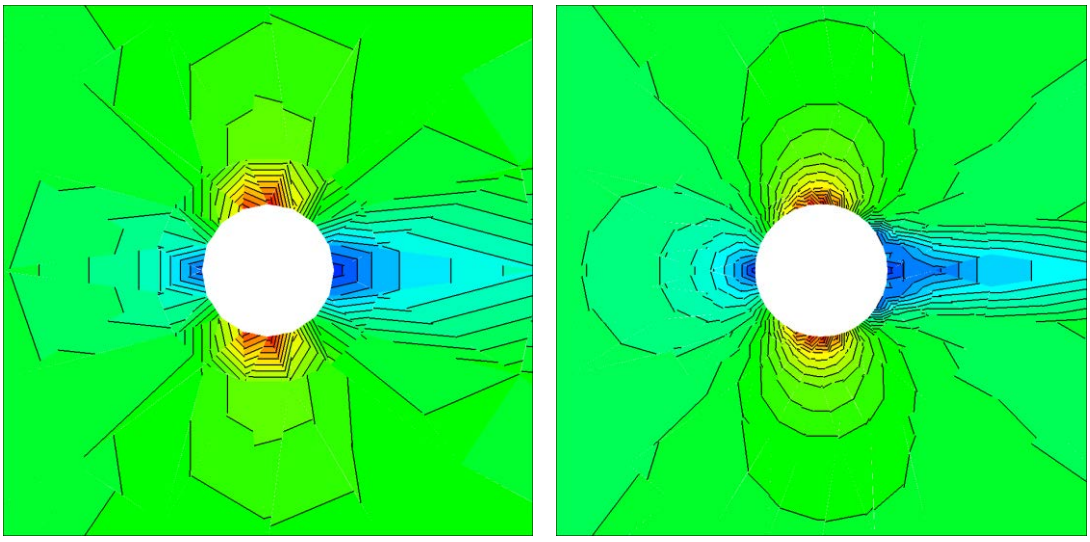


Figure 1. Successively refined meshes of cylinder, 16×4 (top left), 32×8 (top right), 64×16 (bottom left) and 128×32 (bottom right)



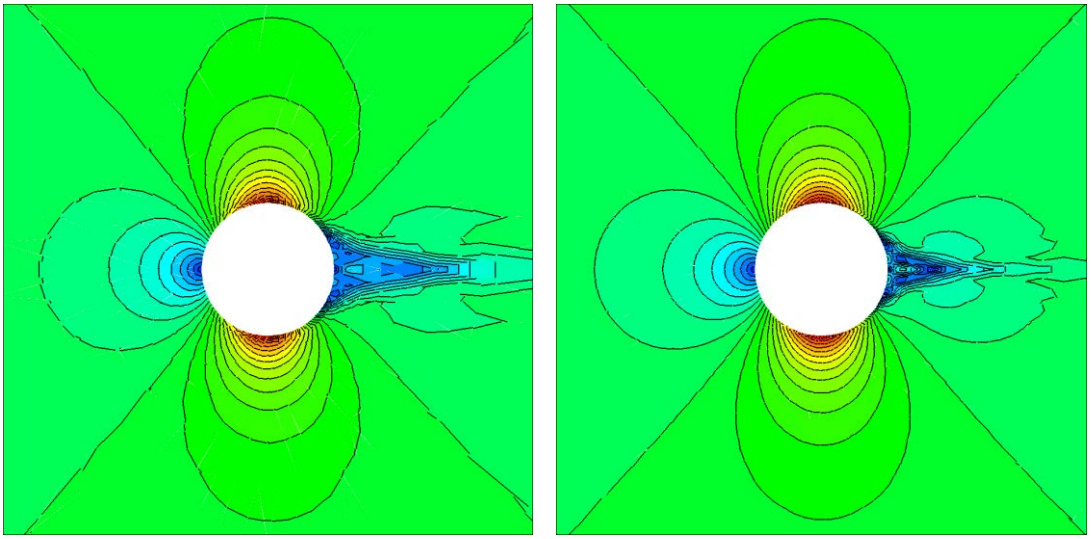


Figure 2. Mach contour using mesh 16×4 (top left), 32×8 (top right), 64×16 (bottom left) and 128×32 (bottom right), with straight-edge triangular elements

Mach contours using these meshes of first order DGM are shown in Fig.2, ΔM=0.038. Theatrically, the contour should have a symmetric pattern, but all of these results show unphysical wakes along the downstream of cylinder, and these unphysical wakes will not disappear with the successive refinement of meshes. This indicates that straight edged elements could not fit curved boundaries and cause wrong solutions.

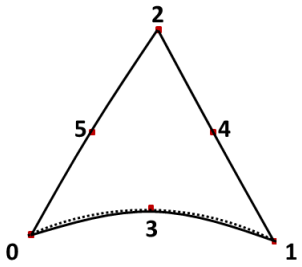


Figure 3. 6-node triangular element (solid) fitting a curved boundary (dashed)

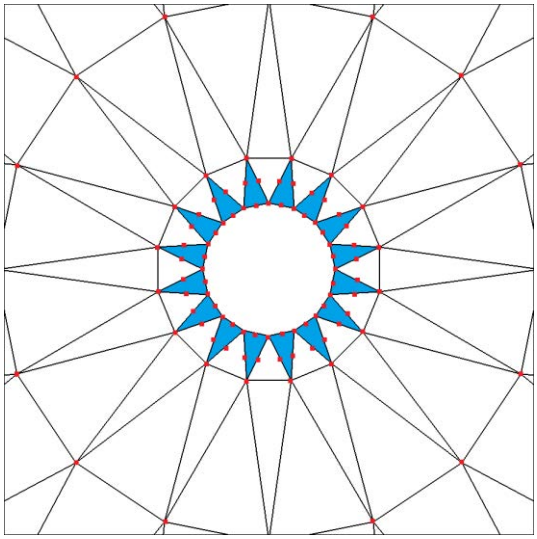


Figure 4. Mesh with 6-node curved triangular elements (blue) near boundary

In order to generate meshes for curved boundaries, the high order 6-node curve edged triangular element is introduced, and it is shown in Fig.3. With curved elements at cylinder

boundary, a new mesh is generated (Fig.4). This mesh contains both curve edged elements for boundary fitting and straight edged elements for space filling.

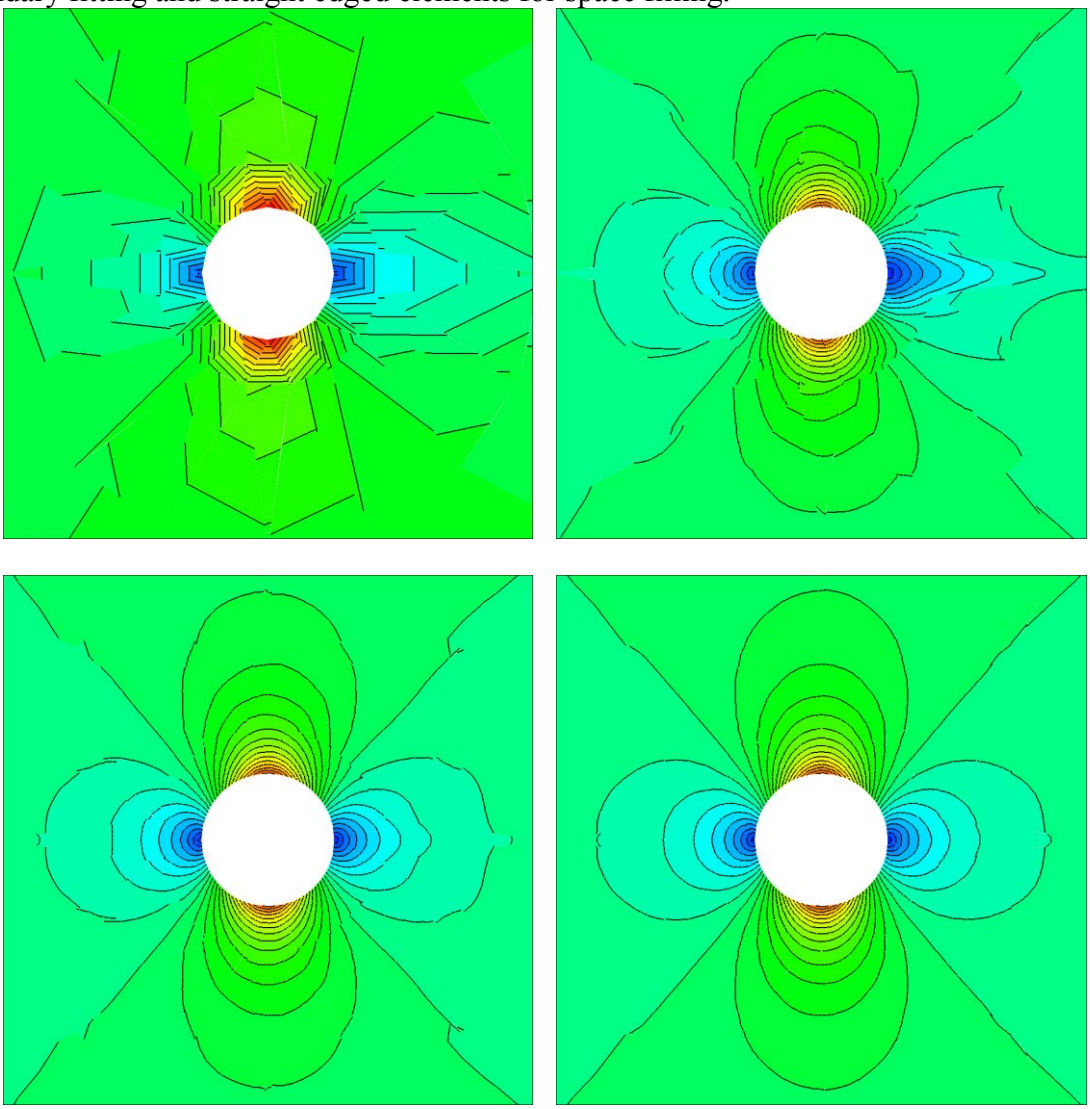


Figure 5. Mach contour on 16×4 mesh with p=1 (top left) p=2 (top right), p=3 (bottom left) and p=4 (bottom right), with curved boundary elements

Mach contours obtained with up to fourth order DGM using this new mesh are shown in Fig.5. The results show that, with the utilizations of curved elements near cylinder boundary, numerical dissipations are reduced, and the symmetric pattern of Mach contour is obtained with high order DGM, which implies the correct discretization of curved boundaries with high order curved elements is crucial in high order DGM computations.

Conclusions

In this study, a discontinuous Galerkin method with curved boundary treatment is presented. The curved solid boundary is fitted with high order 6-node triangular elements. The results of subsonic cylinder flows show that, DGM is very sensitive to the treatment of curved boundaries, high order curved elements must be utilized to archive its designed accuracy on domains with curved physical boundaries.

References

[1] Reed, W. H. and Hill, T. R. (1973) Triangular mesh methods for the neutron transport equation, *Los Alamos Scientific Laboratory Report*, LA-UR-73-479.

[2] F. Bassi and S. Rebay. (1997) High-Order Accurate Discontinuous Finite Element Solution of the 2D Euler Equations, *Journal of Computational Physics* **138**, 215–285.

[3] Yidong. Xia, Hong. Luo, Seth. Spiegel and Megan. Frisbey. (2013) A Parallel Implicit Reconstructed Discontinuous Galerkin Method for the Compressible Flows on 3D Arbitrary Grids, *21st AIAA Computational Fluid Dynamics Conference*, AIAA 2013-3062.

[4] Toro, E. F. (2009) Riemann Solvers and Numerical Methods for Fluid Dynamics –A Practical Introduction, Springer, New York

[5] Hong. Luo, Joseph. H. Baum and Rainald. Lohner. (2006) A fast, p-Multigrid Discontinuous Galerkin Method for Compressible Flows at All Speeds, *44th AIAA Aerospace Sciences Meeting and Exhibit*, AIAA 2006-110.

[6] Lilia. Krivodonova. and Marsha. Berger. (2006) High-Order Accurate Implementation of Solid Wall Boundary Conditions in Curved Geometries, *Journal of Computational Physics* **211**, 492–512.

Effect of stent designs on the paravalvular regurgitation of transcatheter aortic valve implantation

***Jin Chang[1], Liu Rong-hui[1], Zhong Sheng-ping[2], †Wang Li-zhen[1], and †Fan Yu-bo[1.3]**

[1] Key laboratory for biomechanics and Mechanobiology of Ministry of Education, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China
[2] KingstronBio(Changshu) Co.,Ltd,Chang Shu，215500, China
[3] National Research Center for Rehabilitation Technical Aids, Beijing 100176, China).

*Presenting author: mz327@163.com
†Corresponding author: lizhenwang@buaa.edu.cn, yubofan@buaa.edu.cn

Abstract

Objective Transcatheter aortic valve impantation(TAVI) rapidly developed in recent decade, however, paravalvular aortic regurgitation (AR) , as a complication, significantly influences the morbidity and mortality after TAVI. In this study, it was evaluated that effect of stent design on the paravalvular regurgitation of transcatheter aortic valve implantation based on numerical simulations. **Methods** Three self-expanding transcatheter aortic valve stent designs model were developed base on the commercial products. Three stents had the same inflow end but different outflow end. Stent1 had more struts on the outflow end which is similar first generation stent. Stent 2 had a sparser outflow end. Stent 3 had most sparse struts on the outflow end. They were radial compressed and implanted into the representative calcified human aortic root models. Then the effect of different stent design on the aortic root stresses, stent deformations and the gaps between the stent and aortic root was analyzed in order to understand the relationship to paravalvular regurgitation. **Results** The same inflow end and the different degree of sparse outflow end was found in the three stent, which induced the different stent deformation, a different stress on the calcified plaques and aortic root, and a different gap between the stent and aortic root. Both an excessively dense or sparse outflow end design resulted in a lager stent deformation, a higher stress calcified plaques and aortic root, and a lager gap, which led to more serious paravalvular regurgitation. Stent 2 had a moderate sparse outflow end obtained the smallest deformation, the lowest stress, and the smallest gaps, indicate this design will lead to a low risk of paravalvular regurgitation. **Conclusions** An excessively dense or a sparse outflow end would result in a larger stent deformation indexes, a higher stress calcified plaques and aortic root, and a lager gap, and led to serious paravalvular regurgitation. This study provided the guidance for design of transcatheter aortic valve. And it would help predict the clinical outcome after implanted.

Keywords: Transcatheter aortic valve implantation; Finite element model; Stent design; Paravalvular regurgitation

1 Introduction

Transcatheter aortic valve implantation (TAVI) have been first used in human in 2002[1], and developed rapidly in the last decade. It has been broadly used to treat aortic valve stenosis of impossible surgery and high-risk patients [2-3]. However, paravalvular aortic regurgitation (AR) significantly influences the morbidity and mortality after TAVI. Clinical follow-up indicated moderate or severe AR post-TAVI was about 6% to 21% [4]. It is considerably

higher than that surgical valve replacement. 1-year mortality in patients with moderate or severe AR was 60.0%, compared with 19.6% in patients with none/mild AR [5]. Median survival time in patients with at least moderate AR was 1.7 years, compared with 3.4 years in patients with mild/trivial/no AR [6]. Recently study indicated that mild AR was also associated with an increased hazard ratio for mortality [7]. So it is very important to investigate the influencing factor and method to reduce the incidence of the paravalvular regurgitation.

Numerical simulation had been used to analyze TAVI, especially in patient-specific simulations of transcatheter aortic valve implantation [8]. Stent apposition and deformation, stress on valve leaflets, stent, and calcified plaque were studied to evaluate the ability of anchoring by finite element analysis [9-11]. It was indicated that lager calcified spot stress will lead to risk of stroke [11]. Aortic rupture location had been investigated and got the same outcome in clinic [12]. It was also simulated the influence of the asymmetric coaptation, anchoring position and angle, calcification position, pattern and size to paravalvular regurgitation and clinical outcome in previous studies [9][11][13-14]. Recently, a workflow had been developed to simulate the potential leakage prior to the implantation to help decide the best implant type, size and position [15].

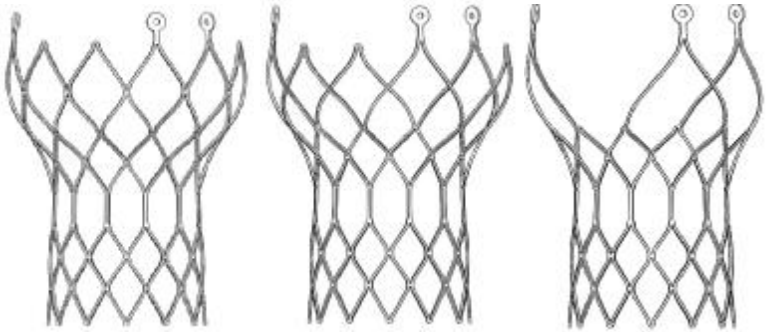
There are little studies was reported that the effect of the stent designs on the paravalvular regurgitation based on numerical simulation. Therefore, it would be analyzed the deformation of stent, the stress of aortic wall and calcified plaque, the gap of stent and aortic wall after TAVI. And evaluate the effect of stent designs on paravalvular regurgitation of different calcification patterns in this study.

2 Methods

1.1 The three stent design models

Three self expanded stent designs was established, and all the stent designs come from KingstronBio (Changshu) Co, Ltd, China. Self expanded stents was widely used as transcatheter aortic valve stents, such as CoreValve and Portico transcatheter aortic valve. The diameter of the outflow was in generally bigger than the inflow diameter for self expanded stent. The smaller inflow end anchored on the aortic valve root, and the bigger outflow end archored on the aorta ascends. The three stents had the same inflow end construction, but with different degree of sparseness for outflow end design, as shown in Fig.1. The outside diameter of the inflow end is 26mm for all the stents. Stent1 had more struts on the outflow end which is similar first generation stent. Stent 2 had a sparser outflow end. Stent 3 had most sparse struts on the outflow end.

The stents were structured by Solidworks(Dassault Systemes, France), divided into hexahedral element from adopting Hyoermesh 12.0(Altair Engineering USA), and imported into ABAQUS 6.13(SIMULIA, USA). This study adopted a linear-elastic, isotropic material model to simplify the superelasticity of the Nitinol stent materials [14]. TheYoung modulus, poisson ratio and density of Nitinol were set 50000Mpa, 0.3 and 6450 kg/m3 [16]. Because of the negligible effect of sealing skirt and leaflet compared to Nitinol alloy [17], the skirt and leaflet were not included in this investigation.



Stent 1 Stent 2 Stent 3

Figure 1 Three stent designs with different sparse outflow design

1.2 Aortic root model

The healthy aortic root model was shown in Fig. 2, include the left ventricular outflow tract (LVOT), aortic sinus, aortic heart valve and ascending aorta [17-18]. Aortic sinus includes right coronary sinus, left coronary sinus, and non-coronary sinus. The thickness was set 2.5mm for all the LVOT, aortic sinus, aortic anulus and aorta ascends, meshed by C3D10M tetrahedral element. The diameter of the aortic annulus was set as 24mm. Leaflet thickness was set as 0.3mm [19], and meshed by S4R shell element. All the materials were set linear-elastic, isotropic material. Young modulus, Poisson ratio and density of Nitinol were set to 2Mpa, 0.45 and 2000 kg/m3, respectively.

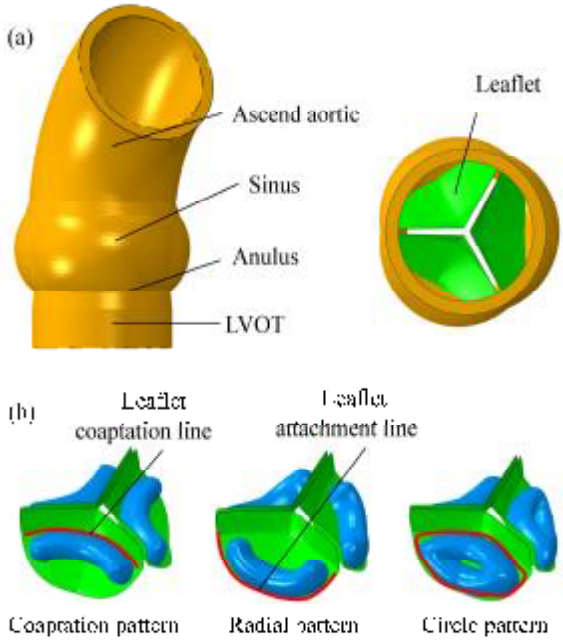


Figure.2 The model of (a) the health aortic root; (b) the three calcification patterns

According to correlation analysis, we established three calcification pattern: coaptation pattern, radial pattern, and circle pattern (Fig. 2b).Calcification plaque was arc shaped and located along the coaptation line for coaptation pattern; calcification plaque was along the attachment line for the radial pattern, and circle pattern was the combined pattern of the two, which represented the severe calcification. The maximum thickness of calcification plaque is 4.5mm, and the volume of the three plaques was 1085mm^3, 1083mm^3 and 1696mm^3, respectively. The material of the calcification plaque were set linear-elastic, isotropic material. Young modulus, Poisson ratio and density were set 12.6MPa,0.3 and 2000kg/m3, respectively[13].C3D10M tetrahedral element was adopted for calcification plaque.

1.3 Boundary and loading conditions

The deployment process was shown in Fig.3.A rigid cylinder shell was adopted to control the stent crimp and release. Firstly crimp the stent into the cylinder shell. Constrain the inflow end to move along the longitudinal center line during the crimp process. The compressive pressure was loaded on the leaflet surface to make the leaflet open, withdrawn the shell to release the stent and the pressure on the leaflet simultaneously. No friction was set between the stent and shell. The stent would contact with leaflet, aortic wall, and calcification plaque until the stent was fully deployed. Frictional coefficient was set 0.2 among the stent and leaflet, aortic wall, and calcification plaque. The distance of the stent bottom under the aortic anulus was set 4.5mm. No friction was set between the stent and shell.

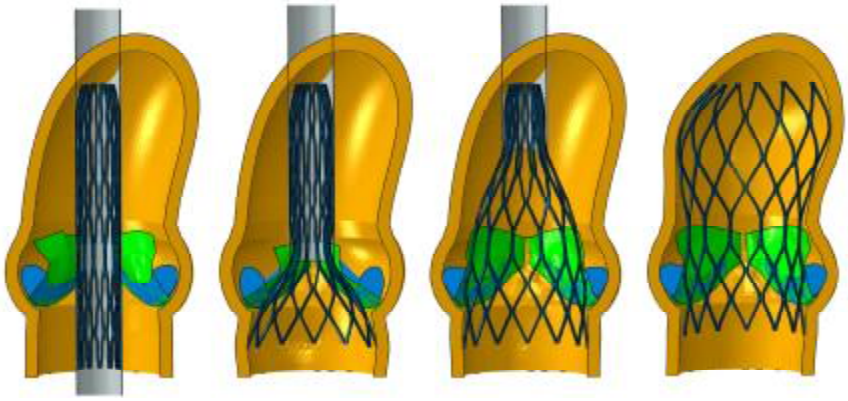


Fig.3 Steps of the valve stent deployment

The stent deformation, Von Mises stress and the maximum principle stress of aortic wall and calcification plaque, gaps between the stent and aortic wall was analyzed. Six planes was set on the position near the leaflet. The coordinate of the stent on the plane was extracted, and geometrical center and distance from the centre to the strut were calculated using the MATLAB R2014b (MathWorks, USA). The maximum and minimum distances were defined as r_{max} and r_{min}. The ratio of the maximum and minimum distance, $r_{max}/r_{min,}$ was defined as stent defomation index e. Stent defomation index indicated the comformity of the stent deformation. Fig. 4 showed the position of the 6 planes and r_{max} and r_{min} of the plane 1.

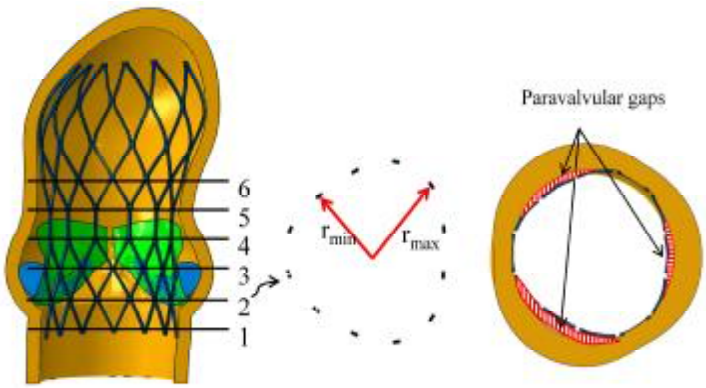


Fig.4 Measure of deployed valve stent distortion and paravalvular gaps (R: right coronary sinus, L: left coronary sinus, N: non-coronary sinus)

Paravalvular regurgitation would happen when there was a bad contact between the stent and aortic root. The degree of the contact was calculated by the gaps area of the stent and aortic root [9]. It was found that gaps located under the right coronary sinus, left coronary sinus, and non-coronary sinus, therefore, the three gaps were calculated individually and then got the total gap. Fig. 4 showed the gaps in plane 2.

3 Results

3.1 Deformation of stent

As Fig.5 shown, the stent deformation index for all stents and calcification pattern, increased from the inflow end to the outflow end. It arrived on the peak value at plane 3 or plane 4, and then decreased. However, there are many distinguish on each planes for different stents and calcification patterns. Stent 2 had the minimum stent deformation index at plane 2 for all the calcification patterns.

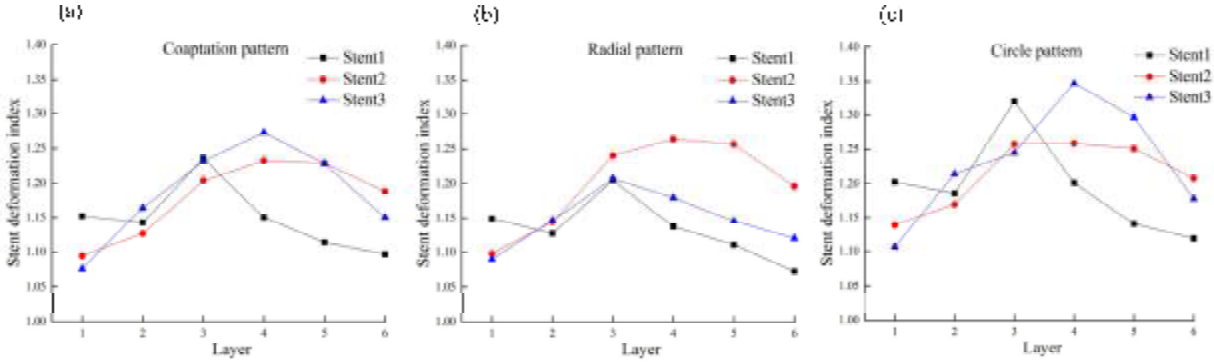


Fig. 5 Stent deformation index of (a)Coaptation pattern; (b)Radial pattern;(c)Circle pattern

3.2 Stress for calcification and aortic wall

Table 1 provided the peak value of von mises stress and maximum principle stress of aortic wall and calcification plaque. Stent 2 had the lowest peak value of the von mises stress and the maximum principal stress for all three calcification patterns. Stent 3 had the highest stress for all the patterns.

Table 1 Max stress of aortic wall and calcification plaque

Pattern and stent	Aortic wall		Calcification plaque	
	Max. von mises	Max. principle sress	Max. von mises	Max. principle stress
Coaptation pattern				
1#	2.3	0.9	7.2	2.8
2#	1.0	0.7	6.9	2.6
3#	2.1	2.4	8.6	10.5
Radial pattern				
1#	1.8	2.0	10.5	9.3
2#	1.5	1.8	9.9	8.6
3#	1.7	2.0	10.7	9.2
Circle pattern				
1#	1.7	0.8	7.9	5.1
2#	1.1	0.8	5.7	4.3
3#	2.0	2.3	11.4	11.4

Table 2 described the location of the maximum stress. The aortic wall maximum stress of all stents happened at leaflet commissure, which located on the plane 3-4. For the radial and circle pattern, the calcification plaque maximum stress of all stents happened on the plane 3. However, for the coaptation pattern, the maximum stress of stent1 and 2 located on the plane 3-4.

Table 2 Position of the max. Von mises stress point.

Stent and pattern	Calcification plaque	Aortic wall
coaptation pattern		
1#	plane 3-4	plane 3-4
2#	plane 3-4	plane 3-4
3#	plane 3	plane 3-4
Radial pattern		
1#	plane 3	plane 3-4
2#	plane 3	plane 3-4
3#	plane 3	plane 3-4
Circle pattern		
1#	plane 3	plane 3-4
2#	plane 3	plane 3-4
3#	plane 3	plane 3-4

Fig.6 was the stress distribution diagram of stent 1 for different calcification patterns. It showed the location of the maximum stress for aortic wall and calcification plaques.

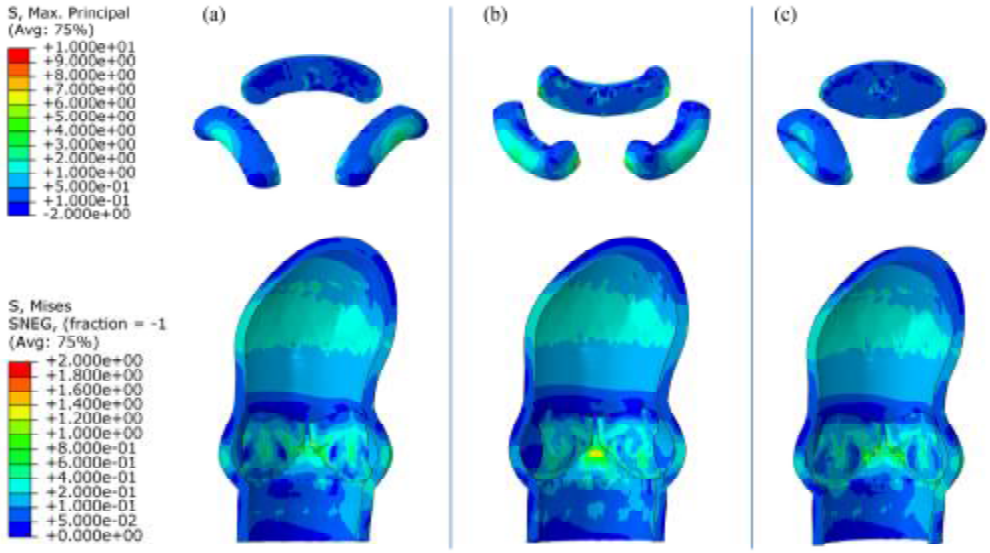


Fig. 6 Calcification plaque and aortic wall stress distribution diagram of (a)Coaptation pattern; (b)Radial pattern;(c)Circle pattern

3.3 Paravalvular Gaps

Fig. 7 showed the paravalvular gaps of all the stents and calcification patterns on plane 2. Stent 1 had the maximum total paravalvular gaps for all three calcification patterns. Stent 2 and 3 had the similar total paravalvular gaps. The total gap for radial pattern was larger than other two patterns. The coaptation pattern had the smallest total gaps. The gap near the right coronary sinus is larger than which near left and no coronanry sinus. The gap near right coronary sinus of stent 1 was significantly larger than other two stents for all three

calcification patterns.

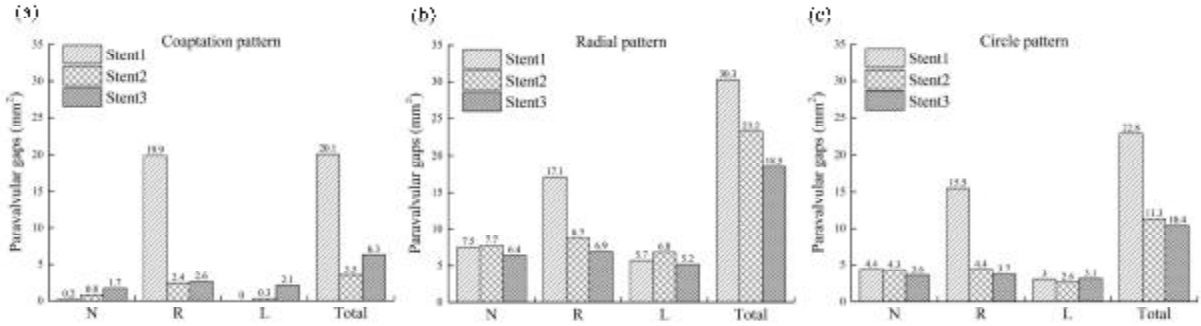


Fig. 7 Paravalvular gaps of (a)Coaptation pattern; (b)Radial pattern;(c)Circle pattern

4 Discussion

Transcatheter aortic valve impantation(TAVI) rapidly developed in recent decade. Paravalvular aortic regurgitation (AR) as a complication significantly influences the morbidity and mortality after TAVI [5-7]. In this study, it was evaluated that effect of stent design on the paravalvular regurgitation of transcatheter aortic valve implantation based on numerical simulations. Three self-expanding transcatheter aortic valve stent designs model were developed base on the commercial products. Three stents had the same inflow end but different outflow end. They were radial compressed and implanted into the representative calcified human aortic root models. The effect of different stent design on the aortic root stresses, stent deformations and the gaps between the stent and aortic root was analyzed and discussed in follow to reveal the relationship to paravalvular regurgitation.

4.1 Deformation of stent

Stent deformation was commonly found after TAVI. It commonly appeared triangle and ellipse [20]. Noncircular stent would lead to a higher stress on the leaflet [21-22].

In this study, stent 1 had the maximum stent deformation index for all three calcification patterns at plane 1. Three stents had the same design except for the outflow end, therefore, the different degree of sparseness caused the influence to deformation index.

The stent anchored on plane 2 which was the plane of aortic annulus. The deformation of stent on this plane was simultaneously affected by stent design and calcification plaque. Stent 1 had a higher deformation index than stent 2 for coaptation and circle patterns on this plane. It indicated that the deformation of stent 1 was more inhomogeneous compared to stent 2. The sparseness caused a critical influence relative to the calcification. However, the stent 3 had the maximum index for the two calcification patterns. It may be the reason that the stent 3's excessively sparse outflow end made the support force obviously decrease, resulting in a serious inhomogeneous deformation. This indicated calcification may be has the primary effect to the deformation. All the three stents had the similar deformation for the circle pattern. For the circle pattern, the calcification plaque along the coaptation line and the attachment line commonly effect the stent deformation. The degree of sparse for the outflow had a little influence for circle pattern.

Stent 1 still had a higher index than stent 2 on plane 3 for coaptation and radial patterns. This still indicated that deformation of stent 1 was more inhomogeneous than stent2. The degree of the sparseness still caused a critical influence relative to the calcification. Calcification plaque mainly located near the plane 3 for coaptation pattern, therefore, calcification plaque made an important effect to stent 3 and result in a higher deformation index. Calcification plaque was far away from this plane, so there is a lower deformation index of stent 2 and 3 for circle patterns. There was a similar index for all the three stents for radial pattern, which indicated the calcification made a primary effect to radial pattern.

The plane 4, 5 and 6 were far away from the calcification plaque, the deformation index was determined by degree of the sparseness of stent outflow end. Less sparseness outflow end made the bend aortic easily straighten, therefore, got a more concentricity deformation. However, if the outflow end was too sparse, there was no strut at the maximum deformation spot. And result in a lower deformation index, such as stent 3 on plane 6. These planes had a little effect to the paravalvular regurgitation.

Stent 2 had the lowest index for three calcification patterns. Higher deformation index easily lead to lager paravalvular gaps and result in more serious paravalvular regurgitation. There was a smaller difference of the stents for calcification patterns with leaflet calcification, which may indicated that stent 2 possess more excellent performance for less serious calcification patient.

4.2 Stress on calcification and aortic wall

The peak value of stress of calcification plaque and aortic wall located on the plane near plane 3 for all stents and calcification patterns. Location of maximum stent deformation index was close to this plane. This may be the reason of the peak stress.

Stent 2 had the minimum stress of calcification plaque and aortic wall. Rigid outflow end of stent 1 led to a higher deformation index and higher stress for calcification plaque and aortic wall. Stent 3 had the maximum peak value of stress. This may be the reason that the excessively sparse outflow end of stent 3 made the support force obviously decrease, and lead to a higher deformation, eventually result in a higher stress than other two stents.

Higher stress would make the stent easily to anchor on the aortic annulus [9]. And made the stent firmly contact with the aortic wall, then result in a smaller gap between the stent and aortic wall, eventually lead to lower paravalvular regurgitation. However, a higher stress will lead to a risk for aortic fracture and calcification plaque break. A breaken calcification plaque would lead to a risk of stroke. There was less report about stent immigration for self-expand stent, but the incidence was about 1.5-6%% for stroke [23]. It was better when there was a stress lower stress, but a smaller gap between the stent and aortic wall. So stent 2 was the best choice with a lower stroke and lower paravalvular regurgitation.

4.3 Paravalvular Gaps

The gap between stent and aortic wall could be used to predict the paravalvular regurgitation [9-11]. The influence of ascend aortic angle and paravalvular regurgitation had been investigated [24].

In this study, the paravalvular gap of right coronary sinus was the maximum for all the calcification patterns and stents. This may be the reason that the orient of the bent ascend aortic made the location has the lowest interact between the stent and aortic, induced the maximum gaps. Stent 1 had the maximum total gaps. Stent 2 and stent3 had the similar, but lower total gaps. However, stent 3 was a little higher than stent 2.

The location of coronary sinus for stent 1 had the obviously lager gaps than the other two, which led to a total larger gaps than others. This indicated there would be a seriously paravalvular regurgitation after implantation for stent 1. Rigid outflow end of stent 1 caused the stent not be compliance with the bend ascend aortic, result in larger gaps. As discussed in above, the deformation index of stent 2 and 3 on plane 2 was mainly decided by calcification, higher deformation index of stent 3 resulted in a larger gap than stent 2.

There was similar regularity for radial pattern. However, there is a small difference among the three stents. Calcification plaque located on plane 2 caused a critical effect to the gap for radial pattern and circle pattern. Therefore, radial pattern lead to a larger gap than coaptation pattern and circle pattern. The gap of circle pattern was lower than radial pattern, but lager

than coaptation pattern. This maybe the calcification plaque on the leaflet decreased the effect of the calcification plaque which located along the attachment line.

Stent 2 had the minimum gap for all location and had the minimum total gaps. Stent 2 also had a smaller gap when the calcification patterns had a little calcification volume, such as coaptation patterns. All investigations indirectly clarified why the new generation transcatheter aortic valve had a more sparseness outflow end than the first generation, such as Corevalve Elout R and Portico.

5 Conclusions

In this study, it illuminated that an excessively dense or a sparse outflow end would result in a larger stent deformation indexes, a higher stress calcified plaques and aortic root, and a lager gap, and led to serious paravalvular regurgitation.

Overall, stent 2 had the most excellent performance compared to other two stents. Stent 2 had morderate sparse outflow end. This made the stent easily be compliance with bend ascend aortic, and had the enough support force. All of these made the stent 2 had the identical deformation, lowest calcification plaque stress and smallest gaps. Especially, stent 2 had a more excellent performance when there was a little calcification, or there was a calcification plaque on the leaflet. This indicated stent 2 was more suitable for lower risk patient.

This study provides guidance for design of transcather aortic valve, and help to predict the clinical outcome.

Acknowledgement

The work was supported by the National Natural Science Foundation of China (11421202) and National key research and development program in China (2016YFC1102202) and the 111 Project (B13003).

References

[1] Alain Cribier, Helene Eltchaninoff, Assaf Bash, et,al; Percutaneous Transcatheter Implantation of an Aortic Valve Prosthesis for Calcific Aortic Stenosis First Human Case Description. Circulation 2002; 106:3006-3008.

[2] Leon MB, Smith CR, Mack M, et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. N Engl J Med 2010; 363:1597– 607.

[3] Smith CR, Leon MB, Mack MJ, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. N Engl J Med 2011; 364:2187–98.

[4] Michael Gotzmann, Michael Lindstaedt, and Andreas Mügge; From pressure overload to volume overload: Aortic regurgitation after transcatheter aortic valve implantation; Am Heart J 2012;163:903-11

[5] Mariuca Vasa-Nicotera, Jan-Malte Sinning, Derek Chin; et al. Impact of Paravalvular Leakage on Outcome in Patients After Transcatheter Aortic Valve Implantation. JACC: Cardiovascular interventions. VOL. 5, NO. 8,2012.

[6] Stefan Toggweiler, Karin H. Humphries, May Lee, et,al., 5-Year Outcome After Transcatheter Aortic Valve Implantation. J Am Coll Cardiol; Vol. 61, No. 4, 2013.

[7] Ganesh Athappan, Eshan Patvardhan, Murat Tuzcu, et.al, Incidence, Predictors, and Outcomes of Aortic Regurgitation After Transcatheter Aortic Valve Replacement, J Am Coll Cardiol, Vol. 61, No. 15, 2013.

[8] Vy P, Auffret V, Badel P, et al. Review of patient-specific simulations of transcatheter aortic valve implantation. Int J Adv Eng Sci Appl Math, 2015, 8(1): 2-24.

[9] Morganti S, Conti M, Aiello M, et al. Simulation of transcatheter aortic valve implantation through patient-specific finite element analysis: two clinical cases J Biomech, 2014, 47(11): 2547-2555.

[10] Auricchio F, Conti M, Morganti S, et al. Simulation of transcatheter aortic valve implantation: a patient-specific finite element approach Comput Methods Biomech Biomed Engin, 2014, 17(12): 1347-1357.

[11] Sturla F, Ronzoni M, Vitali M, et al. Impact of different aortic valve calcification patterns on the outcome of transcatheter aortic valve implantation: A finite element study. J Biomech, 2016, 49(12): 2520-2530.

[12] Wang Q, Kodali S, Primiano C, et al. Simulations of transcatheter aortic valve implantation: implications for aortic root rupture [J]. Biomech Model Mechanobiol, 2015, 14(1): 29-38.

[13] Morganti S, Brambilla N, Petronio AS, et al. Prediction of patient-specific post-operative outcomes of TAVI procedure: The impact of the positioning strategy on valve performance . J Biomech, 2016, 49(12): 2513-2519.
[14] Russ C, Hopf R, Hirsch S, et al. Simulation of transcatheter aortic valve implantation under consideration of leaflet calcification [C]. Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. IEEE, 2013: 711-714.
[15] Bart Bosmans, NeleFamaey, EvaVerhoelst,et.al., A validated methodology for patient specific computational modeling of self-expandable transcatheter aortic valve implantation. J Biomech 2016,49 (13) 2824–2830
[16] Tzamtzis S, Viquerat J, Yap J, et al. Numerical analysis of the radial force produced by the Medtronic-CoreValve and Edwards-SAPIEN after transcatheter aortic valve implantation (TAVI) . Med Eng Phys, 2013, 35(1): 125-130.
[17] Conti CA, Votta E, Della Corte A, et al. Dynamic finite element analysis of the aortic root from MRI-derived parameters. Med Eng Phys, 2010, 32(2): 212-221.
[18] Haj-Ali R, Marom G, Ben Zekry S, et al. A general three-dimensional parametric geometry of the native aortic valve and root for biomechanical modeling. J Biomech, 2012, 45(14): 2392-2397.
[19] Halevi R, Hamdan A, Marom G, et al. Progressive aortic valve calcification: three-dimensional visualization and biomechanical analysis. J Biomech, 2015, 48(3): 489-497.
[20] Zegdi R, Ciobotaru V, Noghin M, et al. Is it reasonable to treat all calcified stenotic aortic valves with a valved stent? Results from a human anatomic study in adults..J Am Coll Cardiol, 2008, 51(5): 579-584.
[21] Gunning PS, Vaughan TJ, McNamara LM. Simulation of self expanding transcatheter aortic valve in a realistic aortic root: implications of deployment geometry on leaflet deformation . Ann Biomed Eng, 2014, 42(9): 1989-2001.
[22] Sun W, Li K, Sirois E. Simulated elliptical bioprosthetic valve deformation: implications for asymmetric transcatheter valve deployment [J]. J Biomech, 2010, 43(16): 3085-3090.
[23] Daneault B, Kirtane AJ, Kodali SK, et al. Stroke associated with surgical and transcatheter treatment of aortic stenosis: a comprehensive review [J]. J Am Coll Cardiol, 2011, 58(21): 2143-2150.
[24] Sherif MA, Abdel-Wahab M, Stocker B, et al. Anatomic and procedural predictors of paravalvular aortic regurgitation after implantation of the Medtronic CoreValve bioprosthesis . J Am Coll Cardiol, 2010, 56(20):1623-1629.

A Two-Phase Flow Model for Aerogel in a Non-Equilibrium Process

†Dia Zeidan[1], Eric Goncalves[2] and Lucy T. Zhang[3]

[1]School of Basic Sciences and Humanities, German Jordanian University, Amman, Jordan
[2]ISAE-ENSMA, Institut Pprime, UPR 3346 CNRS, Poitiers, France
[3]Department of Mechanical, Aerospace and Nuclear Engineering, Rensselaer Polytechnic Institute, Troy, USA

†Corresponding and presenting author: dia.zeidan@gju.edu.jo

ABSTRACT

This paper presents the extension of hyperbolic and conservative two-phase flow model to a mixture of porous media containing nanofluids which is known as an aerogel. We focus mainly on the application of non-equilibrium mixture behaviour between phases in one space dimension. The governing equations are solved by finite volume techniques using Godunov methods of centred-type. Special emphasis is given to important and unsteady non-linear phenomena such as shock propagation in low densities aerogels. Simulation results are compared with other methods providing a remarkable agreement. Results show the good capabilities of this mixture formulation in the resolution of discontinuities in aerogel problems. This provides some insights into the fundamental properties of aerogels and helps to better understand some of the inherent difficulties in quantifying them using two-phase flow processes.

Keywords: Multiphase flows, Non-equilibrium, Nanofluids, Porous media, System of conservation laws, Riemann solution, Numerical simulation

Introduction

Aerogel is very common in industrial processes and its applications are already in use within the energy efficiency and pharmaceutical industries. In the meantime, many aerogels rely on light porous materials for which it may be densified into silica glasses by thermal processes (see, for example, [1, 2, 3, 16, 21]). During the last century, the nuclear and aerospace industries pushed strong research activity on the area. Their efforts have been aimed at the clarification of the mechanisms taking place during this complicated physical situation. In general, aerogels are solids with high void fractions, that is, porosities along with high surface areas, and possess very low densities and low conductivities. See, for example, [4, 5, 14] and references therein Typically, attention is given to the combination of high void fractions and very small pores to provide aerogels along with their extreme properties such as very low solid density and low thermal conductivity. These are mainly based on experimental analysis [6, 10, 11, 13]. However, there have been various gel-derived materials numerical simulations on the basis of simple models known as, for instance, reaction-limited cluster aggregation (RLCA) [18], diffusion-limited cluster aggregation (DLCA) [12] and diffusion-limited aggregation (DLA) [9]. Given such developments, it is of interest to investigate aerogels from mathematical and numerical point of views. Since aerogels belong to the family of nanoporous materials, it can be considered as a porous media containing nanofluids. Accordingly, one can use single or two-phase flow approach to study such phenomena. In the single-phase flow approach, the nanosolids can be simply fluidized where the relative velocity between phases is considered negligible. However, in the two-phase flow approach, this relative velocity may not be zero due to phase

interactions. In this regard, there exist a large number of articles concerning two-phase flow models where the nanofluids are treated as a simplified mixture of a base fluid and nanoparticles (see, for example, [7, 16, 20]). Most these articles are indicated that the two-phase mixture approach is more precise than single-phase approach. Further, these articles also have mainly examined the physical aspects of nanofluids rather than theoretical and numerical aspects of such models. As far it goes, only limited number of articles address the issues of the application of two-phase flow equations in nanofluids. These articles, however, used turbulent and laminar steady state two-phase flow models without taking care of the mathematical features of such models. This paper aims to study aerogels by considering a recently developed two-phase flow model (see, for example, [15, 19]). To this end, an unsteady system based on mixture formulations is presented. The aerogel is composed of nanofluids in which the dispersed phase is a gas. The governing partial differential equations (PDEs) are three mixture conservation of mass, momentum and energy. This is accomplished by three balance equations for gas void fraction, gas mass fraction and relative velocity between the gas and nanofluids. The single set of equations do not need any physical artificial stabilizing terms due to their conservativity and hyperbolicity features. Further, the deeply coupled equations are resolved on the basis of the Riemann problem using Godunov methods of centred type [17]. This solution becomes appropriate since it leads to making the relative velocity visible which is an advanced computational tool for aerogels. The model provides successful results at high and low phase velocities with low and high void fractions. It is found that the inclusion of nanofluids into the base fluid produce physically realistic solutions for strong relative motion between the nanofluids and gas phases. Results are compared with other available numerical methods producing accurate, efficient and free from numerical dissipation and dispersion computations. Simulation and test results show that the model equations can effectively simulate non-equilibrium aerogel, which may broaden the possible application areas of aerogel such as energy storage and biological tissues. The model equations and their numerical discretization for the non-equilibrium behaviour is given in the following section. Verification results are discussed in the section after followed by the conclusions section.

Governing Equations and Numerical Implementation

The aerogel considered here is composed of nanofluids in which the dispersed phase is a gas. The governing equations are formulated upon the conservation laws for mixture mass and mixture momentum along with a balance equation for the relative velocity between the two phase system in a single set of equations. For the investigation within the scope of this paper, the two-phase flow treatment is essentially a two-phase mixture model with isentropic conditions which significantly simplifies the mathematical formulation for the aerogel thereby reduces the computational costs. Within the context of mixture formulations, the total density, ρ, total and relative velocities, u and u_r, are discontinuous at the time-dependent interface with void fraction of the gas phase is always between 0 and 1. The time-dependent equations for mixture mass, mixture momentum and relative velocity are as follows:

$$\frac{\partial}{\partial t}(\rho) + \frac{\partial}{\partial x}(\rho u) = 0, \tag{1}$$

$$\frac{\partial}{\partial t}(\rho u) + \frac{\partial}{\partial x}(\rho u^2 + P + \rho c(1-c)u_r^2) = \mathcal{S}, \tag{2}$$

$$\frac{\partial}{\partial t}(u_r) + \frac{\partial}{\partial x}\left(u u_r + (1-2c)\frac{u_r^2}{2} + \psi(P)\right) = \pi, \tag{3}$$

where t is the time, x is the spatial coordinate, c is the gas mass void fraction, P represent the mixture pressure and $\psi(P)$ is a function that connect the different phases through the following relation

$$\psi(P) = e_g + \frac{P_g}{\rho_g} - e_{nf} - \frac{P_{nf}}{\rho_{nf}}. \tag{4}$$

The terms $\mathcal{S}$ and π in the mixture momentum and relative velocity equations approximate the interphase exchange processes. In addition to that, constituting relationships need to be added to system (1)-(3) so that one can predict the content of the aerogel. Closure laws account for the nanofluid physical properties and the thermodynamic behaviour of the two-phase system between gas and nanofluid phases. Under the interest that the two-phase system is considered as an aerogel, the density, velocity, relative velocity, pressure of the aerogel are calculated by the following mixture laws:

$$\begin{aligned} &\rho = \alpha_g\rho_g + (1-\alpha_g)\rho_{nf} \quad \text{and} \quad \rho u = \alpha_g\rho_g u_g + (1-\alpha_g)\rho_{nf}u_{nf}, \\ &P = \alpha_g P_g + (1-\alpha_g)P_{nf},\ c = \alpha_g\rho_g\,\rho^{-1} \quad \text{and} \quad u_r = u_g - u_{nf}, \end{aligned}$$

where subscripts (g) and (nf) denote gas and nanofluid, respectively. Further, the physical properties of the nanofluid are defined as follows:

$$\rho_{nf} = \alpha_s\rho_s + (1-\alpha_s)\rho_{bf} \quad \text{and} \quad u_{nf}\rho_{nf} = \alpha_s\rho_s u_s + (1-\alpha_s)\rho_{bf}u_{bf},$$

where indexes (s) and (bf) refer to the solid and base fluid, respectively, and the void fraction of solid and base fluid agree with $\alpha_s + \alpha_{bf} = 1$. For the system closure, the stiffened equation of state (EOS) is used for each phase as

$$P_j = K_j\left(\frac{\rho_j}{\bar{\rho}_j}\right)^{\gamma_j} - P_\infty.$$

The subscript $j = g$ is for the gas phase while $j = nf$ for the nanofluid phase and γ_j, K_j, P_∞ and $\bar{\rho}_j$ are constant parameters to be specified for each phase.
It should be noted that system (1)-(3) has been intensively studied and independent of the numerical methods being employed to resolve it. See, for example, [8, 15, 19], and the references therein. This is due to the fact that the governing equations of the system inherit conservative form as well as well-posedness forming an initial-boundary-value problem that describe different physical phenomena of interests. Furthermore, the aerogel equations (1), (2) and (3) are discretized herein by finite volume Godunov-type approach. Within such approach, the governing equations incorporate the resolution of the Riemann problem for computing the interface fluxes using the following time-marching formula [17]

$$\mathbb{U}_i^{n+1} = \mathbb{U}_i^n - \frac{\Delta t}{\Delta x}\left[\mathbb{F}_{i+\frac{1}{2}} - \mathbb{F}_{i-\frac{1}{2}}\right] + \Delta t\,\mathbb{S}_i, \tag{5}$$

which update the aerogel variables to a new time step. In (5), subscript (i) represents the cell index, the superscript (n) is the time level, Δx and Δt are the cell size and time step, respectively. $\mathbb{F}_{i\pm\frac{1}{2}}$ are the numerical fluxes through the left and right interfaces of cell (i) and $\mathbb{S}_i$ is the source terms that are evaluated at the cell centre. These fluxes are calculated by means of the solution of a Riemann problem with appropriate time-centred left and right input states. In the results shown below, Godunov methods of centred-type such as the second-order Slope Limiter

Centred (SLIC) scheme are employed in the numerical resolution of system (1)-(3). The SLIC scheme provides a high-resolution of large-gradient regions that are free from spurious oscillations. For the numerical background and details of the SLIC scheme, the readers may refer to [17].

Results and Discussion

The numerical simulations are performed by solving the governing equations (1)-(3) with two different test problems using finite volume approach within the aerogel. The first test problem, Test 1, involves the following initial data for the Riemann problem

$$\begin{aligned}
\left(\alpha_g, \rho_g, u_g\right)_L &= \left(0.5, 2.0, -0.1\right) \\
\left(\alpha_s, \rho_s, u_s, \rho_{bf}, u_{bf}\right)_L &= \left(0.7, 8.0, -2.0, 0.1, -2.0\right) \\
\text{if} \quad x &\leq 0, \\
\left(\alpha_g, \rho_g, u_g\right)_R &= \left(0.5, 2.0, 0.1\right) \\
\left(\alpha_s, \rho_s, u_s, \rho_{bf}, u_{bf}\right)_R &= \left(0.7, 8.0, 2.0, 0.1, 2.0\right) \\
\text{if} \quad x &> 0,
\end{aligned}$$

corresponding to rarefaction waves traveling in opposite direction separated by a contact discontinuity. In the second test problem, Test 2, we consider a collision of two symmetric shock waves, weak, and a trivial contact discontinuity with the following initial data

$$\begin{aligned}
\left(\alpha_g, \rho_g, u_g\right)_L &= \left(0.01, 3.0, 0.1\right) \\
\left(\alpha_s, \rho_s, u_s, \rho_{bf}, u_{bf}\right)_L &= \left(0.0001, 7.0, 0.4, 2.0, 0.4\right) \\
\text{if} \quad x &\leq 0, \\
\left(\alpha_g, \rho_g, u_g\right)_R &= \left(0.01, 3.0, -0.1\right) \\
\left(\alpha_s, \rho_s, u_s, \rho_{bf}, u_{bf}\right)_R &= \left(0.0001, 7.0, -0.4, 2.0, -0.4\right) \\
\text{if} \quad x &> 0.
\end{aligned}$$

The simulations results are shown in figures 1 and 2. In all results, it is not possible to solve the governing equations analytically due to the existence of several non-linear properties and closure relations involved for the current interest in aerogel. Thus, we produce a high-resolution numerical solution for the Riemann problem to calculate the reference solution on a very fine mesh of 10000 cells by using the total variation diminishing (TVD) slope limiter centre (SLIC) scheme in the computational domain $[-10, 10]$. Simulations are also carried out for the aerogel with transmissive boundary conditions along with a CFL number of 0.9 through the SUPERBEE limiter in the course of the SLIC scheme. The numerical resolutions (symbols) are compared with the reference solutions (solid lines) as well as with numerical solutions provided by other numerical methods available in the literature on a coarse mesh of 100 cells. In the test cases, $\gamma_g = 1.4$, $\bar{\rho}_g = 1.0$ kg/m^3, $K_g = 1.0$ Pa and $P_\infty = 0$, $\gamma_{nf} = 2.8$, $\bar{\rho}_{nf} = 1.0$ kg/m^3, $K_{nf} = 1.0$ Pa and $P_\infty = 1$ Pa. Results for Test 1 are shown in figure 1 for the mixture two-phase flow variables using three different numerical methods at a time $t = 0.8$ ms. The solution for this test problem consists of left and right rarefactions propagating in opposite direction. Figure 1 shows that the resolution for the relative velocity jump across the middle wave with lower densities whereas the mixture density, mixture velocity and mixture pressure are not for rarefaction

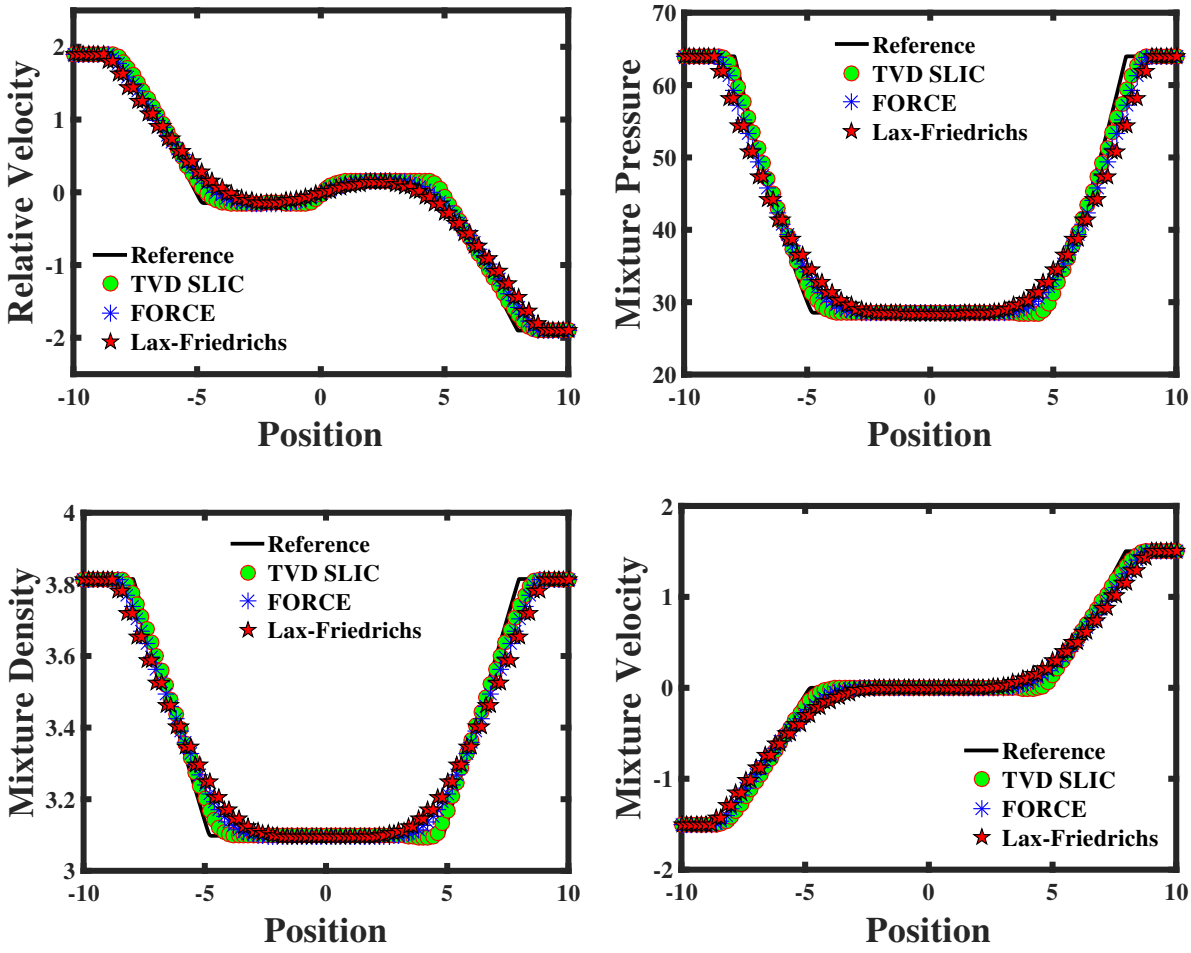


Figure 1: Test 1: Expansion tube problem in an aerogel at time $t = 0.8$ ms. The TVD SLIC, first-order centered (FORCE) and Lax-Friedrichs methods are compared with the reference solution results. Coarse meshes, symbols, are provided on 100 cells and very fine meshes of 10000 cells for the solid lines. The waves seen from left to right, a fast left rarefaction, a contact discontinuity and a fast right rarefaction.

waves propagating in an aerogel. The relative velocity jump indicates the possible sudden jump of gas and nanofluid velocities across the middle wave. It is also noted that a good agreement with the reference solutions for all the three methods and the model is able to deal with low density test case producing rarefaction waves. In figure 2, the results from a collision with an aerogel of low phase densities are displayed at a time $t = 2.3$ ms. The solution for the aerogel contains a left shock wave, a contact discontinuity and a right shock wave. It is observed that the mixture flow variables remain constant across the middle wave, however, the relative velocity jump discontinuously as in Test 1. This leads to the fact that lower phase densities slowly increase the relative motion between the gas and nanofluid phases during the collision. Again, the numerical results compared favourably with the reference solutions and capable of producing oscillation-free profiles at discontinuities. We conclude that the present model eqautions and the associated methods can automatically treat aerogel as a two-phase flow system even with different physical situations.

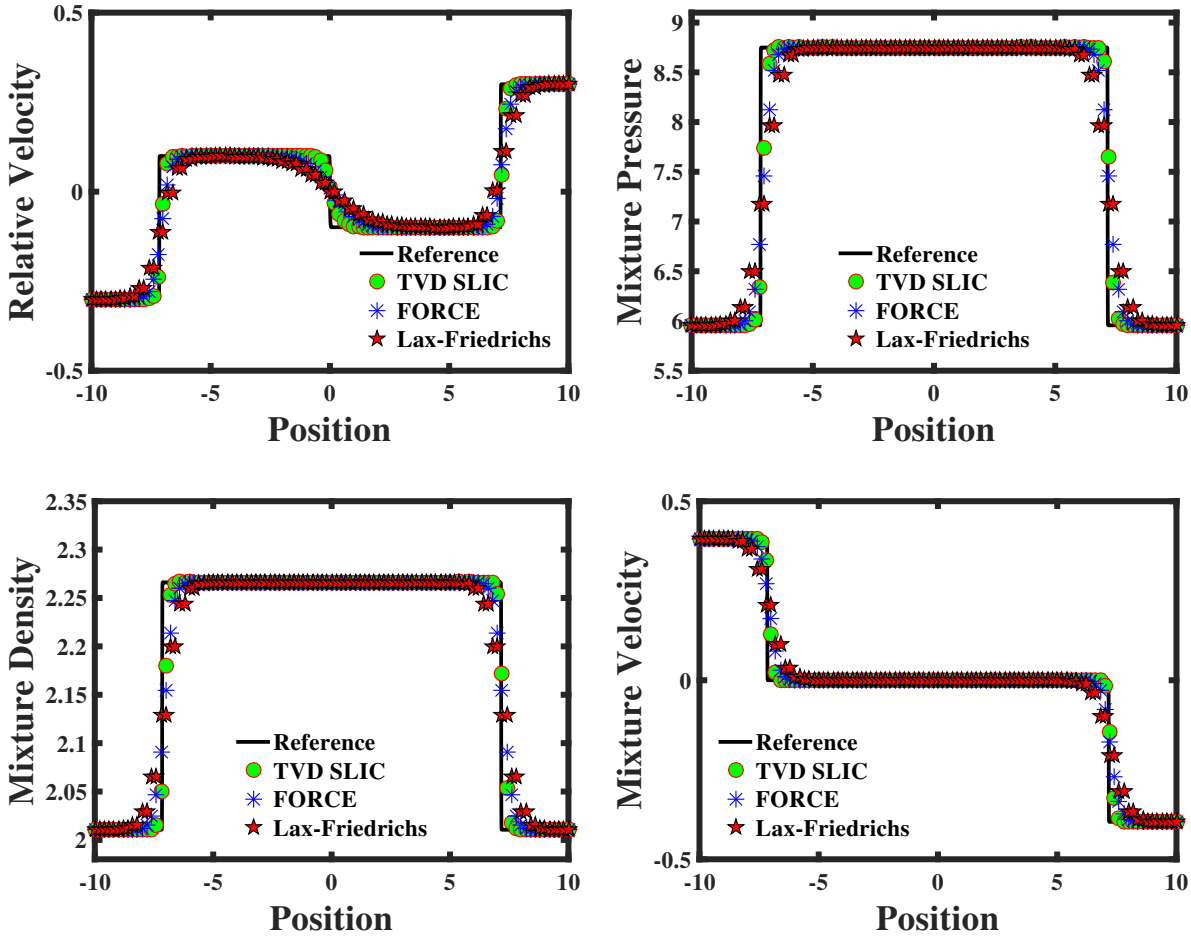


Figure 2: Test 2: The numerical results of the shock tube problem in an aerogel at time $t = 2.3$ ms. Comparison of the three different numerical methods using a coarse mesh of 100 cells (symbols) and the reference solution on a very fine mesh of 10000 cells (solid lines).

Concluding Remarks

A mixture two-phase flow model has been proposed for the simulation of porous media containing nanofluids used widely in aerogel production. This is motivated by the relative motion demanded to cope with the two-phase system arising in aerogels. Godunov methods with gradually growing levels of complexity in aerogels are employed to solve the governing equations in a one-dimensional domain representing the conceptual low-densities two-phase flow problems. This indicates that the extension of such mixture model to aerogel seems worthwhile as these type of hyperbolic problems are much better documented.

Acknowledgment

This work is supported by the German Jordanian University through the project SEED-SNRE 7-2014.

References

[1] Aegerter, M.A., Prassas, M. (2011) *Advances in Sol-Gel Derived Materials and Technologies*, Springer.

[2] Akimov, Y.K. (2003) Fields of application of aerogels (Review). *Instruments and Experimental Techniques* **46**, 287–299.

[3] Chaput, F., Lecomte, A., Dauger, A., Boilot, J.P. (1989) Preparation and structure of aluminosilicate aerogels. *Chem. Mater.* **1**, 199–201.

[4] Gurikov, P.A., Kolnoochenko, A.V., Menshutina, N.V. (2009) 3D reversible cellular automata for simulation of the drug release from aerogel-drug formulations. *Computer Aided Chemical Engineering* **26**, 943–947.

[5] Fricke, J. (1986) Thermal transport in porous superinsulations in aerogels. *Springer Proceedings in Physics* **6**, 94–103.

[6] Ibrahim, M., Biwole, P.H., Achard, P., Wurtz, E. (2015) Aerogel-based materials for improving the building envelopes thermal behavior: a brief review with a focus on a new aerogel-based rendering. *Energy Sustainability Through Green Energy*, 163–188.

[7] Kalteh, M., Abbassi, A., Saffar-Avval, M., Harting, J. (2011) Eulerian-Eulerian two-phase numerical simulation of nanofluid laminar forced convection in a microchannel. *International Journal of Heat and Fluid Flow* **32**, 107–116.

[8] La Spina, G., de'Michieli Vitturi, M., Clarke, A.B. (2017) Transient numerical model of magma ascent dynamics: application to the explosive eruptions at the Soufrière Hills Volcano. *Journal of Volcanology and Geothermal Research* **336**, 118–139.

[9] Ma, H.-S., Prévost, J.-H., Jullien, R., Scherer, G.W. (2001) Computer simulation of mechanical structureproperty relationship of aerogels. *J.Non-Cryst.Solids* **285**, 216–221.

[10] McNaught, A.D., Wilkinson, A. (2007) *Compendium of Chemical Terminology*, IUPAC Goldbook, PAC, 791801, Blackwell Science.

[11] Morales, R.V., da Cunha, C.R., Rambo, C.R. (2014) A complex network approach for the growth of aerogels. *Physica A: Statistical Mechanics and its Applications* **406**, 131–138.

[12] Paul, M. (1983) Formation of fractal clusters and networks by irreversible diffusion-limited aggregation. *Phys. Rev. Lett.* **51**, 1119–1122.

[13] Platzer, W.J., Bergkvist, M. (1993) Bulk and surface light scattering from transparent silica aerogel. *Solar Energy Mater Solar Cells* **31**, 243–251.

[14] Richter, K., Norris, P.M., Chang, C.L. (1995) Aerogels: applications, structure and heat transfer phenomena. *In Review on Heat Transfer, V. Prasad, Y. Jaluria, and G. Chen, Eds.* **6**, 61–114.

[15] Romenski, E., Resnyansky, A.D., Toro, E.F. (2007) Conservative hyperbolic formulation for compressible two-phase flow with different phase pressures and temperatures. *Quart. Appl. Math.* **65**, 259–279.

[16] Sehaqui, H., Zhou, Q., Berglund, L.A. (2011) High-porosity aerogels of high specific surface area prepared from nanofibrillated cellulose (NFC). *Composites Science and Technology* **71**, 1593–1599.

[17] Toro, E.F. (2009) *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*, Springer.

[18] Woignier, T., Phalippou, J., Pelous, J., Courtens, E. (1990) Different kinds of fractal structures in Si02 aerogels. *J. Non-Cryst. Solids* **121**, 198–204.

[19] Zeidan, D. (2016) Assessment of mixture two-phase flow equations for volcanic flows using Godunov-type methods. *Applied Mathematics and Computation* **272**, 707–719.

[20] Zhao, L., Yang, S., Bhatia, B., Strobach, E., Wang, E.N. (2016) Modeling silica aerogel optical performance by determining its radiative properties. *AIP Advances* **6**, 025123-1–8.

[21] Zhou, X., Torabi, M., Lu J., Shen, R., Zhang, K. (2014) Nanostructured energetic composites: synthesis, ignition/combustion modeling, and applications. *ACS Appl Mater Interfaces* **6**, 3058–74.

Behavior study of the generalized-strain mesh-free method (GSMF)

†Wilber Vélez[1], Tiago Oliveira[2], and Artur Portela[1]

[1,] Department of Civil and Environmental Engineering, University of Brasília, BR.

*Presenting author: wilbervelez@hotmail.com
†Corresponding author: aportela@unb.br

Abstract

The generalized-strain mesh-free (GSMF) local method, it is derived through a weighted-residual formulation that leads to the work theorem of structures theory. In a local region, the work theorem establishes an energy relationship between a statically – admissible stress field and an independent kinematically – admissible strain field.

In the formulation of the GSMF, the local form of work theorem is simply an integration – free formula. The Moving Least Squares (MLS) approximation of the elastic field is used to construct the trial function in this local meshless formulation. GSMF has a highly computational efficiency leading to accurate numerical results in two-dimensional elasticity problems.

This paper is concerned with the size effect of the configuration parameters of the local support domain and the local weighted-residual or quadrature/collocation domain on GSMF. A comparison of the error in energy and displacement is presented for five different regular nodal distribution, in order to solve the Timoshenko cantilever beam problem and the infinite plate with circular hole. The results are compared with the exact solution and optimal parameters have been determined.

Keywords: Generalized-Strain Mesh-Free (GSMF), work theorem, Moving Least Squares (MLS), local support domain, local weighted-residual domain.

Introduction

There are different numerical modeling techniques and the most popular and widely used is the Finite Element Method (FEM). Although recently meshless methods are becoming more used due to their accuracy and performance in numerical analysis. Different methods have been proposed [1-7], some of them are derived from a weak – form formulation on global domain and others from local sub – domains. The weighted – residual method is the basis for the meshless formulation [8].

The Generalized – Strain Mesh – free (GSMF) formulation presented by [9], expressed that the work theorem generates a weak form that is completely integration free, working as a weighted-residual weak – form collocation. This formulation has two important parameters associated to support and quadrature/collocation domain, this parameters have a greater influence in the problem solution obtained by Meshfree methods.

The size local support domain (α_s) and the weak – form domain or local quadrature/collocation domain (α_q) are very important meshless parameter, both related to accuracy and computational efficiency. A comparative study of the effect of these size parameters using the Meshless Local Petrov – Galerkin method (MLPG) for the solution of a cantilever beam was presented by [10] and, using the same method, a similar study for the cantilever beam and the infinite plate with circular hole was carried out by [11].

This paper presents a numerical comparison of the Generalized-strain Mesh-free (GSMF) formulation for the size effect of the configuration parameters of the local support domain and the local quadrature/collocation domain; for five different regular nodal distribution of the Timoshenko cantilever beam and one nodal distribution of the infinite plate with circular hole. The results obtained in this work, both for the energy and displacement, are compared with the exact solution presented for 2D problems plane stress case and important conclusions are presented in the end.

MLS Approximation

Let Ω be the domain of a body with boundary Γ and let $N = \{X_1, X_2, .., X_N\} \in \Omega$ be a set of scattered nodal points that represents a meshless discretization. Some of them are located on the boundary Γ where Ω_S, represented as Ω_P, Ω_Q and Ω_R, is the local compact support of a node X_t, represented as X_P, X_Q and X_R. Ω_x is the domain of definition of a sampling point X and Ω_q is the local weak-form domain or quadrature domain of node X_i, as represent in Fig. 1.

Circular or rectangular local supports centered at each nodal point can be used. In the region of a sampling point X, the domain of definition of MLS approximation is the subdomain Ω_x, where the approximation is defined.

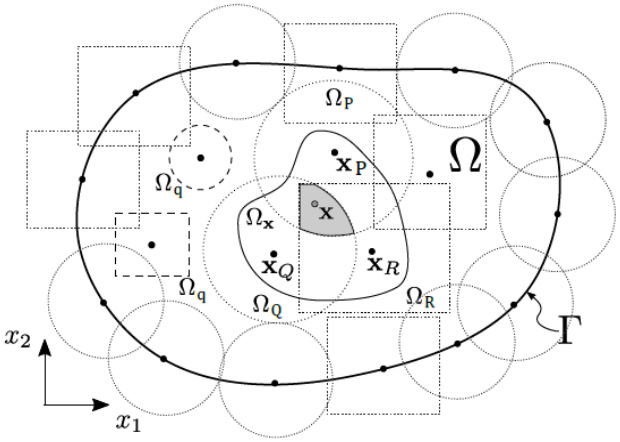


Figure 1. Representation of a global domain Ω and boundary Γ in a meshless discretization, with Xi nodes distributed within the body.

Shape Functions

Let Ω_x be the domain of definition of the MLS approximation, in a neighborhood of a sampling point x. To approximate the displacement $u(x) \in \Omega_x$, over a number of scattered nodes $X_i \in \Omega$, $i = 1, 2, .., n$, where the nodal parameters $\hat{u}_i$ are defined, the MLS approximation is given by

$$u^h(x) = \mathbf{p}^T(x)a(x), \tag{1}$$

for $x \in \Omega_x$, in which

$$\mathbf{p}^T(x) = \left[p_1(x), p_2(x), ..., p_m(x)\right], \tag{2}$$

is a vector of the complete monomial basis of order m and a(x) is the vector of unknown coefficients $a_j(x), j = 1, 2, ..., m$ that are functions of the space coordinates $x = \left[x_1, x_2\right]^T$, for 2-D problems.

The coefficient vector a(x) is determined by minimizing the weighted discrete L_2 norm

$$J(x)=\frac{1}{2}\sum_{i=1}^{n}w_i(x)\left[u^h(x_i)-\hat{u}_i\right]^2=\frac{1}{2}\sum_{i=1}^{n}w_i(x)\left[\mathbf{p}^T(x_i)a(x)-\hat{u}_i\right]^2, \tag{3}$$

with respect to each term of a(x), in which $w_i(x)$ is the weight function associated with the node x_i, with the compact support that is $w_i(x)>0$, for all x in the support of $w_i(x)$. In the Fig. 1 is represented the compact support of the MLS weight functions associated with a few nodes. Finding the extremum of *J(x)* with respect to each term of a(x), leads to

$$\mathbf{A}(x)\mathrm{a}(x)=\mathbf{B}(x)\hat{\mathbf{u}}, \tag{4}$$

where,

$$\mathbf{A}(x)=\sum_{i=1}^{n}w_i(x)\mathbf{p}(x_i)\mathbf{p}^T(x_i), \tag{5}$$

$$\mathbf{B}(x)=\left[w_1(x)\mathbf{p}(x_1),w_2(x)\mathbf{p}(x_2),...,w_n(x)\mathbf{p}(x_n)\right] \tag{6}$$

and

$$\hat{\mathbf{u}}=\left[\hat{u}_1,\hat{u}_2,...,\hat{u}_n\right]. \tag{7}$$

Solving equation (4) for a(x) yields

$$\mathrm{a}(x)=\mathbf{A}^{-1}(x)\mathbf{B}(x)\hat{\mathbf{u}}, \tag{8}$$

Provided $n \geq m$, for each some point x, condition defined in the MLS approximation.

$$u^h(x)=\sum_{i=1}^{n}\phi_i(x)\hat{u}_i, \tag{9}$$

in which

$$\phi_i(x)=\sum_{i=1}^{n}p_j(x)\left[\mathbf{A}^{-1}(x)\mathbf{B}(x)\right]_{ji} \tag{10}$$

The shape function of the MLS approximation for the node x_i is represented Figure 2.

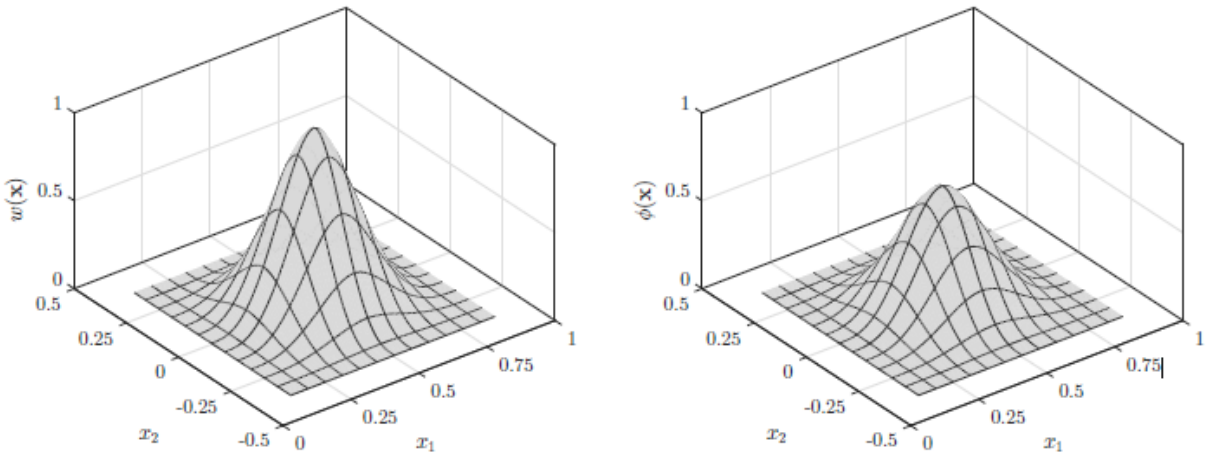


Figure 2. Respectively the typical weight function and shape function of the MLS approximation

The MLS shape functions are not nodal interpolants which means mathematically that Since $\phi_i(x_j)\neq\delta_{ij}$. Since $\phi_i(x)$ vanishes for *x* not in the local domain of the node x_i, is preserved the local

character of the MLS approximation. The nodal shape function is complete up to the order of the basis. The smoothness of the nodal shape function is determined by the smoothness of the basis and of the weight function. The spatial derivatives of the shape function $\phi_i(x)$ are giving by

$$\phi_{i,k}(x) = \sum_{j=1}^{m}\left[p_{j,k}\left(\mathbf{A}^{-1}\mathbf{B}\right)_{ji} + p_j\left(\mathbf{A}^{-1}\mathbf{B}_{,k} - \mathbf{A}^{-1}\mathbf{A},_k\,\mathbf{A}^{-1}B\right)_{ji}\right], \tag{11}$$

in which $()_{,k} = \partial() / \partial x_k$.

Weight Functions

In the Figure 2 the weight functions are represented, in Eq. (3) are introduced for each node x_i, have a compact support for all x, which defines the subdomain where $w_i(x) > 0$. This paper considers rectangular compact supports with weight functions defined as

$$w_i(x) = w_{i_x}(x) w_{i_y}(x) \tag{12}$$

the weight function is given by the quartic spline function

$$w_{i_x}(x) = \begin{cases} 1 - 6\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^2 + 8\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^3 - 3\left(\dfrac{d_{i_x}}{r_{i_x}}\right)^4 & for\ 0 \le d_{i_x} \le r_{i_x} \\ 0 & for\ d_{i_x} > r_{i_x} \end{cases} \tag{13}$$

And

$$w_{i_y}(x) = \begin{cases} 1 - 6\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^2 + 8\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^3 - 3\left(\dfrac{d_{i_y}}{r_{i_y}}\right)^4 & for\ 0 \le d_{i_y} \le r_{i_y} \\ 0 & for\ d_{i_y} > r_{i_y} \end{cases} \tag{14}$$

in which $d_{i_x} = \|x - x_i\|$ and $d_{i_y} = \|y - y_i\|$. The parameters r_{i_x} and r_{i_y} represent the size of the support for the node i, respectively in the x and y directions.

Elastic Field

The elastic field is approximated at a sampling point x. Considering Eq. (9) the displacement and strain components

$$\mathbf{u} = \begin{bmatrix} u^h(x) \\ v^h(x) \end{bmatrix} = \begin{bmatrix} \phi_1(x) & 0 & \dots & \phi_n(x) & 0 \\ 0 & \phi_1(x) & \dots & 0 & \phi_n(x) \end{bmatrix} \begin{bmatrix} \hat{u}_1 \\ \hat{v}_1 \\ \vdots \\ \hat{u}_n \\ \hat{v}_n \end{bmatrix} = \mathbf{\Phi}\hat{\mathbf{u}} \tag{15}$$

$$\varepsilon = \mathbf{L}\mathbf{u} = \mathbf{L}\mathbf{\Phi}\hat{\mathbf{u}} = \mathbf{B}\hat{\mathbf{u}} \tag{16}$$

in which geometrical linearity is assumed in the differential operator **L** and thus,

$$\mathbf{B}=\begin{bmatrix} \phi_{1,1} & 0 & \cdots & \phi_{n,1} & 0 \\ 0 & \phi_{1,2} & \cdots & 0 & \phi_{n,2} \\ \phi_{2,1} & \phi_{1,1} & \cdots & \phi_{n,2} & \phi_{n,1} \end{bmatrix} \tag{17}$$

Stress and traction components are

$$\boldsymbol{\sigma}=\mathbf{D}\varepsilon=\mathbf{DB}\hat{\mathbf{u}} \tag{18}$$

and

$$\mathbf{t}=\mathbf{n}\boldsymbol{\sigma}=\mathbf{nDB}\hat{\mathbf{u}} \tag{19}$$

in which **D** is the matrix of the elastic constants and **n** is the matrix of the components of the unit outward normal, defined as

$$\mathbf{n}=\begin{bmatrix} n_1 & 0 & n_2 \\ 0 & n_2 & n_1 \end{bmatrix} \tag{20}$$

Eq. (15) to (19) show that the variables of the elastic field are defined in terms of the nodal unknowns $\hat{u}$, for all point x ϵ Ω_x.

Local Form of the Work Theorem

The development for the local form of the work theorem introduced in [12] is presented in this section. Let Ω be the domain of a body and Γ its boundary, subdivided in Γ_u and Γ_t that is $\Gamma=\Gamma_u\cup\Gamma_t$; in the Fig. 3 show the nodal points *P, Q* and *R* have corresponding local domains Ω_P, Ω_Q, and Ω_R.

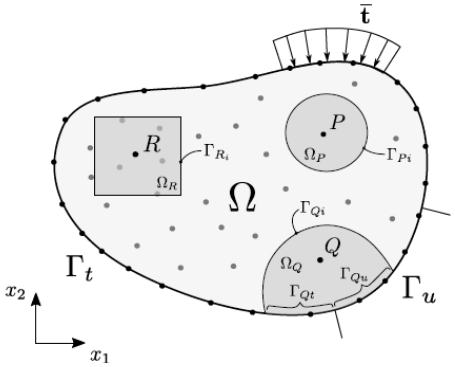


Figure 3. Meshless discretization of the global domain Ω and the local domains Ω_P, Ω_Q and Ω_R, with boundary $\Gamma = \Gamma_u \cup \Gamma_t$ represented.

The mixed fundamental boundary value problem of linear electrostatics aims to determine the distribution of stress σ, strains ε and displacements u throughout the body, when it has constrained displacements $\bar{u}$ defined on Γ_u and is loaded by an external system of distributed surface and body forces with densities denoted by $\bar{\mathbf{t}}$ on Γ_t and b in Ω, respectively.

The entire admissible elastic field is the solution of the posed problem that simultaneously satisfies the kinematic admissibility and the static admissibility. If this solution exists, it can be shown that it is unique, provided linearity and stability of the material are admitted [12, 13].

The general work theorem establishes an energy relationship between any statically – admissible stress field and any kinematically – admissible strain field that can be defined in the body. Derived as a weighted residual statement, the work theorem serves as a unifying basis for the formulation of numerical models Continuum Mechanics [14].

In the domain of the body, consider a statically-admissible stress field that is

$$\mathbf{L}^T\sigma + b = 0 \tag{21}$$

in the domain Ω, with boundary conditions

$$t = n\sigma = \bar{t} \tag{22}$$

on the static boundary Γ_t, σ: Stress components, L: Matrix differential operator, t: represent the traction components, $\bar{\mathbf{t}}$: Prescribed tractions values, n: Unit normal components to the boundary.

In the global domain Ω, consider an arbitrary local subdomain Ω_Q, centered at the point Q, with boundary $\Gamma = \Gamma_{Qi} \cup \Gamma_{Qt} \cup \Gamma_{Qu}$, Γ_{Qi}: Interior local boundary and Γ_{Qt} and Γ_{Qu}: Local boundaries that respectively share a global boundary. Due to its arbitrariness, this local domain can be overlapping with other similar subdomains. For the local domain Ω_Q the strong form of the weighted-residual equation is written as

$$\int_{\Omega_Q} \left(\mathbf{L}^T\sigma + b\right)^T \mathbf{W}_\Omega d\Omega + \int_{\Gamma_{Qt}} (t - \bar{t})^T \mathbf{W}_\Gamma d\Gamma = 0 \tag{23}$$

in which W_Ω and W_Γ are arbitrary weighting functions defined, respectively in Ω and on Γ. When the domain term of Eq. (23) is integrated by parts, the following local weak form of the weighted residual equation is obtained

$$\int_{\Gamma_Q} (\mathbf{n}\boldsymbol{\sigma})^T \mathbf{W}_\Omega d\Gamma - \int_{\Omega_Q} (\boldsymbol{\sigma}^T \mathbf{L}\mathbf{W}_\Omega - \mathbf{b}^T \mathbf{W}_\Omega) d\Omega + \int_{\Gamma_{Qt}} (\mathbf{t} - \bar{\mathbf{t}})^T \mathbf{W}_\Gamma d\Gamma = 0 \tag{24}$$

which now requires continuity of $\mathbf{W}_\Omega$, as an admissibility condition for integrability. For the sake of convenience, the arbitrary weighting function $\mathbf{W}_\Gamma$ is chosen as

$$\mathbf{W}_\Gamma = -\mathbf{W}_\Omega \tag{25}$$

on the boundary Γ_{Qt}. Thus, Eq. (24) leads to

$$\int_{\Gamma_Q} (\mathbf{n}\boldsymbol{\sigma})^T \mathbf{W}_\Omega d\Gamma - \int_{\Omega_Q} (\boldsymbol{\sigma}^T \mathbf{L}\mathbf{W}_\Omega - \mathbf{b}^T \mathbf{W}_\Omega) d\Omega + \int_{\Gamma_{Qt}} (\mathbf{t} - \bar{\mathbf{t}})^T \mathbf{W}_\Gamma d\Gamma = 0 \tag{26}$$

Consider further an arbitrary kinematically-admissible strain field ε^*, with continuous displacements u* and small derivatives, in order to assume geometrical linearity, defined in the global domain that is

$$\varepsilon^* = \mathbf{L}\mathbf{u}^* \tag{27}$$

in the domain Ω, with boundary conditions

$$\mathbf{u}^* = \bar{\mathbf{u}} \tag{28}$$

on the kinematic boundary Γ_u.

When the continuous arbitrary weighting function $\mathbf{W}_\Omega$, is defined as

$$\mathbf{W}_\Omega = \mathbf{u}^* \tag{29}$$

the weak form (26), of the weighted residual equation, becomes

$$\int_{\Gamma_Q-\Gamma_{Qt}-\Gamma_{Qu}} \mathbf{t}^T\mathbf{u}^* d\Gamma + \int_{\Gamma_{Qu}} \mathbf{t}^T\bar{\mathbf{u}}^* d\Gamma + \int_{\Gamma_{Qt}} \bar{\mathbf{t}}^T\mathbf{u}^* d\Gamma - \int_{\Omega_Q} \left(\boldsymbol{\sigma}^T\mathbf{L}\mathbf{u}^* - \mathbf{b}^T\mathbf{u}^*\right) d\Omega = 0 \tag{30}$$

which can be written in a compact form as

$$\int_{\Gamma_Q} \mathbf{t}^T\mathbf{u}^* d\Gamma + \int_{\Omega_Q} \mathbf{b}^T\mathbf{u}^* d\Omega = \int_{\Omega_Q} \boldsymbol{\sigma}^T\varepsilon^* d\Omega \tag{31}$$

This equation is the starting point of kinematically admissible formulations of the local mesh-free method presented in this paper. Equation (31) which expresses the static-kinematic duality, is the local form of the well-known work theorem, the fundamental identity of solid mechanics [16].

It is the important to notice that the stress field $\boldsymbol{\sigma}$, is any one that satisfies equilibrium with the applied external forces **b** and **t**, which is not necessarily the stress field that actually settles in the body. Also, the strain field ε^*, is any one that is compatible with the constrains $\mathbf{u}^* = \bar{\mathbf{u}}$, which is not necessarily the strain field that actually settles in the body. These two fields are not connected function $\mathbf{W}_\Omega$ they are completely independent. For that reason, Eq. (31) can be used under the only assumption of geometrical linearity. It is the independence of the two admissible fields of the Eq. (31) that allows generation of different meshfree methods, when the strain field is locally defined through different options, as carried out in this paper.

A final important remark, worth of mentioning, is that the local domain Ω_Q, is any arbitrary subdomain Ω, of the body.

Modeling Strategy

In the local meshfree methods different formulations can be derived when the arbitrary kinematically – admissible field ε^*, is locally defined in the work theorem, Eq. (31). In the following section, simple kinematically – admissible local fields will be used to derive the meshless formulation presented in this paper, the Generalized-Strain Mesh-Free (GSMF) formulation. On the other hand, the statically – admissible local field $\boldsymbol{\sigma}$, will be always assumed as the elastic field that actually settles in the body. Not only satisfying static admissibility, through Eq. (21) and (22), but also satisfying kinematic admissibility in the elastic field defined as

$$\varepsilon = \mathbf{L}\mathbf{u} \tag{32}$$

In the domain Ω, with boundary conditions

$$\mathbf{u} = \bar{\mathbf{u}} \tag{33}$$

on the kinematic boundary Γu; in the which the displacements u, are assumed continuous with small derivatives, in order to allow for geometrical linearity of the strain field ε. Therefore, Eq. (33) must be enforced in the numerical model, in order to provide a unique solution of the posed problem.

For a meshless discretization of the body, the local weak-form domain or quadrature domain Ω_Q, centered at the node *Q*, can be defined in this paper as a rectangular or circular subdomain, as represented in Fig. 3.

Generalized-Strain Formulation

The Equation (31), the kinematically – admissible displacement field u*, was assumed as a continuous function leading to a regular integral function that is the kinematically-admissible strain field ε^*. For more information about the method, see [9].

However, this continuity assumption on u*, enforced in the local form of the work theorem, is not required but can be relaxed by convenience, provided ε^* can be useful as a generalized function, in the sense of the theory of distributions [15]. Hence, this formulation considers that the kinematically-admissible displacement field is a piece-wise continuous function, defined in terms of the Heaviside step function and therefore the corresponding kinematically-admissible strain field is a generalized function, defined in terms of Dirac delta function.

For the sake of the simplicity, in dealing with Heaviside and Dirac delta functions in a two-dimensional coordinate space, consider a scalar function d, defined as

$$d = \left\| \mathrm{x} - \mathrm{x}_Q \right\| \quad \text{that is} \quad \begin{cases} d = 0 & \text{if} \quad \mathrm{x} \equiv \mathrm{x}_Q \\ d > 0 & \text{if} \quad \mathrm{x} \neq \mathrm{x}_Q \end{cases} \tag{34}$$

which represents the absolute-value function of the distance between a field point x and a particular reference point x_Q, in the local domain $\Omega_Q \bigcup \Gamma_Q$ assigned to the field node Q. Therefore, this definition always assumes $d = d(\mathrm{x},\mathrm{x}_\mathrm{Q}) \geq 0$, as a positive or null vale, in this case when-ever x and x_Q are coincident points. It is important to remark that, in Eq. (34), neither the field point x nor the reference point x_Q is necessarily a nodal point of the local domain.

For a scalar coordinate, $d \supset d(\mathrm{x},\mathrm{x}_\mathrm{Q})$, the Heavise step function can be defined as

$$H(d) = \begin{cases} 1 \text{ if } d \leq 0 \quad (d = 0 \text{ for } \mathrm{x} \equiv \mathrm{x}_\mathrm{Q}) \\ 0 \text{ if } d > 0 \qquad \text{that is } \mathrm{x} \neq \mathrm{x}_\mathrm{Q} \end{cases} \tag{35}$$

in which the discontinuity is assumed at x_Q and consequently, the Dirac delta function is defined with the following properties

$$\delta(d) = H'(d) = \begin{cases} \infty \text{ if } d = 0 \qquad \text{that is } \mathrm{x} \equiv \mathrm{x}_\mathrm{Q} \\ 0 \text{ if } d \neq 0 \quad (d > 0 \text{ for } \mathrm{x} \neq \mathrm{x}_\mathrm{Q}) \end{cases} \quad \text{and} \quad \int_{-\infty}^{+\infty} \delta(d)\mathrm{d}d = 1 \tag{36}$$

in which H'(d) represents the distributional derivate of H(d). Note that the derivate of H(d), with respect to the coordinate x_i, can be defined as

$$H(d)_{,i} = H'(d)d_{,i} = \delta(d)d_{,i} = \delta(d)n_i \tag{37}$$

Since the result of this equation is not affected by any particular value of the constant n_i, this constant will be conveniently redefined later on.

Kronecker delta function can be defined through Heaviside step function as

$$\Delta(d) = H(d) + H(-d) - 1 = \begin{cases} 1 \text{ if } d = 0 \quad \text{that is } \mathrm{x} \equiv \mathrm{x}_\mathrm{Q} \\ 0 \text{ if } d > 0 \quad \text{that is } \mathrm{x} \neq \mathrm{x}_\mathrm{Q} \end{cases} \tag{38}$$

Which has the distributional derivative always null that is

$$\Delta'(d) = \delta(d) - \delta(-d) = \delta(d) - \delta(d) = 0 \tag{39}$$

as a consequence of the symmetry of Dirac delta function.

Now considerer that d_l, d_j and d_k represent the distance function d, defined in Eq. (34), for corresponding field points x_l, x_j and x_k. Then, the kinematically – admissible displacement field can be defined as a linear combination of Kronecker delta function evaluations at an arbitrary number of collocation points, conveniently arranged in the local domain $\Omega_Q \bigcup \Gamma_Q$ of the field node Q, that is

$$\mathrm{u}^*(\mathrm{x}) = \left[\frac{L_i}{n_i} \sum_{l=1}^{n_i} \Delta(d_l) + \frac{L_t}{n_t} \sum_{j=1}^{n_t} \Delta(d_j) + \sum_{k=1}^{n_\Omega} \Delta(d_k) \right] e \tag{40}$$

in which $\mathrm{e} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$ represents the metric of the orthogonal directions; n_i, n_t and n_Ω represent the number of collocation points, respectively on the local interior boundary $\Gamma_{Q_i} = \Gamma_Q - \Gamma_{Q_t} - \Gamma_{Q_u}$ with length L_i, on the local static boundary Γ_{Q_t} with length L_t and in the local domain Ω_Q with area S. This assumed displacement field u*(x), a discrete rigid-body unit displacement defined at collocation points, schematically represent in Fig. 4, conveniently leads to a null

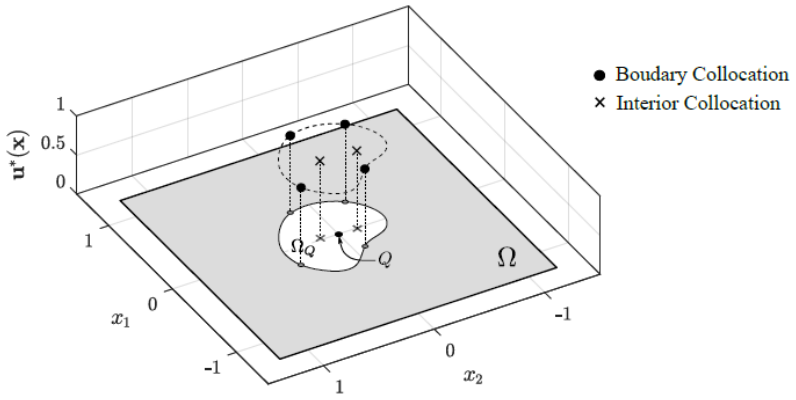


Figure 4. Schematic representation of the displacement u*(x) **of Eq. (38), a discrete rigid-body unit displacement defined at collocation points, of the Generalized-Strain Mesh-free formulation, for a local domain associated with a field node** *Q*.

generalized strain field that is

$$\varepsilon^*(\mathrm{x}) = 0 \tag{41}$$

as a consequence of Eq. (39). The local work theorem, Eq. (31), can be written as

$$\int_{\Gamma_Q - \Gamma_{Qt}} \mathbf{t}^T \mathbf{u}^* d\Gamma + \int_{\Gamma_{Qt}} \overline{\mathbf{t}}^T \mathbf{u}^* d\Gamma + \int_{\Omega_Q} \mathbf{b}^T \mathbf{u}^* d\Omega = \int_{\Omega_Q} \boldsymbol{\sigma}^T \varepsilon^* d\Omega \tag{42}$$

which, after considering the assumed displacement and the strain components of the kinematically-admissible field, respectively Eq. (40) and (41), leads to

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}\int_{\Gamma_Q-\Gamma_{Qt}}\mathbf{t}^T\Delta(d_l)\mathrm{e}\mathrm{d}\Gamma+\frac{L_t}{n_t}\sum_{j=1}^{n_t}\int_{\Gamma_{Qt}}\bar{\mathbf{t}}^T\Delta(d_j)\mathrm{e}\mathrm{d}\Gamma+\frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\int_{\Omega_Q}\mathbf{b}^T\Delta(d_k)\mathrm{d}\Omega=0 \tag{43}$$

Now considering the properties of Kronecker delta function, defined in Eq. (38), the Eq. (43) simply leads to

$$\mathrm{e}^T\left[\frac{L_i}{n_i}\sum_{l=1}^{n_i}t_{\mathrm{x}_l}+\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{\mathrm{x}_j}+\frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathrm{x}_k}\right]=0 \tag{44}$$

and finally to

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}t_{\mathrm{x}_l}=-\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{\mathrm{x}_j}-\frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathrm{x}_k} \tag{45}$$

Equation (45) states the equilibrium of tractions and body forces, pointwisely defined at collocation points, as schematically represented in Fig. 5; obviously, the pointwise version of the Euler – Cauchy stress principle.

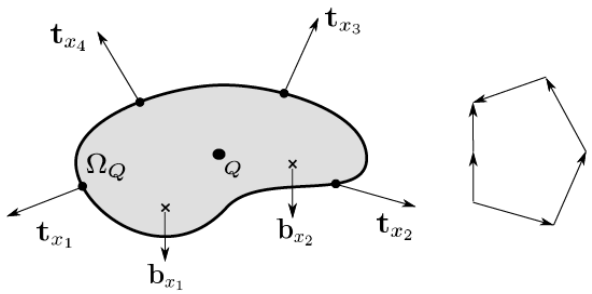


Figure 5. Schematic representation of the equilibrium of tractions and body forces of Eq. (44), pointwisely defined at collocation points of a local domain associated with a field node *Q*, of the Generalized-Strain Mesh-free formulation.

This is the equation used in the GSMF formulation which, therefore, is free of integration. Since the work theorem is a weighted – residual weak form, it can be easily seen that this integration – free formulation is nothing else other than a weighted – residual weak – form collocation.

Equations (45), of the Generalized – Strain Mesh – free formulation, can be derived from another kinematically – admissible displacement field, directly defined in terms of Heaviside step function, see [8].

Discretization of Eq. (45) is carried out with the MLS approximation, Eq. (15) to (19), for the local domain Ω_Q, in terms of the nodal unknowns $\hat{\mathbf{u}}$, thus leading to the system of two linear algebraic equations

$$\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{n}_{\mathrm{x}_l}\mathbf{D}\mathbf{B}_{\mathrm{x}_l}\hat{\mathbf{u}}=-\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{\mathrm{x}_j}-\frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{\mathrm{x}_k} \tag{46}$$

that can be written as

$$\mathbf{K}_Q\hat{\mathbf{u}}=\mathbf{F}_Q \tag{47}$$

in which K_Q, the nodal stiffness matrix associated with the local Ω_Q, is a *2x2n* matrix given by

$$\mathbf{K}_Q=\frac{L_i}{n_i}\sum_{l=1}^{n_i}\mathbf{n}_{\mathrm{x}_l}\mathbf{D}\mathbf{B}_{\mathrm{x}_l} \tag{48}$$

and F_Q is the respective force vector given by

$$\mathbf{F}_Q = -\frac{L_t}{n_t}\sum_{j=1}^{n_t}\bar{\mathbf{t}}_{x_j} - \frac{S}{n_\Omega}\sum_{k=1}^{n_\Omega}\mathbf{b}_{x_k} \tag{49}$$

Consider that the problem has a total of *N* field nodes *Q,* each one associated with the respective local region Ω_Q. Assembling Eq. (47), for all *M* interior and static – boundary field nodes leads to the global system of *2M x 2N* equations

$$\mathbf{K}\hat{\mathbf{u}} = \mathbf{F} \tag{50}$$

Finally, the remaining equations are obtained from the *N – M* boundary field nodes on the kinematic boundary. For a field node on the kinematic boundary, a direct interpolation method is used to impose the Kinematic boundary condition as

$$u_k^h(\mathrm{x_j}) = \sum_{i=1}^{n}\phi_i(\mathrm{x_j})\hat{u}_{ik} = \bar{\mathbf{u}}_k \tag{51}$$

Or, in matrix form as

$$\mathbf{u}_k = \Phi_k\hat{\mathrm{u}} = \bar{\mathbf{u}}_k \tag{52}$$

with k = 1, 2, where $\bar{\mathbf{u}}_k$ is specified nodal displacement component. Equations (51) are directly assembled into the global system of equations (50).

Numerical Examples

This section presents some numerical results for Cantilever beam and the Plate with a circular hole for different nodal configurations. The effects of the size of local support and quadrature/collocation domain are analyzed and compared with the exact solution.

For a generic node i, the size of the local support ΩS and the local domain of integration Ωq are respectively given by

$$r\Omega_S = \alpha_S c_i, \tag{53}$$

$$r\Omega_q = \alpha_q c_i, \tag{54}$$

in which Ci represents the distance of the node i, to the nearest neighboring node. For the analysis performed in this paper different values were considered for the local support domain size (α_s), which vary from 4.0 to 10.0 with 0.5 increments, and the local quadrature/collocation domain size (α_q) which vary from 0.4 to 0.9999 with 0.05 increments.

Displacement and energy norms can be used for error estimation. These norms can be computed, respectively as

$$\|u\| = \left[\int_\Omega u^T u d\Omega\right]^{1/2} \tag{55}$$

$$\|\varepsilon\| = \left[\frac{1}{2}\int_\Omega \varepsilon^T D\varepsilon d\Omega\right]^{1/2} \tag{56}$$

The relative error for $\|u\|$ and $\|\varepsilon\|$ is given, respectively by

$$r_u = \frac{\|u_{num} - u_{exact}\|}{\|u_{exact}\|} \tag{57}$$

$$r_{\varepsilon} = \frac{\left\| \varepsilon_{num} - \varepsilon_{exact} \right\|}{\left\| \varepsilon_{exact} \right\|} \tag{58}$$

Cantilever Beam

A Cantilever beam showed in Fig. 6, is subjected to a parabolic traction at the free end. The principal properties are presented in Table 1 and the problem is solved for plane stress case.

Table 1. Properties of Cantilever Beam

Parameters	Values
Height, D	12 m
Length, L	48 m
Thickness, t	1 m
Load, P	1000 N
Modulus of Elasticity, E	30 MPa
Poisson`s Ratio, ν	0.3

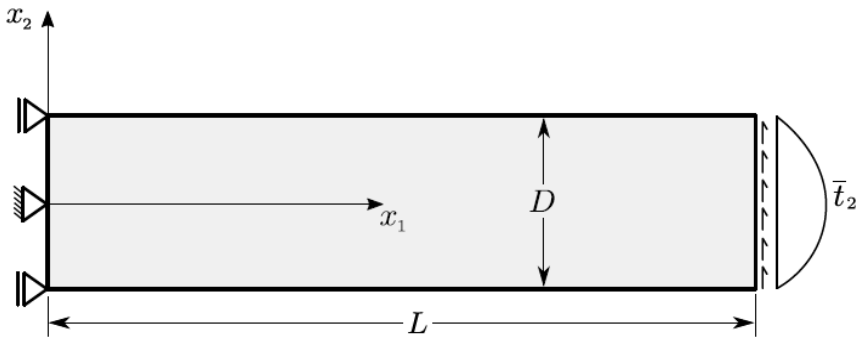


Figure 6. Cantilever beam

The parabolic traction and the moment of inertia is given by

$$\bar{t}_2(x_2) = -\frac{P}{2I}\left(\frac{D^2}{4} - x_2^2\right), \tag{59}$$

$$I = \frac{D^3}{12} \tag{60}$$

The exact solution of the problem is given by [19]. The equations for the exact displacement are:

$$u_1(x_1, x_2) = -\frac{Px_2}{6EI}\left[(6L - 3x_1) + (2+\nu)\left(x_2^2 - \frac{D^2}{4}\right)\right] \tag{61}$$

$$u_2(x_1, x_2) = \frac{P}{6EI}\left[3\nu x_2^2(L - x_1) + (4+5\nu)\frac{D^2 x_1}{4} + (3L - x_1)x_1^2\right] \tag{62}$$

The GSMF was used for solving this problem, five regular nodal distributions were considered with a discretization of 33 x 5 = 165 nodes, 65 x 9 = 585 nodes, 97 x 13 = 1261 nodes, 129 x 17 = 2193 nodes and 193 x 25 = 4825 nodes. The first discretization is showed in the Fig. 7.

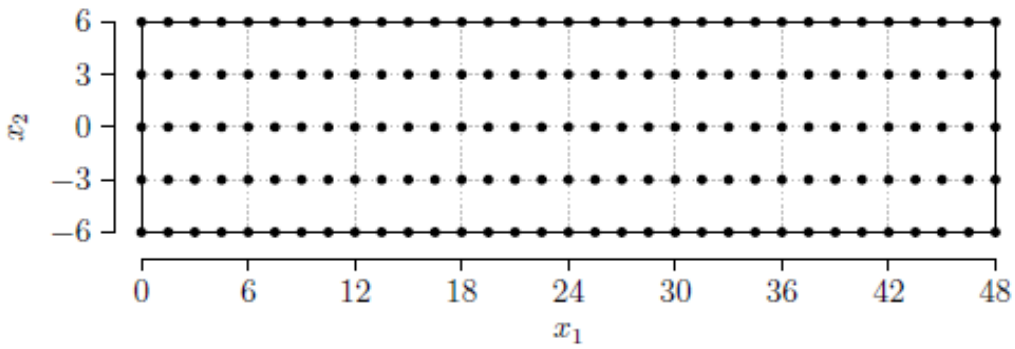


Figure 7. The regular nodal distribution of 33 x 5 = 165 nodes.

Influence of the local support domain size (α_s)

This parameter must be greater than 1.0, the reason is that for the small values, the algorithm of MLS approximation may be singular and the shape function cannot be constructed, because there is not enough nodes for the interpolation. The influence of α_s in the solution is obtained when the α_q is fixed.

Figure 8 shows the variation of relative error as a function of the size of the local support domain with 13 ratios which vary from 4.0 to 10.0 with 0.5 increments, and $\alpha_q = 0.5$.

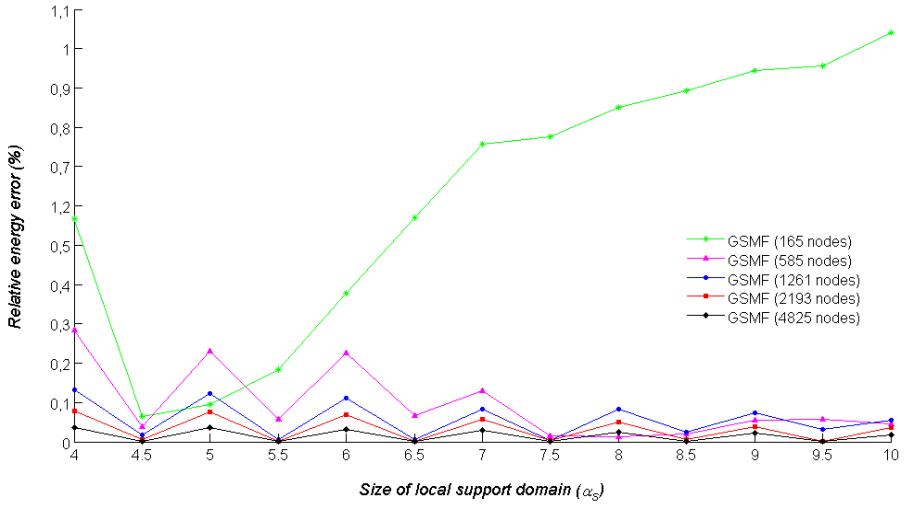


Figure 8. Analysis of the mesh - free discretization parameter α_s with $\alpha_q = 0.5$ defined in equation (53), carried out with five regular discretization of the cantilever – beam, with 33 x 5 = 165, 65 x 9 = 585, 97 x 13 = 1261, 129 x 17 = 2193 and 193 x 25 = 4825 nodes.

Figure 8 shows that the value of 4.5 for the local support domain (α_s) presents low relative energy errors in the five regular discretization. The same graphic is obtained for the displacement.

Influence of the local quadrature/collocation domain size (α_q)

This parameter must be less than 1.0. The reason is to ensure that the local sub – domains of the internal nodes are entirely within the solution domain, without being intersected by the global boundary. The influence of α_q is obtained when the α_s is fixed.

Figure 9 shows the variation of relative energy error as a function of the size of the local quadrature/collocation domain with 13 ratios, which vary from 0.4 to 0.9999 with 0.05 increments, and fixed on $\alpha_q = 4.5$.

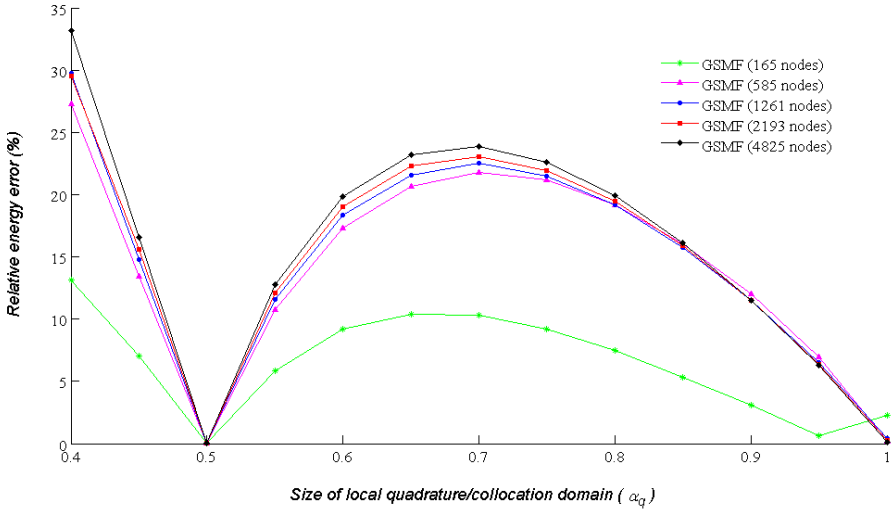


Figure 9. Analysis of the mesh - free discretization parameter α_q with $\alpha_s = 4.5$ defined in equation (54), carried out with five regular discretization of the cantilever – beam, with 33 x 5 = 165, 65 x 9 = 585, 97 x 13 = 1261, 129 x 17 = 2193 and 193 x 25 = 4825 nodes.

Figure 9 shows that the value of 0.5 for the local support domain (α_s) presents low relative energy errors in the five regular discretization. The same graphic is obtained for the displacement.

The Figure 10 and 11 presented the relative energy error suface for the nodal regular discretization with 97 x 13 = 1261 nodes, in one direction is presented the variation from 0.4 to 0.9999 for the size of local quadrature/collcation domain (α_q) and the other diretion the variation from 4 to 10 for the local support domain (α_s).

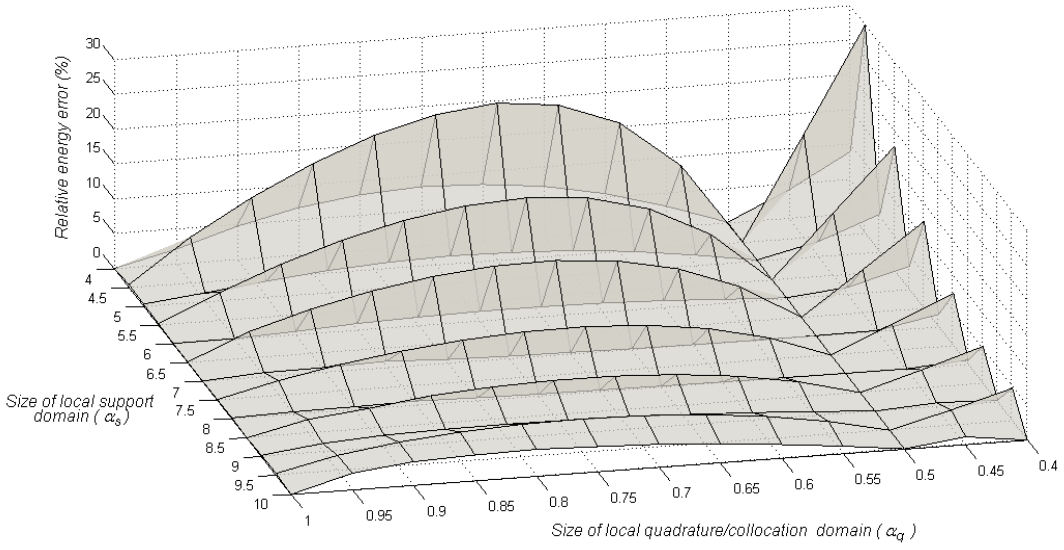


Figure 10. Analysis of the mesh - free discretization for parameters α_q and α_s, carried out by the regular discretization of the cantilever – beam, with 97 x 13 = 1261 nodes. Especial focus for the α_q parameter.

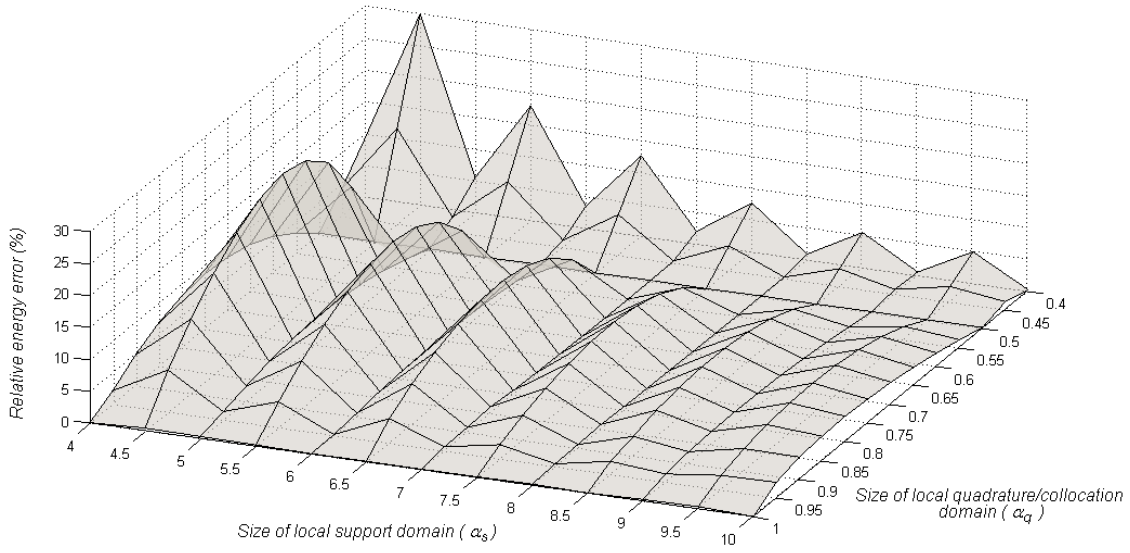


Figure 11. Analysis of the mesh - free discretization for parameters α_q and α_s, carried out by the regular discretization of the cantilever – beam, with 97 x 13 = 1261 nodes. Especial focus for the α_s parameter.

The Figure 11 and 12 shows a relative energy error surface in percentage for two different angles, the error for $\alpha_q = 0.5$ is always the lowest; as for the α_s there are different values, but the one that presents a greater efficiency is 4.5, because it results in a smaller CPU time consumption. Similar results were obtained for the relative displacement error.

Plate with a circular hole

Consider an infinite plate with a centered circular hole under unidirectional unit tension along the x1 direction, as represented in Fig. 12. Due to the symmetry of the problem about the horizontal and vertical axes, only a portion of the upper right quadrant of the plate is considered. The modeled section of the plate has dimensions b x b and the center circle has a radius a = 1, with b = 5a. The principal properties are presented in Table 2 and the problem is solved for plane stress case.

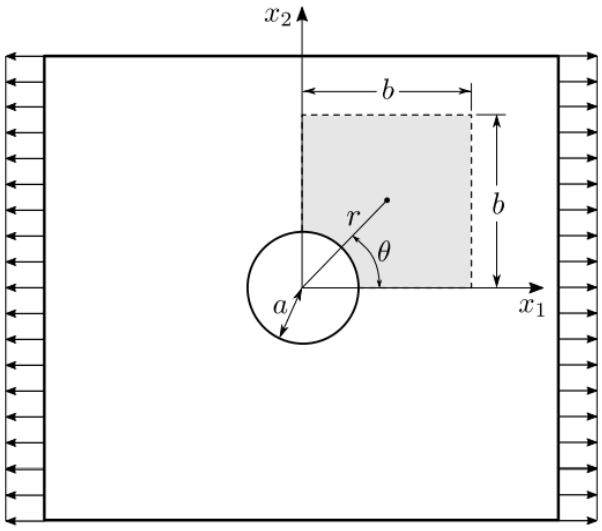


Figure 12. Plate with a hole.

Table 2. Properties of Plate with Circular Hole

Parameters	Values
Hole Radius, a	1 m
Height, D	5 m
Length, L	5 m
Thickness, t	1 m
Modulus of Elasticity, E	100000 Pa
Poisson`s Ratio, ν	0.25

The exact stress distribution in the plate is given by

$$\sigma_{11}(r,\theta) = 1 - \frac{a^2}{r^2}\left(\frac{3}{2}\cos 2\theta + \cos 4\theta\right) + \frac{3}{2}\frac{a^4}{r^4}\cos 4\theta \tag{63}$$

$$\sigma_{22}(r,\theta) = -\frac{a^2}{r^2}\left(\frac{1}{2}\cos 2\theta - \cos 4\theta\right) - \frac{3}{2}\frac{a^4}{r^4}\cos 4\theta \tag{64}$$

$$\sigma_{12}(r,\theta) = -\frac{a^2}{r^2}\left(\frac{1}{2}\cos 2\theta - \cos 4\theta\right) - \frac{3}{2}\frac{a^4}{r^4}\cos 4\theta \tag{65}$$

Where r and θ are the usual polar coordinates, centered at the center of the hole. A plane-stress state is considered which leads to the following displacements

$$u_1(r,\theta) = -\frac{\cos\theta}{2r^3E}\left[4a^4\cos^2\theta(1+\nu)(1-r^2) - 3a^4(1+\nu) + (ar)^2(1-3\nu) - 2r^4\right] \tag{66}$$

$$u_2(r,\theta) = -\frac{\sin\theta}{2r^3E}\left[4a^4\cos^2\theta(1+\nu)(1-r^2) - a^4(1+\nu) + (ar)^2(\nu-3) + 2r^4\nu\right] \tag{67}$$

The bottom and left edges of the plate are assumed as kinematic boundaries, with displacements specified on the bottom u_2 (x_1, x_2 = 0) and left edges (u_1 (x_1 = 0, x_1 = L, x_2) = 0). The rigid and top edges are assumed as static boundaries, loaded by tractions computed from the stresses of the exact solution (63, 64 e 65) as $t_j = \sigma_{ij,}$ in wich n_i represents the components of the unit outward normal to the edge of the plate.

The solve this problem, the plate was discretized by [17] with 9 nodes in the tangential direction and 11 nodes in the radial direction, see Fig. 13, with circular local supports. MLS approximation was considered, with the second order polynomial basis. GSMF was applied with 1 boundary collocation point per quadrant of the circular local domain.

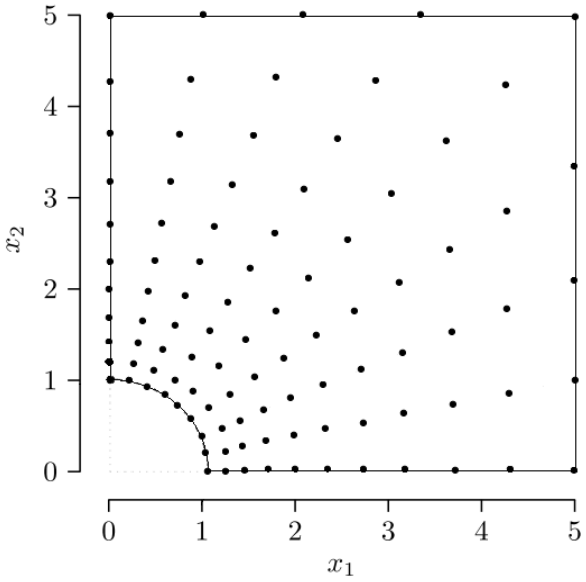


Figure 13. Nodal distribution for the infinite plate of 11x9 = 99 nodes.

The Figure 14 and 15 presented the relative energy error suface for the nodal regular discretization with 9 x 11 = 99 nodes, in one direction is presented the variation from 0.2 to 0.55 for the size of local quadrature/collocation domain (α_q) and the other diretion the variation from 3.5 to 6.0 for the local support domain (α_s).

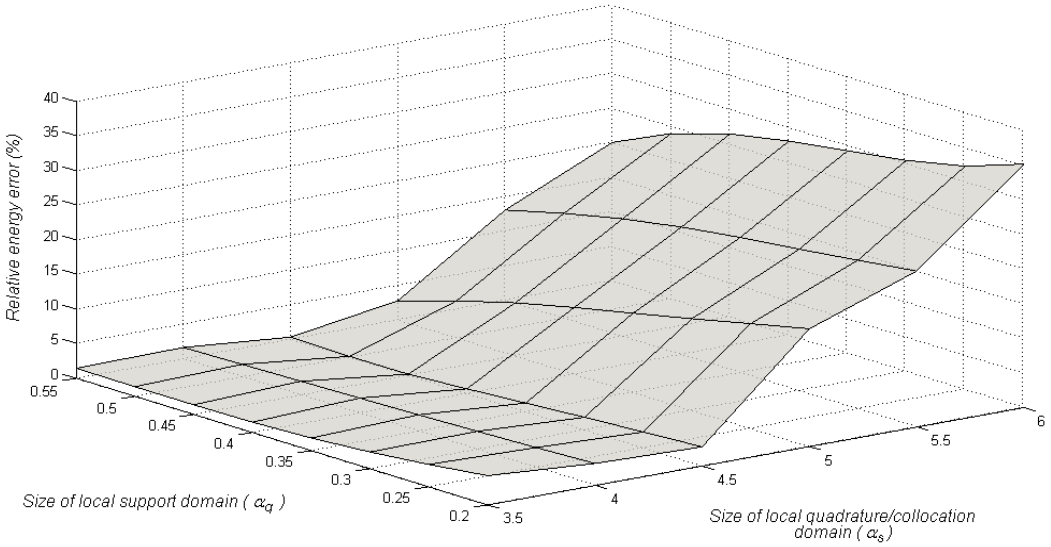


Figure 14. Analysis of the mesh - free discretization for parameters α_q and α_s, carried out by the regular discretization of the plate with a circular hole, with 9 x 11 = 99 nodes. Especial focus for the α_s parameter.

The Figure 14 and 15 shows the relative energy error surface in percentage for two different angles, the error for the α_s = 4.5 is always the least, for the α_s there are different values but the one that presents greater efficiency is 0.5 because present less CPU time consumption. Similar results were obtained for the relative displacement error.

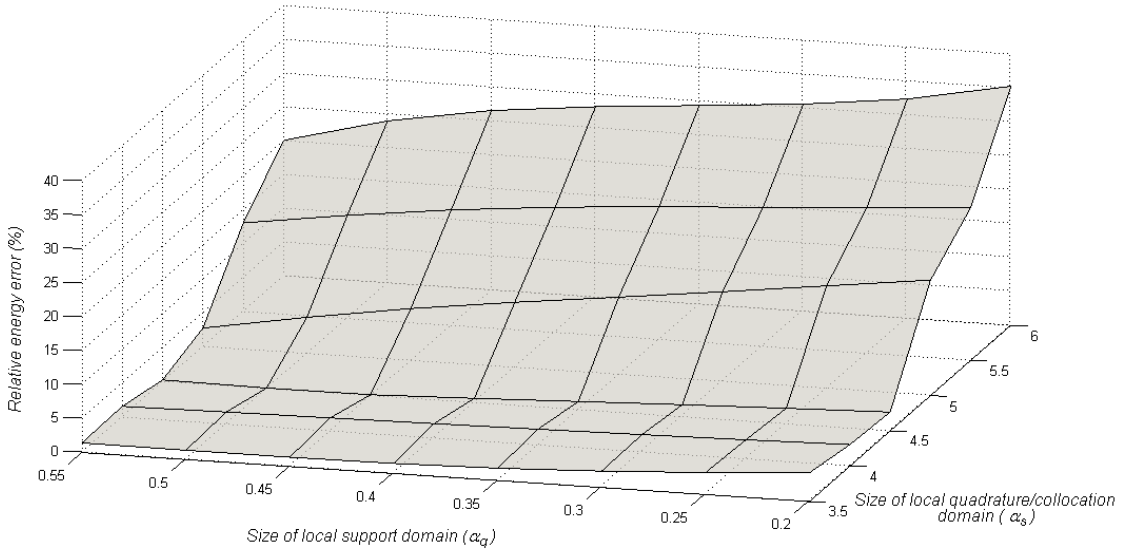


Figure 15. Analysis of the mesh - free discretization for parameters α_q and α_s, carried out by the regular discretization of the plate with a circular hole, with 9 **x** 11 = 99 nodes. Especial focus for the α_q parameter.

Conclusions

The size of local support domain (α_s) and local quadrature/collocation domain (α_q) affect the accuracy and performance of GSMF local method. For the different nodal discretization this parameters have a great influence in the relative energy and displacement error; but it is especially small for one particular values.

The special study realized for the cantilever beam discretized with 97 x 13 = 1261 nodes, presented greater variation in the relative energy error for different values of local support domain (α_s) and local quadrature/collocation domain (α_q). The parameter $\alpha_s = 0.5$ presented the lowest values independent of α_q value.

In this study were found certain values for the local support domain (α_s) and local quadrature/collocation domain (α_q) that generate less errors when the f method is applicate, these values are $\alpha_s = 4.5$ and $\alpha_q = 0$. It is necessary study others cases and other discretization especially for the plate with a circular hole.

Acknowledgments

The program PECC – *Pós-Graduação em Estruturas e Construção Civil*, Department of Civil and Environmental Engineering, Faculty of Technology, University of Brasília and CNPq – *Brazilian National Counsel of Technological and Scientific Development* for his PhD scholarship.

References

[1] Belytshko, T. Lu, Y. Y. and Gu, L. (1994) Element – free Galerkin methods, *International Journal for Numerical Methods in Engineering*, Vol. 20, 1081 – 1106.

[2] Atluri, S. N. and Shen, S. (2002) The Meshless Local Petrov – Galerkin (MLPG) Method: A simple and Less-costly Alternative to the Finite Element and Boundary Element Methods, *CMES: Computer Modeling in Engineering and Sciences*, Vol **3**(1), 11-51.

[3] Sladek, J. Sladek V. and Zhang, C. (2003) Application of Meshless Local Petrov – Galerking (MLPG) Method to Electrodynamic Problems in Continuously Non – Homogeneous Solids. *Computer Modeling in Engineering & Sciences*, Vol. **4**,637-648.

[4] Atluri, S. N. Han, Z. D. and Rajendram, A. M. (2004) A New Implementation of the Meshless Finite Volume Method Through the MLPG Mixed Approach, *CMES: Computer Modeling in Engineering and Sciences*, Vol **6**, 491-513.

[5] Alturi, S. N. (2004) The Meshless Method, (MLPG) for domain & BIE Discretizations, *Tech Science Press, Forsyth*.

[6] Atluri, S. N. Liu H. T. and Han Z. D. (2006) Meshless Local Petrov – Galerkin (MLPG) Mixed Collocation Method for Elasticity Problems, *CMES: Computer Modeling in Engineering and Sciences*, Vol **14**, 141-152.

[7] Liu, G. R. (2009) Meshfree methods moving beyond the finite element method. CRC Press, 2nd edition.

[8] Finalyson, B. A. (1972) The Method of Weighted Residuals and Vibrational Principles. Academic Press.

[9] Oliveira, T. and Portela, A. (2016) Weak – Form Collocation – a Local Meshless Method in Linear Elasticity, *Engineering Analysis with Boundary Elements*, 73: 144 – 160.

[10] Moussaoui, A. and Bouziane, T. (2013) Comparative Study of the Effect of the Parameters of Sizing Data on Results by the Meshless Methods (MLPG), *World Journal of Mechanics*, 3, 82-87.

[11] Erdayi, D. C. (2014) Meshless Local Petrov – Galerkin Method for Plane Elasticity Problems, Master Thesis, Middle East Technical University, Turkey, 2014.

[12] Fredholm, I. (1906) Solution d'un probléme fondamental de la théorie de l'éasticité, Arkiv for Matematik Astronomi och Fysk 2, 28(1):1-8.

[13] Fichera, G. (2006) Linear Elliptic Diferencial Systems and Eigenvalue Problems. Springer.

[14] Brebbia, C. A. and Tottenham, H. (1985) Variational Basis of Aproximate Models in Continuum Mechanics. Southampton and Springer Verlag.

[15] Gedalf, I. M. and Shilov, G. E. (1964) Generalized Functions, volume 1. Academic Press.

[16] Timoshenko, S. P. and Goodier, J. N. (1970). *Theory of Elasticity*, 3rd edn, McGraw-Hill, New York, USA.

[17] Alturi, S. N. and Zhu, T. L. (2000) The meshless local Petrov – Galerkin (MLPG) approach for solving problems in elasto – statics, *Computacional Mechanics*, 25:169-179.

Numerical Simulation of Internal Flow Field Characteristics for a Ducted Propeller

***Jie Gong [1], †C.Y. Guo [1], T.C. Wu [1], K.W. Song [1], and J.F. Lin [1]**

[1]College of Shipbuilding Engineering, Harbin Engineering University, China

*Presenting author: gongjie13@hrbeu.edu.cn
†Corresponding author: guochunyu_heu@outlook.com

Abstract

The internal flow field characteristics of a ducted propeller set in a propeller open water test have been investigated numerically. Detached Eddy Simulation (DES) approach is employed to track the main features of the vortical structures of the internal wake field. The Spalart & Allmaras turbulence model and unsteady sliding mesh technique have been used in the simulations. Appropriate grid divided topology meeting the y^+ requirement and boundary conditions set to satisfy all research objectives are discussed. The results show that the numerical hydrodynamic performance results are in good agreement with the model test. The peak value of the surface pressure fluctuation occurs in the blade passing frequency (BPF) and its multiples in the frequency domain. DES model can effectively simulate the different vortex structures of the ducted propeller, such as shear layer vortices of the duct, blade tip vortices, blade shedding vortices and blade root vortex. Tip vortex structures perform strong deformation due to the interaction between the duct and the propeller. The morphology of tip vortex changes and the vorticity distribution covers a larger area of the inner surface of the duct. All above contributes to the recovery of the propeller wake vortex energy and increases the efficiency of the propulsion.

Keywords: Ducted propeller, DES, Pressure fluctuation, Vortex structure.

1. Introduction

Ducted propeller is widely used in the propulsion system for various engineering vessels. The duct could provide protection for the propeller and improve the efficiency of the propulsion system. In addition, the ducted propeller gets a better performance for cavitation and radiation noise, especially in heavy load conditions. Research [1] shows that the circulation flow that distributes near the inner surface of the duct is the essential reason for the production of duct thrust, while the duct is helpful to recover the vortex energy induced by the propeller. In order to further understand the mechanism of the hydrodynamic performance of the ducted propeller, the flow characteristics of the internal flow field of ducted propeller (especially vortex characteristics) are necessary analyzed.

Most of the early numerical calculations are based on potential flow theory [2]-[3], Su et al. [4] and Xie et al. [5] has calculated the hydrodynamic performance and internal flow field of ducted propeller by using the panel method theory which is based on velocity potential. It is concluded that the computational efficiency of potential flow method is considerable, but it is difficult to simulate the complex flow accurately. Recently, numerous numerical methods based on viscous flow have been applied in the simulations for ducted propellers [6]-[9]. Based on solving the Reynolds-averaged Navier-Stokes equations (RANS), the flow field around the ducted propeller and the characteristics of the wake vortex are predicted by Hu et al. [10] and Shi et al. [11]. Bhattacharyya et al. performed extensive studies on predicting the

hydrodynamic performance of ducted propeller, including the manner in which different types of ducts affect propeller performance [12], scale effects of the ducted propellers [13], and the application of computational fluid dynamics (CFD) in the design of ducted propellers [14]. Gaggero et al. [15] used the RANS method to perform numerical predictions on cavitating tip vortices in ducted propellers, and the meshing and refinement topology used in this study are highly valuable as a reference for numerical studies as they attempted to capture the details of the flow field.

In this paper, numerical simulation of ducted propeller for open-water test is performed basing on the detached eddy simulation (DES) method. In order to confirm the accuracy and reliability of numerical calculation, the calculated hydrodynamic performance is verified by comparing with model test results. The detailed numerical model of grid type, turbulence model and computational methods has been described. The objective of this study focus on the analysis of the distribution characteristics of the internal flow field of the ducted propeller, as well as the structural composition, morphology and distribution characteristics of vortices of internal flow fields following the effects of complex interferences.

2. Computational details

2.1 Governing equations

The numerical simulations of viscous flow field is performed by the integration of the RANS equation. The motion of incompressible Newton fluids is integrated with the continuity equation and momentum conservation equation [16] as follows:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial(\rho u_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j) = \frac{\partial p}{\partial x_j} + \frac{\partial}{\partial x_j}(\mu \frac{\partial u_i}{\partial x_j} - \rho \overline{u_i' u_j'}) + S_j \tag{2}$$

where u_i and u_j denote the time-averaged values of velocity components (i, j = 1, 2, 3), p denotes time-averaged pressure, μ denotes the dynamic viscosity coefficient, $\rho \overline{u_i' u_j'}$ denotes the Reynolds stress term, and S_j denotes the generalised source term of the momentum equation.

The DES model is a hybrid numerical model that combines the strengths of both RANS and LES methods. In all DES simulations, the Spalart & Allmaras one-equation [17] turbulence model is employed to resolve turbulence structures in the wake. The finite volume based segregated flow solver of the STAR-CCM+ CFD code is utilised to simulate the flow-field characteristics of the internal vortices generated by the ducted propeller under various advance coefficients. In the DES simulation, the SIMPLE algorithm is adopted to solve the pressure-velocity coupling equation. Specifically, convection, diffusion, and time terms are discretised by using the 2nd order upwind scheme, central differencing scheme, and 2nd order implicit time discretisation, respectively. In the unsteady DES simulation, the Courant-Friedrichs-Lewy (CFL) condition is used to balance the temporal and the spatial scales.

2.2 Computational conditions

The propeller is a constant-pitch 4-bladed propeller with a diameter of D = 0.240 m, a hub ratio of 0.190, and a propeller area ratio of 0.417. The Wageningen 19A duct is integrated with a tip clearance of 2 mm. The geometry of the studied ducted propeller is shown in Figure 1. An O-xyz Cartesian coordinate system with the origin O located at the geometrical centre of the propeller is built in the simulation. The positive direction of the x-axis is defined as the uniform inflow direction, while the positive directions of the y and z axes are determined by the right-hand rule.

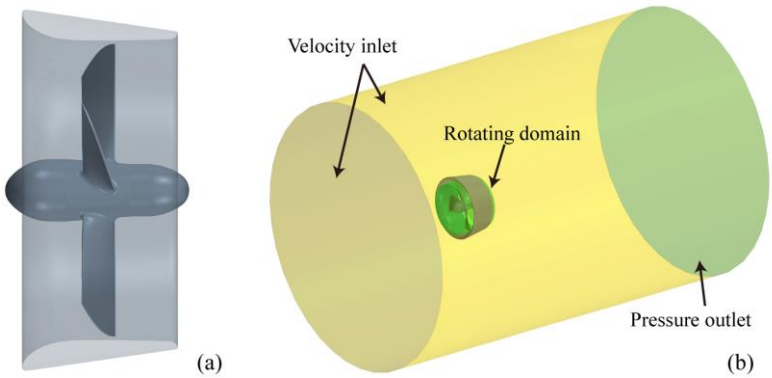


Figure 1. Geometric model and calculation domain

The computational domain is a cylinder with a diameter of 8*D*, and extends 2*D* upstream and 8*D* downstream from the origin O, respectively. The sliding mesh method is used in the simulation of the propeller rotating motion, hence, the computational domain is divided into static and rotating sub-domains. The boundary conditions of the computational domain are set as follows: the inlet and the side boundary of the cylinder are both set to the velocity inlet; the outlet is set to the pressure outlet; the surfaces of the propeller and the duct are set to the no-slip wall.

2.3 Mesh characteristics

A proper discretisation of the computational domain is extremely important to improve the accuracy of the numerical simulation. It is necessary for the cell size of the computational domain to transition in a gradual manner to avoid numerical dissipations. In the study, the rotational sub-domain is discretised by using a polyhedral mesh while the static sub-domain is discretised by using a trimmed mesh.

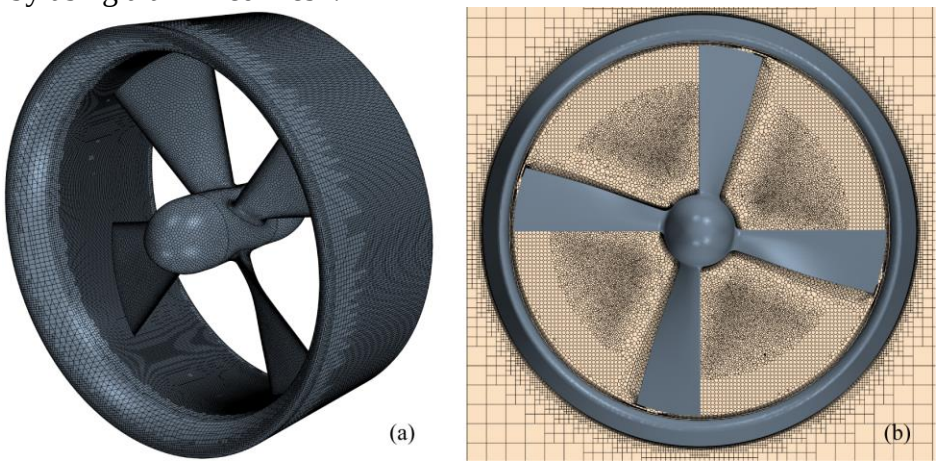


Figure 2. Surface mesh of ducted propeller and details of mid-ship section

The meshes of two sub-domains possess a same base size (BS) of 0.1 m. A volume refinement (size = 2.5% BS) is implemented in the annular cylindrical region of the blades' tips, while a gap refining region with an even finer grid size (size = 1.25% BS) is employed in the gap between the blades' tips and the duct (shown in Figure 2). The blade surfaces are refined along the feature lines of the leading and trailing edges (size = 0.25% BS). In addition, a hub vortex refining region is set up behind the hub to capture the formation and evolution of the hub vortex. The total number of cells of the computational domain corresponds to 9.51 million while the static and rotational sub-domains account for 7.36 million and 2.15 million. In order to satisfy the requirements of the DES algorithm, the y^+ value of the walls is restricted to $y^+ < 1$.

3. Results and discussion

3.1 Verification of hydrodynamic performance

The hydrodynamic coefficients of ducted propellers are defined as follows:

$$J = \frac{Vx}{nD} \tag{3}$$

$$Ktp = \frac{T_p}{\rho n^2 D^4}, Ktd = \frac{T_d}{\rho n^2 D^4} \tag{4}$$

$$Kq = \frac{Q_0}{\rho n^2 D^5} \tag{5}$$

$$\eta_0 = \frac{J}{2\pi}\frac{Ktp + Ktd}{Kq} \tag{6}$$

where, J is the advanced coefficient, Vx is the local axial velocity, n is the rotate speed, D is the diameter of the propeller, Tp and Td are the thrusts of propeller and duct, Q_0 is the torque of propeller，ρ is the density of fluid. Ktp and Ktd *are* the thrust coefficients of propeller and duct, respectively. Kq is the torque coefficient of propeller and η_0 is the efficiency of the ducted propeller.

Three advance coefficients, 0.3, 0.4 and 0.5, were chosen to allow for a comparison between different loads. A constant rotational speed, n = 11 revolutions per second (RPS) was selected in the simulations, and the time step was set to 1° /step. The fluid density was kept constant at ρ = 997.56 kg m^{-3}, and the Reynolds number is the same as that of the model experiment, i.e. Re = 3.5~3.6×10^5.

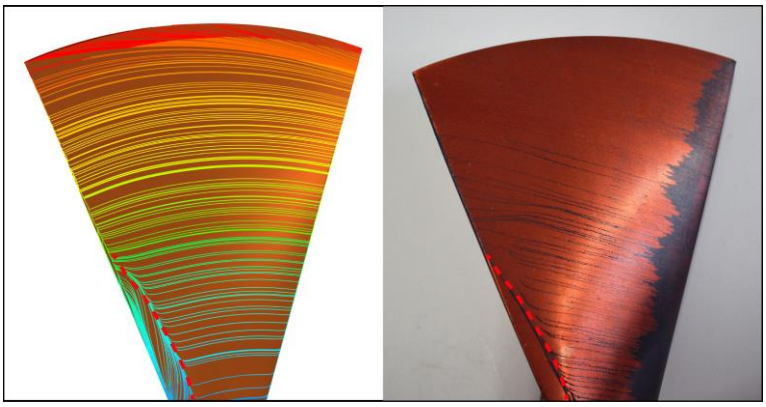


Figure 3. Streamlines of blade surface; Left: CFD, Right: EFD

Compared with the experiment results, the maximum relative error among all coefficients is the duct thrust coefficient, which has reached 4.42%. The periodic rotation of the impeller has caused the thrust fluctuation of duct. All thrust coefficients, torque coefficients of the propeller are in good agreement with the experimental results, the errors are within 3%. Results above show that the computational accuracy of the numerical model is acceptable, and the analysis of the characteristics of the internal flow field and vortex characteristics are of relative reliability.

Table1. Comparison of numerical calculation (CFD) and model test results (EFD)

J	Ktp			Ktd			$10Kq$		
	CFD	EFD	Er.(%)	CFD	EFD	Er.(%)	CFD	EFD	Er.(%)
0.3	0.2017	0.1985	1.61	0.0842	0.0881	-4.42	0.3369	0.3401	-0.94
0.4	0.1806	0.183	-1.31	0.0565	0.0581	-2.75	0.308	0.316	-2.53
0.5	0.1615	0.1664	-2.94	0.0287	0.0279	2.86	0.2876	0.2915	-1.33

3.2 Internal flow field characteristics

Typical profiles are selected to analyze the characteristics of the internal flow field. Two cross sections (y-z plane) are selected at the front disk (x=-0.5R) and the back disk (x=0.3R) in the

internal region, and all the coordinates are normalized by the radius of the propeller. The transient induced axial velocity and vorticity distribution are analyzed in detail.

The presence of the duct enhanced the suction effect in front of the propeller, and the axial speed in the front disk changed obviously. Analysis along the radial direction, the axial induced velocity (*Vx-in*) increased first and then decreased (Figure 4 (a)), which is caused by the retardant of the duct surface and the hub. The axial induced velocity has shown obvious periodicity fluctuation for different radius in the front disk during one period. There are 4 peaks of the induced axial velocity for all the radius (0.4*R*, 0.6*R* and 0.8*R*), which is the same with the number of blades. In addition, as closer to the outside of the disk, the induced velocity fluctuation amplitude is larger, and the position of 0.8*R* generates the most intense fluctuation obviously.

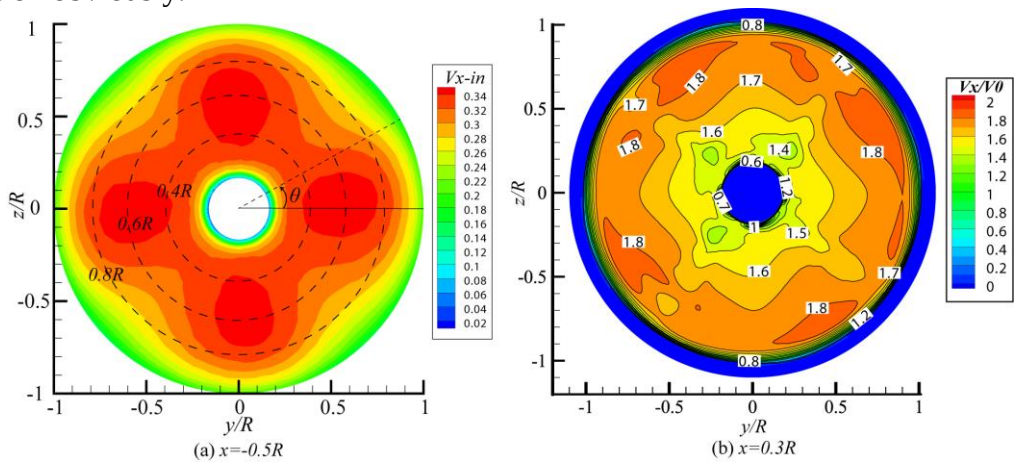


Figure 4. Axial velocity distribution of the front and back disks (*x*=-0.5*R* and *x*=0.3*R*)

The dimensionless axial velocity (*Vx/V0*) distribution for the back disk (x=0.3*R*) is shown in Figure 4 (b). It can be observed that acceleration of the blades is reflected in the range of 0.5-0.8*R*, and the transient axial velocity distributes in the range of 1.4~1.8. The vorticity distribution could reflect the composition of the internal vortex system of the duct propeller. The shear flow of the inner surface of the duct generates shear layer vortices, and the shear layer vortices show a ring shape in the cross section. On the other hand, the vortex system induced by the propeller are described as follows: blade tip vortex (Figure 5(a)-I), blade shedding vortex (Figure 5(a)-II) and blade root vortex (Figure 5(a)-III), and the vortices cores distributed near the maximum value of vorticity. Furthermore, the turbulent viscosity (Figure 5(b)) in the back disk reflects the eddy diffusion motion. It is obvious that the turbulent viscosity is larger in the blade tip and root zones, and which means the turbulent viscosity distribution is consistent with the vortex distribution characteristics.

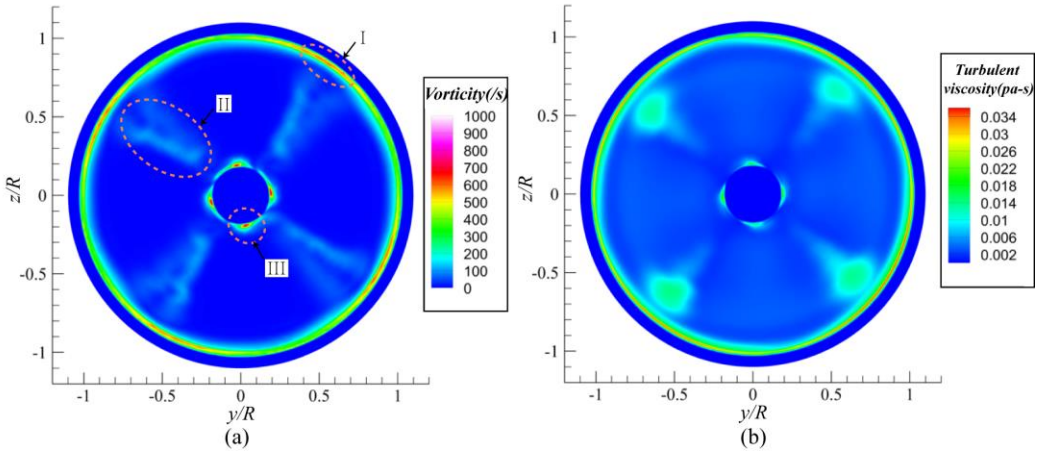


Figure 5. Typical quantities distributions of the back disk(*x*=0.3*R*); (a) vorticity; (b) turbulent viscosity

Meanwhile, by setting the probe P_1 at the x=0 cross section of the inner surface of the duct, the variation of pressure pulsation at P_1 position for operation conditions is monitored. For the time domain, there are 4 peak and valley pressure value during one period, and the pulsation amplitude is directly related with the blade position. The flow field changes sharply and the hydrodynamic performance of the ducted propeller is unsteady. For the frequency domain, the pressure peaks appear at the blade passing frequency (BPF) and its multiples, the continuous peaks are progressively shifted with respect to the integer BPF. The pressure pulsation characteristics of the inner surface of the duct relate to the actual working conditions of ducted propellers.

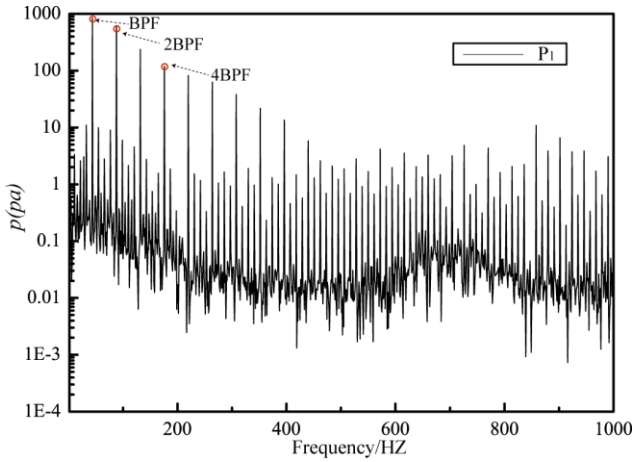


Figure 6. Pressure fluctuation of probe P_1 in the frequence domain

3.3 Vortex characteristics analysis

The vortex characteristics of internal flow fields and its distribution affect the hydrodynamic performance of the ducted propeller directly. The horizontal section (z=0) is selected to analyze the distribution of velocity and vorticity along the stream direction. As shown in Figure 7, the acceleration effect of the ducted propeller on the passing flow is obvious in the velocity field. Because of the fliud viscous action, there is a low velocity zone diatributing at the trailing edge of the duct. On the other hand, there distributes large areas of vorticity energy near the duct inner surface(Figure 8), the vortices are generated by the pressure difference before and after the blade surfaces. The strong interaction between the propeller and the duct has made the vortices structures more complicated(Figure 8-I). It is relatively clear to recognize the blade root vortex (Figure 8-II) and the hub vortex (Figure 8-III). In addition, the shear layer vortex of the outer surface and the internal vortices interfere with each other near the duct tailing region, and the vortex energy transfer to the downstream. There exists continuous elliptical vorticity distributions (Figure 8-IV) in the near wake field.

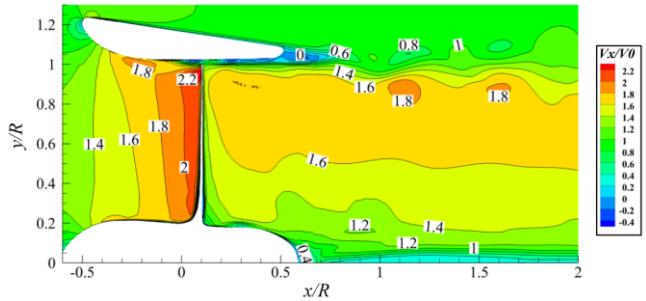


Figure 7. Instantaneous velocity distribution of the horizontal section (z=0)

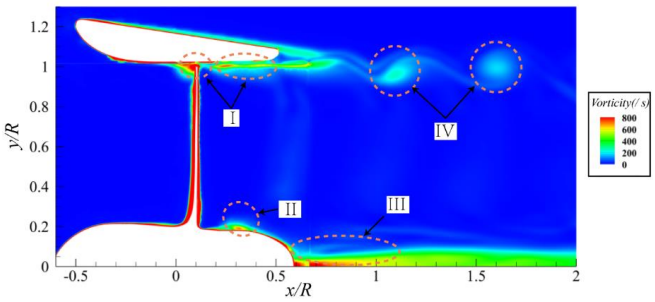


Figure 8. Instantaneous vorticity distribution of the horizontal section (z=0)

The volume render method is employed to analyze the instantaneous spatial vorticity distributions. As shown in Figure 9, four flake blade shedding vortices are generated near the blade trailing edges, and the energy diffuses rapidly in the stream. The different root vortices develop spirally and gather near the hub region gradually. The tip vortex distributes on the inner surface of the duct, and the morphology of the tip vortex has changed significantly. There is a larger area of vortices distributing on the inner wall, which causes stronger circulations and produces the hydrodynamic thrust of the duct.

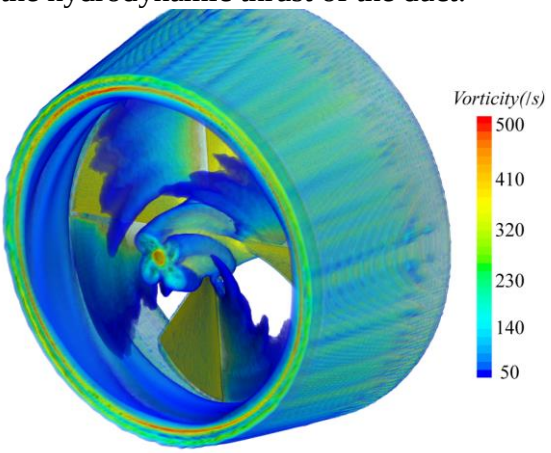


Figure. 9 Three-dimensional vorticity distribution in internal flow field of ducted propeller

The internal characteristics of tip vortex deformation are discriminated. The contour projection is in the shape of “W”, and there are two eddy centres in the internal flow field, which is obviously different from the vorticity distribution of a non-ducted propeller. Furthermore, it is found that the vorticity value of two eddy centres gradually decreased along the flow direction.

4. Conclusion

The DES method has been implemented in the simulations of the internal flow field characteristics and vortex structures of the ducted propeller. The results indicate that DES method can obtain a reasonable numerical accuracy and capture the flow characteristics well. The conclusions are as follows:

1. The internal flow field vortex system of the ducted propeller includes the shear layer vortex of the duct, the blade tip vortex, shedding vortex, root vortex and the hub vortex.
2. The existence of the duct directly affects the distribution of propeller tip vortex. A larger area of vortices distributing on the inner wall of the duct is helpful to recover the vortex energy so as to improve the propulsion efficiency.

Acknowledgements

This project was supported by the National Natural Science Foundation of China (Grant Nos. 51379043, 51209048 and 51409063), and thanks for the technical support provided by Dr. Xu Yongze and Dou Pengfei of Siemens Ltd. China.

Reference

[1]Go, J.S., Yoon, H.S. and Jung, J.H. (2017) Effects of a duct before a propeller on propulsion performance. Ocean Engineering, 136, 54-66.
[2] WANG, G.Q. and LIU, X.G. (2006) Prediction of unsteady performance of ducted propellers by potential based panel method, Journal of Ship Mechanics, 10(1): 36-42.
[3] HU, J., HUANG, S., MA, C. and XIE, X.S. (2009) Several Influence Factors for the Inner Flow Field of Ducted Propeller, JOURNAL OF TIANJIN UNIVERSITY, 42(4):340-344.
[4] SU, Y.M., LIU, Y.B., SHEN, H.L. and DOU, F.X. (2012) Surface panel method-based numerical calculation for predicting ducted propeller performances, JOURNAL OF HUAZHONG UNIVERSITY OF SCIENCE AND TECHNOLOGY. NATURE SCIENCE, 08:57-61.
[5] XIE, X.S., HUANG, S., HU, J., and MA, C. (2009) Inner flow field calculations for ducted propellers, JOURNAL OF HARBIN ENGINEERING UNIVERSITY, 01: 7 -12.
[6] ZHOU, J.W., LI, F.Z. and MEI, L. (2017) Peak efficiency analysis of non-cavitation ducted propeller, Journal of Harbin Institute of Technology, 04: 149-155.
[7] LIU, Y.B., SHU, Y.M., ZHAO J.X. and ZHANG, H. (2014) Optimal design of a ducted propeller based on the circulation theory and pump theory, JOURNAL OF HARBIN ENGINEERING UNIVERSITY, 11: 1307-1313.
[8] RAO, Z.Q., LI, W. and YANG, J.C. (2014) The Effect of Stator Parameters on Performance of Ducted Propeller with Pre-swirl Stators, Journal of Shanghai Jiaotong University, 11:1307-1313.
[9] HU, J., WANG, N. and HU, Y. (2017) Performance comparison of accelerating duct and decelerating duct, Journal of Beijing University of Aeronautics and Astronautics, 02: 240-252.
[10] HU, J., WANG, N. and HU, Y. (2017) Numerical Study on the Hydrodynamic performance of Ducted Propellers, JOURNAL OF HARBIN ENGINEERING UNIVERSITY, 06:1-7.
[11] SHI, L.P., XIONG, Y., YANG, Y.and WANG, B. (2016) Numerical simulation of the flow around a ducted propeller using Reynolds-averaged Navier-Stokes equations, JOURNAL OF HARBIN ENGINEERING UNIVERSITY, 03: 344-348.
[12] Bhattacharyya, A., Krasilnikov, V. and Steen, S. (2016) Scale effects on open water characteristics of a controllable pitch propeller working within different duct designs, Ocean Engineering, 112: 226-242.
[13] Bhattacharyya, A., V. Krasilnikov, and Steen, S. (2015) Scale Effects on a 4-Bladed Propeller Operating in Ducts of Different Design in Open Water, Fourth International Symposium on Marine Propulsors smp'15, USA.
[14] Bhattacharyya, A., V. Krasilnikov, and Steen, S. (2016) A CFD-based scaling approach for ducted propellers, Ocean Engineering, 123:116–130.
[15] S. Gaggero, G. Tani, M. Viviani, and F. Conti. (2014) A study on the numerical prediction of propellers cavitating tip vortex, Ocean Engineering, 92:137-161.
[16] GONG, J., GUO, C.Y. and ZHANG, H.P. (2017) Numerical Analysis of Impeller Flow Field of Waterjet Self-Propelled Ship Model, Journal of Shanghai Jiaotong University, 03:326-331.
[17] Philippe R. Spalart. (2009) Detached-Eddy Simulation, Annual Review of Fluid Mechanics, 41:181-202.

Reproducibility of optical coherence tomography imaging based measurements of the fibrous cap thickness in lipid-enrich atheroma

†Chunliu He[1,2], *Zhiyong Li[1,2,3], Jiaqiu Wang[3], Yuxiang Huang[1,2], Tongjing Zhu[1,2], Yuehong Miao[1,2]

[1] State Key Laboratory of Bioelectronics, Southeast University, Nanjing 210096, P.R. China.

[2]School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, P.R. China.

3School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4000, Australia.

† Presenting author: Chunliu He, Southeast University, Nanjing China;

*Corresponding authors: Zhiyong Li, Southeast University, Nanjing China, email: zylicam@gmail.com

Abstract

Rupture of the thin fibrous cap of the atherosclerotic plaque is the primary cause of acute coronary syndrome accounting for more than half of all cardiovascular deaths. Fibrous cap thickness (FCT) is seen as critical to plaque vulnerability. In this study, the intra-observer reproducibility of FCT and the correlation analysis between FCT and intravascular optical coherence tomography (IVOCT) images features were implemented to find the relationship between FCT of lipid-enrich plaques and images information by two observers. We performed IVOCT pullbacks in consecutive series on 20 patients and selected 102 images containing lipid-enrich plaques. Firstly, region of interests (ROIs) were extracted by an unsupervised fuzzy c means clustering (FCM) stage. Then, 32 features, which are associated with the structural and biochemical changes of tissue within the ROIs, were carried out using First order statistics (FOS), Gray level co-occurrence matrix (GLCM), Neighborhood gray tone difference matrix (NGTDM), Invariant moment (IM), Fractal dimension (FD) and Shape features (SF). Finally, the FCT in grayscale IVOCT images were manually measured by two independent observers. The intraclass correlation coefficient (ICC) was 0.80 for two different observers. The image features with ROI region and FCT showed a high correlation coefficient for both observers (r=0.88, p<0.001 and r= 0.91, p<0.001, respectively). The results suggest that the features of IVOCT images based FCT measurements may be useful to quantify the plaque cap thickness and vulnerability.

Keyword: atherosclerotic plaque; intravascular optical coherence tomography; fibrous cap thickness; coronary plaque vulnerability; intra-observer reproducibility

1. Introduction

Coronary atherosclerotic plaque rupture is a major cause of acute coronary syndrome (ACS) [1-3]. Thin-capped fibroatheroma (TCFA) is recognized as a precursor for plaque rupture. The pathologic features of TCFA are a large lipid-enrich necrotic core (the maximum lipid arc>90^o), a thin fibrous cap, and macrophage infiltration into the cap [2-8]. Postmortem studies have shown that a fibrous cap thickness (FCT) (<65um) prone to rupture, the critical threshold was widely accepted [9-11]. The composition and morphology of atherosclerotic plaques are considered to be more important in determining the risk of acute syndromes than the degree of

luminal stenosis [12]. Therefore, detection and quantification of FCT of lipid-enrich atherosclerotic plaque are important for the assessment of plaque vulnerability in order to prevent acute events and monitor interventional treatments.

Intravascular imaging modalities such as intravascular ultrasound (IVUS) and angiography do not have ability to accurately quantify some of the critical components of a vulnerable plaque such as FCT and macrophage content. Intravascular optical coherence tomography (IVOCT), however, is a unique high axial resolution (~10μm) imaging modality capable of characterizing these important morphological features of atherosclerotic plaque. IVOCT has demonstrated its capacity in the identification and quantification of FCT in clinical practice [8, 13, 14].

According to the published consensus standards for IVOCT images, the plaque lipid core is a signal-poor region within an atherosclerotic plaque, with poorly delineated borders, and little or no signal backscattering. In contrast, the fibrous cap has a relatively homogeneous signal with high backscattering. Several semi-automatic and fully-automatic methods have been used to segment lipid and fibrous components by a supervised segmentation based on pixels [15, 16]. The two major drawbacks that hinder such image analysis are: (1) the procedure is cumbersome and time-consuming because of the large number of data points, and (2) manual segmentation as the gold standard are subject to a certain degree of variability between different analysts. Therefore, an unsupervised method based on FCM algorithm was introduced in the study to resolve the poorly delineated borders of the lipid core.

The purpose of this study was to analyze reproducibility of FCT measurements *in vivo*, which were achieved by two independent observers. In addition, we determined the correction coefficient and statistically significant between FCT and IVOCT images features that might mimic lipid-enrich coronary atherosclerosis plaques to assess influence of feature set in quantization FCT.

2. Materials and Methods

2.1 Image dataset

All 33 IVOCT clinical pullbacks of 20 patients were taken from Affiliated Drum Tower Hospital, Nanjing University between December 2015 and December 2016. The IVOCT images were acquired by using a commercially available Fourier Domain OCT (FDOCT) system (2.7F C7-XR, St. Jude Medical, St. Paul, Minnesota) and C7 Dragonfly catheter (St. Jude). The system is equipped with a near-infrared laser light source with a central wavelength of 1310 nm and full-width-at-half-maximum bandwidth of 80 nm. The imaging system provides an axial resolution ~10 um and a lateral resolution of ~30um in biological tissues. Scan parameters were set as 100 frames/sec, 54,000 A-scans/sec, pullback speed of 20 mm/sec, pullback length of ~54.2 mm. This study was approved by the institutional human ethics committee. All the patients have given explicitly informed consent. IVOCT images including lipid-enrich plaques from all pullbacks were selected from all databases. Out of these images, only segments containing lipid-enrich plaques were selected based on the published consensus standards [5] and the improvement of standard interpretation algorithm [17]. Total of 102 images were selected for analysis.

2.2 Manual measurement fibrous cap thickness

Measurement FCT of atherosclerotic plaques is difficult because of its complex structures with discrete individual components, especially lipid-enrich plaque with seriously diffuse border. Therefore, observers of measurement were familiar with related work and had a deep knowledge of American College of Cardiology clinical expert consensus document on standards. 102 images were analyzed by two expert observers using an OCT system software (LightLab Imaging Inc., Westford, Massachusetts). The representative FCT measurements of IVOCT images in lipid-enrich plaque from two observers is shown in Figure 1. For each plaque, both observers selected the same images from the IVOCT run and measured the thinnest FCT two times, from which the final measurement value of FCT was calculated by averaging.

Observer 1 Observer 2

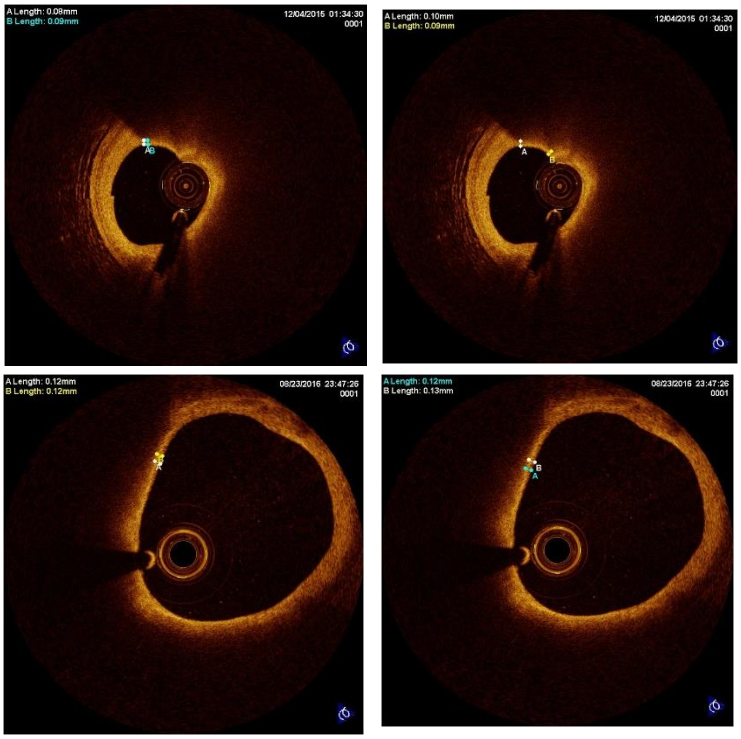


Figure 1. The representative IVOCT images for measurements of FCT by observer 1 and observer 2

3. Image Analysis

3.1 Pre-processing

Consider the IVOCT images in polar coordinates (θ, r) where θ is angle and r is depth. $I(i, j)$ represents intensity of each pixel at row i and column j. Ring-area (RA) and Lumen are automatically segmented by processing the following four steps. The results of each step are shown in Figure 2.

1. Remove guide-wire and artifacts;

$$I\left(:,1:\underset{1\le j\le n}{max}\left(\bar{I}\left(i,j\right)\right)+wc\right)=0 \quad (1)$$

$$I\left(\underset{1\le i\le m}{min}\left(\bar{I}\left(i,j\right)\right)\pm wr,:\right)=0 \quad (2)$$

where $\bar{I}\left(i,j\right)$ refers to the average of pixel value, and wc and wr are the thresholds. The parameter values $wc=50$ and $wr=10$ used in the paper were determined based on catheter size.

2. Binarization images processed by adaptive threshold OSTU's method algorithm and by morphological connect neighborhood and area constraint [18];
3. Lumen was automatically segmented by connecting the nonzero pixels, interpolating pixels of full zero row, and then expanding lumen to 1mm to take into account the limited penetration depth of OCT system;
4. The polar images were subsequently converted to a cartesian coordinate in order to reconstruct an image that preserved the true morphology of visualization.

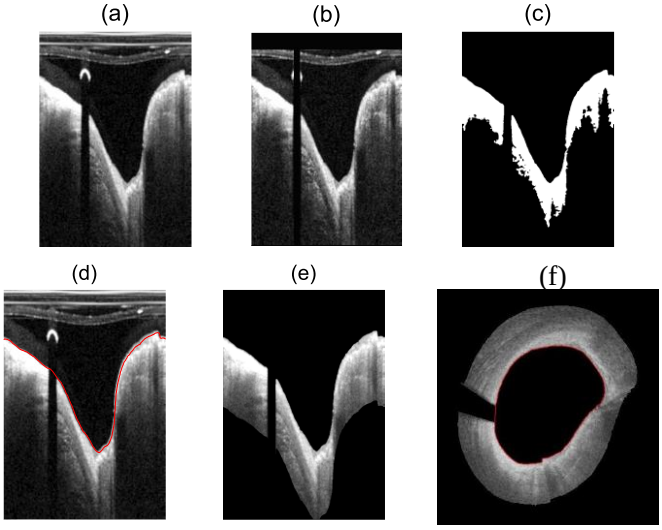


Figure 2. Illustration of the using fully-automated segmentation procedure. Image (a) shows original raw polar domain image; Image (b) show the guide-wire and catheter artifacts remove result partially; Image (c) illustrates the application of the Otsu's method, morphological operations and the area constrain; Image (d) shows lumen segmentation result; Image (e) and (f) show the RA segmentation results before and after scan-conversion respectively.

3.2 Region of interest (ROI) extraction

Compared to fibrous cap, necrotic lipid core exhibits a lower signal density and a more heterogeneous back-scattering [19, 20]. Lipid core area has the following major characteristics: diffusely bordered, signal-poor regions with overlying signal-rich bands. In this paper, FCM method was selected to extract the cap of fibrous components [21]. Once the cap of fibrous components was segmented, the lipid core borders were subsequently obtained by arc angle of lumen contours. In the paper, the research problem with the green contour model were formulated as shown in Figure 3(a). The contour of a cap of fibrous component in the 2-D image was represented by two curves along *x-axis* and y-*axis* in Figure 3 (b) and (c). A simple polynomial curve fitting algorithm was proposed in order to smooth two curves. Next, the key

problem was to locate the two points pointed by the white arrow to extract ROI. We used the simple geometric constraints: the catheter center set as an origin, four equal regions were divided, the same arc angle in Figure 3 (b) and (c) are the points indicated by white arrow in Figure 3 (a). Figure 4 gives three representative results of the ROI in different pullbacks.

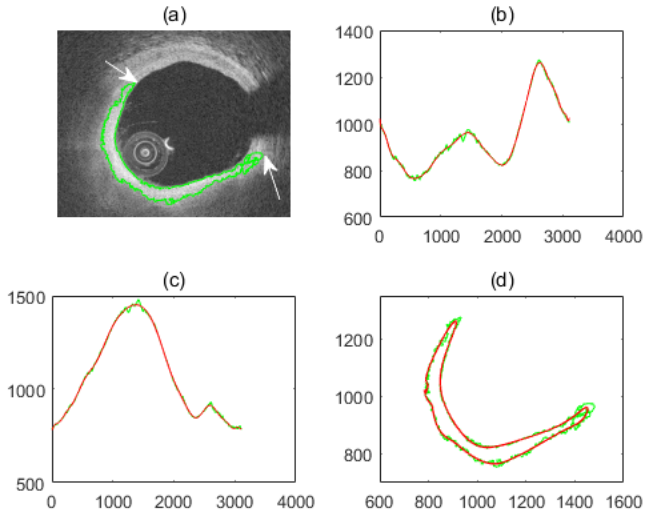


Figure 3. The cap of fibrous component extraction algorithm using the FCM combination with geometric constraint. The green contour of image (a) shows the cap segmentation result based on the FCM algorithm. Image (b) and (c) display the fitting results using polynomial curve fitting algorithm. The green line and red line represent the row and column index value before and after polynomial curve fitting, respectively. Image (d) shows the cap contours before and after polynomial curve fitting.

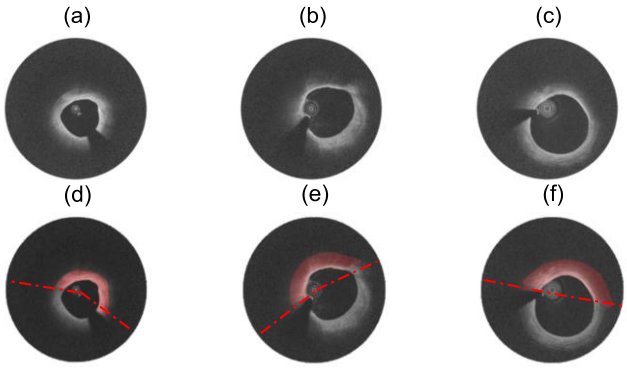


Figure 4. Representative results of the ROI on three frames from different pullbacks. Image (a), (b) and (c) show the log image with lipid-enrich plaques in the cartesian coordinate. Image (c), (d) and (f) show the ROIs (red regions) corresponding to the image (a), (b) and (c), respectively.

3.3 Feature extraction

Texture features and shape parameters were extracted from ROIs. Texture refers to the spatial interrelationship and arrangement of the basic elements of an image [22, 23]. Texture features have to be derived from the gray images because the spatial interrelationships and the arrangements of the image pixels are seen as variations in the intensity patterns or gray tones visually. Although it is easy for humans to recognize different kinds of textures, it is quite a

difficult task to define and interpret the textures automatically by computer algorithm. Shape is also an important feature for medical image [12]. In this paper, six different feature sets composing of a total 32 features were listed in table 1. The implementation details for the texture feature and shape parameters and referred papers are shown below:

Table 1. Feature sets information and corresponding references.

Feature sets	Feature name	Reference
FOS	mean, variance, median, skewness, kurtosis	[23]
GLCM	correlation, contract, dissimilarity, energy, entropy, homogeneity, maximum probability.	[24]
NGTDM	busyness, contrast, complexity, coarseness, texture length	[25]
IM	I1,I2,I3,I4,I5,I6,I7	[26]
FDTA	H^1 , H^2 , H^3 , H^4	[23] [27]
SP	eccentricity, perimeter, majoraxislength (mal), minoraxislength (mil)	[23]

3.4 Statistical analysis

Inter-observer agreement and intra-observer reproducibility estimates were analyzed using the two paired t-test, intraclass correlation coefficient (ICC), and Bland Altman analyses estimating 95% limits of agreement (LOA). LOA was defined as mean 1.96±SD of absolute difference by Bland–Altman method. Generally, an ICC <0.4, between 0.4-0.75, and >0.75 indicates poor, moderate, and excellent agreement, respectively [28]. Initially, univariate linear regressions were performed between each thickness measure and IVOCT image features. Direct linear regression was appropriate here, because the IVOCT images sampling interval was far more than 0.2 mm and the data at nearby points were independent. In addition, for each of the 102 thickness measures, multivariate linear regressions were performed against all 32 image features. Multiple correlation coefficient between variables was estimated using Pearson's correlation coefficient (*r*). For all test, a two tailed *p* value<0.05 was considered statistically significant. All statistical analysis was performed with SPSS statistical software (IBM SPSS Statistics for Windows, Version 19.0. IBM Corp., Armonk, New York).

4. Result

The ICCs (0.99 for observer 1 and 0.99 for observer 2) of FCT measurement showed excellent agreement and reproducibility (ICC=0.80 between observer 1 and 2). FCT11, FCT12 represented two measurement results by observer 1, FCT21, FCT22 were the measurement results by observer 2, mFCT1 and mFCT2 were the mean by observer 1 and 2. The Bland–Altman plots showed LOAs of different FCT measurements from two observers (Figure 5). The LOAs for FCT11 vs FCT21, FCT11 vs FCT22, FCT12 vs FCT21 and FCT12 vs FCT22 were 110, -55um (*p*=0.016), 130, -43um (*p*=0.03), 100, -66um (*p*=0.008) and 110, -53um (*p*=0.015), respectively.

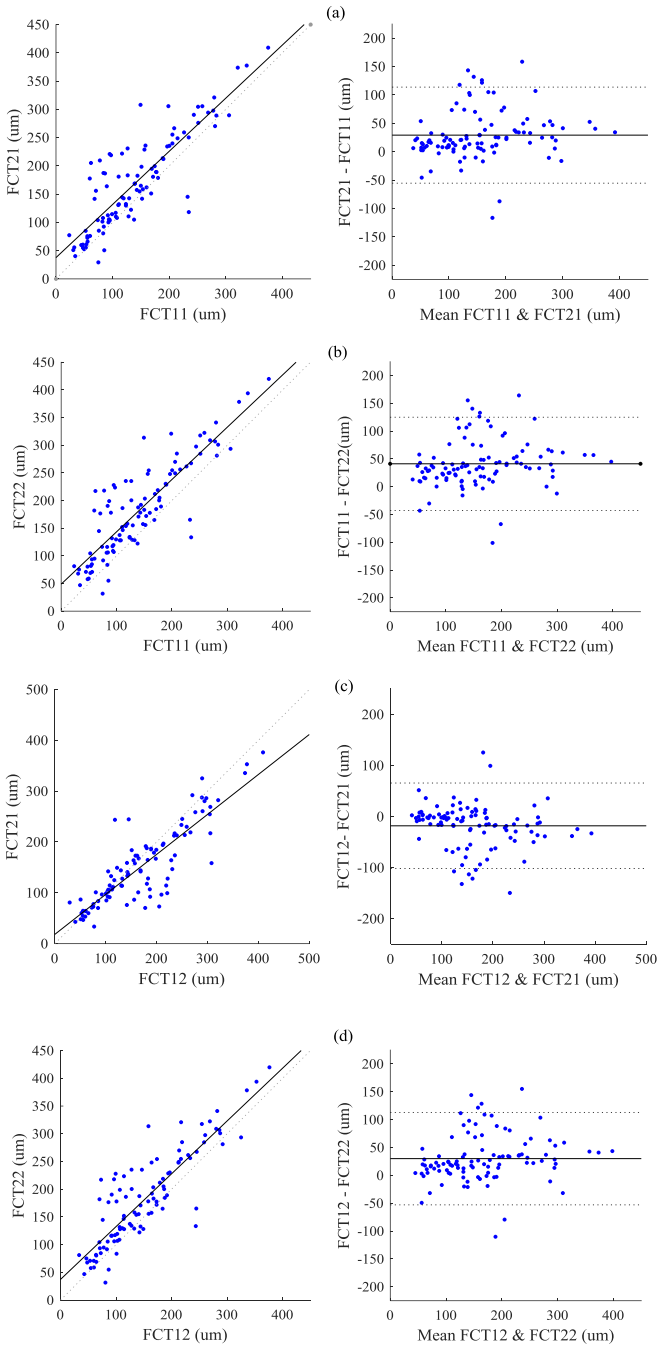


Figure 5. Comparison of the FCT measured by observer 1 versus observer 2 (left panels) Bland-Altman test for two observers in measurement of FCT (right panels).

Table 2. The correction coefficient of mFCT and univariate image features by two observers

	mFCT1		mFCT2			mFCT1		mFCT2	
feature name	*r*	*p*-value	*r*	*p*-value		*r*	*p*-value	*r*	*p*-value
mean	0.62	<0.001	0.64	<0.001	texture length	0.39	<0.001	0.38	<0.001
variance	0.48	<0.001	0.49	<0.001	I1	0.53	<0.001	0.54	<0.001
median	0.45	<0.001	0.48	<0.001	I2	0.13	**0.1848**	0.22	0.03
skewness	0.48	<0.001	0.51	<0.001	I3	0.33	<0.001	0.33	<0.001
kurtosis	0.52	<0.001	0.53	<0.001	I4	0.03	**0.73**	0.05	**0.59**
correlation	0.59	<0.001	0.61	<0.001	I5	0.01	**0.97**	0.13	**0.19**
contract	0.58	<0.001	0.61	<0.001	I6	0.36	<0.001	0.28	<0.001
dissimilarity	0.59	<0.001	0.62	<0.001	I7	0.47	<0.001	0.50	<0.001
energy	0.60	<0.001	0.64	<0.001	H^1	0.16	**0.12**	0.08	**0.43**
entropy	0.49	<0.001	0.48	<0.001	H^2	0.16	**0.12**	0.08	**0.41**
homogeneity	0.33	<0.001	0.35	<0.001	H^3	0.45	<0.001	0.44	<0.001
maximum probability	0.34	<0.001	0.34	<0.001	H^4	0.01	**0.95**	0.06	**0.56**
busyness	0.23	0.01	0.25	0.0112	eccentricity	0.17	**0.09**	0.27	<0.001
contrast	0.37	<0.001	0.36	<0.001	perimeter	0.08	**0.42**	0.16	**0.10**
complexity	0.06	**0.52**	0.16	**0.1**	mal	0.16	**0.11**	0.26	0.01
coarseness	0.39	<0.001	0.39	<0.001	mil	0.30	0.0021	0.41	<0.001

Table 2 reports statistically significant (*p*) and Pearson correction coefficient (*r*) between univariate feature and mFCT. The correction coefficient is generally low, where the lowest and highest values are 0.62 (mean) and 0.01 (I5 and H^4) from observer 1 and 0.64 (mean and energy) and 0.05 (I4) from observer 2. Bold *p*-values represent no statistically significant between two

variable values.

Similarly, Table 3 shows statistically significant (*p*) and Pearson correction coefficient (*r*) between multivariate feature sets and mFCT. Statistically significant results were observed in both two groups i.e. individual group feature set and the fusion feature set. The lowest correction coefficient of individual group appeared at shape parameter group (feature set 6), which were 0.48 and 0.58 for observer 1 and 2, respectively. The phenomenon was in turn confirmed in Table 2 that the correction coefficient of four shape parameters were overall lower than others. In the contrast, the highest correction coefficient of individual group was observed in feature set 2, which were 0.78 and 0.80, respectively. For both groups, Pearson correlation coefficient of the fusion feature sets for observers 1 and 2 were 0.88 and 0.91, which better than any individual group feature set.

Table 3. The correction coefficient of mFCT and multivariate image features by two observers

Feature set	mFCT1		mFCT2	
	r	*p*-value	*r*	*p*-value
FOS	0.67	<0.001	0.68	<0.001
GLCM	0.78	<0.001	0.80	<0.001
NGTDM	0.68	<0.001	0.74	<0.001
FD	0.74	<0.001	0.72	<0.001
IM	0.52	<0.001	0.62	<0.001
SP	0.48	<0.001	0.58	<0.001
Fusion feature sets	0.88	<0.001	0.91	<0.001

5. Discussion

It is an important role of FCT as indicators of vulnerable plaques which could potentially guide appropriate surgical treatment such as percutaneous coronary intervention (e.g., balloon angioplasty or stent placement). Therefore, there is a strong desire to treat these lesions before they cause harm. The reliable examination of these indicators of atherosclerotic plaques will ultimately determine the clinical value of IVOCT, depending on the application of meaningful and reproducible methods. The main findings of the present study are the excellent inter-observer agreement of the manual assessment of FCT and excellent intra-observer reproducibility in the FCT measurement. In addition, the high correction between the feature of ROIs and mFCT measured by two observers, which show that IVOCT image feature is able to provide more information in quantization FCT to promote both the computer-aided routine clinical use and analysis of large-scale data sets from clinical trials in vulnerable plaque.

The current accepted universal method for assessing FCT *in vivo* using IVOCT images is based on single measurement of the thinnest portion of the fibrous cap [14, 29]. In practice, the extensive clinical image data *in vivo* were usually analyzed manually by expert analysts. Indeed, the excellent inter-observer agreement of IVOCT images to measure the FCT manually, have been previously reported. Kim et al. [30] performed first *in vivo* investigation in the inter-observer agreement (ICC=0.99) and intra-observer reproducibility (ICC=0.49) of FCT by 4 independent observers. Subsequently, Gerbaud et al [31] reported intra-observer reproducibility

of FCT was moderate (ICC=0.48). In the present study, excellent inter-observer agreement resulted for FCT measurement, with ICC of 0.99 was reached in the analysis and was similar to literature previously. Greatly, excellent intra-observer reproducibility (ICC=0.80) was achieved for FCT measurement, higher than the result of the previous mentioned studies. Recently, Kini et al [17] studied intra-observer reproducibility before and after developing the lesion assessment criteria with 170 pullbacks. The result shown that a significantly higher level for FCT measurement, with ICC of 0.82 compared with the observed in our study of reproducibility *in vivo* measurement. However, these independent observers extensively learned the development of standard interpretation criteria formulated which significantly provided the level of intra-observer reproducibility. The lower intra-observer reproducibility in our study may, in part, be explained by the heterogeneity in the coronary plaques imaged. Indeed, the patients in our study are more likely to have lipid-enrich, complex plaques, with a higher potential for intra-observer variability. Although learning the standard interpretation algorithm, a limited pullback data may cause a low learning outcomes result. Therefore, more data are more likely to represent a true reproducibility value, based on the current commercial available IVOCT systems.

Although others' and our studies had been certified the FCT measurement may be repeatable by independent observer manually, few literates focus on the interrelationship between IVOCT image feature information and FCT. Such an idea will help in enhancing the significance of noninvasive coronary artery tests in the identification of FCT and assessment the risk factors of stroke. Thus, in the study, we first analyzed the correct coefficient and statistically significant between FCT and the six group image features based on the priori knowledge that the more higher the correction coefficient, the better elucidate the texture feature was used to quantify FCT.

The results in this study (Table 2 and Table 3) indicate significant relationships between feature sets and FCT. The *r* value of the univariate regressions indicate that only several single texture feature factor are dominant in determining FCT, there are mean, contract, dissimilarity, energy, entropy, homogeneity, maximum probability (Table 2). On the other hand, the individual feature set of the multivariate regressions are all highly significant, and the correlation coefficients are substantially higher as well (Table 3). Thus, thickening seems to be influenced by multiple aspects of the texture feature and shape parameters. This is to be expected, the texture features are postulated to act through their influence on the spatial interrelationships and arrangement of the gray image, and it is reasonable that each of these FCT (the minimum distance implied in the spatial arrangement) would be influenced by texture feature and shape parameters. Best feature sets were the GLCM feature set, followed by the FDTA. In general, all individual feature set performed in a range of about 0.52-0.78 and 0.62-0.80 for observer 1 and 2, except of the shape parameters that performed much worse. In order to enhance the influences of feature set, the six feature sets were combined, by connecting the feature one by one. Fusing results of the six different feature sets, improved the correction results obtained by the individual feature sets, reaching an average correction coefficient of 0.88 and 0.91 for the observer 1 and 2. The benefits of fusion results are more obvious in the case where there is no dominant best feature sets, as the case with the features extracted from the lipid-enrich plaque images in this study. It

is noteworthy in this respect that the signs of the regression coefficients in the univariate and multivariate regressions in Tables 2 and 3 are consistent.

Study Limitations

In multivariate regression analysis, correlation among the independent variables is one common problem. The problem may be an influential factor if the primary purpose of the regression is to identify important explanatory variables that might play a causal role. The estimated regression coefficients for such correlated variables can be different. This problem was not involved and discussed in our case. Feature selection method with deleting the possible correlations between the independent variables are suggested in the future research.
In computer vison analysis, efficiency measured by the computational time is another common problem. Computational times for preprocessing, lumen segmentation, scan-conversion and ROI extraction were recorded by matlab code, especially scan-conversion spent a long time (two hours for 271 images) in the study. As such, further coding and implementation in a faster language (e.g. C/C++) would significantly reduce computational time, possibly achieving the analysis of a multiple IVOCT images in a time a few minutes.

Lack of histology data as the golden standard in the FCT measurement is the third problem. Given that IVOCT manual FCT measurement of atherosclerotic plaques is subject to some inter-observer variability, the use of a third reader is always required in case of disagreement between two readers. As a matter of fact, only FCT measurement using a large series of histological samples would be able to give more objective and detailed results. However, even if histology can provide a stronger ground truth, the correct registration with IVOCT images can be a challenge due to histological slice thickness and helicoidal IVOCT data acquisition [32, 33]. Therefore, a large amount of histological data would be required to achieve enough statistical analysis result, which is not currently available.

Conclusion

We discussed the variabilities between observers for quantifying the FCT in human coronary arteries based on IVOCT using manual measurement. Intra-reproducibility result demonstrate that FCT was repeated by manual measurements in lipid-enrich atherosclerotic plaque. In addition, we analyzed the relationship between FCT and image feature. The regression result demonstrated the fusion feature played an important role in quantification FCT for online identification of high-risk plaques.

References

[1] Yonetsu T, Kakuta T, Lee T, Takahashi K, Kawaguchi N, Yamamoto G, et al. In vivo critical fibrous cap thickness for rupture-prone coronary plaques assessed by optical coherence tomography. *European Heart Journal* 2011; 32(10): 1251-1259.

[2] Habara M, Nasu K, Terashima M, Ko EH, Yokota D, Ito T, et al. Impact on optical coherence tomographic coronary findings of fluvastatin alone versus fluvastatin plus ezetimibe. *American Journal of Cardiology* 2014; 113(4): 580-587.

[3] Cardoso L, Weinbaum S. Changing views of the biomechanics of vulnerable plaque rupture: a review. *Annals of Biomedical Engineering* 2014; 42(2): 415-431.

[4] Di Vito L, Yoon JH, Kato K, Yonetsu T, Vergallo R, et al. Comprehensive overview of definitions for optical coherence tomography-based plaque and stent analyses. *Coronary Artery Disease* 2014; 25(2): 172-185.

[5] Falk E, Nakano M, Bentzon JF, Finn AV, Virmani R. Update on acute coronary syndromes: the pathologists' view. *European Heart Journal* 2013; 34(10): 719-746.

[6] Fujii K, Hao H, Shibuya M, Imanaka T, Fukunaga M, Miki K, et al. Accuracy of OCT, grayscale IVUS, and their combination for the diagnosis of coronary TCFA: an ex vivo validation study. *JACC: Cardiovasc Imaging* 2015; 8(4): 451-460.

[7] Jang IK, Tearney GJ, MacNeill B, Takano M, Moselewski F, Iftima N, et al. In vivo characterization of coronary atherosclerotic plaque by use of optical coherence tomography. *Circulation* 2005; 111(12): 1551-1555.

[8] Virmani R, Burke AP, Farb A, Kolodgie FD. Pathology of the vulnerable plaque. *Journal of the American College of Cardiology* 2006; 47(8 Suppl): 13-18.

[9] Tanaka A, Imanishi T, Kitabata H, Kubo T, Takarada S, Tanimoto T, et al. Lipid-rich plaque and myocardial perfusion after successful stenting in patients with non-ST-segment elevation acute coronary syndrome: an optical coherence tomography study. *European Heart Journal* 2009; 30(11): 1348-1355.

[10] Tearney GJ, Jang IK, Bouma BE. Optical coherence tomography for imaging the vulnerable plaque. *Journal of Biomedical Opticst* 2006; 11(2): 021002.

[11] Tian J, Dauerman H, Toma C, Samady H, Itoh T, et al. Prevalence and characteristics of TCFA and degree of coronary artery stenosis: an OCT, IVUS, and angiographic study. *Journal of the American College of Cardiology* 2014; 64(7): 672-680.

[12] Wang Z, Liu N, Zhang L, Li X, Han X, Peng Y, et al. Real-time elastography visualization and histopathological characterization of rabbit atherosclerotic carotid arteries. *Ultrasound in Medicine and Biology* 2016; 42(1): 176-184.

[13] Jang IK, Bouma BE, Kang DH, Park SJ, Park SW, Seung KB, et al. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. *Journal of the American College of Cardiology* 2002; 39(4): 604-609.

[14] Kubo T, Imanishi T, Takarada S, Kuroi A, Ueno S, Yamano T, et al. Assessment of culprit lesion morphology in acute myocardial infarction: ability of optical coherence tomography compared with intravascular ultrasound and coronary angioscopy. *Journal of the American College of Cardiology* 2007; 50(10): 933-939.

[15] Wang Z, Chamie D, Bezerra HG, Yamamoto H, Kanovsky J, Wilson DL, et al. Volumetric quantification of fibrous caps using intravascular optical coherence tomography. *Biomedical Optics Express* 2012; 3(6): 1413-1426.

[16] Athanasiou L S, Bourantas C V, Rigas G, et al. Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of Biomedical Opticst*, 2014, 19(2): 026009.

[17] Kini AS, Vengrenyuk Y, Yoshimura T, Matsumura M, Pena J, Baber U, et al. Fibrous cap thickness by optical coherence tomography in vivo. *Journal of the American College of Cardiology* 2017; 69(6): 644-657.

[18] Koga S, Ikeda S, Yoshida T, Nakata T, Takeno M, Masuda N, et al. Elevated levels of systemic pentraxin 3 are associated with thin-cap fibroatheroma in coronary culprit lesions assessment by optical coherence tomography and intravascular ultrasound. *JACC: Cardiovascular Interventions* 2013; 6(9): 945-954.

[19] Prati F, Guagliumi G, Mintz GS, Costa M, Regar E, Akasaka T, et al. Expert review document part 2: methodology, terminology and clinical applications of optical coherence tomography for the assessment of interventional procedures. *European Heart Journal* 2012; 33(20): 2513-2520.

[20] Prati F, Regar E, Mintz GS, Arbustini E, Di Mario C, Jang IK, et al. Expert review document on methodology, terminology, and clinical applications of optical coherence tomography: physical principles, methodology of

image acquisition, and clinical application for assessment of coronary arteries and atherosclerosis. *European Heart Journal* 2010; 31(4): 401-415.

[21] Chamie D, Bezerra HG, Attizzani GF, Yamamoto H, Kanaya T, Stefano GT, et al. Incidence, predictors, morphological characteristics, and clinical outcomes of stent edge dissections detected by optical coherence tomography. *JACC: Cardiovascular Interventions* 2013; 6(8): 800-813.

[22] Kato K, Yonetsu T, Jia HB, Abtahian F, Vergallo R, Hu SN, et al. Nonculprit coronary plaque characteristics of chronic kidney disease. *Circulation:Cardiovascular Imaging* 2013; 6(3): 448-471.

[23] Christodoulou CI, Pattichis CS, Pantziaris M, Nicolaides A. Texture-based classification of atherosclerotic carotid plaques. *IEEE Transactions on Medical Imaging* 2003; 22(7): 902-912.

[24] Kalyan K, Jakhia B, Lele RD, Joshi M, Chowdhary A. Artificial neural network application in the diagnosis of disease conditions with liver ultrasound images. *Advances in Bioinformatics* 2014; 708279.

[25] Araki T, Ikeda N, Shukla D, Jain PK, Londhe ND, Shrivastava VK, et al. PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: a link between carotid and coronary grayscale plaque morphology. *Computer Methods and Programs in Biomedicine* 2016; 128: 137-158.

[26] Yoshikawa D, Ishii H, Kurebayashi N, Sato B, Hayakawa S, Ando H, et al. Association of cardiorespiratory fitness with characteristics of coronary plaque: Assessment using integrated backscatter intravascular ultrasound and optical coherence tomography. *International Journal of Cardiology* 2013; 162(2): 123-128.

[27] Qiu WB, Chen Y, Li X, Yu YY, Cheng WF, Tsang FK, et al. An open system for intravascular ultrasound imaging. *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control* 2012; 59(10): 2201-2209.

[28] Radu MD, Yamaji K, Garcia-Garcia HM, Zaugg S, Taniwaki M, Koskinas KC, et al. Variability in the measurement of minimum fibrous cap thickness and reproducibility of fibroatheroma classification by optical coherence tomography using manual versus semi-automatic assessment. *EuroIntervention* 2016; 12(8): 987-997.

[29] Kume T, Akasaka T, Kawamoto T, Okura H, Watanabe N, Toyota E, et al. Measurement of the thickness of the fibrous cap by optical coherence tomography. *American Heart Journal* 2006; 152(4).

[30] Kim SJ, Lee H, Kato K, Yonetsu T, Xing, L, Zhang S, and Jang IK. Reproducibility of in vivo measurements for fibrous cap thickness and lipid arc by OCT. *JACC: Cardiovascular Imaging* 2012, 5(10): 1072-1074.

[31] Gerbaud E, Weisz G, Tanaka A, Kashiwagi M, Shimizu, Wang L, et al. Multi-laboratory inter-institute reproducibility study of IVOCT and IVUS assessments using published consensus document definitions. *European Heart Journal Cardiovascular Imaging* 2016, 17(7):756-764.

[32] Rieber J, Meissner OG, Babaryka S, Reim M, Oswald A, Koenig TM et al. Diagnostic accuracy of optical coherence tomography and intravascular ultrasound for the detection and characterization of atherosclerotic plaque composition in ex-vivo coronary specimens: a comparison with histology *Coronary Artery Disease* 2006, 17(5): 425–430.

[33] Meissner OA, Rieber J, Babaryka G, Oswald M, Reim S, Siebert U, Redel T et al. Mueller- Lisse U., Intravascular optical coherence tomography: comparison with histopathology in atherosclerotic peripheral artery specimens, *Journal of vascular & interventional radiology* 2006, 17(2): 343–349.

Linear Dynamic Reanalysis Using Frequency-Shift Combined Approximations

***Guikai Guo[1] , and †Fei Cheng[1]**

[1]School of Mechanical Science and Engineering, Jilin University, PRC.

*Presenting author: ggk@jlu.edu.cn

†Corresponding author: chengfei@jlu.edu.cn

Abstract

In the structural dynamic optimization procedure, many repeated analyses are conducted to evaluate dynamic performance of successively modified structural designs. It is noted that, the reduction of degrees of freedom is more important for computational effort in dynamic problems than in static problems. This paper focuses on the reanalysis for structural dynamic problem in the framework of combined approximations (CA) method. A new procedure for structural dynamic reanalysis is developed based on iteration and inverse iteration method with frequency-shift in mode superposition method, and linear combination acceleration is also used to reduce the high computational cost of structural reanalysis. Frequency-shift factor is calculated first, and then combined approximations method corrected by the given factor allows calculating higher modes accurately. After higher modes are obtained, mode expansion and dynamic response can be described accurately. Numerical example is presented to demonstrate the accuracy of the proposed method. Excellent results can be obtained when large modifications are made.

Keywords: Frequency-Shift, Combined Approximations, Mode Superposition, Dynamic Reanalysis.

Introduction

In order to make a design structure satisfy the predetermined demands, such as the structural dynamic design procedure, usually the designer will modify the structure repeatedly. The dynamic responses changed by the modifications of parameters on the structure. In the structural optimization, dynamic analyses are repeated in successively modified structure design procedure. Research of how to reduce the computational cost has made sense.

Reanalysis technology was established to evaluate responses of changed structures without complete analysis in process of design and optimization[1]. Reanalysis of structure for displacements and stresses have been discussed since the 20th century[2]. Combined Approximations (CA) approach is one of the most effective methods for solving static displacement equations[3]. After CA method was founded, extended CA methods were proposed[4]. IFU method is proposed for general low-rank local modifications, including boundary modifications and non-boundary modifications[5].

Reanalysis methods for vibration problems have been presented since the early 21st century[6]. Kirsch grafted the CA approach to solve eigenproblems [7]. Combining CA and Rayleigh quotient, an extended CA method of eigenproblem for large changes was presented

by Chen[8]. A Modified Combined Approximations(MCA) method for solving large-scale structure dynamic problem was discussed[9]. With a suitable frequency shift coefficient, FSCA approach allowed to calculate higher modes accurately[10].

Some studies in the literature have approached the structural dynamic reanalysis problem for large perturbations in the structural parameters. A method for the dynamic reanalysis of structures subjected to deterministic or stochastic loads is presented by Cacciola [11]. The CA approach, developed originally for linear static reanalysis, is also used for dynamic reanalysis of structures by Kirsch[12, 13]. The approach is based on the integration of several concepts and methods, including series expansion, reduced basis, matrix factorization, and Gram-Schmidt orthogonalizations. Based on epsilon-algorithm, the dynamic response reanalysis method has been developed by Chen[14]. In his computational process, the Neumann series expansion was used to construct the vector sequence for epsilon-algorithm iterative form.

In this study, a linear dynamic reanalysis process using FSCA method is proposed. The formulations of mode superstition based on CA method with frequency shift are expressed, and then the application of this algorithm to a truck body finite element analysis is described. Conclusions are discussed at last.

Linear dynamic analysis by mode superposition

Linear dynamic analysis consider the equations of motion for a system subjected to external dynamic forces

$$\mathbf{M}\,{}^{t}\ddot{\mathbf{u}} + \mathbf{C}\,{}^{t}\dot{\mathbf{u}} + \mathbf{K}\,{}^{t}\mathbf{u} = {}^{t}\mathbf{R} \tag{1}$$

where $\mathbf{M}$ is the mass matrix, $\mathbf{C}$ is the damping matrix and $\mathbf{K}$ is the stiffness matrix. The displacement vector ${}^{t}\mathbf{u}$, the velocity vector ${}^{t}\dot{\mathbf{u}}$, the acceleration vector ${}^{t}\ddot{\mathbf{u}}$ and the load vector ${}^{t}\mathbf{R}$ are functions of the time t.

In practical analysis, the common procedures can be solved by mode superposition method, where the equilibrium equations are transformed in to a form in which only limited modes are considered. In this approach a change of basis from the finite element nodal displacements to the eigenvectors of the generalized eigenproblem is preformed prior to the time integration. The following transformation is used.

$${}^{t}\mathbf{u} = \mathbf{\Phi}\,{}^{t}\mathbf{X} \tag{2}$$

where $\mathbf{\Phi}$ is $n \times n$ transformation matrix and components of ${}^{t}\mathbf{X}$ are the modal coordinates. With the transformation, new system stiffness, damping and mass matrices are obtained, which are much smaller scale than those in the original system. And then, a new dynamic equation is obtained.

$$\mathbf{\Phi}^{T}\mathbf{M}\mathbf{\Phi}\,{}^{t}\ddot{\mathbf{X}} + \mathbf{\Phi}^{T}\mathbf{C}\mathbf{\Phi}\,{}^{t}\dot{\mathbf{X}} + \mathbf{\Phi}^{T}\mathbf{K}\mathbf{\Phi}\,{}^{t}\mathbf{X} = \mathbf{\Phi}^{T}\,{}^{t}\mathbf{R} \tag{3}$$

where

$$\tilde{\mathbf{M}} = \mathbf{\Phi}^{T}\mathbf{M}\mathbf{\Phi} \qquad \tilde{\mathbf{C}} = \mathbf{\Phi}^{T}\mathbf{C}\mathbf{\Phi} \qquad \tilde{\mathbf{K}} = \mathbf{\Phi}^{T}\mathbf{K}\mathbf{\Phi} \qquad {}^{t}\tilde{\mathbf{R}} = \mathbf{\Phi}^{T}\,{}^{t}\mathbf{R} \tag{4}$$

The transformation matrix $\mathbf{\Phi}$ is the orthogonalized modal displacement solution of the free vibration.

$$\mathbf{I} = \mathbf{\Phi}^T \mathbf{M} \mathbf{\Phi} \qquad \mathbf{\Lambda} = \mathbf{\Phi}^T \mathbf{K} \mathbf{\Phi} \tag{5}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{\Lambda}$ is the spectral matrix. Then the equilibrium equation correspond to the orthogonalized modal displacements is described.

$${}^t\ddot{\mathbf{X}} + \mathbf{\Phi}^T \mathbf{C} \mathbf{\Phi}\, {}^t\dot{\mathbf{X}} + \mathbf{\Lambda}\, {}^t\mathbf{X} = \mathbf{\Phi}^T\, {}^t\mathbf{R} \tag{6}$$

When damping effects are not considered, Eq.(6) becomes

$${}^t\ddot{\mathbf{X}} + \mathbf{\Lambda}\, {}^t\mathbf{X} = \mathbf{\Phi}^T\, {}^t\mathbf{R} \tag{7}$$

The individual equation is of the form

$${}^t\ddot{\mathbf{X}}_i + \lambda_i\, {}^t\mathbf{X}_i = \boldsymbol{\varphi}_i^T\, {}^t\mathbf{R} \tag{8}$$

The initial conditions at time 0 are obtained by Eq.(9).

$${}^0\mathbf{X} = \mathbf{\Phi}^T \mathbf{M}^{(0)} \mathbf{u} \qquad {}^0\dot{\mathbf{X}} = \mathbf{\Phi}^T \mathbf{M}^{(0)} \dot{\mathbf{u}} \tag{9}$$

Mode reanalysis by FSCA method

In practice, natural frequency has the same mean with eigenvalue in mathematics. The equation of the first m eigenvalues and eigenvectors can be expressed:

$$\underset{n\times n}{\mathbf{K}^{(0)}}\ \underset{n\times m}{\mathbf{\Phi}^{(0)}} = \underset{n\times n}{\mathbf{M}^{(0)}}\ \underset{n\times m}{\mathbf{\Phi}^{(0)}}\ \underset{m\times m}{\mathbf{\Lambda}^{(0)}} \tag{10}$$

where $\mathbf{\Lambda}^{(0)}$ denotes the matrix of the first m eigenvalues and $\mathbf{\Phi}^{(0)}$ is the corresponding matrix of first m eigenvectors, *n* is DoFs for the initial system. Assuming there are changes in the stiffness and mass matrices, respectively.

$$\mathbf{K} = \mathbf{K}^{(0)} + \Delta\mathbf{K} \qquad \mathbf{M} = \mathbf{M}^{(0)} + \Delta\mathbf{M} \tag{11}$$

The eigenproblem of the changed structure can be rearranged:

$$\underset{n\times n}{\mathbf{K}}\ \underset{n\times m}{\mathbf{\Phi}}\ \underset{m\times m}{\mathbf{\Lambda}^{-1}} = \underset{n\times n}{\mathbf{M}}\ \underset{n\times m}{\mathbf{\Phi}} \tag{12}$$

where $\mathbf{\Lambda}$ denotes the matrix of the first m eigenvalues and $\mathbf{\Phi}$ is the corresponding matrix of first m eigenvectors for the changed structure.

Eq.(12) is rearranged using a frequency-shift factor:

$$\underset{n\times m}{\mathbf{\Phi}} = [(\mathbf{M} - \mu^{-1}\mathbf{K})^{-1}\ \underset{n\times n}{\mathbf{K}}]\ \underset{n\times m}{\mathbf{\Phi}}\ [\mathbf{\Lambda}^{-1} - \mu^{-1}\ \underset{m\times m}{\mathbf{I}}\ ] \tag{13}$$

Given an initial $\mathbf{\Phi}^{(i)}$, we can compute $\mathbf{\Phi}^{(i+1)}$ by solving iterative formula as Eq.(14).

$$\underset{n\times m}{\mathbf{\Phi}^{(i+1)}} = (\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}}\underset{n\times m}{\mathbf{\Phi}^{(i)}}(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I}) \tag{14}$$

Assuming that a linear expression of $\mathbf{\Phi}^{(i)}$, where $i=0,1,\cdots,s-1$, can be close to the exact solutions, the linear expression is given:

$$\begin{aligned}
\underset{n\times m}{\mathbf{\Phi}_c} &= a_0\underset{n\times m}{\mathbf{\Phi}^{(0)}} + a_1\underset{n\times m}{\mathbf{\Phi}^{(1)}} + a_2\underset{n\times m}{\mathbf{\Phi}^{(2)}} + \cdots + a_{s-1}\underset{n\times m}{\mathbf{\Phi}^{(s-1)}} \\
&= a_0\underset{n\times m}{\mathbf{\Phi}^{(0)}} + a_1(\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}}\underset{n\times m}{\mathbf{\Phi}^{(0)}}(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I}) + a_2((\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}})^2\underset{n\times m}{\mathbf{\Phi}^{(0)}}(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I})^2 \\
&\quad + \cdots + a_{s-1}((\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}})^{s-1}\underset{n\times m}{\mathbf{\Phi}^{(0)}}(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I})^{s-1} \\
&= [\underset{n\times m}{\mathbf{\Phi}^{(0)}}, (\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}}\underset{n\times m}{\mathbf{\Phi}^{(0)}}, ((\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}})^2\underset{n\times m}{\mathbf{\Phi}^{(0)}}, \cdots, ((\mathbf{M}-\mu^{-1}\mathbf{K})^{-1}\underset{n\times n}{\mathbf{K}})^{s-1}\underset{n\times m}{\mathbf{\Phi}^{(0)}}] \\
&\quad \cdot\left[a_0\mathbf{I}, a_1(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I}), \quad a_2(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I})^2 \quad \cdots \quad a_{s-1}(\mathbf{\Lambda}^{-1}-\mu^{-1}\mathbf{I})^{s-1}\right]^T \\
&= \underset{n\times ms}{\mathbf{R}}\ \underset{ms\times m}{\mathbf{X}}
\end{aligned} \tag{15}$$

Premultiplying Eq.(12) by $\mathbf{R}^T$, a condensed equation is got and expressed in the following form:

$$[\mathbf{R}^T\ \underset{ms\times ms}{\mathbf{K}}\ \mathbf{R}]\ \underset{ms\times m}{\mathbf{X}} = [\mathbf{R}^T\ \underset{ms\times ms}{\mathbf{M}}\ \mathbf{R}]\ \underset{ms\times m}{\mathbf{X}}\ \underset{m\times m}{\mathbf{\Lambda}} \tag{16}$$

The matrices $[\mathbf{R}^T\ \mathbf{K}\ \mathbf{R}]$ and $[\mathbf{R}^T\ \mathbf{M}\ \mathbf{R}]$ of the condensed system are much smaller than those in the initial system. So we can calculate a new $ms\times ms$ system in Eq.(16) instead. The computing time can be greatly reduced.

Frequency shift consideration

For the purpose of improving the accuracy of the higher modes calculation and eliminate the numerical errors, the approximate modes and basis vectors are recalculated using Gram-Schmidt orthogonalizations in FSCA method

The advantage of the shift factor is that more accuracy results are obtained. In FSCA method, to improve the accuracy of higher modes calculation, the highest mode vector is chosen to generate the frequency shift factor.

$$\mu^{(i+1)} = \frac{\underset{n\times 1}{\boldsymbol{\varphi}_m^{(i)T}}\ \underset{n\times n}{\mathbf{K}}\ \underset{n\times 1}{\boldsymbol{\varphi}_m^{(i)}}}{\underset{n\times 1}{\boldsymbol{\varphi}_m^{(i)T}}\ \underset{n\times n}{\mathbf{M}}\ \underset{n\times 1}{\boldsymbol{\varphi}_m^{(i)}}} \tag{17}$$

where $\underset{n\times 1}{\varphi_m^{(i)}}, i=0,\cdots,s-1$ is the highest mode in the i th iteration. Considering the increasing computational cost for $\mu^{(i+1)}$ calculations, the Rayleigh quotient Eq.(18) is chosen for the frequency-shift factor in FSCA method instead of Eq.(17). The numerical example demonstrates that the frequency-shift factor is effective.

$$\mu = \frac{\underset{n\times 1}{\boldsymbol{\varphi}_m^{(0)T}} \underset{n\times n}{\mathbf{K}} \underset{n\times 1}{\boldsymbol{\varphi}_m^{(0)}}}{\underset{n\times 1}{\boldsymbol{\varphi}_m^{(0)T}} \underset{n\times n}{\mathbf{M}} \underset{n\times 1}{\boldsymbol{\varphi}_m^{(0)}}} \tag{18}$$

Numerical example

A truck body dynamic reanalysis numerical example, as shown in Fig.1, is given to demonstrate the accuracy of the FSCA method for large scale dynamic reanalysis. The objective is to evaluate the response of the seat for loading on the left rear wheel. The truck body contains 1896 shell and solid elements, 1944 nodes and 11664 degrees of freedom. The time step is 0.1s. The Young's modulus of the material is $E = 2.1\times10^{11}\,Pa$; the mass density is $\rho = 7.8\times10^{3}\,kg/m^{3}$; the Poisson's ratio is 0.3. Assuming that the response of the initial structure is known, the modified response has been evaluated by the FSCA approach with only 3 basis vectors.

The resulting vertical displacement, velocity and acceleration at the seat installation point are shown in Fig.2-4. It is observed that good agreement is obtained between solutions of the FSCA formulations and Lanczos formulations of the modified structure.

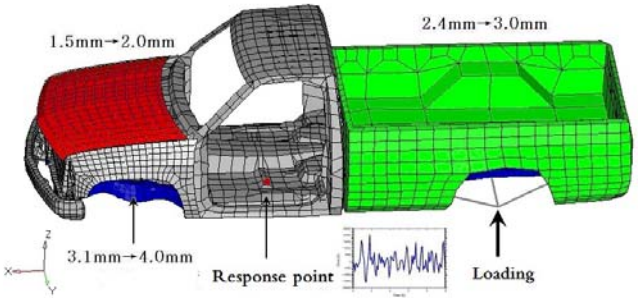


Figure 1. Modifications of truck body

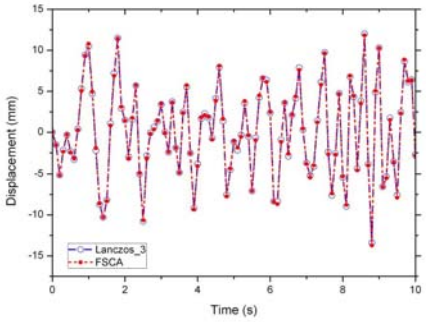


Figure 2. Comparison of displacement responses

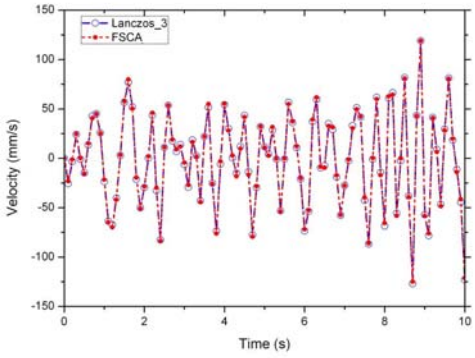


Figure 3. Comparison of velocity responses

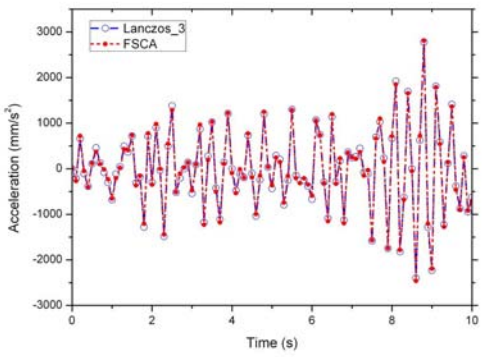


Figure 4. Comparison of acceleration responses

Conclusions

In this study, a new reanalysis technique, the FSCA method has been developed for dynamic reanalysis with respect to improve the solution accuracy in case where global large modifications are made. Numerical example is shown for the demonstrations of accuracy in this work. It can be seen that the accuracy approximate solutions were achieved with FSCA method with large changes.

When general optimization problems are considered, a lot of research has been performed to reduce the computational cost in repeated analysis of modified structures. It is expected that the FSCA method could reduce the overall computational cost in problems where repeated analyses are needed.

Acknowledgement

This work is supported by Project of National Science Foundation of China (NSFC) (Grant No. 11502092).

References

[1] Kirsch U. (2010) Reanalysis and sensitivity reanalysis by combined approximations, *Struct Multidiscip Optim* **40**, 1-15.

[2] Phansalkar SR. (1974) Matrix iterative methods for structural reanalysis, *Comput Struct* **4**, 779-800.

[3] Kirsch U. (2003) A unified reanalysis approach for structural analysis, design, and optimization, *Struct Multidiscip Optim* **2003**, 67-85.

[4] Zuo W, Yu Z, Zhao S, Zhang W. (2012) A hybrid Fox and Kirsch's reduced basis method for structural static reanalysis, *Struct Multidiscip Optim* **46**, 261-272.

[5] Huang GX, Wang H, Li GY. (2016) An exact reanalysis method for structures with local modifications, *Struct Multidiscip Optim* **54**, 499-509.

[6] Chen SH, Yang XW, Lian HD. (2000) Comparison of several eigenvalue reanalysis methods for modified structures, *Struct Multidiscip Optim* **20**, 253-259.

[7] Kirsch U. (2003) Approximate vibration reanalysis of structures, *AIAA J* **41**, 504-511.

[8] Chen SH, Yang XW. (2000) Extended Kirsch combined method for eigenvalue reanalysis, *AIAA J* **38**, 927-930.

[9] Zhang G, Nikolaidis E, Mourelatos ZP. (2009) An efficient re-analysis methodology for probabilistic vibration of large-scale structures, *J Mech Design* **131**,

[10] Xu T, Guo G, Zhang H. (2011) Vibration reanalysis using frequency-shift combined approximations, *Struct Multidiscip Optim* **44**, 235-246.

[11] Cacciola P, Impollonia N, Muscolino G. (2005) A dynamic reanalysis technique for general structural modifications under deterministic or stochastic input, *Comput Struct* **83**, 1076-1085.

[12] Kirsch U, Bogomolni M, Sheinman I. (2007) Efficient Dynamic Reanalysis of Structures, *Journal of Structural Engineering* **133**, 440-448.

[13] Kirsch U, Bogomolni M, Sheinman I. (2006) Nonlinear dynamic reanalysis of structures by combined approximations, *Computer Methods in Applied Mechanics & Engineering* **195**, 4420-4432.

[14] Chen SH, Ma L, Meng GW. (2008) Dynamic response reanalysis for modified structures under arbitrary excitation using epsilon-algorithm, *Comput Struct* **86**, 2095-2101.

Probabilistic fracture toughness prediction of composite materials

†Yan Li[1]*, Min Zhou[2]

[1]Department of Mechanical and Aerospace Engineering, California State University, Long Beach, USA.

[2]The George W. Woodruff School of Mechanical Engineering, School of Materials Science and Engineering Georgia Institute of Technology, Atlanta, GA 30332-0405, U.S.A.

*Presenting author: yan.li@csulb.edu

†Corresponding author: yan.li@csulb.edu

Abstract

One of the biggest challenges in material sensitive design is to predict the variation of key material properties such as strength and fracture toughness. It has been proved that the stochastic nature of microstructure is the primary reason for fracture toughness scatter. Although Weibull distribution has been widely used to determine the probability of material fracture, its role has been confined to fitting fracture toughness data rather than providing predictive insight of material fracture toughness and the magnitude of scatter. Besides, the Weibull parameters which are obtained through curve fitting carry little physical significance. In this paper, an integrated computational and analytical model is developed to predict fracture toughness in a statistical sense. The Weibull distribution parameters are correlated with the statistical measures of microstructure characteristics and the statistical characterization of the competition between crack deflection and crack penetration at matrix/reinforcement interfaces. The approach and model will lead to more reliable material design through microstructure tailoring.

Keywords: Microstructure, Fracture toughness, Weibull distribution

Introduction

Fracture toughness of a composite material is not a deterministic property. This is primarily due to the stochastic nature the crack-microstructure interactions [1-3]. Most of the existing probabilistic models for fracture toughness prediction only consider near crack-tip stress states [4-6]. Information regarding microstructure of its microstructure as well as the activation of different fracture mechanisms during characteristics and failure mechanisms associated with the crack propagation process is not explicitly included in the model formulations. He and Hutchinson [7] used the energy criterion to analyze the behavior of a semi-infinite crack perpendicular to an infinite planar interface in a symmetrically loaded, isotropic bi-material. They concluded that crack deflection occurs when

$$J_d / J_p > \Phi_{in} / \Phi_p. \tag{1}$$

Here, J_d and J_p denote the energy release rate for crack deflection and crack penetration, respectively. Similarly, Φ_{in} and Φ_p denote the surface energy of the interface and reinforcement, respectively.

Gupta et al. [8] extended He and Hutchinson's work to anisotropic materials and developed a strength criterion for crack deflection and validated their analysis using laser spallation experiments. Subsequently, Martinez and Gupta [9] improved the criterion such that it does not require any assumption concerning crack extension ratio by using a quasi-static approximation and by assuming that deflection occurs under constant loading. Although these models reveal some of the fundamental relations that govern the behavior of cracks as they approach interfaces, the analyses concern the interaction between a single crack and an infinite, flat interface. These criteria cannot be directly applied to real composite materials analysis due to the following reasons. First of all, the reinforcements in real composite materials have finite size. Therefore, the interface cannot be considered as infinite. Besides, it has been proved that the shape of reinforcements also influence the activation of different fracture mechanisms. The shape of reinforcements needs to be quantified and included in the criterion as well.

Based on the previous work, Li and Zhou [10] further extend He and Hutchinson's criterion by including the effects of finite reinforcement size, reinforcement shape and distribution in a two-phase composite material. The criterion is parameterized by

$$U = \frac{1-\beta^2}{a_0\left(1-\alpha\right)}\left[\left|c\right|^2 + \left|h\right|^2 + 2R_e\left(ch\right)\right]\overline{\rho}^{a_1}e^{\left(\frac{a_2}{s}\right)} - \frac{\Phi_{in}}{\Phi_p}, \tag{2}$$

to determine the activation of the two competing failure mechanisms. Specifically, interface debonding, which is activated by crack deflection, is predicted when $U > 0$. Otherwise, crack penetration induced reinforcement cracking will be activated instead. In the above relation, $\overline{\rho}$ is the roundness of the reinforcement. s represents the characteristic reinforcement size. Φ_{in} and Φ_p are the surface energies of the interface and reinforcement, respectively.

Based on the previous work, an integrated computational and analytical model is introduced which allows the possible range of fracture toughness values to be predicted as function of microstructure. The Weibull distribution parameters are directly correlated to the two-point correlation functions as well as the quantification of fracture mechanisms. These relations can be used for material reliability design by controlling the fracture toughness scatter through microstructure tailoring.

Cohesive Finite Element Method (CFEM) based fracture toughness prediction

The edge-cracked square specimen under Mode I tensile loading in Fig. 1 has a size length of 3.65 mm. The microstructure region has a length of 2 mm, width of 1 mm and a pre-crack length of 0.73 mm. A boundary velocity of $\mathrm{v} = 5$ mm/s is imposed at the top and bottom edges. The remaining edges of the specimen are traction-free. Conditions of plane strain are assumed to prevail. In the following analysis, we consider TiB_2 reinforced Al_2O_3 composites.

Cohesive elements are embedded along the edge of each bulk element in the microstructure region. Bilinear traction separation law is employed. This law is derived from a potential Φ which is a function of separation vector Δ through a state variable defined as

$\lambda = \sqrt{\left(\Delta_n/\Delta_{nc}\right)^2 + \left(\Delta_t/\Delta_{tc}\right)^2}$. This variable describes the effective instantaneous state of mixed-mode separations. Here, $\Delta_n = \mathbf{n}\cdot\boldsymbol{\Delta}$ and $\Delta_t = \mathbf{t}\cdot\boldsymbol{\Delta}$ denote, respectively, the normal and tangential components of $\boldsymbol{\Delta}$, with $\mathbf{n}$ and $\mathbf{t}$ being unit normal and tangent vectors. Δ_{nc} is the critical normal separation at which the cohesive strength of an interface vanishes under conditions of pure normal deformation ($\Delta_t = 0$). Similarly, Δ_{tc} is the critical tangential separation at which the cohesive strength of an interface vanishes under conditions of pure shear deformation ($\Delta_n = 0$). λ tracks instantaneous mixed-mode separations during both loading and unloading. Apparently, $\lambda = 0$ corresponds to $\Delta = 0$ (undeformed state or fully unloaded state) and $\lambda \geq 1$ implies complete separation, i.e. total debonding of the cohesive surface pair.

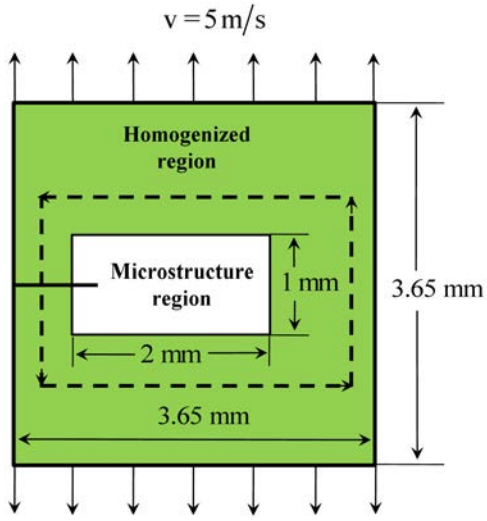


Fig. 1 Specimen configuration used in the analysis.

CFEM models with traction-separation laws with finite initial stiffness have two competing requirements on element size. The upper bound requires that the element size must be small enough to accurately resolve the stress distribution inside the cohesive zones at crack tips. The lower bound, on the other side, requires the cohesive surface induced stiffness reduction be small, such that the wave speed in the solid is not significantly affected due to the presence of the cohesive surfaces. For the conditions of this paper, the preferred range of the element size is $7\,\mu m \ll h \ll 14\,\mu m$, allowing the convergence criterion in Tomar et al. [11] to be satisfied.

For brittle materials, the fracture toughness K_{IC} is related to the energy release rate J_{IC} as

$$K_{IC} = \sqrt{J_{IC}\frac{\bar{E}}{1-\bar{\nu}^2}}, \tag{3}$$

where $\bar{E}$ and $\bar{\nu}$ are the effective Young's modulus and effective Poisson's ratio of the heterogeneous material, respectively. $\bar{E}$ and $\bar{\nu}$ are estimated by using the Mori-Tanaka method as

$$\begin{cases} \bar{E} = \dfrac{9\bar{K}\bar{\mu}}{3\bar{K}+\bar{\mu}} \text{ and} \\ \bar{\nu} = \dfrac{3\bar{K}-2\bar{\mu}}{6\bar{K}+2\bar{\mu}}, \end{cases} \tag{4}$$

where $\bar{K}$ and $\bar{\mu}$ are effective bulk and shear moduli. $\bar{K}$ and $\bar{\mu}$ are calculated according to

$$\begin{cases} \bar{K} = K_0 + \dfrac{f(K_1-K_0)(3K_0+4\mu_0)}{3K_1+4\mu_0} \text{ and} \\ \bar{\mu} = \mu_0 + \dfrac{5f\mu_0(\mu_1-\mu_0)(3K_0+4\mu_0)}{3K_0(3\mu_0+2\mu_1)+4\mu_0(2\mu_0+3\mu_1)}. \end{cases} \tag{5}$$

Here, K_r and μ_r represent the bulk and shear modulus, respectively for Al_2O_3 ($r=0$) and TiB_2 ($r=1$).

To account for inertia effects, a fully dynamic deformation formulation is used. Within this framework, the path-independent *J*-integral is (Moran & Shih [12])

$$J = \int_\Gamma \left[\left(\int_0^t \boldsymbol{\sigma} : \mathrm{d}\boldsymbol{\varepsilon} + \frac{1}{2}\rho\dot{\mathbf{u}}\cdot\dot{\mathbf{u}} \right) \mathrm{d}x_2 - \boldsymbol{t}\cdot\frac{\partial \mathbf{u}}{\partial x_1}\mathrm{d}s \right] + \int_A \left(\rho\ddot{\mathbf{u}}\cdot\frac{\partial \mathbf{u}}{\partial x_1} - \rho\dot{\mathbf{u}}\cdot\frac{\partial \dot{\mathbf{u}}}{\partial x_1} \right) \mathrm{d}A, \tag{6}$$

where $\boldsymbol{t}$ is the traction on a surface with normal **N**, **u** is the displacement, $\boldsymbol{\varepsilon}$ denotes the strain and ρ is the mass density.

The *J* value in eqn. (6) is integrated along an arbitrary closed contour as shown in the dashed line in Fig. 1. Six snap shots of the crack propagation process in a microstructure with circular TiB_2 reinforcement at a loading velocity of $v = 5$ mm/s is illustrated in Fig. 2(a). The corresponding histories of *J* and *K* are shown in Fig. 2(b). Fracture initiates in the Al_2O_3 matrix at 105.0 μs, this event defines the initiation toughness K_{IC}^{i}. The crack is arrested by a TiB_2 particle and pauses at the Al_2O_3/TiB_2 interface for approximately 42.5 μs. During the pause, *J* increases rapidly. At approximately 149.2 μs, as a result of the higher level of driving force *J*, the crack penetrates the TiB_2 particle. Subsequently, the crack propagates rapidly, causing *J* (and therefore *K*) in Fig. 2 (b) to plateau for the remainder of the analysis. The average value of *K* during this period is taken as the propagation toughness K_{IC}.

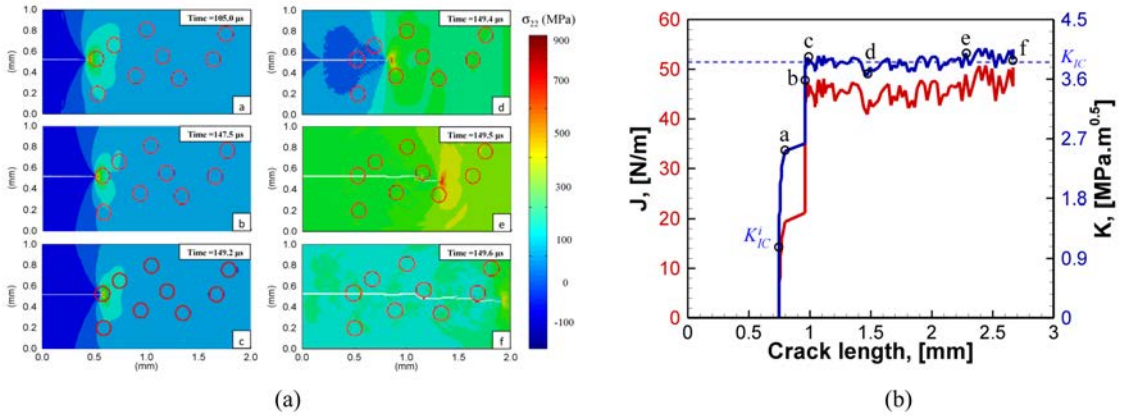


Fig. 2 History of (a) crack propagation in Al_2O_3/TiB_2 and the evolution of corresponding J and K.

From the energy point of view, a crack would grow when the energy available in the elastic stress field reaches the energy required to form new fracture surfaces. For crack propagation in a composite material as shown in Fig. 3, new crack surfaces can be created in the matrix, along the interface and in the enforcement. Therefore, J_{IC} can be stated as

$$\begin{aligned} J_{IC} &= \frac{\partial U_f}{\partial A} \approx \frac{\left(\Phi_{in}L_{in} + \Phi_m L_m + \Phi_p L_p\right)t}{Wt} \\ &= \frac{L}{W}\left(\Phi_{in}\frac{L_{in}}{L} + \Phi_m \frac{L_m}{L} + \Phi_p \frac{L_p}{L}\right) \\ &= \xi\left(R, f\right)\left(\Phi_{in}H_{in} + \Phi_m H_m + \Phi_p H_p\right), \end{aligned} \tag{7}$$

where U_f is the total energy released. $A = Wt$ is the total projected crack surface area with W and t being the crack projection length and specimen thickness, respectively. $\xi = L/W$ is a function which captures the tortuosity of the entire crack path. Based on the microstructure configuration discussed before, ξ depends on the R and f. The detailed calculation of H_{in}, H_m and H_p are discussed in detail in Li and Zhou [13].

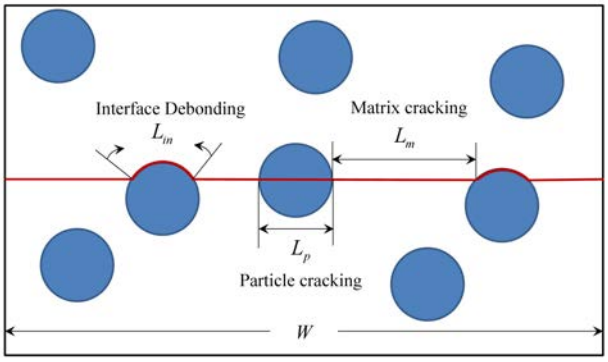


Fig. 3 Schematic illustration of crack lengths associated with different mechanisms in two-phase composite materials.

Probabilistic fracture toughness analysis

A typical two-parameter Weibull distribution function is in the form of

$$P_f = 1 - \exp\left[-\left(\frac{K}{K_0} \right)^m \right]. \tag{8}$$

Here P_f is the probability of fracture. K and K_0 are the fracture toughness K_{IC} measured from experiments and the normalization factor, respectively. m is defined as the shape parameter. The parameters m and K_0 are obtained through a linear regression fit to N data points of K. In order to have a good statistical representation of the stochastic fracture process, $N \geq 20$ is preferred.

In most of the existing probabilistic models, the fracture toughness data is obtained first and then then fitted by Weibull distribution function [14-16]. The problem of these probabilistic models is that they do not allow the scatter of fracture toughness data to be predicted prior to the experimental testing. Without the material sensitivity information, it is hard to determine the number of tests required to obtain a good estimate of probability of material fracture.

Although Weibull parameters in Eqn. (8) are fitting parameters which carry little physical significance, their correlations with microstructure characteristics and fracture mechanisms can provide valuable insight to material sensitive design without doing repeated experimental testing.

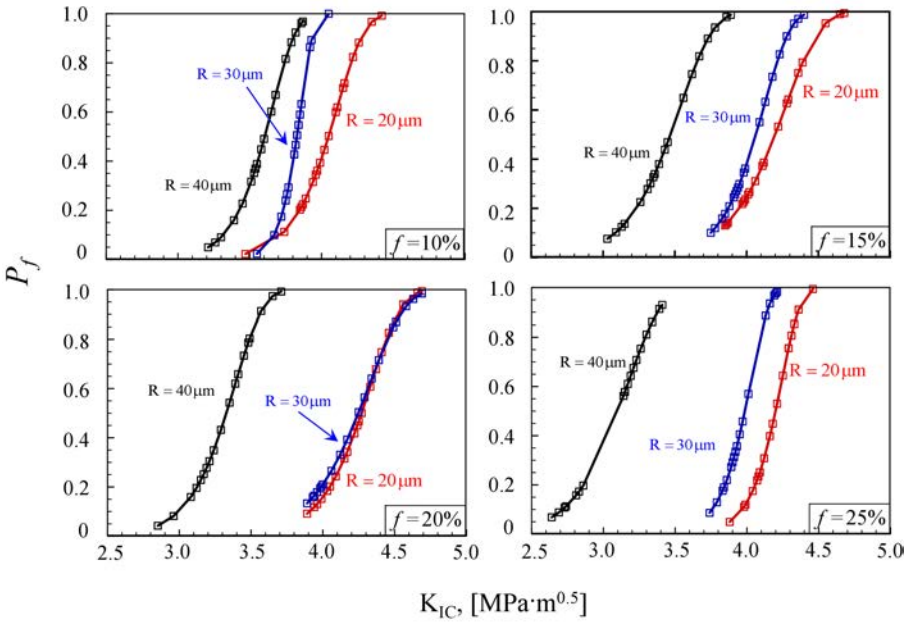


Fig. 4 Fracture probability distribution predicted from CFEM simulations [2] for microstructures with randomly distributed non-overlapping circular particles.

In the following discussions, the K values in Eqn. (8) are calculated from Eqn.(3), Eqn. (6) and Eqn. (7). Microstructures with non-overlapping circular reinforcements are considered. Fig. 4 compares the probability of fracture P_f for microstructures with systematically varying particle radius ($R = 20\,\mu\text{m}$, $30\,\mu\text{m}$ and $40\,\mu\text{m}$) and volume fraction (f = 10%, 15%, 20% and 25%). It is observed that microstructures with smaller radius tend to have higher fracture toughness and lower probability of fracture for all the volume fractions considered. The same trend is observed from the CFEM (Cohesive Finite Element Method) calculations [10] in Fig. 4, where 20 microstructures with same combination of R and f are considered in the analysis. As show in Fig. 5, when f is kept as a constant, the slope of fracture probability curve becomes steeper as R increases, leading to less fracture toughness scatter. The opposite trend is observed when R is fixed while f is increased from 10% to 25%. This indicates that microstructures with fine particles and high volume fractions will have higher-order uncertainties due to the large fracture toughness variation. However, it should be noted that this type of combination also yield higher level fracture toughness values at the same time. This trend has been reported in a few research studies [17-19]. From the microstructure design perspective, it is not surprising that shifting up the fracture toughness values will lead to larger scatter band. First of all, composites materials are toughened through crack-particle interactions. Generally speaking, more interactions during the crack propagation process will lead to higher fracture resistance. If a crack does not encounter any reinforcement, the choice of crack path is very limited. The fracture toughness of the composite material is very close to the fracture toughness of matrix material which is considered as the lower bound fracture toughness. Besides, more crack-particle interactions can be created by increasing the volume fraction of particles. As discussed previously, the effective toughening mechanism during crack-particle interaction is crack deflection induced interface debonding. This requires fine particles in addition to high volume fraction. Crack deflection, which contributes to enhancing the level of fracture toughness, also provides the crack with more opportunities in choosing the path. The scatter is intensified when more crack-particle interactions are included. This explains why microstructure configurations which lead to higher level of fracture toughness also have larger fracture toughness scatter.

It is also noted that the fracture toughness values predicted from Eqn. (3) and (7) is larger than the values predicted from CFEM results. In Fig. 5, the range of K_{IC} values predicted from the analytical model is approximately from $2.7\ \text{MPa}\sqrt{\text{m}}$ to $8.4\ \text{MPa}\sqrt{\text{m}}$. In contrast, the range of K_{IC} values predicted from CFEM calculations only spans from $2.7\ \text{MPa}\sqrt{\text{m}}$ to $4.7\ \text{MPa}\sqrt{\text{m}}$ as shown in Fig. 4. With lower bound K_{IC} predictions being consistent, the analytical model predicts a much higher upper bound value. This is because the two-point correlation functions employed in the analytical model quantify the possibility of crack-particle interactions in the entire microstructure region. However, the crack propagation in CFEM simulations is primarily localized in a small region near the pre-crack plane. The crack does not have the opportunity to interact with particles which are far away from the pre-crack plane. To quantitatively understand how localization of crack propagation influences the level

of fracture toughness and the magnitude of scatter, another set of calculations are carried out by considering the interactions of particles within the local region of each microstructure instantiation employed in CFEM calculations .

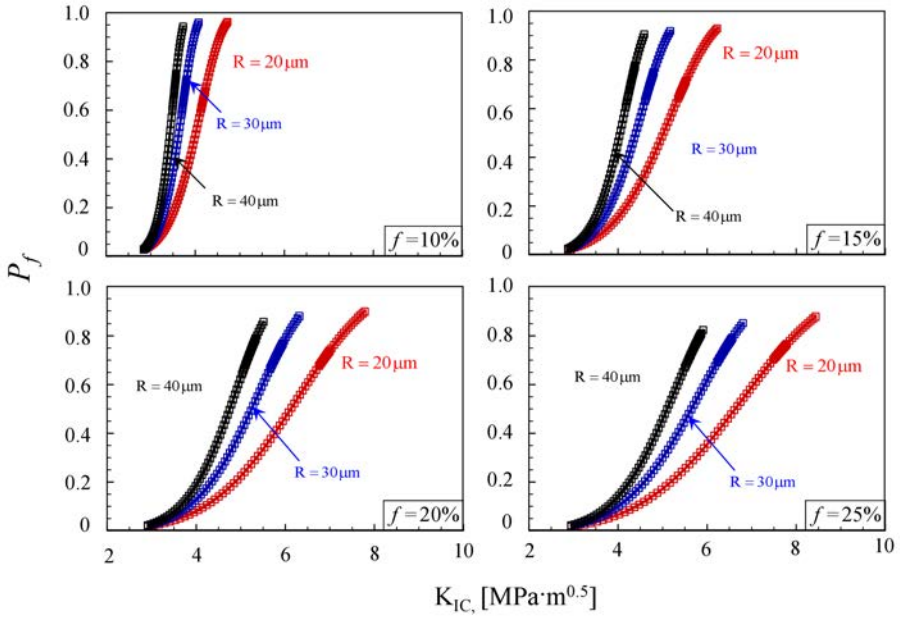


Fig. 5 Fracture probability distribution predicted from analytical model for microstructures with randomly distributed non-overlapping circular particles.

Fracture toughness scatter is quantified by the shape parameter m. It can be inferred from Eqn. (8) that if the magnitude of the scatter is large, then m is small and vice versa. Specially, $m \to \infty$ is expected if there is no scatter. Theoretically, it can be achieved only when $f = 0$ or $f = 1$ as the microstructure is purely matrix phase or reinforcement phase.

Fig. 6 compares m values predicted from the analytical model and CFEM framework. The solid lines and dashed lines represent analytical solutions with the entire microstructure and local microstructure region, respectively. CFEM results are illustrated by dots. Microstructures considered here have volume fraction f ranging from 0 to 30%. Three particle sizes with $R = 20\,\mu\text{m}$, $R = 30\,\mu\text{m}$ and $R = 40\,\mu\text{m}$ are employed and represented by blue, red and black color, respectively. As demonstrated in Fig. 6, a much higher level of m values are predicted for analytical solutions considering the local microstructure region. This means there is smaller fracture toughness scatter when the crack only propagates in the localized microstructure region. It makes sense that a lower level of m values are predicted when the entire microstructure region is considered since the interactions of crack with all the particles in the microstructure region are included. From microstructure design prospective, m values predicted by considering the entire microstructure region are very conservative since the crack-particle interaction is usually localized when the reinforcements are well bonded with the matrix. Therefore, analysis with local and entire microstructure region can serve as the upper limit and lower limit of m as represented by dashed and solid lines, respectively. Despite discrepancies in m, both predictions share the same trends. First of all, the increase

of volume fraction f leads to decreased m. The larger scatter of fracture toughness is observed due to more intensified crack-particle interactions during crack propagation. The decrease of R can have the same effect as small particles promote interface debonding and create more uncertainties in choosing the crack path. It is noted that the CFEM predictions of m all fall between the upper and lower bound with the similar trends as observed from the analytical predictions. It is also observed that the CFEM predictions are closer to the upper bound m as f increases. When $f = 25\%$, m values predicted from CFEM calculations are very close to the upper bound prediction especially when particle size is small. It can be inferred that the crack-particle interactions in local microstructure region is representative of the entire microstructure region when reinforcements with small size and large volume fractions are considered. As shown in Fig. 6, the discrepancy between the lower bound curves and upper bound curves becomes smaller with increasing f and decreasing R. Once f increases to 100% and R decreases to 0, both the upper and lower bound curves will saturate to $m \to \infty$ as the microstructure becomes pure reinforcement phase. This means m will not continue to decrease as f increases. After f reaches a critical value, the trend reverses. Although the critical f cannot be predicted because the analytical model developed here only considers non-overlapping circular reinforcements, the trends observed from the above analysis are still valid for most engineering cases.

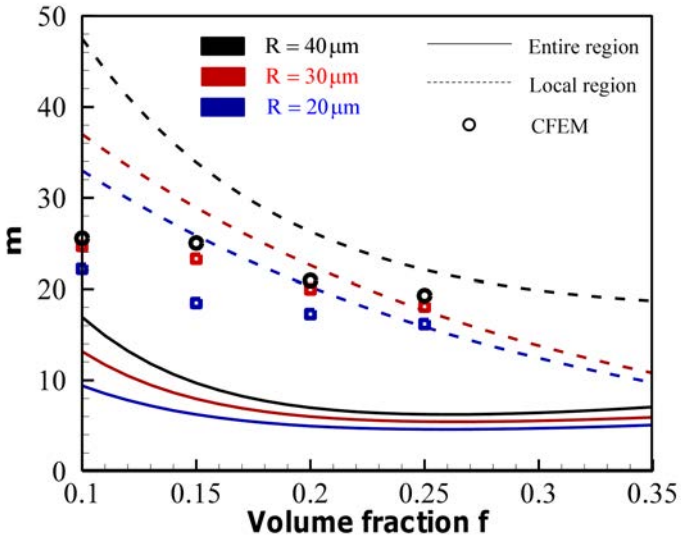


Fig. 6 Effect of microstructure attributes on m predicted from CFEM model and analytical model considering the entire and local microstructure region, respectively.

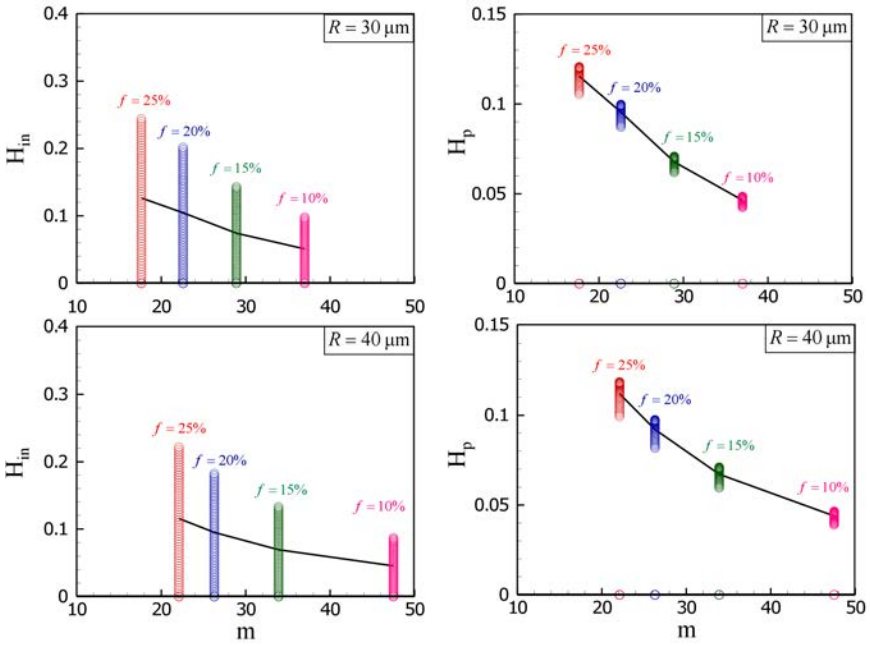


Fig. 7 Effect of proportions of interface debonding H_{in} and particle cracking H_p on m under different values of particle size and volume fraction.

Fig. 7 summarizes the scatter of H_{in} and H_p under different particle size and volume fraction, and compares them with the scatter of fracture toughness m. H_{in}, H_p and m are all predicted from the analytical model considering the local microstructure region. The solid black line in each sub-figure connects the average value of H_{in} or H_p under each volume fraction. It is noted that the scatter of fracture toughness primarily comes of the scatter of H_{in}. The increase in volume fraction f and decrease in particle size R can lead to higher average H_{in} and higher H_{in} as well. Compared with H_{in}, H_p is less sensitive to R and f. The increase in volume fraction f and particle size R has limited effect on the average value of H_p and its scattering.

Conclusions

In conclusion, the most effective way to improve the fracture toughness of two-phase composite material is to increase crack tortuosity by promoting interface debonding. This can be achieved by introducing refined second-phase reinforcements with adequate volume fraction. It should be noted that the decrease in reinforcement size and increase in volume fraction also enhance the sensitivity of the material system as larger fracture toughness scatter is observed at the same time. The analytical model developed here provide a way to estimate the upper and lower limit of fracture toughness by considering microstructure attributes and fracture mechanisms involved in the failure process. The prediction of Weibull parameter m

as shown in Fig. 6 can be employed as a reference of fracture toughness scatter for material sensitive design of two-phase composite materials.

References

[1] K. Strecker, S. Ribeiro, R. Oberacker, and M. J. Hoffmann, "Influence of microstructural variation on fracture toughness of LPS-SiC ceramics," *International Journal of Refractory Metals & Hard Materials,* vol. 22, pp. 169-175, 2004.

[2] Y. Li and M. Zhou, "Prediction of fracture toughness of ceramic composites as function of microstructure: I. Numerical simulations," *Journal of the Mechanics and Physics of Solids,* vol. 61, pp. 472-488, Feb 2013.

[3] S. Lee, T. H. Kim, and D. Kwon, "Microstructural Analysis of Fracture-Toughness Variation in 2xxx-Series Aluminum-Alloy Composites Reinforced with Sic Whiskers," *Metallurgical and Materials Transactions a-Physical Metallurgy and Materials Science,* vol. 25, pp. 2213-2223, Oct 1994.

[4] T. L. Becker, R. M. Cannon, and R. O. Ritchie, "Statistical fracture modeling: crack path and fracture criteria with application to homogeneous and functionally graded materials," *Engineering Fracture Mechanics,* vol. 69, pp. 1521-1555, Sep 2002.

[5] B. Z. Margolin, A. G. Gulenko, and V. A. Shvetsova, "Probabilistic model for fracture toughness prediction based on the new local fracture criteria," *International Journal of Pressure Vessels and Piping,* vol. 75, pp. 307-320, Apr 1998.

[6] A. Valiente, J. Ruiz, and M. Elices, "A probabilistic model for the pearlite-induced cleavage of a plain carbon structural steel," *Engineering Fracture Mechanics,* vol. 72, pp. 709-728, Mar 2005.

[7] H. Ming-Yuan and J. W. Hutchinson, "Crack deflection at an interface between dissimilar elastic materials," *International Journal of Solids and Structures,* vol. 25, pp. 1053-1067, // 1989.

[8] V. Gupta, J. Yuan, and D. Martinez, "Calculation, Measurement, and Control of Interface Strength in Composites," *Journal of the American Ceramic Society,* vol. 76, pp. 305-315, 1993.

[9] D. Martínez and V. Gupta, "Energy criterion for crack deflection at an interface between two orthotropic media," *Journal of the Mechanics and Physics of Solids,* vol. 42, pp. 1247-1271, 1994/08/01 1994.

[10] Y. Li and M. Zhou, "Prediction of fracturess toughness of ceramic composites as function of microstructure: II. analytical model," *Journal of the Mechanics and Physics of Solids,* vol. 61, pp. 489-503, 2// 2013.

[11] V. Tomar, J. Zhai, and M. Zhou, "Bounds for element size in a variable stiffness cohesive finite element model," *International Journal for Numerical Methods in Engineering,* vol. 61, pp. 1894-1920, Nov 21 2004.

[12] B. Moran and C. F. Shih, "A general treatment of crack tip contour integrals," *International Journal of Fracture,* vol. 35, pp. 295-310, 1987.

[13] Y. Li and M. Zhou, "Prediction of fracture toughness scatter of composite materials," *Computational Materials Science,* vol. 116, pp. 44-51, Apr 15 2016.

[14] W. Weibull, "A statistical theory of the strength of materials," *Ing. Vetenskaps Akad. Handl.151.,* pp. 1-45, 1939.

[15] K. Wallin, "Macroscopic Nature of Brittle-Fracture," *Journal De Physique Iv,* vol. 3, pp. 575-584, Nov 1993.

[16] K. Wallin, K. Torronen, R. Ahlstrand, B. Timofeev, V. Rybin, V. Nikolaev, *et al.*, "Theory Based Statistical Interpretation of Brittle-Fracture Toughness of Reactor

Pressure-Vessel Steel 15x2m-Phi-a and Its Welds," *Nuclear Engineering and Design*, vol. 135, pp. 239-246, Jun 1992.

[17] K. V. Logan, "Composite Ceramics, Final Tehnical Report, final technical report," ed: USSTACOM DAAEO7-95-C-R040, 1996.

[18] A. G. Evans, "Perspective on the Development of High-Toughness Ceramics," *Journal of the American Ceramic Society*, vol. 73, pp. 187-206, Feb 1990.

[19] T. Lin, A. G. Evans, and R. O. Ritchie, "Statistical-Analysis of Cleavage Fracture Ahead of Sharp Cracks and Rounded Notches," *Acta Metallurgica*, vol. 34, pp. 2205-2216, Nov 1986.

Numerical simulation of galvanic corrosion by boundary node method

†*Sanshan Tu, Hongqi Yang, Liang Zhou, and Yi Huang

School of Naval Architecture, Dalian University of Technology, Dalian, China.

*Presenting author: tss71618@163.com
†Corresponding author: tss71618@163.com

Abstract

The electric potential in the electrolyte obeys Laplace equation in galvanic corrosion problems. In this research, we developed the moving Kriging interpolation based boundary node method (MKIBNM) for galvanic corrosion problems to predict the corrosion rate and track the moving boundary of the corroding constituent. The numerical results obtained from MKIBNM are compared with the experimental results.

Keywords: Laplace equation; Galvanic corrosion; Moving Kriging interpolation; Boundary node method;

1 Introduction

The conventional methods such as finite element method (FEM) and boundary element method (BEM) require elements to construct the shape function, while a meshless method constructs the shape function entirely based on scattered nodes with greater flexibility and higher precision.

The meshless methods have been used in many engineering problems such as fluid mechanics problems [1], solid mechanics problems [2] [3], Elasticity Problems [4] [15], viscoelasticity problems [5], heat conduction problems [6] [20] and elastodynamic problems [14] [16].

Boundary integral equation (BIE) based meshless methods are an important part of meshless methods. The boundary node method (BNM) is firstly proposed by Mukherjee et al. [7]-[10] based on the moving least square (MLS) method [11] and BIE. Unlike many other 'domain' type meshless methods, the BNM only requires scattered nodes on the 1-D bounding curve of a 2-D area or the 2-D bounding surface of a 3D body and a simple boundary cell structure for numerical integration. The idea of BNM has been developed by many other researchers such as Zhu et al. [12] [13], Liew et al.[14]-[16], Ren et al. [17] [18] and Li et al. [19] with different approximating methods.

Approximating methods play an important role in meshless methods. The moving Kriging interpolation (MKI) method [20], also known as the radial point interpolation method (RPIM) [2], is an important approximating method to construct shape functions in meshless methods. MKI has partition of unity property, consistency property and high approximation precision. Besides, the shape functions constructed by MKI possess Kronecker delta property. Then, the essential boundary condition can be imposed directly and easily. Li et al. [19] have proposed the moving Kriging interpolation-based boundary node method (MKIBNM) by combing moving Kriging interpolation (MKI) with BIE for potential problems. MKIBNM has a high precision and can directly impose boundary conditions. Therefore, we chose MKIBNM to simulate galvanic corrosion.

Galvanic corrosion is an example of a process undergoing electro dissolution that aggravates metal corrosion and limits the widespread use of alloys such as the use of magnesium alloys in the automotive industry. Researchers have investigated galvanic corrosions by both experiments and numerical simulations [21]-[25]. We, for the first time, developed the MKIBNM method to investigate the corrosion behavior of galvanic couples. MKIBNM can easily track the moving boundary of the moving boundary of the corroding constituent

without re-meshing. We use the case and experimental results from reference [21] [22] to demonstrate the accuracy and flexibility of BNM.

2 Moving kriging interpolation (MKI) on 2D boundary

In 2D boundary meshless methods, we use one parametric coordinate s to approximate boundary function values. The shape function $\boldsymbol{N}(s)$ in MKI is defined by

$$\boldsymbol{N}(s) = \boldsymbol{p}(s)\boldsymbol{A} + \boldsymbol{r}^T(s)\boldsymbol{B} \tag{1}$$

where $\boldsymbol{p}(s) = [p_1(s), p_2(s), \cdots, p_m(s)]$, $p_j(s) = s^{j-1}$ $(j = 1, 2, \cdots, m)$ are monomial basis functions， m is the number of terms of basis. $\boldsymbol{r}(s)$ is the $n \times 1$ vector of correlation between point at s and the given nodes

$$\boldsymbol{r}(s) = [R(s, s_1), R(s, s_2), \cdots, R(s, s_n)]^{\mathrm{T}} \tag{2}$$

$\boldsymbol{A}$ and $\boldsymbol{B}$ are the temporary matrices

$$\begin{cases} \boldsymbol{M} = \boldsymbol{P}^T \boldsymbol{R}^{-1} \boldsymbol{P} \\ \boldsymbol{A} = \boldsymbol{M}^{-1} \boldsymbol{P}^T \boldsymbol{R}^{-1} \\ \boldsymbol{B} = \boldsymbol{R}^{-1}(\boldsymbol{I} - \boldsymbol{P}\boldsymbol{A}) \end{cases} \tag{3}$$

$\boldsymbol{P}$ is the $n \times m$ matrix that has basis function values at the given nodes.

$$\boldsymbol{P} = \begin{bmatrix} p_1(s_1) & p_2(s_1) & \cdots & p_m(s_1) \\ p_1(s_2) & p_2(s_2) & \cdots & p_m(s_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(s_n) & p_2(s_n) & \cdots & p_m(s_n) \end{bmatrix} \tag{4}$$

$\boldsymbol{R}$ is the $n \times n$ matrix of correlation between the given nodes

$$\boldsymbol{R} = \begin{bmatrix} 1 & R(s_1, s_2) & \cdots & R(s_1, s_n) \\ R(s_2, s_1) & 1 & \cdots & R(s_2, s_n) \\ \vdots & \vdots & \ddots & \vdots \\ R(s_n, z_1) & R(s_n, s_2) & \cdots & 1 \end{bmatrix} \tag{5}$$

The correlation function $R(s_i, s_j)$ takes the form of Gaussian function in this work.

$$R(s_i, s_j) = \exp(-\frac{\omega r_{ij}^2}{d_m^2}) \tag{6}$$

where $r_{ij} = \left| s_i - s_j \right|$, d_m is the minimum distance between any two nodes on the local boundary, $\omega > 0$ is a correlation parameter and $\omega = 0.03 \sim 0.2$ is recommended [30].
$\boldsymbol{N}(s)$ has the following properties:
Kronecker delta property

$$N_I(s_J) = \delta_{IJ} \tag{7}$$

, consistency property

$$\sum_{i=1}^{n} N_i(s) s_i^j = s^j \quad (j = 0, 1, \cdots, m-1) \tag{8}$$

and partition of unity property

$$Sum(\boldsymbol{N}(\mathrm{s})) = \sum_{i=1}^{n} N_i(s) = 1 \tag{9}$$

The Kronecker delta property enables the boundary condition to be imposed directly and the results to be obtained directly, and the consistency property ensures high accuracy of the approximation.

3 Boundary node method for galvanic corrosion

3.1 Governing equations

The equation governing the potential distribution and the current flow in the electrolyte can be derived from charge conservation. The continuity equation requires that the current per unit volume, $\boldsymbol{J}$, relates to the charge, q, by

$$\nabla \boldsymbol{J} = \frac{\partial q}{\partial t} \tag{10}$$

Taking into account the relationship of electric field intensity, $\boldsymbol{E}$,

$$\boldsymbol{E} = -\nabla \phi \tag{11}$$

and Ohms law,

$$\boldsymbol{I} = \sigma \boldsymbol{E} \tag{12}$$

where σ is the conductivity of the electrolyte, the continuity equation transforms to

$$\nabla(\sigma \nabla \phi) = -\nabla(\frac{\partial q}{\partial t}) \tag{13}$$

Galvanic corrosion is a very slow process, thus we can make the following assumptions:
1. The electrolyte solution is well mixed that the conductivity is isotropic, σ is a constant.
2. The solution is electro-neutral, $\frac{\partial q}{\partial t} = 0$.

With the above assumptions, Eq. (13) can be simplified as,

$$\nabla^2 \phi = 0 \tag{14}$$

Therefore, for a uniform, isotropic electrolyte, the potential obeys the Laplace equation.

3.2 Boundary condition

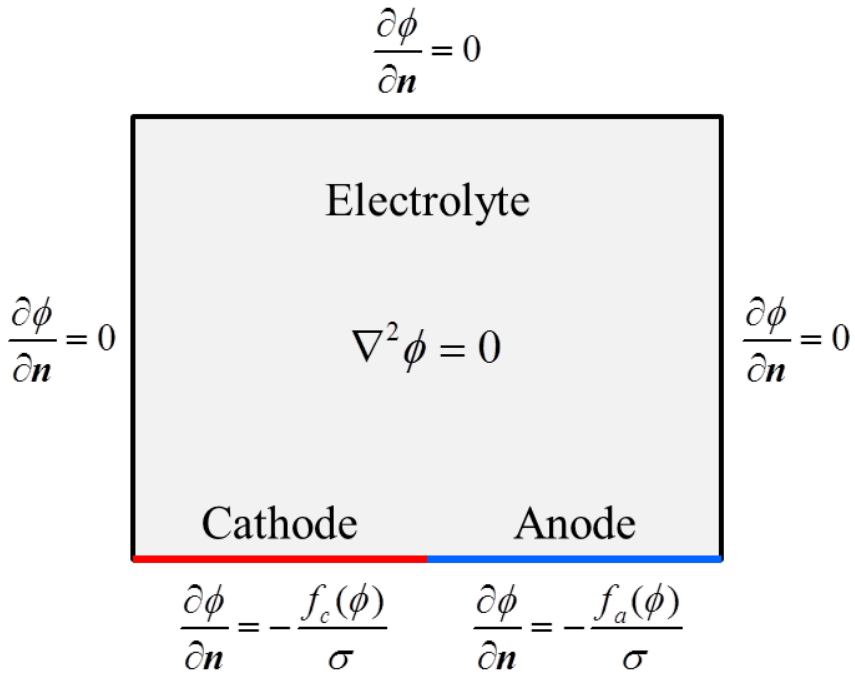


Figure 1. Schematic of the mathematical model of galvanic corrosion

Eq. (14) can be solved with the boundary conditions shown schematically in Figure 1. The boundary conditions at the anode and the cathode surfaces are vital to predict the corrosion rates. The boundary condition applied at the anode surface Γ_a is:

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = -\frac{j}{\sigma} = -\frac{f_a(\phi)}{\sigma} \tag{15}$$

where σ is the electrical conductivity of the electrolyte solution and $f_a(\phi)$ is the current density determined by anodic species. $f_a(\phi)$ is a piecewise linear interpolation function which is obtained from the polarization curve of the anodic species. Thus, we use a piecewise linear interpolation approach to handle non-linear boundary conditions.

Similarly, the boundary condition applied at the cathode surface Γ_c is

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = -\frac{f_c(\phi)}{\sigma} \tag{16}$$

The boundary condition applied at the insulation surface and the electrolyte-air interface Γ_{ins} is

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = -\frac{j}{\sigma} = 0 \tag{17}$$

3.3 BNM formulation

In summary, the governing equation and the boundary conditions are

$$\begin{cases} \Delta\phi = \dfrac{\partial^2 \phi}{\partial x^2} + \dfrac{\partial^2 \phi}{\partial y^2} = 0 & \text{in } \Omega \\ \dfrac{\partial \phi}{\partial \boldsymbol{n}} = -\dfrac{f_a(\phi)}{\sigma} & \text{on } \Gamma_a \\ \dfrac{\partial \phi}{\partial \boldsymbol{n}} = -\dfrac{f_c(\phi)}{\sigma} & \text{on } \Gamma_c \\ \dfrac{\partial \phi}{\partial \boldsymbol{n}} = 0 & \text{on } \Gamma_{ins} \end{cases} \tag{18}$$

where ϕ is the potential, $j = -\sigma \dfrac{\partial u}{\partial \boldsymbol{n}}$ is the current density along the normal of the boundary, the whole boundary is $\Gamma = \partial\Omega = \Gamma_a + \Gamma_c + \Gamma_{ins}$. The integral representation of the solution for Eq. (18) is

$$c(\boldsymbol{\xi})\phi(\boldsymbol{\xi}) + \int_\Gamma \frac{\partial \phi^*(\boldsymbol{x}-\boldsymbol{\xi})}{\partial \boldsymbol{n}} \phi(\boldsymbol{x}) \mathrm{d}\Gamma = \frac{-1}{\sigma} \int_\Gamma \phi^*(\boldsymbol{x}-\boldsymbol{\xi}) j(\boldsymbol{x}) \mathrm{d}\Gamma \tag{19}$$

where $\boldsymbol{\xi}$ denotes the given source point on the boundary,

$\boldsymbol{x}$ denotes a filed point on the boundary,

$\phi^*(\boldsymbol{x}-\boldsymbol{\xi}) = -\dfrac{1}{2\pi} \ln|\boldsymbol{x}-\boldsymbol{\xi}|$ is the fundamental solution of Laplace's equation,

$\dfrac{\partial \phi^*}{\partial \boldsymbol{n}}$ is the normal derivative of ϕ^* on the boundary,

$c(\boldsymbol{\xi})$ is a coefficient related to the boundary smoothness.

The boundary Γ is divided into background cells Γ_i $(i = 1, 2, 3, \mathrm{L}, m)$ for numerical integration, and the discrete form of Eq. (19) is

$$c(\boldsymbol{\xi})u(\boldsymbol{\xi})+\sum_{i=1}^{m}\int_{\Gamma_i}\frac{\partial\phi^*(\boldsymbol{x}-\boldsymbol{\xi})}{\partial\boldsymbol{n}}\phi(\boldsymbol{x})\mathrm{d}\Gamma=\frac{-1}{\sigma}\sum_{i=1}^{m}\int_{\Gamma_i}\phi^*(\boldsymbol{x}-\boldsymbol{\xi})\,j(\boldsymbol{x})\mathrm{d}\Gamma \tag{20}$$

$u(\boldsymbol{x})$ and $q(\boldsymbol{x})$ are approximated by MKI.

$$\begin{cases} u=\boldsymbol{N}(s)\boldsymbol{\Phi}=\sum_{k=1}^{n}\varphi_k(s)\phi_k \\ j=\boldsymbol{N}(s)\boldsymbol{Q}=\sum_{k=1}^{n}\varphi_k(s)\,j_k \end{cases} \tag{21}$$

To track the moving boundary, $\boldsymbol{x}$ is also approximated by MKI.

$$\boldsymbol{x}=\boldsymbol{N}(s)\boldsymbol{X}=\sum_{k=1}^{n}\varphi_k(s)\boldsymbol{x}_k \tag{22}$$

Substituting Eq. (21) into Eq. (20) yields

$$c(\boldsymbol{\xi})\phi(\boldsymbol{\xi})+\sum_{i=1}^{m}\int_{\Gamma_i}\frac{\partial\phi^*(\boldsymbol{x}-\boldsymbol{\xi})}{\partial\boldsymbol{n}}\boldsymbol{N}(s)\boldsymbol{\Phi}\mathrm{d}\Gamma=\frac{-1}{\sigma}\sum_{i=1}^{m}\int_{\Gamma_i}\phi^*(\boldsymbol{x}-\boldsymbol{\xi})\boldsymbol{N}(s)\boldsymbol{J}\mathrm{d}\Gamma \tag{23}$$

Employing numerical methods for the integrals in Eq. (23), at every nodes, we can obtain the linear algebraic equations.

$$\boldsymbol{C\Phi}+\hat{\boldsymbol{H}}\boldsymbol{\Phi}=\boldsymbol{GJ} \tag{24}$$

where

$$\boldsymbol{C}=diag(c(\boldsymbol{\xi}_1),c(\boldsymbol{\xi}_2),\mathrm{L}\ ,c(\boldsymbol{\xi}_n)) \tag{25}$$

$$\hat{\boldsymbol{H}}_i=\sum_{k=1}^{m}\int_{\Gamma_k}\frac{\partial\phi^*(\boldsymbol{x}-\boldsymbol{\xi}_i)}{\partial\boldsymbol{n}}\boldsymbol{N}(s)\mathrm{d}\Gamma \tag{26}$$

$$\boldsymbol{G}_i=\frac{-1}{\sigma}\sum_{k=1}^{m}\int_{\Gamma_k}\phi^*(\boldsymbol{x}-\boldsymbol{\xi}_i)\boldsymbol{N}(s)\mathrm{d}\Gamma \tag{27}$$

Let

$$\boldsymbol{H}=\boldsymbol{C}+\hat{\boldsymbol{H}} \tag{28}$$

Eq. (24) can then be rewritten as

$$\boldsymbol{H\Phi}=\boldsymbol{GQ} \tag{29}$$

Normally, we do not directly compute $\boldsymbol{C}$, because the diagonal elements of $\boldsymbol{H}$ can be computed by constant potential method.

$$H_{ii}=-\sum_{k=1,\mathrm{k}\neq i}^{n}H_{ik} \tag{30}$$

Finally, we can solve Eq. (29) with the boundary conditions in Eq. (18) and obtain the nodal values of potential and current density on the boundary.

4 Case study

The MKIBNM method, for the first time, is applied to investigate the corrosion behavior of AE44 (Mg, Anode)–mild steel (Cathode) couple which is exposed to 1.6 wt.% NaCl (electrolyte) solution. The numerical model is shown in Figure 2. The basis function of MKI is $\boldsymbol{p}(s)=[1,s,s^2]$ and the correlation function is Gaussian function with $\omega=0.1$. The polarization data in Figure 1 from reference [22] are used as the boundary condition for the anode and the cathode surfaces. The conductivity of the electrolyte σ is 2.5 S/m.

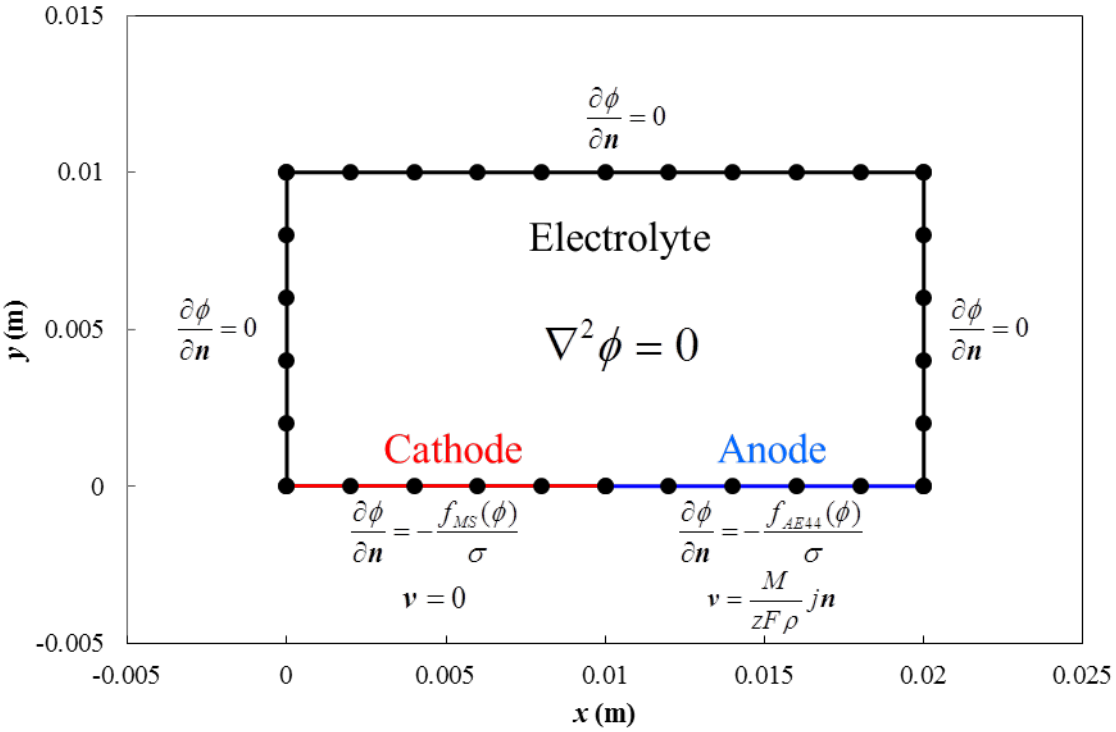


Figure 2. Numerical model of BNM for the corrosion of the galvanic couple.

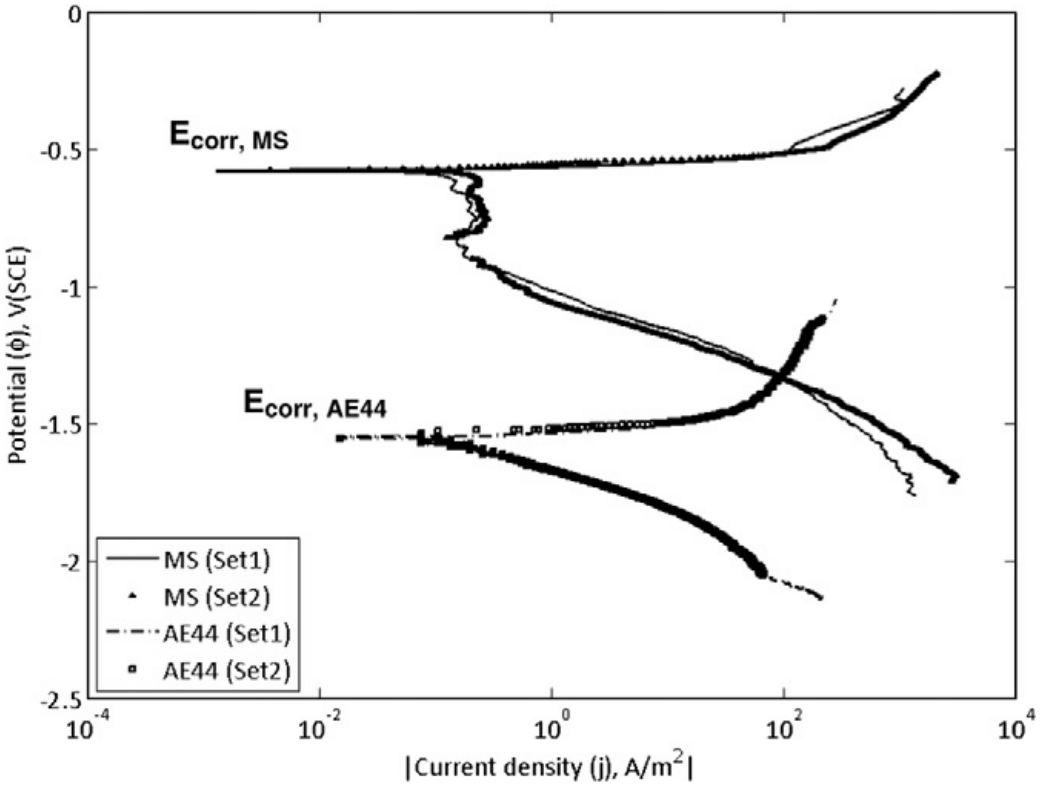


Figure 3. Polarization behaviour of mild steel and AE44 [22].

The initial current density along the anode and the cathode surfaces is plotted in Figure 4. It can be seen that the initial peak current density at the anodic region predicted by the model is 84.8 A/m^2 and the current density gradually decreases to around 29.1 A/m^2 away from the junction. The initial current density obtained from the experiment [21] [22] is also plotted in Figure 4 for comparing. The peak anodic current density at the junction of the couple obtained from the experiments is about 81.6 A/m^2. Thus, the numerical estimate of the peak anodic

current density is within 4% of that obtained from the experiment. The current density profile over the anodic and the cathodic regions obtained from MKIBNM agrees well with the profile obtained from the experiments.

The corrosion rate can be estimated from the anodic current density using Eq.(31).

$$C_R = \frac{M}{zF\rho} j \tag{31}$$

Where F is the Faraday constant, 96485.34 C/mol,

M is the atomic mass, 26.82 g/mol,

z is the electron number, 2,

ρ is the densityfor the anode AE44, 1820 kg/m^3,

C_R is corrosion rate,

j is the current density.

In the numerical model, the position of the moving anode surface can be evaluated from the transient current density.

$$(dx, dy) = \boldsymbol{v}\delta t = \frac{M}{zF\rho} j\boldsymbol{n}\delta t \tag{32}$$

The profile of the anode surface for AE44–mild steel couple obtained from the numerical model with time step $\delta t = 1\text{h} = 3600\text{s}$ after 3 days of constant exposure to the electrolyte solution is shown in Figure 5. It can be seen that a 1.8 mm deep pit at the AE44 side of the galvanic couple is predicted by the numerical model. The surface profile predicted using the numerical model can also be validated using the immersion experiment [21] [22]. The surface profile of AE44–mild steel couple after 3 days of immersion in 1.6 wt.% NaCl solution is shown in Figure 5. A 2.0 mm deep pit is formed at the junction. Thus, the numerical prediction of the pit depth is within 10% of that obtained from the immersion experiment for AE44 – mild steel couple.

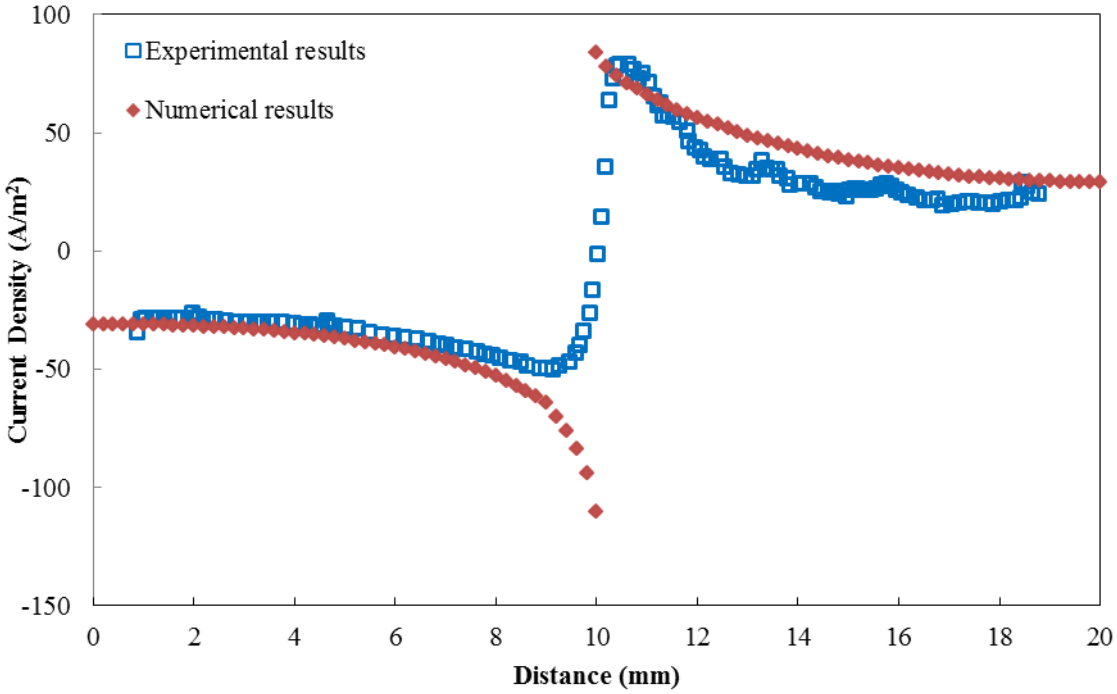


Figure 4 The initial spatial current density variation predicted using the numerical model and obtained from experiments

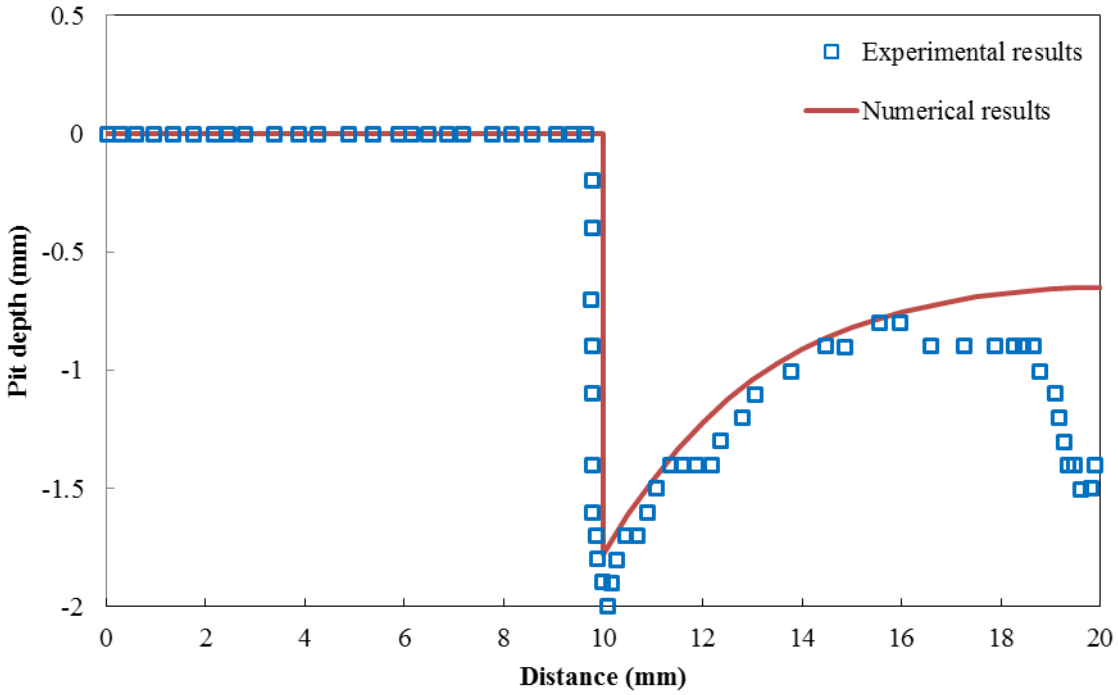


Figure 5 The surface profile predicted using the numerical model and obtained from the immersion experiment

5 Conclusions

1. A MKI-based BNM method developed in this work is capable of tracking a moving boundary during galvanic corrosion and can handle non-linear boundary conditions.
2. The numerical estimate of the peak anodic current density is within 4% of that obtained from the experiment. The numerical prediction of the pit depth is within 10% of that obtained from the experiment.
3. The MKI-based BNM is an effective and flexible method to simulate the process of galvanic corrosions.

Acknowledgement

This research was sponsored by Key Projects in the National Science & Technology Pillar Program during the Thirteenth Five-year Plan Period of China (No. 2016ZX05057012).

References

[1] Liu, G. R. and Liu, M. B. (2003) *Smoothed Particle Hydrodynamics — A Meshfree Particle Method*, World Scientific, Singapore.
[2] Cui, X. Y., Feng, H., Li, G. Y. and Feng, S. Z. (2015) A cell-based smoothed radial point interpolation method (CS-RPIM) for three-dimensional solids, *Engineering Analysis with Boundary Elements* **50**, 474-485
[3] Zheng, B. J. and Dai, B. D. (2011) A meshless local moving Kriging method for two-dimensional solids, *Applied Mathematics and Computation* **218**, 563-573
[4] Peng, M. J., Liu, P. and Cheng, Y. M. (2009) The Complex Variable Element-Free Galerkin (Cvefg) Method for Two-Dimensional Elasticity Problems, *International Journal of Applied Mechanics* **1**, 367-385
[5] Cheng, Y. M., Li, R. X. and Peng, M. J. (2012) Complex variable element-free Galerkin method for viscoelasticity problems, *Chinese Physics B* **21**, 090205
[6] Deng, Y. J., Liu, C., Peng, M. J. and Cheng, .Y. M. (2015) The Interpolating Complex Variable Element-Free Galerkin Method for Temperature Field Problems, International Journal of Applied Mechanics 7, 1550017
[7] Mukherjee, Y. X. and Mukherjee, S. (1997) The boundary node method for potential problems, *International Journal for Numerical Methods in Engineering* **40**, 797-815

[8] Chati, M. K., Mukherjee, S. and Mukherjee, Y. X. (1999) The boundary node method for three-dimensional linear elasticity, *International Journal for Numerical Methods in Engineering* **46**, 1163-1184
[9] Kothnur, V. S., Mukherjee, S. and Mukherjee, Y. X. (1999) Two-dimensional linear elasticity by the boundary node method, *International Journal of Solids and Structures* **36**, 1129-1147
[10] Chati, M. K. and Mukherjee, S. (2000) The boundary node method for three-dimensional problems in potential theory, *International Journal for Numerical Methods in Engineering* **47**, 1523-1547
[11] Lancaster, P. and Salkauskas, K. (1981) Surfaces generated by moving least squares methods, *Mathematics of Computation* **37**, 141-158
[12] Zhu, T., Zhang, J. and Atluri, S. N. (1998) A meshless local boundary integral equation (LBIE) method for solving nonlinear problems, *Computational Mechanics* **22**, 174-186
[13] Zhu, T., Zhang, J. D. and Atluri, S. N. (1998) A local boundary integral equation (LBIE) method in computational mechanics, and a meshless discretization approach, *Computational Mechanics* **21**, 223-235
[14] Liew, K. M., Cheng, Y. M. and Kitipornchai, S. (2005) Boundary element-free method (BEFM) for two-dimensional elastodynamic analysis using Laplace transform, *International Journal for Numerical Methods in Engineering* **64**, 1610-1627
[15] Liew, K. M., Cheng, Y. M. and Kitipornchai, S. (2006) Boundary element-free method (BEFM) and its application to two-dimensional elasticity problems, *International Journal for Numerical Methods in Engineering* **65**, 1310-1332
[16] Liew, K. M. and Cheng, Y. M. (2009) Complex variable boundary element-free method for two-dimensional elastodynamic problems, *Computer Methods in Applied Mechanics and Engineering* **198**, 3925-3933
[17] Ren, H. P., Cheng, Y. M. and Zhang, W. (2009) An improved boundary element-free method (IBEFM) for two-dimensional potential problems, *Chinese Physics B* **18**, 4065-4073
[18] Ren, H. P., Cheng, Y. M. and Zhang, W. (2010) An interpolating boundary element-free method (IBEFM) for elasticity problems, *Science China-Physics Mechanics & Astronomy* **53**, 758-766
[19] Li, X. G., Dai, B. D. and Wang, L. H. (2010) A moving Kriging interpolation-based boundary node method for two-dimensional potential problems, *Chinese Physics B* **19**, 120202
[20] Gu, L. (2003) Moving Kriging interpolation and element-free Galerkin method, *International Journal for Numerical Methods in Engineering* **56**, 1-11
[21] Deshpande, K. B. (2010). Experimental investigation of galvanic corrosion: comparison between svet and immersion techniques. *Corrosion Science*, **52**(9), 2819-2826.
[22] Deshpande, K. B. (2010). Validated numerical modelling of galvanic corrosion for couples: magnesium alloy (ae44)–mild steel and ae44–aluminium alloy (aa6063) in brine solution. *Corrosion Science*, **52**(10), 3514-3522.
[23] Trinh, D., Dauphin, D. P., Mengesha, T. U., Kish, J. R., & Mauzeroll, J. (2012). Influence of edge effects on local corrosion rate of magnesium alloy/mild steel galvanic couple. *Analytical Chemistry*, **84**(22), 9899-906.
[24] Jia, J. X., Song, G., and Atrens, A. (2007). Experimental measurement and computer simulation of galvanic corrosion of magnesium coupled to steel. *Advanced Engineering Materials*, **9**(1-2), 65-74.
[25] Murer, N., Oltra, R., Vuillemin, B., & Néel, O. (2010). Numerical modelling of the galvanic coupling in aluminium alloys: a discussion on the application of local probe techniques. *Corrosion Science*, **52**(1), 130-139.

Fluid-structure interaction eigenvalue analysis by using a coupled FE-BE solver

***Changjun Zheng[1], Chuanxing Bi[1], Haibo Chen[2], and Chuanzeng Zhang[3]**

[1]Institute of Sound and Vibration Research, Hefei University of Technology, China.
[2]Department of Modern Mechanics, University of Science and Technology of China, China
[3]Department of Civil Engineering, University of Siegen, Germany.

*Presenting & Corresponding author: cjzheng@hfut.edu.cn

Abstract

The vibration behavior of thin elastic structures is noticeably influenced by the surrounding water, which represents a heavy fluid. In this case, the feedback of the fluid pressure onto the structures cannot be neglected and a strong coupling scheme between the structural domain and the fluid domain is required. In this paper, a coupled finite element and boundary element (FE-BE) solver is developed for the modal analysis of three-dimensional submerged elastic structures. The structures are modeled by means of the finite element method (FEM). The compressibility of the surrounding fluid is taken into consideration, and thus the Helmholtz equation is used as the governing equation and solved by using the boundary element method (BEM). The resulting nonlinear eigenvalue problem (NEVP) is converted into a small linear one by using a contour integral method. A numerical example is finally given to demonstrate the effectiveness and applicability of the developed method.

Keywords: Fluid-structure interaction, Modal analysis, Coupled FE-BE method, Nonlinear eigenvalue problem, Contour integral method

Introduction

Finding resonances allows the designers to anticipate the structural vibration and to ensure that the resonance frequencies are distinct from those of the vibrating sources. In engineering applications, it is common to apply the FEM to perform the modal analysis of a structure *in vacuo* due to the high flexibility and applicability of the FEM to large-scale models. However, when the structure is submerged in a heavy fluid, e.g., water, a strong interaction between the structural domain and the fluid domain occurs and noticeably alters the resonance frequencies, especially for thin elastic structures [1]. Numerical simulations of the vibro-acoustic behavior of submerged structures usually require dealing with the fluid-structure interaction (FSI) since the feedback of the fluid pressure onto the structures can not be neglected for a heavy fluid. Thus, a scheme which takes the effect of the fluid on the structure into account should be used. Although some techniques such as the perfectly matched layer (PML) can be used to simulate the infinite fluid domains, the FEM still has some troubles in solving exterior problems, for instance the questions related to the position, size and parameter settings of the PML, and also the consequently big discretized model. By contrast, the BEM is much more favorable for the numerical solution of exterior problems since only the boundary of the structural domain has to be discretized and the Sommerfeld radiation condition can be satisfied automatically by the choice of the fundamental solution [2]. As a result, the coupled FE-BE methods are usually preferred for the numerical solution of the FSI problems [3][4].

In the numerical modal analysis of submerged structures, the fluid is sometimes assumed to be incompressible and hence modeled by the Laplace equation for simplicity [5]. The effect of the fluid on the structure can be regarded as adding mass and then a generalized eigenvalue

problem (GEVP) which is easy to solve can be obtained. However, when the compressibility of the fluid is taken into account and thus the fluid is modeled by the Helmholtz equation, the resulting eigenvalue problem is nonlinear since the frequency parameter appears nonlinearly in the boundary integral formulations of the Helmholtz equation. To solve such a NEVP, one scheme is to set up a GEVP by treating the term involving wave number in the Helmholtz equation as a non-homogeneous term. The fundamental solution of the Laplace equation is applied in the boundary integral formulations instead of that of the Helmholtz equation. The volume integrals caused by the non-homogeneous term can be transformed into boundary integrals by means of various methods, such as the dual reciprocity method [6] and the radial integration method [7]. Some other schemes which are presented for instance in [8][9] are based on the polynomial approximations of the coupled FE-BE coefficient matrix.

In addition to the approaches mentioned above, a group of methods based on contour integrals [10]-[12] have been recently developed. Through the use of these methods, a NEVP can be easily converted into a GEVP whose dimension is much smaller than the original NEVP. The eigenvalues lying inside a domain enclosed by a prescribed contour path can then be extracted by solving the small GEVP. The conversion is achieved directly by solving a series of linear systems of equations along the contour path. Since these systems of equations are independent and in the similar form as the ones arising in the response analysis, the big advantages of the contour integral methods are that they are very easy to be implemented and more suitable to be parallelized effectively. So far some of these methods have already been applied to solve some NEVPs in engineering applications. For instance, the method proposed by Asakura *et al*. [10] has been applied to solve the acoustic eigenvalue problems in [13] and to conduct the band structure analysis of phononic crystals in [14]. Kimeswenger *et al*. [15] analyzed the approximation of an FSI eigenvalue problem and used Beyn's method [11] for the numerical solution of the discretized NEVP. In this paper, a coupled FE-BE solver is developed for the modal analysis of three-dimensional submerged structures. The resulting NEVP is converted into a small linear one by using the contour integral method proposed by Asakura *et al*. [10]. Numerical implementation of the method in the FSI eigenvalue problems is given and some discussions are also given to further improve the efficiency and effectiveness of the method. A numerical example is employed finally to demonstrate the applicability and effectiveness of the developed FSI modal analysis method.

Formulation

In this section, a NEVP is first formulated for the modal analysis of an elastic structure which is submerged in an infinite fluid domain. The NEVP is then converted into a small GEVP by using a contour integral method proposed in [10]. Numerical implementation of the method in the modal analysis of submerged elastic structures is given in detail and some discussions are also given to improve the efficiency and avoid missing the resonance frequencies of interest.

FSI eigenvalue problems

Modal analysis of an elastic structure which is submerged in an infinite compressible inviscid fluid domain is discussed in this paper. If the structure is subjected to a time-harmonic load, we can derive an FEM system of equations in the frequency domain as

$$\left(\mathbf{K}-\omega^2\mathbf{M}\right)\mathbf{u}=\mathbf{f}_s+\mathbf{f}_f \tag{1}$$

where $\mathbf{K}$ and $\mathbf{M}$ are the global stiffness and mass matrices of the structure, ω is the circular frequency, $\mathbf{u}$ is the nodal displacement vector, $\mathbf{f}_s$ and $\mathbf{f}_f$ are the vectors with respect to the nodal values for the structural excitation force and fluid interaction force, respectively.

Because the compressibility of the fluid is taken into account in this paper, the propagation of time-harmonic acoustic waves in the fluid is described by the Helmholtz equation, which can be recast into a Kirchhoff-Helmholtz boundary integral equation (HBIE). It is widely known that the BEM based on the HBIE suffers from the fictitious eigenfrequency problem or the non-unique solution difficulty at the eigenfrequencies of the associated interior problems [16]. To overcome this difficulty, the Burton-Miller formulation [16] which is a linear combination of the HBIE and its normal derivative is adopted in this paper. Discretizing the HBIE or the Burton-Miller formulation and collecting the equations for all collocation points allow us to obtain a BEM system of equations as

$$\mathbf{Hp} = i\rho\omega\mathbf{Gv} + \mathbf{p}_i \tag{2}$$

where **p** and **v** are the vectors with respect to the nodal values for the sound pressure and the normal velocity on the fluid-structure interface, **H** and **G** are the BEM coefficient matrices corresponding to **p** and **v**, i is the imaginary unit, ρ is the mass density of the fluid, and $\mathbf{p}_i$ is the vector for the incident wave on the fluid boundary.

In the coupled FE-BE method for the numerical analysis of the FSI problems, Eqs. (1) and (2) have to be linked up via the coupling conditions across the fluid-structure interface to obtain a fully coupled system of equations. Firstly, considering the continuity of the normal surface velocity on the interface, we obtain

$$\mathbf{v} = -i\omega\mathbf{L}^{-1}\mathbf{T}_{fs}\mathbf{u} \tag{3}$$

where $\mathbf{L} = \int_{\Gamma_I} \mathbf{N}_f^T\mathbf{N}_f \, \mathrm{d}\Gamma$ and $\mathbf{T}_{fs} = \int_{\Gamma_I} \mathbf{N}_f^T\mathbf{n}\mathbf{N}_s \, \mathrm{d}\Gamma$, $\mathbf{N}_f$ and $\mathbf{N}_s$ are the BEM and FEM interpolation functions for the fluid and structural domains, respectively. $\mathbf{N}_f^T$ is the transpose of $\mathbf{N}_f$ and $\mathbf{n}$ is the unit normal vector on the fluid-structure interface Γ_I. In addition, the interaction force vector $\mathbf{f}_f$ represents the effect of the sound pressure on the structure and can be calculated by

$$\mathbf{f}_f = \mathbf{T}_{sf}\mathbf{p} \tag{4}$$

where $\mathbf{T}_{sf} = \mathbf{T}_{fs}^T = \int_{\Gamma_I} \mathbf{N}_s^T\mathbf{n}\mathbf{N}_f \, \mathrm{d}\Gamma$.

An appropriate scheme to generate a fully coupled system of equations is to substitute the FEM system into the BEM system with the use of Eqs. (3) and (4) to generate

$$\mathbf{Ap} = \mathbf{Bf}_s + \mathbf{p}_i \tag{5}$$

where $\mathbf{A} = \mathbf{H} - \mathbf{B}\mathbf{T}_{sf}$, $\mathbf{B} = \mathbf{G}\mathbf{W}$, $\mathbf{W} = \rho\omega^2\mathbf{L}^{-1}\mathbf{T}_{fs}\mathbf{A}_s^{-1}$ and $\mathbf{A}_s = \mathbf{K} + i\omega\mathbf{C} - \omega^2\mathbf{M}$.

The coupled system, i.e., Eq. (5) is fundamental in the numerical analysis of the FSI problems. However, as its coefficient matrix **A** involves the frequency parameter implicitly, we obtain a NEVP in the FSI modal analysis for finding the eigenpairs (λ_j, ϕ_j^p) that satisfy

$$\mathbf{A}(\lambda_j)\phi_j^p = 0 \tag{6}$$

where λ_j is the eigenvalue and ϕ_j^p represents the corresponding eigenvector with respect to the sound pressure on the fluid-structure interface.

Eq. (6) has non-trivial solutions when the determinant of $\mathbf{A}(\lambda_j)$ is equal to zero. In general, it is not an easy task to solve such a problem directly, therefore, a contour integral method is employed next to convert such a NEVP into a small GEVP which is much easier to deal with.

Contour integral method

The contour integral method proposed by Asakura *et al*. [10] and usually referred to as the block Sakurai-Sugiura (bSS) method is introduced here. It is a projection method which can extract eigenvalues while preserving their multiplicities in a domain enclosed by a positively oriented Jordan curve. In this method, the projection is performed through two Hankel matrices $\mathbf{H}_1, \mathbf{H}_2 \in \mathbb{C}^{KL\times KL}$, which are formed by

$$\mathbf{H}_1 = \left[\mathbf{M}_{j+l-2}\right]_{j,l=1}^{K} \text{ and } \mathbf{H}_2 = \left[\mathbf{M}_{j+l-1}\right]_{j,l=1}^{K} \tag{7}$$

where the moments $\mathbf{M}_l \in \mathbb{C}^{L\times L}$ are defined by

$$\mathbf{M}_l = \frac{1}{2\pi i}\oint_C z^l \mathbf{V}^H \mathbf{A}^{-1}(z)\mathbf{V}\,\mathrm{d}z, \quad l = 0,1,\cdots,2K-1, \tag{8}$$

and *C* is a positively oriented closed Jordan curve in the complex plane, **V** is a nonzero matrix chosen as random, $\mathbf{V}^H$ is the conjugate transpose of **V**, **A** is the coefficient matrix of Eq. (5), *K* and *L* are positive integers.

It has been proved mathematically in [10] that the eigenvalues of the linear matrix pencil ($\mathbf{H}_2$, $\mathbf{H}_1$) are identical to those of the original NEVP lying inside *C*. After obtaining the eigenpairs (λ_j, ψ_j) of the matrix pencil, we can calculate the eigenvectors for the original problem by

$$\phi_j^p = \mathbf{S}\psi_j \tag{9}$$

where $\mathbf{S} = [\mathbf{S}_0, \mathbf{S}_1, \cdots, \mathbf{S}_{K-1}]$, and

$$\mathbf{S}_l = \frac{1}{2\pi i}\oint_C z^l \mathbf{A}^{-1}(z)\mathbf{V}\,\mathrm{d}z, \quad l = 0,1,\cdots,K-1, \tag{10}$$

It is found that the original NEVP has been converted into a GEVP whose dimension is much smaller than the original problem through the bSS method. The conversion can be achieved readily and directly by solving a series of linear system of equations, i.e., **AX**=**V**, which are independent and similar to the one used in the normal FSI analysis. Thus, the method is very easy to be implemented and suitable to be parallelized effectively. Next, the implementation of the bSS method in the numerical modal analysis of submerged structures is presented.

Numerical modal analysis of submerged structures

It is found from Eqs. (8) and (10) that two sets of contour integrals in the form of

$$\mathbf{I}_l = \frac{1}{2\pi i}\oint_C z^l f(z)\,\mathrm{d}z \tag{11}$$

have to be evaluated in numerical computation. In Eq. (11), $f(z) = \mathbf{V}^H\mathbf{A}^{-1}(z)\mathbf{V}$ for $\mathbf{M}_l$ and $f(z) = \mathbf{A}^{-1}(z)\mathbf{V}$ for $\mathbf{S}_l$. When the eigenvalues of interest are located in an interval of $[\lambda_{\min}, \lambda_{\max}]$ and the contour path *C* is chosen from a family of ellipse of $z = \gamma + \rho(\cos\theta + i\zeta\sin\theta), \theta \in [0, 2\pi]$, Eq. (11) can be shifted, scaled and approximated by the *N*-point trapezoidal rule to produce

$$\hat{\mathbf{I}}_l = \frac{1}{N}\sum_{j=1}^{N}\rho\left(\frac{z_j-\gamma}{\rho}\right)^l\left(\zeta\cos\theta_j + i\sin\theta_j\right) f\left(z_j\right) \tag{12}$$

where $z_j = \gamma + \rho(\cos\theta_j + i\zeta\sin\theta_j)$, $\theta_j = (2\pi/N)(j-1/2)$, $\gamma = (\lambda_{\max}+\lambda_{\min})/2$, $\rho = (\lambda_{\max}-\lambda_{\min})/2$ and ζ is a scaling factor. When ζ is set to 1, the contour path *C* turns into a circle.

As the results of using Eq. (12), $\hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2$ which are the shifted and scaled approximations of the two Hankel matrices $\mathbf{H}_1$, $\mathbf{H}_2$ can be obtained. In order to calculate the eigenpairs of the matrix pencil $(\hat{\mathbf{H}}_2, \hat{\mathbf{H}}_1)$, a singular value decomposition (SVD) is performed on $\hat{\mathbf{H}}_1$ to obtain

$$\hat{\mathbf{H}}_1 = \mathbf{P}\boldsymbol{\Sigma}\mathbf{Q}^H \tag{13}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are unitary matrices, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \cdots, \sigma_{KL})$ and $\sigma_1, \sigma_2, \cdots, \sigma_{KL}$ are nonnegative real numbers in descending order.

The original NEVP then can be converted into an ordinary linear eigenvalue problem to find the eigenpairs $(\hat{\lambda}_j, \hat{\psi}_j)$ of the following matrix:

$$\hat{\mathbf{H}}_3 = \mathbf{P}^H \hat{\mathbf{H}}_2 \mathbf{Q}\boldsymbol{\Sigma}^{-1} \tag{14}$$

After obtaining the eigenpairs of $\hat{\mathbf{H}}_3$, the original eigenvalue λ_j can be recovered by

$$\lambda_j = \gamma + \rho\hat{\lambda}_j \tag{15}$$

and the corresponding eigenvector ϕ_j^p with respect to the sound pressure on the fluid-structure interface can be calculate by

$$\phi_j^p = \mathbf{S}\mathbf{Q}\boldsymbol{\Sigma}^{-1}\hat{\psi}_j \tag{16}$$

The eigenvector ϕ_j^u with respect to the structural nodal displacements can then be obtained by solving

$$\mathbf{A}_s \phi_j^u = \mathbf{T}_{sf} \phi_j^p \tag{17}$$

In order to omit small singular values which bring irrelevant results, a threshold δ is used in the original bSS method [10] to truncate the SVD of $\hat{\mathbf{H}}_1$. However, this brings a difficulty to specify a proper value for δ. A large δ may cause a possibility of missing some eigenvalues of interest, while a small δ may not filter out irrelevant results totally. Moreover, it is pointed out by Sakurai *et al*. [17] that it is not necessary to take a large *N* to reduce the quadrature error in calculating contour integrals. However, it is found from our numerical experiments that irrelevant results cannot be filtered out by using a small δ, like $\delta = 10^{-12}$ used in [10], even with a large *N*. The reason is that not only the quadrature error of calculating contour integrals exists in engineering applications but also some other errors, e.g., the errors from the modeling, mesh discretization and solution of systems of equations. To filter out all irrelevant results, a large *N* (i.e., more integration points) is usually required to achieve a better performance of the filter, but unfortunately this makes the computational cost increase quickly.

In this paper, in order to truncate the SVD effectively and efficiently, the gaps between the singular values of $\hat{\mathbf{H}}_1$, i.e., $\sigma_1, \sigma_2, \cdots, \sigma_{KL}$ are tested and the SVD of $\hat{\mathbf{H}}_1$ is truncated first at the biggest gap. $\hat{\mathbf{H}}_3$ in Eq. (14) is then truncated according to the truncation of $\hat{\mathbf{H}}_1$, and as a result a set of eigenvalues denoted by Λ_{gap} can be separated out by finding the eigenvalues of the truncated matrix of $\hat{\mathbf{H}}_3$. At the same time, the SVD of $\hat{\mathbf{H}}_1$ is also truncated by a threshold δ and another set of eigenvalues denoted by Λ_{thr} can be obtained. The components in the two sets, i.e., Λ_{gap} and Λ_{thr}, are then checked. If every number in Λ_{gap} is inside the contour and every number in the difference set $\Lambda_{thr} \setminus \Lambda_{gap}$ (i.e., $\{\lambda \in \Lambda_{thr} \mid \lambda \notin \Lambda_{gap}\}$) is outside the contour, the numbers in Λ_{gap} are taken as the final numerical solutions. Otherwise, a large *N* is required. Thus, it is found that the threshold δ is now used only to check if the truncation at the biggest gap is reasonable or not, and to make sure that no eigenvalue of interest is missed. Moreover, the present truncation scheme can be treated as a stopping criterion, and thus a small *N* can be

given initially in the bSS method and increased gradually until the final solutions are obtained. To further improve the efficiency, the integration points of the trapezoidal rule are evenly distributed with respect to the angle θ and doubled for the increase of *N*. As a result only the systems of equations at the new integration points need to be solved and the rest have already been solved at the previous steps.

Numerical example

An elastic spherical shell structure is employed in this section as a numerical example to show the effectiveness and applicability of the present numerical tool for the modal analysis of three-dimensional submerged structures. The shell structure is made of steel, and the material properties for the structure and the surrounding water are listed in Table 1. The structure has the outer radius of a = 5.0m and the thickness of h = 0.05m.

Table 1. Material properties for the structures and water

Density (structures)	ρ_s	7800	kg/m^3
Young's modulus (structures)	E	210	GPa
Poisson's ratio (structures)	v	0.3	-
Density (water)	ρ	1000	kg/m^3
Speed of sound (water)	C_f	1482	m/s

In the numerical analysis, the structure is modeled into a finite element mesh with 600 shell elements, which corresponds to 6492 DOFs. The fluid-structure interface is discretized into a boundary element mesh with 600 discontinuous quadratic elements, which corresponds to 4800 DOFs. Resonance frequencies in an interval of [34.0, 82.0]Hz are calculated, so that an elliptical path with $\gamma = (58.0, 0)$, $\rho = 24.0$ and $\zeta = 0.05$ can be employed as the contour path. The parameters used in the modified block SS method are set as K = 4, L = 15 and $\delta = 10^{-12}$. The computation terminates automatically at N = 16, and the computed eigenfrequencies are listed in the left part of Table 2. It is observed that the eigenfrequencies whose multiplicities are equal to or larger than one can both be extracted by using the present numerical tool. In addition, it is observed that the imaginary parts of these numerical eigenfrequencies are all negative, which implies that they are physically related to the radiation damping. This example is also analyzed numerically in [9], where a polynomial approximation method is used to solve the underlying NEVP. The calculated eigenfrequencies therein are 55.84-1.18*i*, 70.48-0.31*i* and 80.59-0.042*i*, and the multiplicities for them are 5, 7 and 9, respectively. It can be found that the numerical eigenfrequencies obtained by the present numerical tool are very close to the numerical results presented in [9].

The fluid-loaded modes (sometimes also called the wet modes) of the spherical shell structure, which are obtained by the developed FSI eigensolver are illustrated in Fig. 1. In Table 2, the computed eigenfrequencies of the *in vacuo* structure are also given, where the finite element based eigenvalue problem is solved by the bSS method (indicated by SS-FEM) and ANSYS, respectively. Another elliptical path with $\gamma = (137.0, 0)$, $\rho = 17.0$ and $\zeta = 0.05$ is used in the SS-FEM, and the parameters utilized in the bSS method are set the same as the ones used above. The computation terminates automatically at N=32. The computed eigenfrequencies are listed in the right part of Table 2, and their mode shapes are similar to the fluid-loaded modes of the eigenfrequencies listed in the left part of Table 2. It can be observed that the real parts of the numerical results obtained by the SS-FEM are equal to the numerical results obtained by ANSYS, and the imaginary parts are very small and can be neglected directly. Furthermore,

as can be seen, the fluid has a significant influence on the eigenfrequencies of the submerged elastic structure. All frequencies are lowered due to the fluid, that is because the surrounding fluid acts like an adding mass to the submerged structure. Accordingly, such variations of the eigenfrequencies necessitate the solution of the coupled eigenvalue problem.

Table 2. Eigenfrequencies of the spherical shell structure

i	Frequencies (Hz, with fluid)	Frequencies (Hz, no fluid)	
	SS-FEM-BEM	SS-FEM	ANSYS
1	$55.84 - 1.18i$	$120.91 - 3.07\times10^{-14}i$	120.91
2	$55.84 - 1.18i$	$120.91 - 2.36\times10^{-14}i$	120.91
3	$55.84 - 1.18i$	$120.91 - 4.16\times10^{-14}i$	120.91
4	$55.84 - 1.18i$	$120.91 + 2.74\times10^{-14}i$	120.91
5	$55.84 - 1.18i$	$120.91 + 1.30\times10^{-13}i$	120.91
6	$70.48 - 0.31i$	$143.22 - 2.41\times10^{-13}i$	143.22
7	$70.48 - 0.31i$	$143.22 + 4.57\times10^{-13}i$	143.22
8	$70.48 - 0.31i$	$143.22 + 1.91\times10^{-13}i$	143.22
9	$70.48 - 0.31i$	$143.22 + 9.27\times10^{-14}i$	143.22
10	$70.48 - 0.31i$	$143.22 + 2.32\times10^{-13}i$	143.22
11	$70.48 - 0.31i$	$143.22 + 9.42\times10^{-13}i$	143.22
12	$70.48 - 0.31i$	$143.22 + 5.34\times10^{-13}i$	143.22
13	$80.59 - 0.042i$	$152.10 - 6.35\times10^{-13}i$	152.10
14	$80.59 - 0.042i$	$152.11 + 1.26\times10^{-13}i$	152.11
15	$80.59 - 0.042i$	$152.11 + 3.58\times10^{-14}i$	152.11
16	$80.59 - 0.042i$	$152.11 - 6.37\times10^{-14}i$	152.11
17	$80.60 - 0.042i$	$152.12 + 1.76\times10^{-13}i$	152.12
18	$80.60 - 0.042i$	$152.12 + 1.69\times10^{-13}i$	152.12
19	$80.60 - 0.042i$	$152.12 - 6.41\times10^{-14}i$	152.12
20	$80.61 - 0.042i$	$152.13 - 2.45\times10^{-14}i$	152.13
21	$80.61 - 0.042i$	$152.13 - 5.88\times10^{-15}i$	152.13

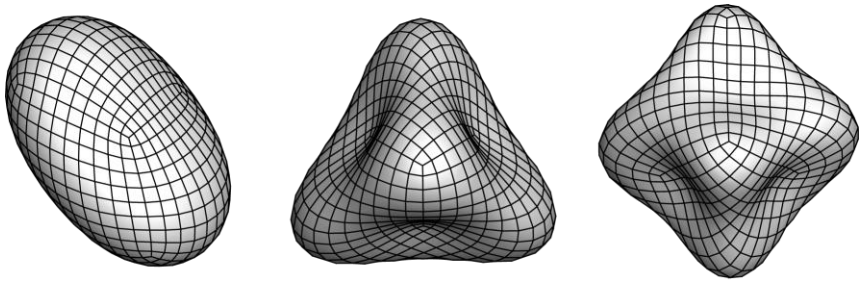


Figure 1. Mode shapes of the submerged spherical shell structure

Conclusions

In this paper, the numerical modal analysis of three-dimensional submerged elastic structures is carried out by using a coupled FE-BE solver. The submerged structure is modeled by the FEM. The compressibility of the surrounding infinite fluid domain is taken into account and hence the Helmholtz equation serves as the governing equation and is solved by the BEM. A contour integral method proposed by Asakura *et al*. [10] is employed to convert the resulting NEVP into a small GEVP. Numerical implementation of the method in the FSI eigenvalue

problems is given in detail. In order to improve the efficiency of the method and also avoid missing the eigenvalues of interest, a novel scheme for the truncation of the small singular values is presented. This scheme is then used as a stopping criterion, and thus a small number of integration points for the numerical quadrature of contour integrals can be given initially and then increased gradually until the final solutions are obtained. The effectiveness and applicability of the present numerical tool for the modal analysis of submerged structures are shown by a numerical example of an elastic spherical shell structure. All eigenfrequencies of the spherical shell structure are lowered due to the surrounding fluid, and the variations of the eigenfrequencies necessitate the solution of the coupled eigenvalue problem.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (No. 11402071).

References

[1] Junger, M. C. and Feit, D. (1986) *Sound, structures, and their interaction*, MIT press, Cambridge.
[2] Ciskowski, R. D. and Brebbia, C. A. (1991) *Boundary element methods in acoustics*, Computational Mechanics Publications and Elsevier Applied Science, Southampton.
[3] He, Z. C., Liu, G. R., Zhong, Z.H., Zhang, G. Y. and Cheng, A. G. (2011) A coupled ES-FEM/BEM method for fluid-structure interaction problems, *Engineering Analysis with Boundary Elements* **35**, 140-147.
[4] Chen, L. L., Zheng, C. J. and Chen, H. B. (2014) FEM/wideband FMBEM coupling for structural-acoustic design sensitivity analysis, *Computer Methods in Applied Mechanics and Engineering* **276**, 1-19.
[5] Junge, M., Brunner, D. and Gaul, L. (2011) Solution of FE-BE coupled eigenvalue problems for the prediction of the vibro-acoustic behavior of ship-like structures, *International Journal for Numerical Methods in Engineering* **87**, 664-676.
[6] Nardini, D. and Brebbia, C. A. (1983) A new approach to free-vibration analysis using boundary elements, *Applied Mathematical Modelling* **7**, 157-162.
[7] Gao, X. W. (2002) The radial integration method for evaluation of domain integrals with boundary-only discretization, *Engineering Analysis with Boundary Elements* **26**, 905-916.
[8] Li, S. (2005) A state-space coupling method for fluid-structure interaction analysis of plates, *Journal of Acoustical Society of America* **118**, 800-805.
[9] Peters, H., Kessissoglou, N. and Marburg, S. (2013) Modal decomposition of exterior acoustic-structure interaction, *Journal of Acoustical Society of America* **133**, 2668-2677.
[10] Asakura, J., Sakurai, T., Tadano, H., Ikegami, T. and Kimura, K. (2009) A numerical method for nonlinear eigenvalue problems using contour integrals, *JSIAM Letters* **1**, 52-55.
[11] Beyn, W. J. (2011) An integral method for solving nonlinear eigenvalue problems, *Linear Algebra and its Applications* **436**, 3839-3863.
[12] Xiao, J., Meng, S., Zhang, C. and Zheng C. (2016) Resolvent sampling based Rayleigh-Ritz method for large-scale nonlinear eigenvalue problems, *Computer Methods in Applied Mechanics and Engineering* **310**, 33-57.
[13] Zheng, C. J., Gao, H. F., Du, L., Chen, H. B. and Zhang, C. (2016) An accurate and efficient acoustic eigensolver based on a fast multipole BEM and a contour integral method, *Journal of Computational Physics* **305**, 677-699.
[14] Gao, H., Xiang, J., Zheng, C., Jiang, Y. and Matsumoto, T. (2015) BEM-based analysis of elastic banded material by using a contour integral method, *Engineering Analysis with Boundary Elements* **53**, 56-64.
[15] Kimeswenger, A., Steinbach, O. and Unger, G. (2014) Coupled finite and boundary element methods for fluid-solid interaction eigenvalue problems, *SIAM Journal on Numerical Analysis* **52**, 2400-2414.
[16] Burton, A. J. and Miller, G. F. (1971) The application of integral equation methods to the numerical solution of some exterior boundary-value problems, *Proceedings of the Royal Society of London. Series A*, **323**, 201-210.
[17] Sakurai, T., Futamura, Y. and Tadano H. (2013) Efficient parameter estimation and implementation of a contour integral-based eigensolver, *Journal of Algorithms & Computational Technology*, **7**, 249-269.

Numerical modeling of a hybrid GFRP-concrete beam subjected to low-velocity impact loading

†*Z. Li[1], A. Khennane[1], P. J. Hazell[1], and A. Remennikov[2]

[1]School of Engineering and Information Technology, The University of New South Wales, Canberra, Australia.
[2]School of Civil, Mining and Environmental Engineering, University of Wollongong, Australia.

*Presenting author: zongjun.li@student.adfa.edu.au
†Corresponding author: zongjun.li@student.adfa.edu.au

Abstract

This study reports on the results of a numerical investigation on the impact behaviour of a hybrid GFRP-concrete beam subjected to low-velocity impacts with different impact energy levels. A non-linear finite element model has been developed using the commercial software ABAQUS/Explicit to simulate the impact performance of the hybrid beam under dynamic loading. The numerical results, including the dynamic loading history, failure modes and impact performance are compared to the experimental results. The agreement between the numerical and experimental results indicates that the developed numerical model is capable of analysing the impact behaviour of such hybrid GFRP-concrete system.

Keywords: Numerical modeling, Finite element method, Failure modes, Impact performance

Introduction

Fiber reinforced polymer (FRP) composite materials have been extensively used due to their excellent anti-corrosion performance. Composite materials in the form of pultruded glass fiber reinforced polymer (GFRP) profiles have a great potential since they are economically affordable through the pultrusion process, which offers the best productivity/cost ratio of all the composites fabrication processes [1]. However when used as structural elements on their own, pultruded profiles suffer from the susceptibility to instability, the high deformability and the brittle failure [2][3]. On the other hand, when combined in a hybrid form with concrete, they have shown to offer excellent performance as floors [4], bridge decks [5][6], beams [7]-[11], and rail sleepers [12] when subject to static loadings.

However, their dynamic performance has yet to be assessed. If these hybrid beams are to be used as bridge girders or railway sleepers, their impact behavior has to be investigated. The aim of this study therefore is to investigate the behavior of a hybrid beam made of a rectangular hollow pultruded profile filled with concrete using a numerical model because experimentation alone does not reveal the failure modes of the concrete hidden inside the pultruded profile. For this purpose, the numerical investigation will help to analyse the development and propagation of damage in the concrete.

Experimental program

To verify the numerical modeling of the hybrid GFRP-concrete beam, the experimental tests of this hybrid beam are reported briefly here. Six hybrid beams, which consisted of rectangular hollow section pultruded GFRP profiles filled with high performance concrete, were tested along the weak axis. The hybrid beams were subjected to a concentrated impact load by a cylindrical impactor for three ascending impact energy levels. The total length and the span length of each beam is 2000 mm and 1440 mm respectively. The schematic diagram of the test setup is shown in Fig. 1.

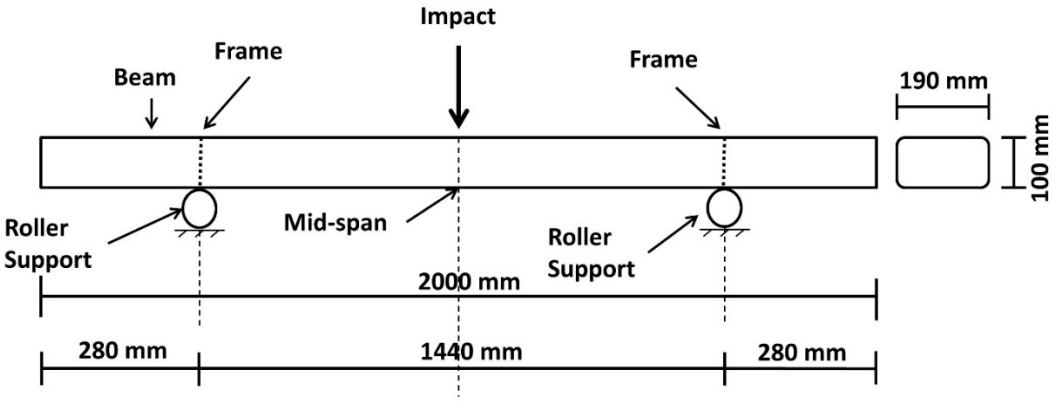


Figure 1. Schematic diagram of experimental setup

Numerical modeling

Finite element method

The experimental results can only provide the impact characteristics and performance for the hybrid GFRP-concrete beam. Details such as concrete damage or debonding between the concrete block and the pultruded profile are not visible. For this purpose, a non-linear finite element (FE) model was developed using the commercial software ABAQUS/Explicit to study these phenomena and predict the overall response of the hybrid GFRP-concrete beam subjected to low-velocity impacts. The impactor was modeled using discrete rigid elements (R3D4). A mass of 592 kg was assigned to the impactor as per experimental conditions. The concrete block is modelled using linear eight-node three dimensional solid elements with reduced integration (C3D8R). As reported in Li et al. (2017), the through-thickness properties of pultruded GFRP composites are negligible in low-velocity impact cases, therefore all the effective layers of the GFRP composites were modeled by using eight-node quadrilateral in-plane general-purpose continuum shell elements (SC8R) [13]. The orientation of the fibers in each layer was assigned accordingly to the corresponding coordinate systems. The mesh density was chosen to 20 mm × 20 mm for the hybrid beam members on the basis of mesh sensitivity analysis in terms of computational time and convergence solution. Three different initial impact velocities were imposed to the impactor to simulate the impact events. The boundary conditions are shown in Fig. 2 according to the experimental setup. The interaction between the impactor and the hybrid beam was simulated by surface-to-surface contact pairs. The mechanical constraint formulation was enforced using the kinematic contact algorithm, and the friction coefficient was set to 0.5. The interaction between the pultruded profiles and the concrete was also simulated by surface-to-surface contact pairs but with 0.3 as the friction coefficient factor.

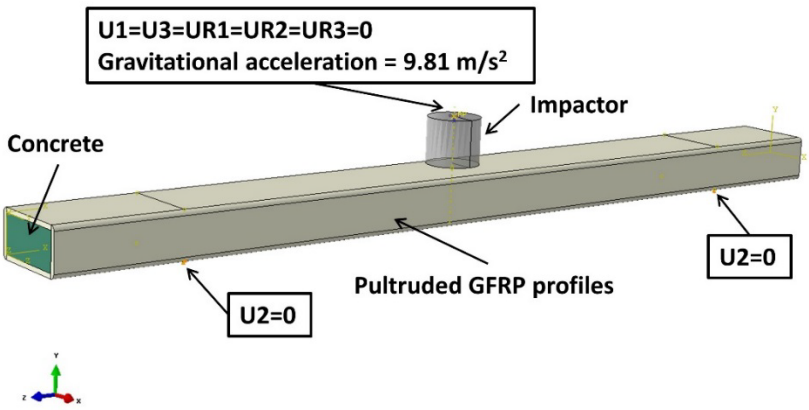


Figure 2. Assembled FE model with boundary conditions

Material models

Two material models of high performance concrete and pultruded GFRP composites were used in this study. The concrete section of the hybrid beam was modeled by the use of concrete damaged plasticity model (CDPM) and, the pultruded profiles were modeled using the FRP composite material model (Hashin damage model). For the compressive behaviour of concrete, two stress-strain relationship models were used to represent the ascending and descending zone respectively. The expression developed by Popovics (1973), modified by Thorenfeldt et al. (1987), were used to describe the compressive hardening behaviour (ascending zone), as shown in Eq. (1) [14][15].

$$\frac{f_c}{f_c^{'}} = \frac{\varepsilon_c}{\varepsilon_c^{'}} \cdot \frac{n}{n-1+\left(\frac{\varepsilon_c}{\varepsilon_c^{'}}\right)^{nk}} \tag{1}$$

The expression proposed by Wee et al. (1996) was used to describe the strain softening behaviour (descending zone), as shown in Eq. (2) [16].

$$f_c = f_c^{'}\left\{\frac{k_1^{'}\beta\left(\frac{\varepsilon_c}{\varepsilon_c^{'}}\right)}{k_1^{'}\beta-1+\left(\frac{\varepsilon_c}{\varepsilon_c^{'}}\right)^{k_2^1\beta}}\right\} \tag{2}$$

The tension stiffening behaviour of the concrete was defined with the post-failure stress as a function of cracking strain in this study. The compressive and tensile stress-strain curves, shown in Fig. 3 and Fig. 4, were implemented in CDPM.

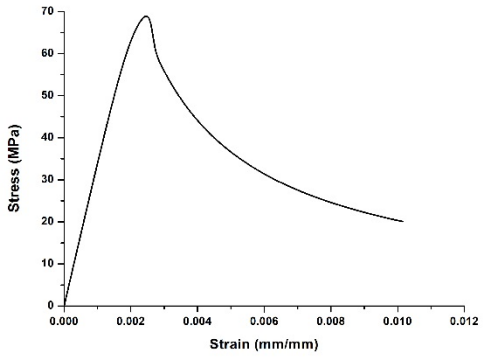


Figure 3. Compressive stress-strain curve

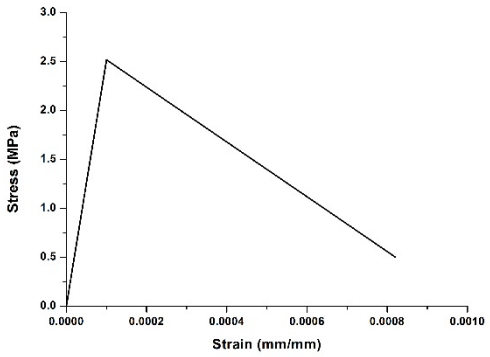


Figure 4. Tensile stress-strain curve

The FRP damage model for fiber-reinforced composites were used to predict the onset of failure and post-failure development of the pultruded GFRP profiles. Details of mechanisms of CDPM and fiber-reinforced material model can be found in ABAQUS documentation [17]. The material properties and input parameters are shown in Table 1-4 below.

Table 1. Material properties of pultruded GFRP profiles [13]

$E_1[MPa]$	$E_2[MPa]$	$G_{12}=G_{13}^{*}=G_{23}^{*}[MPa]$	ν_{12}	$\varepsilon_f^T=\varepsilon_f^{C*}$	$\varepsilon_t^T=\varepsilon_t^{C*}=\varepsilon^{S*}$
28870	3505	2980	0.21	0.011	0.013

Table 2. Damage initiation parameters of pultruded GFRP profiles [13]

$X^T[MPa]$	$X^C[MPa]$	$Y^T[MPa]$	$Y^C[MPa]$	$S^L[MPa]$	$S^{T*}[MPa]$
301.198	310.785	29.78	31.97	33.0	33.0

Table 3. Damage evolution parameters for pultruded GFRP profiles [13]

$G^*_{ft,c}[N/mm]$	$G^*_{fc,c}[N/mm]$	$G^*_{mt,c}[N/mm]$	$G^*_{mc,c}[N/mm]$
55.0	95.0	11.5	11.5

Table 4. Material properties of concrete

Cross-section, [mm]	Modulus of elasticity, [MPa]	Poisson's ratio	Maximum compressive strength, [MPa]	Maximum tensile strength, [MPa]
170 × 80	33930	0.18	69.84	2.52

*Assumed

Numerical results

Failure modes

The predicted results were compared to the experimental ones. The numerical results identified four main failure modes. Taking the 450 mm dropping height case (impact velocity = 2.97 m/s) as an example, as shown in Fig. 5, localised concrete crushing occurred at the impact area. Moreover, fractural cracks can be found between the mid-span and the supports. The concrete fractured into several parts due to the impact loading. Shear cracks on the corners of the profiles can be observed in both the numerical and experimental results as shown in Fig. 6. The fourth and the last failure mode is the slip of the concrete. Debonding between the concrete and the profiles occurred in both the numerical and experimental results during the tests as shown in Fig. 7.

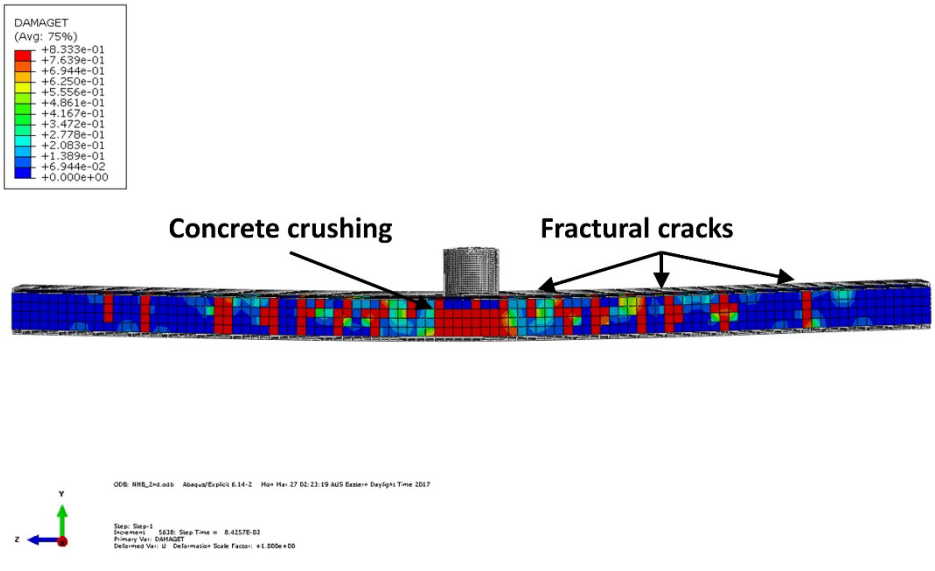


Figure 5. Fractural cracks of the concrete inside

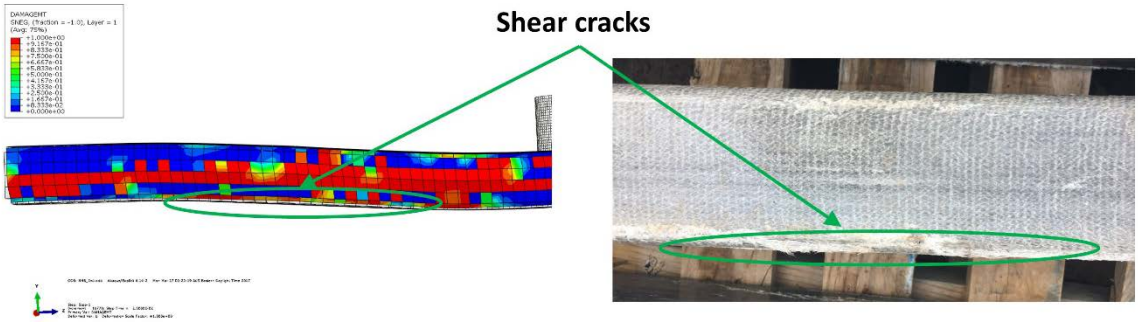


Figure 6. Shear cracks on the corners of the profiles for both numerical and experimental observations

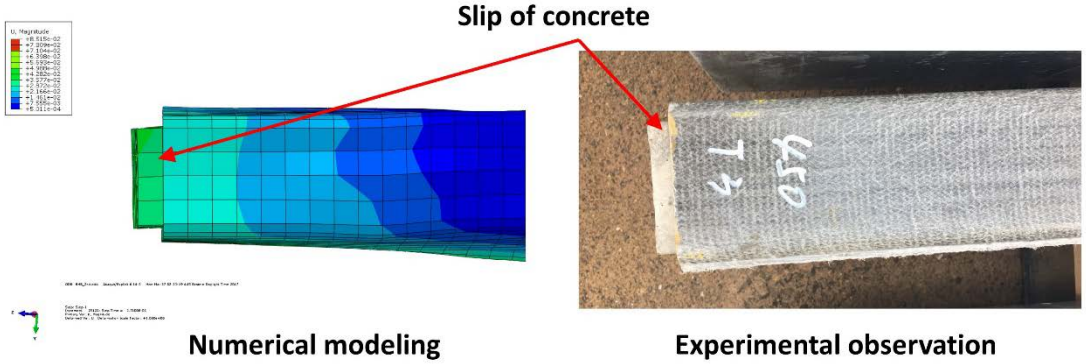


Figure 7. Slip of concrete for both numerical and experimental observations

Loading history and impact performance

The comparisons of the numerical and experimental results for three loading histories are shown in Fig. 8. A reasonably good agreement is achieved between the experimental results and numerical ones. The typical loading histories for all impact energy levels can be divided into two stages, inertial resistance stage and dynamic bending resistance stage. Just after contact is initiated between the impactor and the top surface of the hybrid beam, the first stage is represented by a significant rapid increase in load to the maximum value, and dropping back to zero. In this stage, the impact force is represented by a rapid, short peak of inertial force due to the striking drop mass on the contact zone. The inertial force increases and then decreases quickly as the velocity of the hybrid beam increases. In the numerical prediction, the impactor was modeled as a rigid part instead of a steel impactor. This could explain why the predicted values in this stage are relatively higher than the experimental ones. The true impact resistance of the hybrid beam is represented by the second stage. In this stage, the hybrid beam starts to carry the impact load until failure occurs. Multiple failure mechanisms occur in this stage, including the debonding failure between the concrete and the profiles, shear cracks on the corners of the profiles and the fracture of the concrete. The average maximum experimental impact loads recorded in this stage are 84.8, 100.3 and 110.2 kN, respectively for the three ascending energy levels. The corresponding predicted values are 96.9, 104.9 and 105.5 kN respectively. The average difference between experimental results and numerical ones is only 7 %. The numerical predictions are found to corroborate the experimental results in terms of failure modes and impact performance.

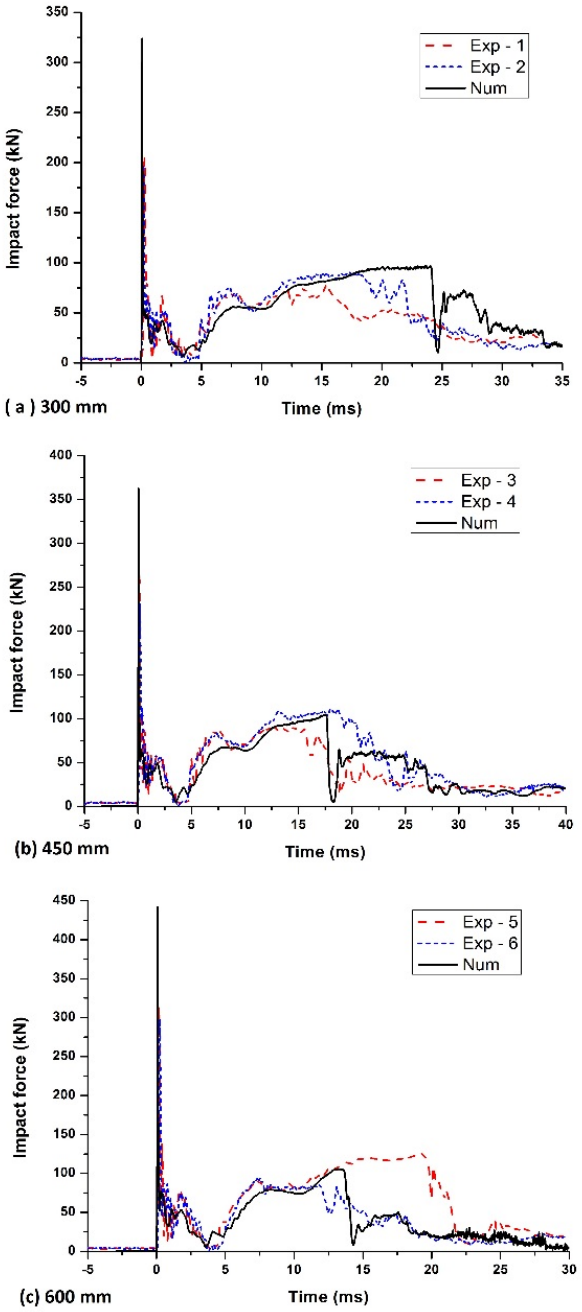


Figure 8. Comparison of loading histories for three ascending energy levels

Conclusions

A numerical investigation of the impact resistance of hybrid GFRP-concrete beams subjected to low-velocity impacts has been conducted. Three dimensional FE models were developed to predict the impact behaviour of this hybrid system. Multi-nonlinear stress-strain relationships were combined to describe the material models. The numerical results including the dynamic loading histories, failure modes and impact performance are compared and verified with the experimental results. The agreement between numerical and experimental results indicates that the developed numerical model is capable of analysing the impact behaviour of hybrid GFRP-concrete systems.

References

[1] Zureick, K. and Scott, D. (1997) Short-term behavior and design of fiber-reinforced polymeric slender members under axial compression, ASCE, *Journal of Composites for Construction* **1(4)**, 140-149.

[2] Correia, J. R., Branco, F. A. and Ferreira, J. G. (2007) Flexural behaviour of GFRP-concrete hybrid beams with interconnection slip, *Composite Structures* **77(1)**, 66-68.

[3] Gonilha, J. A., Correia, J. R. and Branco, F. A. (2013) Dynamic response under pedestrian load of a GFRP-SFRSCC hybrid footbridge prototype: Experimental tests and numerical simulation, *Composite Structures* **95**, 453-463.

[4] Correia, J. R., Branco, F. A. and Ferreira, J. (2009) GFRP-concrete hybrid cross-sections for floors of buildings, *Engineering Structures* **31(6)**, 1331-1343.

[5] He, J., Liu, Y., Chen, A. and Dai, L. (2012) Experimental investigation of movable hybrid GFRP and concrete bridge deck, *Construction and Building Materials* **26(1)**, 49-64.

[6] Mendes, P. J. D., Barros, J. A. O., Sena-Cruz, J. M. and Taheri, M. (2011) Development of a pedestrian bridge with GFRP profiles and fiber reinforced self-compacting concrete deck, *Composite Structures* **93(11)**, 2969-2982.

[7] Deskovic, N., Triantafillou, T. C. and Meier, U. (1995) Innovative design of FRP combined with concrete: short term behaviour, ASCE, *Journal of Structural Engineering* **121(7)**, 1069-1078.

[8] Canning, L., Hollaway, L. and Thorne, A. M. (1999) Manufacture, testing and numerical analysis of an innovative polymer composite/concrete structural unit, *Proceedings-institution of civil engineers, structures and buildings* **134**, 231-241.

[9] Hulatt, J., Hollaway, L. and Thorne, A. (2003) Short term testing of hybrid T beam made of new prepreg material, *Journal of Composites for Construction* **7(2)**, 135-144.

[10] Khennane, A. (2009) Manufacture and testing of a hybrid beam using a pultruded profile and high strength concrete, *Australian Journal of Structural Engineering* **10(2)**, 145-155.

[11] Chakrabortty, A., Khennane, A., Kayali, O. and Morozov, E. (2011) Performance of outside filament-wound hybrid FRP-concrete beams, *Composites Part B: Engineering* **42(4)**, 907-915.

[12] Ferdous, W., Manalo, A., Khennane, A. and Kayali, O. (2015) Geopolymer concrete-filled pultruded composite beams – Concrete mix design and application, *Cement and Concrete Composites* **58**, 1-13.

[13] Li, Z., Khennane, A., Hazell, P. J. and Brown, A. D. (2017) Impact behaviour of pultruded GFRP composites under low-velocity impact loading, *Composite Structures* **168**, 360-371.

[14] Popovics, S. (1973) A numerical approach to the complete stress-strain curve of concrete, *Cement and Concrete Research* **3**, 583-599.

[15] Thorenfeldt, E., Tomaszewicz, A. and Jensen, J. J. (1987) Mechanical properties of high strength concrete and application in design, *Proceeding of the Symposium of Utilization of high strength concrete*, Trondheim, 149-159.

[16] Wee, T. H., Chin, M. S. and Mansur, M. A. (1996) Stress-strain relationship of high-strength concrete in compression, *Journal of Materials in Civil Engineering* **8(2)**, 70-76.

[17] ABAQUS, ABAQUS Documentation, Dassault Systèmes, Providence, RI, USA, 2011.

Simulation and experimental validation of hydraulic collecting in deep-ocean mining

***Guocheng Zhao[1,2], †Longfei Xiao[1,2], Weijie Zhao[1], and Yangrui Cheng[3]**

[1]State Key Laboratory of Ocean Engineering, Shanghai Jiao Tong University, China
[2]Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, China
[3]State Key Laboratory of Exploitation and Utilization of Deepsea Mineral Resources, China

*Presenting author: guocheng.zhao@sjtu.edu.cn
†Corresponding author: xiaolf@sjtu.edu.cn

Abstract

Mineral-collecting technology is the core technology of deep-ocean mining, which has a decisive influence on the feasibility and profitability of technology application. Hydraulic collecting method is considered as one of the most promising options to collect the ores in deep-ocean mining system. Therefore, the characteristics of flow field and suction force of single spherical ore particle in the hydraulic collecting was investigated, which have a significant effect on collecting efficiency.

The flow around a spherical manganese nodule placed under a collecting tube was investigated by using Detached Eddy Simulation (DES) and Large Eddy Simulation (LES) based on Shear Stress Transport (SST) model, respectively. A collecting test system was constructed in the laboratory and a series of 285 tests with various working parameter combinations were conducted. The sphere remains still during collecting and the vertical suction force was measured by a force transducer which is connected to the sphere and able to catch the vertical force with high precision. It is verified that the vertical force prediction for single ore particle in collecting condition by simulation with DES/LES method based on SST model is feasible and sufficiently precise, by comparing the results of numerical simulation and tests.

An insight into understanding the characteristics of flow field in hydraulic collecting for deep-ocean mining was provided by comprehensive analysis on dynamic change regularity of wake vortex structures and pressure distribution. The drag coefficients of the sphere with various working parameters were calculated based on test results and defined as approximate equation using dimensionless quantities.

Simulation and experimental validation of spiral flow collecting were also conducted and the values of vertical suction force were measured and compared to that of ordinary flow collecting. The study reveals that the value of vertical force of the sphere in spiral flow is greater than that in ordinary flow under the same test condition. One reason for that is the spiral flow brings about an increase of the transfer distance of low pressure. The results of this study can provide reference and support for mechanism study and optimal design of the hydraulic collecting system for deep-ocean mining.

Keywords: deep-ocean mining, hydraulic collecting, spiral flow, vertical force, simulation, experimental validation

Introduction

The world's growing economy demands more mineral extracted from the ocean. This demand requires the development of the deep-ocean mining technology. Polymetallic nodules are deposited over the ocean floor at water depth of 4,000～6,000 m. They are mostly spherical or ellipsoidal, with their long-axis length varying from 2 cm to 10 cm and a density of 2,100

(Liu et al., 2014)[1]. However, a profitable exploitation of deep-ocean mining is feasible on the premise that there is a nodule collector with maximum collecting capacity of 140 kg of wet nodules per second (Herrouin et al., 1989)[2]. As a result, effective collecting of manganese nodules out of sediment upper layer of deep seafloor is not only one of the key processes of the deep-ocean mining technology, but also the beginning of an economic and environmentally acceptable mining operation.

To pick up these nodules, a variety of collecting methods such as hydraulic methods, mechanical methods and hybrid collection methods have been developed. Commercial production must achieve high sweep efficiency (Chung, 1985)[3]. A sea test (the OMI Test, 1978) showed that hydraulic methods had a higher collecting efficiency than mechanical methods. It is also found that the hydraulic methods have better adaptability to the variation of seabed height than other methods (Zhao et al., 1995)[4].

In the process of hydraulic methods, polymetallic nodules are collected due to the force induced by local high-velocity flow. A number of established computational and experimental results about these methods are available. Hong et al. (1999) dealt with experimental approaches for enhanced understanding of the hydraulic performance of a hybrid pick-up device[5]. The experiments were conducted in a 2-D flume tank. By parametric experiments they found position and shape of baffle plates are significant factors for effective design of hydraulic nodule lifter. Yang et al. (2003) discussed major parameters and their influence on the performance of the pick-up device by tests[6]. The results showed the hydraulic pick-up device with proper dimension and parameters could get high pick-up rate and low content sediments. Chen and Jian (1996) carried out an experimental study on hybrid collection methods[7]. The results showed the relationship between the pick-up efficiency and other key factors.

The flow field of hydraulic collecting for deep-ocean mining is unique and worth of study. Lim et al. (2015) analyzed flow field characteristics with outflow discharge from a collecting device in deep seawater by FLUENT[8]. He revealed seawater velocity and streamline distributions along with complicated flow characteristics downstream including nodule particles behavior.

In contrast to the research about hydraulic collecting, much more studies on the flow around a sphere in a uniform flow field have been done. Achenbach (1972) had experiments on the flow past spheres in the Reynolds number range $5\times104\leq$ Re $\leq6\times104$[9]. He compared his results with other available data and pointed out the dependence of friction forces on Reynolds number. Tsutsui (2008) had an experimental study of the flow field and the aerodynamic force on a sphere above a plane[10]. The surface pressures on the sphere and the plane were measured by inclined multi-tube manometers and the results were compared with photographs showing the flow visualization of the sphere.

Johnson and Patel (1999) investigated the flow of an incompressible viscous fluid past a sphere at different Reynolds numbers by DES (Detached Eddy Simulation)[11]. They calculated the instantaneous and mean flow field around the sphere as well as the wake vortex, and validated their numerical results by flow-visualization experiments. On the basis of their analysis, a mechanism driving the transition to unsteady flow was proposed. In the study of Constantinescu and Squires (2003), LES (Large Eddy Simulation) and DES were applied to predict and investigate the flow around a sphere at a Reynolds number of 10,000 in the subcritical regime[12]. Comparison of the computed results with experimental data showed

that LES and DES were able to simulate the flow around a sphere.

A number of analyses on hydraulic lifting have also been attempted. Chung et al. (1998) conducted experimental investigation to study the shape effect of solids on pressure drop in a 2-phase vertically upward transport[13]. Their research showed that the spherical particles required a larger V_{min}/V_T ratio. An experimental investigation about the effects of particle sizes and concentration on pressure gradient or drag was conducted by Chung et al. (2001)[14]. They found that contrary to the conventional perception, 80-100 mesh (0.18-0.15 mm) sands had smaller drag and pressure gradient at enough high Reynolds numbers, as compared to the larger 8-10 (2.36-2.00 mm) and 30-40 (0.6-0.425 mm) mesh sands.

In this study, the flow around a sphere placed at various heights below the bottom of collecting tube was investigated by using DES based on the SST model. Meanwhile, the test system has been constructed in the laboratory and a series of 285 tests with various working parameter combinations were conducted. The vertical force of the sphere was measured and compared with the simulation results. The drag coefficients of the sphere with various working parameters were calculated based on test results and defined as approximate equation using dimensionless quantities. A type of collecting tube with spiral deflectors installed inside was designed, with which spiral flow collecting tests were conducted and the values of vertical suction force were measured and compared to that of ordinary flow collecting.

Numerical Method

In hydraulic collecting, polymetallic nodules will be separated from the sea floor and sucked into the collector when fluidic force overweighs their own gravity and viscous force of sea mud. The schematic diagram is shown in Fig. 1. Unlike the moving equipment in the sea, the collecting tube in our experiments is still.

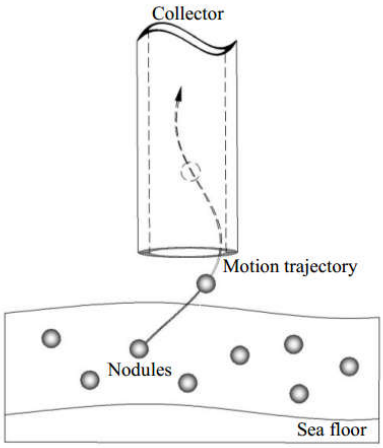
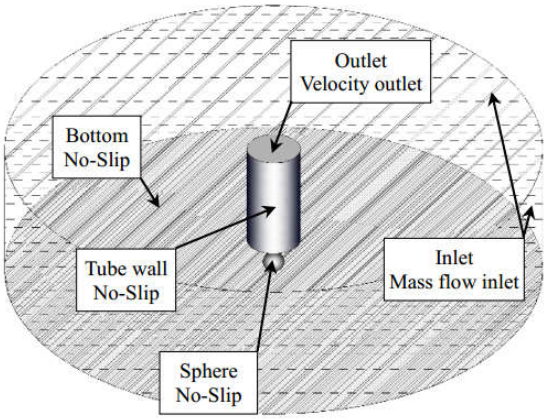


Figure 1. Schematic diagram of hydraulic collecting methods

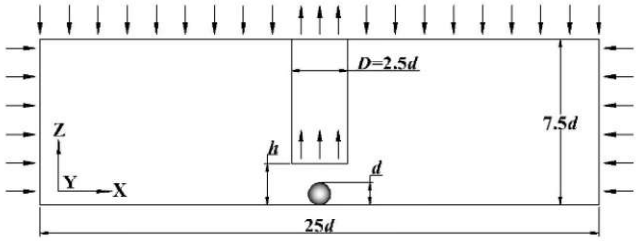
Computational Domain and Boundary Conditions

According to the physical environment of hydraulic collecting, the computational domain and boundary conditions can be set as shown in Fig. 2 (a)~(b). The collecting height h is the distance between the bottom of the collecting tube and the bottom of the sphere. To simplify the simulation, the sphere is fixed on the bottom of the domain. For flow field analysis, the diameter of the sphere is chosen as d=40mm, which is the typical long-axis length of manganese nodules, and the diameter of the collecting. As for the computational domain, the diameter is 25d, and the height is 7.5d. The impact of the boundary condition on the flow field

below the collecting tube can be neglected because the size of the computational domain is sufficiently large.



(a) Computational domain and boundary conditions for analysis



(b) Cutaway view of the computational domain

Figure 2. Computational domain and boundary conditions

Equations of SST Model

The SST-DES is a hybrid RANS/LES model which employs Reynolds-Averaged Navies-Stokes (RANS) in the regions near boundary layers and Large-Eddy Simulation (LES) in the separated region. And the closure model in DES is based on a modification to the Spalart-Allmarars(S-A) one-equation model[15]. When the flow field is simulated by DES, better simulation accuracy can be obtained with less mesh number.

The two equations of Shear-Stress Transport-model (SST model) can be presented as:

$$\frac{\partial(\rho k)}{\partial t}+u_i\frac{\partial(\rho k)}{\partial x_i}=P_k-\frac{\rho k^{\frac{3}{2}}}{l_{k-\omega}}+\frac{\partial}{\partial x_i}\left[\left(\mu_l+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right] \tag{1}$$

$$\frac{\partial(\rho\omega)}{\partial t}+u_i\frac{\partial(\rho\omega)}{\partial x_i}=C_\omega P_\omega-\beta_\omega\rho\omega^2+\frac{\partial}{\partial x_i}\left[\left(\mu_l+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_i}\right]$$
$$+2\rho(1-F_1)\frac{1}{\sigma_{\omega_2}}\frac{1}{\omega}\frac{\partial k}{\partial x_i}\frac{\partial\omega}{\partial x_i} \tag{2}$$

and the eddy viscosity is given by

$$\mu_t = \min\left[\frac{\rho k}{\omega}, \frac{a_1 \rho k}{\Omega F_2}\right] \tag{3}$$

where P_k and P_ω are induced by turbulence and are constant in model equations according to the results in Menter (1993)[16].

In the dissipation term of k equation of SST model, the parameter of turbulence scale $l_{k-\omega}$ can be obtained by

$$l_{k-\omega} = k^{1/2} / \beta_k \omega \tag{4}$$

The parameter $l_{k-\omega}$ is replaced by $\min(l_{k-\omega}, C_{DES}\Delta)$ in DES method, where $\Delta = \max(\Delta x, \Delta y, \Delta z)$ is the longest side length of the mesh and C_{DES} is equal to 0.65. In the regions near boundary layers, where $l_{k-\omega}$ is less than or equal to Δ, the DES model can be viewed as SST model. In the separated region, where $l_{k-\omega}$ becomes greater than C_{DES}, the DES model can be considered as LES model.

Experimental Program

Experimental Setup

In order to provide some corroborating evidence for the simulation results, the experimental system is set up to measure the vertical force. Fig. 3 and Fig. 4 show the experimental setup. The water tank used for the experiments was 2.5m long, 1.5m wide and 1.0m deep. Fresh water was sucked from the tank by a pump to simulate the fluid field near the collecting tube. And a frequency transformer was used to give exact adjustments to the rate of the flow through the pump. Precise control of the distance between the bottom of the collecting tube and the center of the sphere was performed by using an ER50-C10 6-dof industrial robot.

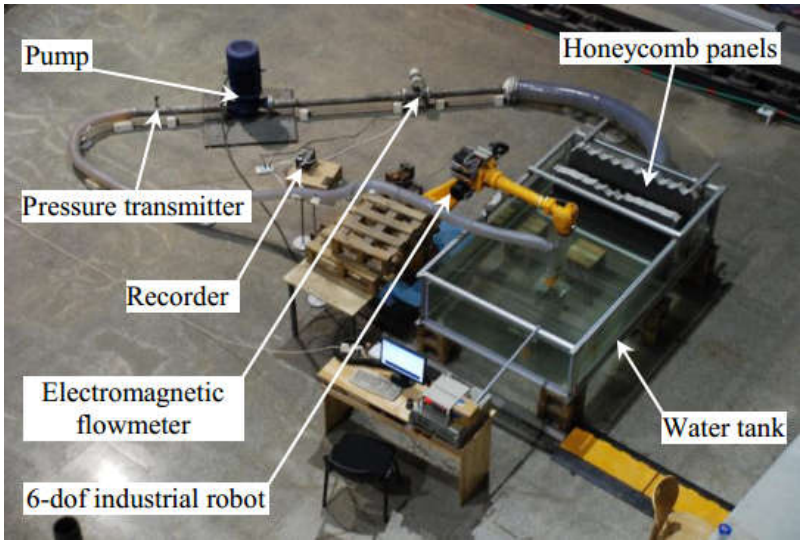


Figure 3. Overall experimental setup

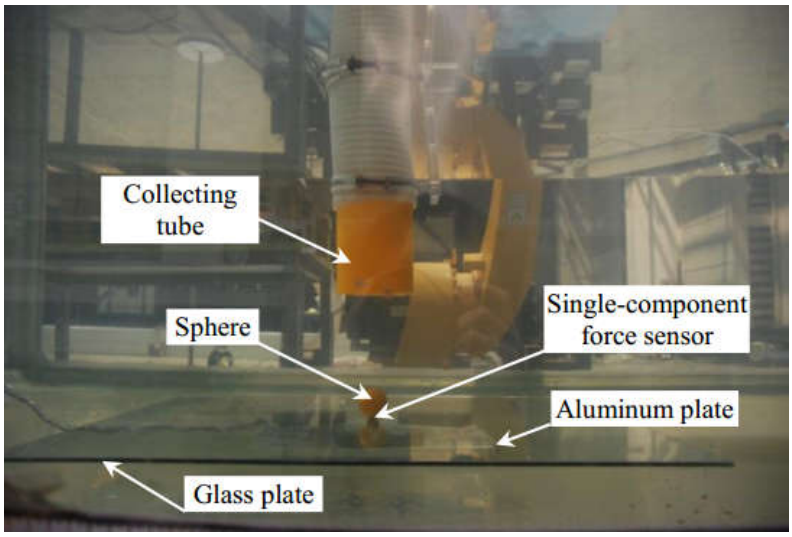


Figure 4. Details of the water tank

The creep and measurement error could be contained within 2‰ by using a LTR-S-5NSA16 single-component force sensor, which meets the requirement of the test. The sensor was fixed on an aluminum plate at the bottom of the water tank. In order to prevent the influence of variation of flow field on the accuracy of sensor measurement, a piece of square glass with the side length of 0.3m and a 7 mm-diameter hole in the center was placed between the sensor and the sphere. Therefore, the sensor was separated from the dynamic fluid flow.

Test Cases

Collecting tubes with different diameters (D=75mm, 100mm, 125mm) were tested under conditions of different collecting heights (varied from 0.050m to 0.080m). In each condition, several sphere sizes were chosen (d=32mm, 36mm, 40mm) and the vertical force exerted on the sphere was measured respectively under different flow velocities of 1.0m/s, 1.2m/s, 1.4m/s, 1.6m/s, 1.8m/s and 2.0m/s. These velocities were chosen to include conditions that hydraulic force on the sphere is smaller or bigger than F_0 (F_0 is equal to the sphere's gravity minus its buoyancy). In order to calculate the critical vertical start force on nodules of identical size, the densities of water and sphere were selected as 1,000 kg/m^3 and 2,100 kg/m^3 (the typical density of manganese nodules on the seabed) respectively. For example, F_0 is 0.3616 N when d=40mm.

For the collecting velocity from 1.0=m/s to 2.0m/s, the Reynolds number based on d (diameter of the sphere) is from around 30,000 to 60,000 when d=40mm. The Re range in actual flow of engineering operation involves some specific collecting conditions such as collecting heights, structures of the collector and local flow velocity of the nodules. There is some possible overlap between our Re range and the actual Re range.

Results and Discussion

Vertical Force

The time-history curves and mean values of vertical force at collecting velocity of 1m/s and 2m/s are shown in Fig. 5. The mean values of numerical results match well with those of experimental measurements.

The oscillation amplitudes of experimental measurements are much larger than those of numerical results, probably because there is some interference in the experiment process such as fluctuations of the free surface, flux changes in the pump and vibration noise.

Due to the periodicity and symmetry of the interference, errors can be offset and accurate mean values can be obtained if we measure for a long time and get enough data. The transverse force also makes a difference on the measurements, but the magnitude of the transverse force is relatively small compared to the vertical force. Therefore, its influence on the transducer can be ignored.

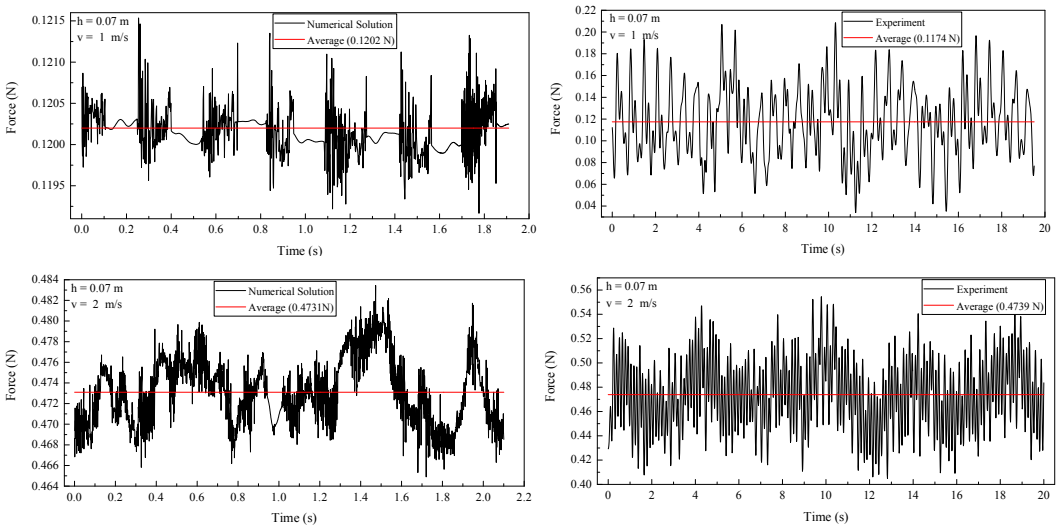


Figure 5. Time-history curves and mean values of vertical force of numerical results and experimental measurements at collecting velocity of 1m/s and 2m/s

(*D*=100mm, *d*=40mm, *h*=0.07m)

Fig. 6 shows the mean values of force on the sphere at different collecting heights and collecting velocities. The overall qualitative behaviors of numerical simulated results and experimental results are in good agreement with each other. Among all the cases, the max difference between test result and CFD result is 5.85%. The comparison validates the feasibility and high accuracy of forecasting the characteristics of the force on an ore particle collected on the seafloor by numerical method. Meanwhile, accurate vertical force measurement can be obtained by the experimental system, which provides a feasible solution for experiments of the special flow around a sphere.

As the collecting height gets smaller, results of the experiments are less than those of the numerical solutions. The arrangement of the sensor may be the cause of the phenomenon. The flow velocity below the sphere would increase when the collecting tube gets closer to the sphere. This would enhance the impact of the sensor placed under the sphere on the experimental measurements. The arrangement of the sensor does not only slightly reduce the area of force of the sphere, but also have a little interference on the fluid field at the bottom of the sphere. This explains the differences between two series of results. The interference of the arrangement of the sensor can be decreased by placing the sensor under the glass plate and extending its tip through a small hole at the center of the plate. In this way, the error could be decreased as much as possible.

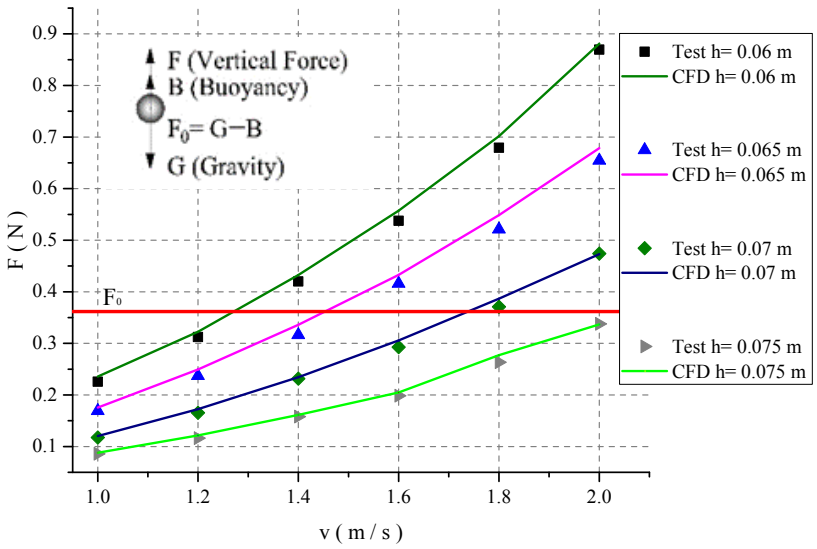


Figure 6. Variation of the mean values of vertical force on the sphere under different collecting heights and collecting velocities (*D*=100mm, *d*=40mm)

Flow Field

The velocity distribution and pressure distribution for collecting height 0.07 m and collecting velocity 2.0m/s are shown in Figs. 7~9. In the collecting flow field, the velocities are relatively larger and the streamlines are more chaotic above the sphere than those velocities and streamlines in other regions, which leads to the forming of turbulent areas in the wake vortex and low pressure areas upon the sphere. Vertical force is induced with the pressure differences along the sphere.

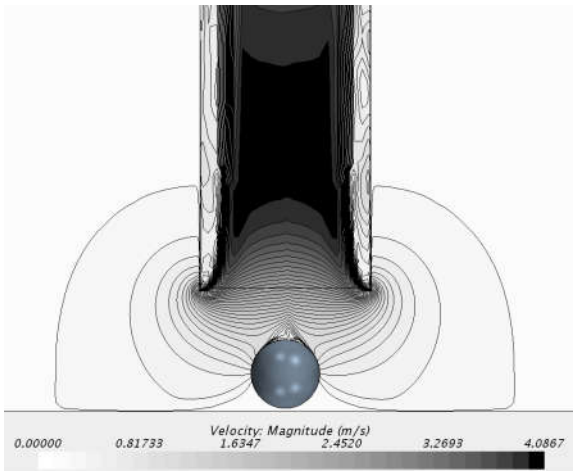


Figure 7. Velocity distribution (*D*=100mm, *d*=40mm, *h*=0.07m, *v*=2.0m/s)

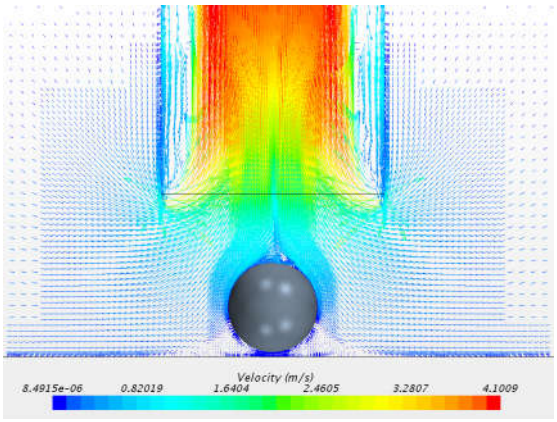


Figure 8. Velocity vector (*D*=100mm, *d*=40mm, *h*=0.07m, *v*=2.0m/s)

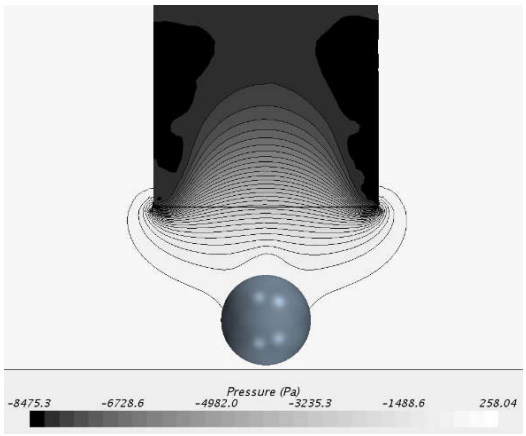


Figure 9. Pressure distribution (*D*=100mm, *d*=40mm, *h*=0.07m, *v*=2.0m/s)

An insight into understanding the characteristics of spherical ore particles is provided in Fig. 10 by the comprehensive analysis on the dynamic change regularity of wake vortex structures and the pressure distribution in one circulation. The repeated growth and burst of the wake vortex is the major reason of the amplitude oscillation of the vertical force.

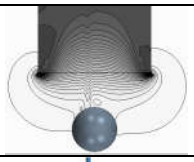
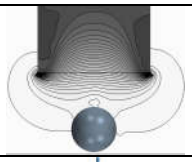
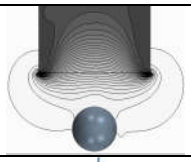
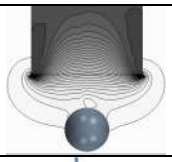
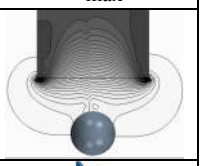
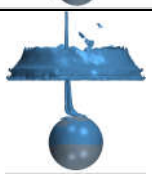
	F_{max}	$F_{max} \rightarrow F_{min}$	F_{min}	$F_{min} \rightarrow F_{max}$	F_{max}
Pressure distribution					
Vortex structure					

Figure 10. Development of the pressure distribution and vortex structures

Dimensionless Analysis

Dimensional analysis is applied to find the correlation between the force on the sphere and other experiment data. It is helpful to apply dimensional analysis to analyze the correlation among different parameters in experiment and thus to indicate the cause and effect among them. We can also find out the significance of different parameters with the help of

dimensional analysis so as to make it clear what factor plays a key role in the problem. Therefore, it is conducive to deep research of the problem.

In order to discuss the drag coefficient at high Reynolds number, dimensionless quantity C_d is introduced. The physical quantities that determine the drag force are ρ, v, d, μ, D, and h. The dimensionless quantity is obtained by combining these quantities, and C_d is expressed as:

$$C_d = \frac{F}{(1/2)\rho v^2 A} = f\left(\frac{\rho v d}{\mu}, \frac{h}{d}, \frac{D}{d}\right) \tag{5}$$

In addition, we set kinematic coefficient of viscosity as 1.31×10^{-6} N·S/m^2 in calculations because the water temperature is about 10 degrees Celsius. The relations between drag coefficients and Reynolds numbers, ratios of h/d and D/d were investigated in this paper, based on the results of a series of 285 tests.

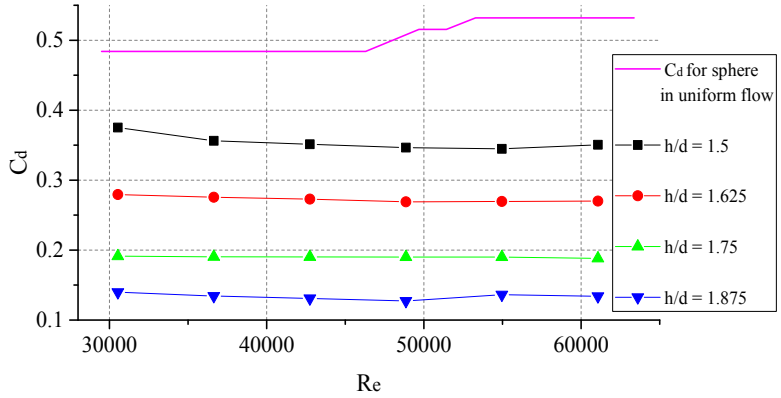
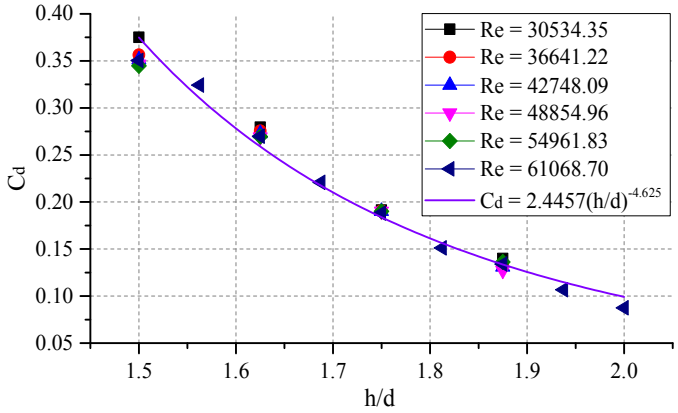
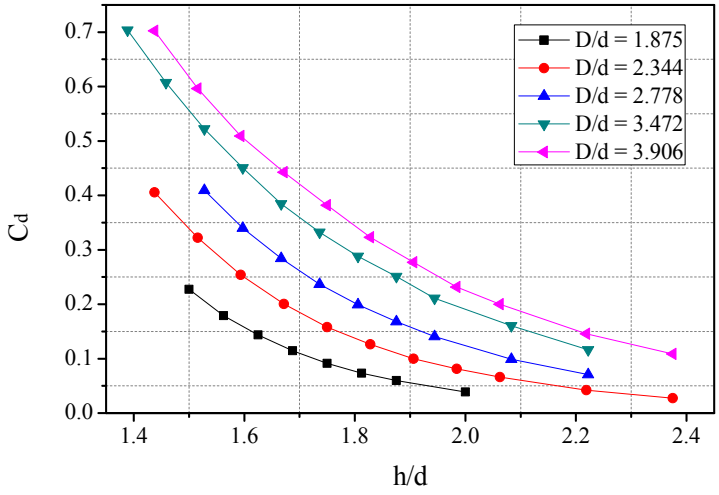


Figure 11. Relations between drag coefficients and Reynolds numbers (*D*=100mm, *d*=40mm)

As Fig. 11 shows, the drag coefficients for sphere in uniform flow are bigger than those coefficients of our numerical results at the same Reynolds numbers. One reason for this phenomenon is that the pressure gradient changes more violently, another cause is that the detached points of the sphere's wake in the numerical cases are father than that point in uniform flow. Fig. 11 also reveals that the drag coefficients decrease at the same Reynolds number as ratios of h to d grow.



(a) C_d – Re, h/d



(b) C_d - h/d, D/d

Figure 12. Relations between drag coefficients and Re, *h*/*d* and *D*/*d*

As Fig. 12(a) demonstrates, C_d decreases with the increase of h/d in the range of 1.5 to 2.0 but not significantly varies with Re numbers. Fig. 12(b) shows that C_d increases with the increase of D/d when h/d is constant.

The relationship is combined to obtain the following:

$$C_d = 2.8525\left(h/d\right)^{3.3559\left[\ln\left(D/d\right)-2.4084\right]} \tag{6}$$

However, there is no significant relationship between C_d and Reynolds numbers.

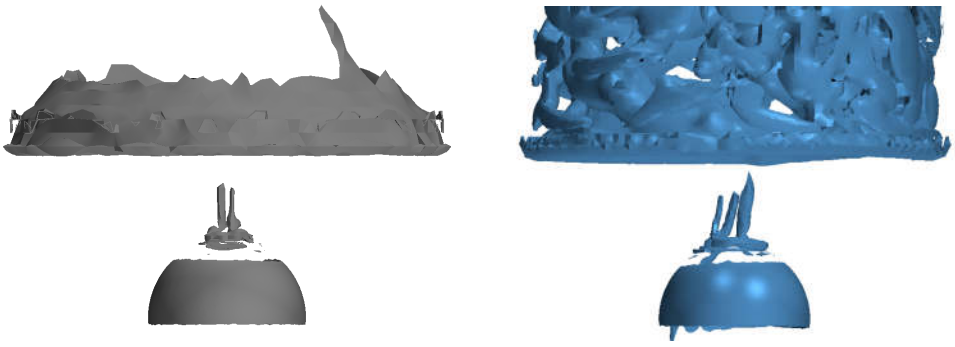
Comparison of Vertical Force in Ordinary and Spiral Flow

A type of collecting tube with spiral deflectors installed inside is also designed, which is able to induce spiral flow inside and under the tube. The numerical simulation results of vertical force in ordinary and spiral flow are shown as Table 1 when diameter of the collecting tube D=100mm, tube height above the sphere h=60mm, and diameter of the sphere d=36mm.

Table 1. Vertical force in ordinary flow and spiral flow (*D*=100mm, *h*=60mm, *d*=36mm)

Flow velocity v (m/s)	Vertical force F/N	
	Ordinary flow	Spiral flow
1.0	0.1385	0.1543
1.2	0.2053	0.2223
1.4	0.2834	0.3070
1.6	0.3803	0.4142
1.8	0.4753	0.5228
2.0	0.5849	0.6349

It can be concluded that spiral flow exerts stronger suction force on the sphere as the value of vertical force in the collecting tube with spiral deflectors is much greater. One reason for that is the spiral flow brings about an increase of the scale of vortex structure as is shown in Fig.13.



(a) ordinary flow (b) spiral flow

Figure 13. Comparison of the scale of vortex structure

(*D*=100mm, *h*=60mm, *d*=36mm, *v*=1.0m/s)

As to recognizing the vortex structure, Q criteria was applied, where Q is defined as

$$Q = -\frac{1}{2}\left(\|\mathbf{S}\|^2 - \|\mathbf{\Omega}\|^2\right)$$

S , **Ω** represent strain tensor and rotation tensor respectively. In this case Q was chosen as equal to 5000. Fig. 13 indicates that under that condition the scale of wake vortex structure in spiral flow is significantly larger than that in ordinary flow. It means that the scale of low pressure area in the wake flow of the sphere is greater. Therefore, the differential pressure is much larger and the suction force induced by spiral flow is supposed to be much stronger compared to the ordinary flow.

In order to validate the results, model tests with the ordinary collecting tube and the tube with spiral deflectors were conducted and the suction forces were measured respectively. The relations between vertical suction force *F* and tube height *h* and flow velocity *v* in ordinary and spiral flow collecting are shown as Fig. 14, based on test results.

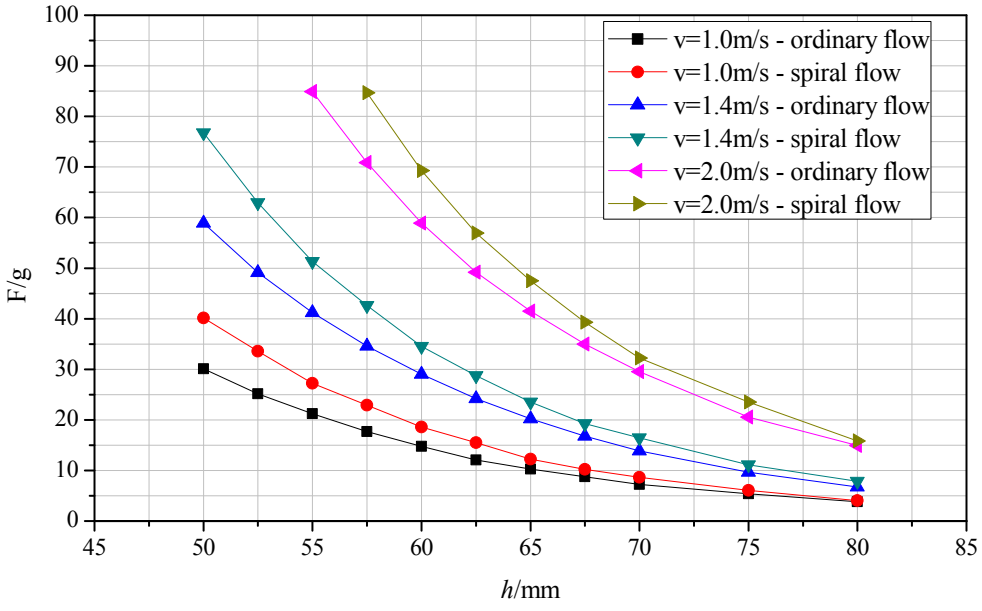


Figure 14. Relations between suction force and tube height and flow velocity in ordinary and spiral flow collecting (*D*=100mm, *d*=36mm)

From Fig. 14 it can be indicated that F increases with the decrease of h, more significantly when h is smaller. Also, it suggests that the suction force of spiral flow is larger than that of ordinary flow at equal flow velocity or flow rate. The ratio reaches about 1.3 at h=50mm according to experiment results.

Conclusions

The flow around a sphere was simulated with detached-eddy simulation (DES) method based on shear stress transport (SST) model, and the wake flow behind the spheres was investigated. Meanwhile, the test system for hydraulic collecting, which is corresponding to the model of CFD, has been set up in the laboratory. A series of 285 tests with various working parameter combinations were conducted. The vertical force of spherical particles is measured by the force transducer. Tests of spiral flow collecting were also conducted and the vertical force results were compared to that of ordinary flow. The following are the main conclusions from this work:

1. The vertical force prediction for single ore particle in collecting condition based on numerical simulation with DES method based on SST model is feasible and sufficiently precise. Also the test system layout is reasonable to be able to catch vertical force with high precision. Therefore, it is a promising method to design the new collector by numerical simulation and test more complex collector by the test system.

2. When h/d is in the range of 1.5 to 2.0 and D/d in the range of 1.875 to 3.906, C_d can be concluded as following:

$$C_d = 2.8525\left(h/d\right)^{3.3559\left[\ln\left(D/d\right)-2.4084\right]}$$

And the drag coefficients for sphere in uniform flow are bigger than those coefficients of our numerical results at the same Reynolds numbers due to different changes of the pressure gradient and the delay of the detached point of the sphere's wake.

3. The pressure distribution pattern has a significant influence on the vertical force of the sphere. The variation of the wake vortex is the dominant factor of force vibration.

4. The value of vertical suction force of spiral flow is greater than that of ordinary flow at equal flow velocity. Spiral flow brings about an increase of the scale of wake vortex. Therefore, the scale of low pressure area in the wake flow of the sphere is much larger, which forms larger differential pressure and induces stronger suction force on the sphere compared to the ordinary flow.

The results of this study can provide reference and support for mechanism study and optimal design of the hydraulic collecting system for deep-ocean mining. In future research, more working parameter combinations are expected to be designed so as to improve and verify the non-dimensional equation. And the behaviors of the flow and vortices will be observed during flow visualization.

References

[1] Liu, S., Liu, C., and Dai, Y., (2014). "Status and Progress on Researches and Developments of Deep Ocean Mining Equipments," Journal of Mechanical Engineering, Vol 50, No 2, pp 8-18.
[2] Herrouin, G, Lenoble, JP, Charles, C, Mauviel, F, Bernard, J, .and Taine, B (1989). "A Manganese Nodule

Industrial Venture Would Be Profitable: Summary' of a 4-Year Study in France", Proceedings of Offshore Technology Conference.
[3] Chung, J. S. (1985). "Advances in manganese nodule mining technology." Marine Technology Society Journal, Vol 19, No 4, pp 39-44.
[4] Zhao, S., and Liu, F., (1995) "Study of Deep-sea Mining Technique In Germany." Metal mine, No 6, pp 14-17.
[5] Hong, S., Choi, J. S., Kim, J. H., & Yang, C. K. (1999). Experimental study on hydraulic performance of hybrid pick-up device of manganese nodule collector. In The Third ISOPE Ocean Mining Symposium. International Society of Offshore and Polar Engineers. Goa, India, pp 69-77.
[6] Yang, N., & Tang, H. (2003). "Several considerations of the design of the hydraulic pick-up device." In The Fifth ISOPE Ocean Mining Symposium. International Society of Offshore and Polar Engineers. Tsukuba, Japan, pp 119-122.
[7] Chen, X., and Jian, Q., (1996). "An Experimental Study On Hybrid Collection Method For Deep-sea Nodule Mining." Mine R & D, Vol 16, No 4, pp 1-4.
[8] Lim, S. J., Kim, J. W., Jung, S. T., Cho, H. Y., and Lee, S. H. (2015). "Deep Seawater flow Characteristics Around the Manganese Nodule Collecting Device." Procedia Engineering, Vol 116, pp 544-551.
[9] Achenbach, E. (1972). "Experiments on the flow past spheres at very high Reynolds numbers." Fluid Mech, Vol 54, No 3, pp 565-575.
[10] Tsutsui, T. (2008). "Flow around a sphere in a plane turbulent boundary layer." Journal of Wind Engineering and Industrial Aerodynamics, Vol 96, No 6, pp 779-792.
[11] Johnson, T. A., and Patel, V. C. (1999). "Flow past a sphere up to a Reynolds number of 300." Journal of Fluid Mechanics, Vol 378, pp 19-70.
[12] Constantinescu, G. S., and Squires, K. D. (2003). "LES and DES investigations of turbulent flow over a sphere at Re= 10,000." Flow, Turbulence and Combustion, Vol 70, No 1-4, pp 267-298.
[13] Chung, J. S., Yarim, G., & Savasci, H. (1998). "Shape effect of solids on pressure drop in a 2-phase vertically upward transport: silica sands and spherical beads." In The Eighth International Offshore and Polar Engineering Conference. International Society of Offshore and Polar Engineers. Montreal, Canada, Vol 1, pp 58-65.
[14] Chung, J. S., Lee, K., Tischler, A., & Yarim, G. (2001). "Effect of particle size and concentration on pressure gradient in two-phase vertically upward transport." In The Fourth ISOPE Ocean Mining Symposium. International Society of Offshore and Polar Engineers. Szczecin, Poland, pp 132-138.
[15] Spalart, P. R., Jou, W. H., Strelets, M., & Allmaras, S. R. (1997). "Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach." Advances in DNS/LES, Vol 1, pp 4-8.
[16] Menter, F. R. (1993). "Zonal two equation k-turbulence models for aerodynamic flows." AIAA paper, pp 2906.
[17] Hunt, J. C., Wray, A. A., and Moin, P. (1988). "Eddies, streams, and convergence zones in turbulent flows." Studying Turbulence Using Numerical Simulation Databases, pp 193-208.
[18] Kida, S., Goto, S., and Makihara, T. (2002). "Life, structure, and dynamical role of vortical motion in turbulence." National Institute for Fusion Science, Theory and Computer Simulation Center, pp 1-2
[19] Ting, L., and Klein, R. (1991). "Viscous vortical flows." Viscous Vortical Flows, Vol 374.

Shape optimum design of shear panel damper made of low yield steel

†*Y. Liu[1], Y. Nishio[2], M. Shimoda [3]and M. Shimazu [4]

[1] Department of Mechanical Engineering, Sojo University, Japan.
[2] Graduate School of Mechanical Engineering, Sojo University, Japan.
[3] Department of Advanced Science and Technology, Toyota Technological Institute, Japan
[4]Department of Architecture, Sojo University, Japan.

*Presenting author: liuyang60212@gmail.com
†Corresponding author: liu@mec.sojo-u.ac.jp

Abstract

In this paper, a shape optimization approach is presented for a shear panel damper made of low yield steel to improve the deformation ability. A minimization problem of maximum cumulative equivalent plastic strain, an index of the deformation ability of the shear panel damper is formulated subject to a constraint of total absorbing energy. The response surface methodology as well as the design of experiment technique is applied to the optimization process. In this study, finite element analysis with isotropic/kinematic hardening model is adopted to simulate the cyclic elasto-plastic behavior instead of experimental approach, and the numerical solutions are validated by comparing with previous experimental results. With the numerical analysis, the shape parameters effects are investigated and second order polynomials are fitted to obtain the regression equations for the maximum cumulative equivalent plastic strain and the total absorbing energy. The final optimal shape is determined by using the established regression equations. The shape optimization approach can substantially improve the deformation capacity of the shear panel damper.

Keywords: Shear panel damper, Shape optimization, Deformation ability, Energy absorption, Response surface method

Introduction

Shear panel damper (SPD) made of low yield steel has high level of passive energy dissipation capacity as a consequence of inelastic deformation of the low yield steel, and has been received considerable interest in the last two decades. When installed into building and bridge structures, it is expected to partially divert the input seismic energy into the SPDs and effectively reduce the seismic responses of the structures under strong earthquake loads. To be a type of hysteretic damper, properly design of the SPD (stiffened or unstiffened) is strongly required to sustain high deformation capacity and repetition durability for low cycle fatigue under cyclic seismic loading, especially for the application in bridge structures, which demand large range of shear deformation. If the SPD is designed unreasonably, the clacks should initiate at edges or corners of the shear panel in the early stage of cyclic loading due to the stress concentration, and grow along with cycles, that will decrease the energy absorption capacity drastically. In focusing on improving the deformation capacity and repetition durability, recently, some experimental and analytical researches have been carried out by varying the panel shape or installing the stiffeners on the left and right sides of the SPD for the application in bridge structures [1][2][3]. However, most of researches are confined to be empirical methods or analytical researches dealing with direct problems, the shape optimization of obtaining the maximization of deformation capacity and total absorbing energy has not been studied.

The studies on optimization of elastic and elastoplastic structures have been extensively investigated during the past 30 years, and a number of useful algorithms and methodologies are developed (e.g. [4][5]). As a practical and effective optimum design methodology for nonlinear problem, the response surface methodology (RSM) combined with the design of experiment (DOE) technique is currently applied to nonlinear design optimization problems, such as optimization problems of crushing energy absorbing of the automobile body and box-type column structures [6]. In this paper, a shape optimum design of the SPD was studied by the response surface approximation methodology and the technique of design-of-experiment. Since deformation capacity of the SPD can be evaluated by the maximum cumulative

equivalent plastic strain at a cyclic shear deformation. Instead of experimental approach, the cyclic behavior of SPD subjected to cyclic loading is studied by sophisticated finite element method (FEM) with isotropic/kinematic hardening model, and a comparison between numerical simulation and experimental result was made and precision of the numerical simulation was confirmed. Then an orthogonal array is employed to arrange the design point using the technique of design of experiment, and numerical simulations of the SPDs with various shape parameters were carried out. Based on the numerical results, the influences of the shape parameters to responses of the maximum cumulative equivalent plastic strain and the energy absorbing behaviors are investigated to obtain the regression equations of the two objective functions. Finally, the response surface methodology was adopted to solve the shape optimization problem, and to obtain the optimal shape parameters of the SPD.
In our previous study, free edges on the left and right sides of the SPD were optimized to improve the deformation ability. The geometry of the reduced left and right edges was represented by cubic Bezier curves as a linear combination of Bezier polynomial [7]. In this study, a shape optimum design is carried out to determine the optimal shape parameters of holes in the panel when all of the free edges are invariable.

Numerical Analysis

The initial shape of SPD, which is a 156×156 mm square plate with uniform plate thickness of t_w=12 mm, is shown in Figure 1. The upper and lower edges of the panel are groove welded to plates. Cyclic lateral load was applied at the upper plate, and the loading history is shown in Figure 2, in which the increments of the shear displacement in each loading cycle are $\pm 1\delta_y$, where δ_y=5 mm is the shear yield displacement corresponding to the 0.2% offset yield stress of the material.

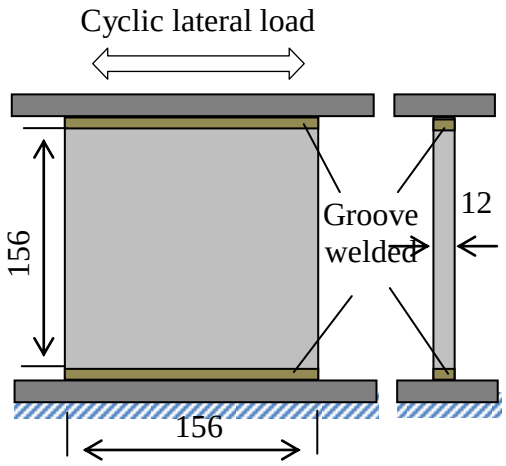


Figure 1. Initial shape and boundary conditions

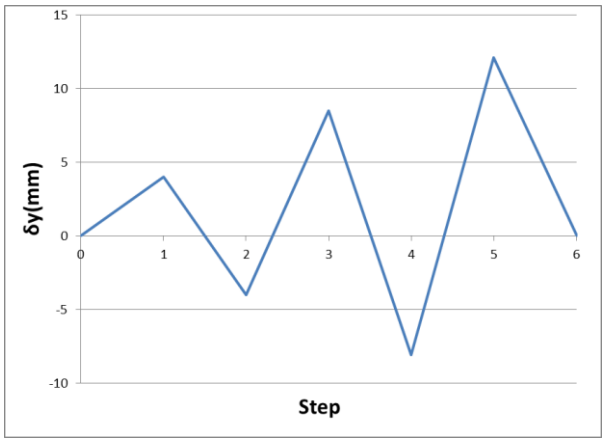


Figure 2. Loading history

The material properties of low-yield 100(LY100) steel were measured by tensile coupon test and the obtained stress-strain curves are shown in Figure 3. The yield strength defined 0.2% offset value of LY100 is 80.1 N/mm^2 and elongation reaches 60%, which is about three times of SS400 mild steel.
The cyclic elastoplastic behavior of SPD subjected to cyclic loading is simulated by ABAQUS with a combined isotropic/kinematic hardening model [7][8], which was employed as constitutive law to describe the material cyclic behavior accurately. The combined hardening model consists of two components: a nonlinear kinematic hardening component and an isotropic hardening component.

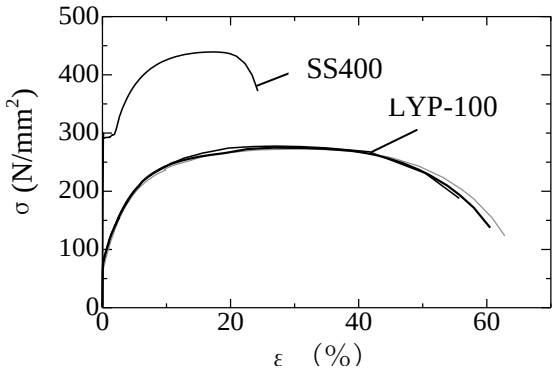
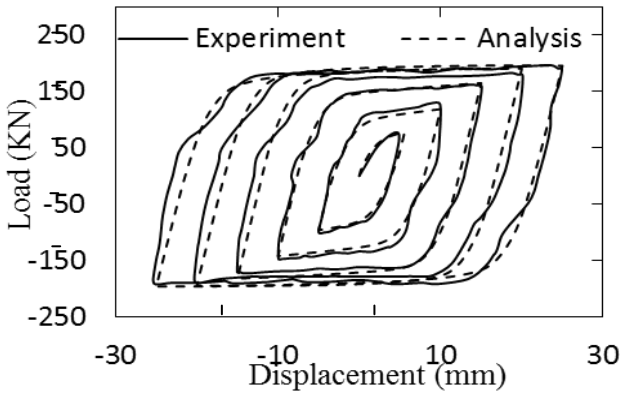


Figure 3. Stress-strain curves for SS400 and LY100 tension coupons

Figure 4 shows typical hysteretic curve of the shear load versus displacement of the initial shape of SPD compared with corresponding experimental curve. As shown in Figure 4, hysteretic curve obtained from the analysis agree generally well with those from the experiment. It is clarified that the present analysis with the combined isotropic/kinematic hardening model can accurately predict the cyclic elastoplastic behavior of the LY100 SPD. Figure 5 shows the accumulated equivalent plastic strain distribution by FEM simulation in the 4th cycle loading, and remarkable strain concentration at the panel corners can be observed.



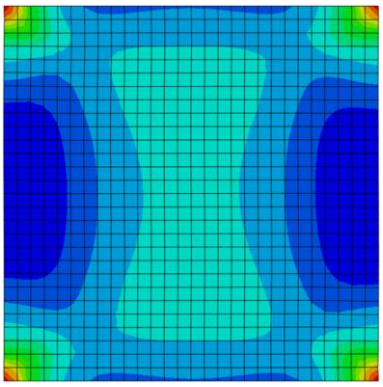


Figure 4. Hysteretic load-displacement curves Figure.5 *CEPS* distribution of the initial SPD

DESIGN OPTIMIZATION

In this study, with the aim of minimizing the maximum cumulative equivalent plastic strain ($CEPS_{max}$) subjected to a constraint of total absorbing energy (E), the shape parameters r_1, r_2, θ of holes in the panel (as shown in Figure. 6) are taken as design variables. An orthogonal array in the design of experiment is employed to assign analysis points in simulating the cyclic elastoplastic behavior of SPD. Based on the numerical results of the cyclic elastoplastic analysis, the response surface approximation technique is applied to generate the regression equations of the $CEPS_{max}$ and the E in terms of the design variables that are evaluated to be significant at high levels for the response by means of analysis of variance. Then, the optimization problem is formulated as:

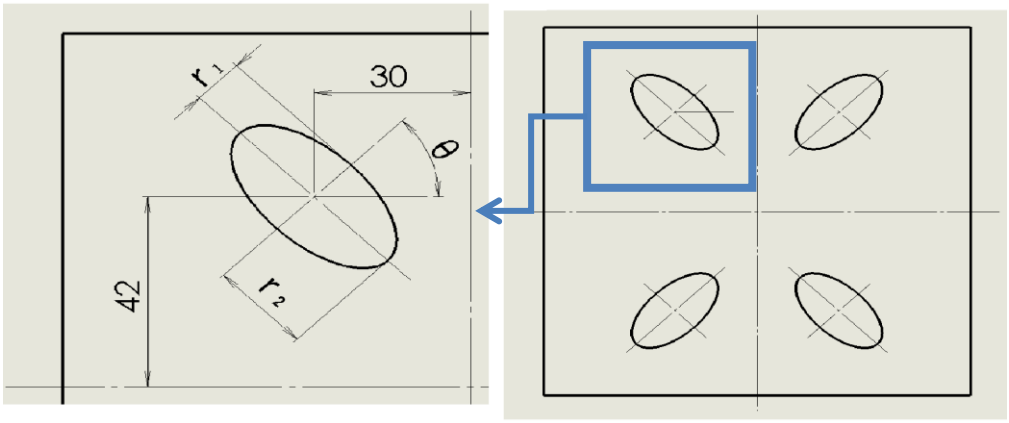


Figure 6. Shape parameterization

$$\begin{aligned} &\text{Design variables : } r_1,\ r_2,\ \theta \\ &\text{Objective function : Minimizing } CEPS_{\max} \\ &\text{Constraint condititon : } E \geqq k \cdot E_{\text{in}} \end{aligned} \tag{1}$$

where $CEPS_{\max}$ indicates the maximum cumulative equivalent plastic strain, and E_{in}, E indicate the total absorbing energy of the initial shape and the optimized shape respectively. k indicates the coefficient of the allowable lower bound and is taken as 80 % and 50% in this study.

Optimization process of the first iteration

The experimental design levels of the process variables in the first iteration are shown in Table 2. As shown in Table 2, an orthogonal array L_{27} is employed to arrange the design point, and results of $CEPS_{\max}$ and E are obtained by the cyclic elastoplastic analysis at each design point under the same cyclic lateral load shown in Figure 2.

Based on the numerical solutions of the cyclic elasto-plastic analysis in the first iteration as shown in Table 2, the response surface regression procedure was employed to fit the polynomial Equation to the numerical analysis results, and the maximum cumulative equivalent plastic strain ($CEPS_{\max}$), the total absorbing energy E are approximated in the form of orthogonal polynomials as:

$$\begin{aligned} CEPS_{\max}(r_1, r_2, \theta) = {} & 0.750 - 1.641\times10^{-2} r_1 + 3.956\times10^{-4} r_1^2 + 3.955\times10^{-2} r_2 - 9.811\times10^{-4} r_2^2 \\ & - 6.370\times10^{-4}\theta + 4.664\times10^{-5}\theta^2 - 6.41\times10^{-3} r_1 r_2 + 1.561\times10^{-4} r_1 r_2^2 \\ & + 2.431\times10^{-4} r_1^2 r_2 - 7.227\times10^{-6} r_1^2 r_2^2 + 1.996\times10^{-3} r_1\theta - 2.309\times10^{-5} r_1\theta^2 \\ & - 6.811\times10^{-5} r_1^2\theta + 7.193\times10^{-7} r_1^2\theta^2 - 1.934\times10^{-3} r_2\theta + 1.680\times10^{-5} r_2\theta^2 \\ & + 5.989\times10^{-5} r_2^2\theta - 4.419\times10^{-7} r_2^2\theta^2 \end{aligned} \tag{2}$$

$$\begin{aligned} E(r_1, r_2, \theta) = {} & 6418.812 - 66.951 r_1 - 0.932 r_1^2 - 155.879 r_2 + 3.656 r_2^2 \\ & + 37.193\theta - 0.429\theta^2 + 13.020 r_1 r_2 - 0.631 r_1 r_2^2 \\ & - 0.481 r_1^2 r_2 + 2.094\times10^{-2} r_1^2 r_2^2 - 4.410 r_1\theta + 4.466\times10^{-2} r_1\theta^2 \\ & + 0.121 r_1^2\theta - 1.086\times10^{-3} r_1^2\theta^2 - 0.636 r_2\theta + 1.350\times10^{-2} r_2\theta^2 \\ & + 4.486\times10^{-2} r_2^2\theta - 8.311\times10^{-4} r_2^2\theta^2 \end{aligned} \tag{3}$$

Table 2. Design levels numerical solutions for SPDs

Design Point	r_1 (mm)	r_2 (mm)	θ(°)	$CEPS_{max}$	E (KJ)
1	10	10	15	0.6017	4.91
2	10	10	45	0.6017	4.91
3	10	10	75	0.6017	4.91
4	10	15	15	0.5615	4.50
5	10	15	45	0.5478	4.52
6	10	15	75	0.5707	4.44
7	10	20	15	0.5312	4.06
8	10	20	45	0.5408	4.15
9	10	20	75	0.6918	3.85
10	15	10	15	0.5509	4.36
11	15	10	45	0.5388	4.36
12	15	10	75	0.5275	4.41
13	15	15	15	0.4737	3.91
14	15	15	45	0.4737	3.91
15	15	15	75	0.4737	3.91
16	15	20	15	0.4567	3.44
17	15	20	45	0.4587	3.40
18	15	20	75	0.5001	3.29
19	20	10	15	0.5711	3.72
20	20	10	45	0.5266	3.75
21	20	10	75	0.4706	3.89
22	20	15	15	0.4972	3.27
23	20	15	45	0.4133	3.27
24	20	15	75	0.4375	3.37
25	20	20	15	0.4152	2.75
26	20	20	45	0.4152	2.75
27	20	20	75	0.4152	2.75

Table3 Design levels and optimization result (k=0.8)

		1st	2nd	3rd	4th	5th	6th
Design levels	r_1(mm)	10 15 20	5 10 15	7 12 17	7 12 17	7 12 17	10 12 14
	r_2(mm)	10 15 20	8 13 18	5 10 15	5 10 15	6 11 16	8 10 12
	θ (°)	15 45 75	30 45 60	34 44 54	44 54 64	44 54 64	44 54 64
Optimization result	r_1(mm)	10.2	12.3	12.3	11.9	12.4	12.4
	r_2(mm)	13.0	10.3	10.1	10.9	10.1	10.2
	θ (°)	45.0	43.9	54.0	54.0	54.0	54.0
	$AEPS_{max}$	0.5555	0.5398	0.5430	0.5495	0.5367	0.5348
	E (KJ)	4.66	4.63	4.68	4.65	4.66	4.65

To obtain more precise approximated response surface, analysis points are selected for the 1st, 2nd, 3rd, 4th, 5th and 6th iteration of the optimization process, and the optimum results of the 6 iterations are shown in Table 3. The obtained optimal shape of SPD is shown in Figure 7. Figure 8 shows accumulated plastic distribution, which is simulation result of the cyclic elastoplastic behavior of the optimized SPD subjected to cyclic loading. The numerical estimation of $CEPS_{max}$ in the optimized SPD is 0.5348, which decreased by 25% comparing with the initial shape. It is obvious that the optimal shape can substantially increase the

deformation capacity of SPD. Furthermore, when the coefficient of the allowable lower bound k is set to k=0.5, the obtained optimal shape of SPD is shown in Figure 8, and the $CEPS_{max}$ was reduced to 0.3662, which is 48.2 % down than value of the SPD with the initial shape.

Conclusions

In this study, a shape optimum design of shear panel damper made of low yield steel was carried out to determine the optimal shape parameters of holes in the panel. The combination between the response surface method and the design of experiment technique and the cyclic elastoplastic behavior simulation of shear panel was employed as the optimization methodology. It was confirmed that the shape optimization approach can effectively obtain the optimal shape of the shear panel damper. In the optimal shape, the maximum cumulative equivalent plastic strain is decreased significantly, that can substantially improve the deformation capacity of the shear panel damper.

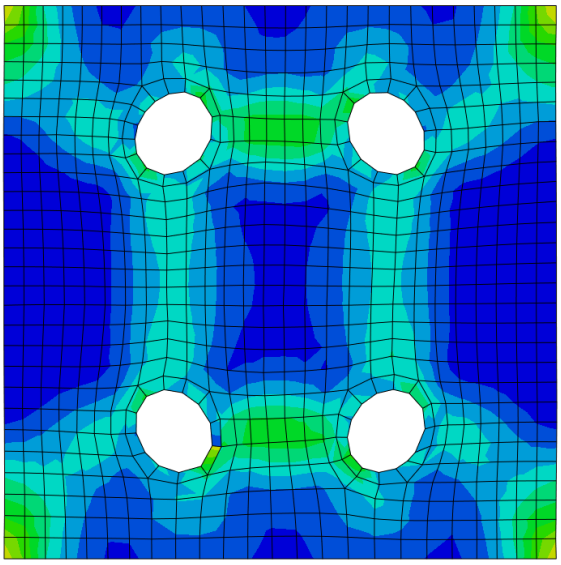


Figure 7. Optimal shape of SPD (k=0.8)

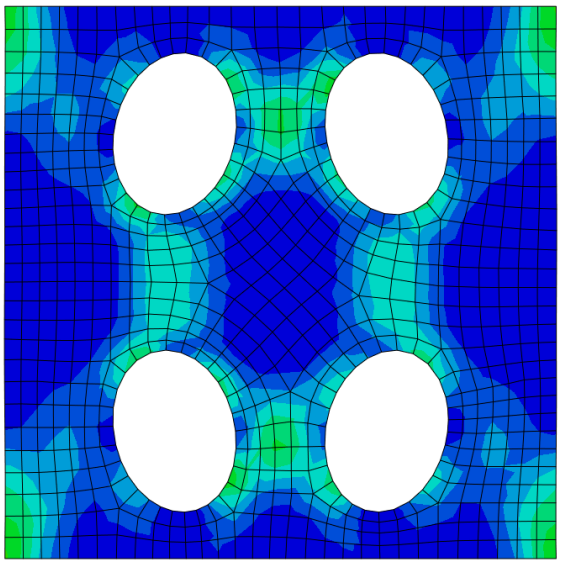


Figure 8. Optimal shape of SPD (k=0.5)

References

[1] Y. Liu, T. Aoki, T. Takaku, Y. Fukumoto (2007) Cyclic loading tests of shear panel damper made of low yield steel, *J. of Structural Eng., JSCE*., 53A, 560-567.

[2] T. Aoki, Y. Liu, T. Takaku, and Y. Fukumoto (2007) Experimental investigation of tapered shear-type seismic devices for bridge bearings, *Pacific Structural Steel Conference 2007, Steel Structures in Natural Hazards, Wairakei, New Zealand*, 1, 111-117.

[3] Y. Koike, T. Yanaka, T. Usami, H.B. Ge, S. Oshita, D. Sagou, and Y. Uno (2008) An Experimental Study on Developing High-Performance Stiffened Shear Panel Dampers, *J. of Structural Engineering*, *JSCE*, 54A, 372-381.

[4] Masatoshi Shimoda, Hideyuki Azegami, Toshiaki Sakurai, "Traction Method Approach to Optimal Shape Design Problems", *Journal of Passenger Cars*, *SAE 1997 Transactions*, pp. 2355-2365, 1998.

[5] H. Ihara, H. Azegami, M. Shimoda (1999) Solution to Shape Optimization Problems Considering Material

Nonlinearity", *Computer Aided Optimization Design of Structures VI*, Chap. 3, Sec. 4, pp. 87-96.

[6] J. Han, K. Yamazaki, (2000) A study on maximization of dynamic crushing energy absorption of square tubes with and without stiffener, *JSME Int, J. Ser. A*, 43 , 138-145.

[7] Y. Liu, M. Shimoda (2013) Shape optimization of shear panel damper for improving the deformation ability under cyclic loading, *Structural and Multidisciplinary Optimization*, 48, 427-435.

[8] ABAQUS, Inc., ABAQUS/Analysis user's manual-version 6.5

[9] R. M. Myers, D. C. Montgomery (2002) Response Surface Methodology, John Wiley & Sons, New York.

Experimental and Numerical Study on the Wake Field of a 76000 DWT Panamax Bulker

*** Tie-cheng Wu[1], †C.Y. Guo[1], W.Z. Luo [1], J. Gong [1], Y.Z. Xu [2], and W.X. She[1]**

[1]College of Shipbuilding Engineering, Harbin Engineering University, China
[2]CD-adapco software technology (shanghai) co ltd, shanghai, China

*Presenting author: wutiechengship@126.com
†Corresponding author: guochunyu_heu@outlook.com

Abstract

In the present study, the wake field of a ship model was measured using a Stereo towed under water particle image velocimetry (PIV) system to identify the flow characteristics around a hull with a high block coefficient. Meanwhile, a numerical model is established based on computational fluid dynamics to provide expansion information for the experimental measurement. Unsteady Reynolds-averaged Navier-Stokes (URANS) method is employed to track the main features in the wake field. The bare hull wake was measured for a 76000 DWT Panamax Bulker model, at a Froude number of 0.167 based on the length between perpendiculars (L_{PP}).

The measurement results show that the 76000 DWT Panamax Bulker has a similar U shaped stern as in a large vessel, such as the KVLCC, and incoming flows through the U shaped stern gradually produce a strong stern bilge vortex. The axial velocity of the propeller plan has a "hook-shaped" velocity contour at the rotating bilge vortex. In addition, under the bilge vortex, there is a propeller boss cap vortex in the opposite direction of rotation of the bilge vortex, which is close to the longitudinal section in the center disc of the ship. The presence of the hook-like bilge vortex and the counter-rotating boss cap vortex has a positive effect on improving the rotation efficiency of the propeller. The numerical results are consistent with the experimental results. By the analysis of three-dimensional flow fields, the same velocity contours (U shaped and hook-shaped) and similar vortex structures are observed. The numerical calculation has advantages in performing the spatial evolution of the complex wake field. The combination of numerical calculation method and the experimental method could enrich the wake field study for large vessels, and make the results more reliable.

Keywords: Nominal wake field; URNAS; Flow field characteristics; Stereo particle image velocimetry (SPIV); Vorticity Field

1. INTRODUCTION

With the aim of conducting numerical simulations of a ship's performance using CFD and experimental fluid dynamics (EFD), Zhang (2010) applied the Reynolds-Averaged Navier-Stokes (RANS) method to the study of the viscous flow field of a KCS. Ahmed *et al.* (2009) applied the viscous and potential flow methods to the simulation of the free surface flow around a very large crude carrier (VLCC) hull. Wilson, *et al.* (2006) used the unsteady RANS method (URANS) to simulate the viscous flow and motions of a surface combatant ship. Wan, *et al.* (2010) applied the level set method to the study of the viscous flows around a Wigley ship. Longo and Stern (2002) and Xing (2012) applied the CFD and EFD methods to the study of the hydrodynamic performance of a ship subject to ODCs (amplitude motion and maneuvering condition). During the Gothenburg 2010 Workshop on Numerical Ship Hydrodynamics, 33 special interest groups devised 18 test cases based on various ship

models, including the KCS, KRISO very large crude carrier (KVLCC), and the DTMB 5415. Larsson, *et al.* (2010, 2014) collated the data from these cases and then carried out statistical computations and analysis. Kim, *et al.* (2001) of KRISO used a pitot tube and other equipment to measure and study the surrounding flow fields and free-surface wave-making resistance of the KCS and KVLCC. Joon *et al.* (2003) used particle image velocimetry (PIV) to measure the turbulent flow fields at various sections of the KCS for which *Lpp* = 1.5. Lee, *et al.* (2009) employed PIV to determine the wake field of a particular container ship under the fully loaded and in-ballast conditions.

Relative to experimental research, CFD technology could realize major resource savings. With the introduction of a specific calculation model, we could perform a more accurate simulation analysis for problems related to ship and ocean engineering. In addition, with improvements in computer technology and CFD technology, the results of simulation and calculation accuracy are becoming more precise.

2. Methodology

2.1 Model Geometries and Test conditions

A model of the HEU model, a 76000 DWT Panamax Bulker, was used as the test case for this study (Figure 1). PIV measurements of stern flow fields was carried out by HEU, and used for validating the CFD simulations. The measurement model parameters and operating conditions are listed in Tables 1 and 2, respectively.

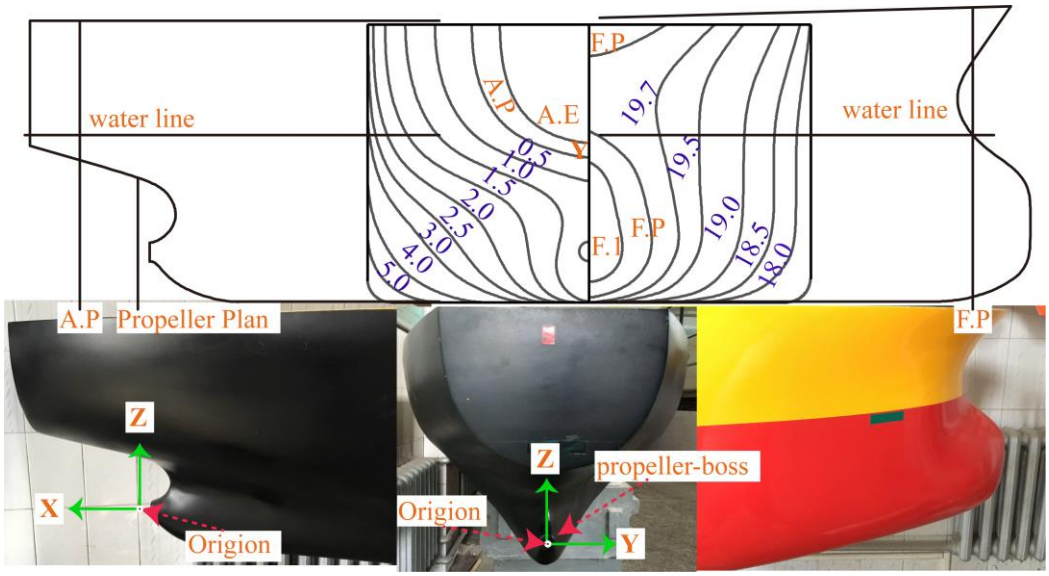


Fig. 1 76000 DWT Panamax Bulker model and local coordinate system for wake field measurements

Table 1 Geometry data and C/EFD test for HEU model and full scale model of hull.

Principal dimension Parameter		Full Scale	HEU Model
Length between perpendiculars	L_{pp}(m)	217	4.82
Bean	B(m)	32.25	0.72
Draft	T(m)	12.40	0.28
Design speed	U(m/s)	7.716	1.15
Scale factor	λ	1	40.000
Propeller radius	R(m)	3.375	0.075

Table 2 Operating conditions for calculation and test of HEU model.

Operating condition	CFD	EFD

Fore draft (FP/m)	0.28	0.28
After draft (AP/m)	0.28	0.28
Trim angle (°)	0	0
Ship posture	Even keel	Even keel
fluid medium	Water, air (all conditions)	
Velocity (m/s)	1.15 (all conditions)	
Waves	Flat wave (all conditions)	

2.2 Testing facility and SPIV system

Tests were conducted in the Harbin engineering university towing tank, which is 108 m long, 7 m wide, and 3.5 m deep. The main bridge and sub-bridge were installed on a towing carriage with a maximum speed of 6.5 m/s and a speed control tolerance less than 0.003 m/s. The motion capture system and the PIV system were attached to bridge and sub-bridge for data measurement.

The present SPIV is a Dantec Dynamics Inc. (Skovlunde, Denmark) custom-designed, under water SPIV system. The SPIV system arrangement form is similar with the one referred to in the literature (Falchi et al., 2014;Yoon et al.,2015),which consists of a fully underwater probe with a modular structure facilitating assembly in different configurations.

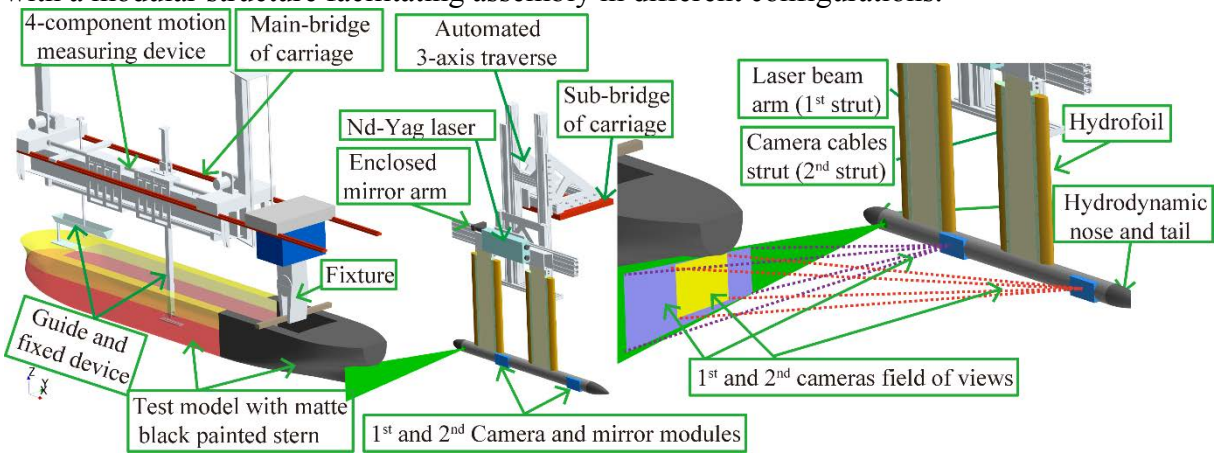


Fig.2 SPIV set up with test model

2.3 Governing Equations

The motion-compliant continuity and momentum conservation equations for incompressible Newtonian fluid motion（David, 1994) are as follows:

$$\frac{\partial \rho}{\partial t}+\frac{\partial\left(\rho u_i\right)}{\partial x_i}=0 \tag{1}$$

$$\frac{\partial\left(\rho u_i\right)}{\partial t}+\frac{\partial}{\partial x_j}\left(\rho u_i u_j\right)=-\frac{\partial p}{\partial x_j}+\frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}-\rho \overline{u_i' u_j'}\right)+S_j \tag{2}$$

where u_i and u_j are the time-averaged values (i, j = 1, 2, 3) of the velocity component; p is the time-averaged value of the pressure; ρ is the fluid density; μ is the dynamic viscosity coefficient; $\rho \overline{u_i' u_j'}$ is the Reynolds stress term; and S_j is the source term.

2.4 Turbulence Model and Treatment of Free Surface

In the analysis, a computational finite volume method was used to analyze a case 76000 DWT Panamax Bulker by using a segregated flow solver. The formulation is fully coupled and based on pressure, with second-order upwind spatial discretization being used for the

convective flux terms and second-order central discretization being used for the diffusion terms. An implicit pseudo time-marching scheme is used to determine an unsteady-state solution. Preconditioning is used to make this approach suitable for low-speed, isothermal flows (Weiss, 1995). A point-implicit (Gauss-Seidel) linear system solver is used with algebraic multigrid acceleration to solve the resulting discrete linear system at each iteration. The shear stress transport $k-\omega$ model (Menter, 1994) was used in this study to simulate the strong adverse pressure gradient flow field and consider the impact of the shear force exerted by the wall of the model. The free surface was modeled using the two-phase volume of fluid (VOF) technique (Hirt, 1981).

2.5 Wave Damping

A numerical wave beach was established at the inlet, outlet, and side boundaries of the tank to avoid any unphysical reflections. Waves can be damped by introducing resistance to vertical motion. The method devised by Choi and Yoon (2009) adds a resistance term to the equation for *w*-velocity:

$$S_z^d = \rho\left(f_1 + f_2|w|\right)\frac{e^{\kappa}-1}{e^{1}-1}w \tag{3}$$

with

$$\kappa = \left(\frac{x - x_{sd}}{x_{ed} - x_{sd}}\right)^{n_d} \tag{4}$$

Where x_{sd} is the starting point for wave damping (propagation in the *x*-direction); x_{ed} is the end point for wave damping (boundary); f_1, f_2, and nd are the parameters of the damping model; and *w* is the vertical velocity component.

Those unphysical reflections are mainly due to two reasons: Interaction between the waves and the edge where the mesh size suddenly changes (that is, the border of two different yet adjacent volumetric mesh refinements); Interaction between the waves and the boundaries of the computational domain.

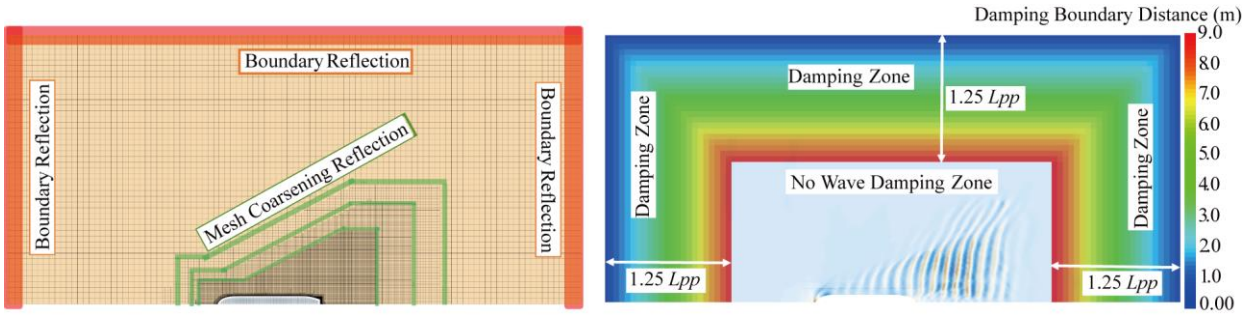


Fig. 3 Unphysical reflections of mesh refinement around the hull and the boundaries of the domain (Left); Wave Damping Zone and Length (Right)

Due to those considerations, the first source of reflection is the wake refinement around the hull. Considering that usually 3 different volumetric refinements are used in order to gradually coarsen the mesh, ideally the wave damping should start before the coarsest grid level is reached. Indeed there would be no point in starting to damp after this refinement because one of the possible sources of reflection will not be included into the damping zone. Damping "as close as possible to the hull" is not a good choice because the waves generated around the hull give a non-negligible contribution to the drag, and damping these waves would lead to the wrong result.

In conclusion, the Wave Damping must start at a location where the mesh is still fine, but not too close to the hull, in order to avoid any influence on the wave pattern. To reach that

condition, it is important to size the 3 wake refinements in order to still guarantee a sufficiently fine mesh far from the hull and let the damping start inside the last one.

2.6 CFD Mesh

The cut meshes are generated using STAR-CCM+, Through the use of arbitrarily shaped volumetric control regions, the volume mesh was refined in different ways, as illustrated in Figure 3, which shows different cross-sectional views. The first is the overall mesh for the model. A Kelvin mesh refinement was defined in the region of the Kelvin wave system to capture the wave pattern and better resolve the water/air layer interaction. A robust automated prism layer meshing algorithm was used to capture the boundary layer, with a two-layer all $y+$ wall treatment. The $y+$ values were kept in the 30–60 range (Pope, 2001).

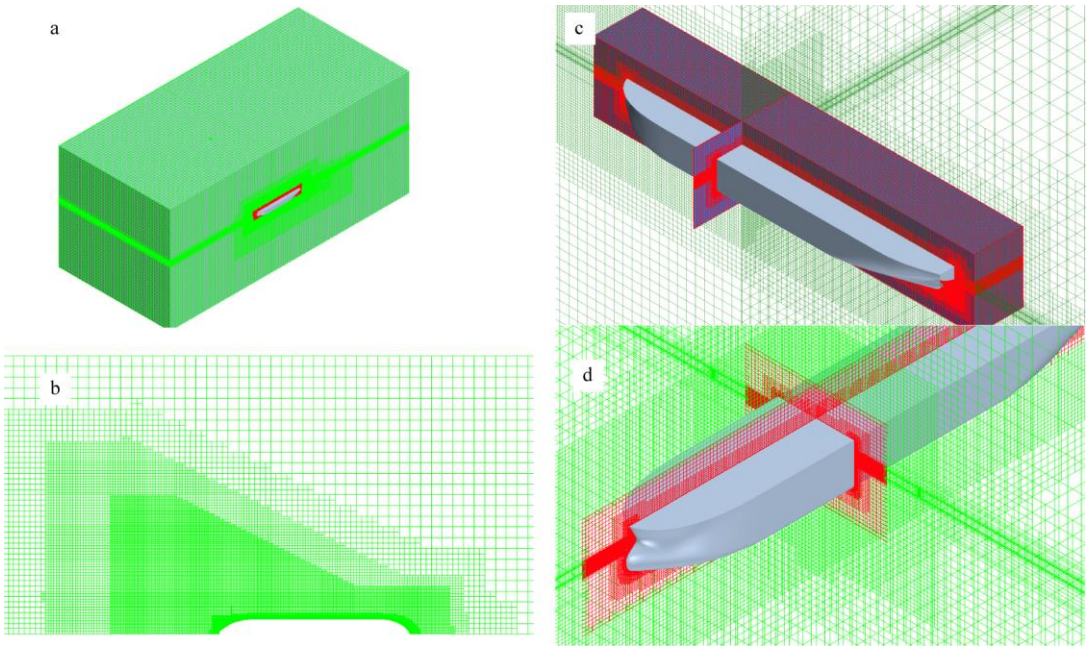


Fig. 4 Computational meshes. a) Overall mesh; b) Mesh area of Kelvin wave system; c) Mesh area of hull region; d) Prismatic boundary layer mesh of the bow.

2.7 Computational Domain and Boundary Conditions

The computational domain is $-2.0Lpp \leqq x \leqq 4.0Lpp$, $0.0Lpp \leqq y \leqq 2.5Lpp$, and $-1.5Lpp \leqq z \leqq 1.25Lpp$. The coordinate system for comparison is fixed to $x = 0.0$ (FP) on the undisturbed water plane. The boundary conditions are listed in Table 3.

Table 3 Boundary parameters.

Boundary	Boundary Parameter
Inlet	Inlet with a flat vof wave defined volume fraction, flow speed = 1.15 m/s
Outlet	Outlet with hydrostatic pressure of flat vof wave
Bottom/Top	Same as inlet
Ship	Wall with no slip condition
Symmetry plane	Along centerline of hull
Side	Same as Symmetry plane
Overset Boundary	Overset Interface

3. RESULTS and DISCUSSION

3.1 Velocity Wake Field at Stern

Fig. 5-8 show the results of the global wake field of a 76000DWT Panamax Bulker, where the u/U (v/U, w/U, S/U) velocity contour is on the left, $\Delta u/U = 0.1$ ($\Delta v/U = 0.05$, $\Delta w/U = 0.05$,

$\Delta S/U$ = 0.05), and the right side of the bulk flow field is the streamline chart for measuring sections. The predicted results for these global wake field and the available experimental results are in good agreement

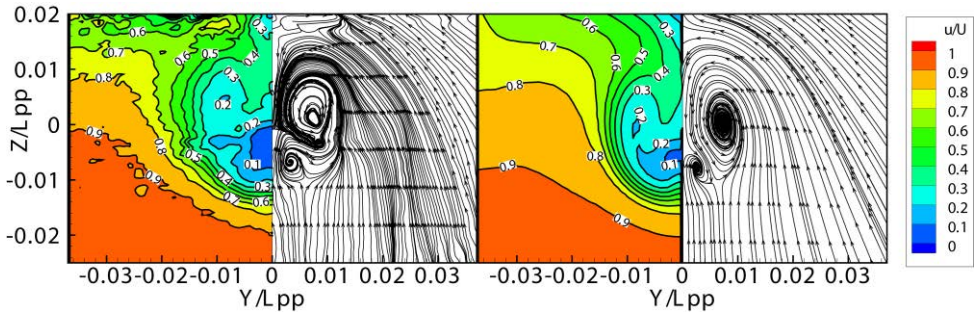


Fig.5 Nominal wake field of *u/U* and stream lines : EFD (left) and CFD (right)

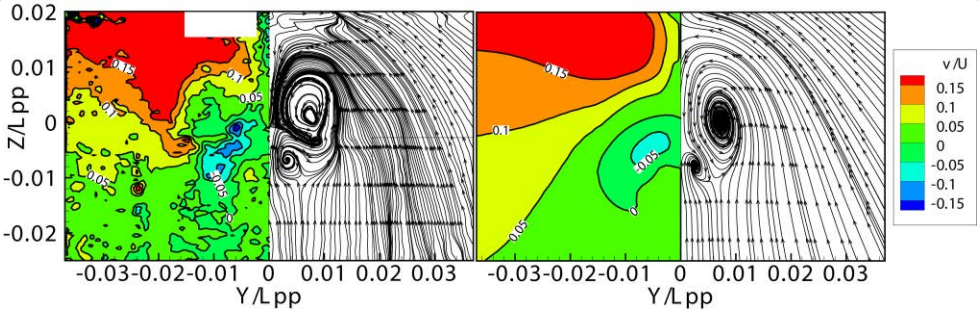


Fig.6 Nominal wake field of *v/U* and stream lines : EFD (left) and CFD (right)

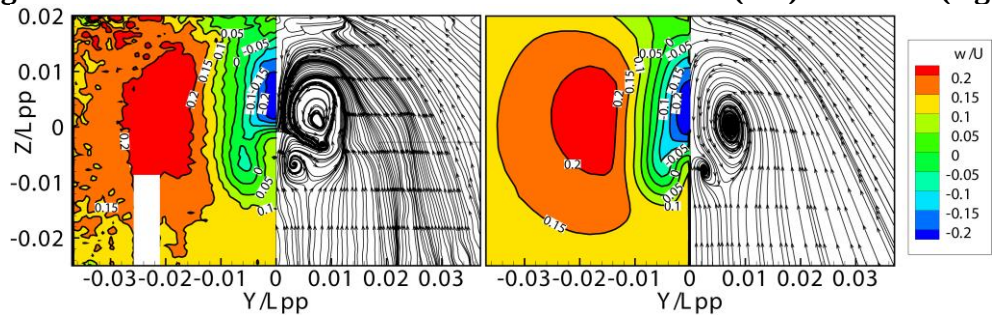


Fig.7 Nominal wake field of *w/U* and stream lines : EFD (left) and CFD (right)

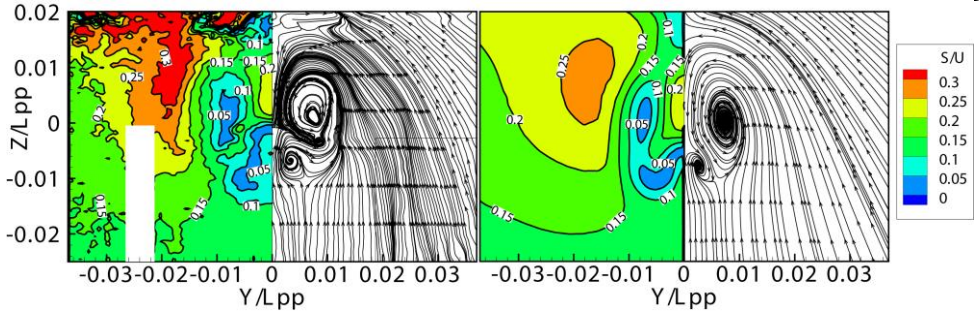


Fig.8 Nominal wake field of *S/U* and stream lines : EFD (left) and CFD (right)

Fig. 5 shows that the 76000 DWT Panamax Bulker has a similar U shaped stern (Fig. 1) as in a large vessel, such as the KVLCC (Seo et al.,2016), and incoming flows through the U shaped stern gradually produce a strong stern bilge vortex. The axial velocity of the disc surface obviously has a "hook-like" velocity contour at the rotating bilge vortex with a velocity value of u/U = 0.3. The distortion of this "hook-like" velocity contour is due to the fact that the fluid with low kinetic energy is transmitted by the strong longitudinal bilge vortex to weaken the original speed near the center of the hull to form a local hook-like structure. In addition, under the bilge vortex, there is a propeller boss cap vortex in the opposite direction of rotation of the bilge vortex, which is close to the longitudinal section in

the center disc of the ship. The presence of the hook-like bilge vortex and the counter-rotating boss cap vortex has a positive effect on improving the rotation efficiency of the propeller (Kim et al.,2001). Fig. 6 shows that the positive and negative velocities correspond to the direction of rotation in the bilge vortex and propeller cap vortex. Fig. 7 shows that the *w/U* velocity contour has a regular hierarchical distribution, the positive and negative velocities correspond to the direction of rotation in the bilge vortex and propeller cap vortex. Fig. 8 shows that The intersection point of the "0" *v/U* contour in Fig. 6 and "0" *w/U* contour in Fig. 7 is the rotation center of the bilge vortex and propeller boss cap vortex .

3.2 Vorticity Field at Stern

Figure 8-f shows the contour map of vorticity. For planar data, gradients in the x-direction cannot be calculated, so only rotation around the x-axis can be determined:

$$\omega_x = \frac{\partial W}{\partial y} - \frac{\partial V}{\partial z} \tag{5}$$

As shown in Figs. 8-e and 8-f, there are two sources of vortices: the bilge vortex and the hub cap vortex. The bilge vortex (counter-clockwise on the starboard; clockwise on the port side) is generated from the ship hull and is then shed into the propeller plane. Thus, the vortex is located around the outer boundary layer (u/U = 0.3–0.6), covering a large area. Positive vorticity, corresponding to a counter-clockwise rotating vortex, is produced. The vortex shedding around the shaft is the hub cap vortex. The hub cap vortex (clockwise on the starboard; counter-clockwise on the port side) forms inside this inner wake field with $u/U <$ 0.2. The bilge vortex is above the shaft and the vortex in the low-speed area is below the shaft. The simulated Vortices Field at Stern matches the results of the experiment well.

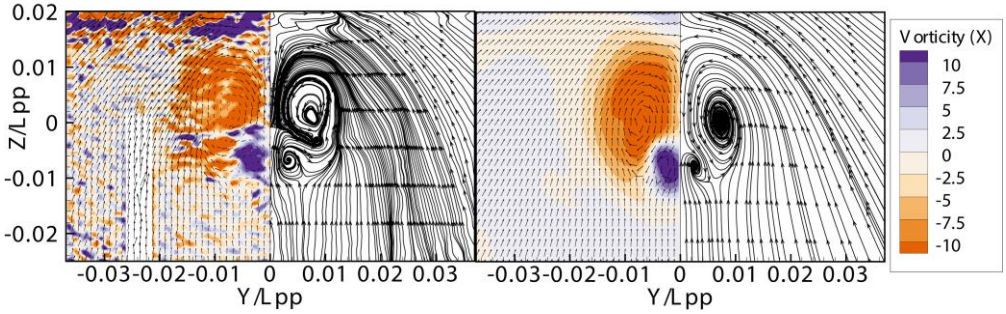


Fig.9 Nominal wake field of S/U and stream lines : EFD (left) and CFD (right)

4. CONCLUSIONS

Through numerical simulations of the viscous flow field around a 76000 DWT Panamax Bulker, the following conclusions were drawn:

The PIV measurement results show that the axial velocity of the propeller plan has a "hook-shaped" velocity contour at the rotating bilge vortex. In addition, under the bilge vortex, there is a propeller boss cap vortex in the opposite direction of rotation of the bilge vortex, which is close to the longitudinal section in the center disc of the ship. The use of PIV to measure a ship's nominal wake field is an important method for directly measuring its detailed turbulent field.

The predicted wake field results and available experimental results are in good agreement. The results show that the mesh and numerical simulation methods are valid.

Acknowledgements

This project was supported by the National Natural Science Foundation of China (Grant Nos. 51379043, 51209048 and 51409063), and thanks for the technical support provided by Dr. Xu Yongze and Dou Pengfei of CD-adapco software in Siemens Ltd. China.

References

Ahmed, Y. and Soares, C.G., 2009. Simulation of free surface flow around a VLCC hull using viscous and potential flow methods. *Ocean Engineering*, 36, 691-696.

Choi, J. and Yoon, S.B., 2009. Numerical simulations using momentum source wave-maker applied to RANS equation model. *Coastal Engineering*, 56(10), 1043-1060.

David C. Wilcox, 1994. Turbulence Modeling for CFD. DOW Industries, Inc. La Canada, California, 15-19p.

Falchi, M., Felli, M., Grizzi, S., Aloisio, G., Broglia, R., & Stern, F., 2014. SPIV measurements around the delft 372 catamaran in steady drift. Exp. Fluids, 55(11), 1-20.

Hirt, C.W. and Nichols, B.D., 1981. Volume of fluid /VOF/ method for the dynamics of free boundaries. *Journal of Computational Physics*, 39(81), 201-225.

Kim, W.J.; Van, S.H., and Kim, D.H., 2001. Measurement of flows around modern commercial ship models. *Experiments in Fluids*, 31(5), 567-578.

Larsson, L. and Stern, F., 2014. Numerical ship hydrodynamics-an assessment of the Gothenburg 2010 Workshop. Springer, Dordrecht, 84-85p.

Larsson, L.; Stern, F., and Visonneau, M., 2013. CFD in Ship Hydrodynamics—Results of the Gothenburg 2010 Workshop. *Computational Methods in Applied Sciences*, 30(6), 237-259.

Lee, J.Y.; Bu, G.P., and Sang, J.L., 2009. PIV measurements of hull wake behind a container ship model with varying loading condition. *Ocean Engineering*, 36(5), 377–385.

Longo, J. and Stern, F., 2002. Effects of Drift Angle on Model Ship Flow. *Experiments in Fluids*, 32, pp 558-569.

Menter, F.R., 1994. Two-equation eddy-viscosity turbulence models for engineering applications. *Aiaa Journal*, 32(8), 1598-1605.

Pope, S.B., 2001. Turbulent Flows. UK, Cambridge University Press, 806p.

Seo, J., Seol, D.M., Han, B., Rhee, S.H., 2016. Turbulent wake field reconstruction of VLCC models using two-dimensional towed underwater PIV measurements. Ocean Eng. 118, 28–40.

Sang Joon, L.; Min-Seok, K., and Choung-Mook, L., 2003. PIV velocity field measurements of flow around a KRISO 3600TEU container ship model. *Journal of Marine Science & Technology*, 8(2), 76-87.

Wan, D.C.; Shen, Z.R., and Ma, J., 2010. Numerical simulations of viscous flows around surface ship by level set method. *Journal of Hydrodynamics*, 22(5), supplement, 271-277.

Weiss, J.M. and Smith W.A., 1995. Preconditioning applied to variable and constant density flows. *Aiaa Journal*, 33(11), 2050-2057.

Wilson, R.V.; Carrica, P.M., and Stern, F., 2006. Unsteady RANS method for ship motions with application to roll for a surface combatant. *Computers & Fluids*, 35, 501-524.

Xing, T.; Bhushan, S., and Stern, F., 2012. Vortical and Turbulent Structures for KVLCC2 at Drift Angle 0, 12, and 30 Degrees. *Ocean Engineering*, 55, 23-43.

Yoon, H., Longo, J., Toda, Y., & Stern, F., 2015. Benchmark CFD validation data for surface combatant 5415 in PMM maneuvers – part ii: phase-averaged stereoscopic PIV flow field measurements. Ocean. Eng.109, 735-750.

Zhang, Z.R., 2010. Verification and validation for RANS simulation of KCS container ship without/with propeller. *Journal of Hydrodynamics,* 22(5), supplement: 932-939.

Three-dimensional Simulation of Liquid Sloshing in an Elastic Tank

Youlin Zhang[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: sir.zhyl@163.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

Abstract

Sloshing in a liquid tank of huge size could potentially induce the structural vibration and fatigue damage. Though many numerous researches of sloshing have been conducted to characterize the impact phenomenon, the problem still remains to be addressed since the neglect of hydro-elastic behaviors of structure which will presence in the real phenomenon. In this paper a hybrid method has been developed to study the three-dimensional (3D) liquid sloshing with the consideration of structural elasticity. The improved moving particle semi-implicit (MPS) method is employed to simulate the evolution of 3D flow. The finite element method (FEM) is employed to calculate the vibration of the flexible tank wall. The MPS and FEM methods are coupled with a partition strategy within the fully Lagrangian system. Then, the sloshing in a 3D elastic tank is numerically investigated and results are compared with those corresponding to a 3D rigid tank. The effects of the structural elasticity on the sloshing behaviors are discussed.

Keywords: Particle method; Moving Particle Semi-Implicit (MPS); finite element method (FEM); Fluid structure interaction (FSI); Sloshing; MLParticle-SJTU solver.

Introduction

Fluid structure interaction (FSI) is omnipresent in nature and in many engineering fields. For instance, the sloshing phenomenon occurred in a partially loaded oil tanker or liquid natural gas ship is a typical FSI problem involving multi-physic, yet interrelated liquid, gas and solid domains interact with each other as a unit [1]. For this intricate problem, it's hard to achieve analytical solution whereas laboratory experiment is limited in scope [2]. Considering the fundamental physics involved in the problem can be obtained by numerical simulations, active numerical researches have been carried out in the field of FSI over the past two decades and multiple numerical models were developed [3].

Conventionally, the FSI problems are solved with the fluid field modeled in an arbitrary-Lagrangian-Eulerian (ALE) formulation while the structure field modeled in a Lagrangian formulation. For this approach, grids are necessary to tessellate the solution domain. As the structure undergoes large deformations, the fluid mesh may get highly distorted, especially for a 3D FSI simulation. Although the re-meshing or mesh-updating techniques can be employed to improve the mesh quality as the solution is advanced, extra charge of computation time is unavoidable [4]. Furthermore, the distorted mesh is detrimental for the accuracy of free surface which plays a crucial role in the sloshing phenomenon.

In the recent decade, the fully Lagrangian approaches for both fluid and structure fields are utilized to model the FSI problems since they are flexible in dealing with structural deformation, tracking of free surface, and without having to cope with the nonlinear

convective term which appears in the momentum equation in the Eulerian framework. Till to now, several representative Lagrangian methods, such as the smoothed particle hydrodynamics (SPH) method [5], the particle finite element method (PFEM) [6], the material point method (MPM) [7], etc. have been proposed for fluid domain analysis while the FEM method is employed for the structure domain analysis. According to the prior results, disordered pressure fields of fluid are observed, although these methods have shown the great potential for the practical FSI problems involving with motion of fluid or structure particles, surface waves and water splashing. Comparatively, another representative Lagrangian method, the moving particle semi-implicit (MPS) method which is originally proposed by Koshizuka and Oka for incompressible flow [8], is able to achieve smooth fluid pressure field since lots of improvements were proposed to suppress the numerical unphysical pressure oscillation [9]-[11]. In the nearly few years, the MPS method has been introduced into the FSI problems [12]-[16], and results shown that this method is stable and reasonable accurate for simulating nonlinear FSI problems. Hence, the MPS method is employed for the computation of fluid domain of the FSI problem in this paper.

Indeed, all the aforementioned Lagrangian methods for the FSI problems are implemented within two-dimensional space [17]-[22]. To address the practical FSI problems, it's essential to extend these methods into 3D space. However, it's a time consuming task of simulation while the structural domain is dispersed by grids with the nodes coincide with the fluid particles on the interface. Normally a much larger mesh size compared to the size of fluid particle is accurate enough to simulate the structure field. In the present work, the MPS and FEM coupled method is developed for 3D FSI problems, and an interpolation scheme is proposed for the communication on the isomerous interface where the size of structural boundary grids differs from the size of fluid particles. Then, the MPS-FEM coupled method is applied to the practical problem of violent sloshing flow interacting with elastic tank walls, and influence of structural elasticity on the sloshing phenomenon is comparatively investigated.

Numerical methods

In the present study, the FSI problem is numerically studied by a partitioned coupled approach, of which the flow equations and the structural equations are solved separately. Here, the fluid domain is calculated by our in-house particle solver MLParticle-SJTU [23]-[26] based on improved MPS method and the structural domain is calculated by the FEM method.

Fluid solver based on MPS method

Governing equations for incompressible viscous fluid in Lagrangian system are

$$\nabla \cdot \boldsymbol{V} = 0 \tag{1}$$

$$\frac{D\boldsymbol{V}}{Dt} = -\frac{1}{\rho}\nabla P + \nu\nabla^2\boldsymbol{V} + \boldsymbol{g} \tag{2}$$

where $\boldsymbol{V}$, t, ρ, P, ν and $\boldsymbol{g}$ represent the velocity vector, time, water density, pressure, kinematic viscosity and the gravity acceleration vector, respectively.

In particle method, governing equations should be expressed by the particle interaction models based on the kernel function. Here, the kernel function presented by Zhang et al.[23] is employed.

$$W(r)=\begin{cases}\dfrac{r_e}{0.85r+0.15r_e}-1 & 0\le r<r_e\\ 0 & r_e\le r\end{cases} \tag{3}$$

where r is distance between particles and r_e is the effect radius.

The particle interaction models, including the differential operators of gradient, divergence and Laplacian, are defined as

$$<\nabla\phi>_i=\frac{dim}{n^0}\sum_{j\neq i}\frac{\phi_j+\phi_i}{|\boldsymbol{r}_j-\boldsymbol{r}_i|^2}(\boldsymbol{r}_j-\boldsymbol{r}_i)\cdot W(|\boldsymbol{r}_j-\boldsymbol{r}_i|) \tag{4}$$

$$<\nabla\cdot\boldsymbol{\Phi}>_i=\frac{dim}{n^0}\sum_{j\neq i}\frac{(\boldsymbol{\Phi}_j-\boldsymbol{\Phi}_i)\cdot(\boldsymbol{r}_j-\boldsymbol{r}_i)}{|\boldsymbol{r}_j-\boldsymbol{r}_i|^2}W(|\boldsymbol{r}_j-\boldsymbol{r}_i|) \tag{5}$$

$$<\nabla^2\phi>_i=\frac{2dim}{n^0\lambda}\sum_{j\neq i}(\phi_j-\phi_i)\cdot W(|\boldsymbol{r}_j-\boldsymbol{r}_i|) \tag{6}$$

where ϕ is an arbitrary scalar function, $\boldsymbol{\Phi}$ is an arbitrary vector, *dim* is the number of space dimensions, n^0 is the initial particle number density for incompressible flow, λ is a parameter defined as

$$\lambda=\frac{\sum_{j\neq i}W(|\boldsymbol{r}_j-\boldsymbol{r}_i|)\cdot|\boldsymbol{r}_j-\boldsymbol{r}_i|^2}{\sum_{j\neq i}W(|\boldsymbol{r}_j-\boldsymbol{r}_i|)} \tag{7}$$

which is introduced to keep the variance increase equal to that of the analytical solution [8].

The incompressible condition of MPS method is represented by keeping the particle number density constant. In each time step, there are two stages: first, temporal velocity of particles is calculated based on viscous and gravitational forces, and particles are moved according to the temporal velocity; second, pressure is implicitly calculated by solving a Poisson equation, and the velocity and position of particles are updated according to the obtained pressure. The Pressure Poisson Equation (PPE) in present MPS solver is defined as

$$<\nabla^2P^{n+1}>_i=(1-\gamma)\frac{\rho}{\Delta t}\nabla\cdot\boldsymbol{V}_i^*-\gamma\frac{\rho}{\Delta t^2}\frac{<n^*>_i-n^0}{n^0} \tag{8}$$

where γ is a blending parameter with a value between 0 and 1. The range of $0.01\le\gamma\le0.05$ is better according to numerical experiments conducted by Lee et al.[27] In this paper, $\gamma=0.01$ is adopted for all simulations.

For the MPS method, pressure of the fluid domain is closely affected by the accuracy of free surface detection. In present solver, we employ a free surface detection method by Zhang et al.[23] and defined as

$$<\boldsymbol{F}>_i=\frac{dim}{n^0}\sum_{j\neq i}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}(\boldsymbol{r}_i-\boldsymbol{r}_j)W(\boldsymbol{r}_{ij}) \tag{9}$$

where the vector function $\boldsymbol{F}$ represents the asymmetry of arrangements of neighbor particles. Particle satisfying

$$<|\boldsymbol{F}|>_i\ >\ 0.9|\boldsymbol{F}|^0 \tag{10}$$

is considered as free surface particle, where $|\boldsymbol{F}|^0$ is the initial value of $|\boldsymbol{F}|$ for surface particle.

Structure solver based on FEM method

In present study, the FEM method is employed to solve the deformation of structure which is governed by the equations expressed as

$$\mathbf{M}\,\ddot{\boldsymbol{y}} + \mathbf{C}\,\dot{\boldsymbol{y}} + \mathbf{K}\,\boldsymbol{y} = \boldsymbol{F}(t) \tag{11}$$

$$\mathbf{C} = \alpha_1 \mathbf{M} + \alpha_2 \mathbf{K} \tag{12}$$

where **M**, **C**, **K** are the mass matrix, the Rayleigh damping matrix, the stiffness matrix of the structure, respectively. $\boldsymbol{F}$ is the external force vector acting on structure, and varies with computational time. $\boldsymbol{y}$ is the displacement vector of structure. α_1 and α_2 are coefficients which are related with natural frequencies and damping ratios of structure.

To solve the structural dynamic equation, another two group functions should be supplemented to set up a closed-form equation system. Here, Taylor's expansions of velocity and displacement developed by Newmark [28] are employed:

$$\dot{\boldsymbol{y}}_{t+\Delta t} = \dot{\boldsymbol{y}}_t + (1-\gamma)\ddot{\boldsymbol{y}}_t \Delta t + \gamma \ddot{\boldsymbol{y}}_{t+\Delta t}\Delta t \quad , \quad 0 < \gamma < 1 \tag{13}$$

$$\boldsymbol{y}_{t+\Delta t} = \boldsymbol{y}_t + \dot{\boldsymbol{y}}_t \Delta t + \frac{1-2\beta}{2}\ddot{\boldsymbol{y}}_t \Delta t^2 + \beta \ddot{\boldsymbol{y}}_{t+\Delta t}\Delta t^2 \quad , \quad 0 < \beta < 1 \tag{14}$$

where β and γ are important parameters of the Newmark method, and selected as β=0.25, γ=0.5 for all simulations in present paper. The nodal displacements at $t = t+\Delta t$ can be solved by the following formula [29]:

$$\bar{\mathbf{K}}\,\boldsymbol{y}_{t+\Delta t} = \bar{\boldsymbol{F}}_{t+\Delta t} \tag{15-a}$$

$$\bar{\mathbf{K}} = \mathbf{K} + a_0\mathbf{M} + a_1\mathbf{C} \tag{15-b}$$

$$\bar{\boldsymbol{F}}_{t+\Delta t} = \boldsymbol{F}_t + \mathbf{M}(a_0\boldsymbol{y}_t + a_2\dot{\boldsymbol{y}}_t + a_3\ddot{\boldsymbol{y}}_t) + \mathbf{C}(a_1\boldsymbol{y}_t + a_4\dot{\boldsymbol{y}}_t + a_5\ddot{\boldsymbol{y}}_t) \tag{15-c}$$

$$\begin{aligned} &a_0 = \frac{1}{\beta\Delta t^2},\ a_1 = \frac{\gamma}{\beta\Delta t},\ a_2 = \frac{1}{\beta\Delta t},\ a_3 = \frac{1}{2\beta} - 1, a_4 = \frac{\gamma}{\beta} - 1, \\ &a_5 = \frac{\Delta t}{2}\left(\frac{\gamma}{\beta} - 2\right),\ a_6 = \Delta t(1-\gamma),\ a_7 = \gamma\Delta t \end{aligned} \tag{15-d}$$

where $\bar{\mathbf{K}}$ and $\bar{\boldsymbol{F}}$ are so-called effective stiffness matrix and effective force vector, respectively. Finally, the accelerations and velocities corresponding to the next time step are updated as follows.

$$\ddot{\boldsymbol{y}}_{t+\Delta t} = a_0(\boldsymbol{y}_{t+\Delta t} - \boldsymbol{y}_t) - a_2\dot{\boldsymbol{y}}_t - a_3\ddot{\boldsymbol{y}}_t \tag{16}$$

$$\dot{\boldsymbol{y}}_{t+\Delta t} = \dot{\boldsymbol{y}}_t + a_6\ddot{\boldsymbol{y}}_t + a_7\ddot{\boldsymbol{y}}_{t+\Delta t} \tag{17}$$

Data interpolation on the interface between fluid and structure domain

For the simulation of 3D FSI problems based on aforementioned MPS-FEM coupled method, the space of fluid domain will be dispersed by particles while the space of structural domain

will be dispersed by grids. In general, the fine particles should be arranged within the fluid domain to keep a satisfactory precision for the fluid analysis. By contrast, the much coarser grids could be accurate enough for the structure analysis, which indicates that the fluid particles are not coincided with the structural nodes on the interface between the fluid and structure domain. Hence, the isomerous interface between the two domains may result in the challenge of data exchange in the process of FSI simulation. In the present study, special data interpolation technique is implied to apply the external force carried by the fluid particles onto the structural nodes and update the positions of boundary particles corresponding to the displacements of structural nodes.

For the transformation of force from the fluid domain to the structural boundary, the schematic diagram of the technique is shown in Figure 1 and the procedure of interpolation can be summarized as below.

(1) The mapping relationship between the boundary particle and the structural element will be established while the particle is arranged within the element at the initial time instant.

(2) The external force Q_j acting on the structural boundary is calculated by the formula

$$Q_j = P_j \cdot l_0^2 \quad (i = 1,2,3,4;\ j = 1,2\cdots npe) \tag{18}$$

where P_j is the pressure of the boundary particle obtained from the fluid domain, l_0 is the initial distance between neighbor particles, *npe* is the number of particles on the interface.

(3) The force Q_j carried by the boundary particle j is divided into four parts and assigned onto the four nodes of the element s by the formula (19) with the help of the interpolation vector $\boldsymbol{N}$, which is consisted of the shape functions N_k, N_{xk}, N_{yk}.

$$\boldsymbol{F}_s^e = [\boldsymbol{F}_{s,1}^e\ \boldsymbol{F}_{s,2}^e\ \boldsymbol{F}_{s,3}^e\ \boldsymbol{F}_{s,4}^e] = \boldsymbol{N}^{\mathrm{T}} Q_j\ \ (s = 1,2\cdots ne) \tag{19}$$

$$\boldsymbol{N} = \begin{bmatrix} N_1 & N_{x1} & N_{y1} & N_2 & N_{x2} & N_{y2} & N_3 & N_{x3} & N_{y3} & N_4 & N_{x4} & N_{y4} \end{bmatrix} \tag{20}$$

$$\begin{aligned} N_k &= \frac{1}{8}(1+\xi_k\xi)(1+\eta_k\eta)(2+\xi_k\xi+\eta_k\eta-\xi^2-\eta^2) \\ N_{xk} &= -\frac{1}{8}b\eta_k(1+\xi_k\xi)(1+\eta_k\eta)(1-\eta^2) \\ N_{yk} &= \frac{1}{8}a\xi_k(1+\xi_k\xi)(1+\eta_k\eta)(1-\xi^2) \end{aligned} \tag{21}$$

$$\xi = \frac{x}{a}, \xi_k = \frac{x_k}{a}, \eta = \frac{y}{b}, \eta_k = \frac{y_k}{b} \quad (k = 1,2,3,4) \tag{22}$$

where $\boldsymbol{F}_s^e$ is the force vector regarding the element s, a and b are the half values of the width and height of the element, respectively.

(4) Finally, the equivalent nodal force $\boldsymbol{F}_I$ corresponding to the node I is obtained by the summation of force components regarding to the four neighbor elements.

$$\boldsymbol{F}_I = \boldsymbol{F}_{r,3} + \boldsymbol{F}_{s,4} + \boldsymbol{F}_{m,1} + \boldsymbol{F}_{n,2} \tag{23}$$

where $\boldsymbol{F}_{s,4}$ is the force component contributed by the element s. Schematic program of the neighbor elements adjoining the node I and the concept of nodes numbering within element are shown as Figure 2.

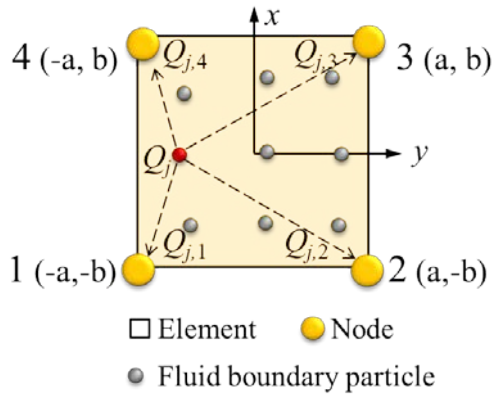


Figure 1.Schematic program of the force interpolation within element

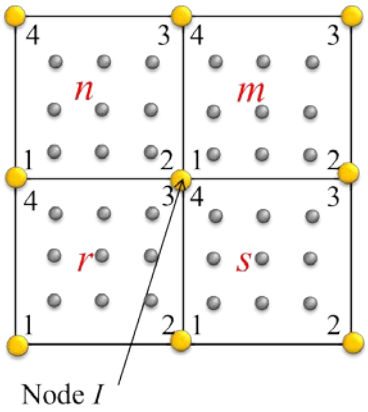


Figure 2. Schematic program of the neighbor elements adjoining the node I

Besides, the fluid boundary which is consisted of particles will deforms corresponding to the deformation of structural boundary. The deflection values of boundary particles w can be obtained by the interpolation based on the shape functions $\boldsymbol{N}$ and the nodal displacements $\boldsymbol{\delta}$.

$$w = \boldsymbol{N}\boldsymbol{\delta} \tag{24}$$

$$\boldsymbol{\delta} = [w_1 \quad \theta_{x1} \quad \theta_{y1} \quad w_2 \quad \theta_{x2} \quad \theta_{y2} \quad w_3 \quad \theta_{x3} \quad \theta_{y3} \quad w_4 \quad \theta_{x4} \quad \theta_{y4}]^{\mathrm{T}} \tag{25}$$

where w_i is the linear displacement of node i, θ_{xi} and θ_{yi} are the angular displacements around the axis x and y, respectively.

Numerical Simulations

In the ship and ocean engineering, the sloshing phenomenon in a partially filled liquid tank is of great importance in assessing the strength of structure and has been intensively studied in the past a few decades. However, most contributions are focused on the mechanism of the nonlinear phenomenon regarding the rigid tank, and the elasticity of tank walls, which plays an important role in the sloshing phenomenon, have not been taken into account.

In this study, the aforementioned MPS-FEM coupled method is employed to simulate the interaction between sloshing flow and three dimensional elastic tank. The influence of structural elasticity on the sloshing phenomenon will be investigated by comparing against the phenomenon in a rigid tank.

Numerical setup

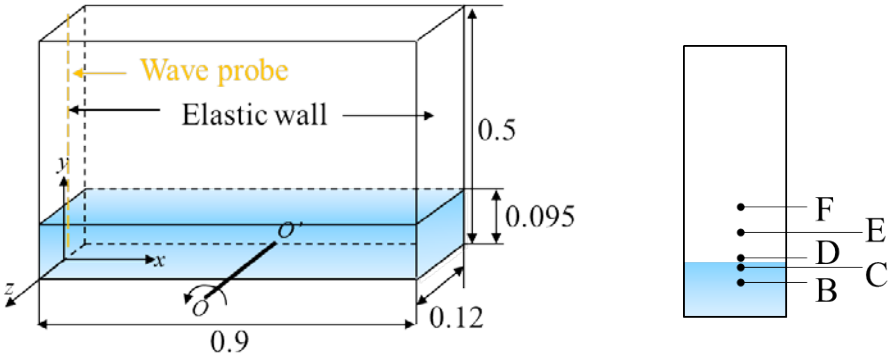
Figure 3 shows the schematic diagram of the 3D computational model. The tank is free to roll around the axis *O-O'* which is the symmetry axis of the floor. The tank is forced to roll harmoniously with the governing equation of motion defined as

$$\theta(t) = \theta_0 \sin(\omega t) \tag{10}$$

where $\theta(t)$ is the rotation angle of the tank, the excitation amplitude θ_0 is set to 4 degrees, the angular frequency of rotation ω is set to 3.857 rad/s. To investigate the climb of water on the

lateral wall of tank, the wave probe is mounted at the point A (0.01, 0, 0). In addition, the vibration of left lateral wall will be measured at the point B (0, 0.05, 0), C (0, 0.08, 0), E (0, 0.15, 0), F (0, 0.2, 0), and the impact pressure will be recorded at the point D (0, 0.095, 0).

In the present simulations, the 3D computational model is dispersed by particles with an initial spacing size (l_0) of 0.005 m for both rigid and elastic tanks. To calculate the structural responses of the elastic walls which would experience the sloshing impact loads, the lateral tank walls are dispersed by elements with the spacing size of 0.01 m. Detailed parameters for both fluid and structural analysis are presented in Table 1. Herein, the Rayleigh's damping has been taken into account for the structural analysis by setting the factor of mass-proportional contribution α_1 as 0.0128 while the factor of stiffness-proportional contribution α_2 as 5.01e^{-7}.



(a) tank (b) measuring points

Figure 3. Schematic diagram of the rolling tank with elastic lateral walls (Unit: m)

Table 1. Simulation parameters of numerical cases

Fluid parameters	Values	Structural parameters	Values
Fluid density (kg/m^3)	1000	Structure density (kg/m^3)	1800
Kinematic viscosity (m^2/s)	5×10^{-5}	Young's modulus (GPa)	10
Gravitational acceleration (s/m^2)	9.81	Poisson's ratio	0.3
Particle spacing (m)	0.005	Element size (m)	0.01
Number of fluid particles	74106	Damping coefficients α_1	0.025
Total number of particles	229816	Damping coefficients α_2	0.0005
Time step size (s)	1×10^{-4}	Time step size (s)	1×10^{-4}

Impact loads on lateral walls

The elasticity of tank walls can give rise to the difference of the impact loads acting on the lateral walls between the elastic and rigid tanks. As shown in Figure 4, the pressure time histories corresponding to rigid tank and elastic tank are measured at the point D. For the pressure in a rigid tank, the well-known character of the impact events, "church roof shape", is observed. For the pressure in the tank with elastic lateral walls, the roof shape of the impact pressure signal shows much different features comparison against that regarding rigid tank. For instance, the peaks of the impact pressure are less than 2200 Pa, which are obviously smaller than those regarding the rigid tank. Furthermore, the pressure curve presents much larger amplitude oscillation, as shown in Figure 4 (b). According to the enlarged signal of

pressure, four peaks and three valleys can be observed within one cycle of tank's roll motion. The oscillation of pressure should be closely linked to the vibration of elastic wall.

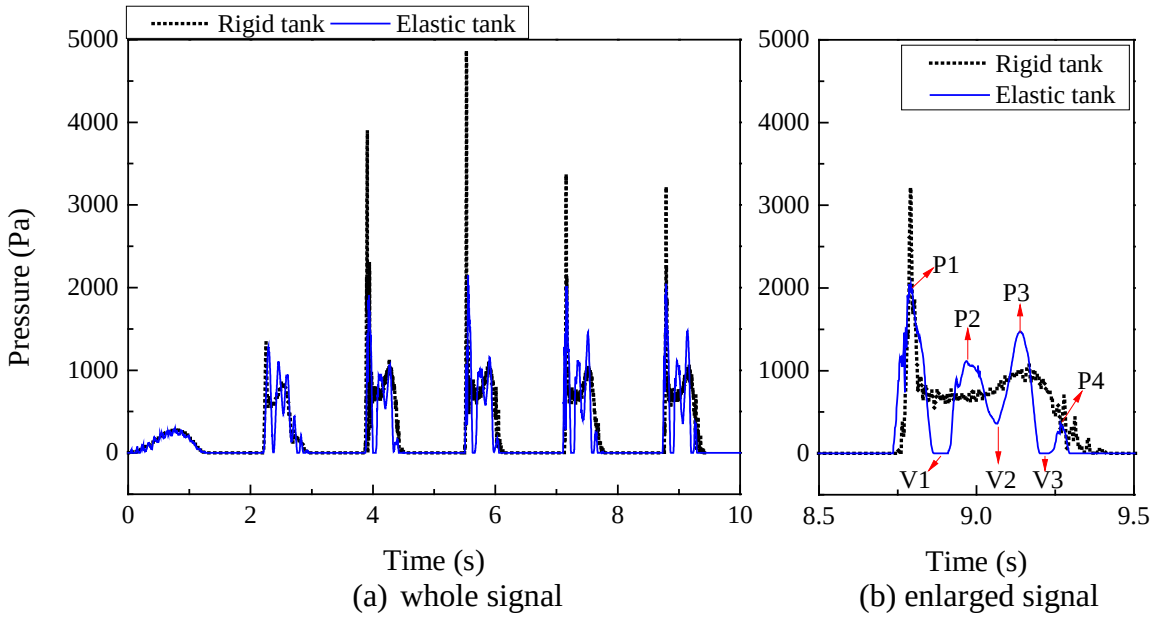


Figure 4. Time histories of pressure at the measuring point D

Climb of water front on lateral walls

The elasticity of tank walls can also lead to the difference of free surface evolutions between the elastic and rigid tanks. Figure 5 shows the time histories of water levels at the measuring point A. According to the figure, the water level regarding the elastic tank is much lower than that of rigid tank. Herein, the water level corresponding to the rigid tank is marked as "level 1" with the value 0.335 m, and the peaks of the curve is 0.5 m which indicates that the water particles hit the roof at a certain time instant since the splashing of water front. In contrast, the water level regarding the elastic tank is marked as "level 2" with the value 0.24 m, and the curve is featured with no pulsing signal which indicates that the splashing phenomenon of water front would not be observed in the region above the measuring point.

Figure 6 shows the climbs of water fronts on the lateral walls in the front view. At the instant t_1, the fluid particles distribute evenly over the rigid wall after the front of sloshing wave impacting onto the lateral wall, while those cluster at the area A with the shape "O" on the elastic wall. At the instant $t_2 = t_1+0.1$ s, the jet water climbs along the rigid wall and the distribution of fluid particles is homogeneous in the z direction. In contrast, the distribution of jet water along the elastic wall is uneven and presents in the "V" form.

To obtain a more clear understanding of the difference of the free surface between the elastic and rigid tanks, the climbs of water fronts on the lateral walls are shown as Figure 7 in the side view. At the instant t_1, the jet is generated after the impact event and turns to climb upward along the rigid wall. In comparison, the direction of the jet water inclines to inside of the tank since the deformation of the wall. At the instant t_2, the water front regarding the rigid wall climbs up to roof of the tank without gap between the water surface and the lateral wall. However, the triangular pocket may exist at the upper corner of the elastic tank while the free surface touches the roof.

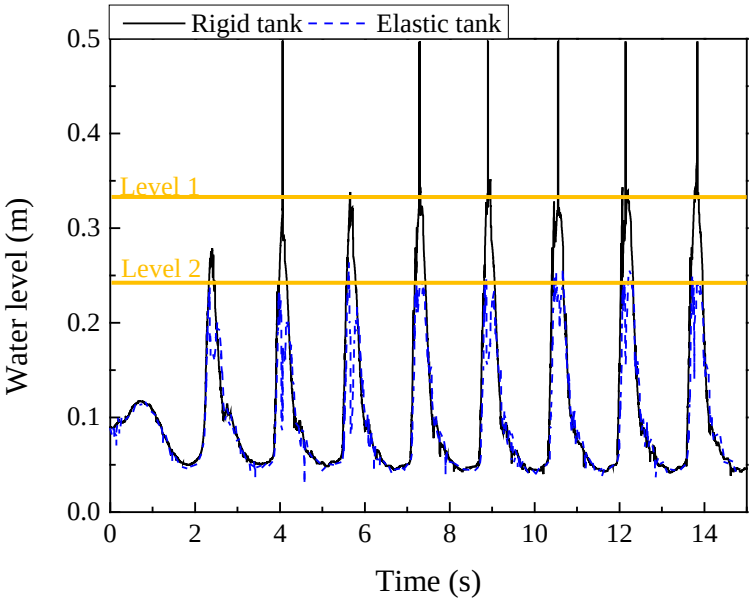
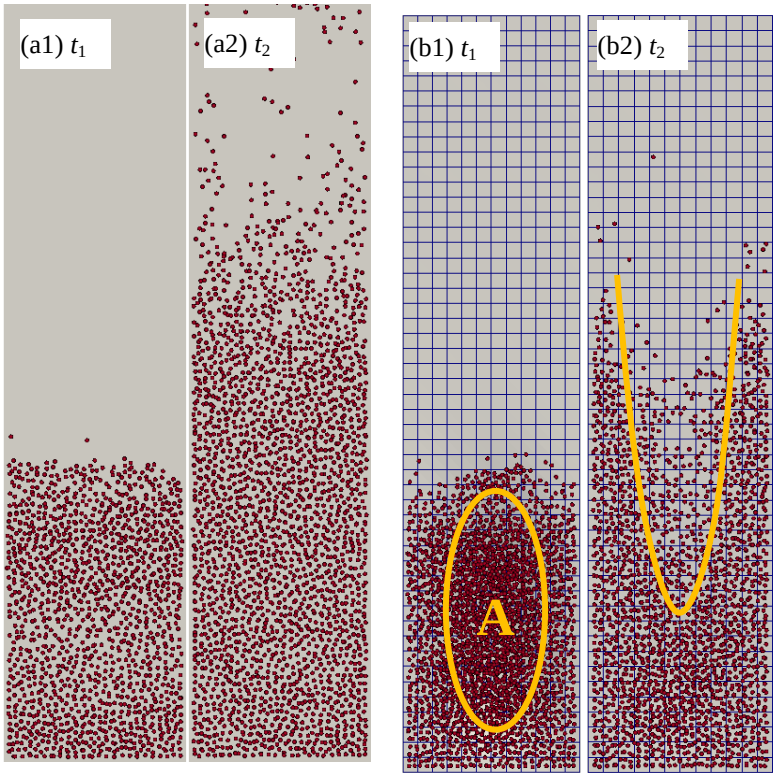
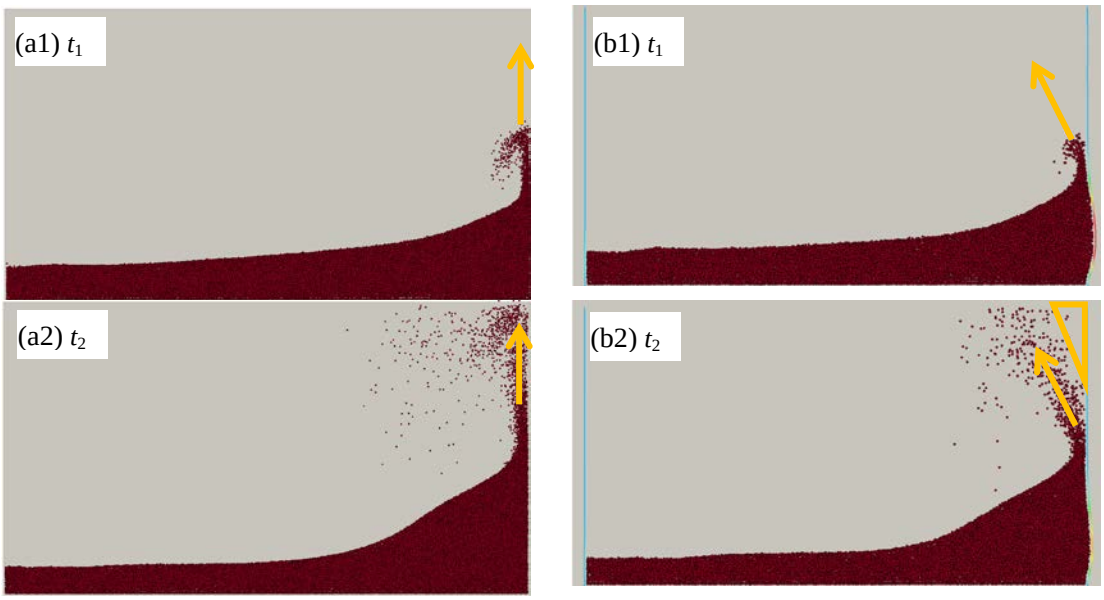


Figure 5. Evolution of water level at point A



(a) Rigid tank (b) Elastic tank

Figure 6. Climb of water front on the lateral wall (front view)



(a) Rigid tank (b) Elastic tank

Figure 7. Climb of water front on the lateral wall (side view)

Deformation of elastic walls

Figure 8 shows the time histories of displacements of measuring points which mounted on left wall of the elastic tank. The similar character of the structural oscillations regarding different measuring points can be observed. According to the trends of the curves, the large amplitude vibrations present periodically and following with small amplitude vibrations. Remarkably, the oscillation amplitude of the measuring point C is much larger than that of other points away from it, which proofs that the elastic wall deforms with 3D feature, as exhibited in Figure 9.

Figure 10 shows the relationship between structural vibration and impact pressure. Herein, the time instant when the pressure going up drastically is marked as t_{impact} which denotes the start of the sloshing impact event, and the time instant when the pressure drops to zero is marked as t_{end} which denotes the end of the impact event. The trend of pressure is in gear with that of structural vibration during the impact stage, which indicates that the impact pressure is sensitive to the vibration of tank wall. It can be inferred that the fluid particles may be drove away from the elastic wall and a gap would generate between the fluid and the lateral wall during the interaction of sloshing flow and the tank. As a result, the pressure which is measured by the contribution of neighbor fluid particles would rapidly reduce and result in the valleys of the pressure signal. In addition, the elastic wall vibrates with the amplitude decreasing gradually beyond the impact stage, which is induced by the joint effects of the structural damping and elastic restoring force.

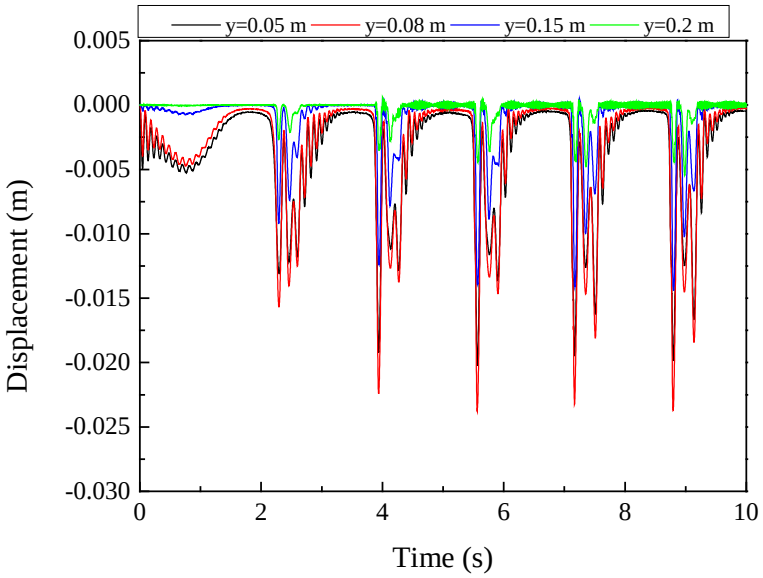


Figure 8. Time history of structural displacement at measuring points B, C, D and E

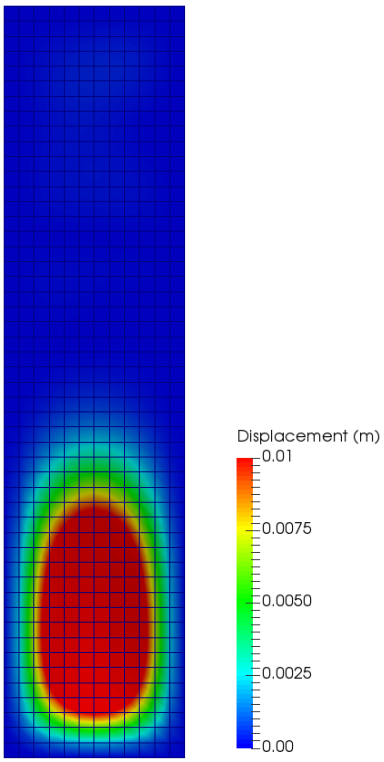


Figure 9. Deformation of elastic wall

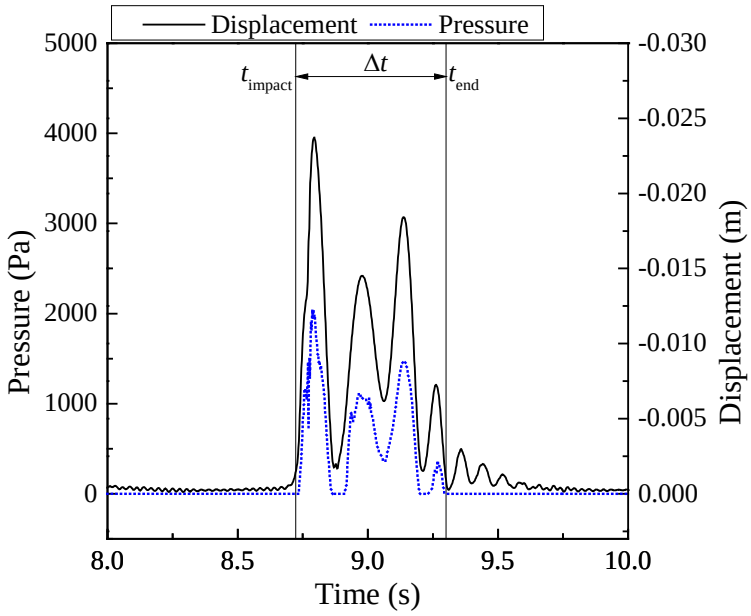


Figure 10. Relationship between structural vibration and impact pressure

Conclusions

In the present study, the in-house solver MLParticle-SJTU based on the MPS-FEM coupled method is developed for 3D FSI problems. The mathematical equations for the MPS and FEM methods are introduced and an interface interpolation approach for data transformation between fluid and structure domains is proposed. With the help of the present FSI solver, the tentative investigation of 3D sloshing problem with the consideration of structural deformation can be successfully conducted. According to the numerical results, the influence of structural elasticity on the sloshing phenomenon can be observed. For instance, the elasticity of tank wall can give rise to the large amplitude oscillation of pressure which is in gear with that of structural vibration during the impact event. The lateral wall deforms in the form of cambered surface while the sloshing wave impacting onto it. The climb height of water front on the elastic wall is much lower than that regarding the rigid tank. The particle distribution of jet water presents the "V" form over the elastic wall while that is homogeneous over the rigid wall. Generally, the study present in this paper shows that the present MPS-FEM coupled method is a promising numerical tool for simulating highly non-linear liquid sloshing in an elastic tank.

Acknowledgement

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Landajuela, M., Vidrascu, M., Chapelle, D. and Fernandez, M. A. (2017) Coupling schemes for the FSI forward prediction challenge: comparative study and validation, *International Journal for Numerical Methods in Biomedical Engineering*, e02813.

[2] Hou, G., Wang, J. and Layton, A. (2012) Numerical methods for fluid-structure interaction-A review, *Commun. Comput. Phys.* **12** (2), 337–377.

[3] Ryzhakov, P. (2017) A modified fractional step method for fluid–structure interaction problems, *Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería* **33**(1), 58–64.

[4] Farahani, M. H., Amanifard, N. and Hosseini, S. M. (2009) A Fluid-Structure Interaction Simulation by Smoothed Particle Hydrodynamics, *Engineering Letters* **17**(1), 30–35.

[5] Monaghan, J. J. (1994) Simulating Free Surface Flows with SPH, *Journal of Computational Physics* **110**(2), 399–406.

[6] Idelsohn, S. R., Onate, E. and Del, P. F. (2004) The Particle Finite Element Method: A Powerful tool to Solve Incompressible Flows with Free-surfaces and Breaking Waves, *International Journal for Numerical Methods in Engineering* **61**(7), 964–989.

[7] Li, J. G., Hamamoto, Y., Liu, Y. and Zhang, X. (2014) Sloshing impact simulation with material point method and its experimental validations, *Computers & Fluids* **103**, 86–99.

[8] Koshizuka, S. and Oka, Y. (1996) Moving particle semi-implicit method for fragmentation of incompressible fluid, *Nuclear Science and Engineering* **123**, 421–434.

[9] Khayyer, A. and Gotoh, H. (2011) Enhancement of stability and accuracy of the moving particle semi-implicit method, *Journal of Computational Physics* **230**, 3093–3118.

[10] Khayyer, A. and Gotoh, H. (2013) Enhancement of performance and stability of MPS mesh-free particle method for multiphase flows characterized by high density ratios, *Journal of Computational Physics* **242**, 211–233.

[11] Khayyer, A., Gotoh, H. and Shimizu, Y. (2017) Comparative study on accuracy and conservation properties of two particle regularization schemes and proposal of an optimized particle shifting scheme in ISPH context, *Journal of Computational Physics* **332**, 236–256.

[12] Zhang, Y. L., Tang, Z. Y. and Wan, D. C. (2016) *MPS-FEM Coupled Method for Interaction between Sloshing Flow and Elastic Structure in Rolling Tanks*, Proceedings of the 7th International Conference on Computational Methods (ICCM2016), Berkeley, USA, 1493-6106-1-PB.

[13] Zhang, Y. L., Chen, X. and Wan, D. C. (2016) MPS-FEM coupled method for the comparison study of liquid sloshing flows interacting with rigid and elastic baffles, *Applied Mathematics and Mechanics* **37**(12), 1359–1377. (In Chinese)

[14] Zhang, Y. L., Chen, X. and Wan, D. C. (2016) *Simulation of Dam-Break Flow Interaction with Elastic Gate by MPS-FEM Coupled Method*, Proceedings of the 11th Asian Computational Fluid Dynamics Conference (ACFD2016), Dalian, China, 390–394.

[15] Mitsume, N., Yoshimura, S., Murotani, K. and Yamada, T. (2014a) MPS-FEM partitioned coupling approach for fluid-structure interaction with free surface flow, *International Journal of Computational Methods* **11**(4), 4157–4160.

[16] Mitsume, N., Yoshimura, S., Murotani, K. and Yamada, T. (2014b) Improved MPS-FE Fluid-Structure Interaction Coupled Method with MPS Polygon Wall Boundary Model, *Comput. Model. Eng. Sci* **101**(4), 229–247.

[17] Liao, K. P. and Hu, C. H. (2013) A coupled FDM-FEM method for free surface flow interaction with thin elastic plate, *J. Mar. Sci. Technol* **18**, 1–11.

[18] Hou, G., Wang, J. and Layton, A. (2012) Numerical Methods for Fluid-Structure Interaction - A Review, *Commun. Comput. Phys* **12**(2), 337–377.

[19] Longatte, E., Verremana, V. and Souli, M. (2009) Time marching for simulation of fluid-structure interaction problems, *Journal of Fluids and Structures* **25**, 95–111.

[20] Sun, Z., Xing, J. T., Djidjeli, K. and Cheng, F. (2015) *Coupling MPS and Modal Superposition Method for Flexible Wedge Dropping Simulation*, Proceedings of the Twenty-fifth International Ocean and Polar Engineering Conference, Kona, Big Island, Hawaii, USA, 144–151.

[21] Onate, E., Idelsohn, S. R., Celigueta, M. A. and Rossi, R. (2006) *Advances in the particle finite element method for fluid-structure interaction problems*, Proceedings of 1st South-East European Conference on Computational Mechanics (SEECCM-06), Kragujevac, Serbia and Montenegro.

[22] Idelsohn, S. R., Marti, J., Limache, A. and Onate, E. (2008) Unified Lagrangian formulation for elastic solids and incompressible fluids: Application to fluid-structure interaction problems via the PFEM, *Comput. Methods Appl. Mech. Eng* **197**, 1762–1776.

[23] Zhang, Y. X., Wan, D. C. and Hino, T. (2014) Comparative study of MPS method and level-set method for sloshing flows, *Journal of hydrodynamics* **26**(4), 577–585.

[24] Zhang, Y. L., Tang, Z. Y. and Wan, D. C. (2016) Numerical Investigations of Waves Interacting with Free Rolling Body by Modified MPS Method, *International Journal of Computational Methods* **13**(4), 1641013.

[25] Tang, Z. Y., Zhang, Y. L. and Wan, D. C. (2016) Multi-resolution MPS method for free surface flows, *International Journal of Computational Methods* **13** (4), 1641018.

[26] Zhang, Y. L. and Wan, D. C. (2017) Numerical study of interactions between waves and free rolling body by IMPS method, Computers and Fluids, https://doi.org/10.1016/j.compfluid.2017.03.019.

[27] Lee, B. H., Park, J. C., Kim, M. H., Jung, S. J., Ryu, M. C. and Kim, Y. S. (2010) Numerical simulation of impact loads using a particle method, *Ocean Engineering* **37**, 164–173.

[28] Newmark, N. M. (1959) A method of computation for structural dynamics, *Journal of the engineering mechanics division* **85**(3), 67–94.

[29] Hsiao, K. M., Lin, J. Y. and Lin, W. Y. (1999) A consistent co-rotational finite element formulation for geometrically nonlinear dynamic analysis of 3-D beams, *Comput. Methods Appl. Mech. Engrg* **169**, 1–18.

Seismic design loads of truss arch frames supported by RC columns with ceilings subjected to vertical and horizontal earthquake motions

†*K. Ishikawa[1] and M. Iso[1]

[1]Department of Architecture and Civil Engineering, University of Fukui, Japan.

*Presenting author: ishikawa@u-fukui.ac.jp
†Corresponding author: ishikawa@u-fukui.ac.jp

Abstract

Spatial structures such as a gymnasium and an exhibition hall often use ceilings because of enhancing sound effects and reducing heating bills. Although the ceiling members fell down on a large scale due to the seismic motion according to the past great earthquake disaster reports, structural engineers particularly do not carry out the seismic design. The study gives structural engineers the equivalent static loads for the design of the earthquake-proof design of the ceiling system. In particular, it is significant to investigate the dynamic behavior and the applied seismic loads for the complicated vibration of the long span arch building structures with RC columns.

Keywords: Truss arch frame, RC column, Knee brace, Vertical and horizontal earthquake motions, Earthquake-proof design, Equivalent static load, Dominant natural mode

Introduction

Long span building structures sometimes set up ceiling systems for insulation to reduce heating bills and soundproofing to use as an auditorium. Although the earthquake-proof design for the non-structural member of the ceiling system is necessary to prevent the fall due to the vibration, the design is not carried out as well as the other non-structural member. It is also known by reports on the past earthquake disasters that ceiling boards of spatial structures fall down due to the dynamic response of the roof. In particular, the ceiling board is absolutely dangerous to human life staying in the building.

The purpose of this study is to investigate the earthquake response such as the acceleration, the velocity, the displacement and the axial force of long span arch truss frames subjected to vertical and horizontal earthquake motions. Based on the dynamic response, the practical calculation method is shown to predict the equivalent static load for the earthquake-proof design of the ceilings as well as the long span truss arch frames. As long as the accuracy verification of the equivalent static load is concerned, the collapse mechanism and the axial stress of the frames subjected to the static loads show a good agreement with the earthquake response analyses subjected to the earthquake motions.
In the practical calculation method, the distribution and the value of the equivalent static load are calculated by means of the participation vector and the earthquake acceleration response spectrum. The static earthquake-proof design is able to carry out using the equivalent seismic loads practically. In general, it is not easy to predict the distribution of the acceleration response because of the complicated dynamic behavior combined with the vertical and horizontal response of the arch shape beam. This is a reason why the earthquake static load is required for the earthquake-proof design using the static analysis.

Long Span Truss Arch Frames with Ceilings supported by RC Columns Subjected to Vertical and Horizontal Earthquake Motions

Seismic Design Coefficient of Ceilings

Seismic design coefficients of arch beams and ceilings are neccesary to calucalate the design seismic load applied at the beam and ceiling for their safety verification.
The ceiling response (E) is shown in Fig.1 is induced by the floor response (D) due to ground motion (B) with amplification characteristics of the surface layer (the layer B) in relation to predominant periods of the layer. The input ground motion (C) for the design is to be defined for the engineering bedrock (the layer A), with the shear wave velocity being about 400m/s or more.
The inertia force F_H must be set up considering the most important factors as follows:

$$F_H=K_HW \tag{1}$$

$$K_H=Z\beta_Hk_HK_0 \tag{2}$$

Where K_H=the seismic design coefficient, W=the weight of the ceiling, Z=the seismic hazard zoning coefficient, k_H=the seismic design coefficient determined by the roof amplification ratio, β_H=the seismic design coefficient determined by ceiling amplification ratio, and K_0=the standard seismic design coefficient.

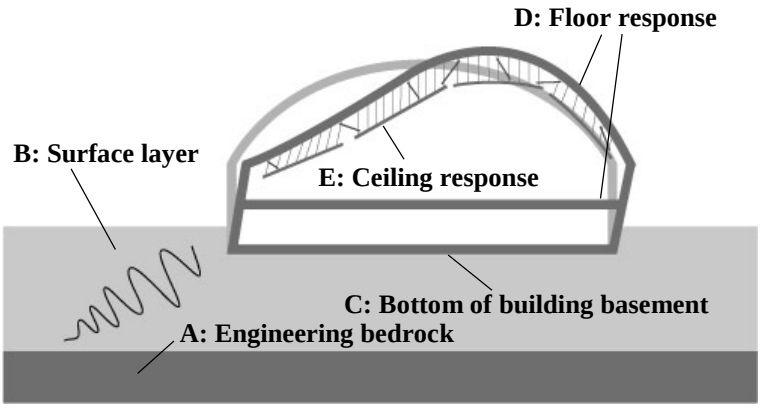


Figure 1. Response evaluation of the arc beam and the ceiling

Analysis Model

The elastic earthquake response analysis is carried out in order to investigate the amplification characteristics at the ceiling board using the analysis model shown in Fig.2. The analysis model deals with the truss arch frame which has around 20m in the span, 7.9 m in the ridge height and 12.19 m in the eaves height. The span-depth ratio α is taken to be 20. The fundamental natural period of the arch structure results in 0.372 second as a result of the eigenvalue analysis. The ceiling has a symmetry system which is divided into 4 parts as shown in Fig.2. The Rayleigh damping is used for the analysis and the damping ratio is taken to be 2%. The Newmark β method is used for solving the vibration equation. As far as a boundary condition is concerned, the bottom of the RC column is set up to be a fix support. The bottoms of two steel truss columns are set up to an anchor bolt and a steel base plate attached by the RC column at the node A and G as shown in Fig.2.

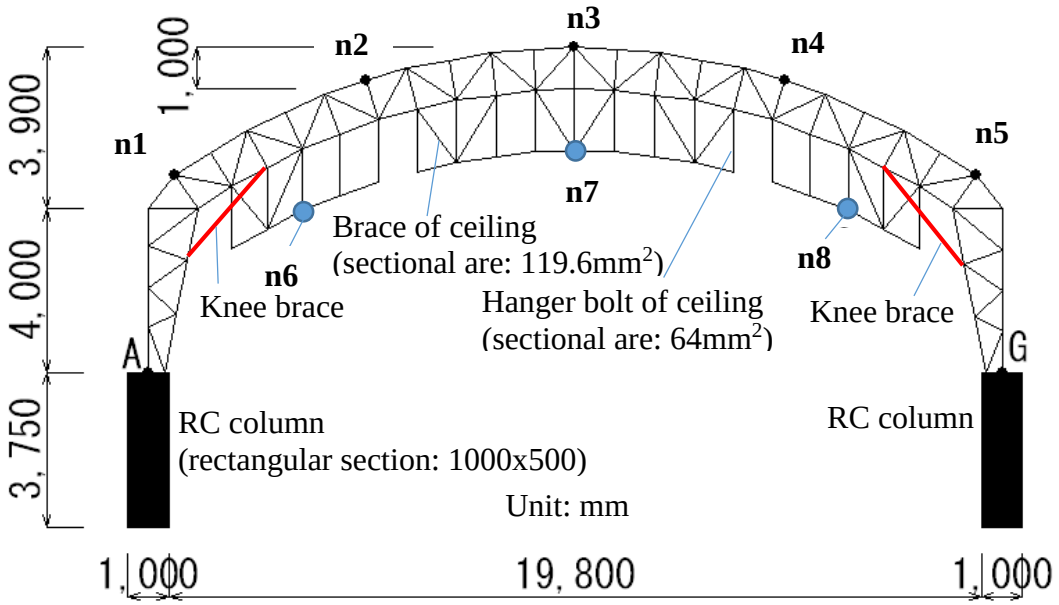


Figure 2. Truss arch frame with ceiling and knee brace

Table 1. Member sectional area (mm^2) of the truss arch frame and the ceilings

arch beam member	upper chord	lower chord	diagonal web	post web	beam and column connection
n1-n2, n4-n5	2724	5952	1745.4	2444	5842
n2-n4	2724	1745.4	1160.4	2444	

column member	outer chord	inner chord	diagonal web	knee brace
left	5842	5842	3800	5842
right	5842	5842	3800	5842

Response of Arch beams and Ceilings Based on The Vibration Characteristics of Long Span Truss Arch Frames supported by RC Columns

Estimation Method of the Maximum Response Acceleration and the Static Equivalent Static Seismic Load applied at the Beams and the Ceilings

In the earthquake-proof design of the truss arch frame and the ceiling, the design seismic load is calculated considering both of the roof and the ceiling amplification with respect to the vibration characteristics of the truss arch frame. On the other hand, the methodology and the calculation standard are not enough to calculate the seismic design load. The practical estimation method is proposed by means of the seismic response spectrum and the participation factor.

The estimation equation of the maximum response acceleration of each node in the truss arch frame and the ceilings is proposed and verified with respect to the accuracy in comparison with the dynamic numerical analysis.

The maximum acceleration of the node "j" in the "i"th vibration mode is given as follows:

$$Acc_{ij} = \beta_i S_A(T_i, h_i) D_{ij} \quad (1)$$

Where Acc_{ij} is the estimated acceleration of the node "j". β_i is the modal participation coefficient of the "i"th vibration mode. $S_A(T_i , h_i)$ is the acceleation response spectrum with respect to the " i" th natural period T_i and damping ratio h_i. And D_{ij} is the "i"th vbration mode value of the node "j" .

The earthquake-proof design of building structures is carried out by means of the static analysis using the seismic load and the load distribution. The static seismic load distribution is also used in the design of the long span truss arch structure. The horizontal distribution of the arch beam is affected by the first vibration mode value and the vertical distribution is affected by the second vibration mode, respectively.

The horizontal static seismic load "F_j^H" at the node "j" are given by using the first modal participation coefficient β_1 and the first mode value D_{1j}^H.

$$F_j^H = m_j \beta_1 S_A^H(T_1 , h_1) D_{1j}^H \tag{2}$$

The horizontal static seismic load "F_j^V" at the node "j" are given by using the second modal participation coefficient β_2 and the second mode value D_{2j}^V.

$$F_j^V = m_j \beta_2 S_A^V(T_2 , h_2) D_{2j}^V \tag{3}$$

The horizontal seismic load of the design static is given by

$$F_j^H = S_j^H C_0^H \sum_{i=1}^{N} w_i \tag{4}$$

Where

$C_o^H = S_A^H(T_1, h_1)/g$, $S_j^H = \frac{m_j \cdot D_{1j}^H}{\sum_{i=1}^{N} m_i}$

m_j is mass of node "j", D_{1j} is the first vibration mode value at node "j", w_i is the weight at node "i", $S_A^H(T_1 , h_1)$ is the acceleation response spectrum (Fig.5) with respect to the first natural period T_1 and damping ratio h_1=0.02 and g is the gravity acceleration.

The vertical seismic load of the design static is given by

$$F_j^V = S_j^V C_0^V \sum_{i=1}^{N} w_i \tag{5}$$

Where

$C_o^V = S_A^V(T_2, h_2)/g$, $S_j^V = \frac{m_j \cdot D_{2j}^V}{\sum_{i=1}^{N} m_i}$

m_j is mass of node "j", D_{2j} is the second vibration mode value at node "j", w_i is the weight at node "i", $S_A^V(T_2 , h_2)$ is the vertical acceleation response spectrum (Fig. 5) with respect to the second natural period T_2 and damping ratio h_2=0.02 and g is the gravity acceleration.

Vibration Characteristics such as Natural Period and Mode of Truss Arch Frame with Ceiling

The two natural periods and the vibration modes of the truss arch frame with the ceiling and the knee brace are obtained by means of the eigenvalue analysis, respectively. The first and the second natural periods such as T_1 and T_2 are shown in Fig.3. The two corresponding vibration modes are also shown in Fig.3, respectively. It is seen that the horizontal vibration shape of an arch beam appears in the first vibration mode. On the other hand, the vertical vibration shape of an arch beam appears in the second vibration mode. The study focuses on the horizontal and

vertical earthquake response of the arch beam and the ceilings subjected to horizontal and vertical earthquake motions.

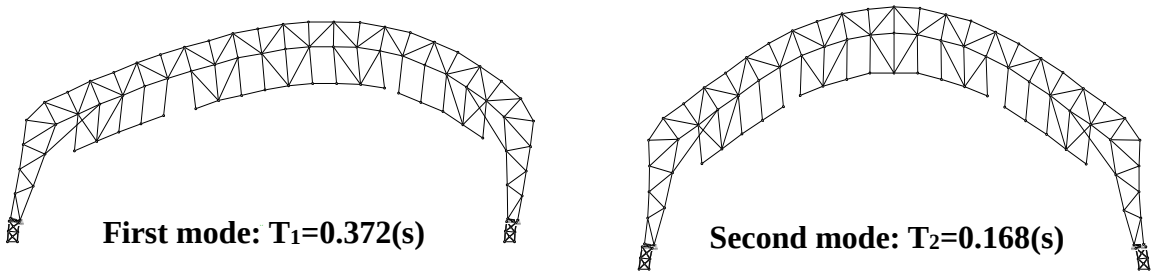


Figure 3. The first and second vibration modes and natural periods T_1 and T_2

Elastic Dynamic Analyses of Truss Frames Subjected to Earthquake Motions

The elastic earthquake response analysis of the truss arch frame with the ceiling and the knee brace is carried out to verify the accuracy of the estimation value of the seismic load applied at the arch beam and the ceiling by means of the proposed method. The numerical integration method uses the Newmarkβ method in the vibration equation. Since it has been known that the case of β=1/4 will be unconditionally stable for most nonlinear problems, β=1/4 is used in this study. The Rayleigh damping is used and both of the first and second damping factors are taken to be 0.02.

Input Vertical and Horizontal Artificial Earthquake Motions on Surface Ground

In conventional earthquake-response analysis, the most common approaches to use waves observed either at the ground surface of a certain location, or at the basement or ground floor of a building as the input ground motion.
The basic thinking behind setting up input ground motion for the seismic design and the analyses is described in this section. There are two basic focal points:
(1) Designing is to basically a two-phase design procedure, Level 1 (moderate earthquake motion) and Level 2(severe earthquake motion), with seismic design carried out for these earthquake inputs.
(2) The input ground motion for design calculates considering the amplification of the surface layer from the basement layer with the shear wave velocity such as about 400m/s or more.
The dynamic analysis of the truss structure with ceiling is carried out using the vertical and horizontal motions. The artificial seismic waves with the phase of the five observed earthquake motions are used. The surface layer amplification is considered in the waves. The motion fits the target acceleration response spectrum in Fig.4 of the damage limit artificial earthquake motion with a phase characteristic (Level 1) and the damping factor 0.05. The peak ground accelerations of the vertical and the horizontal earthquake motions are shown in Table 1 respectively. The average value of the horizontal and vertical motions is 119 and 62 cm/sec^2 respectively.

Table 2. PGA (Peak ground acc.: cm/sec^2) of the used input earthquake motion

Used phase characteristic	PGA	Used phase characteristic	PGA
El Centro-NS (1940)	112	El Centro-UD (1940)	62
Taft-EW (1952)	129	Taft-UD (1952)	57
Hachinohe-NS (1968)	120	Hachinohe-UD (1968)	75
Tohoku-NS (1978)	102	Tohoku-UD (1978)	64
Kobe-NS (1995)	133	Kobe-UD (1995)	52

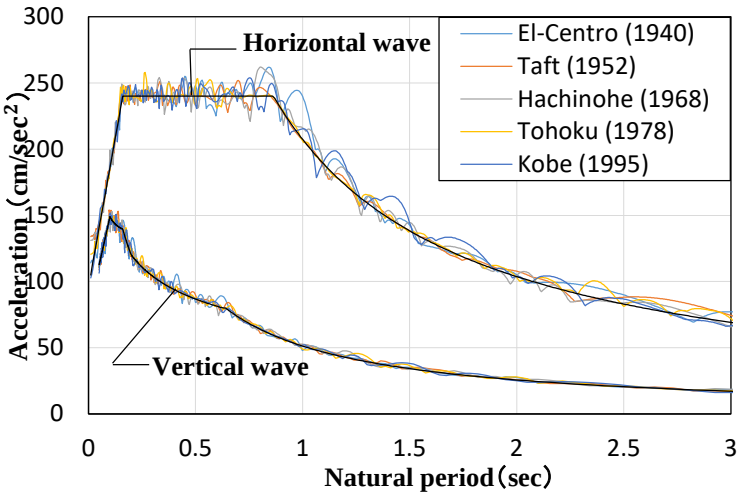


Figure 4. Target acceleration response spectrum (damping factor h=0.05) of the damage limit artificial earthquake motion with a phase characteristic (Level 1) (moderate earthquake motion)

The acceleration response spectrum $S_A(T, h=0.02)$ in case of in case of the damping factor h=0.02 is shown in Fig.5 using the input horizontal and vertical earthquake motions. The spectrum such as $S_A^H(T, h)$ and $S_A^V(T, h)$ are used in the proposed equations (4) and (5).

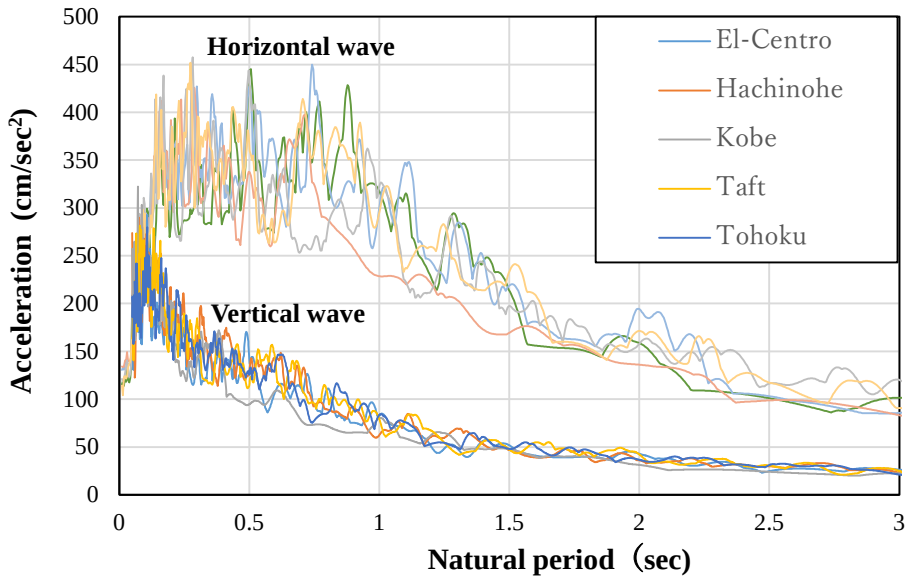


Figure 5. Acceleration response spectrum (damping factor h=0.02) of the damage limit artificial earthquake motion with a phase characteristic (Level 1) (moderate earthquake motion)

Comparison of the Analyses and the Proposed Method

The earthquake response analyses are carried out to verify the accuracy of Eqs. (4) and (6) by using the damage limit earthquake motions with the phase of the observed earthquake motions. It is seen in Figs. 6 and 7 that the estimation values of the horizontal and vertical maximum response acceleration by means of Eqs. (1), (4) and (6) show a good agreement with the time history analyses in the both cases of α=20. It is noted that the estimation values are calculated by using just the first modal participation coefficient and the first vibration mode value.

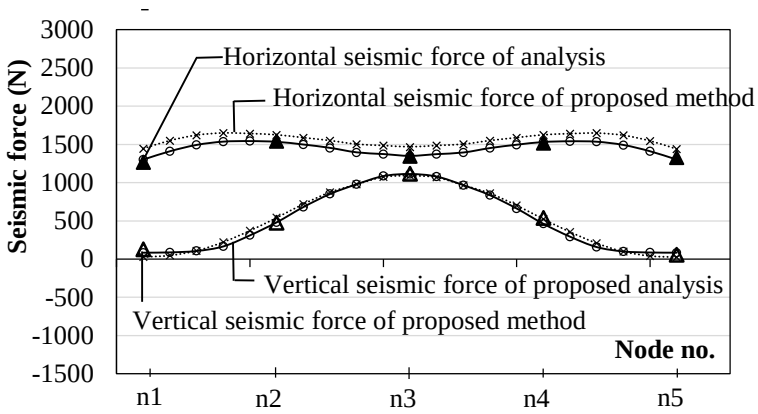


Figure 6. Distribution of equivalent seismic force at the arch truss beam

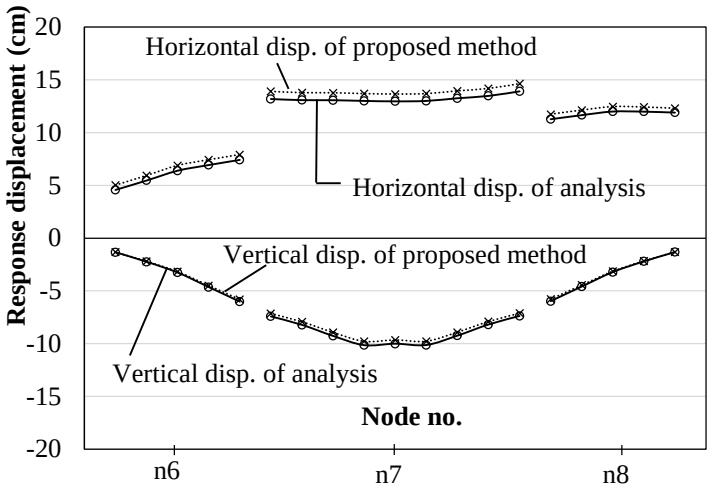


Figure 7. Distribution of seismic force at the ceiling

Conclusions

The characteristic response of the arch beam and the ceiling occurs in the truss arch frame subjected to vertical and horizontal earthquake motions. The design seismic coefficient of the applied load distribution considering the vertical response effect is also actually necessary for the safety study in the earthquake proof design. The study proposes the estimation method of calculating the design seismic coefficient for the applied load for the earthquake-proof design of the long span truss arch frame. The coefficient is easily obtained by means of the eigenvalue analysis. The accuracy is also verified by a good agreement with the earthquake response analysis of the truss arch frames with the ceiling and the knee brace subjected to the damage limit artificial vertical and horizontal earthquake motions. The proposed method of predicting the equivalent static seismic loads can be used to obtain the stress and the deformation for the seismic design.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 15K06290.

References

[1] Ishikawa, K. (2009) Effects of resonance between spatial structures and ceiling systems on the seismic response", Proceedings of the International Association for Shell and Spatial Structures (IASS) Symposium 2009, Valencia Evolution and Trends in Design, Analysis and Construction of Shell and Spatial Structures Universidad Politecnica de Valencia, Spain Alberto DOMINGO and Carlos LAZARO (eds.).

Breaking Wave Simulations of High-speed Surface Combatant using OpenFOAM

Jianhua Wang[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: tjuwjh@126.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

Abstract

Wave breaking on ship bow, shoulder and stern waves is of great importance for the hydrodynamic performance of high-speed surface ships. In the present work, an unsteady VOF based RANS method is used to resolve and investigate wave breaking around a high-speed surface combatant advancing in calm water. High resolution VOF method with artificial compression technique in OpenFOAM is used to accurately resolve ship bow waves as well as induced free surface scars. In order to provide physical understanding of ship breaking waves, simulations over two ship speeds, i.e. Fr=0.35, Fr=0.41, are carried out. The benchmark ship model DTMB 5415 is used for all the simulations and extensive experimental flow data (provided by INSEAN and IIHR) are available to validate the CFD results. All the computations are carried out by in-house CFD solver naoe-FOAM-SJTU. Predicted ship resistance, wave elevations and flow velocities around ship hull are presented and compared with the experimental measurement. Wave breaking on ship bow waves and stern waves are observed. The induced free surface scars are also resolved and its relation with breaking waves is also analyzed. The CFD solution of ship resistance shows good agreement with the experiment. Comparisons of velocity components at cross planes show that the present VOF based RANS method can accurately predict the wake region associated with breaking wave.

Keywords: Breaking wave, naoe-FOAM-SJTU solver, ship hydrodynamics, DTMB 5415

Introduction

Towing ship in calm water is one of the most fundamental studies in the research of ship hydrodynamics. Despite of the high accuracy of the resistance prediction, it is still challenging to accurately resolve the breaking wave phenomenon, which has long been recognized. Extensive experiments have been performed to try to give the physical understanding of the breaking wave mechanism and CFD validation experimental data. Baba[1] identified a new component of viscous resistance from the experimental study and some theoretical approaches. The author noted that large contribution of resistance is due to the wave breaking of ship bow waves. Duncan[2] conducted measurements for the surface profile and the vertical distributions of velocity of a two-dimensional hydrofoil. The measurements were used to resolve the drag on the foil into two parts: one associate with the turbulent breaking region and the other associate with the remaining non-breaking wavetrain. The measurement showed that the drag associate with breaking reach more than three times the maximum drag that could theoretically be obtained with non-breaking waves. Kayo and Takekuma[3] investigated bow wave breaking phenomenon around full ship models by

velocity field measurements and by a flow visualization technique. By use of dye particles placed in front of the bow of full form, the authors find that there exists a shear flow on the free surface. Dong et al.[4] performed experimental study using particle-image-velocimetry (PIV) measurements and free surface visualizations around a ship model at two different speeds, i.e. Fr=0.28 and Fr=0.45. Wave breaking phenomena was observed and the breaking wave associated with vorticity was further discussed. Roth et al.[5] applied the same approach in the experimental study for a 7-meter-long ship model at Froude number of 0.30. Negative vorticity was found at the toe of the wave and positive vorticity appeared on the top of the wave and at the ship boundary. Longo and Stern[6] performed mean velocity measurements using a five-hole Pitot and wave elevation measurements using capacitance wires and point gauges for the static drift condition showing the presence of a bow wave breaking induced vortex on the windward side of the model. Olivieri et al.[7] performed experimental study for the wave breaking of model DTMB 5415, where scars and vortices induced by ship bow and shoulder wave breaking is analyzed. Through measurements of several ship speeds, Fr=0.35 was selected for the intensive study due to the large extents of quasi-steady plunging bow and spilling shoulder wave breaking.

Despite the extensive study through experiment, numerical investigations for the ship breaking waves have also been used as an alternative way to predict and analyze the ship wave breaking phenomena. Wilson et al.[8] applied unsteady single-phase level set RANS method to resolve and investigate bow wave breaking around a surface combatant advancing in calm water, including induced vortices and free surface scars. Good agreement was achieved for both velocity components and axial vorticity at four cross planes and it showed that the CFD approach can accurately predict the detailed flow associate with breaking bow wave. Moraga et al.[9] proposed a sub-grid air entrainment model for breaking bow waves and applied for the simulation of naval surface ship DTMB 5415 and Athena. The model compared favorably with data at laboratory scale and also presented the right trends at full-scale. Marrone et al.[10,11] studied ship wave breaking patterns using both 2D+t and 3D meshless SPH simulations. Through comparison with the experimental data, the proposed schemes were proved to be robust and accurate in simulating ship wave breaking. Landrini et al.[12] presented splashing bow wave simulations using a hybrid BEM-SPH method. A domain-decomposition strategy was adopted which combines two Lagrangian methods, where a potential-flow solution, given by a boundary element method (BEM), follows the jet evolution up to the breaking and then initiates a rotational solution, provided by a smoothed particle hydrodynamics (SPH) technique. Noblesse et al.[13] reviewed the recent results about the overturning and breaking bow wave regimes, and the boundary that divides these two basic flow regimes. Questions and conjectures about the energy of breaking ship bow waves, and free-surface effects on flow circulation, are also noted.

Previous numerical studies on the ship wave breaking are mostly based on the level set approach or the meshless Lagrangian method. In the present work, high resolution Volume of Fluid (VOF) method is used to accurately resolve the large deformation of free surface. The main framework of this paper goes as following. The first part is the numerical methods, where VOF method and numerical schemes are presented. The second part is the geometry model and test conditions. Then comes the simulation part, where wave breaking simulations are present at different Froude numbers. In this part, extensively comparisons are performed against the experimental measurements including ship resistance, wave patterns and wake profiles at longitudinal slices. Finally, a conclusion of this paper is drawn.

Numerical methods

Governing equations

The in-house CFD solver naoe-FOAM-SJTU[14–16], developed on open source platform OpenFOAM, applied in this study solves the Reynolds-Averaged Navier-Stokes equations for unsteady turbulent flows and VOF method is used to capture free surface around the complex geometry models. The URANS equations are written as a mass conservation equation and a momentum conservation equation:

$$\nabla \cdot \boldsymbol{U} = 0 \tag{1}$$

$$\frac{\partial \rho \boldsymbol{U}}{\partial t} + \nabla \cdot \left[(\rho \boldsymbol{U} - \boldsymbol{U}_g)\boldsymbol{U} \right] = -\nabla p_{\mathrm{d}} - \boldsymbol{g} \cdot \boldsymbol{x} \nabla \rho + \nabla \cdot (\mu_{eff} \nabla \boldsymbol{U}) + (\nabla \boldsymbol{U}) \cdot \nabla \mu_{eff} + f_\sigma \tag{2}$$

where $\boldsymbol{U}$ is the fluid velocity field and $\boldsymbol{U}_g$ is the velocity of mesh points; $p_d = p - \rho \boldsymbol{g} \cdot \boldsymbol{x}$ is the dynamic pressure, obtained by subtracting the hydrostatic component from the total pressure; ρ is the mixture density of the two-phase fluid; $\boldsymbol{g}$ is the gravity acceleration; $\mu_{\mathrm{eff}} = \rho(\nu + \nu_t)$ is the effective dynamic viscosity, in which ν and ν_t are the kinematic viscosity and kinematic eddy viscosity respectively, the latter one is obtained by the two-equation shear stress transport turbulence model $SST\ \ k-\omega$ [17]; f_σ is a source term due to surface tension.

VOF method and surface tension

For the wave breaking simulations, the free surface capture method plays an important role in the accuracy of predicted results. In the present work, VOF method with bounded compression technique[18] is applied to capture free surface and the transport equation is expressed as:

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot \left[\left(\boldsymbol{U} - \boldsymbol{U}_{\mathbf{g}} \right) \alpha \right] + \nabla \cdot \left[\boldsymbol{U}_{\boldsymbol{r}} \left(1 - \alpha \right) \alpha \right] = 0 \tag{3}$$

where α is volume of fraction, 0 and 1 represent that the cell is filled with air and water respectively and $0 < \alpha < 1$ stands for the interface between two-phase fluid. $\boldsymbol{U}_{\boldsymbol{r}}$ in Eqn. (3) is the velocity field used to compress the interface and it only takes effect on the free surface due to the term $(1-\alpha)\alpha$.

According to the literature concerning wave breaking, small scale wave breaking is strongly influenced by surface tension. The role played by the surface tension is quite different for breaking and non-breaking waves since the surface tension pressure jump depends on the magnitude of the radius of curvature of the free surface. In order to reappear the wave patterns of the experiment, the surface tension is taken account in the present simulation and the surface tension term mentioned in Eqn. (2) is expressed as:

$$f_\sigma = \sigma \kappa \nabla \alpha \tag{4}$$

where σ stands for the surface tension, κ is the curvature of free surface and it is defined as:

$$\kappa = -\nabla \cdot \boldsymbol{n} = -\frac{\sum_f \boldsymbol{S}_f \cdot \boldsymbol{n}_f}{V_i} \tag{5}$$

V_i represents the volume of cell i, $\sum_f \boldsymbol{S}_f$ stands for the sum of value on each face of cell.

Finite volume method (FVM) with fully unstructured grids is used to transform the RANS and VOF equations from physical space into computational space. The merged PISO-SIMPLE (PIMPLE) algorithm is applied to solve the coupled equations for velocity and pressure field. The Semi-Implicit Method for Pressure-Linked Equations (SIMPLE)

algorithm allows to couple the Navier-Stokes equations with an iterative procedure. And the Pressure Implicit Splitting Operator (PISO) algorithm enables the PIMPLE algorithm to do the pressure-velocity correction. Detailed description for the SIMPLE and PISO algorithm can be found in Ferziger and Peric[19] and Issa[20]. Near wall treatment wall functions are applied to the moving wall boundary. In addition, several built-in numerical schemes in OpenFOAM are used in discretizing and solving the partial differential equations (PDE). The convection terms are discretized by a second-order TVD limited linear scheme, and the viscous terms are approximated by a second-order central difference scheme. A second-order backward scheme is used for temporal discretization except the VOF advection equation, where Euler scheme is adopted.

Geometry model and test conditions

Geometry model

The geometry model of interest is the surface combatant model DTMB 5415 and the numerical simulations are conducted using 5.72 m replica (INSEAN Model 2340). The geometry model of the ship model is shown in Fig. 1, and its principle parameters are listed in Table 1. Extensive experiments have been conducted for the ship model under various Froude numbers and photo study is also available through the experimental measurements[7]. The available experimental data, i.e. ship resistance, wave patterns and flow velocities around ship hull, can be used to validate our present computational results.

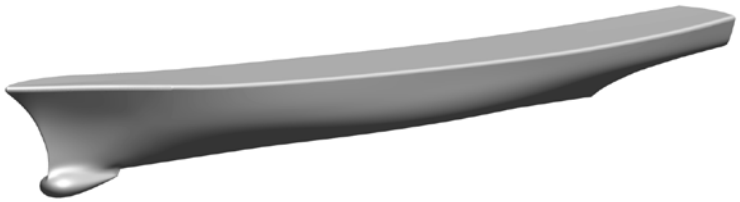


Fig. 1 Geometry model of DTMB 5415 (INSEAN Model 2340)

Table 1 Principle dimensions of DTMB 5415

Main particulars		Model scale	Full scale
Scale factor	λ	24.824	1
Length between perpendiculars	$L_{pp}(m)$	5.719	142.0
Maximum beam of waterline	$B_{WL}(m)$	0.768	19.06
Draft	$T(m)$	0.248	6.15
Displacement volume	$\Delta(m^3)$	0.554	84244
Wetted surface area	$S_0\ (m^2)$	4.786	2972.6
Metacentric height	$GM(m)$	NA	1.95
Moment of inertia	K_{xx}/B_{WL}	NA	0.37
	$K_{yy}/L_{WL}, K_{zz}/L_{WL}$	0.252	0.25

According to the literature concerning wave breaking, the phenomena are strongly connected with turbulence generation, which, in other words, is a viscous phenomenon. Hence, the turbulence parameters should be considered carefully in the numerical simulations. Table 2

summarizes the water quality and physical quantities adopted in the experiments and simulations.

Table 2 Physical quantities in experiment and simulation

Parameters	Symbol	Experiment	Simulation	Full scale
Water density	$\rho\ (kg/m^3)$	998.5*	998.5	1030*
Kinematic viscosity	$\nu\ (m^2/s)$	1.09×10^{-6}	1.09×10^{-6}	1.17×10^{-6}
Surface tension	$\sigma\ (N/m)$	0.0734	0.0734	0.0734
Gravity acceleration	$g\ (m/s^2)$	9.8033*	9.81	9.806*

*Data taken from literature.

Test conditions

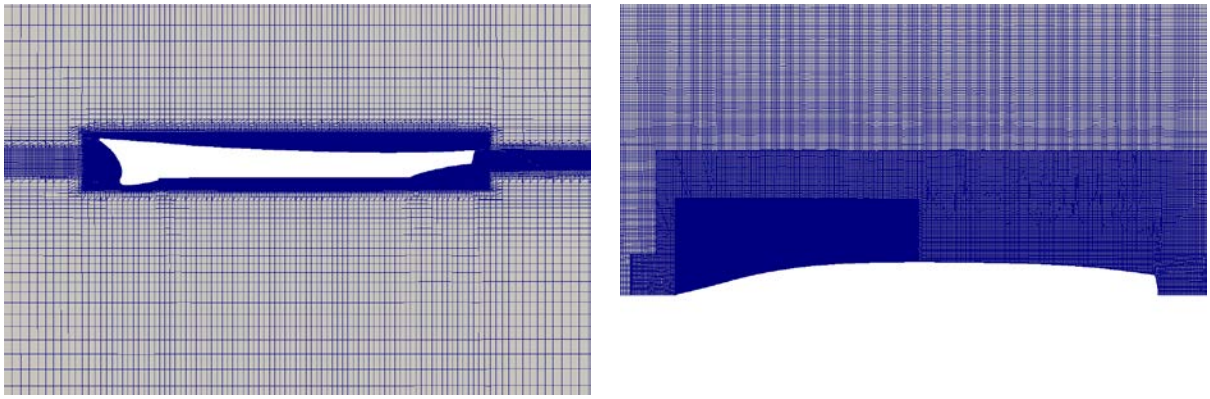
The present work is for the wave breaking simulation DTMB 5415 bare hull model. Two approaching speeds corresponding to Froude number of $Fr = 0.35$ and $Fr = 0.41$, are taken into account to investigate the wave breaking behavior. During the simulations, the model was held fixed with sinkage and trim set to the values previously determined in unrestrained conditions[21]. The simulation conditions are tabulated in Table 3.

Table 3 Simulation conditions for DTMB 5415

Froude number	Speeds (m/s)	Trim(deg)	Sinkage(Lpp)
0.35	2.621	0.069	0.0032
0.41	3.071	-0.421	0.0047

Computational overview

Fully unstructured grids used in this paper are generated by *snappyHexMesh* with the background grid generated by *blockMesh*, both are pre-processing utilities provided by OpenFOAM. In order to accurately capture the breaking wave phenomenon, several refinement regions have been adopted and the grid distribution is shown in Fig. 2. As the ship model is fixed with corresponding trim and sinkage, only half of the flow domain is modeled. The total grid number in the present simulation is 18.7 million.



a) Profile of grid distribution b) Grid arrangement at free surface

Fig. 2 Grid distribution around ship hull

Simulation results and analysis

The computations are carried out on a HPC cluster (IBM nx360M4) in Shanghai Jiao Tong University, which consist of 20 CPUs per node and 64GB accessible memory (Intel Xeon E5-2680v2 @2.8 GHz). 128 processors are assigned to calculate the wave breaking cases under different ship speeds. The time step was set to $\Delta t = 0.001s$, and time to complete the computation was approximately 136 wall clock hours and 17369 CPU hours with about 35000 time steps for the wave breaking simulation.

Fr=0.35

As mentioned in Table 3, the simulation case for Fr=0.35 is under ship speeds of $U = 2.621m/s$ and the ship model is fixed with the initial sinkage of 0.0032Lpp and trim of 0.069deg. Two grids are adopted in the simulation of Fr=0.35 case, one has 18.7 million cells as shown in Fig. 2b and another one has 12 million cells with no refinement in the bow region. Fig. 3 shows the wave pattern with different grids. Significant difference can be observed for the bow waves surrounded by the red box, which indicates that the density of the grid in the bow wave region can strongly affect the bow wave breaking phenomena.

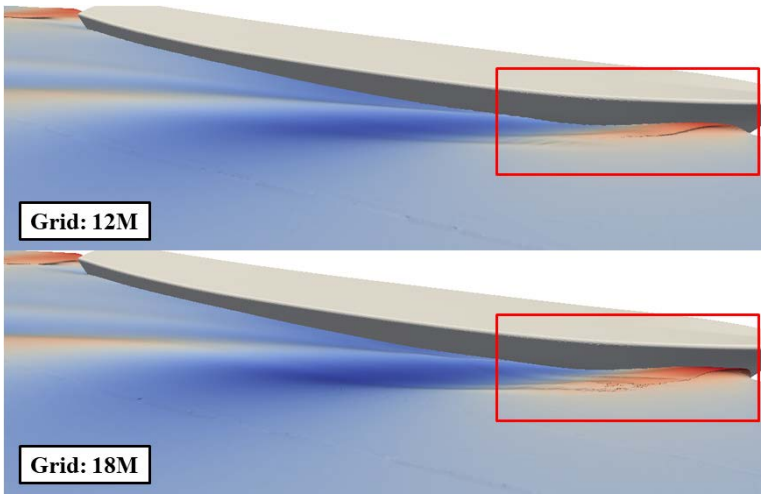
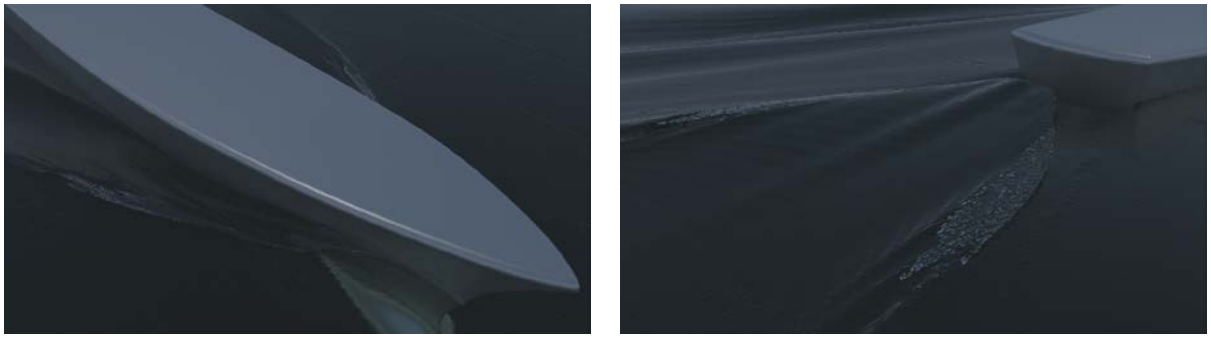


Fig. 3 Wave pattern with different grids

Table 4 Total resistance comparison at Fr=0.35

Cases	Grid number (Million)	Total resistance (N)	Error (%)
Fine grid	18.7	78.33	-2.86
Coarse grid	12.4	78.15	-3.09
Experiment[21]	--	80.64	--

Despite the different wave pattern with fine and coarse grids, the predicted ship resistance agrees very well with the experiment, which indicates that the grid density has little influence on the resistance when grid number is over 10 million for RANS computation. The finer mesh can give promising result of the bow wave pattern according to the experimental measurements, where two scars can be formed due to the evolution of the bow wave. As shown in Fig. 4a, the bow wave breaking phenomena can be obviously observed, and furthermore, the stern wave breaking can also be captured shown in Fig. 4b.



a) Bow wave breaking b) Stern wave breaking

Fig. 4 Simulation results of bow and stern wave breaking at Fr=0.35

The simulation for present Fr=0.35 case and Fr=0.41 talked in the next section adopt the finer grids with about 18.7 million cells. In order to validate the present CFD results, detailed flow velocities at two cross sections are presented and compared with the available experimental measurements[7]. The two cross sections ($x/L_{pp}=0.15$, $x/L_{pp}=0.40$) correspond to the bow and shoulder waves and flow patterns, respectively. Fig. 5 and Fig. 6 illustrate the nondimensionalized mean velocity components (i.e. $(u,v,w)=(U_x,U_y,U_z)/U$) at $x/L_{pp}=0.15$ and $x/L_{pp}=0.40$, respectively. Left figures are the CFD simulation results and right column are the experimental measurements.

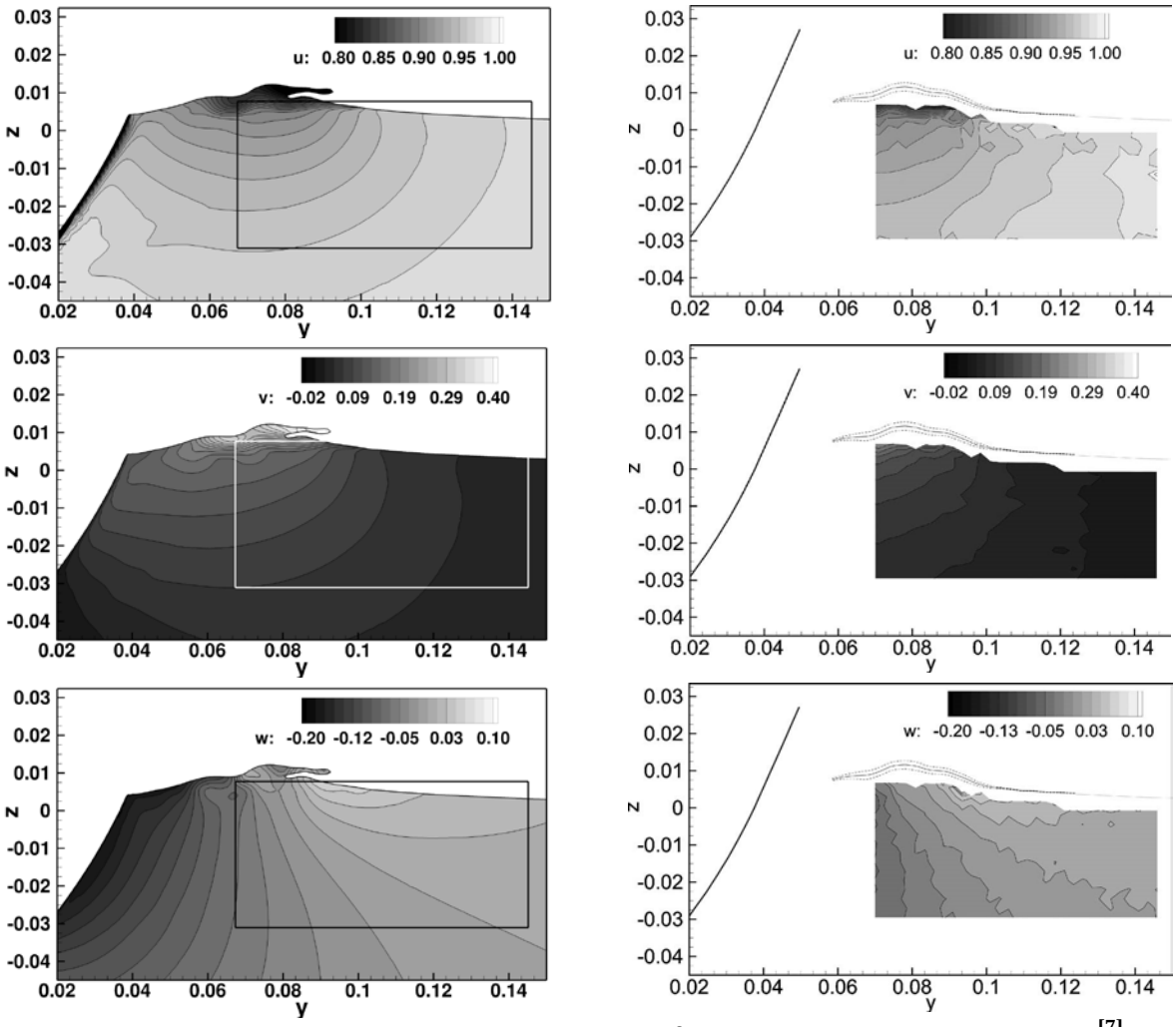


Fig. 5 Velocity components at $x/L_{pp}=0.15$ (left: CFD results, right: experiment[7])

The rectangle box on the CFD figures stands for the area measured experimentally. The agreement between CFD and EFD is satisfactory in both sections for the three-velocity components. It can be obviously seen that at section $x/L_{pp}=0.15$, the bow wave shows a plunging type breaker, while at section $x/L_{pp}=0.40$, the shoulder wave appears as spilling type breaker. The same phenomena have also been observed from the experiment measurement. Apart from the bow wave and shoulder wave that were measured in the experiments, the stern wave breaking shown in Fig. 4b can also be resolved. The wave breaking is formed as plunging type, which is the same with the bow wave breaking. Furthermore, the stern wave breaking can be extended to a larger area according to the present simulation.

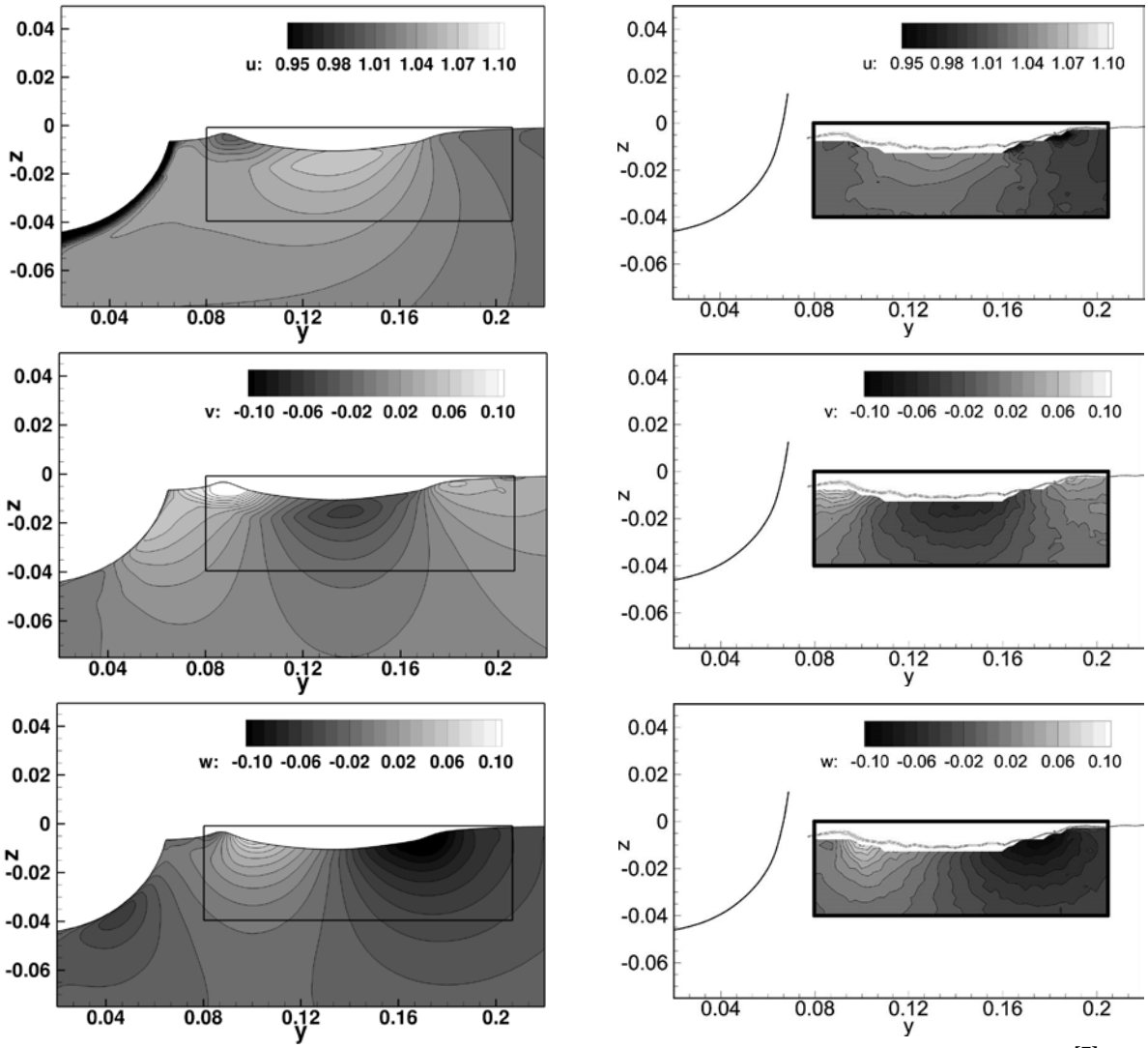


Fig. 6 Velocity components at $x/L_{pp}=0.40$ (left: CFD results, right: experiment[7])

Scars in bow wave breaking of the experiments and present CFD results are shown in Fig. 7. It can be clearly seen with the overturning bow wave and the formation of the scars in the CFD slices of free surface at bow region. Dashed lines show the predicted scars in the bow wave and agree well with the measurements, which show that the URANS approach with VOF method can be robust and reliable in predicting the bow wave breaking of high speed naval ships.

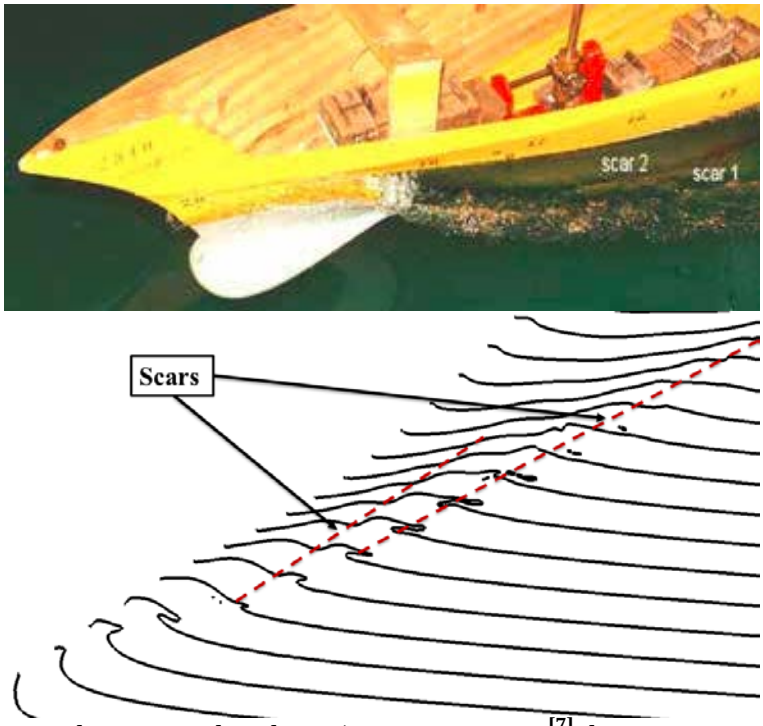


Fig. 7 Scars in bow wave breaking (top: experiment[7], bottom: present result)

Fr=0.41

The higher ship speed with Froude number of 0.41 has also been simulated with the same grids and numerical setup of the Fr=0.35 case. Since there is no available experimental flow data, only total resistance is compared with the measurement. Present CFD prediction of total resistance is just over-estimated by 1.59% compared with the towing tank result[21], which confirms that the CFD simulation is reliable.

Table 5 Total resistance comparison at Fr=0.41

Cases	Grid number (Million)	Total resistance (N)	Error (%)
CFD	18.7	155.02	1.59
Experiment[21]	--	152.59	--

Three velocity components at two cross sections ($x/L_{pp}=0.15$, $x/L_{pp}=0.40$) are shown in Fig. 8. Left column shows the bow wave breaking region and right column is the shoulder wave breaking region. At section $x/L_{pp}=0.15$, the free surface shows more unsteady than that of Fr=0.35. This phenomenon is mainly caused by the ship speeds, where the bow wave formed earlier and the wave height is much larger. The bow wave breaking of ship model appears in the form of a plunging breaker with evidence of the overturning of the wave crest. Air entrainment is considerably increased in this case. At section $x/L_{pp}=0.40$, the shoulder wave also differs a lot with that of Fr=0.35 case, where the shoulder wave is merging with the hull. Furthermore, the bow wave can also be found and it forms a spilling type wave breaker as shown in the shoulder wave section.

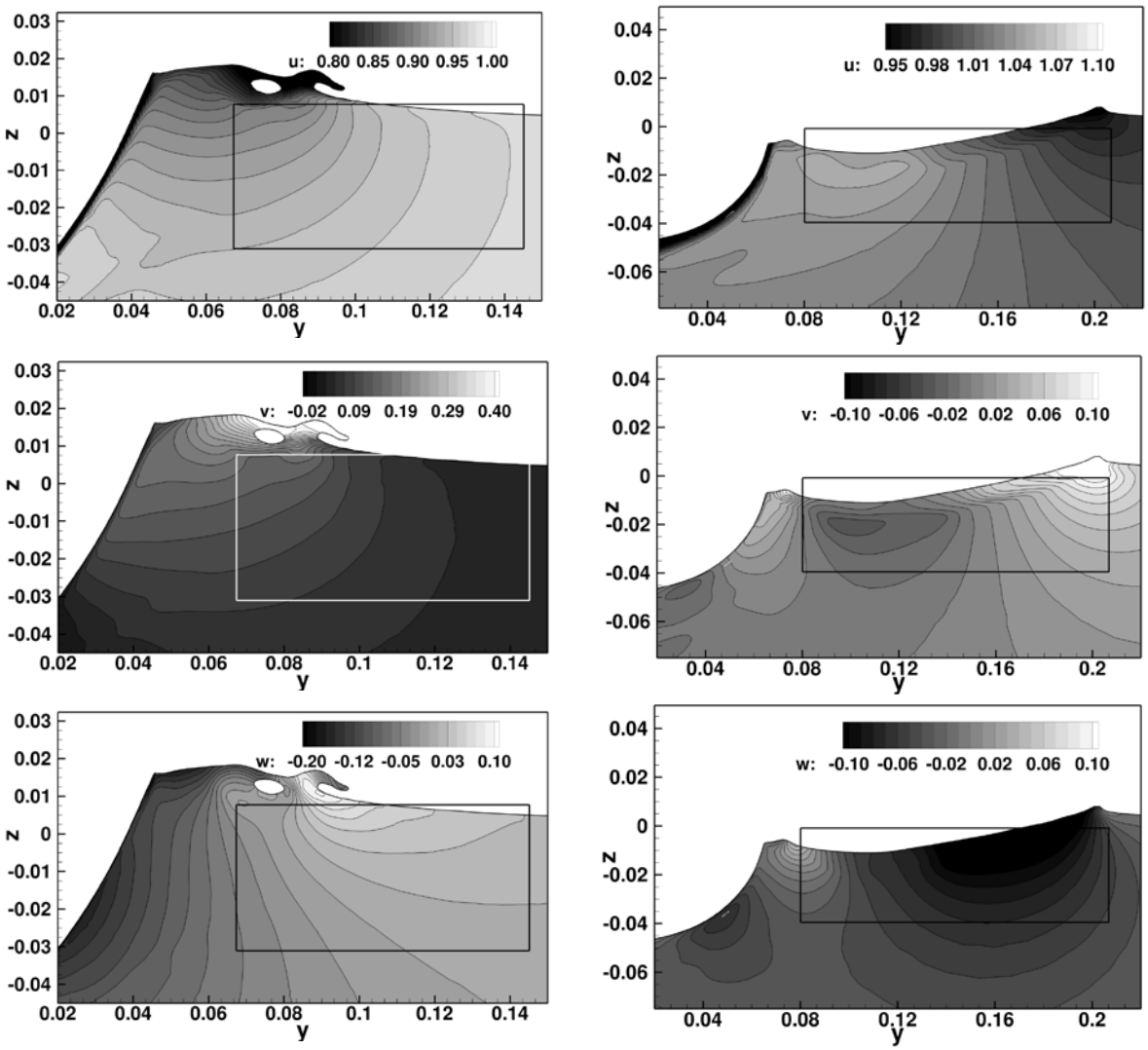
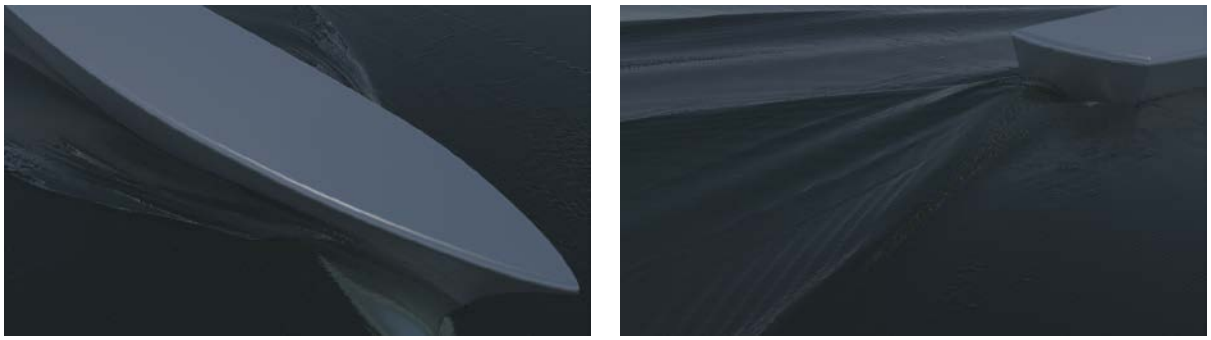


Fig. 8 Velocity components (left: $x / L_{pp} = 0.15$, right: $x / L_{pp} = 0.40$)

Bow wave evolution can be better recognized in Fig. 9a and Fig. 10. The unsteady plunging breaker extends to a larger area compared with Fig. 4a and the scars is also more obvious. Fig. 9b shows the stern wave at Fr=0.41 and from the figure we can see that the wave pattern differs a lot with Fig. 4b. A steady state is presented and the stern wave crest is higher than that at Fr=0.35.



a) Bow wave breaking b) Stern wave breaking

Fig. 9 Simulation results of bow and stern wave breaking at Fr=0.41

Scars at bow wave breaking are illustrated in Fig. 10. There are two obvious overturning of the bow wave, thus result in severe air entrainment. Different from the Fr=0.35 case, the overturning bow wave appears earlier and the air entrainment extends to a larger range in longitudinal direction.

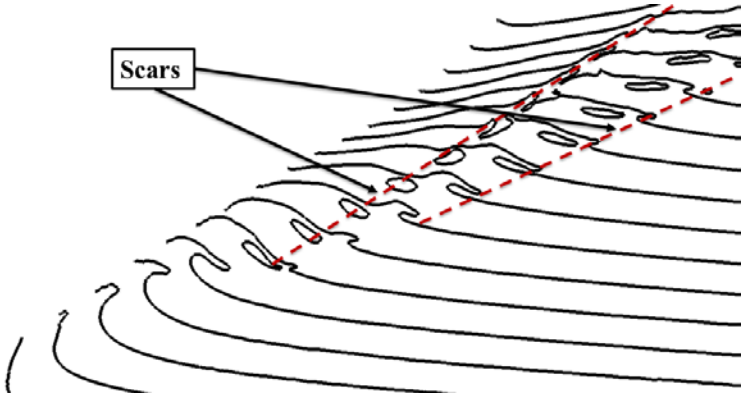


Fig. 10 Scars in bow wave breaking at Fr=0.41

Conclusions

This paper presents the breaking wave simulations of surface combatant DTMB model 5415 at two different speeds, i.e. Fr=0.35, Fr=0.41. All the simulations are performed using in-house CFD solver naoe-FOAM-SJTU. During the simulation, high resolution VOF method with artificial compression technique in OpenFOAM is used to accurately resolve ship bow waves as well as induced free surface scars.

Predicted ship resistance agrees very well with the towing tank measurement. Two grids are adopted for the simulation of wave breaking at Fr=0.35. Both grids give promising results of ship resistance, while the finer grid can better resolve the bow breaking phenomenon. Three velocity components at two cross sections, $x/L_{pp}=0.15$ and $x/L_{pp}=0.40$ corresponding to the bow wave and shoulder wave region, are compared with the available experiment measurements. Good agreement is achieved and bow wave evolution has also been presented and compared with the photo taken using high speed cameras. Plunging breaker for the bow wave and spilling type for the shoulder wave is captured by both CFD and experiment measurement. Furthermore, the CFD simulation also gives the stern wave at Fr=0.35, which shows unsteady plunging type wave breaking. Fr=0.41 shows much difference for the wave patterns with the Fr=0.35 case. Air entrainment can be obviously observed and the bow wave breaking is more unsteady. Shoulder wave is merging to the hull and the bow wave can affect a larger area in the longitudinal direction. Two overturning waves can be captured and thus result in the severe air entrainment.

Future work will focus on the detailed analysis with the vortices associate with the breaking waves. More work will be done to do the wave breaking simulations for the container ship KCS under high speeds, which will be presented at the 2020 CFD workshop in ship hydrodynamics.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai

Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Baba, E. (1969) A New Component of Viscous Resistance of Ships, *Journal of the Society of Naval Architects of Japan* **125**, 23–34.
[2] Duncan, J.H. (1983) The breaking and non-breaking wave resistance of a two-dimensional hydrofoil, *Journal of Fluid Mechanics* **126**, 507–520.
[3] Kayo, Y., and Takekuma, K. (1981) On the Free-Surface Shear Flow related to Bow Wave-Breaking of Full Ship Models, *Journal of the Society of Naval Architects of Japan* **149**, 11–20.
[4] Dong, R.R., Katz, J., and Huang, T.T. (1997) On the structure of bow waves on a ship model, *Journal of Fluid Mechanics* **346**, 77–115.
[5] Roth, G.I., Mascenik, D.T., and Katz, J. (1999) Measurements of the flow structure and turbulence within a ship bow wave, *Physics of Fluids* **11**(11), 3512–3523.
[6] Longo, J., and Stern, F. (2002) Effects of drift angle on model ship flow, *Experiments in Fluids* **32**(5), 558–569.
[7] Olivieri, A., Pistani, F., Wilson, R., Campana, E.F., and Stern, F. (2007) Scars and Vortices Induced by Ship Bow and Shoulder Wave Breaking, *Journal of Fluids Engineering* **129**(11), 1445–1459.
[8] Wilson, R.V., Carrica, P.M., and Stern, F. (2006) URANS simulations for a high-speed transom stern ship with breaking waves, *International Journal of Computational Fluid Dynamics* **20**(2), 105–125.
[9] Moraga, F.J., Carrica, P.M., Drew, D.A., and Lahey, R.T. (2008) A sub-grid air entrainment model for breaking bow waves and naval surface ships, *Computers & Fluids* **37**(3), 281–298.
[10] Marrone, S., Colagrossi, A., Antuono, M., Lugni, C., and Tulin, M.P. (2011) A 2D+t SPH model to study the breaking wave pattern generated by fast ships, *Journal of Fluids and Structures* **27**(8), 1199–1215.
[11] Marrone, S., Bouscasse, B., Colagrossi, A., and Antuono, M. (2012) Study of ship wave breaking patterns using 3D parallel SPH simulations, *Computers & Fluids* **69**, 54–66.
[12] Landrini, M., Colagrossi, A., Greco, M., and Tulin, M.P. (2012) The fluid mechanics of splashing bow waves on ships: A hybrid BEM–SPH analysis, *Ocean Engineering* **53**, 111–127.
[13] Noblesse, F., Delhommeau, G., Liu, H., Wan, D., and Yang, C. (2013) Ship bow waves, *Journal of Hydrodynamics, Ser B* **25**(4), 491–501.
[14] Shen, Z., Cao, H., Ye, H., and Wan, D. (2012) The manual of CFD solver for ship and ocean engineering flows: naoe-FOAM-SJTU, Shanghai Jiao Tong University.
[15] Cao, H., and Wan, D. (2014) Development of Multidirectional Nonlinear Numerical Wave Tank by naoe-FOAM-SJTU Solver, *International Journal of Ocean System Engineering* **4**(1), 52–59.
[16] Shen, Z., Wan, D., and Carrica, P.M. (2015) Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering, *Ocean Engineering* **108**, 287–306.
[17] Menter, F.R., Kuntz, M., and Langtry, R. (2003) Ten years of industrial experience with the SST turbulence model, *Turbulence, Heat and Mass Transfer* **4**(1), 625–632.
[18] Weller, H.G. (2008) A new approach to VOF-based interface capturing methods for incompressible and compressible flow, OpenCFD Ltd, Report TR/HGW/04.
[19] Ferziger, J.H., and Peric, M. (2012) Computational methods for fluid dynamics, Springer Science & Business Media.
[20] Issa, R.I. (1986) Solution of the implicitly discretised fluid flow equations by operator-splitting, *Journal of Computational Physics* **62**(1), 40–65.
[21] Olivieri, A., Pistani, F., Avanzini, A., Stern, F., and Penna, R. (2001) Towing tank experiments of resistance, sinkage and trim, boundary layer, wake, and free surface flow around a naval combatant INSEAN 2340 model, IIHR Report No 421.

Backlash Computation of Harmonic Drive Based on Parametric Solid Finite Element Model

***P.P. Yang[1,2], †X.X. Chen[1,2], J.Z. Xing[1,2], and Y.P. Yao[1,2]**

[1] School of Mechanical Engineering, Tianjin Polytechnic University, China
[2] Tianjin Key Laboratory of Modern Mechatronics Equipment Technology Tianjin, China

*Presenting author: 2462619832@qq.com
†Corresponding author: chenxx@tjpu.edu.cn

Abstract

In order to truly investigate the backlash distribution between tooth profile of flexspline (FS) and circular spline (CS), a finite element model (FEM) of harmonic drive (HD) based on solid element is established by APDL, and backlashes are calculated with the minimum circumferential distance between engaged tooth profiles along tooth height direction. Taking involute tooth profile parameters and structural parameters of FS and CS into consideration, the FEM of planar tooth ring with involute teeth under action of four-rollers wave generator is calculated with contact analysis, and assembly status is obtained. By adjusting the radial position of the roller until the maximum radial displacement on the major axis of the deformed FS reaches the specified value. Backlashes are calculated for all engaged tooth profiles and the backlash distribution curve is obtained. The research indicates that the location of backlash alternately occurs between the addendum of FS and the addendum of CS, and their backlash curves go across twice in the engagement interval. The validity of the backlash computation based on the FEM is higher by the contrastive analysis, which provides a reliable way of backlash computation.

Keywords: Finite Element Model, Parametric Modeling, Harmonic Drive, Backlash Computation

Introduction

HD is a novel gear transmission technology which has been achieved great development in recent decades (Xie, 1979) [1]. It relies on the continuous movement of the deformation wave that caused by wave generator (WG) forcing FS deformation to achieve the movement and power transmission (Ivanov, 1987) [2]. Benefit from its large transmission ratio, high transmission accuracy, small return difference, high transmission efficiency, it is widely used in aerospace, medical equipment, multi-joint robot and other related fields (Shen, 1985) [3].

Ivanov (1987) [2] and Shen (1985) [3] have given the relatively complete research theory in establishing mathematical model of elastic deformation of FS and solving the backlash of the tooth profile. Lu et.al (2009) [4] completed simulation of the whole and local engagement of the harmonic transmission in Matlab, and conducted interference check. Pan (2010) [5] simulated the engagement interval temperature of HD by using ANSYS, and analyzed the variation law of the minimum backlash with the ambient temperature. Yin (2010) [6] made the parameter optimization design of the zero-backlash involute HD and finite element analysis. Taking zero-backlash as the objective function, the optimization of the engagement parameters was completed. Zhang (2012) [7] established the automatic optimization algorithm of tooth profile interference to calculate the minimum backlash, and realized visual simulation of the tooth profile fitting process by using math software. Chen et.al (2011, 2014)[8-9] put

forward the backlash computation method of double-circular-arc tooth profile of HD, and checked interference. Liu et.al (2015) [10] made a 3D tooth profile design based on the cone deformation of FS cylinder, and solved the backlash distribution of the 3D tooth profile. Wang et.al (2017) [11] solved and fitted the conjugate tooth shape of the involute FS, and studied the backlash distribution along tooth height direction from the theoretical point of view.

These theoretical backlash computation methods are based on the assumption of small deformation and that teeth are no deformation. For backlash of tens microns, the experimental measurement is very difficult. In this paper, a parametric FEM of HD is developed. The reasonable assembly status is obtained by iteration, and the iteration can not be stopped until the maximum radial displacement on the major axis of the deformed FS reaches the specified value. Through the definition, extraction and computation of the relevant parameters in the model, the backlashes between engaged tooth profiles of FS and CS are calculated, and the backlash distribution for all engaged teeth is studied.

1 Parameter Analysis

HD has three basic components: FS, CS and WG. First of all, its initial parameters are defined, as shown in Table 1. Then its profile parameters and structural parameters are analyzed.

Table 1. Model parameters

Basic components	FS	CS	WG
Module	m	m	
Number of teeth	z_1	z_2	
Modification coefficient	x_1	x_2	
Roller location angle			β
Contact interference			η
Wall thickness	δ		
Maximum radial displacement	w_0		

(1) The parametric equation of the FS tooth profile is defined as (Shen, 1985) [3]:

$$\begin{cases} x_1 = r_1\left[-\sin\left(u_{a1}-\theta_1\right)+u_{a1}\cos\alpha_0\cos\left(u_{a1}-\theta_1+\alpha_0\right)\right] \\ y_1 = r_1\left[\cos\left(u_{a1}-\theta_1\right)+u_{a1}\cos\alpha_0\sin\left(u_{a1}-\theta_1+\alpha_0\right)\right] \end{cases} \tag{1}$$

Here, u_{a1} is the coordinate parameter of the involute FS, and θ_1 is half of the central angle corresponding to the thickness on the reference circle of FS tooth, and α_0 is the pressure angle, and r_1 is reference circle radius of FS. Among them, u_{a1} are determined by the module, the number of teeth, the pressure angle and the modification coefficient of FS. θ_1 is determined by the module and the modification coefficient of FS. α_0 is the basic constant input. r_1 is determined by the module and the number of teeth.

(2) The parametric equation of the CS tooth profile is defined as (Shen, 1985) [3]:

$$\begin{cases} x_2 = r_2\left[-\sin\left(u_{M2}-\theta_2\right)+u_{M2}\cos\alpha_0\cos\left(u_{M2}-\theta_2+\alpha_0\right)\right] \\ y_2 = r_2\left[\cos\left(u_{M2}-\theta_2\right)+u_{M2}\cos\alpha_0\sin\left(u_{M2}-\theta_2+\alpha_0\right)\right] \end{cases} \tag{2}$$

Here, u_{M2} is the coordinate parameter of the involute CS, and θ_2 is half of the central angle corresponding to the thickness of the reference circle of the CS tooth, and r_2 is reference circle

radius of CS. Among them, u_{M2} is determined by the module, the number of teeth, the pressure angle and the modification coefficient of CS. θ_2 is determined by the module and the modification coefficient of CS. r_2 is determined by the module and the number of teeth.

(3) The structural parameters of FS and the four-rollers wave generator are designated. β is the roller location angle, η is radial displacement of roller (defined as contact interference). δ is the wall thickness, w_0 is maximum radial displacement.

2 Modeling Method

It is considered that HD structure is quite complex, and the design parameters are quite much. So the method of FEM is developed (Wei, 2015) [12], as shown in Fig. 1.

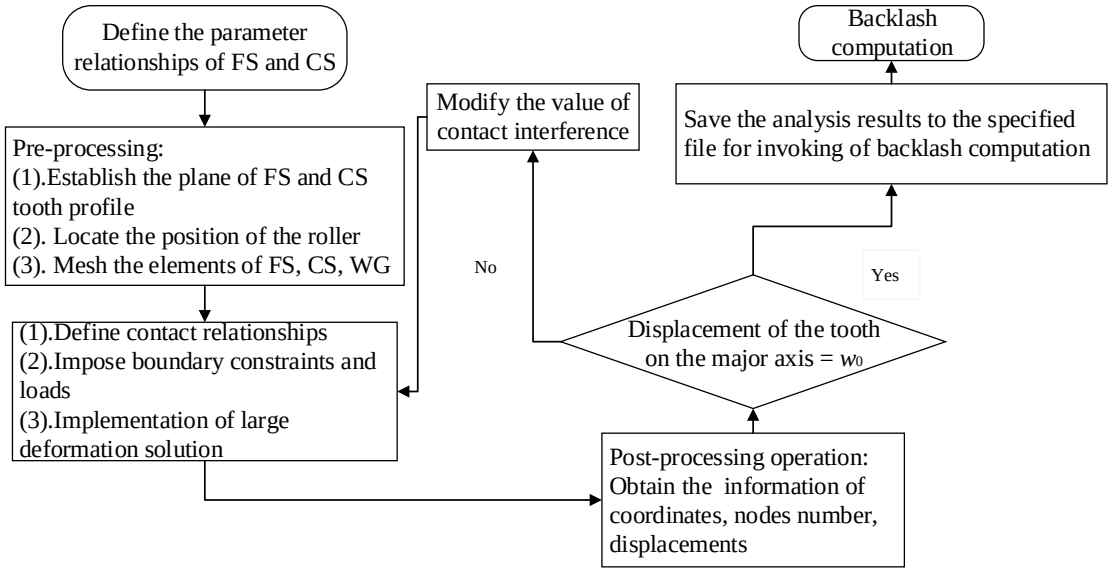


Figure 1. The flow chart of modeling method

Fig. 1 is a complete parametric analysis flow chart of modeling method, including pre-processing, solving, post-processing and iteration. The criterion of iteration is whether the tooth on the major axis of the deformed FS reaches the specified radial displacement. In consideration of the particularity and complexity of harmonic drive, the model is simplified as follows:

(1). Since the length of the ring-shaped FS cylinder is short, it is considered that the deformation along the tooth length is uniform, and it can be studied as a planar engagement problem.

(2). Considering the symmetry of structure and load of the contact model between FS and WG in the assembly status, the 1/4 model of FS and WG is analyzed.

2.1 Establishment of FEM

The establishment of FEM is as follows: (1). The structural parameters and tooth profile parameters of HD are defined before the pre-processing. (2). The bottom-up modeling method is selected. Sex keypoints of tooth profile will be selected, and its coordinates are solved by using the parametric equations (1-2). These keypoints are connected to lines, and the lines constitute a surface, and the surface is mirrored to form a tooth. The fillet is made at the root of tooth, and its radius is half of the tooth space. A tooth is arrayed into a 1/4 model of FS and CS. The coordinates of WG are determined based on the roller location angle and the contact

interference. The element type of FS, CS and WG are chose as plane183. The elastic modulus is 210 GPa, and poisson's ratio is 0.3. (3). In order to analyze deformation of FS, FS is divided by the mapping mesh, which makes the grids fine and regular. CS and WG are divided by the rough free grid to reduce the computation time.

2.2 Solution of FEM

The solution of FEM is as follows: (1). The contact relationship between the FS and the WG is defined as flexible body and rigid body by the contact element. The upper semicircle of WG is defined by the Targe169 target element. The inner wall of FS is used as a flexible contact surface and is defined by the Conta172 contact element. When the contact interference is assigned, it is necessary to obtain the exact value by iterations. When the tooth on the major axis of the deformed FS reaches the specified radial displacement w_0, the iteration stopped. (2). The major axis and minor axis of the deformed FS are subjected to the symmetrical constraint, and the outer surface of WG is subjected to the fixed confinement. (3). In the process of solution, the ratio of the maximum radial deformation and the wall thickness of FS is greater than 0.2, so the large deformation solution is used. (4). During the post-processing phase, the displacement cloud is displayed and the required data is stored in the specified file (Zeng, 2010) [13].

2.3 Case study

The main parameters of HD for electronic equipment: m =0.2, z_1 =140, z_2 =142, x_1 =2.13, x_2 =1.925, η (Eta) =0.112, δ (Delta) =0.3, w_0 =0.2, β (Beta) =30 ° (Shen, 1985) [3]. In order to simplify the work of modifying the command stream file, the customization of the parameter input interface is realized. (As shown in Fig. 2)

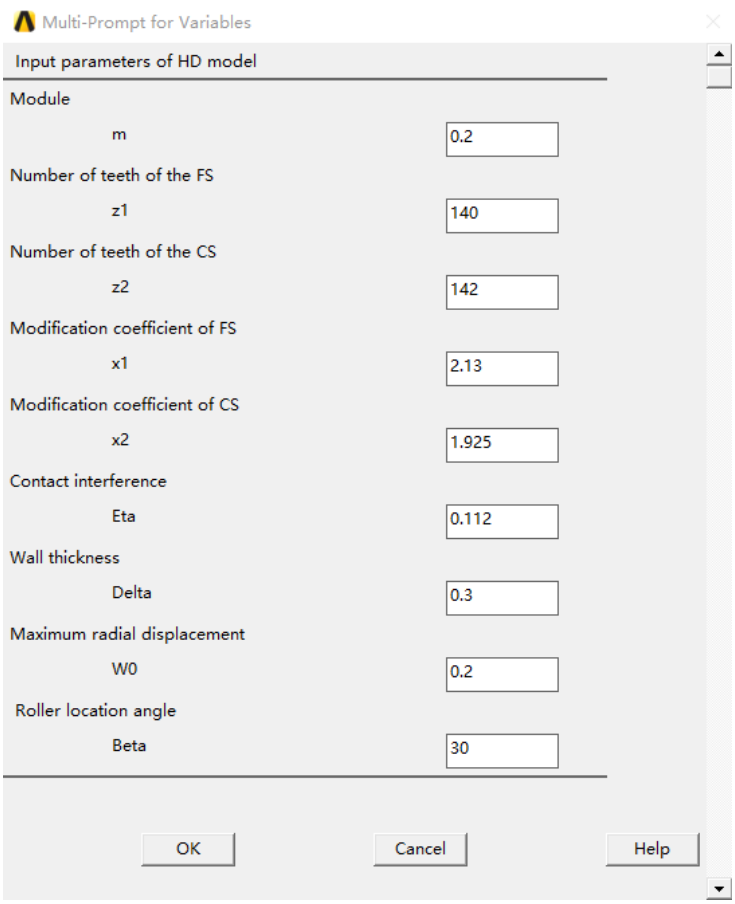


Figure 2. Parameter input interface

Fig. 3 is the FEM of undeformed FS. The nodes number of FS, CS and WG are 30168, 223 and 2808, respectively. And the percentage of the nodes number of FS in the total number is 90.87%. The elements number of FS, CS and WG are 8640, 66 and 720, respectively. And the percentage of the elements number of FS in the total number is 91.66%. Fig. 3 shows that contact relationship is defined between the FS and the four-rollers WG. When the four-rollers WG forces FS to occur deformation, the radial displacement of FEM is shown at the Fig. 4. It shows that the teeth of FS and CS are completely engaged on major axis, and are completely disengaged on minor axis, and are transient state in other areas. And it shows that radial displacement is about 0.2mm on major axis, and radial displacement is about 0.22543mm on minor axis.

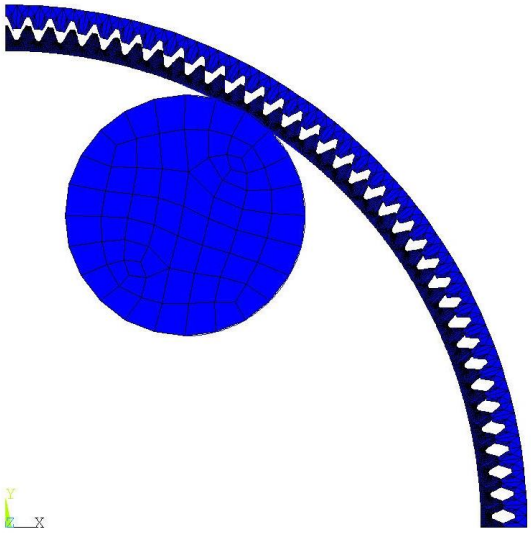


Figure 3. The FEM of the undeformed FS

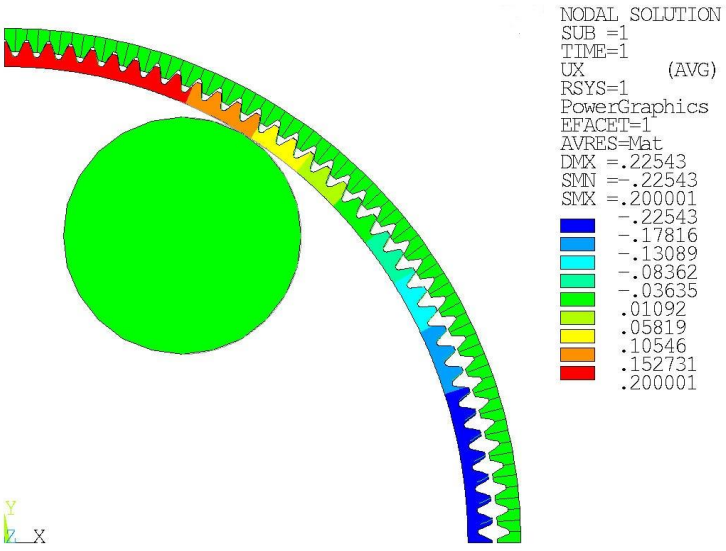


Figure 4. The radial displacement of FEM

3 Computation of backlash along tooth height direction

3.1 Backlash computation process

Most of the studies have considered that the addendum of FS is the position of the minimum distance along tooth height direction, and the backlash is defined as the circumferential distance of the addendum of FS along tooth height direction. But there are not detailed theoretical explanation and experimental data verification for the backlash distribution. In view of the above problems, this paper uses the FEM to calculate the 6 points in the FS tooth profile along tooth height direction and the addendum of CS. The FS tooth profile along tooth height direction were divided into 5 equal parts for points 1-6 (point 1 for the addendum of FS, point 6 for the dedendum of FS), point 7 for the addendum of CS, as shown in Fig. 5.

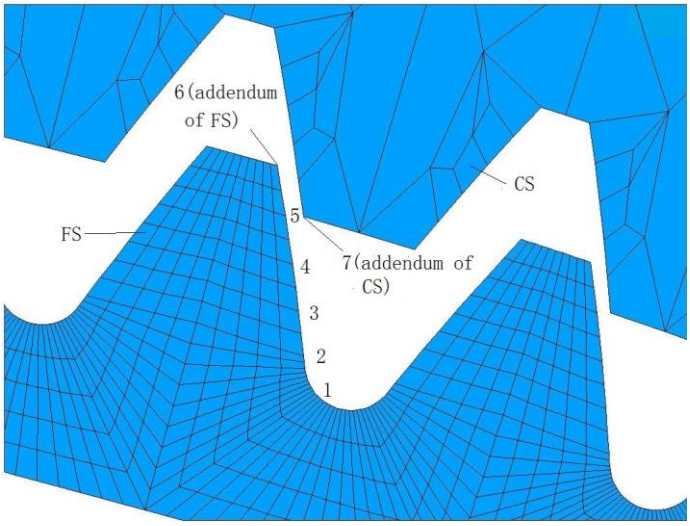


Figure 5. Local engagement status of FS and CS

The computation process of backlash is as follows in the FEM. The computation process of the addendum backlash of FS is described as an example, and the computation process of other points are similar to the addendum of FS.

(1). Definition of parameters. Because the teeth on the cylinder body of FS are reproducible and repeatable, the coordinates, the displacement and the nodes number are defined by array parameter, and cycle statement is used to assign the values. The 1/4 model has a total teeth of $N_Z=z_1/4$, so the cycle number is N_Z. The nodes number is defined as a one-dimensional array of *NZ* rows and one column. The coordinates (x, y) and displacements (x, y) of the nodes are defined as two-dimensional arrays of N_Z rows and two columns. The polar angle, the polar radius of FS nodes and the polar angle of CS nodes are defined as the three-dimensional array of N_Z rows and three columns.

(2). The information such as the coordinates, the displacements, the polar radius, and the polar angle of the addendum of FS should be solved. The coordinates (x, y), the displacements (x, y), the polar radius, and the polar angle of the addendum of FS are assigned to the array parameters by using cycle statement. The above array parameters are saved to the specified path file.

(3). The information such as the polar angle and the coordinates of the engagement point of the addendum of FS in CS involute tooth profile (this point is defined MAFC) should be solved. According to the condition that the polar radius of the addendum of FS equal to the

MAFC's, the coordinates of the MAFC are solved, and the polar angle of the MAFC is calculated. The above data is assigned to the array parameter by using cycle statement and is saved to the specified file.

(4). The circumferential distance between the addendum of FS and the MAFC is solved. The first step is to determine whether the FS teeth and the CS teeth are in the engagement interval. In this paper, we need to determine whether the polar radius of the addendum of FS is bigger than the polar radius of the addendum of CS. If the result is true, it is in the engagement interval. Then the difference between the polar angle of the addendum of FS and the MAFC's is multiplied with the polar radius of the addendum of FS, whose result is the circumferential backlash of the addendum of FS.

3.2 Analysis of the backlash distribution

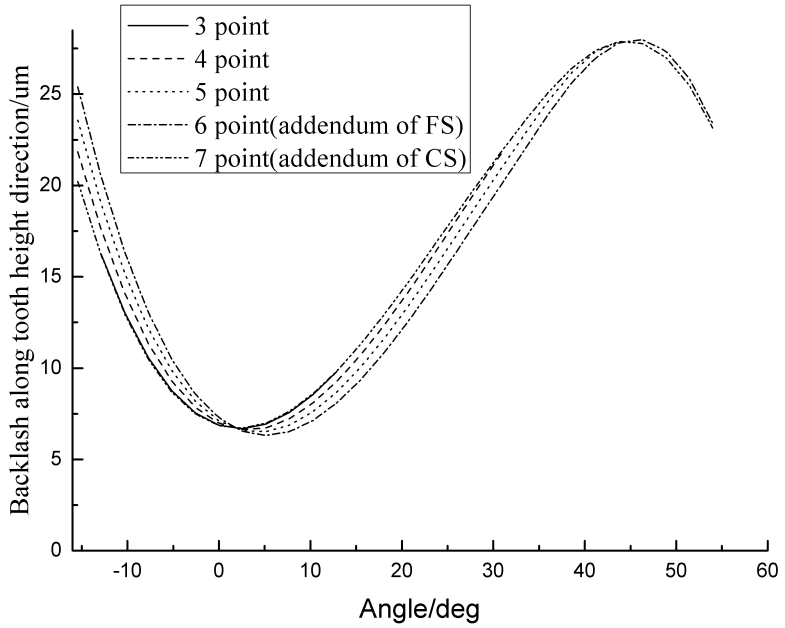


Figure 6. The backlash distribution along tooth height direction

Fig. 6 is the backlash distribution based on computation of FEM. It shows that point 1 and point 2 are close to the dedendum of FS, and do not enter the engagement interval, so there is no backlash value. Most interval of the location of backlash appears alternately between point 6 and point 7. It appears between point 3 and point 5 in φ =1°-2°, but the interval is quite small. Therefore, it is considered that the location of backlash occurs mainly at the addendum of FS or the addendum of CS. Fig. 7 shows the backlash of the addendum of FS and the addendum of CS. At the same time, for verifying the consistency between the FEM and the theoretical algorithm, the backlash curve of the addendum of FS from the article (Shen, 1985) [3] is added in Fig. 7.

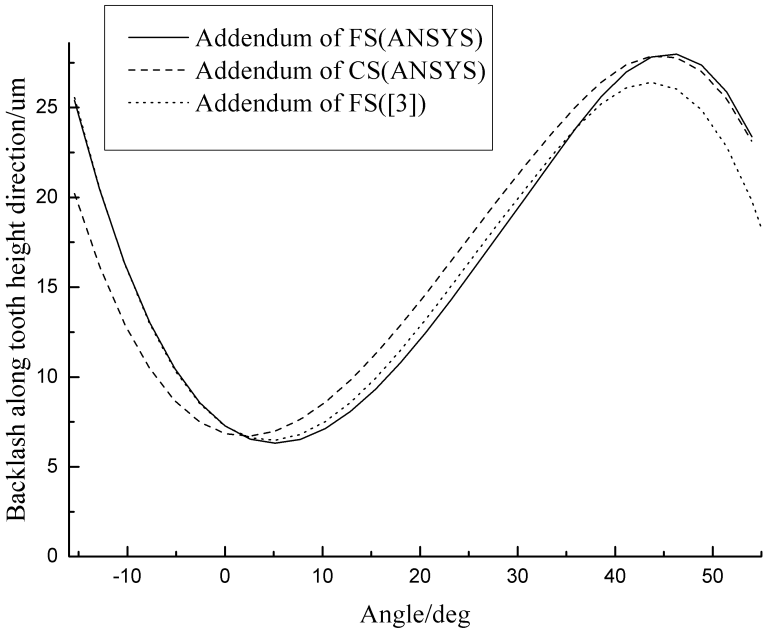


Figure 7. The backlash of the FS and the CS

Consistency analysis between the FEM and the theoretical algorithm (Shen, 1985) [3]: The backlash of addendum of FS from the FEM is basically consistent with that of the theoretical algorithm at near major axis (φ =2 °). There are differences in other areas, especially at near φ =43 °. Because the circumferential displacement is zero in the major axis, the result are consistent. This result verifies the validity of the FEM. In other areas, the circumferential displacements of the two methods are different, which causes backlash deviation. There are many reasons for the deviation, such as small deformation assumption, simplification of normal rotation formula and position formula of the deformed FS (Chen, 2014) [14].

Backlash analysis along tooth height direction: Two backlash curves go across twice at φ =2 °and φ =43 °. The backlash of the addendum of CS is smaller than the addendum of FS in φ =-15 °-2 °and φ = 43 °-55 °. The backlash of the addendum of FS is smaller than the addendum of CS in φ = 2 °-43 °. Throughout the engagement process, the location of minimum backlash appears at the addendum of FS at near φ = 5 °, and the location of maximum backlash appears at the addendum of FS at near φ =43 °.

4 Conclusions

(1) This paper presents a method of backlash computation based on the parametric FEM. Through the analysis of backlash on the major axis of the deformed FS, the validity and consistency between the FEM and the theoretical algorithm are proved. This method provides a reliable way to calculate the backlash.

(2) Two backlash curves of the addendum of FS and CS go across twice at the engagement interval. The first cross appears in the vicinity of the major axis area (φ =2 °-3 °), and the second cross appears in out of the engagement area immediately (φ =43 °-45 °). The backlash of the addendum of CS is the smallest before the first cross and after the second cross. The backlash of addendum of FS is the smallest between the first cross and the second cross. FS and CS belong to tooth surface engagement at two cross intervals, and point engagement at the rest interval.

Acknowledgments

The authors appreciate the support from the National Natural Science Foundation of China (No 51575390), and the Tianjin Research Program of Application Foundation and Advanced Technology (No 14JCYBJC19200).

References

[1] Xie, J R. (1979) The application and development of harmonic drive at home and abroad, *Optical Machinery* **(4)**, 22-31.
[2] Ivanov, M. H. (1987) *The harmonic drive*, 1st edn, Defense Industry Press. Beijing, China.
[3] Shen, Y. W. and Ye, Q. T. (1985) *Theory and design of harmonic drive*. 1st edn, Mechanical Industry Press. Beijing, China.
[4] Lu, Q. H., Liang, Y., Fan, Y. X. and Wang, H. K. (2009) Research on tooth profile interference based on meshing simulation of harmonic drive, *Journal of system simulation* **21(19)**, 6317–6320.
[5] Pan, F., Dong, H. J. and Ge, W. J. (2010) Backlash analysis on the harmonic gear transmission under extremity environment, *Journal of Machine Design* **27(2)**, 58–62.
[6] Yin, Y. (2010) Parameter optimization design and finite element analysis of zero-lateral space involute harmonic gear driving, Yanshan University, Qinhuangdao, China.
[7] Zhang, J. L., Zhou, J., Zhou, F. Q. and Lin, P. (2012) Accurate analysis of harmonic gear clearance and visual simulation, *Computer Simulation* **29(7)**, 292–296.
[8] Chen, X. X., Lin, S. Z., Xing, J. Z. and Liu, Y. S. (2011) Simulation on gear backlash and interference check of harmonic drive with circular-arc teeth profile, *Computer Integrated Manufacturing Systems* **17(3)**, 643–648.
[9] Chen, X. X., Liu, Y. S., Xing, J. Z. and Xu, W. (2014) The parametric design of double-circular-arc tooth profile and its influence on the functional backlash of harmonic drive, *Mechanism and Machine Theory* **73(2)**, 1–24.
[10] Liu, D. H., Xing, J. Z. and Chen, X. X. (2015) Spatial tooth profile design and simulation analysis of harmonic drive with involute tooth profile, *Computer Integrated Manufacturing Systems* **21(3)**, 709–715.
[11] Wang, Y. Q. and Zhang, Z. Y. (2017) Harmonic gear backlash computation of involute tooth profile based on envelope accurate algorithm, *M&E Engineering Technology* **46(03)**, 96–100.
[12] Wei, L. B., Tao, Y. F. and Xu, G. N. (2015) parametric modeling and analysis of finite element based on APDL language, Design and Computation **(1)**, 27–29.
[13] Zeng, P., Lei, L. P. and Fang, G. (2010) *Finite element analysis guide: modeling and analysis of structure*, 1st edn. Mechanical Industry Press. Beijing, China.
[14] Chen, X. X., Liu, Y. S., Xing, J. Z. and Xu, W. (2014) Neutral line stretch of flexspline in harmonic driver, *Journal of Mechanical Engineering* **50(21)**, 189-196.

Unsteady Aerodynamic Simulation of Offshore Wind Turbines with Wave-wind Interaction

Ping Cheng [*], Yong Ai, Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: xiakee@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

The aerodynamic simulation of an NREL-5MW baseline wind turbine is conducted. The three-dimensional Reynolds Averaged Navier-Stokes equations (RANS) are solved, and the k-ω SST turbulence model is used. The in-house code naoe-FOAM-os-SJTU solver based on OpenFOAM and overset grid technique is employed. The wave-wind interaction is studied first. The interaction between wind flow and with wave surface reduces the wind speed below certain height, and a shear wind model is obtained and employed as the non-uniform inflow wind for the unsteady aerodynamic simulation of wind turbine. The aerodynamic simulations of the 5-MW baseline wind turbine under uniform or non-uniform inflow wind are conducted respectively. With shear inflow wind, when the lower-speed layer overlaps the rotating region of turbine blades, the aerodynamic forces on turbine rotor oscillate with 1/3 rotating period.

Keywords: aerodynamic simulation; offshore wind turbine (OWT); overset grid technique; naoe-FOAM-os-SJTU solver

Introduction

As a kind of renewable and sustainable energy source, wind energy shows great potential to solve the worldwide energy and environment crisis. With special and strong advantages over traditional onshore wind turbines, offshore wind turbines (OWT) become more attractive. Accurate aerodynamic prediction of wind turbine is one of the common challenges in design of offshore wind turbines.

Most of the early aerodynamic simulations are conducted with uniform wind condition, while the wind velocity is reduced when the air moves against the surface of water wave (Atcheson, 2016). And shear wind causes oscillations of the aerodynamic forces and power generated by a wind turbine. These oscillations of forces have 1/3 rotating period for a three-bladed wind turbine (Thiringer, 2001). This oscillation on power may effects the output power quality and the control system, and the oscillations of aerodynamic forces could cause additional fatigue damages on the structure. Several studies on this oscillation caused by shear wind have been conducted. Bayne (2000) researched the power oscillation due to turbine blades encountering different wind speed at different vertical positions, and stated that the minimum power occurred when any individual blade pointed directly downwards. Dolan (2006) also study the oscillation on torque and power of a three-bladed wind turbine caused by shear wind and tower shadow effects, similar conclusions were drawn as Bayne did, which stated that the torque and power generated by a three-bladed wind turbine exists oscillation with 3 times the rotational frequency, and the maximum value of oscillating torque and power was observed

when one blade was pointed directly upstream and the minimum value was noted while one blade was pointed directly downstream.

In this paper, the unsteady aerodynamic performance of the NREL-5MW Baseline wind turbine with shear inflow wind is investigated. The three-dimensional Reynolds Averaged Navier-Stokes equations are solved for the aerodynamic numerical simulation. The naoe-FOAM-os-SJTU solver, which is based on OpenFOAM and overset grid technology and developed for ship and ocean engineering problems, is employed. In this paper, the interaction between wave and wind is studied first to get a proper shear wind model for unsteady aerodynamic simulations with non-uniform inflow wind. And aerodynamic simulations in three different inflow wind conditions are conducted. From the simulation, the time series of the unsteady torque and thrust are obtained, together with the detailed information of the wake flow field to clarity the detailed flow filed information.

1 Numerical Method

1.1 Governing Equations

In the numerical simulations in this paper, the incompressible Reynolds-Average Navier-Stokes (RANS) equations are solved, which contain the continuity equations and the momentum equations:

$$\frac{\partial U_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial U_i}{\partial t} + \frac{\partial}{\partial x_j}\left(U_i U_j\right) = -\frac{1}{\rho}\frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j}\left(\nu\frac{\partial U_i}{\partial x_j} - \overline{u_i' u_j'}\right) \tag{2}$$

In these equations, U_i is the averaged flow velocity component, while u_i' is the fluctuation part; ρ is the density of the fluid; p is the pressure; ν is the kinematic viscosity.

But the equations are not closed as they contain more variables than there are equations. In order to meet the closure requirement and solve the above equations, the two-equation turbulence model k-ω SST (Menter, 1994) is employed, and the turbulent kinetic energy k and the turbulent dissipation rate ω can be described as:

$$\frac{\partial}{\partial \mathrm{t}}(\rho k) + \frac{\partial}{\partial x_i}(\rho k\, \mathrm{u}_i) = \frac{\partial}{\partial x_j}(\Gamma_k \frac{\partial k}{\partial x_j}) + \mathrm{G}_k - Y_k + S_k \tag{3}$$

$$\frac{\partial}{\partial \mathrm{t}}(\rho \omega) + \frac{\partial}{\partial x_i}(\rho \omega\, \mathrm{u}_i) = \frac{\partial}{\partial x_j}(\Gamma_\omega \frac{\partial \omega}{\partial x_j}) + \mathrm{G}_\omega - Y_\omega + D_\omega + S_\omega \tag{4}$$

Where, Γ_k and Γ_ω are the effective diffusion coefficients for the turbulent kinetic energy k and the turbulent dissipation rate ω respectively, G_k and G_ω are turbulence generation terms for k and ω, Y_k and Y_ω are turbulent dissipation terms for k and ω, S_k and S_ω are the source term for k and ω, D_ω is the cross-diffusion term for ω.

1.2 Overset Grid Technique

Using overset grid technique, the separated overlapping grids for each part with independent motion are allowed, which makes powerful in simulation of large amplitude motion problems. And the connection among grids of each part is built by interpolation at appropriate cells or points using DCI (domain connectivity information) which is produced by SUGGAR++.

(Noack, R.W. 2005b. Carrica, et al. 2010b). There are four main steps when using DCI in the overset grid technique: The first step is to mark the hole cells which are located outside the simulation domain or of no interest, and exclude them from computation. As shown in Fig.1, in each overset grid, there exist series of cells around hole cells named fringe cells, and for each fringe cell there are several donor cells which provide information from the donor grids, so the second step is to seek for the donor grids of each fringe cell and provide information from the donor grids. The third step is to obtain the value of a variable ϕ of the fringe cell by interpolation using Eq.1 from the donor cells find in the second step.

$$\phi_I = \sum_{i=1}^{n} \omega_i \cdot \phi_i \qquad (5)$$

Where ϕ_I is the value of a variable ϕ of the fringe cell, ϕ_i is the value for the i^{th} donor cell, ω_i is the weight coefficient, which is dimensionless and follows the condition shown in Eq.2:

$$\sum_{i=1}^{n} \omega_i = 1 \qquad (6)$$

And the last step is to optimize the overlapping area and improve the accuracy of interpolation.

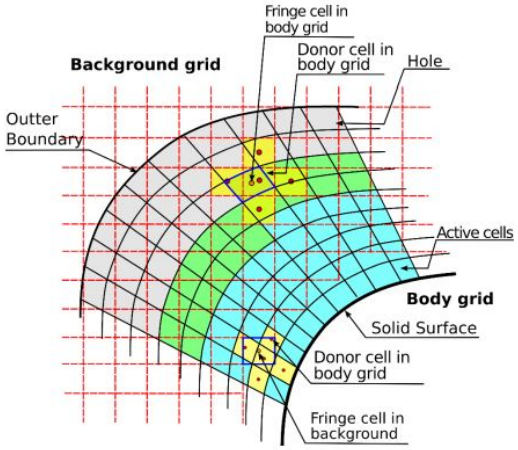


Fig. 1 Diagram of overset grid

2 Physical Model and Grid Structure

2.1 NREL-5MW Baseline Wind Turbine

The representative utility-scale NREL-5MW baseline wind turbine (Jonkman, 2009) is employed in current study. This is a three-bladed wind turbine and the length of each blade is 61.5m. For each blade, the cross sections from blade root to blade tip are composed of a series of DU_XXX and NACA64_XXX airfoils (Lindenburg, 2002). With these detailed data of blade and airfoils, the structural model is built (Bazilevs, 2011) which is shown in Fig.2.

Table 1. Summery of NREL-5MW Baseline Wind Turbine Properties

Rating Power	5MW
Rotor Orientation	Upwind
Rotor Diameter	126m
Hub Diameter	3m
Hub Height	90m
Maximum Rotor	12.1rpm
Maximum Tip Speed	80m/s
Overhang / Shaft Tilt / Precone	5m / 5º / 2.5º

Two overlapping meshes are generated to use the overset grid technique, which are the background mesh grid of the simulation domain generated, and the overlapping grids for the turbine rotor. Both the two set of grid are generated with OpenFOAM including the blockMesh and snappyHexMesh utilities.

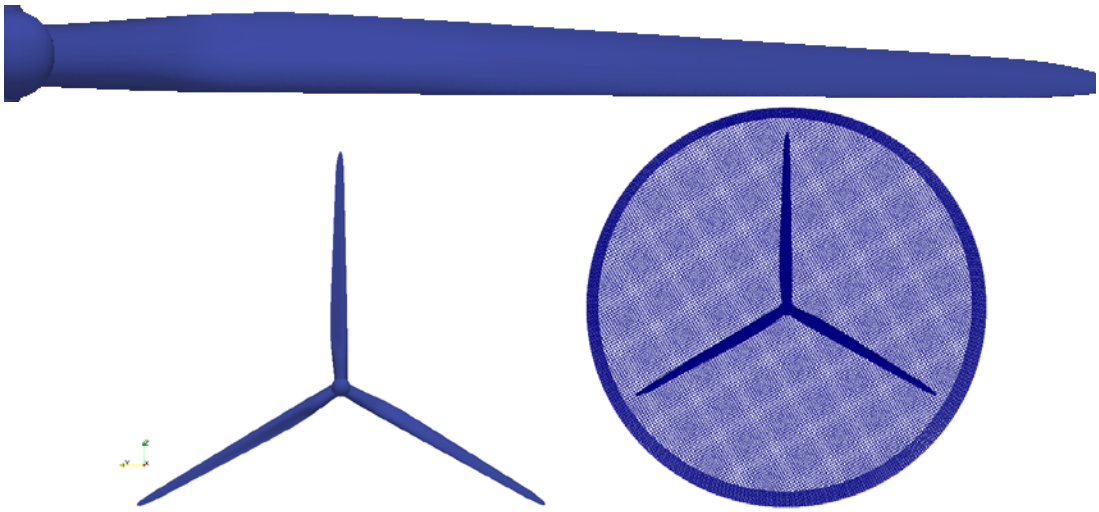


Figure 2. NREL-5MW Baseline Wind Turbine

In the turbine rotor grid system, a cylinder subdomain is built with radius of 72m, and width of 20m. and 3 levels of refinement around the blades are conducted using refineMesh utility in OpenFOAM, and another two levels of refinement are done before surface snapping, and six layers are set up at the surface grid. The total number of cells for this part of grid is about four million.

2.2 Setup of Simulation Domain

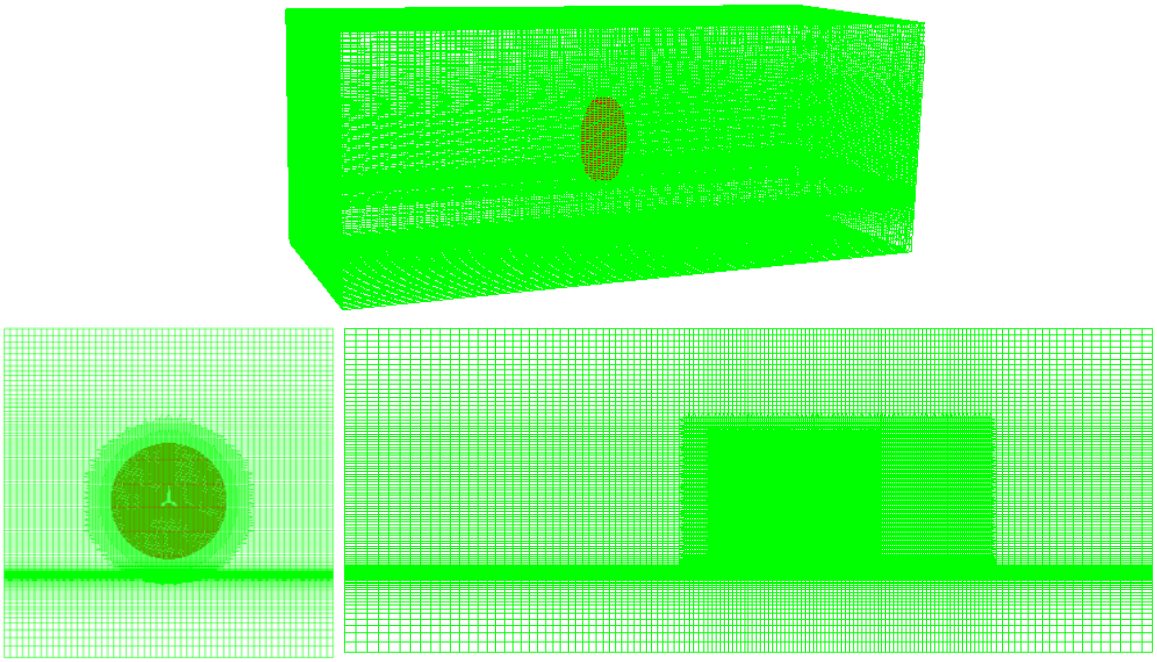


Figure 3. Setup of the simulation domain

According to the structural properties, the simulation domain is generated as a box (X: -500m~500m, Y: -200m~200m, Z: -100m~300m), which is shown in Fig.3. To improve the

simulation accuracy, refinement of the mesh around turbine blades is necessary, and proper mesh refinement in the wake flow field is also very important to capture the flow information in the wake flow. The region of refined mesh is a cylinder with radius of 110m, and length of 350m (-50m~300m). To take the wave-wind interaction into consideration, refinement of mesh near water free surface is also carried out. And the simulation of wave-wind interaction is conducted with this background grid system without refinement of the cylinder region.

3 Shear Wind Model

Most of the aerodynamic studies on the wind turbine are conducted in uniform inflow wind conditions. While with effects of free surface, the air velocity is reduced when moving against surface of wave. A wave-wind interaction study was conducted first to get the shear wind model. In the simulation domain described above, regular wave with period of 10s and wave length of 156m is generated, and the boundary condition of inlet is set as fixed inlet wind with wind speed distribution described as case2 in table.2. After 100s simulation, the steady state of the wave-wind interaction is available.

Fig.3 shows the distribution of wind velocity with increase of height from still water level (z=0) at x=0, where the center of turbine rotor locates. As the free surface changes all the time, which may cause the variation of the flow velocity, velocity distributions at four typical moments during one complete wave period (T0=200s, T0+1/4T, T0+2/4T, T0+3/4T) are presented in Fig.4. Comparing these four sets of data, tiny differences are observed, which mainly lie in lower areas where height is less than 40m. As introduced above about the NREL-5MW wind turbine, the rotational center of this turbine lies 90m above the still water level, and the diameter of rotor blade is 126m. So the lowest point of the rotating blade is above 27m (not equal to 27m because of the shaft tilt of the rotor plane and the precone of each blade introduced in Table.1). So the differences between velocity distributions at different moments during a wave period are negligible. With these data, a fitting curve is achieved as the red line in Fig.4. With wave-wind interaction, the wind speed in lower area decreased obviously, while the wind velocity in the above area gets increased.

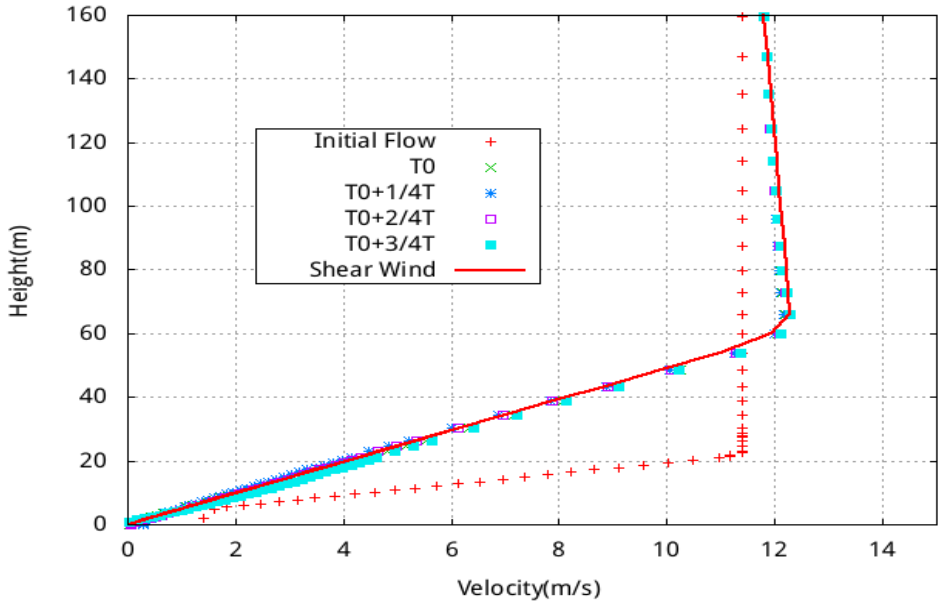


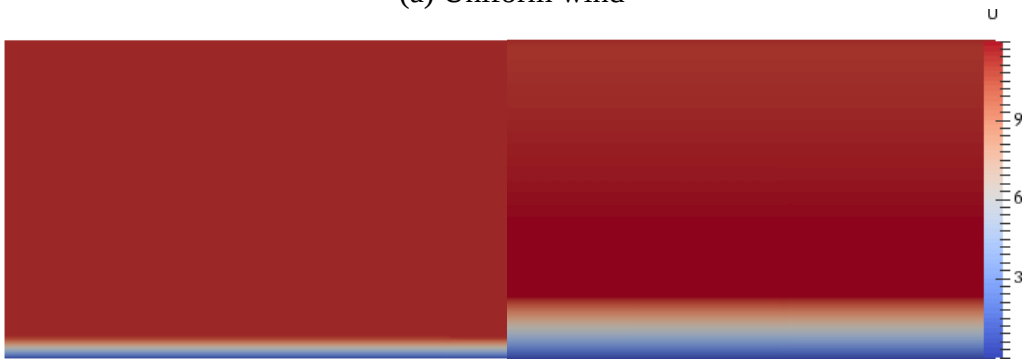
Figure 4. The shear wind model

In this paper, three simulation cases are conducted (Table.2). Uniform inflow wind with wind speed U=11.4m/s is for case1, linear type shear wind described as a piecewise function shown in table.2 is set in case2, and the fitting curve shown in Figure.3 is applied on inlet boundary as well as initialization of the flow field in case3. The initial wind velocity distributions in three cases are shown in Fig.5.

Table 2. Setup of simulation cases with different wind velocity distribution

No.	Cases	Description of inflow wind speed
1	Uniform Wind	U=11.4m/s
2	Linear Shear Wind	$U=\begin{cases}11.4*(20-z)/20 & z<20\\ 11.4 & z\geq 20\end{cases}$ (m/s)
3	Shear Wind with Wave-wind Interaction	Fitting curve ("shear wind") shown in Fig.4

(a) Uniform wind



(b) Linear shear wind (c) Shear wind with wave-wind interaction

Figure 5. Initial wind speed for three cases

4. Results and Discussions

4.1 Aerodynamic Thrust and Torque

From the simulation, the time history of unsteady thrust and torque of the wind turbine are obtained. The time history of thrust and torque are shown in Fig6. Non-dimensional treatment is done on the thrust and torque results by dividing the thrust or torque by the mean value. And the time value t is divided by the rotating period of turbine rotor (T=4.96s).

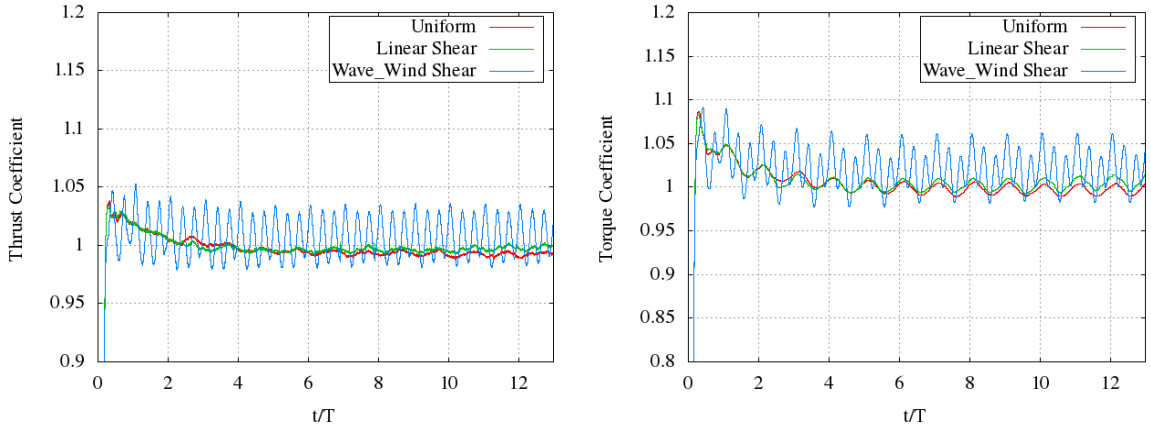
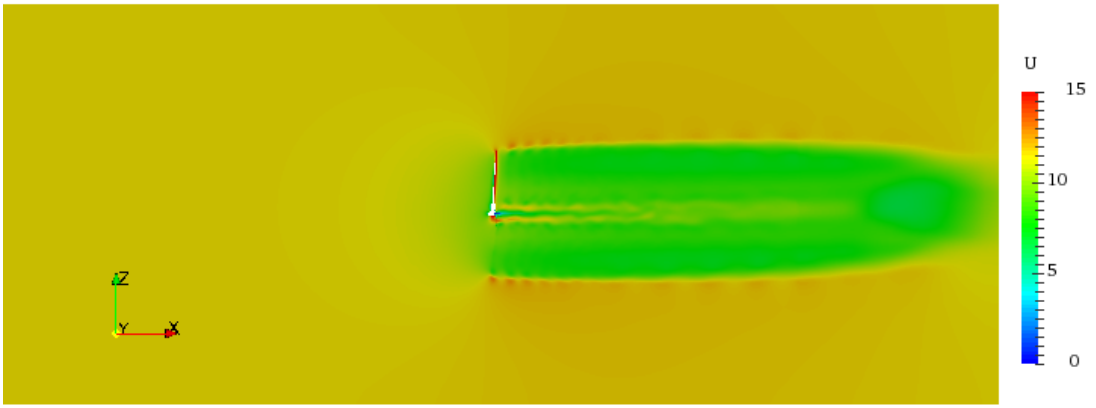


Figure 6.Aerodynamic thrust and torque

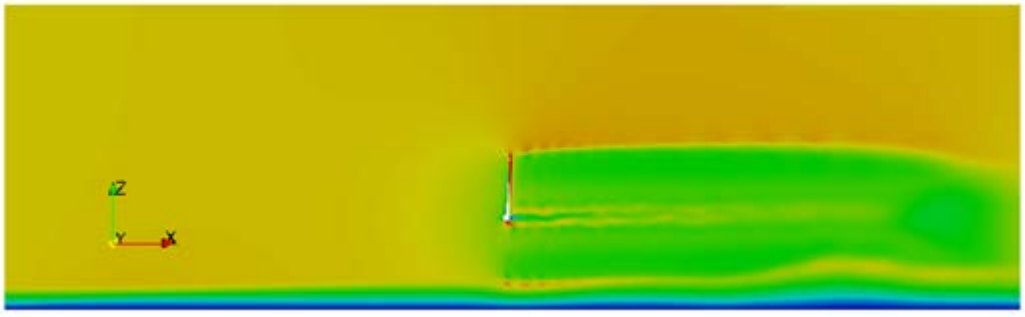
Both the aerodynamic thrust and torque show steady state after about 2~4 rotating periods. The aerodynamic thrust and torque coefficients in both case1 (uniform inflow wind) and case2 (linear shear wind distribution for 20m) shows slight oscillation with same period as turbine rotation, this is believed to be led by the shaft tilt of the rotor plane. The results in case1 and case2 are close, which can be explained with Fig.4 and Fig.5. As illuminated above, the wind speed distributions are the same in the turbine blades rotating areas (z>20m), so the aerodynamic forces are not quite different. Oscillation with much larger amplitude is observed on aerodynamic forces in cases3 (shear wind with wave-wind interaction). In case3, the wind speed is lower than 11.4m/s as height z<60m, while the wind speed grows larger than 11.4m/s in above areas, and turbine blades rotates at the range of 30<z<150. In this case, the aerodynamic forces on a blade rotates through the above area with higher wind speed will be larger than those aerodynamic forces on another blade in lower area with lower wind speed. As the three-blade wind turbine rotates periodically, the aerodynamic forces show periodic oscillating regularity with 1/3 of rotating period. Besides this oscillating regularity, it is also notable that the mean values of the aerodynamic forces in case3 are larger than those in the other two cases. With a close-up view of the curves in Fig.4, the velocity expanding in the height above z=57m is observed on the red line, which means the wind speed is over 11.4m at most of the turbine rotating areas. There's no doubt that the aerodynamic forces grows with the increasing wind speed.

4.2 Flow Field Description

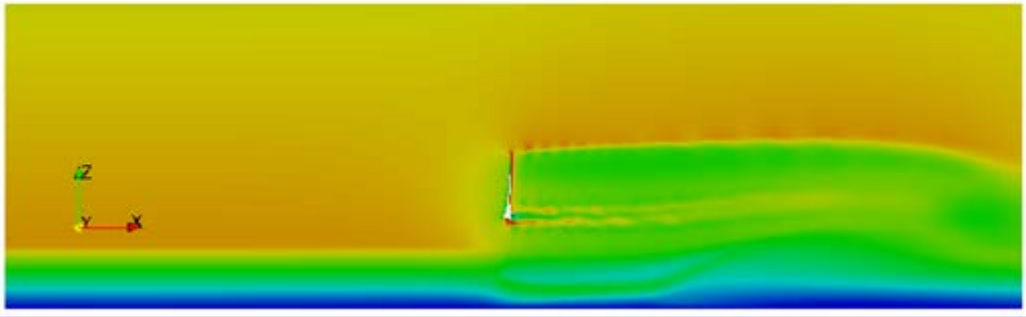
The flow velocity distributions in longitudinal sections are shown in Fig.7. Compared with the initial wind speed shown in Fig.5, the flow velocity shows considerable reduction at the wake flow area as the green regions in the three figures. The upper regions (z>0) of the three figures have slight difference, because the wind speed in these regions are similar, especially for case1 and case2 of which the inflow wind speed are exactly the same in the above areas. Compare with case1, case2 and case3 show stronger interaction between the lower-speed region and the wake flow field of wind turbine. With this interaction, the flow velocity decreases around the lower border of wake flow in case 2 and case 3, in which cases the shape of the borderlines are changed, while the thickness of the lower-speed layer decreases, which means increase of flow velocity in the this lower-speed region.



(a) case1: uniform inflow wind



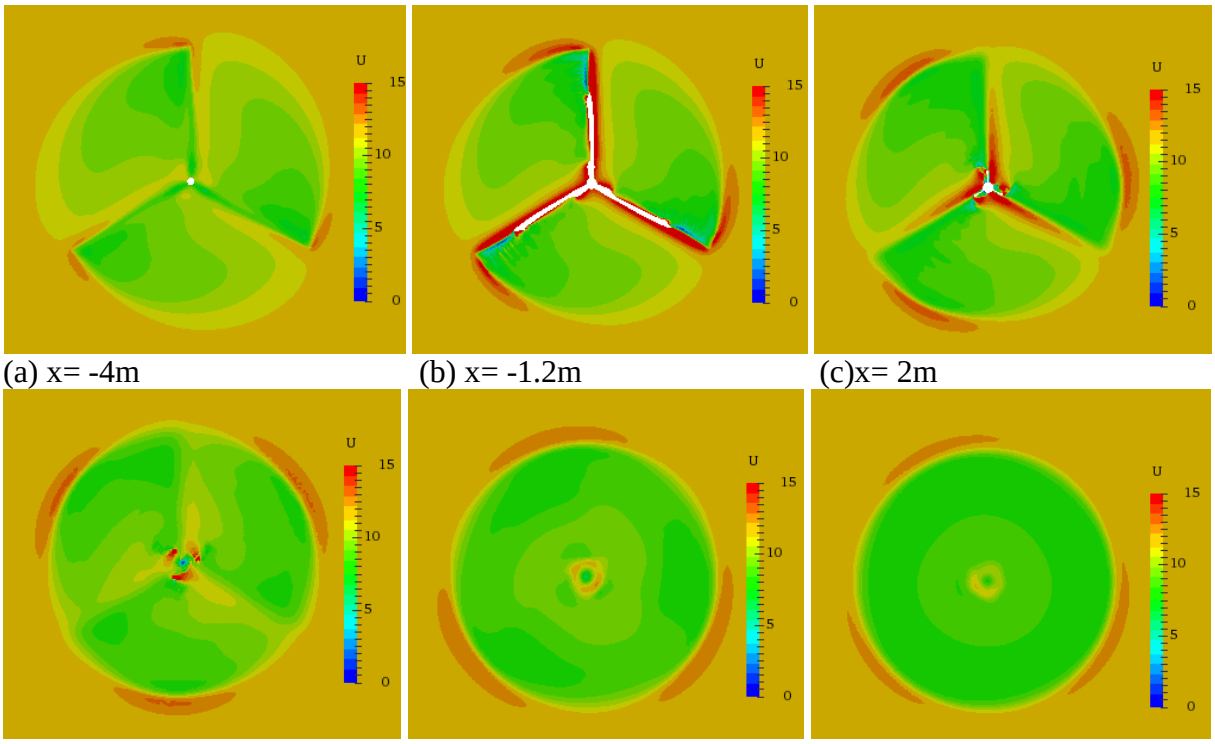
(b) case2: linear shear wind



(c) case3: shear inflow wind with wave-wind interaction

Figure 7. Flow velocity distribution on longitudinal section

Fig.8 shows the flow velocity distribution on cross sections at six different locations along x direction. The location of the cross sections in order from upstream side to downstream region are: (a) x=-4m, (b) x=-1.2m, (c) x=2m, (d) x=5m, (e) x=15m, (f) x=50m, where the rotor center locates at x=0m. As the wind speed in the turbine rotation regions are the same in case1 and case2, the flow velocity distribution looks similar in (a)~(d), while differences appear in the last two sections.



(a) x= -4m (b) x= -1.2m (c)x= 2m

(d) x= 5m (e) x= 15m (f) x= 50m

Figure 8. Flow velocity distribution on cross sections in case1

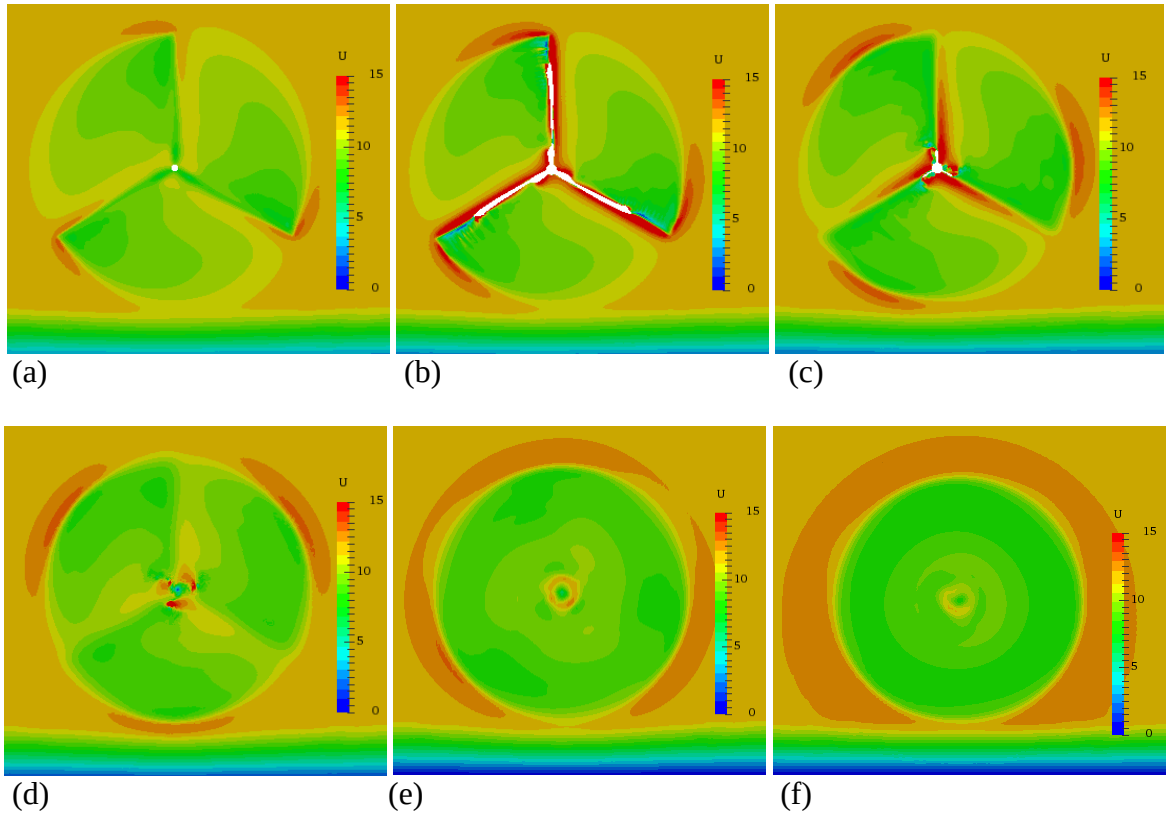
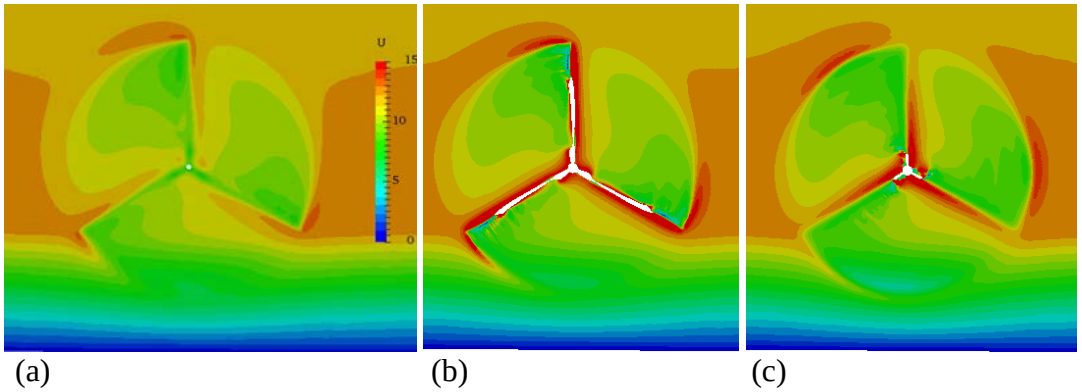
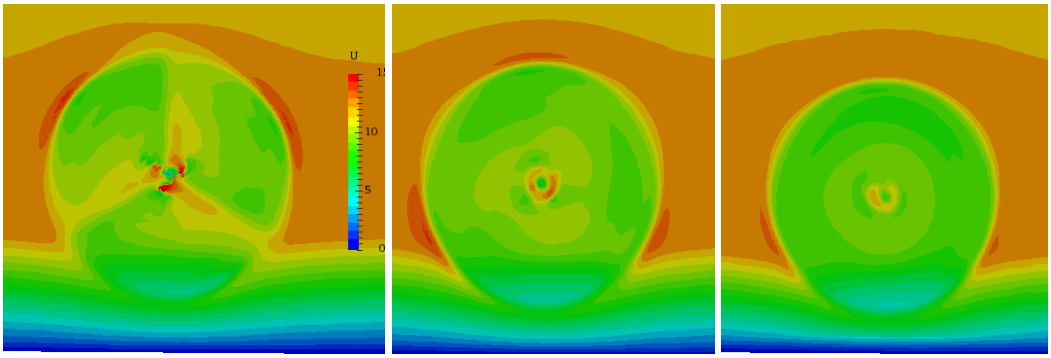


Figure 9. Flow velocity distribution on cross sections in case2

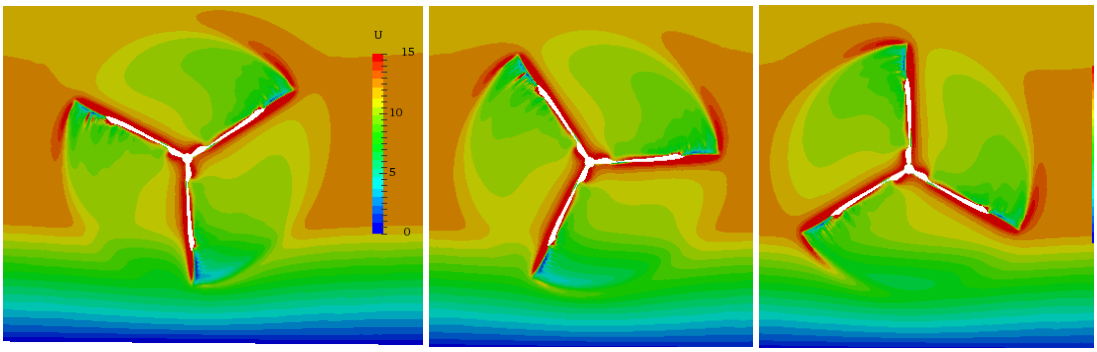
Disturbance of the rotating blade breaks the steady state of initial flow field near blades surface. With viscous effect, flow nearby the blade surface rotates with the blade, which then interacts with neighbor flow and the incoming flow with initial velocity, so the flow velocity in wake flow field decreases and wake vortex propagates to downstream field. In Figure8~11, the flow field are colored with flow velocity, where the red regions mean increased velocity, and green regions mean decrease of velocity. In each figure, red regions occur near the outer edge of wake vortex flow, which contains three pieces in most figures. These outer high-speed regions are the cross section of the spiral tip vortex (the wake vortex structures are shown in Fig.12). With interaction with neighbor flow, the flow velocity in wake vortex decreases slightly as the vortex propagates downstream. In case2 and case3, the flow velocity in spiral vortex near the lower-speed regions is reduced, while the vortex tubes become wider as shown in (d)~(f) in Fig.9~10.





(d) (e) (f)

Figure 10. Flow velocity distribution on cross sections in case3



(a) t=50.8s (b) t=51.2s (c)t=51.6s

Figure 11. Flow velocity distribution on cross sections during 1/6 period

Flow velocity distributions nearby wind turbine at three typical moments are presented in Fig.11, which are selected as the moment when the aerodynamic forces reach the minimum, medium and maximum values. These three moments represents for different azimuth angles of the rotor: (a) t=50.8s, azimuth angle=3659deg, minimum value of aerodynamic forces (b) t=51.2s, azimuth angle= 3689deg, medium value of aerodynamic forces (c) t=51.6s, azimuth angle=3720deg, maximum value of aerodynamic forces. When the rotor rotates to the similar pose as (a), part of one blade submersed in lower-speed layer, which is bound to reduce the aerodynamic forces on this blade. As the three-blade wind turbine rotates at a fixed angular speed, this reduction occurs three times in every rotating period, just as shown in Fig.6.

4.3 Wake Vortex

Wake vortex structure is an important index in aerodynamic analysis, because the wake vortex near the blades affects the aerodynamic properties of the blades. To get a proper wake vortex visualization result, the second invariant of the velocity gradient tensor, Q (Digraskar D A, 2010), is used to capture the iso-surface of the vortex, which is:

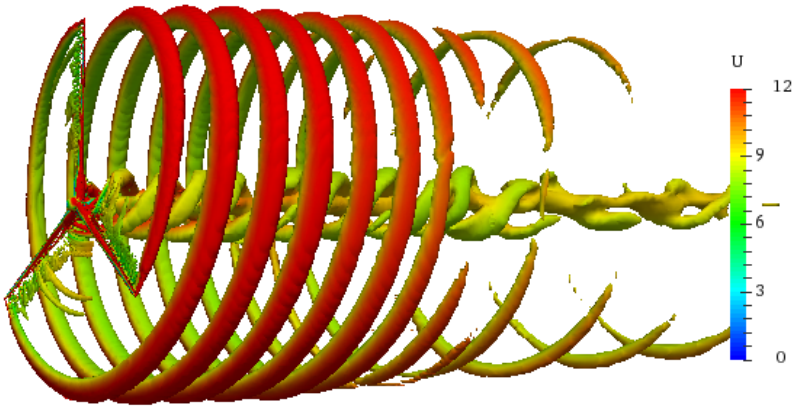
$$Q = \frac{1}{2}(\|\Omega\|^2 - \|S\|^2) \tag{7}$$

Where S is the symmetric part of the velocity gradient tensor and Ω is the anti-symmetric part:

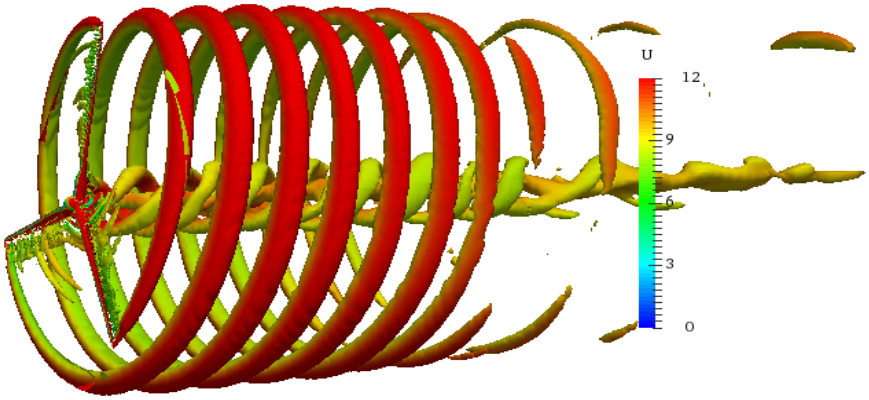
$$S_{ij} = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}),\ \Omega_{ij} = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i}) \tag{8}$$

Fig.12 shows the visualization of wake vortex structures countered by value of Q=0.02. The wake vortex is mainly composed by three spiral tip vortex tubes and three root vortex tubes. The blade tip disturbs flow around its surface, and the rotating flow interacts with its neighbor flow, so the vortex start propagating downstream and the radius of the vortex tube increases

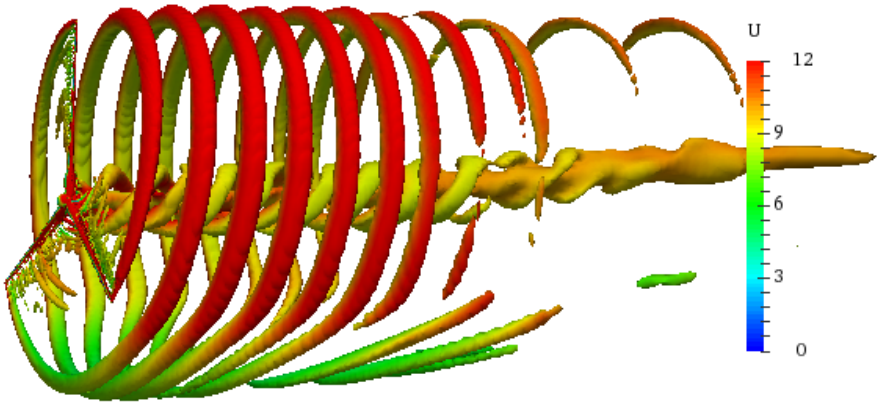
rapidly at beginning of vortex shedding. As discussed above, interaction with lower-speed layer induces the flow velocity in vortex, and makes the tube wider in case2 than that in case1, which accelerates the dissipation of wake vortex at the same time. As the lower-speed layer in case3 is much thicker, part of the blades is immersed in the lower-speed layer. The rotor blades rotates in the same fixed angular speed as in case1 and case2, so the disturbances to wind flow are similar, but the interaction with the lower-speed flow makes the flow velocity decreased significantly in lower-speed layer, as shown in Fig.12(c). The x-component of flow velocity, which mainly comes from the inflow velocity, is the key factor that makes the wake vortex propagating downstream. So the distance between two neighbor vortex tubes narrows down distinctly.



(a) case1: uniform inflow wind



(b) case2: linear shear inflow wind



(c) case3:shear wind with wave-wind interaction

Fig.12 visualization of wake vortex structures

Conclusions

In this paper, the interaction between wave and wind is studied first to get a proper shear wind model for unsteady aerodynamic simulations with non-uniform inflow wind. The wind speed is reduced when passing through the wave surface. The wave surface elevation has slight influence on the flow velocity distribution above certain height (z>30m), so a fitted curve is obtained and employed as distribution of inflow velocity on inlet boundary.

Aerodynamic simulations in three different inflow wind conditions are conducted. Compared with the uniform inflow case, the linear shear wind with lower-speed layer totally below the blades rotating region has negligible effect on the aerodynamic forces of the wind turbine, while the interaction phenomenon between its wake vortex and the neighbor flow is observed. This interaction effect is much pronounced when the lower-speed layer overlaps the rotating region of turbine blades. The aerodynamic forces on the part of blades which is immersed in lower-speed layer are reduced, so the aerodynamic forces of the whole rotor oscillates with 1/3 rotating period.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

Atcheson M. (2016). Floating Offshore Wind Energy.

Bayne S B, Giesselmann M G. (2000). Effect of blade passing on a wind turbine output[C]. Energy Conversion Engineering Conference and Exhibit, 2000.(IECEC) 35th Intersociety. IEEE, 2: 775-781.

Bazilevs Y, Hsu M C, Akkerman I (2011). 3D simulation of wind turbine rotors at full scale. Part I: Geometry modeling and aerodynamics [J]. International Journal for Numerical Methods in Fluids, 2011, 65(1- 3): 207-235.

Carrica P M, Huang J, Noack R (2010) Large-scale DES computations of the forward speed diffraction and pitch and heave problems for a surface combatant [J]. Computers & Fluids, 39(7): 1095-1111.

Digraskar DA (2010). Simulations of flow over wind turbines [J]. Diffraction and pitch and heave problems for a surface combatant. Comput. Fluids 39 (7), 1095–1111.

Dolan D S L, Lehn P W. (2006). Simulation model of wind turbine 3p torque oscillations due to wind shear and tower shadow [J]. IEEE Transactions on energy conversion, 21(3): 717-724.

Jonkman, J, Butterfield, S, Musial, W, Scott, G (2009). Definition of a 5-MW reference wind turbine for offshore system development. Golden, CO: National Renewable Energy Laboratory.

Lindenburg C (2002). Aeroelastic modelling of the LMH64-5 blade [J]. Energy Research Center of the Netherlands, Technical Report No. DOWEC-02-KL-083/0, DOWEC 10083_001.

Menter FR(1994). Two-equation eddy-viscosity turbulence models for engineering applications [J]. AIAA journal, 32(8): 1598-1605.

Noack, RW (2005). SUGGAR: a general capability for moving body overset grid assembly. In: Proceedings of the 17th AIAA Computational Fluid Dynamics Conference, American Institute of Aeronautics and Astronautics, Toronto, Ontario, Canada.

Thiringer T, Dahlberg J A. (2001). Periodic pulsations from a three-bladed wind turbine[J]. IEEE Transactions on Energy Conversion, 16(2): 128-133.

A research of patient-specific flow boundary condition in noninvasive coronary fractional flow reserve

*** Q.Q. Yang[1], † A.K. Qiao[1], Y. Hou[2], and Y. Ma[2]**

[1]College of Life Science and Bioengineering, Beijing University of Technology, Beijing 100124, China
[2]Shengjing Hospital of China Medical University, Shenyang 110001, China

***** Presenting author: yangqq1990@foxmail.com
† Corresponding author: qak@bjut.edu.cn

Abstract

Background: The clinical shortage problems of invasive fractional flow reserve (FFR) can be effectively solved by computed tomography angiography-derived fractional flow reserve (FFR_{CT}), which is a noninvasive functional parameter for the diagnosis of coronary artery disease. However, the accuracy of FFR_{CT} is one of the main problems. The boundary condition is the main factor affecting the accuracy of the FFR_{CT}, while few researches are about this issue.

Object: This study focuses on the settings of the individualized boundary condition for the calculation of FFR_{CT} in order to improve its accuracy.

Methods: A mathematical model of patient-specific flow boundary condition was presented. Firstly, a vessel volume-based method to calculate flow division fraction was presented for the left anterior descending (LAD) artery, the left circumflex (LCX) artery, and the right coronary artery (RCA) based on the "form-function" relationships. Then, the flow division fraction of each coronary outlet could be calculated by combining the shear stress formula of the Hagen-Poiseuille flow, the uniform shear hypothesis and Murray's law. Next, a mathematical model of coronary blood flow at hyperemia (Q_{cor}) was presented. Some independent physiological parameters of coronary blood flow were selected, including the myocardial mass (M_L), diastolic blood pressure (P_D) and heart rate (HR). The model was expressed as: $Q_{cor} = kHR[-0.568\ln(HR) + 3.246]P_D(\frac{M_L}{0.85})^{0.75}$. Finally, the flow rate of each coronary outlet can be calculated by integrating the model and the flow division fraction. Sixteen cases of patients with coronary stenosis were employed for finite element analyses.

Results: (1) The coronary flow divisions for LAD, LCX, and RCA were 32.9%, 20.6%, and 46.5% respectively, and they were almost identical to those from clinical measurement. (2) The mean values of the ratio of total coronary blood flow in cardiac output and myocardial blood flow of 16 patients were 16.97% and 4.07 mL/min/g respectively, which were in accordance with the rule of medical statistics. (3) The diagnostic accuracy of simulation FFR_{CT} was higher than CT alone (85% vs. 65%) with the reference of invasive FFR, and there was a good agreement between FFR_{CT} and FFR.

Conclusions: The coronary FFR_{CT} has good consistency with invasive FFR under the patient-specific flow boundary condition. This study offers a new way for improving FFR_{CT} accuracy, as well as promotes the clinical application of FFR_{CT}.

Keywords: Coronary artery stenosis, Fractional flow reserve, Hemodynamics, Patient-specific boundary condition

Introduction

The clinical shortage problems of invasive fractional flow reserve (FFR) [3] can be effectively solved by computed tomography angiography-derived fractional flow reserve (FFR_{CT}) [1-2], which is a noninvasive functional parameter for the diagnosis of coronary artery disease. FFR_{CT} has been widely concerned by clinicians and researchers in recent years. The diagnostic performance of FFR_{CT} has been investigated experimentally by some multicenter studies, including DISCOVER-FLOW (Diagnosis of Ischemia-Causing Stenoses Obtained Via Noninvasive Fractional Flow Reserve) [4-5]、DeFACTO (Determination of Fractional Flow Reserve by Anatomic Computed Tomographic Angiography) [6-9] and NXT (Analysis of Coronary Blood Flow Using CT Angiography: Next Steps) [10-12], comprising a total of 609 patients and 1050 vessels. The latest trail showed that the accuracy of FFR_{CT} was 81% [11], which is good enough for the result of simulation but not good enough for clinical applications.

The accuracy of FFR_{CT} is the most important problem limiting its clinical application. The main factor affecting the accuracy of the FFR_{CT} is the boundary condition, while few researches are about this issue.

The boundary condition is the driving condition for the coronary blood flow, and directly determines the state of blood flow. The calculation of total coronary blood flow is used for the simplified settings of the boundary condition [1]. In previous studies, the calculation of total coronary blood flow is based on the scaling relation between myocardial mass and coronary blood flow [13]. However, myocardial mass is not the only factor affecting coronary blood flow. There are individual differences in the correlation between myocardial mass and total coronary blood flow. Heart rate and aortic peak pressure were considered as input in the mathematical model of coronary blood flow presented by Arthurs et al [14]. This model can provide patient-specific estimates of coronary blood flow changes between rest and exercise. Except for heart rate and aortic peak pressure, left-ventricular pressure-volume loop was also considered as input parameter which was not easy to detect. So, an accurate and non-invasive method for obtaining total coronary blood flow is a problem that must be faced to improve the accuracy of FFR_{CT} calculation.

In view of the above reasons, this study presents a mathematical model of noninvasive simulation of coronary blood flow, and explores the individualized settings of the boundary condition for FFR_{CT} calculation in order to improve its accuracy.

Methods

Study population

16 patients with steady coronary artery disease were included in our study. The following were reasons for exclusion from the study: (1) myocardial damage or recent myocardial infarction (within one month); (2) left ventricular dysfunction; (3) significant 3-vessel disease; (4) poor quality CT images. All patients underwent coronary computed tomographic angiography (CCTA) and coronary FFR.

Flow division to coronary major branches

Several morphological (diameter, length, and volume) and functional (flow) parameters of the coronary arterial tree in relation to myocardial mass were quantified in animal experiments in vitro by Choy et al [13]. The results showed that arterial volume is linearly related to regional myocardial mass, whereas the sum of coronary arterial branch lengths, vessel diameters, and volumetric flow show a 3/4, 3/8, and 3/4 power-law relationship, respectively. This is

consistent with the earlier experimental results of Seiler et al [15-16]. Based on a physiological observation that longer coronary arteries have more daughter branches feeding a larger mass of cardiac muscle, Lee et al presented a vessel length-based method calculating the coronary flow division over coronary major arteries [17]. Given that the volume parameter contains the diameter and the length, we proposed a vessel volume-based method to calculate flow division over coronary major arteries.

Given that arterial volume is linearly related to regional myocardial mass, the regional myocardial mass of the left anterior descending (LAD) artery, the left circumflex (LCX) artery, and the right coronary artery (RCA) can be described accordingly as equation (1).

$$M_{LAD} = \alpha V_{vessel\ LAD};\ M_{LCX} = \alpha V_{vessel\ LCX};\ M_{RCA} = \alpha V_{vessel\ RCA} \tag{1}$$

where M_{LAD}, M_{LCX}, and M_{RCA} represents the regional myocardial mass of LAD, LCX, and RCA respectively; $V_{vessel\ LAD}$, $V_{vessel\ LCX}$, and $V_{vessel\ RCA}$ represents the vessel volume of LAD, LCX, and RCA respectively; α is a constant scale coefficient.

The volumetric flow showed a 0.75 power-law relationship with myocardial mass, so we can get the equation (2).

$$Q_{LAD} = \beta M_{LAD}^{0.75};\ Q_{LCX} = \beta M_{LCX}^{0.75};\ Q_{RCA} = \beta M_{RCA}^{0.75} \tag{2}$$

where Q_{LAD}, Q_{LCX}, and Q_{RCA} represents the volumetric flow of LAD, LCX, and RCA respectively; β is a constant scale coefficient.

Then, equation (3) can be deduced from the above two equations. Namely, the volumetric flow shows a 0.75 power-law relationship with vessel volume.

$$Q_{LAD} = \alpha\beta V_{vessel\ LAD}^{0.75};\ Q_{LCX} = \alpha\beta V_{vessel\ LCX}^{0.75};\ Q_{RCA} = \alpha\beta V_{vessel\ RCA}^{0.75} \tag{3}$$

So we can deduce the flow division ratio over coronary major arteries as equation (4).

$$Q_{LAD} : Q_{LCX} : Q_{RCA} = V_{vessel\ LAD}^{0.75} : V_{vessel\ LCX}^{0.75} : V_{vessel\ RCA}^{0.75} \tag{4}$$

Flow division to each coronary outlet

The total coronary blood flow is the sum of blood flow over LAD, LCX, and RCA as equation (5).

$$Q_{cor} = Q_{LAD} + Q_{LCX} + Q_{RCA} \tag{5}$$

So the flow division of LAD, LCX, and RCA is described as equation (6) respectively.

$$\left.\begin{aligned} Q_{LAD} &= \frac{V_{vessel\ LAD}^{0.75}}{V_{vessel\ LAD}^{0.75} + V_{vessel\ LCX}^{0.75} + V_{vessel\ RCA}^{0.75}} Q_{cor} \\ Q_{LCX} &= \frac{V_{vessel\ LCX}^{0.75}}{V_{vessel\ LAD}^{0.75} + V_{vessel\ LCX}^{0.75} + V_{vessel\ RCA}^{0.75}} Q_{cor} \\ Q_{RCA} &= \frac{V_{vessel\ RCA}^{0.75}}{V_{vessel\ LAD}^{0.75} + V_{vessel\ LCX}^{0.75} + V_{vessel\ RCA}^{0.75}} Q_{cor} \end{aligned}\right\} \tag{6}$$

Ideally, the blood flow in a vessel is proportional to the 3rd power of the vessel diameter according to Poiseuille's solution and Murray's law [18] as equation (7).

$$Q = \frac{\pi d^3}{4}\sqrt{\frac{\lambda}{\mu}} \tag{7}$$

where Q is the flow rate through a blood vessel, d is its diameter, μ is the fluid viscosity, λ is a constant, and it represents the energy consumed by the metabolism of unit volume.

We set *a*, *b*, and *c* as the diameter of each branch of LAD, LCX, and RCA respectively. The

flow rate of each branch can be inferred from equation (7).

$$\left.\begin{aligned} Q_{LAD-\mathrm{j}} &= Q_{LAD}\frac{a_j^{\,3}}{\sum_{i=1}^{n} a_i^{\,3}} \\ Q_{LCX-\mathrm{j}} &= Q_{LCX}\frac{b_j^{\,3}}{\sum_{i=1}^{m} b_i^{\,3}} \\ Q_{RCA-\mathrm{j}} &= Q_{RCA}\frac{c_j^{\,3}}{\sum_{i=1}^{s} c_i^{\,3}} \end{aligned}\right\} \tag{8}$$

However, the coronary artery model reconstructed from CCTA was not completely matched with Murray's law, thus a revision has been made in equation (8). As shown in Figure 1, the flow division of each branch was calculated in the direction of blood flow according to the classification of vascular branches.

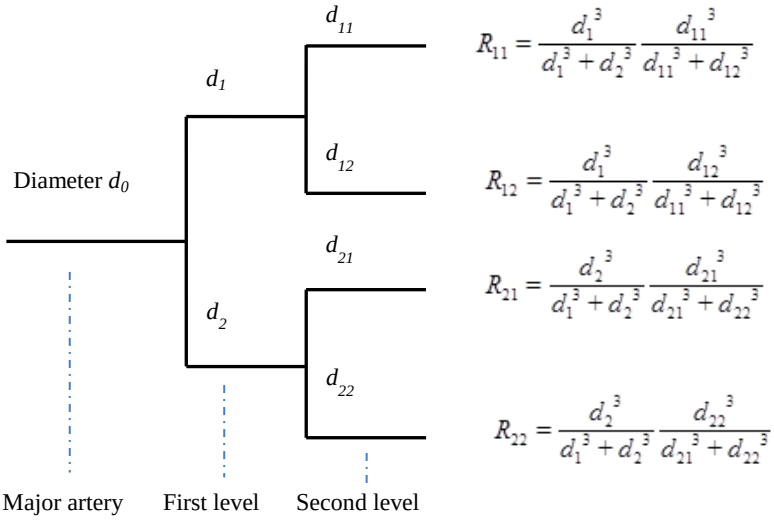


Figure 1. Schematic diagram of diversion ratio calculation (*R* presents diversion ratio)

Model of coronary blood flow

Coronary artery is the vascular system supplying blood for the myocardium. Some vessels buried deep within the myocardium will be pressed during cardiac systole, which will affect coronary blood flow. The blood flow will come to the climax at early diastole, and then decrease slowly [19]. In general, coronary blood flow during left ventricular systole is only 20~30% of those during diastole, and which will be smaller when myocardial contraction strengthens.

Some independent physiological parameters of coronary blood flow were selected based on the discussion above, including the myocardial mass, diastolic blood pressure and heart rate.

Myocardial mass. The scaling laws between coronary blood flow and the myocardial mass can be described as equation (9).

$$Q_{cor} \propto M^{0.75} \tag{9}$$

In a broad sense, the myocardium includes right ventricular myocardium, atrial myocardium and other myocardial tissue besides the left ventricular myocardium. The left ventricular

myocardial mass accounts for about 85% of the total myocardial mass[20-21]. So the scaling law can be revised as equation (10).

$$Q_{cor} \propto \frac{M_L}{0.85}^{0.75} \tag{10}$$

Diastolic blood pressure. The coronary perfusion is mainly affected by the diastolic blood pressure since coronary blood flow perfusion is mainly in diastole. According to Poiseuille's law,

$$Q = \frac{\pi r^4}{8\eta L}\Delta P \tag{11}$$

and coronary anatomy, the perfusion pressure is as equation (12),

$$\Delta P = P_D - P_{Ra} \tag{12}$$

where P_D is the aortic diastolic blood pressure, and P_{Ra} is the right atrial pressure (-2~10 mmHg). For the sake of simplification, the perfusion pressure was set as the aortic diastolic blood pressure in this study. So the relationship between coronary blood flow and the diastolic blood pressure can be described as equation (13).

$$Q_{cor} \propto P_D \tag{13}$$

Heart rate. Effective perfusion time is another factor affecting coronary blood flow. The diastole of a cardiac cycle is the time of coronary blood flow perfusion[19]. Diastolic duration is relatively compressed with the increase of heart rate (Figure 2)[22]. According to the correlation between heart rate and diastolic duration, the perfusion time pre minute can be described as equation (14).

$$T = HR[-0.568\ln(HR) + 3.246] \tag{14}$$

So the relationship between coronary blood flow and the heart rate can be described as equation (15).

$$Q_{cor} \propto HR[-0.568\ln(HR) + 3.246] \tag{15}$$

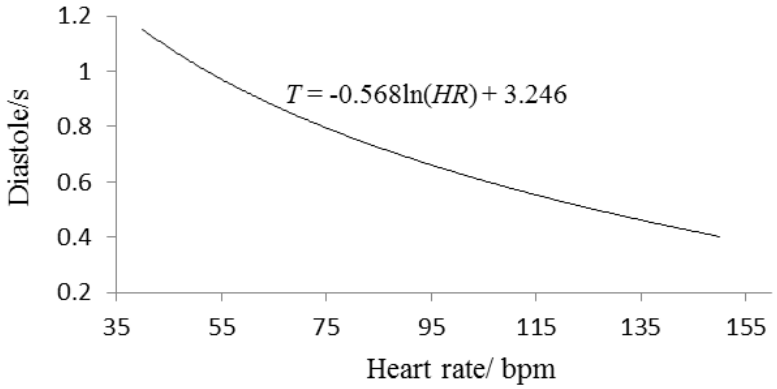


Figure 2. Correlation between left ventricular diastole and heart rate during a cardiac cycle

From the discussion above, the mathematical model of coronary blood flow during hyperemia with adenosine can be established as equation (16).

$$Q_{cor} = kHR[-0.568\ln(HR) + 3.246]P_D(\frac{M_L}{0.85})^{0.75} \tag{16}$$

where *k* is a constant coefficient.

Determination of k

Due to lack of clinical parameters, the reference of coronary blood flow was derived from simulation in this study. 5 patients with mild stenosis and FFR<0.90 served as references to ensure that the simulation was close to the real physiological value.

Figure 3 shows the technical flow of coronary FFR_{CT}. Firstly, three-dimension model of coronary artery was reconstructed from CCTA. Then, the boundary condition was set as the method mentioned above. Blood was modeled as Newtonian fluid to simulate blood flow in the patient-specific coronary artery tree models, and the blood density and dynamic viscosity were constant at 1050 kg/m^3 and 0.0035 pa • s, respectively [23]. The mesh of the geometries was generated using a non-structural mesh with tetrahedron elements. The standard of 1,000,000 elements was appropriate for simulations by mesh independence test. Simulations were carried out using ANSYS Workbench. Steady flow simulation was employed in this study, which reduced the computational cost significantly. It took within 20 minutes of computational time for one case.

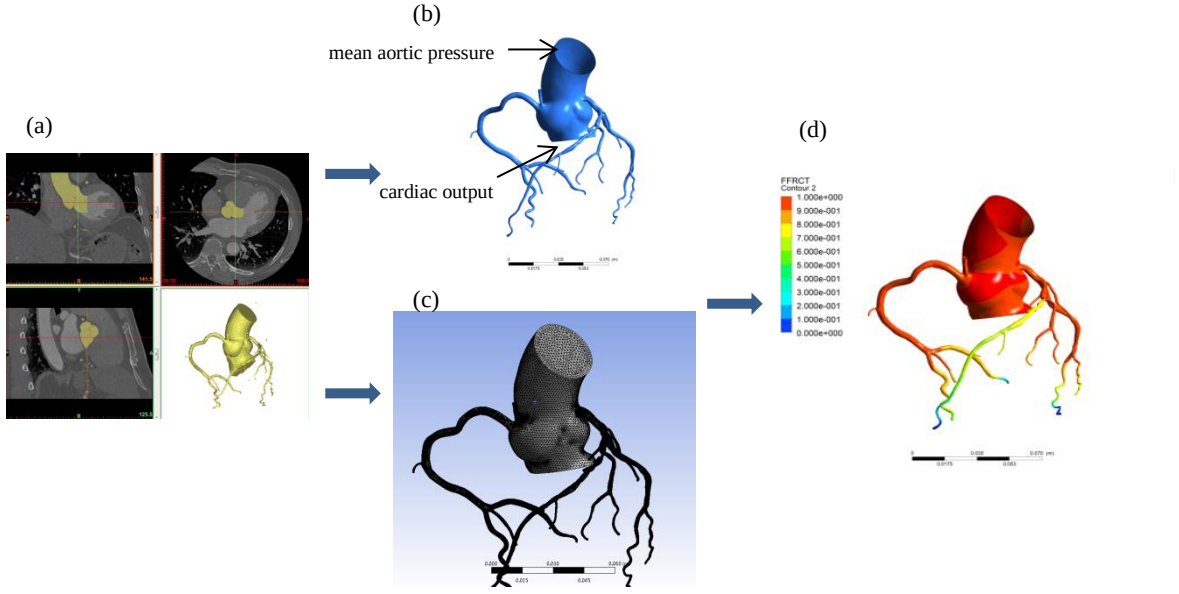


Figure 3. Technical flow chart of coronary FFR_{CT}
(a) 3D reconstruction of coronary artery model based on CCTA; (b) individualized settings of the boundary conditions; (c) meshing of fluid and boundary layer; (d) CFD simulation calculation and post processing

The coronary blood flow was initialized as the product of myocardial mass and myocardial blood flow during hyperemia 3.37 mL/min/g, which is the mean value of human; and then updated as the following steps until the simulation FFR_{CT} matched the invasive FFR.

i. Simulate with the initialization, calculate the coronary FFR_{CT} from the simulation results;

ii. If the simulation FFR_{CT} was larger than invasive FFR, increase the coronary blood flow, otherwise decrease it;

iii. The step-size of adjustment followed by 50, 20, 10, and 5 mL/min, until the simulation FFR_{CT} matched the invasive FFR;

iv. Take the matched one as the reference coronary blood flow.

The reference coronary blood flow was taken into the equation (16) to solve the constant *k*. The mean value of *k* in the 5 patients was taken as the value of *k* in the mathematical model, i.e. *k*=0.003.

Results

Flow division to coronary major branches

Among the 16 patients, 14 cases (88%) were right dominant coronary. The coronary flow division to coronary major branches for the 16 patients was calculated using the vessel volume-based method (Table 1), and they were almost identical to those based on the clinical measurement [24].

Table 1. Comparison of average fraction of coronary flow division over LAD, LCX, and RCA

	LAD (%)	LCX (%)	RCA (%)
Volume-based method	32.9	20.6	46.5
Clinical data	31.1	26.7	41.9

Rationality

Due to lack of clinical coronary blood flow parameters, the ratio of total coronary blood flow in cardiac output (Per_Q) and myocardial blood flow (Q_{myo}) were selected as the evaluation indices to assess the rationality of model of coronary blood flow.

The mean values of Per_Q and Q_{myo} of 16 patients were 16.97% and 4.07 mL/min/g respectively (Table 2 and Table 3), which were in accordance with the rule of medical statistics[25].

Table 2. Summary of Per_Q under model of coronary blood flow

No.	Per_Q (%)	No.	Per_Q (%)
#1	31.85	#9	10.69
#2	17.09	#10	9.67
#3	12.14	#11	14.74
#4	28.79	#12	15.50
#5	10.79	#13	13.35
#6	14.38	#14	14.11
#7	21.58	#15	21.01
#8	21.53	#16	14.22
		Average	*16.97*

Table 3. Summary of Q_{myo} under model of coronary blood flow

No.	Q_{myo} (mL/min/g)	No.	Q_{myo} (mL/min/g)
#1	4.46	#9	3.75
#2	3.94	#10	3.63
#3	4.56	#11	4.61
#4	4.40	#12	4.56
#5	3.68	#13	3.97
#6	2.30	#14	4.79
#7	5.19	#15	3.48
#8	3.94	#16	3.89
		Average	*4.07*

Accuracy

The diagnostic performance of FFR_{CT} was estimated by the reference of clinical coronary FFR, with FFR≤0.80 as threshold.

Bland-Altman method was used to evaluate the consistency of FFR_{CT} and FFR. The 95% confidence interval between FFR_{CT} and FFR ranges from -0.25 to 0.21, and most of the data fall within the interval, indicating that the two indices have good consistency (Figure 4).

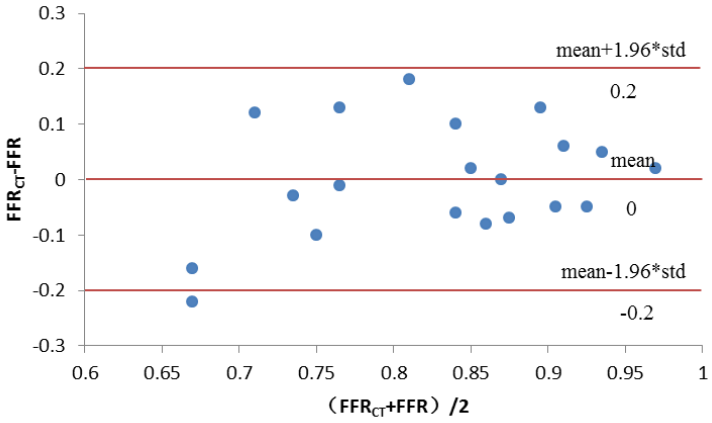


Figure 4. Bland-Altman plot of FFR and FFR_{CT}

Compared with CCTA, which depends only on morphological diagnosis with stenosis rate 50% as ischemic threshold, the diagnostic accuracy of FFR_{CT} was better, i.e. for FFR_{CT} vs CCTA 85% vs 65%, as well as specificity 100% vs. 54.5%, PPV 100% vs. 58.3%, and NPV(78.6% vs. 75% (Table 4).

Table 4. Comparison of the performance between CCTA and FFR_{CT}

Method	Accuracy	Sensibility	Specificity	PPV	NPV
CCTA(50%)	65%	77.8%	54.5%	58.3%	75%
FFR_{CT}(0.80)	85%	66.7%	100%	100%	78.6%

Discussion

This study presented a method of patient-specific flow boundary condition setting orienting to improve the accuracy of coronary FFR_{CT}. The method is simple to operate, and can significantly reduce the time consuming while ensure the accuracy of calculation. A numerical simulation of the calculation takes only 15 minutes with an ordinary PC.

However, there are some limitations in this method. Firstly, the individual difference of coronary microcirculation resistance was not considered. Coronary microcirculation resistance, namely after-loading, which is affected by the activity of the myocardium, affects directly the perfusion of coronary artery. In this study, we supposed that the patients were all with normal coronary microcirculation which had the same response to adenosine and other vasodilator drugs. However, the actual situation is that there is somewhat difference in the coronary microcirculation in different individuals or different regions of individuals, and most of them have different degrees of microcirculatory disturbance [26-27]. Secondly, the model can only be used in the simulation of coronary artery during hyperemia. Whether the model can be applied to rest state still needs to be investigated in-depth. Finally, the simulations in this study are all in steady-state, corresponding to the average condition of hyperemia, so the characteristics of pulsatile blood flow cannot be reflected.

Conclusions

A mathematical method of patient-specific flow boundary condition setting was proposed for the sake of improving the accuracy of coronary FFR_{CT}. By comparing CCTA assessment, coronary FFR_{CT} and invasive FFR, the FFR_{CT} has good consistency with FFR under the patient-specific flow boundary condition. This study offers a new way for improving the accuracy of FFR_{CT}, as well as promoting the clinical application of FFR_{CT}.

Acknowledgement: This work was supported by National Natural Science Foundation of China (11472023, 81301221), and by Program for Liaoning Innovative Research Team in University (LT2014017).

References

[1] Taylor, C. A., Fonte, T. A., & Min, J. K. (2013). Computational fluid dynamics applied to cardiac computed tomography for noninvasive quantification of fractional flow reserve: scientific basis. *Journal of the American College of Cardiology, 61*(22), 2233.

[2] Min, J. K., Taylor, C. A., Achenbach, S., Koo, B. K., Leipsic, J., & Nørgaard, B. L., et al. (2015). Noninvasive fractional flow reserve derived from coronary ct angiography: clinical data and scientific principles. *Jacc Cardiovascular Imaging, 8*(10), 1209-22.

[3] Kleiman, N. S. (2011). Bringing it all together: integration of physiology with anatomy during cardiac catheterization. *Journal of the American College of Cardiology, 58*(12), 1219-1221.

[4] Koo, B. K., Erglis, A., Doh, J. H., Daniels, D. V., Jegere, S., & Kim, H. S., et al. (2011). Diagnosis of ischemia-causing coronary stenoses by noninvasive fractional flow reserve computed from coronary computed tomographic angiograms. results from the prospective multicenter discover-flow (diagnosis of ischemia-causing stenoses obtained via nonin. *Journal of the American College of Cardiology, 58*(19), 1989-97.

[5] Min, J. K., Koo, B. K., Erglis, A., Doh, J. H., Daniels, D. V., & Jegere, S., et al. (2012). Effect of image quality on diagnostic accuracy of noninvasive fractional flow reserve: results from the prospective multicenter international discover-flow study. *Journal of cardiovascular computed tomography, 6*(3), 191.

[6] Min, J. K., Berman, D. S., Budoff, M. J., Jaffer, F. A., Leipsic, J., & Leon, M. B., et al. (2011). Rationale and design of the defacto (determination of fractional flow reserve by anatomic computed tomographic angiography) study. *Journal of Cardiovascular Computed Tomography, 5*(5), 301.

[7] Min, J. K., Leipsic, J., Pencina, M. J., Berman, D. S., Koo, B. K., & Mieghem, C. V., et al. (2012). Diagnostic accuracy of fractional flow reserve from anatomic ct angiography. *Jama the Journal of the American Medical Association, 308*(12), 1237.

[8] Nakazato, R., Park, H. B., Berman, D. S., Gransar, H., Koo, B. K., & Erglis, A., et al. (2013). Noninvasive fractional flow reserve derived from computed tomography angiography for coronary lesions of intermediate stenosis severity results from the defacto study. *Circulation Cardiovascular Imaging, 6*(6):881-889.

[9] Leipsic, J., Yang, T. H., Thompson, A., Koo, B. K., Mancini, G. B., & Taylor, C., et al. (2014). Ct angiography (cta) and diagnostic performance of noninvasive fractional flow reserve: results from the determination of fractional flow reserve by anatomic cta (defacto) study. *American Journal of Roentgenology, 202*(5), 989-94.

[10] Gaur, S., Achenbach, S., Leipsic, J., Mauri, L., Bezerra, H. G., & Jensen, J. M., et al. (2013). Rationale and design of the heartflownxt (heartflow analysis of coronary blood flow using ct angiography: next steps) study. *Journal of Cardiovascular Computed Tomography, 7*(5), 279-288.

[11] Nørgaard, B. L., Leipsic, J., Gaur, S., Seneviratne, S., Ko, B. S., & Ito, H., et al. (2014). The nxt trial (analysis of coronary blood flow using ct angiography: next steps). *Journal of the American College of Cardiology,63*(12), 1145-55.

[12] Miyoshi, T., Osawa, K., Ito, H., Kanazawa, S., Kimura, T., & Shiomi, H., et al. (2015). Non-invasive computed fractional flow reserve from computed tomography (ct) for diagnosing coronary artery disease.*Circulation Journal, 79*(2), 406-12.

[13] Choy, J. S., & Kassab, G. S. (2008). Scaling of myocardial mass to flow and morphometry of coronary arteries. *Journal of Applied Physiology,104*(5), 1281-6.

[14] Arthurs, C. J., Lau, K. D., Asrress, K. N., Redwood, S. R., & Figueroa, C. A. (2016). A mathematical model of coronary blood flow control: simulation of patient-specific three-dimensional hemodynamics during exercise. *American Journal of Physiology - Heart and Circulatory Physiology, 310*(9), H1242-H1258.

[15] Seiler, C., Gould, K. L., & Kirkeeide, R. L. (1992). Basic structure-function relations of the epicardial coronary vascular tree. basis of quantitative coronary arteriography for diffuse coronary artery disease.*Circulation, 85*(6), 1987-2003.

[16] Seiler, C., Kirkeeide, R. L., & Gould, K. L. (1993). Measurement from arteriograms of regional myocardial bed size distal to any point in the coronary vascular tree for assessing anatomic area at risk ☆. *Journal of the American College of Cardiology, 21*(3), 783.
[17] Lee, K. E., Kwon, S. S., Ji, Y. C., Shin, E. S., Choi, J. H., & Kim, S. J., et al. (2016). Estimation of the flow resistances exerted in coronary arteries using a vessel length-based method. *Pflügers Archiv - European Journal of Physiology, 468*(8), 1-10.
[18] Murray, C. D. (1926). The physiological principle of minimum work i. the vascular system and the cost of blood volume. *Proceedings of the National Academy of Sciences, 12*(3), 207.
[19] Abe, M., Tomiyama, H., Yoshida, H., & Doba, N. (2000). Diastolic fractional flow reserve to assess the functional severity of moderate coronary artery stenoses: comparison with fractional flow reserve and coronary flow velocity reserve. *Circulation, 102*(19), 2365-2370.
[20] Yang, M., Shen, X., & Chen, S. (2000). Assessment of the effect of pulmonary hypertension on right ventricular volume and free wall mass by dynamic three-dimensional voxel imaging of echocardiography. *Chinese Journal of Uitrasonography*. *7*:401-404.
[21] Lorenz, C. H., Walker, E. S., Morgan, V. L., Klein, S. S., & Jr, G. T. (1999). Normal human right and left ventricular mass, systolic function, and gender differences by cine magnetic resonance imaging. *Journal of Cardiovascular Magnetic Resonance, 1*(1), 7.
[22] Zhao, Y. X., Liu, J. M., Xu, D. G., Yan, X. B., Lu, L. C., & Xiao, S. Z., et al. (2013). population based study of change trend of the ratio of diastolic to systolic duration during exercise. *Chinese journal of applied physiology, 29*(2), 134.
[23] Zhang, J. M., Zhong, L., Luo, T., Mae, L. A., Huo, Y., & Jonathan, Y., et al. (2016). Simplified models of non-invasive fractional flow reserve based on ct images:. *Plos One, 11*(5), e0153070.
[24] Sakamoto, S., Takahashi, S., Coskun, A. U., Papafaklis, M. I., Takahashi, A., & Saito, S., et al. (2013). Relation of distribution of coronary blood flow volume to coronary artery dominance. *American Journal of Cardiology, 111*(10), 1420-1424.
[25] Uren, N. G., Melin, J. A., De, B. B., Wijns, W., Baudhuin, T., & Camici, P. G. (1994). Relation between myocardial blood flow and the severity of coronary-artery stenosis. *New England Journal of Medicine, 330*(25), 1782-8.
[26] Layland, J. J., Whitbourn, R. J., Burns, A. T., Somaratne, J., Leitl, G., & Macisaac, A. I., et al. (2012). The index of microvascular resistance identifies patients with periprocedural myocardial infarction in elective percutaneous coronary intervention. , *98*(20), 1492-1497.
[27] Erkol, A., Pala, S., Kırma, C., Oduncu, V., Dündar, C., & Izgi, A., et al. (2011). Relation of circulating osteoprotegerin levels on admission to microvascular obstruction after primary percutaneous coronary intervention. *American Journal of Cardiology, 107*(6), 857-62.

Static analysis of functionally graded graphene nanocomposite beams under thermo-electro-mechanical loading

Helong Wu[1], Sritawat Kitipornchai[1], LiaoLiang Ke[2] and *†Jie Yang[3]

[1]School of Civil Engineering, The University of Queensland, St. Lucia, Brisbane, QLD 4072 Australia
[2]Institute of Engineering Mechanics, Beijing Jiaotong University, Beijing 100044, PR China
[3]School of Engineering, RMIT University, PO Box 71, Bundoora, VIC 3083 Australia

*Presenting author: j.yang@rmit.edu.au

†Corresponding author: j.yang@rmit.edu.au

Abstract

The outstanding physical properties, together with the nanoscale effects and interface chemistry, make graphene and its derivatives promising nanofillers to improve the mechanical, thermal and electrical properties of polymeric materials. This paper is concerned with the static analysis of functionally graded multilayer graphene nanocomposite beams that are integrated with two surface-bonded piezoelectric layers and subjected to the combine action of a uniform temperature rise, a constant actuator voltage and a transverse load. The multilayer beam is composed of perfectly bonded graphene nanoplatelet-reinforced composite (GPLRC) layers in which graphene nanoplatelets (GPLs) are randomly oriented and uniformly dispersed in each layer with the weight fraction varying layerwise across the beam thickness. The modified Halpin-Tsai micromechanics model that takes into account the GPL geometry and dimension is used to estimate the effective Young's modulus of GPLRC layers. Within the framework of the first-order shear deformation theory, the governing equations are derived by applying the principle of virtual displacements and then solved by using the differential quadrature method. A comprehensive parametric study is conducted to examine the effects of distribution pattern, concentration, and geometry of GPL, applied voltage, as well as temperature change on the static bending of functionally graded multilayer GPLRC beams. Numerical results show that the bending deflection can be suppressed by applying a negative voltage and distributing more GPLs near the surface layers of the beam. In addition, the effects of GPL weight fraction, aspect ratio and width-to-thickness ratio on the static bending behaviour is highly sensitive to the temperature rise.

Keywords: Thermo-electro-mechanical bending; Graphene nanocomposite; Functionally graded beam; First-order shear deformation theory; Differential quadrature method

Introduction

Graphene [1], a two-dimensional single layer of carbon atoms, has garnered substantial academic and industrial interest due to its superior thermal, electrical and mechanical properties [2-4]. Compared to carbon nanotubes (CNTs), it has comparable tensile strength and Young's modulus [4], but a much larger surface area of up to 2630 m^2g^{-1} [5]. These remarkable physical properties, together with chemical functionalization capacities, make graphene a promising reinforcement material while improving the mechanical properties of polymeric materials. As a result, polymer nanocomposites reinforced with graphene and its derivatives have recently become an emerging research field of extensive efforts [6, 7].

Rafiee *et al* [8] measured and compared the mechanical properties of epoxy nanocomposites reinforced with 0.1wt% graphene platelets (GPLs) and CNTs, respectively. They found that

the Young's modulus, tensile strength and fracture toughness of graphene nanocomposites are observably higher than those of pure epoxy and GPLs are superior in mechanical properties enhancement to CNTs. Following that, Rafiee and his coauthors [9] experimentally investigated the buckling of graphene/epoxy nanocomposite beam structures and observed that the critical buckling load can be significantly increased with addition of only 0.1wt% of GPLs into the epoxy matrix. Most recently, Yang and his coauthors studied the buckling and postbuckling [10], and dynamic instability [11] of functionally graded multilayer graphene nanocomposite beams. Their results revealed that the incorporation of a low content of GPLs into epoxy can significantly increase the buckling and postbuckling strength and reduce the principal unstable region. Feng et al. analyzed the nonlinear bending [12] and free vibration [13] of functionally graded polymer composite beams reinforced with GPLs. They suggested that placing more GPLs with larger surface area and fewer single-layer graphene sheets near the surfaces of the beam is the most effective way to reduce the bending deflection and increase the natural frequencies. Shen *et al* [14] carried out nonlinear bending and thermal postbuckling analyses of functionally graded graphene-reinforced composite laminated beams resting on an elastic foundation and found that the functionally graded graphene reinforcement and temperature variation have pronounced impacts on the nonlinear bending and thermal postbuckling behaviours. Kitipornchai and his coauthors [15, 16] demonstrated that the incorporation of a small amount of GPLs can remarkably reinforce the stiffness of functionally graded porous beams. Smart composite structures, in which the piezoelectric materials are used as actuator or sensors, are of great practical importance for the mechanical behaviour control and structural health monitoring [17, 18]. To the best of the authors' knowledge, no prior work has been done on the mechanical response of piezoelectric graphene nanocomposite structures that are subjected to a combined action of uniform temperature rise, constant actuator voltage and transverse load.

Therefore, this paper is devoted to the analysis of static bending of functionally graded multilayer GPL-reinforced composite (GPLRC) beams integrated with surface-bonded piezoelectric layers under thermo-electro-mechanical loading. Each individual layer is made from a mixture of GPL reinforcements and polymer matrix in which GPLs are randomly oriented and uniformly dispersed. The Young's modulus of GPLRC layers is estimated by the modified Halpin-Tsai model. The linear governing equations are derived based on the first-order shear deformation theory (FSDT) and then converted into a system of linear algebraic equations by using differential quadrature (DQ) method, from which the bending deflection can be determined. Numerical results are presented for piezoelectric multilayer GPLRC beams with different GPL distribution pattern, weight fraction and geometry to explore how these parameters affect the thermo-electro-mechanical bending behaviour. The effects of temperature variation, actuator voltage, slenderness ratio as well as boundary conditions are also discussed in detail.

Theoretical Modelling

Consider a piezoelectric multilayer GPLRC hybrid beam of length L and total thickness H. The hybrid beam consists of two piezoelectric layers of equal thickness h_p that are symmetrically bonded to the top and bottom surfaces of a host beam of thickness h that is composed of perfectly bonded GPLRC layers. It is assumed that GPL reinforcements are uniformly dispersed in each individual layer but its volume fraction has a layer-wise variation across the beam thickness direction. For simplicity, the multilayer GPLRC host beam comprised of an even number of layers is considered here. The GPL volume fraction of the k^{th} layer for different distribution patterns are governed by [10, 11]

U-GPLRC: $V_{\mathrm{GPL}}^{(k)} = V_{\mathrm{GPL}}^{*}$, (1a)

X-GPLRC: $V_{\mathrm{GPL}}^{(k)} = 2V_{\mathrm{GPL}}^{*}\left|2k - N_{\mathrm{L}} - 1\right|/N_{\mathrm{L}}$, (1b)

O-GPLRC: $V_{\mathrm{GPL}}^{(k)} = 2V_{\mathrm{GPL}}^{*}\left(1 - \left|2k - N_{\mathrm{L}} - 1\right|/N_{\mathrm{L}}\right)$, (1c)

A-GPLRC: $V_{\mathrm{GPL}}^{(k)} = V_{\mathrm{GPL}}^{*}\left(2k-1\right)/N_{\mathrm{L}}$, (1d)

where N_{L} is the total number of layers of the GPLRC host beam and k = 1, 2, …, N_{L}. V_{GPL}^{*} is the total GPL volume fraction determined from

$$V_{\mathrm{GPL}}^{*} = \frac{W_{\mathrm{GPL}}}{W_{\mathrm{GPL}} + \left(\rho_{\mathrm{GPL}}/\rho_{\mathrm{m}}\right)\left(1 - W_{\mathrm{GPL}}\right)}, \tag{2}$$

in which W_{GPL} is the total GPL weight fraction; ρ_{GPL} and ρ_{m} are mass densities of the GPL and matrix, respectively.

The modified Halpin-Tsai model that takes into account the effects of GPL geometry and dimension is used to estimate the effective Young's modulus E_{c} of GPLRC [19-21]:

$$E_{\mathrm{c}} = \frac{3}{8}\frac{1 + \xi_{\mathrm{L}}\eta_{\mathrm{L}}V_{\mathrm{GPL}}}{1 - \eta_{\mathrm{L}}V_{\mathrm{GPL}}} \times E_{\mathrm{m}} + \frac{5}{8}\frac{1 + \xi_{\mathrm{T}}\eta_{\mathrm{T}}V_{\mathrm{GPL}}}{1 - \eta_{\mathrm{T}}V_{\mathrm{GPL}}} \times E_{\mathrm{m}}, \tag{3}$$

where η_{L} and η_{T} are expressed as

$$\eta_{\mathrm{L}} = \frac{\left(E_{\mathrm{GPL}}/E_{\mathrm{m}}\right) - 1}{\left(E_{\mathrm{GPL}}/E_{\mathrm{m}}\right) + \xi_{\mathrm{L}}}, \ \eta_{\mathrm{T}} = \frac{\left(E_{\mathrm{GPL}}/E_{\mathrm{m}}\right) - 1}{\left(E_{\mathrm{GPL}}/E_{\mathrm{m}}\right) + \xi_{\mathrm{T}}}, \tag{4}$$

in which E_{GPL} and E_{m} are the Young's moduli of the GPL and matrix, respectively. The GPL geometry factors ξ_{L} and ξ_{T} are defined as

$$\xi_{\mathrm{L}} = 2\left(a_{\mathrm{GPL}}/t_{\mathrm{GPL}}\right) = 2\left(a_{\mathrm{GPL}}/b_{\mathrm{GPL}}\right) \times \left(b_{\mathrm{GPL}}/t_{\mathrm{GPL}}\right), \ \xi_{\mathrm{T}} = 2\left(b_{\mathrm{GPL}}/t_{\mathrm{GPL}}\right), \tag{5}$$

in which a_{GPL}, b_{GPL} and t_{GPL} are the length, width and thickness of GPL nanofillers, respectively.

According to the rule of mixture, the thermal expansion coefficient α_{c} and Poisson's ratio ν_{c} of GPLRC are expressed as

$$\alpha_{\mathrm{c}} = \alpha_{\mathrm{m}}V_{\mathrm{m}} + \alpha_{\mathrm{GPL}}V_{\mathrm{GPL}}, \ \nu_{\mathrm{c}} = \nu_{\mathrm{m}}V_{\mathrm{m}} + \nu_{\mathrm{GPL}}V_{\mathrm{GPL}}, \tag{6}$$

where α_{GPL} and α_{m} are the thermal expansion coefficients, and ν_{GPL} and ν_{m} are the Poisson's ratios, with the subscripts "GPL" and "m" referring to the GPL and matrix, respectively. V_{GPL} and V_{m} are the volume fractions, related by $V_{\mathrm{m}} + V_{\mathrm{GPL}} = 1$.

Based on the FSDT, the displacement field of the beam takes the form of

$$\overline{U}\left(x,z\right) = U(x) + z\psi(x), \ \overline{W}\left(x,\ z\right) = W(x), \tag{7}$$

in which U and W are the longitudinal and transverse displacement components in the mid-plane ($z = 0$); ψ is the angle of rotation of the normal to the mid-plane. In view of Eq. (7) and using the principle of virtual displacements, the linear governing equations of the piezoelectric multilayer GPLRC beam under a combined loading of a uniform temperature rise ΔT, a constant actuator voltage V_z and a transverse load $\bar{q}$ can be derived and expressed in terms of displacements as

$$A_{11}U_{,xx} + B_{11}\psi_{,xx} = 0\,, \tag{8a}$$

$$\kappa A_{55}\left(W_{,xx} + \psi_{,x}\right) - \left(N_x^{\mathrm{T}} + N_x^{\mathrm{E}}\right)W_{,xx} + \bar{q} = 0\,, \tag{8b}$$

$$B_{11}U_{,xx} + D_{11}\psi_{,xx} - \kappa A_{55}\left(W_{,x} + \psi\right) = 0\,, \tag{8c}$$

where the comma denotes the partial derivative with respect to the corresponding coordinate; $\kappa = 5/6$ is the shear correction coefficient. The k^{th} GPLRC layer is located between the points z_k and z_{k+1} in the thickness direction. The stiffness components are defined as

$$\left\{A_{11},\, B_{11},\, D_{11}\right\} = \int_{-H/2}^{-h/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}\{1,\, z,\, z^2\}\mathrm{d}z + \sum_{k=1}^{N_L}\int_{z_k}^{z_{k+1}} \frac{E_{\mathrm{c}}^{(k)}}{1-\left(\nu_{\mathrm{c}}^{(k)}\right)^2}\{1,\, z,\, z^2\}\mathrm{d}z + \int_{h/2}^{H/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}\{1,\, z,\, z^2\}\mathrm{d}z\,, \tag{9a}$$

$$A_{55} = \int_{-H/2}^{h/2} \frac{E_{\mathrm{p}}}{2\left(1+\nu_{\mathrm{p}}\right)}\{1,\, z,\, z^2\}\mathrm{d}z + \sum_{k=1}^{N_L}\int_{z_k}^{z_{k+1}} \frac{E_{\mathrm{c}}^{(k)}}{2\left(1+\nu_{\mathrm{c}}^{(k)}\right)}\{1,\, z,\, z^2\}\mathrm{d}z + \int_{\mathrm{h}/2}^{H/2} \frac{E_{\mathrm{p}}}{2\left(1+\nu_{\mathrm{p}}\right)}\{1,\, z,\, z^2\}\mathrm{d}z\,. \tag{9b}$$

The thermally and electrically forces and moments are given by

$$\left\{N_x^{\mathrm{T}},\, M_x^{\mathrm{T}}\right\} = \int_{-H/2}^{-h/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}\alpha_{\mathrm{p}}\Delta T\left\{1,\, z\right\}\mathrm{d}z + \sum_{k=1}^{N_L}\int_{z_k}^{z_{k+1}} \frac{E_{\mathrm{c}}^{(k)}}{1-\left(\nu_{\mathrm{c}}^{(k)}\right)^2}\alpha_{\mathrm{c}}^{(k)}\Delta T\left\{1,\, z\right\}\mathrm{d}z + \int_{h/2}^{H/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}\alpha_{\mathrm{p}}\Delta T\left\{1,\, z\right\}\mathrm{d}z\,, \tag{10a}$$

$$\left\{N_x^{\mathrm{E}},\, M_x^{\mathrm{E}}\right\} = \int_{-H/2}^{-h/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}d_{31}E_{\mathrm{z}}\left\{1,\, z\right\}\mathrm{d}z + \int_{h/2}^{H/2} \frac{E_{\mathrm{p}}}{1-\nu_{\mathrm{p}}^{\,2}}d_{31}E_{\mathrm{z}}\left\{1,\, z\right\}\mathrm{d}z\,, \tag{10b}$$

in which superscripts "T" and "E" represent the terms contributed by thermal and electrical loadings, respectively. $E_{\mathrm{c}}^{(k)}$, $\nu_{\mathrm{c}}^{(k)}$ and $\alpha_{\mathrm{c}}^{(k)}$ are the values of E_{c}, ν_{c} and α_{c} of the k^{th} GPLRC layer, respectively. E_{p}, ν_{p}, α_{p} and d_{31} are the Young's modulus , Poisson's ratio, thermal expansion coefficient and piezoelectric strain constant of the piezoelectric layer. E_{z} is the electric filed intensity generated by the actuator voltage V_{z}.

The piezoelectric multilayer GPLRC beam is assumed to be either clamped or hinged at the ends ($x = 0, L$), and the associated boundary conditions require

Clamped (C): $U = 0,\ W = 0,\ \psi = 0,$ (11a)

Hinged (H): $U = 0,\ W = 0,\ B_{11}U_{,x} + D_{11}\psi_{,x} - M_x^T = 0\,.$ (11b)

Solution Method

By introducing the following dimensionless quantities:

$$\begin{aligned}&\zeta = x/L,\ \{u, w\} = \{U, W\}/H,\ \varphi = \psi,\ \left\{P^{\mathrm{T}}, M^{\mathrm{T}}\right\} = \left\{N_x^{\mathrm{T}}, M_x^{\mathrm{T}}/H\right\}/A_{110},\\ &\eta = L/H,\ q = \bar{q}L^2/A_{110}H,\ \{a_{11}, a_{55}, b_{11}, d_{11}\} = \left\{A_{11}, \kappa A_{55}, B_{11}/h, D_{11}/H^2\right\}/A_{110},\end{aligned} \tag{12}$$

where A_{110} is the value of A_{11} of a homogeneous beam made from the pure matrix material, the governing equations (8) and associated boundary conditions in Eq. (11) can be rewritten in the dimensionless form as

$$a_{11}u_{,\zeta\zeta} + b_{11}\varphi_{,\zeta\zeta} = 0, \tag{13a}$$

$$a_{55}\left(w_{,\zeta\zeta} + \eta\varphi_{,\zeta}\right) - \left(P^{\mathrm{T}} + P^{\mathrm{E}}\right)w_{,\zeta\zeta} + q = 0, \tag{13b}$$

$$b_{11}u_{,\zeta\zeta} + d_{11}\varphi_{,\zeta\zeta} - \eta a_{55}\left(w_{,\zeta\zeta} + \eta\varphi\right) = 0; \tag{13c}$$

Clamped (C): $u = 0$, $w = 0$, $\varphi = 0$, (14a)

Hinged (H): $u = 0$, $w = 0$, $b_{11}u_{,\zeta} + d_{11}\varphi_{,\zeta} - M^{\mathrm{T}} = 0$. (14b)

According to DQ discretization rule, the unknown displacements and their j^{th} derivatives with respect to x can be approximated as

$$\{u, w, \varphi\} = \sum_{m=1}^{N} l_m(\zeta)\{u_m, w_m, \varphi_m\}, \text{ and } \frac{\partial^j}{\partial\zeta^j}\{u, w, \varphi\}\bigg|_{\zeta=\zeta_i} = \sum_{m=1}^{N} C_{im}^{(j)}\{u_m, w_m, \varphi_m\}, \tag{15}$$

where $l_m(\zeta)$ is the Lagrange interpolation polynomials; $\{u_m, w_m, \varphi_m\}$ are the values of $\{u, w, \varphi\}$ at $\zeta = \zeta_m$; $C_{im}^{(j)}$ is the weighting coefficient of the j^{th} partial derivative and can be calculated from the recursive formula given by Yang *et al* [22]. N is the total number of grid points distributed along the ζ-axis according to a cosine pattern [23].

Keeping Eq. (10) in mind and applying Eq. (15) to the dimensionless governing equations (13) and boundary conditions (14) leads to a system of linear algebraic equations that governs the thermo-electro-mechanical bending of piezoelectric multilayer GPLRC beams and can be expressed in the matrix form as

$$\left(\mathbf{K}_{\mathrm{L}} - \Delta T\mathbf{K}_{\mathrm{T}} - V_{\mathrm{z}}\mathbf{K}_{\mathrm{E}}\right)\mathbf{d} = \mathbf{q} \tag{16}$$

in which **d** implies the unknown displacement vector that is composed of u_i, w_i, φ_i (i = 1, 2, …, N); **q** stands for the transverse load vector; $\mathbf{K}_{\mathrm{L}}$, $\mathbf{K}_{\mathrm{T}}$ and $\mathbf{K}_{\mathrm{E}}$ are $3N\times3N$ constant coefficient matrices. The static bending deflection of piezoelectric multilayer GPLRC beams under thermo-electro-mechanical loading can be obtained by solving Eq. (16).

Results and Discussion

In what follows, the multilayer GPLRC host of thickness h = 0.02 m is made from a mixture of epoxy and GPLs with a length of a_{GPL} = 2.5 μm, width of b_{GPL} = 1.5 μm and thickness of t_{GPL} = 1.5 nm [8]. Their material properties are assumed to be temperature-independent and can be found in [11]. The piezoelectric layers are made of PZT G1195N with E_{p} = 63.0 GPa, ν_{p} = 0.3, α_{p} =0.9×10^{-6} /K and d_{31} = 2.54×10^{-10} m/V. Three typical transverse loads are considered and take the following forms:

$q = q_0$ for uniformly distributed load, (17a)

$q = q_0 \sin(\pi\zeta)$ for sinusoidal load, (17b)

$q = q_0$ at $\zeta = 0.5$, otherwise $q = 0$ for point load, (17c)

where q_0 is the dimensionless transverse load amplitude. The thermo-electro-mechanical bending results are given in the dimensionless form as $w_m = 10^3 W_m E_m / q_0 h$ where w_m is the dimensionless mid-span deflection and W_m (in m) is the transverse displacement at the beam center (i.e. $x = L/2$).

Convergence and comparison studies are first carried out to ensure the accuracy of the present analysis. Table 1 compares the static bending deflection results for the piezoelectric multilayer X-GPLRC beam with varying number of grid points and GPLRC layers. It is seen that convergent results are obtained at N = 7 and N_L = 30. Nonetheless, the tiny difference between the results at N_L = 10 and N_L = +∞ implies that a multilayer GPLRC beam with 10 layers is sufficiently accurate to model an ideal functionally graded beam with smooth variation in both material composition and properties. Take into consideration of the manufacturing difficulties, N_L = 10, as well as N = 9 are used in all the following numerical computations.

The dimensionless mid-span deflections for cross-ply laminated beams are calculated and compared in Table 2 with the analytical results given by Khdeir and Reddy [24]. In addition, the results for uniformly distributed carbon nanotube-reinforced composite (UD-CNTRC) beams under either uniform or sinusoidal transverse load are also given and compared in Table 3 with those of Wattanasakulpong and Ungbhakorn [25]. Good agreements are observed between our results and those reported in the literature.

Table 1 Convergence of dimensionless mid-span deflection w_m for a C-C piezoelectric X-GPLRC beam under a uniform transverse load (W_{GPL} = 0.3%, L/H = 30, h/h_p = 20, ΔT = 0 K, V_z = -200 V).

N (N_L = 10)	w_m	N_L (N = 9)	w_m
3	0.096	4	3.349
5	3.314	6	3.324
7	3.312	10	3.312
9	3.312	20	3.307
11	3.312	30	3.305
13	3.312	+∞	3.305

Table 2 Comparison of dimensionless mid-span deflections for symmetric cross-ply (0°/90°/0°) beams under uniform transverse load

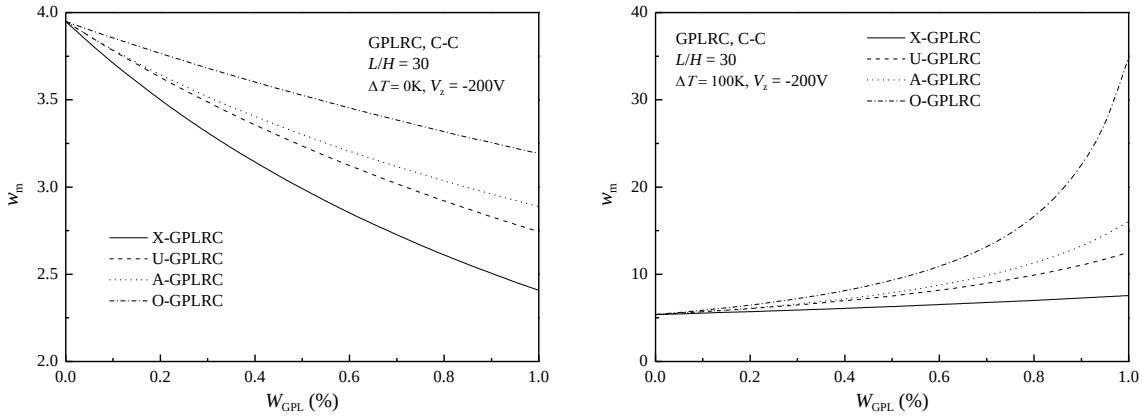
L/H	C-C		C-H		H-H	
	Present	Ref. [24]	Present	Ref. [24]	Present	Ref. [24]
5	1.6293	1.629	1.9216	1.922	2.1464	2.146
10	0.5043	0.504	0.6931	0.693	1.0214	1.021
50	0.1443	0.144	0.2764	0.276	0.6614	0.661

Table 3 Comparison of dimensionless mid-span deflections for H-H UD-CNTRC beams under either uniform or sinusoidal transverse load (*L*/*H* = 20).

Source	Uniform load			Sinusoidal load		
	V_{cn}^{*} = 0.12	V_{cn}^{*} = 0.17	V_{cn}^{*} = 0.28	V_{cn}^{*} = 0.12	V_{cn}^{*} = 0.17	V_{cn}^{*} = 0.28
Present	0.4608	0.3068	0.2028	0.3651	0.2429	0.1608
Ref. [25]	0.461	0.307	0.203	0.365	0.243	0.161

Having validated the formulations and solution method, the thermo-electro-mechanic bending results of piezoelectric multilayer GPLRC beams are presented in Figs. 1-4 and Tables 4 and 5. Unless otherwise stated, the following numerical results are for C-C piezoelectric X-GPLRC beams with W_{GPL} = 0.3%, L/H = 30, h/h_p = 20 subjected to a uniformly distributed transverse load.

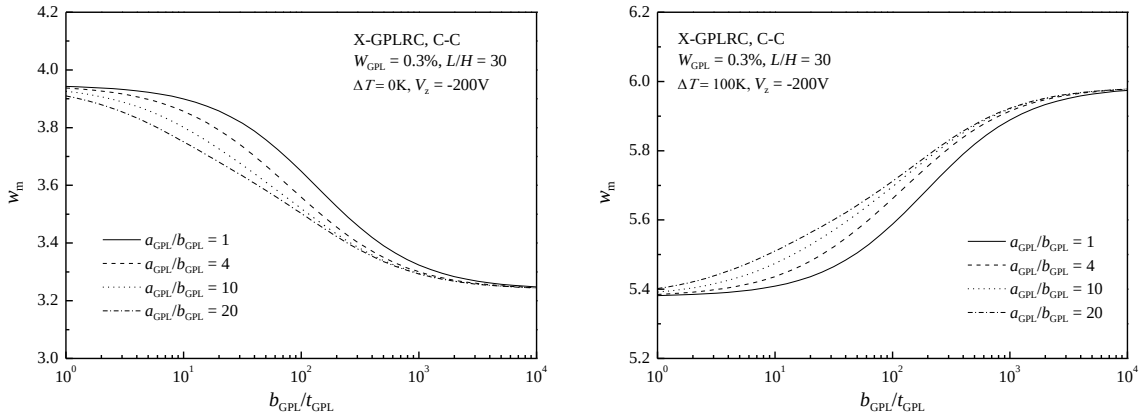
Fig. 1 illustrates the effects of GPL distribution pattern and weight fraction on the thermo-electro-mechanical bending of piezoelectric GPLRC beams under uniform transverse load at ΔT = 0 and 100 K. The results shows that the mid-span deflection at ΔT = 0 K decreases as the GPL weight fraction increases, while this is inversed at ΔT = 100 K where the mid-span deflection becomes prominently lager with an increase in GPL weight fraction. This is because that a higher GPL content results in a higher compressive thermal stress in the beam when subjected a high temperature rise and thereby dramatically lower the beam stiffness. This indicates that the effect of GPL weight fraction on the static bending behaviour is sensitive to the temperature rise. However, the hybrid beam comprised of an X-GPLRC host has the smallest deflection at both ΔT = 0 and 100 K, followed by the beams with U-, A- and O-GPLRC hosts.



(a) (b)

Fig. 1. Effects of GPL distribution pattern and weight fraction on the dimensionless mid-span deflection of piezoelectric GPLRC beams: (a) ΔT = 0 K; (b) ΔT = 100 K.

The influences of GPL geometry and dimension on the thermo-electro-mechanical bending of piezoelectric X-GPLRC beams are examined in Fig. 2 where the a higher value of a_{GPL}/b_{GPL} represents a lager GPL surface area and a greater magnitude of b_{GPL}/t_{GPL} signifies that a GPL contains fewer monolayer graphene sheets. Similarly to the observations in Fig. 1, the mid-span deflection increases at ΔT = 0 K but decreases at ΔT = 100 K as both a_{GPL}/b_{GPL} and b_{GPL}/t_{GPL} increase. It is worthy to note that the effects of a_{GPL}/b_{GPL} and b_{GPL}/t_{GPL} get much less pronounced and the bending deflection tend to be convergent when b_{GPL}/t_{GPL} is greater than 10^3.



(a) (b)

Fig. 2. Effects of GPL geometry and dimension on the dimensionless mid-span deflection of piezoelectric X-GPLRC beams: (a) ΔT = 0 K; (b) ΔT = 100 K.

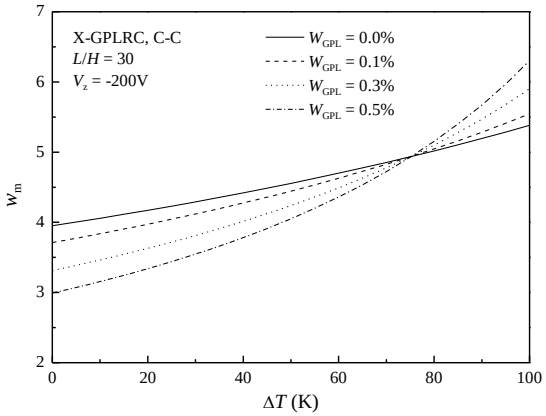
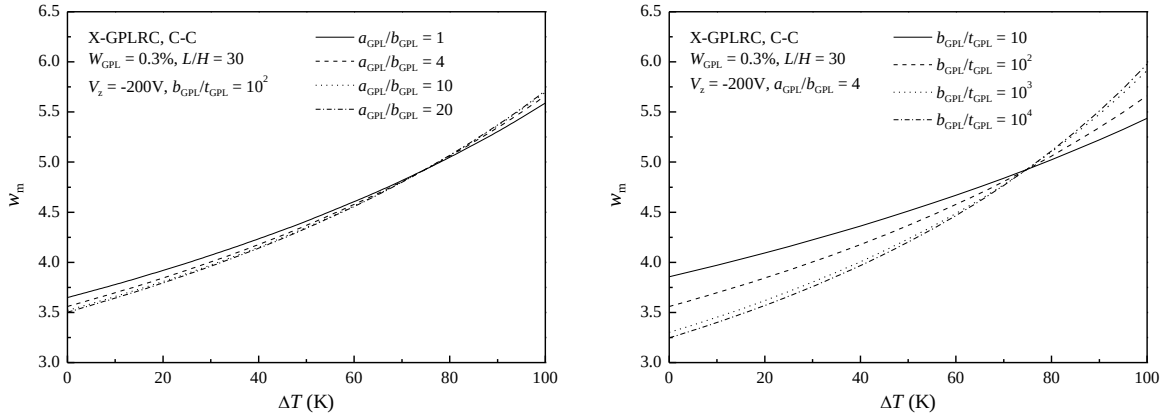


Fig. 3. Dimensionless mid-span deflection versus temperature rise for piezoelectric X-GPLRC beams with different GPL weight fractions.



(a) (b)

Fig. 4. Dimensionless mid-span deflection versus temperature rise for piezoelectric X-GPLRC beams with different GPL (a) aspect ratios and (b) width-to-thickness ratios.

The influence of temperature rise is further studied in Figs. 3 and 4 where the dimensionless mid-span deflection versus temperature rise curves for piezoelectric X-GPLRC beams with different GPL weight fractions and geometries, respectively, are compared. As demonstrated by the figures, the bending deflection is monotonically increased with an increase in

temperature, but is reduced as the GPL weight fraction, aspect and width-to-thickness ratios increase when the temperature rise ΔT is less than 75 K, beyond which the deflection is raised.

Table 4 tabulates the dimensionless mid-span deflections for electrically prestressed piezoelectric X-GPLRC beams with different slenderness ratios and boundary conditions. Compared to the results obtained at V_z = 0 V, the mid-span deflection is decreased at V_z = -200 V but is increased at V_z = 200 V, which indicates that the bending deflection can be suppressed by applying a negative voltage on the piezoelectric layers. This is due to the fact that a negative voltage makes the beam contract and consequently weakens the beam stiffness. In addition, the effect of actuator voltage tends to be more noticeable for the beam with a larger slenderness ratio and soft end supports (i.e. hinged ends).

Table 4. Effect of actuator voltage on the dimensionless mid-span deflection for piezoelectric X-GPLRC beams with different slenderness ratios and boundary conditions (W_{GPL} = 0.3%, ΔT = 0 K).

L/H	C-C			H-H		
	V_z = 200 V	V_z = 0 V	V_z = -200 V	V_z = 200 V	V_z = 0 V	V_z = -200 V
20	0.681	0.680	0.679	3.253	3.231	3.210
30	3.337	3.324	3.312	16.49	16.24	16.00
40	10.44	10.37	10.31	52.59	51.19	49.86

Table 5 lists the static bending results for the piezoelectric GPLRC beams subjected to three different transverse loads with the same amplitude q_0. As expected, the uniformly distributed transverse load leads to the highest mid-span deflection. The results also confirm that among the four GPL distribution patterns, pattern X yields the smallest bending deflection, which means pattern X makes the best use of GPL reinforcements, followed by the patterns U, A and O in order. However, whether the beam with a GPLRC host has a higher or lower bending deflection than the beam with a pure epoxy host is highly dependent on the temperature rise.

Table 5. Effect of load type on the dimensionless mid-span deflection for C-C piezoelectric GPLRC beams with different GPL distribution patterns (W_{GPL} = 0.3%, L/H = 30, V_z = -200 V).

Temperature rise	Load type	Pure epoxy	X-GPLRC	U-GPLRC	A-GPLRC	O-GPLRC
ΔT = 0 K	uniform	3.950	3.312	3.487	3.518	3.682
	sinusoidal	3.337	2.798	2.946	2.973	3.111
	point	1.235	1.036	1.090	1.100	1.151
ΔT = 100 K.	uniform	5.383	5.902	6.488	6.595	7.200
	sinusoidal	4.553	4.996	5.494	5.584	6.097
	point	1.685	1.849	2.033	2.067	2.257

Conclusions

The static bending of piezoelectric functionally graded GPLRC beams under thermo-electro-mechanical loading is investigated based on FSDT. The differential quadrature method is used to obtain piezothermoelastic solutions for the hybrid beams subjected to different types of transverse load. Comprehensive numerical results are presented in both graphical and tabular forms to examine the effects of GPL distribution pattern, weight fraction, and geometry, actuator voltage, as well as temperature rise on the thermo-electro-mechanical bending behaviour. The results show that the effects of GPL content and geometry are highly sensitive

to the temperature rise. Among the four GPL distribution patterns, pattern X with more GPLs distributed near the surface layers gives the smallest bending deflection. In addition, the effects of GPL geometry and dimension tend to be much less pronounced when the GPL width-to-thickness ratio is larger than 10^3. Results also indicate that the static bending deflection can be suppressed by applying a negative voltage on the piezoelectric layers, and this effect is more noticeable for the hinged beam with a greater slenderness ratio.

Acknowledgements

The work described in this paper was fully funded by research grants from the Australian Research Council under Discovery Project scheme (DP140102132, DP160101978) and Linkage Project scheme (LP150100103, LP140100747). The authors are grateful for their financial support.

References

[1] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V, Firsov A A 2004 Electric field effect in atomically thin carbon films *Science* **306** 666-669
[2] Du X, Skachko I, Barker A, Andrei E Y 2008 Approaching ballistic transport in suspended graphene *Nat Nanotechnol* **3** 491-495
[3] Balandin A A, Ghosh S, Bao W, Calizo I, Teweldebrhan D, Miao F, Lau C N 2008 Superior thermal conductivity of single-layer graphene *Nano Lett* **8** 902-907
[4] Lee C, Wei X, Kysar J W, Hone J 2008 Measurement of the elastic properties and intrinsic strength of monolayer graphene *Science* **321** 385-388
[5] Bonaccorso F, Colombo L, Yu G, Stoller M, Tozzini V, Ferrari A C, Ruoff R S, Pellegrini V 2015 Graphene, related two-dimensional crystals, and hybrid systems for energy conversion and storage *Science* **347** 1246501
[6] Young R J, Kinloch I A, Gong L, Novoselov K S 2012 The mechanics of graphene nanocomposites: A review *Compos Sci Technol* **72** 1459-1476
[7] Kim H, Abdala A A, Macosko C W 2010 Graphene/polymer nanocomposites *Macromolecules* **43** 6515-6530
[8] Rafiee M A, Rafiee J, Wang Z, Song H, Yu Z-Z, Koratkar N 2009 Enhanced mechanical properties of nanocomposites at low graphene content *ACS Nano* **3** 3884-3890
[9] Rafiee M, Rafiee J, Yu Z-Z, Koratkar N 2009 Buckling resistant graphene nanocomposites *Appl Phys Lett* **95** 223103
[10] Yang J, Wu H, Kitipornchai S 2017 Buckling and postbuckling of functionally graded multilayer graphene platelet-reinforced composite beams *Compos Struct* **161** 111-118
[11] Wu H, Yang J, Kitipornchai S 2017 Dynamic instability of functionally graded multilayer graphene nanocomposite beams in thermal environment *Compos Struct* **162** 244-254
[12] Feng C, Kitipornchai S, Yang J 2017 Nonlinear bending of polymer nanocomposite beams reinforced with non-uniformly distributed graphene platelets (GPLs) *Compos Part B-Eng* **110** 132-140
[13] Feng C, Kitipornchai S, Yang J 2017 Nonlinear free vibration of functionally graded polymer composite beams reinforced with graphene nanoplatelets (GPLs) *Eng Struct* **140** 110-119
[14] Shen H-S, Lin F, Xiang Y 2017 Nonlinear bending and thermal postbuckling of functionally graded graphene-reinforced composite laminated beams resting on elastic foundations *Eng Struct* **140** 89-97
[15] Kitipornchai S, Chen D, Yang J 2017 Free vibration and elastic buckling of functionally graded porous beams reinforced by graphene platelets *Mater Design* **116** 656-665
[16] Chen D, Yang J, Kitipornchai S 2017 Nonlinear vibration and postbuckling of functionally graded graphene reinforced porous nanocomposite beams *Compos Sci Technol* **142** 235-245
[17] Trollope J E, Burnham K J 2013 Active Buckling Control for Future Lightweight Vehicle Body Structures *Measurement and Control* **46** 315-320
[18] Zhang L, Song Z, Liew K 2016 Optimal shape control of CNT reinforced functionally graded composite plates using piezoelectric patches *Compos Part B-Eng* **85** 140-149
[19] Halpin J, Kardos J 1976 The Halpin-Tsai equations: a review *Polym Eng Sci* **16** 344-352
[20] Derek H. An Introduction to Composite Materials: Cambridge University Press, 1981.
[21] Harris B. Engineering Composite Materials: Institute of metals London, 1986.

[22] Yang J, Kitipornchai S, Liew K 2004 Non-linear analysis of the thermo-electro-mechanical behaviour of shear deformable FGM plates with piezoelectric actuators *Int J Numer Method Eng* **59** 1605-1632
[23] Wu H, Kitipornchai S, Yang J 2015 Free vibration and buckling analysis of sandwich beams with functionally graded carbon nanotube-reinforced composite face sheets *Int J Struct Stab Dy* **15** 1540011
[24] Khdeir A A, Reddy J N 1997 An exact solution for the bending of thin and thick cross-ply laminated beams *Compos Struct* **37** 195-203
[25] Wattanasakulpong N, Ungbhakorn V 2013 Analytical solutions for bending, buckling and vibration responses of carbon nanotube-reinforced composite beams resting on elastic foundation *Comp Mater Sci* **71** 201-208

In-plane free vibration of circular and annular FG disks

†*Y. Yang[1], K.P. Kou[1], C.C. Lam[1]

[1]Department of Civil and Environmental Engineering, University of Macau, Macau, China.

*Presenting and †Corresponding author: yangyang.liju@hotmail.com

Abstract

Analysis of the in-plane free vibration of the circular and annular functionally graded disks by a meshfree boundary-domain integral equation method are presented in this paper. The material properties of the disks are assumed to vary in the radial direction obeying an exponential law. Based on the two-dimensional linear elastic theory, the motion equations of the FG disks are derived by using the static fundamental solutions. Radial integration method as an efficient tool is adopted to treat the domain integrals which raised due to the material inhomogeneous and inertial effects. The natural frequencies and associate mode shapes are calculated for the FG disks with combinations of free and clamped boundary conditions. Parametric studies are also conducted to study the effects of the material gradients, radius ratios and boundary conditions on the frequency of the FG disks.

Keywords: In-plane free vibration, circular and annular FG disks, meshfree boundary-domain integral equation method

Introduction

Circular and annular structures are structural components and commonly used in a wide variety of engineering applications including space structures, electronic components and rotating machinery. A lot of attentions have been focus on the free vibration characteristics of circular structures. For example, Weisensel [1] given the results of an extensive literature search and review of available source of numerical natural frequency information for stationary circular and annular elastic plates, where the information regarding the specific plate theory, boundary conditions, geometric properties and material properties used to determine the natural frequency information. Kirkhope and Wilson [2] applied the finite element method to the stress and vibration analysis of thin rotating discs. Han and Liew [3] presented a numerical analysis of the axisymmetric free vibration of moderately thick annular plates using the differential quadrature method (DQM). Chung et al. [4] derived the governing equation for free vibration of a spinning circular disk by using the variational formulation based upon the Kirchhoff plate theory and von Karman strain one, this was because they found that during the derivation that the governing equation was theoretically valid under the assumption that in-plane deflections were steady and axisymmetric, and that internal forces were linearized while the strains remain nonlinear.

In-plane vibration characteristics are also very important for the circular annular structures. As the circular structures in most applications have direct in-plane forces or in-plane force components due to imperfections in the manufacturing, assembly or alignment of the supporting mounts. However, the studies reported on the in-plane vibrations of circular annular disks is relatively scarce. Ambati [5] carried out the in-plane vibrations of annular rings. Nigh and Olson [6] used a finite element formulation for the analysis of rotating disks in either a body-fixed or a space-fixed co-ordinate system. The in-plane stress distribution resulting from the in-plane body force due to rotation was determined first by a plane stress

finite element analysis. Farag and Pan [7] analyzed the modal characteristics of in-plane vibrations of a solid disk with clamped outer edge. Park [8] derived the frequency equation for the in-plane vibration of the clamped circular plate of uniform thickness by using Hamilton's principle. The in-plane free vibration of an elastic and isotropic disk was studied by Bashmal et al. [9] on the basis of the two-dimensional linear plane stress theory of elasticity. rie et al. [10] examined the in-plane vibrations in circular and annular disks using transfer matrix formulation. Natural frequencies were obtained for several radius ratios of annular disks with combinations of free and clamped conditions at the inner and outer edges but mode shapes were not presented.

With the increased application of the FG structures, vibration analyses of the FG circular and annular structures have attracted intensive research. Based on the first-order shear deformation theory, Francesco Tornabene [11] presented the dynamic behavior of moderately thick FG conical, cylindrical shells and annular plates, which material properties were graded through the thickness direction. Kermani et al. [12] analyzed the three-dimensional free vibrations of multi-directional graded circular and annular plates by the state space based differential quadrature method, which solved dimensionless equations of motion analytically along thickness direction and numerically along radial direction of the plate. Based on the three-dimensional theory of elasticity, the free and forced vibration analysis of FG circular plate with various boundary conditions was carried out by Nie and Zhong [13], the material properties were assumed to be graded in the thickness direction according to an exponential distribution. Three-dimensional free vibration analysis of FG annular plate were done by Dong [14] using the Chebyshev-Ritz method. The material properties also only varied in the thickness direction. The free vibration analysis of FG thick annular plates subjected to thermal environment was studied by Malekzadeh [15] based on the 3D elasticity theory, the material properties were assumed to be temperature dependent and graded in the thickness direction. Most of the mentioned reference were based on the three-dimensional analysis, and the material properties were only graded in the thickness direction. The other in-plane vibration of the FG circular and annular disks are very rare in the literature. Therefore, an accurate analysis of the in-plane vibration of the FG circular and annular disks with varied material properties in space coordinates must be considered.

Thus, in-plane free vibration of the circular and annular functionally graded disks were presented in this paper. The material properties of the disks were assumed to grade in the radial direction. Based on the two-dimensional linear elastic theory, the motion equations of the FG disks could be derived by a meshfree boundary-domain integral equation method, and the radial integration method was applied to transform the domain integrals into boundary integrals. Normalized natural frequencies were obtained to compare with data available in the literature. Mode shapes were presented to illustrate the free vibration behavior of the disks.

Geometrical and material properties of circular and annular FG disks

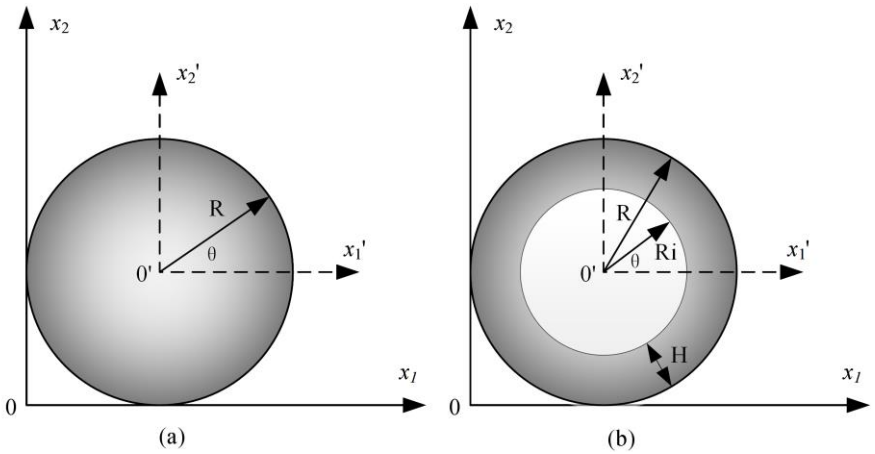


Figure 1 The geometry coordinates of the circular and annular FG disks

The considered circular and annular FG disks are displayed in Fig. 1. R is the radius of the circular disk. For the annular FG disk, the outer and the inner radius are R and R_i respectively, and thickness is H. For commen multi-FG disks, the material properties are a function of polar coordinates ($\check{r}$, θ) and vary continuously in one or more directions. In the present paper, only the material properties of the disks grade in radial directions was considered. The Young's modulus E and the density ρ were assumed to vary continuous along the radial direction obeying an exponential law as shown in Eqs. (1) and (2) and the Poisson's ratio v was taking as a constant [16].

$$E = E_s e^{\beta \check{r}}, \qquad \text{where} \qquad \beta = \frac{1}{R - R_i} ln(\frac{E_e}{E_s}), \tag{1}$$

$$\rho = \rho_s e^{\gamma \check{r}}, \qquad \text{where} \qquad \gamma = \frac{1}{R - R_i} ln(\frac{\rho_e}{\rho_s}), \tag{2}$$

$$\text{where} \qquad \check{r} = \sqrt{(x - x_0)^2 + (y - y_0)^2}, \tag{3}$$

where E_s and ρ_s are the Young's modulus and density at the starting face, E_e and ρ_e are the Young's modulus and mass density at the ending face, β and γ represent the material gradient parameters for Young's modulus and mass density respectively, $\check{r}$ stands for the radial coordinate, x_0 and y_0 are the center Cartesian coordinates of the considered circular and annular disks.

Problem formulation

Based on 2D elasticity theory, considering a FG circular disk, the governing differential equation of the steady-state elastodynamics is expressed as

$$\sigma_{ij,j}(\boldsymbol{x}) + \omega^2 \rho u_i(\boldsymbol{x}) = 0. \tag{4}$$

The stress tensor σ_{ij} and the displacement vector u_i in Eq. (4) are functions of spatial coordinates $\boldsymbol{x}$. By taking the elastostatic displacement fundamental solutions $U_{ij}(\boldsymbol{x}, \boldsymbol{y})$ as the weight function, the weak-form of the equilibrium Eq. (4) can be obtained by

$$\int_\Omega [\sigma_{jk,k} + \omega^2 \rho u_j] \cdot U_{ij} d\Omega = 0. \tag{5}$$

Substituting the generalized Hooke's law $\sigma_{ij} = c_{ijkl}u_{k,l} = \mu(\boldsymbol{x})c^0_{ijkl}u_{k,l}$ and applying the Gauss's divergence theorem, the weak form yields the following boundary-domain integral equations

$$\tilde{u}_i(\boldsymbol{y}) = \int_\Gamma U_{ij}(\boldsymbol{x},\boldsymbol{y})t_j(\boldsymbol{x})d\Gamma - \int_\Gamma T_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Gamma + \int_\Omega V_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Omega + \omega^2\int_\Omega \frac{\rho(\boldsymbol{x})}{\mu(\boldsymbol{x})}U_{ij}(\boldsymbol{x},\boldsymbol{y})\tilde{u}_j(\boldsymbol{x})d\Omega \quad . \tag{6}$$

In Eq. (6), the traction vector $t_i = \sigma_{ij}n_j$ and n_j is the components of the outward unit vector normal to the boundary Γ of the considered layer domain Ω. $\tilde{u}_i$ is the normalized displacement vector associated with the normalized shear modulus $\tilde{\mu}$, as defined by [17]

$$\tilde{u}_i(\boldsymbol{x}) = \mu(\boldsymbol{x})u_i(\boldsymbol{x}), \qquad \tilde{\mu}(\boldsymbol{x}) = \ln[\mu(\boldsymbol{x})]. \tag{7a, b}$$

where the shear modulus $\mu(\boldsymbol{x})$ varies with the coordinates, and is related to the Young's modulus by $\mu(\boldsymbol{x}) = E(\boldsymbol{x})/2(1+\nu)$.

The ratio of the density and shear modulus in Eq. (6) is expressed as $\frac{\rho(\boldsymbol{x})}{\mu(\boldsymbol{x})} = \frac{\rho_s}{\mu_s}e^{(\gamma-\beta)\sqrt{(x-x_0)^2+(y-y_0)^2}}$.

$U_{ij}(\boldsymbol{x}, \boldsymbol{y})$ and $T_{ij}(\boldsymbol{x}, \boldsymbol{y})$ are the elastostatic displacement fundamental solutions for homogeneous, isotropic and linear elastic solids with μ=1 [18].

$$U_{ij} = \frac{-1}{8\pi(1-\nu)}[(3-4\nu)\delta_{ij}\ln(r) - r_{,i}r_{,j}], \tag{8}$$

$$\Sigma_{ijl} = c^0_{rsjl}U_{ir,s} = \frac{-1}{4\pi(1-\nu)r}[(1-2\nu)(\delta_{il}r_{,j} + \delta_{ij}r_{,l} - \delta_{jl}r_{,i}) + 2r_{,i}r_{,j}r_{,l}], \tag{9}$$

$$T_{ij} = \Sigma_{ijl}n_l = \frac{-1}{4\pi(1-\nu)r}[(1-2\nu)(n_ir_{,j} - n_jr_{,i}) + ((1-2\nu)\delta_{ij} + 2r_{,i}r_{,j})r_{,l}n_l], \tag{10}$$

$$V_{ij} = \Sigma_{ijl}\tilde{\mu}_{,l} = \frac{-1}{4\pi(1-\nu)r}[(1-2\nu)(\tilde{\mu}_{,i}r_{,j} - \tilde{\mu}_{,j}r_{,i}) + ((1-2\nu)\delta_{ij} + 2r_{,i}r_{,j})r_{,l}\tilde{\mu}_{,l}], \tag{11}$$

where δ_{ij} is the Kronecker delta. r=|$\boldsymbol{x}$-$\boldsymbol{y}$| is the distance from the field point $\boldsymbol{x}$ to the source point $\boldsymbol{y}$. The elasticity tensor c_{ijkl} can be described in the form of

$$c_{ijkl}(\boldsymbol{x}) = \mu(\boldsymbol{x})c^0_{ijkl}, \quad \text{where} \quad c^0_{ijkl} = \frac{2\nu}{1-2\nu}\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}, \tag{12a, b}$$

where c^0_{ijkl} is the elasticity tensor for the corresponding "fictitious" homogeneous material with μ=1.

One to be noted is that, the $\tilde{\mu}_{,l}$ in Eq.(11) is not constant any more, it yields:

$$\tilde{\mu}_{,l} = \beta\left[(x-x_0)^2 + (y-y_0)^2\right]^{-\frac{1}{2}}(x-x_0), \qquad \text{where } l\text{=1}, \tag{13a}$$

$$\tilde{\mu}_{,l} = \beta\left[(x-x_0)^2 + (y-y_0)^2\right]^{-\frac{1}{2}}(y-y_0), \qquad \text{where } l\text{=2}. \tag{13b}$$

Boundary-domain integral equations for boundary points can be obtained by locating $\boldsymbol{y}$ at the boundary Γ in Eq. (6). It is observed that the two domain integrals in the Eq. (6) arise from the material inhomogeneity and the inertial effect. Radial integration method (RIM) of Gao [19] was employed to transform the domain integrals into boundary integrals over the global boundary. The normalized displacements is approximated by a combination of the radial basis functions and the polynomials of global coordinates as

$$\tilde{u}_i(\boldsymbol{x}) = \sum_A \alpha_i^A \phi^A(R) + a_i^k x_k + a_i^0, \qquad \sum_A \alpha_i^A = 0, \qquad \sum_A \alpha_i^A x_j^A = 0, \tag{14a, b, c}$$

where $\phi^A(R)$ is the radial basis function, α_i^A, a_i^k and a_i^0 are unknown coefficients to be determined, x_k and x_j^A denote the coordinates of the field point $\boldsymbol{x}$ and the application point A respectively. In this analysis, the following 4th order spline-type radial basis function was applied [20]

$$\phi^A(R) = \begin{cases} 1-6\left(\frac{R}{d_A}\right)^2 + 8\left(\frac{R}{d_A}\right)^3 - 3\left(\frac{R}{d_A}\right)^4, 0 \le R \le d_A \\ 0, R \ge d_A \end{cases}, \tag{15}$$

where $R=\|\boldsymbol{x}-\boldsymbol{x}^A\|$ is the distance from the application point A to the field point $\boldsymbol{x}$, and d_A is the support size for the application point A. The two domain integrals of Eq. (6) are transformed into the boundary integrals in the form of [21]

$$\int_\Omega V_{ij}\tilde{u}_j d\Omega = \alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^A d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} F_{ij}^1 d\Gamma + (a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^0 d\Gamma, \tag{16}$$

$$\omega^2 \int_\Omega \frac{\rho}{\mu} U_{ij}\tilde{u}_j d\Omega = \omega^2 \frac{\rho_s}{\mu_s}[\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^A d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} P_{ij}^1 d\Gamma + (a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^0 d\Gamma], \tag{17}$$

where the relation $x_i=y_i+r_{,i}r$ is used to relate x with r. By rewriting Eq. (11) with $V_{ij} = \bar{V}_{ij}/r$, the integral functions in Eqs. (16) and (17) can be expressed as

$$F_{ij}^A = \int_0^r r V_{ij}\phi^A dr = \bar{V}_{ij}\int_0^r \phi^A dr, \; F_{ij}^1 = \int_0^r r^2 V_{ij} dr = \frac{1}{2} r^2 \bar{V}_{ij}, \; F_{ij}^0 = \int_0^r r V_{ij} dr = r\bar{V}_{ij}, \tag{18a,b,c}$$

$$P_{ij}^A = \int_0^r r U_{ij}\phi^A e^{(\gamma-\beta)\sqrt{(x-x_0)^2+(y-y_0)^2}} dr, \qquad P_{ij}^1 = \int_0^r r^2 U_{ij} e^{(\gamma-\beta)\sqrt{(x-x_0)^2+(y-y_0)^2}} dr,$$

$$P_{ij}^0 = \int_0^r r U_{ij} e^{(\gamma-\beta)\sqrt{(x-x_0)^2+(y-y_0)^2}} dr. \tag{19a,b,c}$$

Since $r_{,i}$ in the above radial integrals is a constant, Eqs. (18b, c) can be evaluated analytically and the other integrals are calculated by standard Gaussian quadrature formula [22, 23]. Therefore the displacement boundary integral equations with only boundary integrals are obtained as

$$\begin{aligned} c_{ij}\tilde{u}_j = {} & \int_\Gamma U_{ij} t_j d\Gamma - \int_\Gamma T_{ij}\tilde{u}_j d\Gamma + [\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^A d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} F_{ij}^1 d\Gamma \\ & + (a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} F_{ij}^0 d\Gamma] + \omega^2 \frac{\rho_0}{\mu_0}[\alpha_j^A \int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^A d\Gamma + a_j^k \int_\Gamma \frac{r_{,k}}{r}\frac{\partial r}{\partial n} P_{ij}^1 d\Gamma \\ & + (a_j^k y_k + a_j^0)\int_\Gamma \frac{1}{r}\frac{\partial r}{\partial n} P_{ij}^0 d\Gamma] \end{aligned}. \tag{20}$$

Discretizing of the boundary with boundary elements and collocating the resulting boundary integral equations at all the boundary and internal nodes yield the following $2N_t$ x $2N_t$ generalized eigenvalue system for free vibration analysis,

$$\left[\boldsymbol{K}\right]\left\{\boldsymbol{X}\right\} = \omega^2\left[\boldsymbol{M}\right]\left\{\boldsymbol{X}\right\}. \tag{21}$$

By solving this general eigenvalue equation, the eigenvalue ω and the corresponding mode shapes $\{\boldsymbol{X}\}$ can be obtained numerically.

Numerical verification

A comparison is made on the in-plane free vibration of an isotropic, homogeneous clamped circular plate with the study by Park [14] using Hamilton's principle. In the comparison, the aluminum plate of a radius of 0.5m and a thickness of 5mm, E=71GPa, ρ=2700 kg/m^3 and v=0.33. To validate the analyzed results, the natural frequencies are also computed using the finite element method (FEM). The results of the present method are in good agreement with those of the FEM, which can be seen from the errors listed in Table 1. The present results are also close to the Ref. [8], even though solutions in Ref [8] are not complete, and some values are missing. This may be due to the reason that in Ref. [8], the displacements were assumed as the displacement multiple with a trigonometric function $cosm\theta$ or $sinm\theta$, where m=0, 1, 2…, ∞ was the circumferential wave number. The wave number had a correspondence frequency, if interchanging the trigonometric function $cosm\theta$ or $sinm\theta$, another set of free vibration modes can be obtained.

Table 1 Comparison of natural frequencies of the homogeneous clamped circular disk

Mode	Ref. [8]	Present(Hz)	FEM(Hz)	error(%)
1	3362	3360.1184	3360.3	0.0054
2	-	3360.1340	3360.4	0.0079
3	3835	3832.3110	3830.2	0.0551
4	5219	5214.3752	5211.9	0.0475
5	-	5214.3916	5212.1	0.0439
6	5383	5374.0392	5369.4	0.0863
7	-	5374.0648	5369.7	0.0812
8	6626	6612.1173	6617.5	0.0814
9	6764	6755.7080	6743.5	0.1807
10	-	6755.7498	6743.8	0.1769
11	6939	6911.3853	6911.2	0.0027
12	-	6911.3953	6913.2	0.0261
13	7021	6985.4645	6990.7	0.0749
14	-	8082.0531	8084.4	0.0290
15	-	8082.2690	8089.6	0.0907
16	8130	8129.7628	8091.0	0.4768
17	-	8294.3571	8273.8	0.2478
18	-	8438.0543	8435.4	0.0315
19	-	8438.1857	8437.9	0.0034
20	8489	8499.8673	8512.4	0.1474
21	-	8500.4750	8514.3	0.1626

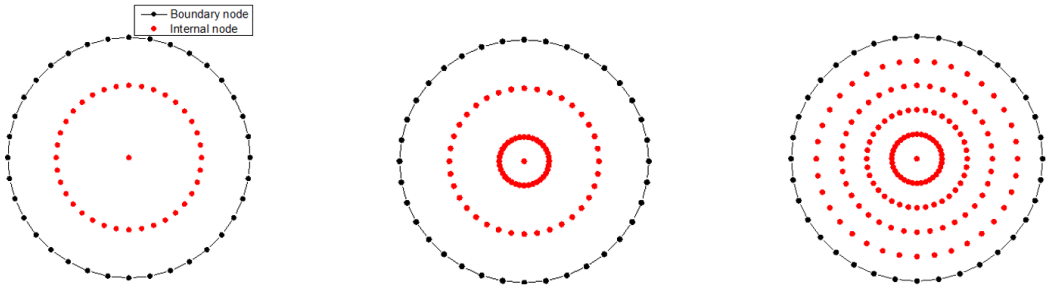
Another comparision is taking on the in-plane free vibration of the isotropic, homogeneous annular disks with several combinations of boundary conditions. The results of the present

method are compared with those available in the literature are tabulated in Table 2. Parameter $\eta=Ri/R$ which is the ratio between inner and outer radii of the disk. The boundary conditions were free in inner and clamped in outer circumfrential and are expressed as F-C, and other boundary conditions are deduced from this. To validate the analyzed results, the natural frequencies are also computed using the FEM. It can be obtained that, the normalized natural frequencies calculated by using the present method agreed well with those of FEM and references. However in Refs. [9] and [10], for F-C annular disks, the fundamental natural frequencies is close to the results of present method and FEM, but for the other two boundary conditions annular disks, the fundamental natural frequencies are close to the second frequencies of present method and FEM.

Table 2 Comparison of non-dimensional fundamental natural frequencies of annular disks

Radius ratio(η)	Boundary conditions	Mode	Present	FEM	Ref. [10]	Ref. [9]
0.2	F-C	1	2.0802	2.0983	2.1040	2.1060
	C-C	1	2.4801	2.4614	-	-
		2	2.7479	2.7347	2.7830	2.8060
	C-F	1	0.3479	0.3444	-	-
		2	0.9085	0.9060	0.9190	0.9400
0.4	F-C	1	2.4365	2.4272	2.5170	2.5220
	C-C	1	3.1449	3.1285	-	-
		2	3.3914	3.3607	3.4290	3.4560
	C-F	1	0.8139	0.8105	-	-
		2	1.2663	1.2631	1.2810	1.2960

Since the in-plane free vibration of the circular FG disk are very rare in the literature, the result computed by the present method is compared with that by using FEM and presented in Table 3. In this example, the center of the disk is x_0=0.5m, y_0=0.5m, and R=0.5m, the material perperties are graded from center steel E=210GPa, ρ=7806kg/m^3 along the radial direction to circumference aluminum E=70GPa, ρ=2707kg/m^3. Three kinds of the boundary and internal nodes distribution are taken to investigate the efficiency of the present method and plotted in Fig. 2. From the comparision, it can be seen that, 36 boundary nodes and 73 internal nodes is sufficent to achieve the convergence results even for high frequencies. Increasing the internal nodes can accurate the results, thus the error between the results of the present method with 36 boundary nodes and 145 internal nodes and those of FEM are vary small. The corresponding vibration modes are also obtained by using the presnet method and darwn in Fig. 3.

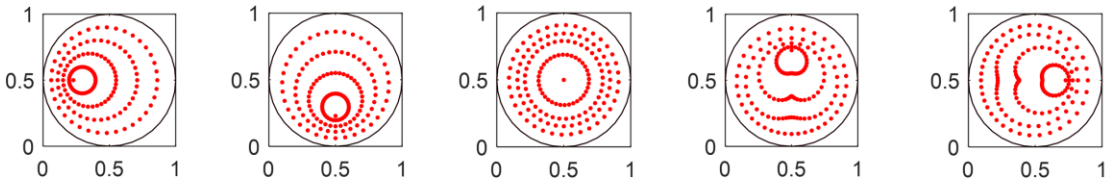


(a) 36B37I (b) 36B73I (c) 36B145I

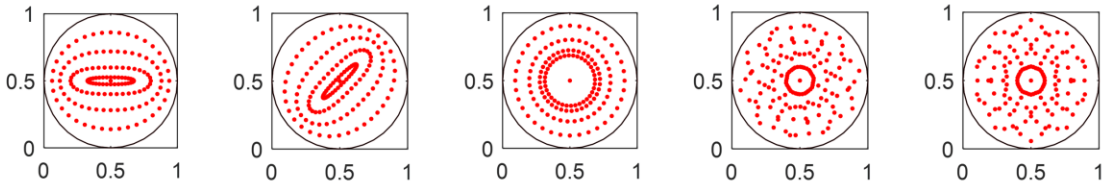
Figure 2 Boundary and internal nodes distribution

Table 3 Comparison of non-dimensional natural frequencies of circular FG disk

Modes	Present			FEM	Error(%)
	36B37I	36B73I	36B145I		
1	1.7544	1.7815	1.7953	1.8196	1.3387
2	1.7545	1.7816	1.7953	1.8196	1.3385
3	2.1044	2.0880	2.0832	2.0843	0.0544
4	3.0696	3.1550	3.1404	3.1494	0.2850
5	3.0697	3.1551	3.1404	3.1494	0.2846
6	3.2924	3.2440	3.2331	3.2314	0.0517
7	3.2924	3.2440	3.2331	3.2314	0.0520
8	4.2514	4.0276	4.0872	4.1249	0.9123
9	4.2514	4.1782	4.2033	4.2041	0.0213
10	4.4116	4.1782	4.2033	4.2041	0.0208



(a)ϖ=1.7953 (b)ϖ=1.7953 (c)ϖ=2.0832 (d)ϖ=3.1404 (e)ϖ=3.1404



(f)ϖ=3.2331 (g)ϖ=3.2331 (h)ϖ=4.0872 (i)ϖ=4.2033 (j)ϖ=4.2033

Figure 3 First ten mode shapes of clamped circular FG disk

Free vibration of the circular and annular FG disks

In this part, the present method is used to analysis the free vibration of the circular and annular FG disks with kinds of combination of boudary conditions. A steel/aluminum FG material is cosidered for all the numerical analyses, thus the matrterial properties can start from steel or aluminum and grade in the radial direction. For circular disks, the material properties grade from steel in the center to the outer edge aluminum are noted as S-A circular disk, A-S circlar disk means the material properties are grading in an opposite order. S-A annular disk presents the material properties grade from steel in the inner edge to outer aluminum edge. Only clmaped boundary support is taken in FG circular disks analyses. For annular FG disks, the free and clamped conditions at the inner and outer edges are noted as F-C, and so on C-F as well as C-C are also considered in the annular FG disks analysis. All the results were all normalized by $\varpi = \omega\left(R - R_i\right)\sqrt{\frac{\rho_{steel}}{E_{steel}}}$.

The first five normalized natural frequencies of clamped circular FG disks are shown in Table 4. To compared the circular FG disks with the homogeneous one, the normalized natural frequencies for the steel and aluminum disks are also listed. It can be seen that, the normalized natural frequencies of S-A(β=-1.0986, γ=-1.0591) < Aluminum(β=0, γ=0) < Steel(β=0, γ=0) < A-S(β=1.0986, γ=1.0591). That is with increasing the material gradients, increase the natural frequencies. For the homogeneous aluminum and steel, the material gradients are all equal to zero, in this case, the harder stiffness leads to the higher natural frequencies. It can be obtained that, by using the corresponding FG materials instead of the homogeneous material, can reduce or increase the natural frequencies of circular disks.

Table 4 First five normalized natural frequencies of clamped circular FG disks

Circular disks	ϖ=1	ϖ=2	ϖ=3	ϖ=4	ϖ=5
S-A	1.7927	1.7927	2.0834	3.1446	3.1446
Aluminum	2.1590	2.1590	2.3282	3.2802	3.2802
Steel	2.2022	2.2022	2.3747	3.3457	3.3457
A-S	2.5701	2.5701	2.6300	3.5125	3.5125

Table 5 First five normalized natural frequencies of kinds of annular FG disks

Radius ratio(η)	Modes	F-C		C-C		C-F	
		A-S	S-A	A-S	S-A	A-S	S-A
0.2	1	2.1794	1.5050	2.3217	1.8726	0.2073	0.4002
	2	2.2668	1.5050	2.6141	2.2271	0.6332	0.9141
	3	2.2668	1.6275	2.6141	2.2271	0.6332	0.9141
	4	2.5965	1.9757	3.1886	2.8651	1.0528	1.5327
	5	2.5965	1.9757	3.1886	2.8652	1.0529	1.5327
0.4	1	1.8157	1.2719	2.1766	1.8283	0.3839	0.6625
	2	2.0013	1.3714	2.3598	2.0260	0.6814	0.9398
	3	2.0013	1.3714	2.3598	2.0260	0.6814	0.9398
	4	2.2527	1.5038	2.8034	2.4727	1.0336	1.4118
	5	2.2528	1.5039	2.8034	2.4727	1.0336	1.4118

	1	1.6742	1.0231	2.1128	1.8465	0.5402	0.8548
	2	1.8051	1.1460	2.1790	1.8917	0.6891	0.9553
0.6	3	1.8124	1.1460	2.1790	1.8917	0.6891	0.9553
	4	1.8126	1.3464	2.3863	2.1216	0.9729	1.2128
	5	1.9151	1.3464	2.3863	2.1216	0.9729	1.2128
	1	1.2539	0.8166	2.0318	1.7238	0.7050	0.9582
	2	1.4467	0.8234	2.0321	1.7931	0.7479	0.9619
0.8	3	1.4550	0.8234	2.1098	1.7931	0.7479	0.9619
	4	1.4551	0.9147	2.2030	1.8054	0.8418	1.0482
	5	1.5376	0.9148	2.2030	1.8057	0.8418	1.0482

The first five normalized natural frequencies of two kinds of annular FG disks with three boundary conditions as well as four radius ratios are also shown in Table 5. It can be seen that, for F-C and C-C annular FG disks, increasing the annular radius ratio would decrease the natural frequencies, and for the same radius ratio, the higher material gradients give rise to higher natural frequencies. But for C-F annular disks, that is clamped at the inner edge and free at the outer edge, increasing the radius ratio would increase the natural frequencies and the smaller material gradients the higher the natural frequencies, these regularities are all oppositing with that of F-C and C-C annular disks. Thus it can be concluded that, the material gradiens, radius ratios and boundary conditions all affect the free vibration of the annular FG disks a lot.

Conclusion

In this paper, the free vibration of the circular and annular FG disks were analyzed by a meshfree boundary-domain integral equation methods. The material properties were graded along the radial direction from center or inner edge to outer edge for circular and annular disks respectively. From the numerical analyses, it can be concluded that, the present method has fast convergence, high efficiency and accuracy.

Acknowledge

This project is supported by the RDAO of University of Macau (RADO Project no: MYRG2015-00142-FST).

References

[1] Weisensel GN. Natural frequency information for circular and annular plates. Journal of Sound and Vibration, 1989, 133(1): 129-137.
[2] Kirkhope J, Wilson GJ. Vibration and stress analysis of thin rotating discs using annular finite elements. Journal of Sound and Vibration, 1976, 44(4): 461-474.
[3] Han JB, Liew KM. Axisymmetric free vibration of thick annular plates, International Jounral of Mechanical Sciences, 1999, 68(41): 1089-1109.
[4] Chung J, Kang NC, Lee JM. A study on free vibration of a spinning disk. Journal of Mechanical Science and Technology, 1996, 10(2): 138-145.
[5] Farag NH, Pan J, Modal characteristics of in-plane vibration of circular plates clamped at the outer edge, Journal of the Acoustical Society of America, 2003, 113: 1935–1946.
[6] Ambati G, Bell JFW, Sharp JCK. In-plane vibrations of annular rings. Journal of Sound and Vibration, 1976, 47(3): 415-432.
[7] Nigh GL, Olson MD. Finite element analysis of rotating disks. Jouranl of Sound and Vibration, 1981, 77(1): 61-78.
[8] Park CI. Frequency equation for the in-plane vibration of a clamped circular plate. Journal of Sound and Vibration 2008, 313; pp: 325-333.

[9] Bshmal S, Bhat R, Rakheja S. In-plane free vibration of circular annular disks. Journal of Sound and Vibration, 2009, 322(1-2): 216-226.
[10] T. Irie, G. Yamada, Y. Muramoto, Natural frequencies of in-plane vibration of annular plates, Journal of Sound and Vibration 97(1984) 171–175.
[11] Tornabene F. Free vibration analysis of functionally graded conical, cylindrical shell and annular plate structures with a four-parameter power-law distribution. Computer Methods in Applied Mechanics and Engineering 2009; 198(37-40), pp: 2911-2935.
[12] Kermani ID, Ghayour M, Mirdamadi HR. Free vibration analysis of multi-directional functionally graded circular and annular plates. Journal of Mechanical Science and Technology 2012, 26(11), pp: 3399-3410.
[13] Nie GJ, Zhong Z. Semi-analytical solution for three-dimensional vibration of functionally graded circular plates. Computer Methods and Applied Mechanical Engineering 2007, 196; pp: 4901-4910.
[14] Dong CY. Three-dimensional free vibration analysis of functionally graded annular plates using the Chebyshev–Ritz method. Material and Design 2008, 29; pp: 1518-1525.
[15] Malekzadeh P, Shahpari SA, Ziaee HR. Three-dimensional free vibration of thick functionally graded annular plates in thermal environment. Journal of Sound and Vibration 2010, 329; pp: 425-442.
[16] Yang Y, Kou KP, Iu VP, Lam CC, Zhang C. Free vibration analysis of two-dimensional functionally graded structures by a meshfree boundary–domain integral equation method. Composite Structures 2014;110(0):342-353.
[17] Gao XW, Zhang C, Sladek J, Sladek V. Fracture analysis of functionally graded materials by a BEM. Composites Science and Technology 2008;68(5):1209-1215.
[18] Gao XW, Davies TG. Boundary Element Programming in Mechanics. Cambridge University Press; 2002.
[19] Gao XW. The radial integration method for evaluation of domain integrals with boundary-only discretization. Engineering Analysis with Boundary Elements 2002;26(10):905-916.
[20] Golberg MA, Chen CS, Bowman H. Some recent results and proposals for the use of radial basis functions in the BEM. Engineering Analysis with Boundary Elements 1999;23(4):285-296.
[21] Yang Y, Lam CC, Kou KP, Iu VP. Free vibration analysis of the functionally graded sandwich beams by a meshfree boundary-domain integral equation method. Composite Structures 2014;117(0):32-39.
[22] Yang Y, Lam CC, Kou KP, Iu VP. Free vibration analysis of two-dimensional functionally graded coated and undercoated substrate structures. Engineering Analysis with Boundary Elements 2015;60:10-17.
[23] Yang Y, Lam CC, Kou KP, Iu VP. Forced vibration analysis of functionally graded beams by the meshfree boundary-domain integral equation method. Engineering Analysis with Boundary Elements 2016;72:100-110.

Static and free vibration of laminated composite plates using higher order cell-based smoothed finite element method with Q8 elements

†Dean Hu ¦ Detao Wan ¦ and Xu Han[1]

1 State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha 410082, P. R. China

Presenting author: hudean@hnu.edu.cn
†Corresponding author: hudean@hnu.edu.cn

Abstract

In this paper, the higher order cell-based smoothed finite element method based on the first-order shear deformation theory is used for the analysis of laminated composite plates. The domain is discretized with eight-node Mindlin plate elements of serendipity family (Q8 elements). Higher order finite element with Q8 elements using the selectively reduced integration is known to alleviate the shear-locking phenomenon. However, it still produces shear-locking phenomenon below a certain thickness-span ratio and also yields poor solutions and sub-optimal convergence rates with distorted meshes. In this paper, we propose a novel approach to eradicate the shear-locking phenomenon and improve the quality of the solutions by employed a linear smoothing technique. Within this technique, each Q8 element is subdivided into eight smoothing cells, and a modified strain is computed over the smoothing cell by using a linear smoothing procedure. The modified bending strain and shear strain are computed by the divergence theorem between the nodal shape functions and their derivatives in Taylor's expansion within each smoothing cell. Several numerical examples indicate that the novel approach can eradicate the shear-locking and also yield more reliable results for the distorted meshes.

Keywords: Laminated composite plates, S-FEM, linear smoothing technique, shear-locking, distorted meshes.

Introduction

Due to the remarkable weight-to-stiffness, strength-to-stiffness characteristics and the advantage of designable, laminated composite plates are widely used in engineering structures as diverse as aerospace, aircrafts, automotive structural parts, civil engineering structures, etc. Many plate theories have been successfully applied to analyze the laminated composite plates. The first-order shear deformation theory is widely used in the analysis of composite plates [1]. In addition, Hinton and Zienkiewicz [2, 3] explicitly indicated that the eight-node Mindlin plate elements of serendipity family (Q8 elements) still suffer from the shear-locking phenomenon below a certain thickness-span ratio and yields poor solutions of distorted meshes.

In order to eliminate shear-locking phenomenon and improve the quality of finite element solutions over simplex elements, recently, Liu et al. [4]-[6] proposed a smoothed finite element method (S-FEM), which is based on the stabilized conforming nodal integration (SCNI) of mesh-free method [7], Note that all the types of S-FEM use finite element with linear interpolants, and the strain smoothing technique over the higher order elements exhibits poor performance [8]. Francis et al. [9] proposed a linear strain smoothing scheme with the framework of CS-FEM to improve accuracy of arbitrary convex polytopes with linear or quadratic interpolants, which was based on the recent work of Duan et al. [10].

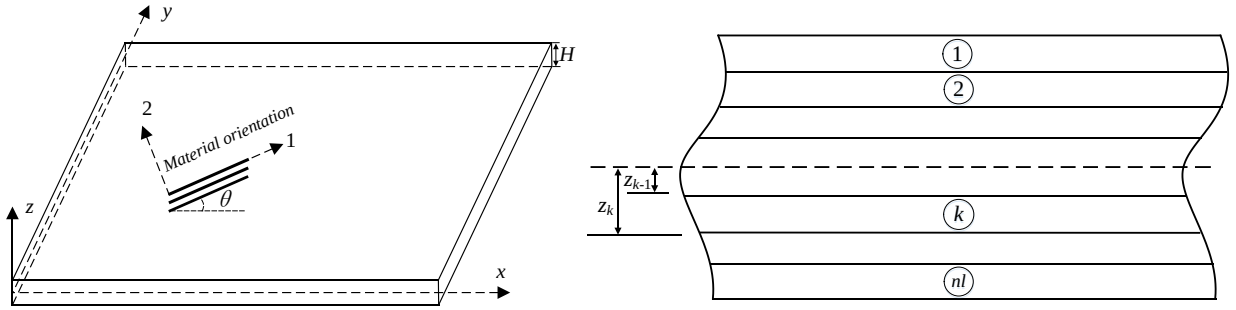
In this paper, with the aim of eliminating the shear-locking phenomenon and improving the accuracy of the solution with distorted meshes, a higher order CS-FEM with eight-node Mindlin plate elements is developed based on the first-shear deformation theory, which is a further

application of the linear smoothing technique. With this technique, as shown in Fig .1, each Q8 element is subdivided into eight smoothing cells, and the modified bending strain and shear strain are computed by the divergence theorem between the nodal shape functions and their derivatives in Taylor's expansion within each smoothing cell. This eliminates the need for isoparametric mapping and all the domain integration are transformed into boundary integration. Meanwhile, the proposed approach can effectively treat the shear-locking phenomenon for both thin and relatively thick plates, and the effect of the mesh distortion to the accuracy can be relieved.

Basic formulations

According to the first-shear deformation theory, the displacements field of the laminated composite plates (shown in **Fig. 1**) can be expressed as

$$\begin{cases} u(x,y,z) = u_0(x,y) - z\beta_x(x,y) \\ v(x,y,z) = v_0(x,y) - z\beta_y(x,y) \\ w(x,y,z) = w_0(x,y) \end{cases} \tag{1}$$



(a) (b)

Fig. 1. (a) Laminated composite plate axes system. (b) Layer details.

The modified strain is given by

$$\tilde{\boldsymbol{\varepsilon}}(\mathbf{x}_k) = \int_{\Omega_s} \boldsymbol{\varepsilon}(\mathbf{x})\mathbf{q}(\mathbf{x})\mathrm{d}\Omega \tag{2}$$

In the process of computing the modified strain matrix $\mathbf{B}_I^b$ and $\mathbf{B}_I^s$, the consistency form should be met by the terms related the shape function $N_I(\mathbf{x})$ and the derivative of shape function $N_{I,i}(\mathbf{x})$ $(i = x, y)$. According to Eq. (2) and implementing the divergence theorem, we can obtain

$$\int_{\Omega_s} N_{I,i}\mathbf{q}(\mathbf{x})\mathrm{d}\Omega = \int_{\Gamma_s} N_I\mathbf{q}(\mathbf{x})n_i\mathrm{d}\Gamma - \int_{\Omega_s} N_I\mathbf{q}_{,i}(\mathbf{x})\mathrm{d}\Omega \tag{3}$$

$$\mathbf{q}(\mathbf{x}) = \mathbf{q}(\mathbf{x}_c) + (x - x_c)\mathbf{q}_{,x}(\mathbf{x}_c) + (y - y_c)\mathbf{q}_{,y}(\mathbf{x}_c) + \mathrm{H.O.T} \tag{4}$$

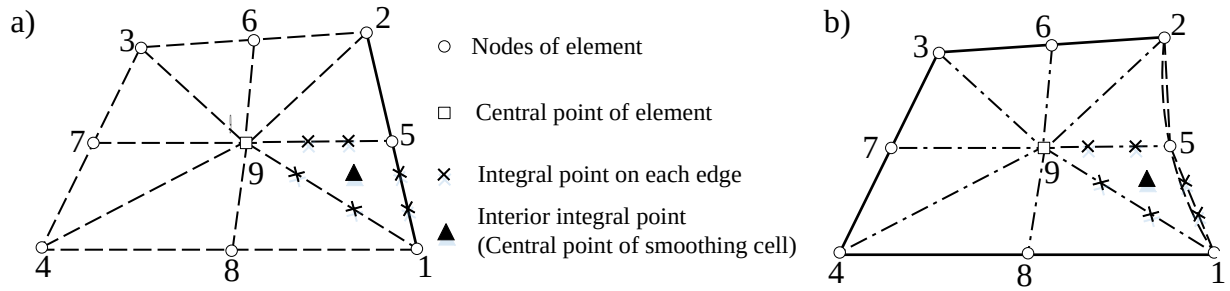


Fig. 2. Division of a Q8 plate element into eight smoothing cells

The evaluation of the modified derivatives $\tilde{N}_{I,x}(\mathbf{x})$ is firstly presented. To begin with, $\tilde{N}_{I,x}(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ are treated by Taylor's expansion, and the expanded form of $\tilde{N}_{I,x}(\mathbf{x})$ and $\mathbf{q}(\mathbf{x})$ are expressed as

$$\tilde{N}_{I,x}(\mathbf{x}) = \tilde{N}_{I,x}(\mathbf{x}_c) + (x - x_c)\tilde{N}_{I,xx}(\mathbf{x}_c) + (y - y_c)\tilde{N}_{I,xy}(\mathbf{x}_c) + \text{H.O.T} \tag{5}$$

$$\mathbf{q}(\mathbf{x}) = \mathbf{q}(\mathbf{x}_c) + (x - x_c)\mathbf{q}_{,x}(\mathbf{x}_c) + (y - y_c)\mathbf{q}_{,y}(\mathbf{x}_c) + \text{H.O.T} \tag{6}$$

where H.O.T is the higher order term. $N_{I,x}(\mathbf{x})$ is replaced by $\tilde{N}_{I,x}(\mathbf{x})$, then, substituting Eqs. (5) and (6) into Eq. (3), whereafter, one interior integral point is used for the left hand side of Eq. (3), and two integral point on each edge of the smoothing cell are employed for the right hand side of Eq. (3), and the subdivision of a Q8 element into smoothing cells is shown in **Fig. 2.** We can obtain

$$\begin{aligned} &\left[\mathbf{q}(\mathbf{x}_c)A\right]\tilde{N}_{I,x}(\mathbf{x}_c) + \left[\mathbf{q}_{,x}(\mathbf{x}_c)I_c^{xx} + \mathbf{q}_{,y}(\mathbf{x}_c)I_c^{xy}\right]\tilde{N}_{I,xx}(\mathbf{x}_c) + \left[\mathbf{q}_{,x}(\mathbf{x}_c)I_c^{xy} + \mathbf{q}_{,y}(\mathbf{x}_c)I_c^{yy}\right]\tilde{N}_{I,xy}(\mathbf{x}_c) \\ &= \int_{\Gamma_s} N_I(\mathbf{x})\mathbf{q}(\mathbf{x})n_x \mathrm{d}\Gamma - \int_{\Gamma_s}\left(\frac{\partial\left(\int N_I(\mathbf{x})\mathbf{q}_{,x}(\mathbf{x})\mathrm{d}x\right)}{\partial x}\right)n_x \mathrm{d}\Gamma \end{aligned} \tag{7}$$

where A is the area of the smoothing cell, and

$$\begin{cases} I_c^{xx} = \int_{\Omega_s}(x - x_c)^2 \,\mathrm{d}\Omega \\ I_c^{xy} = \int_{\Omega_s}(x - x_c)(y - y_c)\,\mathrm{d}\Omega \\ I_c^{xy} = \int_{\Omega_s}(y - y_c)^2 \,\mathrm{d}\Omega \end{cases} \tag{8}$$

Substituting Eq.(4) into Eq.(7) and the two Gauss points on each edge is used for the evaluation of the boundary integral, therefore, the expanded form is expressed as

$$\begin{bmatrix} A & 0 & 0 \\ Ax_c & I_c^{xx} & I_c^{xy} \\ Ay_c & I_c^{xy} & I_c^{yy} \end{bmatrix}\begin{Bmatrix} \tilde{N}_{I,x}(\mathbf{x}_c) \\ \tilde{N}_{I,xx}(\mathbf{x}_c) \\ \tilde{N}_{I,xy}(\mathbf{x}_c) \end{Bmatrix} = \begin{Bmatrix} F_{11}^x \\ F_{12}^x \\ F_{13}^x \end{Bmatrix} \tag{9}$$

where

$$\begin{Bmatrix} F_{11}^x \\ F_{12}^x \\ F_{13}^x \end{Bmatrix} = \begin{Bmatrix} \sum_{k=1}^{N_{eg}}\sum_{G=1}^{2} N_I(\mathbf{x}_G^k)n_x^k W_G \\ \sum_{k=1}^{N_{eg}}\sum_{G=1}^{2}\left(N_I(\mathbf{x}_G^k)x_G^k - \left(\frac{\partial\left(\int N_I(\mathbf{x}_G^k)\mathrm{d}x\right)}{\partial x}\right)\right)n_x^k W_G \\ \sum_{i=1}^{N_{eg}}\sum_{G=1}^{2} N_I(\mathbf{x}_G^k)y_G^k n_x^k W_G \end{Bmatrix} \tag{10}$$

Then, the modified derivatives can be obtained by analytically solving Eq. (9). Therefore, $\tilde{N}_{I,x}(\mathbf{x}_c)$, $\tilde{N}_{I,xx}(\mathbf{x}_c)$ and $\tilde{N}_{I,xy}(\mathbf{x}_c)$ are given as

$$\begin{cases} \tilde{N}_{I,x}(\mathbf{x}_c) = \dfrac{F_{11}^x}{A} \\ \tilde{N}_{I,xy}(\mathbf{x}_c) = \dfrac{I_c^{yy}\left(F_{12}^x - F_{11}^x x_c\right) - I_c^{xy}\left(F_{13}^x - F_{11}^x y_c\right)}{I_c^{xx} I_c^{yy} - \left(I_c^{xy}\right)^2} \\ \tilde{N}_{I,xx}(\mathbf{x}_c) = \dfrac{I_c^{xx}\left(F_{13}^x - F_{11}^x y_c\right) - I_c^{xy}\left(F_{12}^x - F_{11}^x x_c\right)}{I_c^{xx} I_c^{yy} - \left(I_c^{xy}\right)^2} \end{cases} \tag{11}$$

In such a way, the evaluation of the modified derivatives, $\tilde{N}_{I,y}(\mathbf{x}_c)$, $\tilde{N}_{I,yx}(\mathbf{x}_c)$ and $\tilde{N}_{I,yy}(\mathbf{x}_c)$ can be obtained.

After that, the modified bending and shear strain matrices, and the modified derivatives of the bending and shear strain matrices can be respectively expressed as

$$\tilde{\mathbf{B}}_I^b = \begin{bmatrix} \tilde{N}_{I,x} & 0 & 0 & 0 & 0 \\ 0 & \tilde{N}_{I,y} & 0 & 0 & 0 \\ \tilde{N}_{I,y} & \tilde{N}_{I,x} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{N}_{I,x} & 0 \\ 0 & 0 & 0 & 0 & \tilde{N}_{I,y} \\ 0 & 0 & 0 & \tilde{N}_{I,\mathrm{y}} & \tilde{N}_{I,x} \end{bmatrix}, \quad I = 1 \cdots 8 \tag{12}$$

$$\frac{\partial \tilde{\mathbf{B}}_I^b}{\partial x} = \begin{bmatrix} \tilde{N}_{I,xx} & 0 & 0 & 0 & 0 \\ 0 & \tilde{N}_{I,yx} & 0 & 0 & 0 \\ \tilde{N}_{I,yx} & \tilde{N}_{I,xx} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{N}_{I,xx} & 0 \\ 0 & 0 & 0 & 0 & \tilde{N}_{I,yx} \\ 0 & 0 & 0 & \tilde{N}_{I,\mathrm{yx}} & \tilde{N}_{I,xx} \end{bmatrix}, \quad I = 1 \cdots 8 \tag{13}$$

$$\frac{\partial \tilde{\mathbf{B}}_I^b}{\partial y} = \begin{bmatrix} \tilde{N}_{I,xy} & 0 & 0 & 0 & 0 \\ 0 & \tilde{N}_{I,yy} & 0 & 0 & 0 \\ \tilde{N}_{I,yy} & \tilde{N}_{I,xy} & 0 & 0 & 0 \\ 0 & 0 & 0 & \tilde{N}_{I,xy} & 0 \\ 0 & 0 & 0 & 0 & \tilde{N}_{I,yy} \\ 0 & 0 & 0 & \tilde{N}_{I,\mathrm{yy}} & \tilde{N}_{I,xy} \end{bmatrix}, \quad I = 1 \cdots 8 \tag{14}$$

$$\tilde{\mathbf{B}}_I^s = \begin{bmatrix} 0 & 0 & \tilde{N}_{I,x} & -\tilde{N}_I & 0 \\ 0 & 0 & \tilde{N}_{I,\mathrm{y}} & 0 & \tilde{N}_I \end{bmatrix}, \quad I = 1 \cdots 8 \tag{15}$$

$$\frac{\partial \tilde{\mathbf{B}}_I^s}{\partial x} == \begin{bmatrix} 0 & 0 & \tilde{N}_{I,xx} & -\tilde{N}_{I,x} & 0 \\ 0 & 0 & \tilde{N}_{I,\mathrm{yx}} & 0 & \tilde{N}_{I,x} \end{bmatrix}, \quad I = 1 \cdots 8 \tag{16}$$

$$\frac{\partial \tilde{\mathbf{B}}_I^s}{\partial y} == \begin{bmatrix} 0 & 0 & \tilde{N}_{I,xy} & -\tilde{N}_{I,y} & 0 \\ 0 & 0 & \tilde{N}_{I,\mathrm{yy}} & 0 & \tilde{N}_{I,\mathrm{y}} \end{bmatrix}, \quad I = 1 \cdots 8 \tag{17}$$

By using the Taylor's expansion with respect to the each smoothing cell center $\mathbf{x}_c$, smoothed stiffness matrix are rewritten as

$$
\begin{aligned}
(\tilde{\mathbf{K}}^b)^h &= \sum_{C=1}^{nc}\left(\int_{\Omega_s}\tilde{\mathbf{B}}^{b\mathrm{T}}\bar{\mathbf{D}}\tilde{\mathbf{B}}^b\mathrm{d}\Omega\right) \\
&= \sum_{C=1}^{nc}\left(\int_{\Omega_s}\left[\tilde{\mathbf{B}}^{b\mathrm{T}}+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}(x-x_c)+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}(y-y_c)\right]\bar{\mathbf{D}}\left[\tilde{\mathbf{B}}^b+\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}(x-x_c)+\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}(y-y_c)\right]\mathrm{d}\Omega\right) \\
&= \sum_{C=1}^{nc}\left(A\tilde{\mathbf{B}}^{b\mathrm{T}}\bar{\mathbf{D}}\tilde{\mathbf{B}}^{b\mathrm{T}}+I_c^{xx}\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}+I_c^{yy}\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}+I_c^{xy}\left(\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}\right)\right)
\end{aligned}
\tag{18}
$$

$$
\begin{aligned}
(\tilde{\mathbf{K}}^b)^h &= \sum_{C=1}^{nc}\left(\int_{\Omega_s}\tilde{\mathbf{B}}^{b\mathrm{T}}\bar{\mathbf{D}}\tilde{\mathbf{B}}^b\mathrm{d}\Omega\right) \\
&= \sum_{C=1}^{nc}\left(\int_{\Omega_s}\left[\tilde{\mathbf{B}}^{b\mathrm{T}}+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}(x-x_c)+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}(y-y_c)\right]\bar{\mathbf{D}}\left[\tilde{\mathbf{B}}^b+\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}(x-x_c)+\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}(y-y_c)\right]\mathrm{d}\Omega\right) \\
&= \sum_{C=1}^{nc}\left(A\tilde{\mathbf{B}}^{b\mathrm{T}}\bar{\mathbf{D}}\tilde{\mathbf{B}}^{b\mathrm{T}}+I_c^{xx}\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}+I_c^{yy}\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}+I_c^{xy}\left(\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial x}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial y}+\frac{\partial\tilde{\mathbf{B}}^{b\mathrm{T}}}{\partial y}\bar{\mathbf{D}}\frac{\partial\tilde{\mathbf{B}}^b}{\partial x}\right)\right)
\end{aligned}
\tag{19}
$$

where

$$
\tilde{\mathbf{B}}^b=[\tilde{\mathbf{B}}_1^b \quad \tilde{\mathbf{B}}_2^b \quad \cdots \quad \tilde{\mathbf{B}}_8^b] \tag{20}
$$

and

$$
\tilde{\mathbf{B}}^s=[\tilde{\mathbf{B}}_1^s \quad \tilde{\mathbf{B}}_2^s \quad \cdots \quad \tilde{\mathbf{B}}_8^s] \tag{21}
$$

Finally, the smoothed stiffness matrix $\tilde{\mathbf{K}}_{IJ}$ are rewritten as

$$
\tilde{\mathbf{K}}_{IJ}=\tilde{\mathbf{K}}_{IJ}^b+\tilde{\mathbf{K}}_{IJ}^s \tag{22}
$$

Numerical examples and results

Firstly, the shear-locking test is presented. Isotropic simply supported square plate with side length $a=10$ under uniform load P. The deflection is normalized as $\hat{w}=w/(Pa^4/100D)$, where $D=Et^3/12(1-v^2)$. **Fig. 3** shows the locking test results for the isotropic clamped square plate. It can be seen that present method can eradicate the shear-locking.

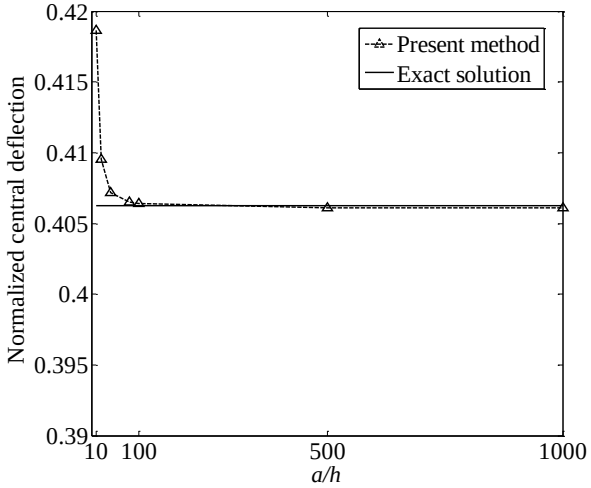
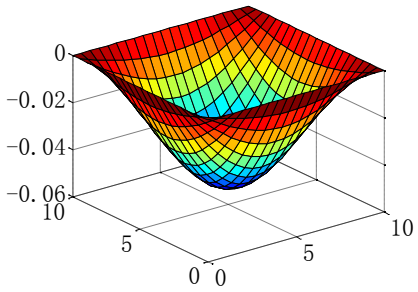


Fig. 3 Shear-locking test for a simply supported square plate

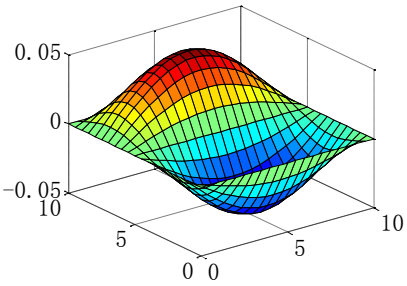
Secondly, two simply supported symmetric cross-ply ($0^\circ/90^\circ/0^\circ$) square plates are considered, the length to thickness ratios is $a/h=100$, The material parameters are $\rho=1643$, $E_{22}=7.6\times10^9$, $E_{11}=25E_{22}$, $G_{12}=G_{13}=0.5E_{22}$, $G_{23}=0.2E_{22}$, and $\nu_{11}=0.25$. Both present results, Liu's and Reddy's solutions are listed in **Table 1** for comparison purpose. Good agreement can be observed for all modes. In addition, the first six modes are shown in **Fig. 4**.

Table 1. Nondimensionalized natural frequencies of simply supported symmetric cross-ply ($0^\circ/90^\circ/0^\circ$) square composite plates with RI ($\bar{\omega}=\omega a^2\sqrt{\rho E_{22}h^2}$)

Mode	Number of elements	$\bar{\omega}$		
		Present method	Liu GR et al.	Reddy
1	8×8	15.1795	15.127	15.183
	16×16	15.1645		
	20×20	15.1656		
	24×24	15.1665		
2	8×8	23.0857	22.658	22.817
	16×16	22.8351		
	20×20	22.8193		
	24×24	22.8132		
3	8×8	41.8609	39.644	40.153
	16×16	40.3884		
	20×20	40.2707		
	24×24	40.2198		
4	8×8	56.1309	55.452	56.210
	16×16	55.9385		
	20×20	55.9479		
	24×24	55.9575		
5	8×8	60.2757	59.289	60.211
	16×16	59.9252		
	20×20	59.9378		
	24×24	59.9510		



Mode 1



Mode 2

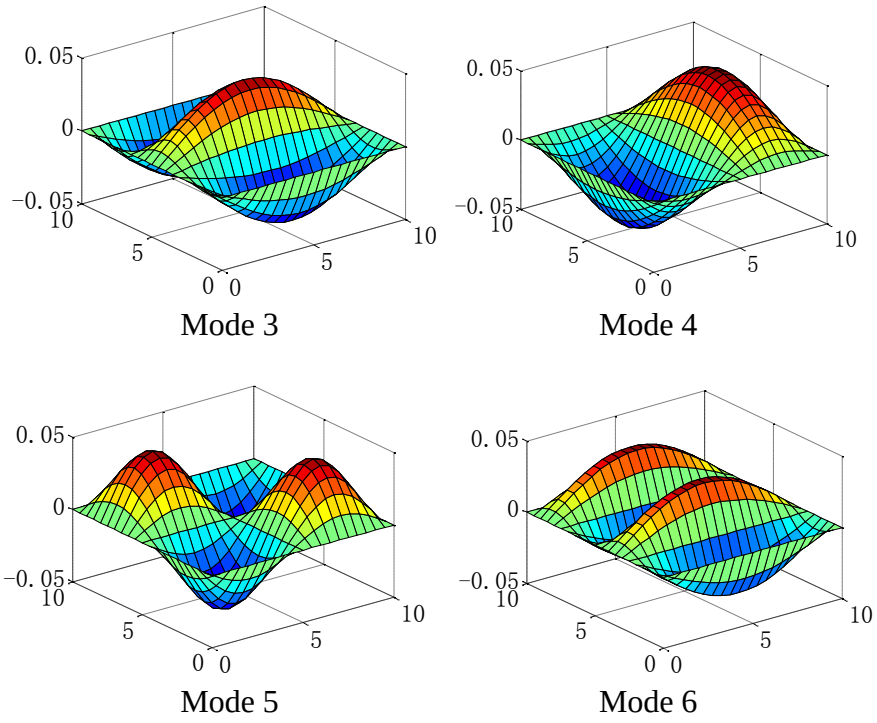


Fig. 4 The first six modes of the simply supported symmetric cross-ply square plates

References

[1] Whitney J. (1969) The effect of transverse shear deformation on the bending of laminated plates, Journal of Composite Materials 3, 534-547.

[2] Zienkiewicz OC, Taylor RL, Too JM. (1971) Reduced integration technique in general analysis of plates and shells, International Journal for Numerical Methods in Engineering, 3, 275-290.

[3] Pugh EDL, Hinton E, Zienkiewicz OC. (1978) A study of quadrilateral plate bending elements with 'reduced' integration, International Journal for Numerical Methods in Engineering,12, 1059-1079.

[4] Liu GR, Dai KY, Nguyen TT. (2007) A smoothed finite element method for mechanics problems, Computational Mechanics, 39, 859-877.

[5] Liu GR, Nguyen TT, Dai KY, Lam KY. (2007) Theoretical aspects of the smoothed finite element method (SFEM), International Journal for Numerical Methods in Engineering, 71, 902-930.

[6] Liu GR, Nguyen-Thoi T, Lam KY. (2009) An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analyses of solids, Journal of Sound and Vibration. 320, 1100-1130.

[7] Chen JS, Yoon S, Wu CT. (2002) A stabilized confirming nodal integration for Galerkin meshfree methods, International Journal for Numerical Methods in Engineering, 53, 2587-2615.

[8] Bordas SPA, Natarajan S, Kerfriden P, Augarde CE, Mahapatra DR, Rabczuk T, Pont SD. (2011) On the performance of strain smoothing for quadratic and enriched finite element approximations (XFEM/GFEM/PUFEM), International Journal for Numerical Methods in Engineering, 86, 637-666.

[9] Francis A, Ortiz-Bernardin A, Bordas SP, Natarajan S. (2016) Linear smoothed polygonal and polyhedral finite elements. International Journal for Numerical Methods in Engineering, 109, 1-28.

[10] Duan Q, Li X, Zhang H, Wang B, Gao X. (2012) Quadratically consistent one-point (QC1) quadrature for meshfree Galerkin methods. Computer Methods in Applied Mechanics and Engineering, 245–246, 256-272.

Acoustic Characteristics of 3-D Membrane-embedded-type Metamaterials

Y. Li[1],† X. M. Wang[1], ⋆Y. L. Mei[2]

[1]School of Mechanical Engineering, Dalian University of Technology, China.

[2] School of Automotive Engineering, Dalian University of Technology, China.

*Presenting author: meiyulin@dlut.edu.cn

†Corresponding author: meiyulin@dlut.edu.cn

Abstract

Researches on manipulating acoustic waves by anisotropic materials, especially, acoustic metamaterials, have received much attention. In this paper, an attempt is made to develop a new kind of 3-D acoustic metamaterial, which is constructed by embedding high-modulus membranes into soft isotropic materials. To begin with, two unit cell models of the acoustic metamaterials are established, which have an embedded single-layer membrane or embedded double-layer membranes. Then dispersion characteristics of the unit cells in the low frequency range are simulated by the finite element method; meanwhile, considering different angles of incidence of elastic waves, we use the transient response full integration method and the asymptotic homogenization method to calculate average velocities of elastic waves propagating in the unit cells. Subsequently, influences of structure parameters of the unit cells on the wave velocities are investigated. In the end, we establish a full wave simulation model of the three-dimensional membrane-embedded-type acoustic metamaterial and illustrate the propagation characteristics of acoustic waves in the metamaterial.

Keywords: Acoustic metamaterials; Membrane-embedded; Dispersion characteristics; Effective material parameters; Wave propagation

Introduction

Recently, researchers pay more attention to studies of manipulating acoustic waves by acoustic metamaterials, including acoustic cloaks, perfect focusing acoustic lenses and acoustic omnidirectional absorbers [1]-[3]. In general, acoustic metamaterials are constructed by periodically embedding local oscillators into base materials to produce local resonance, or embedding materials with high-contrast modulus into acoustic materials to cause multiple scattering of acoustic waves, or directly tailoring solid materials based on homogenization theory. In 2008, Mei, Yang et al. investigate a membrane-type acoustic metamaterial by means of simulation and experiment, where a small mass block is placed at the center of a membrane as a source of local resonance to realize acoustic absorption within the low frequency range [4]. Subsequently, researches on membrane-type acoustic metamaterials mainly highlight mechanical properties of the elastic membranes, such as ultra-thin size and ultra-light weight. And the acoustic characteristics of the metamaterials as well as the locally resonant frequencies and vibration modes can be designed, to some extent, by adjusting sizes of membranes, pre-stress applied in membranes and shapes of mass block [5]-[7]. However, most of the work focuses on acoustic insulation and absorption performance, and there has been very little study of controlling wave propagation paths in the membrane-type acoustic metamaterials. And related studies refer to active acoustic metamaterials, for example, Popa et al. design a class of tunable active acoustic metamaterials by using piezoelectric membranes

in 2013 [8].

Metamaterials with strong acoustic anisotropy can make acoustic propagation paths bend drastically, like curving and distorting, which is significant for practical acoustics problems with limited design spaces. In the paper, a kind of 3-D membrane-embedded-type acoustic metamaterial is constructed by embedding high-modulus membranes into low-modulus materials, where the in-plane stiffness of the embedded membrane is greatly different from the out-of-plane stiffness. In principle, the strong acoustic anisotropy of the metamaterials can be obtained by virtue of compositing isotropic materials with high-contrast modulus, and the directivity of wave propagation and dispersion characteristics of metamaterials can be controlled by tailoring the membranes or adjusting tensions or pre-deformations of the membranes. We mainly address acoustic characteristics of the membrane-embedded-type metamaterials, unit cells of which consist of soft rubber and a single-layer membrane or double-layer membranes. Firstly, dispersion characteristics of the unit cells are investigated in the low frequency range by the finite element method. Secondly, considering different incidence angles of acoustic waves, average velocities of wave propagation are calculated by both the transient response full integration method and the asymptotic homogenization method. Subsequently, influences of structure parameters of the unit cells on the acoustic wave velocities are studied. Finally, numerical examples are given to illustrate the wave propagation from homogeneous rubber materials to anisotropic composite materials consisting of rubber and embedded multi-layer membranes.

Unit cell models and analysis theory

According to the elastic dynamic theory, if an elastic wave propagates in the linear elastic and anisotropic inhomogeneous media, without regard to damping and external excitation, the wave equation can be expressed as

$$\nabla \cdot [\mathbf{C}(\mathbf{r}) : \nabla \mathbf{u}] = \rho(\mathbf{r}) \frac{\partial^2 \mathbf{u}}{\partial t^2} \tag{1}$$

Where $\mathbf{r} = (x, y, z)$ are position vectors, $\mathbf{C}(\mathbf{r})$ and $\rho(\mathbf{r})$ are elasticity tensors and mass-density tensors, respectively; $\mathbf{u} = (u_x, u_y, u_z)$ are displacement vectors; $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ stands for the differential operator; the colon ":" means double dot product; t is the time variable.

For a periodic structure consisting of unit cells, according to Bloch's theorem of periodic differential equations, we have

$$\mathbf{u}(\mathbf{r}) = e^{i(\mathbf{k}\cdot\mathbf{r})}\mathbf{u}_{\mathbf{k}}(\mathbf{r}) \tag{2}$$

Where $i = \sqrt{-1}$, $\mathbf{k} = (k_x, k_y, k_z)$ are wave numbers, $\mathbf{u}_{\mathbf{k}}(\mathbf{r})$ is a periodic function with the same period as one of the periodic structure. Considering periodicity of the structure, in order to reduce the redundant calculation, we just analyze dispersion characteristics of a unit cell. During solving the wave equation by the finite element method, the unit cell is discretized and its eigen circular frequencies ω associated with the eigenmodes $\mathbf{U}$ satisfy

$$(\mathbf{K} - \omega^2 \mathbf{M})\mathbf{U} = 0 \tag{3}$$

Where, $\mathbf{K} = \int \mathbf{B}^{\mathrm{T}}\mathbf{C}(\mathbf{r})\mathbf{B}\,\mathrm{d}V_e$ is the stiffness matrix, $\mathbf{M} = \int \rho(\mathbf{r})\mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}V_e$ is the mass matrix, V_e is the space occupied by the unit cell, $\mathbf{U} = [\mathbf{U}_1 \cdots \mathbf{U}_i \cdots \mathbf{U}_n]^{\mathrm{T}}$ is the displacement matrix,

$\mathbf{U}_i = [u_i \ v_i \ w_i]^{\mathrm{T}}$ $(i = 1, 2, \cdots, n)$ is the node displacement. If we use "$\mathbf{a}$" to represent the lattice constant, the boundary condition for the unit cell can be written as

$$\mathbf{U}(\mathbf{r} + \mathbf{a}) = e^{i(\mathbf{k}\cdot\mathbf{a})}\mathbf{U}(\mathbf{r}) \tag{4}$$

The unit cell is a symmetric structure as shown in Fig.1 or Fig.2, where a single-layer membrane or double-layer membranes are embedded into a soft isotropic material. The cross-section of the membrane is square, and its thickness is h_m; in Fig.1, h_z is the embedded depth of the single-layer membrane; in Fig.2, ΔH is the distance between the embedded double-layer membranes. It is assumed that the membrane is made of iron, and the material parameters are as follows: the modulus of elasticity $E_m = 2.1\times10^{11} Pa$, Poisson's ratio $\upsilon_m = 0.3$ and the density $\rho_m = 7860 kg/m^3$. Meanwhile, we choose rubber as the soft isotropic material, and the material parameters are the modulus of elasticity $E_r = 7.8\times10^{6} Pa$, Poisson's ratio $\upsilon_r = 0.47$ and the density $\rho_r = 980 kg/m^3$.

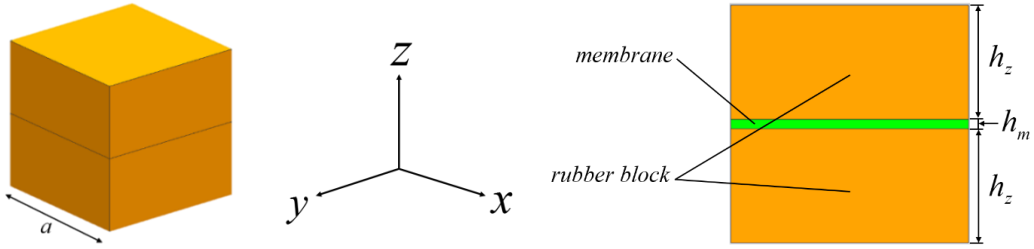


Figure 1. Unit cell with an embedded single-layer membrane

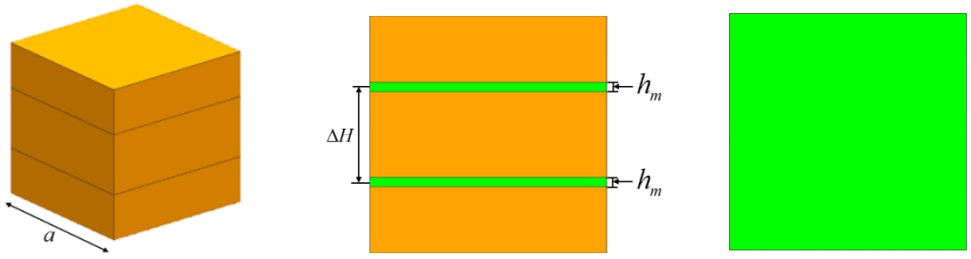


Figure 2. Unit cell with embedded double-layer membranes

Based on Eq.(3) and Eq.(4), after Floquent-Bloch boundary conditions are applied to the unit cell, dispersion characteristics of the unit cell are obtained. During the finite element analysis, in order to control the out-of-plane stiffness of the embedded membrane, we need adjust pre-stress applying to the membrane, so solid elements and membrane elements should be coupled to implement the simulation.

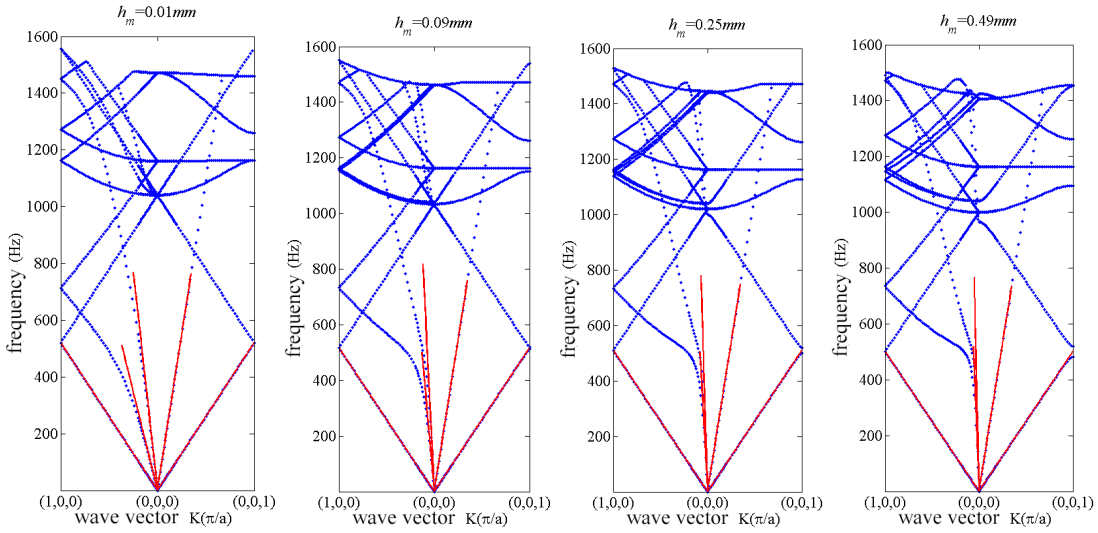
Analysis of dispersion characteristics

We take unit cells in Fig.1 and Fig.2 for examples to simulate dispersion curves of the unit cells with different h_m. Supposing that the pre-stress applied to the membrane along the z axis of Cartesian coordinate system $\{x, y, z\}$ is zero, the pre-stress along the x-axis is $T = 200 N/m$, the lattice constant is $a = 50 mm$, the distance between the double-layer membranes is $\Delta H = 30 mm$, the simulation results are shown in Fig.3(a)-(d) and Fig.4(a)-(d), where wave numbers are defined along the x-axis and the z-axis in the first Brillouin zone.

For the unit cell with a single-layer membrane in Fig.1, Fig.3(a)-(d) display dispersion curves while h_m are $0.01mm$, $0.09mm$, $0.25mm$ or $0.49mm$; similarly, corresponding to the unit cell with double-layer membranes with $\Delta H = 30mm$, Fig.4(a)-(d) display dispersion curves while h_m are $0.01mm$, $0.09mm$, $0.25mm$ or $0.49mm$. In all the figures, every red straight line is the tangent to a fitting curve to a series of data points among simulation results. According to the wave theory, the wave velocity c, the frequency f and the wave numbers $\mathbf{k}$ satisfy

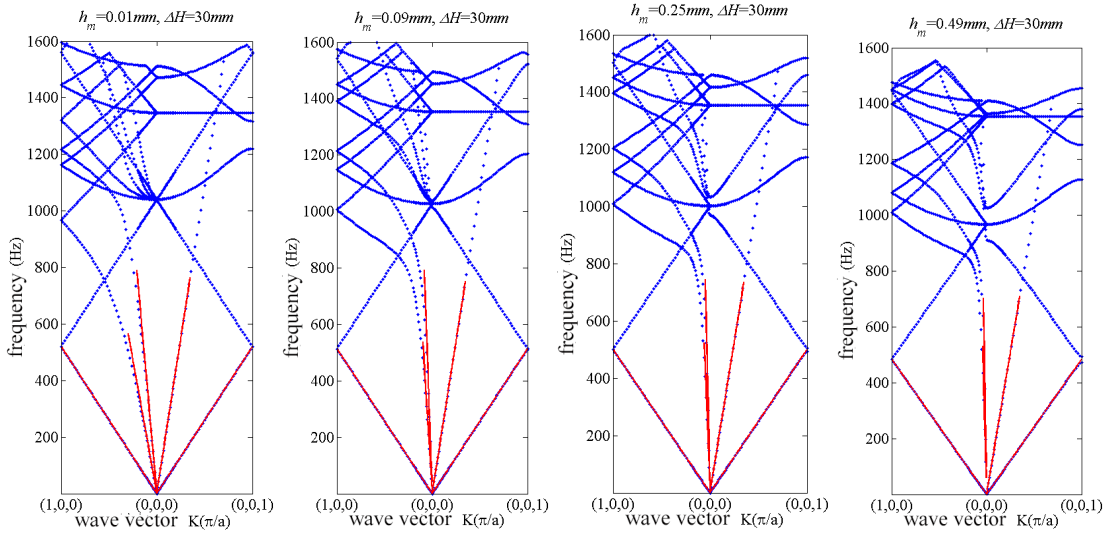
$$c = \frac{2\pi f}{|\mathbf{k}|} \tag{5}$$

Thus, the slope of every red straight line is corresponding to either the shear wave velocity or the longitudinal wave velocity.



(a) (b) (c) (d)

Figure 3. Dispersion curves of the unit cell with a single-layer membrane



(a) (b) (c) (d)

Figure 4. Dispersion curves of the unit cell with double-layer membranes

For an orthotropic material, the equivalent mass density ρ_0 can be calculated by

$$\rho_0 = \frac{\rho_m h_{mm} + \rho_r h_r}{h} \tag{6}$$

In a unit cell, ρ_m is the mass density of the membrane; h_{mm} is the total thickness of the membranes, for a unit cell with a single-layer membrane, $h_{mm} = h_m$, for a unit cell with double-layer membranes, $h_{mm} = 2h_m$; ρ_r is the mass density of rubber; h_r is the total thickness of rubber; h denotes the thickness of the unit cell and satisfies $h = h_r + h_{mm}$. In addition, equivalent modulus of elasticity of the unit cell E can be described as

$$E = \begin{bmatrix} E_{1111} & E_{1122} & E_{1133} & 0 & 0 & 0 \\ & E_{1111} & E_{1133} & 0 & 0 & 0 \\ & & E_{3333} & 0 & 0 & 0 \\ & \text{symmetry} & & E_{2323} & 0 & 0 \\ & & & & E_{1313} & 0 \\ & & & & & E_{1212} \end{bmatrix} \tag{7}$$

Where E is a symmetric matrix, E_{ijkl} stands for tensor of elasticity, and the subscripts $i, j, k, l = 1, 2, 3$ are corresponding to coordinate axes of x, y, z, respectively. Based on Christoffel's equation, we have

$$\left[\Gamma_{ij} - \rho_0 c^2 \delta_{ij}\right]\left[v_j\right] = 0 \tag{8}$$

in terms of

$$\Gamma = \left[\Gamma_{ij}\right] = \begin{bmatrix} E_{1111}n_1^2 + E_{1212}n_2^2 + E_{2323}n_3^2 & \left(E_{1122} + E_{1212}\right)n_1 n_2 & \left(E_{1133} + E_{2323}\right)n_1 n_3 \\ \left(E_{1122} + E_{1212}\right)n_1 n_2 & E_{1212}n_1^2 + E_{1111}n_2^2 + E_{2323}n_3 & \left(E_{1133} + E_{1212}\right)n_2 n_3 \\ \left(E_{1133} + E_{2323}\right)n_1 n_3 & \left(E_{1133} + E_{2323}\right)n_2 n_3 & E_{1212}n_1^2 + E_{1212}n_2^2 + E_{3333}n_3^2 \end{bmatrix} \tag{9}$$

and v_j, $\left(j = 1, 2, 3\right)$ are displacement components of a point in the unit cell, n_1, n_2 and n_3 represent direction cosines of wave numbers.

If $\lambda_{\max}$ is the maximum eigenvalue of matrix Γ, velocities of longitudinal waves in directions $\left(n_1, n_2, n_3\right)$ can be described by the formula

$$c_B = \sqrt{\frac{\lambda_{\max}}{\rho_0}} \tag{10}$$

For unit cells in Fig.1 and Fig.2, setting $a = 50mm$, $h_m = 0.5mm$, $\Delta H = 30mm$, we employ the asymptotic homogenization method (AHM) to calculate equivalent moduli of elasticity $E_{ijkl}, \left(i, j, k, l = 1, 2, 3\right)$, and results are listed in Table 1. In addition, by further analyzing, we figure out average velocities of elastic waves propagating in the unit cells, including longitudinal wave velocities c_{Bx} and c_{Bz}, shear wave velocities c_{Gxy}, c_{Gxz} and c_{Gyz}, which are shown in Table 2. Meanwhile, we analyze dispersion characteristics of the unit cells and compute average velocities in the low frequency range according to dispersion curves (DCM). The results are also in Table 2. By comparison, it can be found that result errors between

AHM and DCM are very little, which verifies the simulation results.

Table 1. Equivalent moduli of elasticity of unit cells / *MPa*

Unit cell	E_{1111}	E_{2222}	E_{3333}	E_{2323}	E_{1313}	E_{1212}	E_{1122}	E_{1133}	E_{2233}
Single-layer membrane	2354.6	2354.6	46.9	2.7	2.7	810.4	733.9	41.6	41.6
Double-layer membranes	4662.3	4662.3	46.9	2.7	2.7	1618.0	1426.2	41.6	41.6

Table 2. Wave velocities calculated by both methods / *m / s*

Unit cell	Methods	c_{Bx}	c_{Bz}	c_{Gxz}	c_{Gxy}	c_{Gyz}
Single-layer membrane	AHM	1498.3	211.4	50.3	879.0	50.3
	DCM	1490.5	210.4	50.1	874.8	50.1
Double-layer membranes	AHM	2042.5	204.8	48.7	1203.2	48.7
	DCM	2024.6	202.9	48.4	1192.8	48.3

Influence of structure parameters of unit cells on wave velocities

In order to investigate acoustic characteristics of the 3-D membrane-embedded-type metamaterials, we can analyze equivalent elastic moduli of unit cells and ratios between the elastic moduli, or study propagation velocities of elastic waves in unit cells and ratios between the velocities. In this section, we will research the influence of structure parameters of unit cells on wave velocities and wave velocity ratios.

Influence of thickness of membranes on wave velocities

For a unit cell with a single-layer membrane, defining velocity ratios of T_I and T_{II} as $T_I = c_{Bx} / c_{Gxz}$, $T_{II} = c_{Bz} / c_{Bx}$, we calculate average velocities of an elastic wave in the unit cell and velocity ratios, where the unit cell has different h_m. The results are shown in Table 3 and Fig.5, where T_I describes the coupling relationship between longitudinal waves and shear waves; T_{II} indicates anisotropic characteristics of the unit cell.

Table 3. Average wave velocities in the unit cell with different h_m / m / s

h_m (mm)	$\sqrt{h_m}$ ($\sqrt{mm}$)	c_{Bx}	c_{Bz}	c_{Gxz}	c_{Gxy}	c_{Gyz}
0.01	0.1	307.85	218.52	52.02831	138.43	51.98913
0.04	0.2	484.45	218.00	51.90377	261.15	51.86466
0.09	0.3	681.86	217.13	51.69815	385.87	51.65909
0.16	0.4	883.83	215.94	51.41430	509.66	51.37521
0.25	0.5	1085.10	214.44	51.05608	631.48	51.01708
0.36	0.6	1283.30	212.64	50.62853	750.64	50.58991
0.49	0.7	1476.90	210.57	50.13427	866.69	50.09949

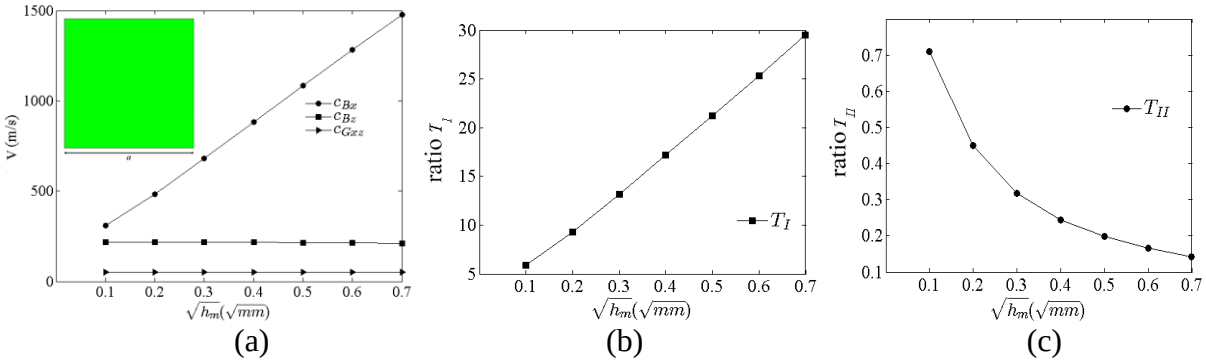
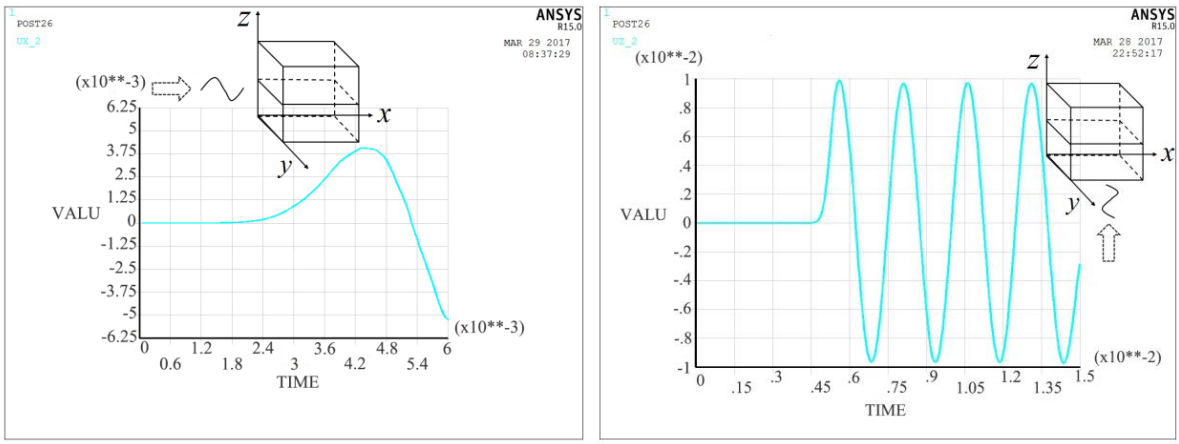


Figure 5. Relationship between wave velocities or wave velocity ratios and $\sqrt{h_m}$

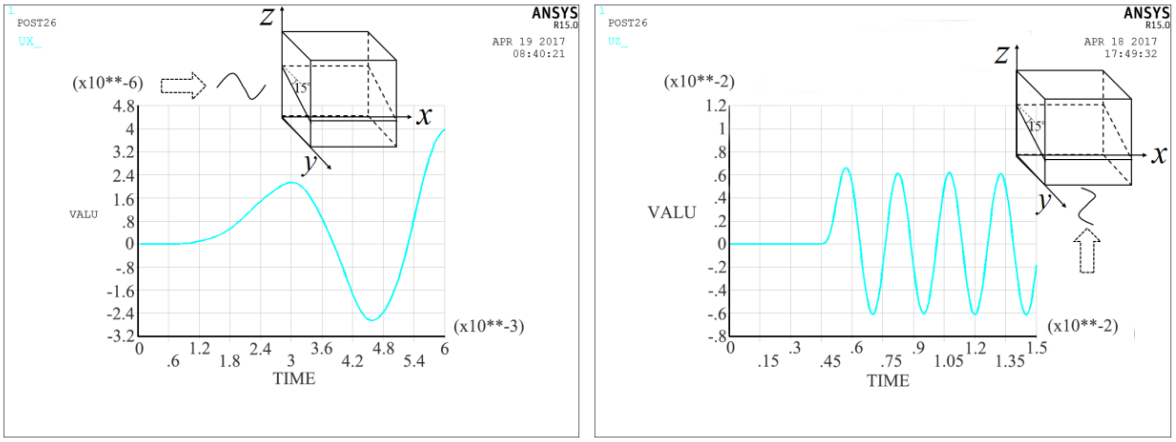
Seeing from Table 3 and Fig.5(a), we find that with the increase in $\sqrt{h_m}$, both the longitudinal wave velocity c_{Bx} and the shear wave velocity c_{Gxy} increase greatly; however, the longitudinal wave velocity c_{Bz} and the shear wave velocities c_{Gxz} and c_{Gyz} change very little. Consequently, the velocity ratio of T_I increases linearly with $\sqrt{h_m}$, and the velocity ratio of T_{II} reduces significantly, as shown in Fig.5(b) and (c). Obviously, the thickness of membranes h_m has a great influence on the anisotropic characteristics of the unit cell.

Influence of inclination angles of membranes on wave velocities

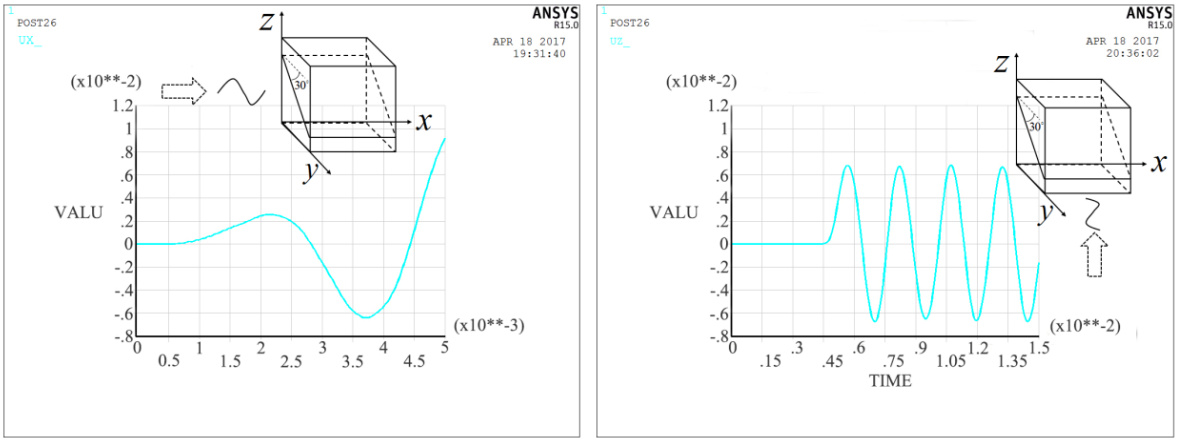
For a unit cell with a single-layer membrane, we set $h_m = 0.5mm$, $a = 50mm$ and use ψ to denote the inclination angle of the membrane, which is defined as the included angle between the normal vector of the membrane and the z axis. It is assumed that incident waves are either along the x-axis or the z-axis, we vary ψ from 0^o to 45^o with an interval of 15^o, and employ the transient response full integration method (TRM) to simulate displacement-time curves. Simulation results are in Fig.6, where Fig.6(a)-(d) are corresponding to different ψ.



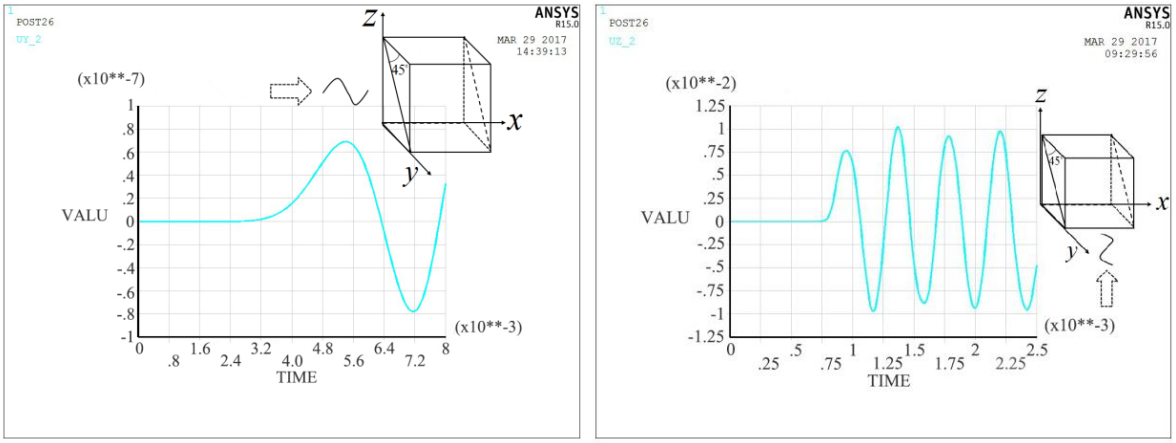
(a) $\psi = 0^\circ$



(b) $\psi = 15^{\circ}$



(c) $\psi = 30^{\circ}$



(d) $\psi = 45^{\circ}$

Figure 6. Displacement-time curves corresponding to different ψ

By further analysis, we calculate the average velocities of elastic waves, including the wave velocity along the x-axis v_x and the wave velocity along the z-axis v_z. Meanwhile, both the asymptotic homogenization method (AHM) and the dispersion characteristic analysis method (DCM) are also used to compute v_x and v_z. And all the wave velocities

calculated by the three methods are listed in Table 4. Seeing from Table 4, we find that wave velocities along the x-axis and the z-axis change smoothly while ψ increases from 0^o to 30^o, and when ψ reaches 45^o, both v_x and v_z change greatly, especially the wave velocities along the z-axis.

Table 4. Average wave velocities in the unit cell with different ψ

ψ	0^o		15^0		30^0		45^0	
Average wave velocity(m/s)	v_x	v_z	v_x	v_z	v_x	v_z	v_x	v_z
AHM	1498	211	1443	210	1521	212	1840	1275
DCM	1491	210	1446	209	1518	204	1741	1232
TRM	1429	211	1418	211	1485	210	1768	1290

Influence of pre-stress applied to membranes on wave velocities

We take unit cells with a single-layer membrane or double-layer membranes for examples, and establish finite element models as shown in Fig.7.

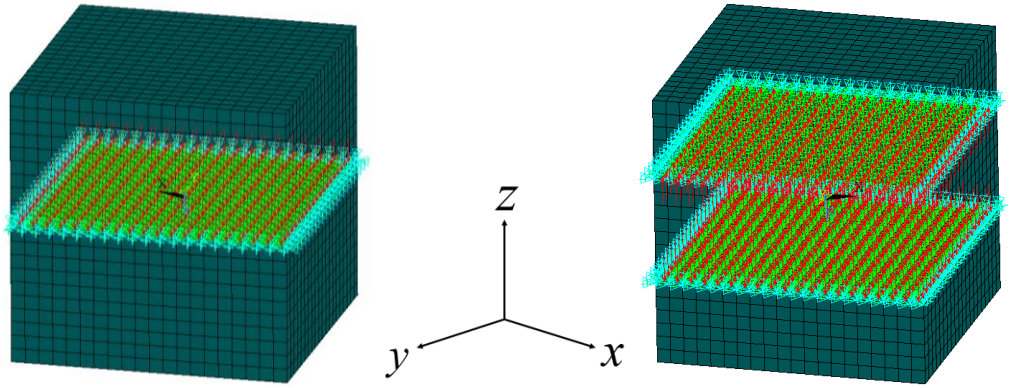
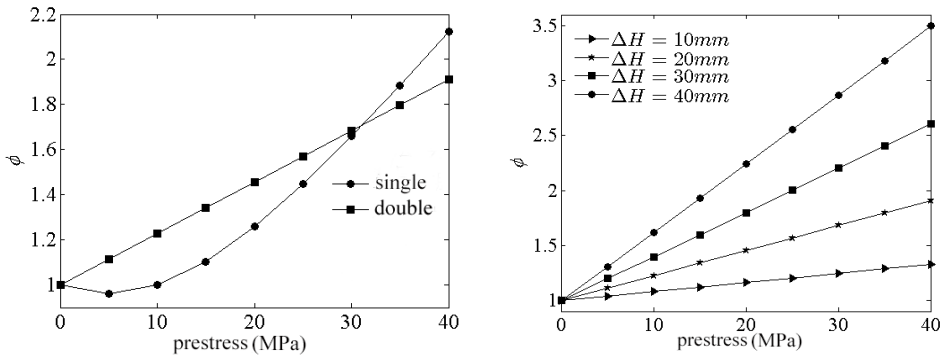


Figure 7. Finite element models of unit cells with pre-stressed membranes

It is assumed that $h_m = 0.5mm$, $a = 50mm$ and the in-plane pre-stress of the membranes is $T = 200N/m$. During the simulation process, a uniformly distributed load in the z axis is applied to the membranes to produce a certain pressure of P. In order to investigate the influence of the pressure on wave velocities, we vary the uniformly distributed load to make P applied to membranes change from $0\,MPa$ to $30MPa$. Here, ϕ is used to denote the velocity ratio, which is defined as $\phi = c_{Bz}/c_{Bz_0}$, where c_{Bz_0} and c_{Bz} stand for the average velocities of longitudinal waves while $P = 0MPa$ and $P \neq 0MPa$, respectively. Fig.8 shows the relationship curves between ϕ and P. In Fig.8(a), one curve is corresponding to the unit cell with a single-layer membrane, and the other corresponding to the unit cell with double-layer membranes. Fig.8(b) shows the simulation results of the unit cell with double-layer membranes, where ΔH is different. Seeing from Fig.8, we find that, for the unit cell with a single-layer membrane, the waves velocities along the z-axis c_{Bz} increases rapidly according to the increase in the pressure of P; for the unit cell with double-layer membranes, c_{Bz} increases linearly with P. In addition, when P reaches $30MPa$, ϕ of the unit cell with a single-layer membrane becomes bigger than that of the unit cell with

double-layer membranes. And meanwhile, for the unit cell with double-layer membranes, the distance between double-layer membranes also has a great influence on wave velocities, and c_{Bz} changes with the increase in ΔH.



(a) (b)

Figure 8. Relationship curves between ratios of the velocity and the pressure

Therefore, for a unit cell with an embedded single-layer membrane, we can adjust the pressure applied to the membrane to control the wave velocity along the z-axis; and for a unit cell with embedded double-layer membranes, we can change P and ΔH to make c_{Bz} reach a required value.

Acoustic characteristics of three-dimensional membrane-embedded-type metamaterials

In order to research acoustic characteristics of three-dimensional membrane-embedded-type metamaterials, we establish a finite element model as shown in Fig.9, which includes two parts, one is the isotropic rubber block and the other is the *N*-layer composite materials, where every layer of composite materials can be considered as a unit cell consisting of rubber and embedded single-layer or multi-layer membranes. Assuming that the incident wave is a plane wave with an incident angle of θ, the thickness of every unit cell is h, the equivalent modulus of elasticity of the unit cell in the ith layer is $E_i\left(i=1,2,\cdots\cdots,N\right)$, the thickness of a membrane is h_m, we simulate the wave propagation in the membrane-embedded-type metamaterial.

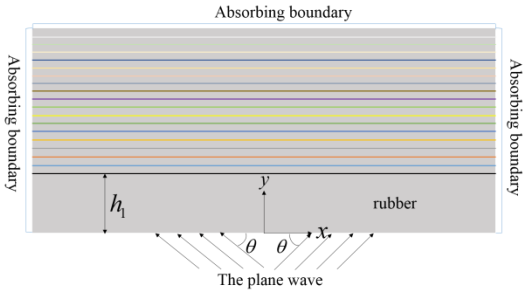
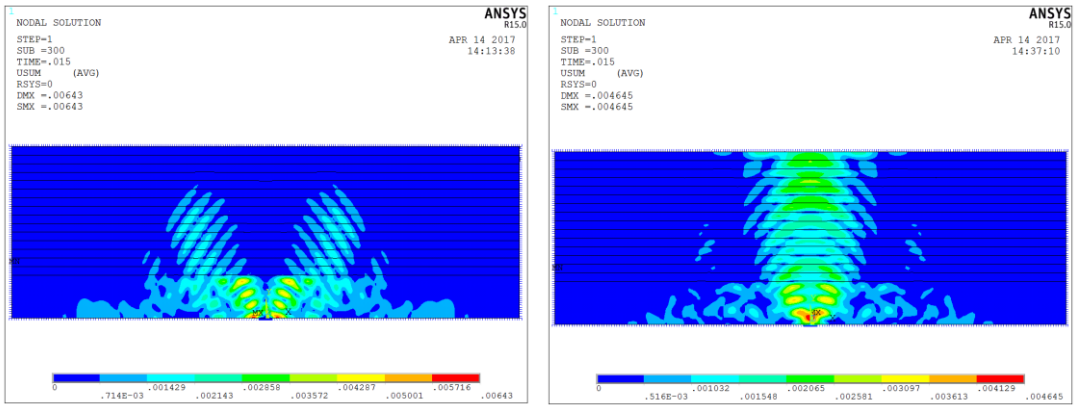


Figure 9. Finite element model

In this case, we set $E_1 = E_2 = \cdots\cdots = E_N$, $h_m = 0.5mm$, the modulus of elasticity of rubber is $E_r = 7.8\times10^6 Pa$, the modulus of elasticity of membrane is $E_m = 2.1\times10^{11} Pa$, and simulation results are shown in Fig.10, where Fig.10(a) and (b) are corresponding to $\theta = 30^o$ and $\theta = 90^o$, respectively. Obviously, in the metamaterial, the wave propagation path bends at the interface between the rubber block and the composite materials, but in the *N*-layer composite

materials, the wave propagation direction does not change. That is because the equivalent modulus of elasticity of every unit cell is same.



(a) $\theta = 30^o$ **(b)** $\theta = 90^o$

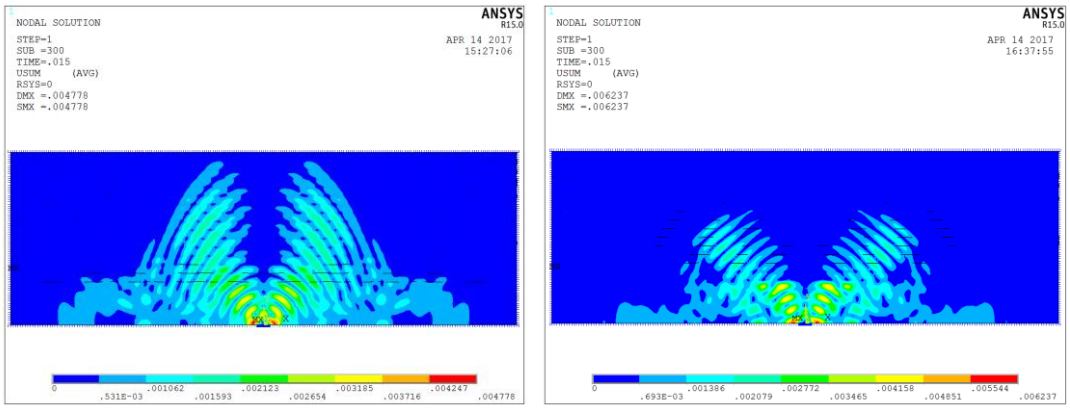
Figure 10. Color nephogram of the wave propagation

Next, in order to further research wave propagation in the membrane-embedded-type metamaterial, $E_i\ (i = 1, 2, \cdots\cdots, N)$ are given in the form of a geometric progression with a scale factor q, and in this way, the equivalent modulus of elasticity of the unit cell in the *ith* layer can be expressed as

$$E_i(y) = E_0 q^{i-1} \tag{11}$$

Where, E_0 is a given value; $y = h_1 + h(i-1)$, and y represents the coordinate position of the unit cell in the *ith* layer, h_1 is the embedded depth of the first layer membrane.

In one case, we set $\theta = 30^o$, $E_0 = 7.8 \times 10^6 Pa$, $q = 10.0$; and in another case, we set $\theta = 30^o$, $E_0 = 2.1 \times 10^{11} Pa$, $q = 0.1$. Simulation results are shown in Fig.11, where Fig.11(a) and (b) are corresponding to $q = 10.0$ and $q = 0.1$, respectively. Clearly, it can be observed that the wave propagation path bends not only at the interface between the rubber block and the composite materials, but also at the interfaces between any two adjacent unit cells, and meanwhile, with the increase in E_i, the deflection angle of wave propagation becomes bigger and bigger.



(a) $q = 10$ **(b)** $q = 0.1$

Figure 11. Color nephogram of the wave propagation

Conclusion

In this paper, we develop a new kind of 3-D membrane-embedded-type acoustic metamaterial, which is constructed by embedding high-modulus membranes into soft isotropic materials. Research shows the acoustic metamaterial has great potential for controlling propagation directions and propagation velocities of acoustic waves. The main conclusions are as follows:

Firstly, for unit cells of the acoustic metamaterial, the thickness of embedded membranes has a great influence on anisotropic characteristics of the unit cells. When the thickness of membranes varies in a certain range, the thicker the membrane is, the stronger the anisotropic characteristic of the unit cell is. Moreover, velocity ratios between longitudinal waves and shear waves almost increase linearly with the square root of the thickness of membranes.

Secondly, for unit cells of the acoustic metamaterial, inclination angles of embedded membranes have a certain influence on propagation velocities of elastic waves. When the inclination angle varies from 0^o to 30^o, wave velocities along the x-axis and the z-axis change smoothly; and when the inclination angle reaches 45^o, the wave velocities change greatly, especially the wave velocities along the z-axis.

Thirdly, for unit cells of the acoustic metamaterial, the pre-stress applied to membranes has a great influence on wave velocities. When the pre-stress varies in a certain range, propagation velocities of longitudinal waves mostly increase linearly with the pre-stress. In addition, for unit cells with multi-layer membranes, the distance between adjacent membranes also has a great influence on wave velocities, and propagation velocities of longitudinal waves change significantly with the increase in the distance between adjacent membranes.

Finally, the propagation directions and propagation velocities of low frequency acoustic waves in the 3-D membrane-embedded-type acoustic metamaterial can be designed by simply analyzing and calculating the equivalent moduli of elasticity of unit cells. Furthermore, the control of acoustic waves can be realized by adjusting structure parameters of every unit cell, including the thickness of membranes, the inclination angles of membranes and the pre-stress applied to membranes.

Acknowledgement

This research was financially supported by the National Science Foundation No.11372059 and No.11272073

References

[1] Andrew N. Norrisa, Acoustic metafluids, J. Acoust. Soc. Am, Vol.125, No. 2, p: 839-849, 2009
[2] Rui-Qi Li, Xue-Feng Zhu, Bin Liang, Yong Li, Xin-Ye Zou, Jian-Chun Cheng，A broadband acoustic omnidirectional absorber comprising positive-index materials, Appl. Phys. Lett., Vol. 99, p:193507（1-3），2011
[3] Mohamed Farhat, Sebastien Guenneau, and Stefan Enoch,Broadband cloaking of bending waves via homogenization of multiply perforated radially symmetric and isotropic thin elastic plates, Physical Review B 85, 020301(R) (2012)
[4] Yang Z, Mei J, Yang M, et al. Membrane-type acoustic metamaterials with negative dynamic mass[J]. Physical Review Letters, 2008, 101(20): 204301
[5] Jun Mei, Guancong Ma, Min Yang, Zhiyu Yang, Weijia Wen, Ping Sheng, Dark acoustic metamaterials as super absorbers for low-frequency sound
[6] Fuyin Ma, Jiu Hui Wu, Meng Huang, Weiquan Zhang and Siwen Zhang. A purely flexible lightweight membrane-type acoustic metamaterial[J]. Journal of Applied Physics, 2015, 48: 175105
[7] F. Langfeldt, J. Riecken, W. Gleine, O. von Estorff. A membrane-type acoustic metamaterial with adjustable acoustic properties[J]. Journal of Sound and Vibration, 2016, 373: 1-18
[8] Bogdan-Ioan Popa, Lucian Zigoneanu, and Steven A. Cummer，Tunable active acoustic metamaterials，Physical Review B，88(2013), 024303

Particle method simulation of violent sloshing under rotational excitation

***M. Luo and C.G. Koh**

Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576

†Corresponding author: ceelm@nus.edu.sg

Abstract

A three-dimensional (3D) numerical model is presented in the framework of Consistent Particle Method (CPM). The 3D gradient and Laplacian operators are derived based on Taylor series expansion, achieving good accuracy and largely alleviating the problem of spurious pressure fluctuation. Validated by our experimental study of water sloshing under rotational excitation, this model is shown to be robust and accurate in long time simulation of violent free surface flows which involve fluid merging and splitting.

Keywords: Particle method, Three-dimensional, Sloshing, Rotational Excitation.

Introduction

Liquid sloshing is a crucial issue in many engineering applications, one of which is the transportation of liquefied natural gas in membrane tanks on ships. The impact force induced by liquid sloshing may destroy the membrane layer and even cause structural failure of tank walls. In addition, large sloshing forces may capsize a ship when large overturning moments are generated. In this context, a better understanding of sloshing phenomenon is essential for the design of cost-effective LNG vessels.

With the rapid development of computer technology, numerical modelling has become increasingly feasible and many numerical algorithms have been developed to simulate sloshing problems. Among them are the mesh-based methods such as Finite Difference Method [1, 2], Finite Volume Method[3] and Finite Element Method[4]. These methods, however, may have some difficulties to model the large and discontinuous fluid motions which are generally involved in violent sloshing. In addition, the recognition or tracking of highly deformed free surface is also a tough issue (complicated and time-consuming) for mesh-based methods although some specialized schemes have been developed such as the Volume of Fluid [5] and the level set method [6].

In recent years, the particle methods such as Smoothed Particle Hydrodynamics (SPH), Incompressible Smoothed Particle Hydrodynamics (ISPH) and Moving Particle Semi-implicit (MPS) method have been developed and extensively used to model liquid sloshing as well as other free surface flows[7-10]. Because of the meshless nature, particle methods have better suitability in modelling merging and splitting of fluid and tracking of free surface. However, one of the challenging issues for these particle methods is the spurious pressure fluctuation. This is mainly caused by that the derivative approximation schemes in these methods which invoke a kernel function introduce numerical errors particularly for irregular particle distribution[11]. To address this issue, a new particle method named Consistent Particle Method (CPM) has recently been proposed [12] by adopting the Generalized Finite Difference scheme[13] to compute the spatial derivatives. Being consistent with Taylor series expansion and eliminating the use of a kernel function, CPM is fundamentally different from SPH, ISPH and MPS in terms of the derivative-approximation scheme. Due to the accurate computation of derivatives, the CPM solution of pressure history shows tremendous improvement over some

other particle methods[12], without the use of artificial viscosity or smoothing technique to remove spurious fluctuation.
In this context, the two-dimensional CPM is further developed into a 3D model. The accuracy of the developed numerical model is demonstrated by our experimental studies of violent water sloshing in a scaled tank under rotational excitation. The wave profiles, sloshing pressures at typical positions and the wave patterns are studied.

Consistent Particle Method

In particle methods, the fluid domain is represented by a collection of discrete Lagrangian particles, each of which carries a fixed mass and moves under external forces mainly arising from gravity and pressure difference. The governing equations are the Navier-Stokes equations in Lagrangian form as follows

$$\frac{1}{\rho}\frac{D\rho}{Dt}+\nabla\cdot\mathbf{v}=0 \tag{1}$$

$$\frac{D\mathbf{v}}{Dt}=-\frac{1}{\rho}\nabla p+\upsilon\nabla^2\mathbf{v}+\mathbf{g} \tag{2}$$

where ρ is the density of fluid, $\mathbf{v}$ the particle velocity vector, p the fluid pressure, μ the dynamic viscosity of fluid and $\mathbf{g}$ the gravitational acceleration.

In CPM, the above equations are solved by a predictor-corrector scheme. In the predictor step, the temporary particle velocities and positions are computed by neglecting the pressure gradient term. In the corrector step, a pressure Poisson equation (PPE) is solved as follows

$$\nabla\cdot\left(\frac{1}{\rho^*}\nabla p^{(k+1)}\right)=\frac{1}{\Delta t^2}\frac{\rho^{(k+1)}-\rho^*}{\rho^{(k+1)}} \tag{3}$$

The incompressibility condition is enforced by setting the fluid density at the current time step ($\rho^{(k+1)}$) to be the initial value (ρ_0). The intermediate fluid density (ρ^*) is evaluated based on the distance-weighted average of the masses of fluid particles including and around the reference particle[14].

In CPM, the 3D gradient and Laplace operators in the governing equations are computed in the following way

$$\frac{\partial p_i}{\partial x}=\sum_{j\neq i}\Big[w_j^2\left(a_1h_j+a_2k_j+a_3l_j+0.5a_4h_j^2+0.5a_5k_j^2+0.5a_6l_j^2+a_7h_jk_j+a_8h_jl_j+a_9k_jl_j\right)\left(p_j-p_i^{'}\right)\Big] \tag{4}$$

$$\frac{\partial^2 p_i}{\partial x^2}=\sum_{j\neq i}\Big[w_j^2\left(d_1h_j+d_2k_j+d_3l_j+0.5d_4h_j^2+0.5d_5k_j^2+0.5d_6l_j^2+d_7h_jk_j+d_8h_jl_j+d_9k_jl_j\right)\left(p_j-p_i\right)\Big] \tag{5}$$

where *i* is the reference particle and *j* the neighbor particles in the influence domain, w_j is the weighting function used in the weighted-least-square scheme and is adopted to be the inverse distance function [12, 13]. It is noted that this weighting function is essentially different from the

kernel function in SPH and ISPH (to approximate the Dirac delta function) and the weighting function in MPS (to specify the quantities diffused from the center particle to its neighbor particles). Similar to the two-dimensional CPM [12], $p_i^{'} = \min\left(p_j\right)$ for $\left\{j \,|\, r_{ij} \le r_e\right\}$ is used in Equation (4) to improve the numerical stability. Similarly, the first derivatives in the *y* and *z* directions can be computed by replacing a_s (s = 1, 2… 9) in Equation (4) with b_s and c_s, whereas the second derivatives in the *y* and *z* directions can be computed by replacing d_s (s = 1, 2, …, 9) in Equation (5) with e_s and f_s. The coefficients $a_s, b_s, c_s, d_s, e_s, f_s$ (s = 1, 2, …, 9) can be obtained as

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_9 \\ b_1 & b_2 & \cdots & b_9 \\ \vdots & \cdots & & \\ f_1 & f_2 & \cdots & f_9 \\ \vdots & \cdots & & \end{bmatrix} = \left(\sum_{j \neq i} C_j\right)^{-1} \tag{6}$$

where

$$C_j = w_j^2 \begin{bmatrix} h_j^2 & h_j k_j & h_j l_j & 0.5h_j^3 & 0.5h_j k_j^2 & 0.5h_j l_j^2 & h_j^2 k_j & h_j^2 l_j & h_j k_j l_j \\ & k_j^2 & k_j l_j & 0.5h_j^2 k_j & 0.5k_j^3 & 0.5k_j l_j^2 & h_j k_j^2 & h_j k_j l_j & k_j^2 l_j \\ & & l_j^2 & 0.5h_j^2 l_j & 0.5k_j^2 l_j & 0.5l_j^3 & h_j k_j l_j & h_j l_j^2 & k_j l_j^2 \\ & & & 0.25h_j^4 & 0.25h_j^2 k_j^2 & 0.25h_j^2 l_j^2 & 0.5h_j^3 k_j & 0.5h_j^3 l_j & 0.5h_j^2 k_j l_j \\ & & & & 0.25k_j^4 & 0.25k_j^2 l_j^2 & 0.5h_j k_j^3 & 0.5h_j k_j^2 l_j & 0.5k_j^3 l_j \\ & & & & & 0.25l_j^4 & 0.5h_j k_j l_j^2 & 0.5h_j l_j^3 & 0.5k_j l_j^3 \\ & & sym & & & & h_j^2 k_j^2 & h_j^2 k_j l_j & h_j k_j^2 l_j \\ & & & & & & & h_j^2 l_j^2 & h_j k_j l_j^2 \\ & & & & & & & & k_j^2 l_j^2 \end{bmatrix}$$

The free surface particles are recognized by the 3D spoke method developed by Luo et al[15]. The fixed particle approach is adopted to model the wall boundaries.

Water sloshing under rotational excitation

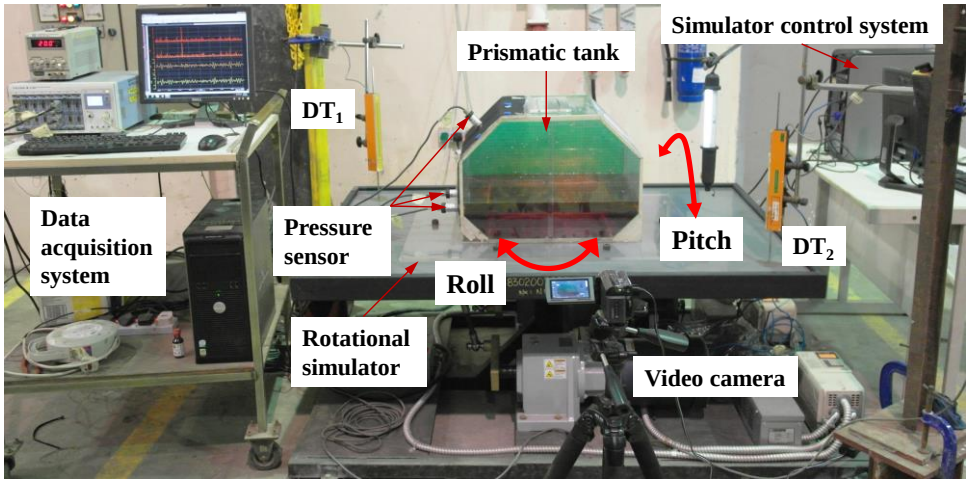


Figure 1. Water sloshing on rotational simulator

CPM is used to simulate more complicated sloshing phenomena under rotational excitations. Experimental studies are conducted on a rotational simulator as shown in Figure 1. The rotations about the *x* and *y* axes are respectively defined as pitch and roll as shown in Figure 2. A scaled tank (the dimension is as shown in Figure 3) fixed to the platform of the rotational simulator with the center of the tank bottom coinciding with the center of the top of the platform (point O' in Figure 2). Based on our parametric studies, the filling depth of d / H = 0.3 generates large sloshing pressure and is adopted as the case study in this section. Estimated by linear wave theory, the fundamental natural frequencies for the sloshing system in the roll and pitch directions are 5.598 rad/s and 7.471 rad/s respectively. Dynamic pressures at P_1 and P_2 are measured by gauge pressure sensors and the sloshing wave motions are captured by a video camera.

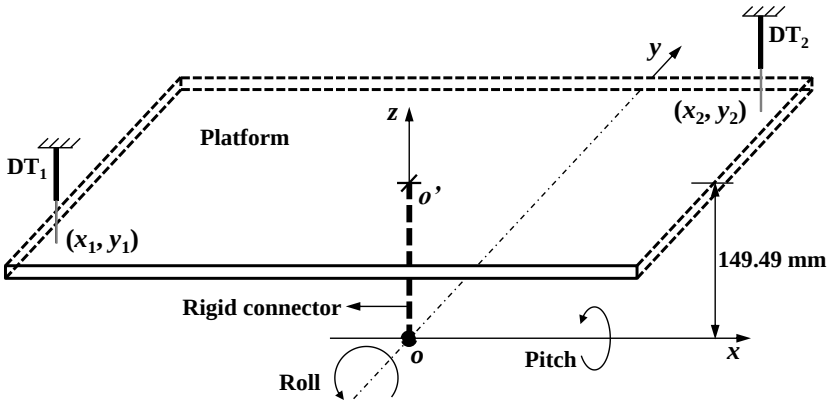
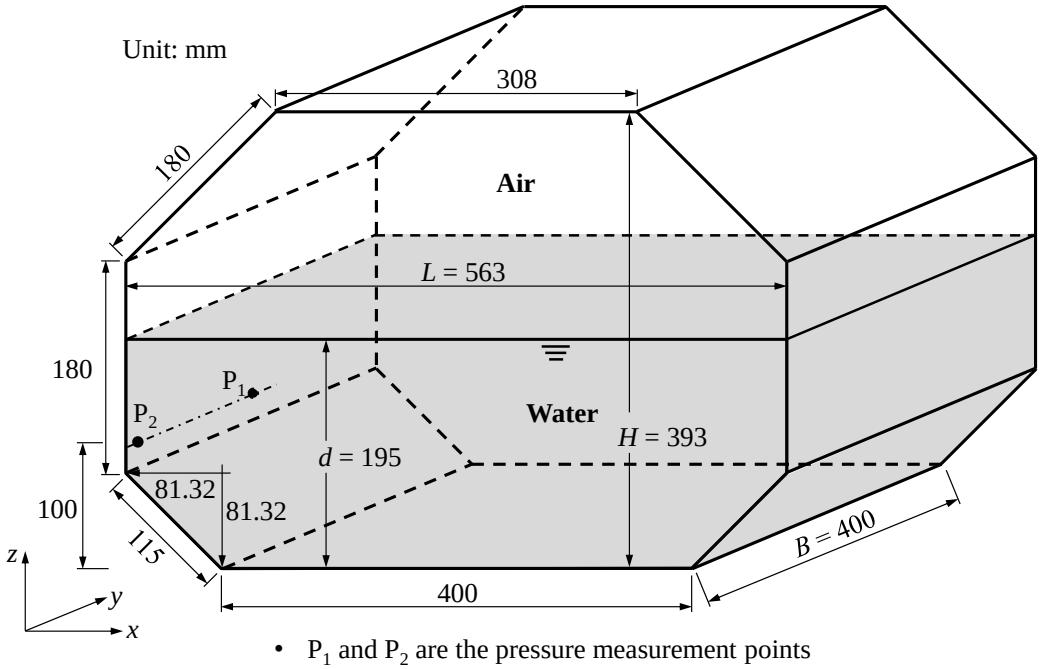


Figure 2. Schematic view of the rotational simulator



• P_1 and P_2 are the pressure measurement points

Figure 3. Geometric dimensions of water tank used in sloshing experiments

The rotational simulator is schematically shown in Figure 2. Two alternating-current motors generate rotations of the rigid platform about the *x* and *y* axes with point O as the pivot. A rigid connector O-O' is perpendicularly fixed to the platform and with length 149.49 mm. To

obtain the actual roll and pitch motions of the simulator, the vertical displacements at two diagonally opposite points on the simulator platform are measured by two vertically positioned displacement transducers (DT_1 and DT_2 in Figure 1 and Figure 2). Both displacement and pressure signals are recorded by a digital oscilloscope.

The excitation frequencies of roll and pitch rotations are set to be the same as the corresponding natural frequencies, i.e. 5.598 rad/s and 7.471 rad/s respectively. The measured rotations for the two directions are as shown in Figure 4. The pitch motion is not sinusoidal because of the control precision of the simulator. However, this does not influence the results since the real motions are used as the input of numerical simulations. In the numerical model, an initial particle spacing of 0.008 m (116,361 particles in total) and fixed time step of 0.0005 s are used. The wave profiles at different time instants are presented in Figure 5, showing generally good agreement between the numerical and experimental results. In the beginning, a swirling wave that rotates anticlockwise (viewed from top) is generated in the tank. Wave breaking occurs on the left wall of the tank at about 1.96 s. After t = 2.40 s, the swirling wave changes its direction to be clockwise which can be seen from the snapshots progressively from 2.40 s to 2.86 s. From about 5.26 s to 5.80 s, the swirling wave becomes anticlockwise again. The swirling wave can be explained by the superposition of the wave components in the length and breadth directions as

$$A(x,y) = A_x \cdot f_{1,0}(x,y) \cdot \sin(\omega_x t + \varphi_x) + A_y \cdot f_{0,1}(x,y) \cdot \sin(\omega_y t + \varphi_y) \tag{7}$$

where $f_{1,0}(x,y)$, A_x, ω_x and φ_x are respectively the first modal shape of the wave, amplitude of wave elevation, and the frequency and phase angle of the wave in the length direction, whereas $f_{0,1}(x,y)$, A_y, ω_y and φ_y are those parameters in the breadth direction. Because ω_x and ω_y are different, the swirling wave changes its rotary direction cyclically (with a "beating" frequency of $|\omega_x - \omega_y|$), which can also be seen from the pressure result presented in Figure 6. This observation is different from the sloshing phenomenon under 1-degree-of-freedom translational excitation[16], in which the rotary direction of swirling is dependent on the initial perturbation of wave and theoretically does not change.

A practical significance is that swirling waves generate large impact force near the corner of the tank as presented in Figure 6, which shows the sloshing pressures at P_1 and P_2. Due to the impact of swirling wave at the tank corner, the impact pressure at P_2 is larger than that at P_1 at the middle of the wall. 3D-CPM captures this phenomenon with the numerical results in generally good agreement with the experimental results. In particular, the large impact peaks at the tank corner are accurately predicted by the CPM as can be seen from the enlarged view of Figure 6. The relative difference between the CPM and experimental results is only 3.9% for the largest pressure peak. Some minor discrepancies exist between the numerical and experimental results of sloshing pressure at the reversals of swirling direction, the reason for which is as follows. The changes of wave swirling direction correspond to the changes of rotational direction of the simulator, which implies larger rotational accelerations. Since the tip of the displacement transducer is in contact with the top surface of the platform, friction force exists between the transducer and the platform when rotational acceleration of the platform is larger, thereby inducing errors in the measured rotational angles (particularly when there is a reversal of rotation). Since the measured rotational angles of the rotational simulator are used as the input of CPM simulation, such errors are brought into the numerical simulation. Although measures were taken to minimise the friction force between the

displacement transducer and the platform, it was not possible to completely eliminate this experimental error.

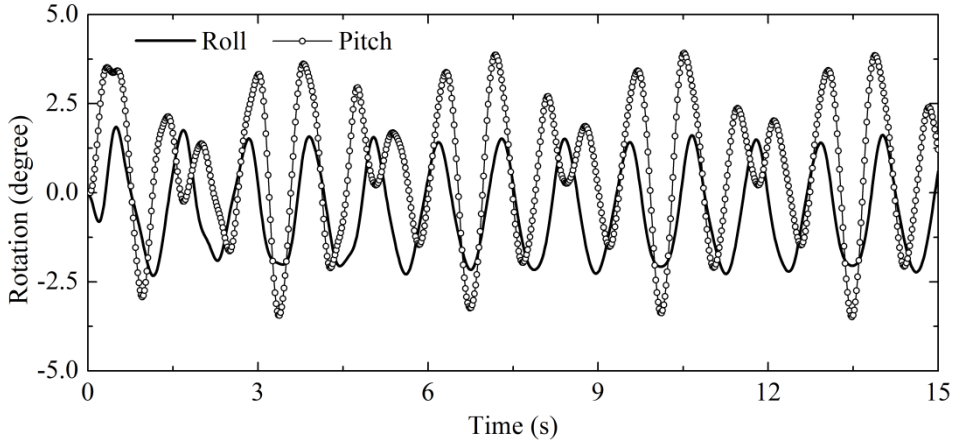


Figure 4. Water sloshing under resonant rotational excitation: roll and pitch angles

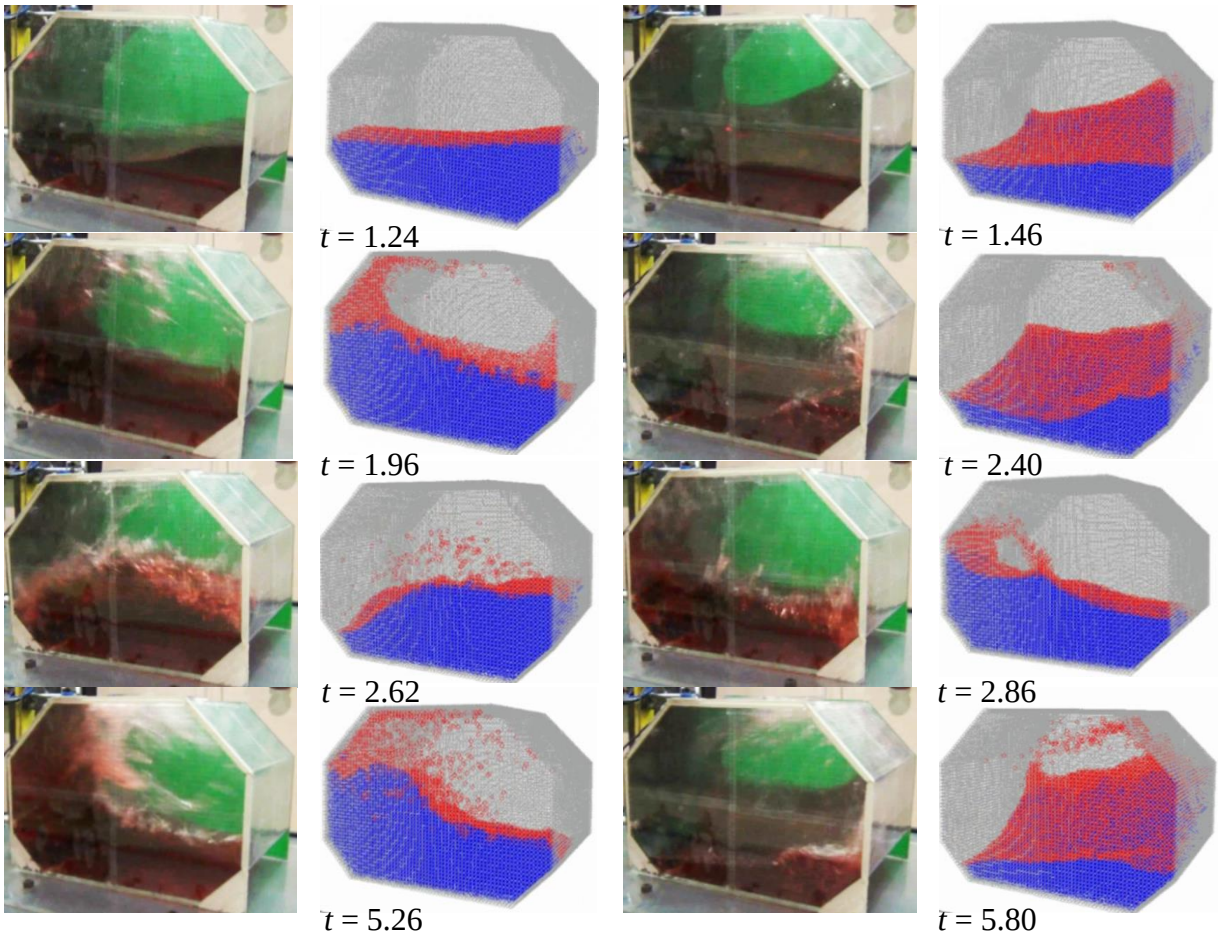


Figure 5. Water sloshing under resonant rotational excitation: wave profiles at typical time instants

Conclusions

In this study, the recently developed CPM is extended to simulate 3D sloshing waves. The 3D spatial derivatives are computed in a way consistent with Taylor series expansion, producing good accuracy even for irregular particle distributions. The 3D CPM is used to simulate the

water sloshing in a scaled tank under rotational excitations. These complex wave motions and the sloshing pressures predicted by the 3D CPM are in fairly good agreement with the experimental results.

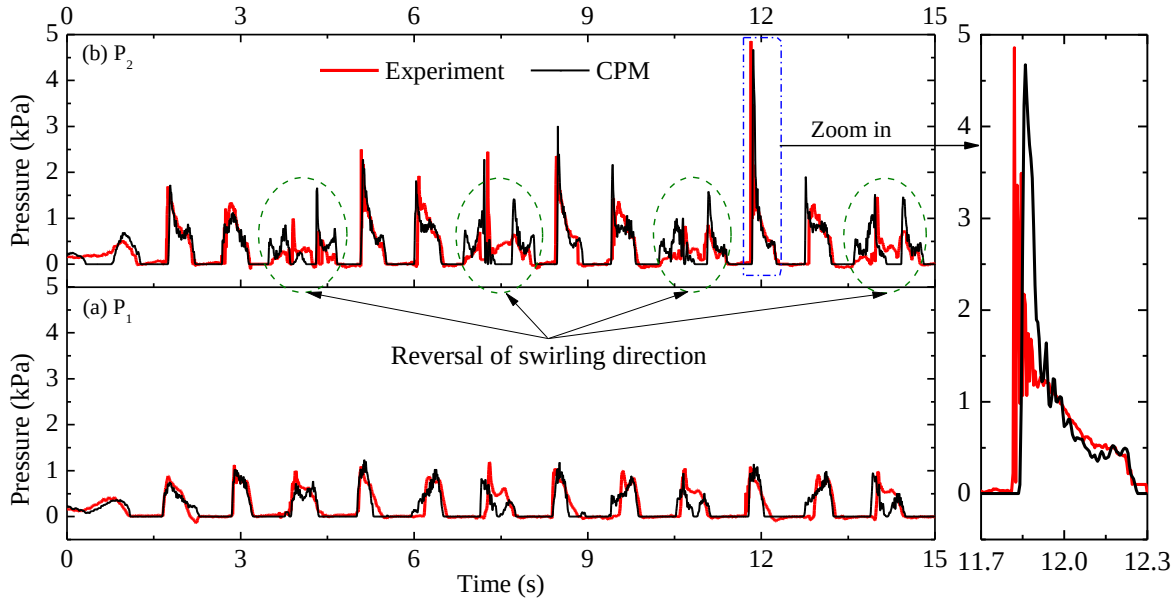


Figure 6. Water sloshing under resonant rotational excitation: pressure histories at P_1 and P_2

Acknowledgement

The authors appreciate the research grant provided by the Singapore Maritime Institute (Project SMI-2014-OF-02) as well as the funding and technical support of Sembcorp Marine Technology Pte Ltd.

References

[1] Liu D, Lin P. A numerical study of three-dimensional liquid sloshing in tanks. Journal of Computational Physics 2008;227:3921-39.
[2] Xue MA, Lin P. Numerical study of ring baffle effects on reducing violent liquid sloshing. Computers & Fluids 2011;52:116-29.
[3] Marsooli R, Wu W. 3-D finite-volume model of dam-break flow over uneven beds based on VOF method. Advances in Water Resources 2014;70:104-17.
[4] Wang CZ, Khoo BC. Finite element analysis of two-dimensional nonlinear sloshing problems in random excitations. Ocean Engineering 2005;32:107-33.
[5] Hirt CW, Nichols BD. Volume of fluid (VOF) method for the dynamics of free boundaries. Journal of Computational Physics 1981;39:201-25.
[6] Sussman M, Smereka P, Osher S. A level set approach for computing solutions to incompressible two-phase flow. Journal of Computational Physics 1994;114:146-59.
[7] Shao S, Lo EYM. Incompressible SPH method for simulating Newtonian and non-Newtonian flows with a free surface. Advances in Water Resources 2003;26:787-800.
[8] Koshizuka S, Nobe A, Oka Y. Numerical analysis of breaking waves using the moving particle semi-implicit method. International Journal for Numerical Methods in Fluids 1998;26:751-69.
[9] Shao JR, Li HQ, Liu GR, Liu MB. An improved SPH method for modeling liquid sloshing dynamics. Computers & Structures 2012;100-101:18-26.
[10] Liu X, Lin P, Shao S. An ISPH simulation of coupled structure interaction with free surface flows. Journal of Fluids and Structures 2014;48:46-61.
[11] Luo M, Koh CG, Gao M, Bai W. A particle method for two-phase flows with large density difference. International Journal for Numerical Methods in Engineering 2015;103:235-55.

[12] Koh CG, Gao M, Luo C. A new particle method for simulation of incompressible free surface flow problems. International Journal for Numerical Methods in Engineering 2012;89:1582–604.
[13] Liszka T, Orkisz J. The finite difference method at arbitrary irregular grids and its application in applied mechanics. Computers & Structures 1980;11:83-95.
[14] Koh CG, Luo M, Gao M, Bai W. Modelling of liquid sloshing with constrained floating baffle. Computers & Structures 2013;122:270-9.
[15] Luo M, Koh CG, Bai W. A three-dimensional particle method for violent sloshing under regular and irregular excitations. Ocean Engineering 2016;120:52-63.
[16] Faltinsen OM, Rognebakke OF, Timokha AN. Classification of three-dimensional nonlinear sloshing in a square-base tank with finite depth. Journal of Fluids and Structures 2005;20:81-103.

The Numerical Investigation on Hydrodynamic Performance of Twisted Rudder during Self-propulsion

Cong Liu[*], Jianhua Wang，Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: maximusliu@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

In the present work, Comparison of Energy Efficiency Performance between a twisted rudder and an ordinary spade rudder is conducted during self-propulsion. The energy-saving mechanism of twisted rudder is revealed by numerical simulation. This study is carried out using our solver naoe-FOAM-SJTU. The governing equations are unsteady Reynolds-Averaged Navier-Stokes (URANS) equations discretized by finite volume method (FVM). Overset technique is employed to handle the propeller rotation. To prepare for self-propulsion simulations, the open water curves of propeller and bare hull resistance are previously obtained by our solver. These two results show good agreement with experimental data. Based on that, the self-propulsion simulations to estimate the speed performance with twisted rudder and spade rudder are carried out. The results show that twisted rudder boosted hull efficiency by 2.4% through reducing the thrust deduction fraction and raising the hull efficiency. The delivered power is also decreased by 3.9% in the model scale comparing with ordinary rudder. The flow field around the propeller- rudder system shows that twisted rudder efficiently retrieves y-direction momentum to x-direction and improves the propulsive performance.

Keywords: twisted rudder; self-propulsion; overset grids; naoe-FOAM-SJTU solver

Introduction

In the context of global warming, the International Maritime Organization (IMO) issued mandatory standard which requires all ships larger than 400 gross tonnage reducing Energy Efficiency Design Index (EEDI) by up to 30% after 2025. To achieve this goal, more efficient rudders which can increase the propulsion are helpful. As we all known, the propeller produces thrust by rotating around its shaft which inevitably induces swirl energy loss. Aiming to retrieve this residual energy, several scholars designed a twisted rudder which can transfer part of rotation energy to propulsion. For example, a benchmark case, Duisburg Test Case(DTC), was designed to equip with a twisted rudder[1]. The energy-saving mechanism of twisted rudder can be explained as follow. The propeller creates a pair of opposing wake flows around the propeller shaft center. Generally, for a clockwise-rotating propeller, wake flow above the shaft centre turns toward port side, then the twisted rudder's leading edge is designed to twist port and vice versa below the shaft. Since the angle of attack of the twisted rudder is smaller than that of the horn-type rudder, the twisted rudder profile can decrease drag and lift to some extent [2]. Twisted rudder can also transfer the cross-stream velocity to stream wise more efficiently and improve the propulsive efficiency. A twisted rudder with a bulb may have 4% less fuel consumption[3]. Sun et al.[4] studied the hydrodynamic characteristics of twisted rudder through both experimental and numerical method. The

energy-saving mechanism and cavitation performance were analyzed without taking the hull body in to account. Kim et al.[2] reported that a Z-twisted rudder, which has a Z-shaped leading edge, with and without a fin may reduce the fuel consumption by 2.35% and 2.95% via self-propulsion model tests. In the paper, CFD method was also used to analyze the fluid characteristics of the twisted rudder. but the effect of free surface was neglected. They assumed that the twisted rudder has no influence on propeller' RPS, so the self-propulsion condition was simplified by fixing the model and keeping propeller's rotational speed at constant value.

In this paper, taking the rudder's effect on PRS into account, the author tries to estimate the speed performance for a ship with twisted and ordinary rudders during self-propulsion. Several validations and preparations need to be done before the self-propulsion simulations. To estimate the wake fraction during the self-propulsion, the open water curves of propeller is firstly calculated by our solver. Then to estimate the friction deduction, the resistance of towed ship without propeller is obtained. Based on that，the propulsion factors for two kinds of rudders are obtained from the self-propulsion simulations using our solver, naoe-FOAM-SJTU. In our solver, a six-degree-of-freedom (6DoF) module coupled with overset grid technique[5] is utilized to simulate the rotation of propeller and predict the motions of ship. This solver has been successfully applied in the region of predicting the ship maneuverability. For example, Wang et al.[6] used our solver to simulate ship self-propulsion with moving propellers and rudders. Furthermore, Wang et al.[7] carried out turning circle simulation by our solver.A twisted rudder and an ordinary spade rudder are compared. The predicted thrust deduction and hull efficiency are validated with the experimental data[1].

This paper is organized as follows: In the first section, a brief introduction of the numerical methods is given. Then an open-water curves and resistance of bare hull are predicted. Next, the simulations of the self-propulsion are following. Then the differences in propulsion factors and hydrodynamic characteristics between two kinds of rudder are analyzed. Finally, a summary of the paper is presented.

Numerical method

Governing Equations

The incompressible Navier-Stokes equations are the governing equations, which can be written as:

$$\nabla \cdot \boldsymbol{U} = \boldsymbol{0} \tag{1}$$

$$\frac{\partial \rho \boldsymbol{U}}{\partial t} + \nabla \cdot \left(\rho\left(\boldsymbol{U} - \boldsymbol{U_g}\right)\boldsymbol{U}\right) = -\nabla p_d - \boldsymbol{g} \cdot \boldsymbol{x}\nabla\rho + \nabla \cdot \left(\mu_{eff}\nabla \boldsymbol{U}\right) \tag{2}$$

where U is fluid velocity field and $\boldsymbol{U}_g$ is the grid velocity; $p_d = p - \rho \boldsymbol{g} \cdot \boldsymbol{x}$ is the dynamic pressure, obtained by subtracting the hydrostatic component from the total pressure; ρ is the mixture density; $\boldsymbol{g}$ is the gravity acceleration; $\mu_{eff} = \rho(\nu + \nu_t)$is effective dynamic viscosity, in which ν and ν_t are kinetic and eddy viscosity, respectively, and ν_t is obtained from turbulence model.

Overset Grid Technique

Overset grid is a grid system composed of multiple blocks of overlapping structured or unstructured grids. In a full overset grid system, a complex geometry is decomposed into a

system of geometrically simple overlapping grids. Boundary information is exchanged between these grids via interpolation of the fluid variables. Through this way overset grid method removes the restrictions of the mesh topology among different objects and allows grids move independently within the computational domain, and can be used to handle with large amplitude motion in the field of ship and ocean engineering. The most critical technique in the overset is the accomplishment of information exchange between grids. Based on the numerical methods from OpenFOAM including the cell-centered scheme and unstructured grids, SUGGAR++ is utilized to generate the domain connectivity information (DCI) for the overset grid interpolation in our solver naoe-FOAM-SJTU. Through this way, the solver can handle with arbitrary motion in the simulation.

Computations and Results

Open water propeller test

Before we move on to the self-propulsion test, the open water curve is predicted using overset grid technique. This step has two purposes. First, it is used to validate the basic dynamic overset grids strategy and implementation by comparison with experimental data. Second, the open water propeller performance is needed for a full CFD prediction of the self-propulsion factors.

The propeller is a fixed-pitch five-bladed propeller with right rotation. This propeller is equipped on Duisburg Test Case(DTC). Table 1 shows its main particulars: propeller diameter D_P ,pitch ratio at 0.7 of propeller radius $P_{0.7}/D_P$, disc ratio A_e/A_0 , chord length $c_{0.7}$ at 0.7 of propeller radius, effective skew angle of propeller blades θ_{eff} and non-dimensional hub diameter d_h/D_P. Figure 1 shows it geometric feature.

Table 1 Propeller parameters

	Model scale	Full Scale
D_P [m]	0.15	8.911
$P_{0.7}/D_P$ [m]	0.959	0.959
A_e/A_0 [m]	0.8	0.8
$c_{0.7}$ [mm]	0.054	3.208
θ_{eff} [°]	31.97	31.97
d_h/D_P [-]	0.176	0.176

Figure 1 view of the propeller

An earth-fixed frame of reference is used where the propeller is moving with constant rotational speeds and towed with a small acceleration to cover a wide range of advance velocity in a single run. The computations use the single-run procedure described in Xing et al.[10] The grid system is composed of two parts: overset grid used to follow the rotation of propeller and background grid used to accommodate the far-field boundary conditions (see Figure 2). The grids interpolated by Suggar++ are illustrated in Fig. 3. The two grids are separately generated by SnappyHexMesh, an automatic mesh-generation utility provided by OpenFOAM. The sizes of these grids are listed in Table 2. The time step is set to $\Delta t = 2 \times 10^{-4} s$.The RPS is fixed at 13.5 n / s and thrust and torque coefficients K_T and K_Q, and efficiency η_0 for each advance coefficient are obtained from the thrust and torque as showed in Figure 4,. In the single-run procedure, the propeller is accelerated from J=0 to1 in 5 s.

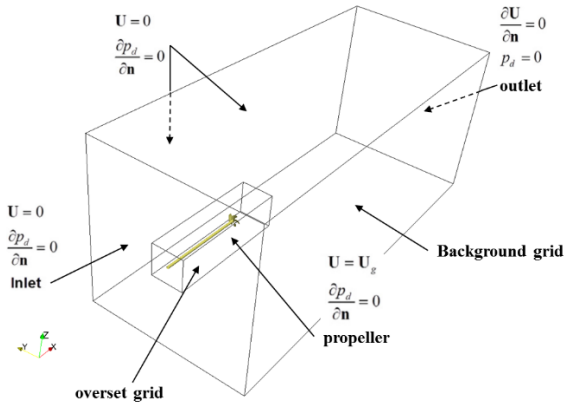


Figure 2 Design of the over set grid system and boundary conditions for the propeller.

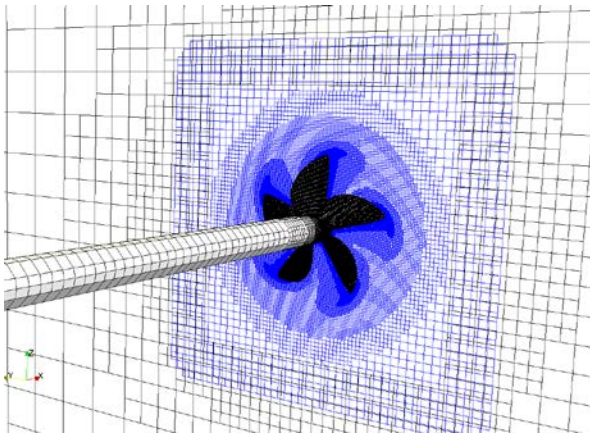


Figure 3 Grids used for open water computations overset

Table 2 mesh size for open water propeller

	background	Overset	Total
Cells number	116,963	3,060,690	3,177,653

Figure 4 shows that the overset technique predicts the curves quite well and results have a good agreement with experimental data, except for J > 0.8 where η_0 is over estimated. Figure 5 illustrates the vortical structure using isosurfaces of Q=200, the second invariant of the velocity gradient tensor. The vortices shed from three position: tips, trailing edges and hub. These results indicate that the overset technique is capable of predicting the open-water curves and will be used to estimate the effective wake friction in the condition of self-propulsion.

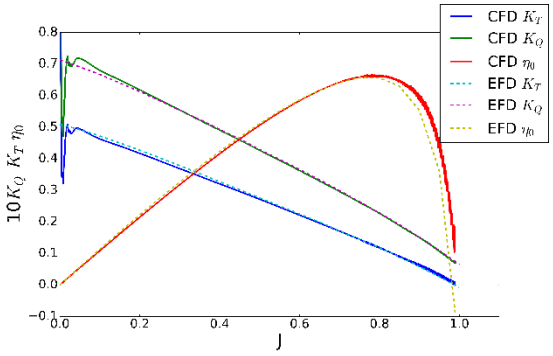


Figure 4 Open water curves for experiments(dash lines),single-run(solid lines)procedures

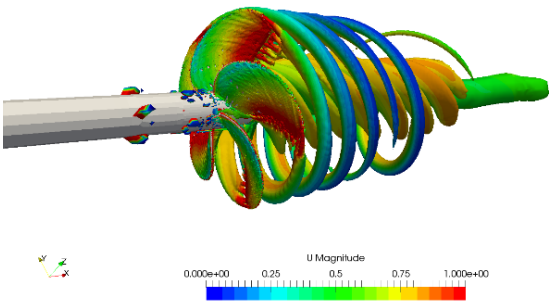


Figure 5 Iso-surfaces of Q=200 at J=0.61

Towed DTC without propeller

Before self-propulsion, the hull resistance should be calculated previously for two reason. In the case of self-propulsion, the hull resistance R_T needs to be obtained, which will be used to estimate the friction deduction and the convergence result of resistance calculation will be used to initialize the initial condition of the self-propulsion to achieve faster convergence.

The model chosen to calculate is Duisburg Test Case(DTC) and the experiments data of hull resistance and self-propulsion can be found in el Moctar[1]. Its characteristics are shown in Figure 6, and its principal particulars are listed in Table 3: Length between perpendiculars L_{pp} , waterline breadth B_{wl} , draught midships T_m, trim angle θ, volume displacement V, block coefficient C_B , wetted surface under rest waterline without appendages S_w .DTC design features a twisted rudder with Costa bulb and a NACA 0018 base profile (see Figure 6, bottom).

Table 3 Main dimensions of DTC in design loading condition

		Model scale	Full scale
L_{pp}	[m]	5.976	355
B_{wl}	[m]	0.859	51
T_m	[m]	0.244	14.5
θ	[]	0	0
V	[m3]	0.827	173467
C_B	[-]	0.661	0.661
S_w	[m2]	6.243	22032

Figure 6 View of Duisburg Test Case hull(top), rudder (bottom))

The speed is 1.668 m/s, corresponding to *Fr=0.218* and the time step is $1 \times 10^{-3}s$. The results compared with experimental data can be find in Table 4 ,including resistance R_T, and its non-dimensional coefficient $C_T = R_T / (0.5 / \rho S_w v^2)$, frictional resistance R_F and its non-dimensional coefficient $C_F = R_F / (0.5 \rho S_w v^2)$, and non-dimensional wave resistance coefficient $C_W = R_W / (0.5 \rho S_w v^2)$. For experimental data, the frictional resistance coefficient C_F was calculated according to ITTC 1957 as $C_F = 0.075 / (log_{10} Re - 2.0)^2$ and the wave resistance coefficient was estimated as $C_W = C_T - (1 + k) C_F$,where form factor $k = 0.094$.

Table 4 Resistance coefficients for DTC bare hull with twisted rudder

	$R_{T\,[N]}$	$R_{F[N]}$	$C_T \times 10^{-3}$	$C_F \times 10^{-3}$	$C_W \times 10^{-4}$
Experimental data	31.83	26.431	3.67	3.047	3.36
Present work	30.19	25.38	3.48	2.928	2.821
Error	-4.39%	-4.14%	-5.46%	-4.06%	-5.66%

The comparisons in Table 4 show quite well agreement between CFD and experimental data. The hull resistance coefficients are underestimated by CFD in the range 4.06%-5.66%. This results show that our solver has the ability to precisely predict the hull resistance. The total hull resistance coefficients C_T obtained by CFD will be used to calculate the friction deduction.

Self-propelled DTC with twisted and spade rudder.

For self-propulsion, DTC is equipped with a directly rotating propeller, which provides the thrust for the ship to advance. The self-propulsion is performed at ship point by adding a friction deduction force of the form

$$F_D = 0.5\rho_M S_{WM} V_M^2 (C_{TM} - C_T) \quad (5)$$

where index M denotes model scale. C_{TM} =3.48×10^{-3} obtained from previous CFD results. The total resistance coefficient C_T is defined at the corresponding temperature. Then F_D=13.41 N

Following Carrica et al.[7], aproportional-integral(PI) controller is employed to adjust the rotational speed of the propeller to achieve the desired ship speed. The instantaneous RPS of the propeller, n, is obtained as $n = Pe + I\int edt$,where P and I are proportional and integral constants, respectively, and e is the error between target ship speed and instantaneous speed, $e = U_{target} - U_{ship}$ .Using the PI controller, we will get the final balance condition at certain RPS of propeller where $T = R_T - F_D$ and $U_{ship} = U_{target}$ .This procedure resembles the self-propulsion experiment using continental method.

The assessment of the rudder performance can be conduct in two ways. The first way is to conduct the assessment at constant propeller's rotational speed n. This method presumes that there is a negligible change in n due to difference in the twisted rudder's design parameter. Through this way, it is convenience to compare the flow fields and the rudders which the resistance of rudder-propeller system is smaller would be considered to have a better rudder performance. In Kim's work, the calculation part was conducted in this indirect way.

But the comparison would best be taken at the self-propulsion point. Through the PI controller, it is possible to directly compare the influence of rudder shape on the propulsive performance at self-propulsion point. For the ships equipped with twisted or ordinary spade rudder， use the PI controller to achieve the target speed and obtain the values of RPS of propeller, propeller thrust and hull efficiency respectively. Then the influence of rudder on propeller performance can be obtain by comparing these values. The merit of this method is that the propeller performance can be compared in an intuitive and practical way. In Kim's work, the experiment part was conducted in this direct way.

In present work, numerical calculation is use to achieve the comparison both at the self-propulsion point and at constant propeller's rotational speed. In computation, the origin of the Earth coordinate system is located at the intersection of the waterline and the ship's bow with the longitudinal x-axis pointing fore to aft, the transversal y-axis pointing from port to starboard, and the vertical z-axis pointing upward. The profile of two kind of rudder is showed in Figure 7 and Figure 8.The spade rudder is obtained by extending the twisted rudder's root profile along the 1/4 chord line and keeping the chord length invariant. The layout of grids system and boundary conditions for self-propulsion calculation can be found in Figure 9.

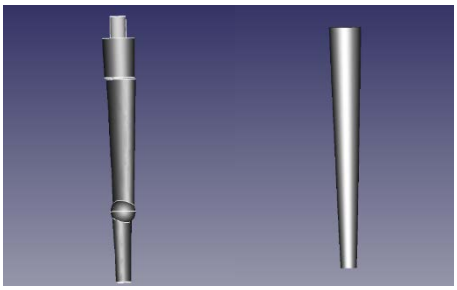


Figure 7 front view of rudders(left: twisted rudder, right: spade rudder)

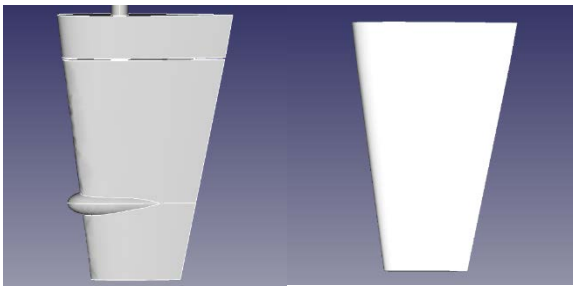


Figure 8 side view of rudders(left: twisted rudder, right: spade rudder)

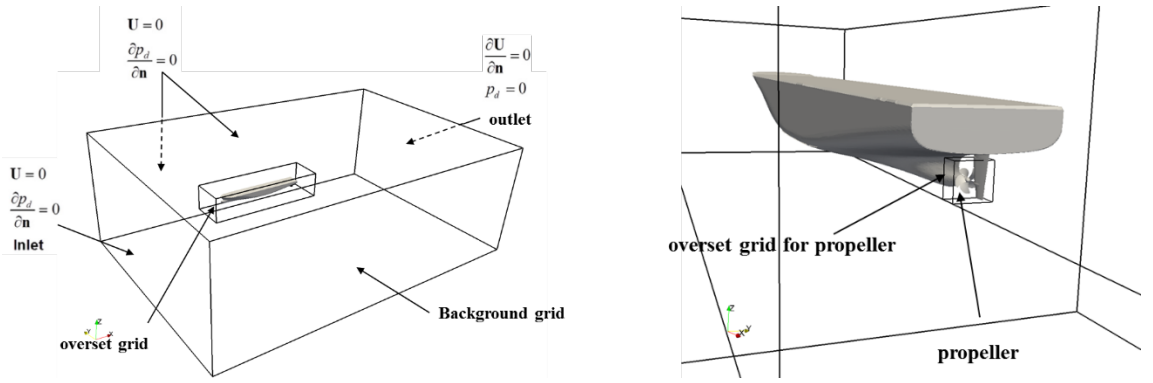


Figure 9 Overset grid system and boundary conditions for self-propelled DTC (left :global view , right: close view)

The time histories of RPS and thrust forces are shown in Figure 10 and Figure 11 respectively. The computation was run for 9s of model scale time. After 6s, the RPS converged to the self-propulsion point.

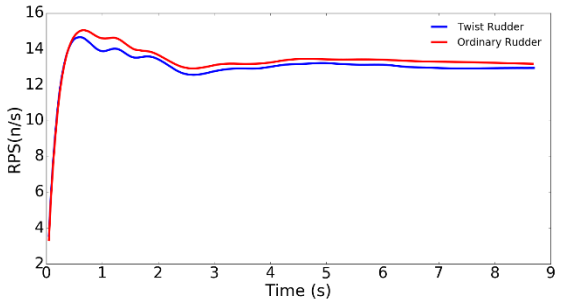


Figure 10 Time history of RPS(blue line represents twisted rudder and red line represents ordinary rudder)

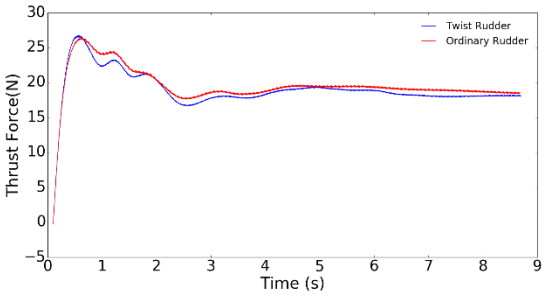


Figure 11 Time history of Thrust forces (blue line represents twisted rudder and red line represents ordinary rudder)

The experimental data of Self-propelled DTC equipped with twisted rudder can be found in el Moctar[1].Table 5 shows the experimental data compared with present work results

Table 5 Resistance coefficients and propulsion factors for self-propelled DTC.

		Present	Experiment data	Error
Resistance Coefficience	C_T	3.48E-03	3.67E-03	5.1%

Thrust deduction	1-t	0.957	0.91	5.2%
Effecticve wake coefficient	1-w	0.762	0.725	-5.1%
Hull efficiency	η	1.256	1.26	0.1%

The RPS of propeller equipped with twisted rudder or ordinary spade rudder at self-propulsion points are 12.8n/s and 13.1n/s, respectively. Due to the twisted rudder, the propeller's RPS decreases by 2.84%. The torque of twisted rudder and ordinary spade are $0.355\,N\cdot M$ and 0.361 $N\cdot M$ corresponding to delivered power P_D are 28.55 W and 29.71 W, respectively. P_D for twisted rudder decreased by 3.9%. Table 6 shows the difference of propulsion factors for two type rudders.

Table 6 Propulsion factors for self-propelled DTC with different rudders

	Thrust coefficient K_T	Thrust Deduction t	Wake fraction ω	Hull efficiency η_H
Twisted rudder	0.202	0.043	0.238	1.256
Spade rudder	0.198	0.058	0.232	1.23
Difference	1.9%	-25.8%	2.59%	2.4%

The main role of the twisted rudder is to lower the t to raise the hull efficiency (η_H). The thrust deduction drop (25.8%) is because the required thrust forces reduces. The rudder causes the wake fraction, ω, to rise 1.8%. The hull efficiency (η_H) rises (2.4%) from the reduced t and increased ω.

To demonstrate the energy-saving mechanisms of twisted rudder, the self- propulsion simulation is conducted in a more restrained way, namely, the ship speed and RPS of propeller are the same for the ships equipped with two type rudder. As we argued before, twisted rudder profiles transferring the cross-stream velocity to stream wise retrieve part of the rotation energy to propulsion. So it is natural to compare the distribution of velocity components around the propeller-rudder system.

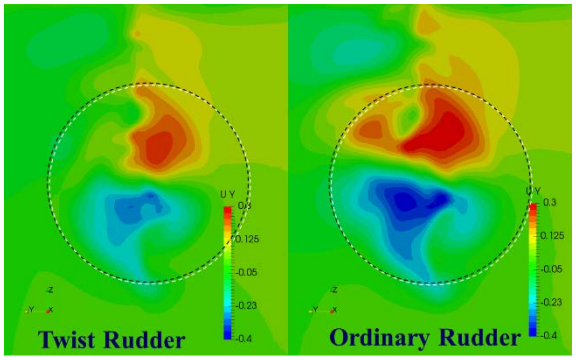


Figure 12 y-direction velocity components contours downstream of the rudder.

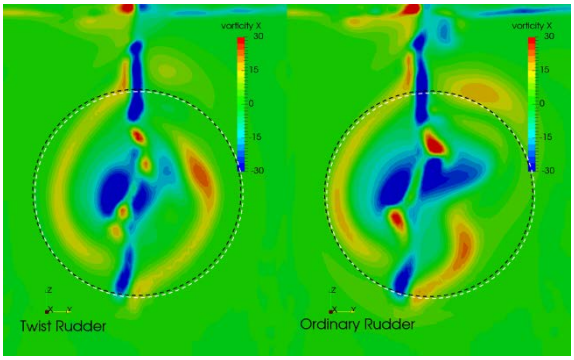


Figure 13 Axial vorticity contours downstream of the rudder

Figure 12 shows the y-direction velocity components contours downstream of the two kinds of rudders. It is evident that the twisted rudder reduces the magnitude of y-direction velocity. Meanwhile the Axial vorticity is also reduced(see Figure 13). Figure 14 and Figure 15 show the y-direction velocity on the horizontal surfaces below and above the shaft centre, respectively. For the twisted rudder，the y-direction velocity magnitude is smaller than the ordinary rudder in front of the region of leading edge. Considering the y-direction velocity has

no contribution to propulsion, the less the y –direction velocity is, the more efficient the propeller is.

Figure 16 shows the x-direction velocity on horizontal surface above the shaft centre. It is obvious that the x-velocity distribution near the twisted rudder are more straight forward than the ordinary rudder. All these velocity components contours indicate that the twisted rudder more efficiently transfers y-direction momentum to x-direction and improves the propulsive performance.

Figure 17 compares the dynamic pressure distribution on the propellers pressure sides for two kinds of rudders. In the case of the twisted rudder, the maximal pressure distribution is more widely found on the pressure side than ordinary rudder which indicates that the propeller produces more thrust force with the twisted rudder at the same RPS.

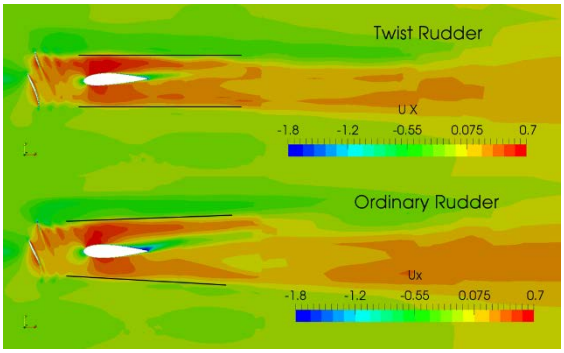


Figure 14 x-direction velocity components contours on horizontal cross-sections above the shaft centre

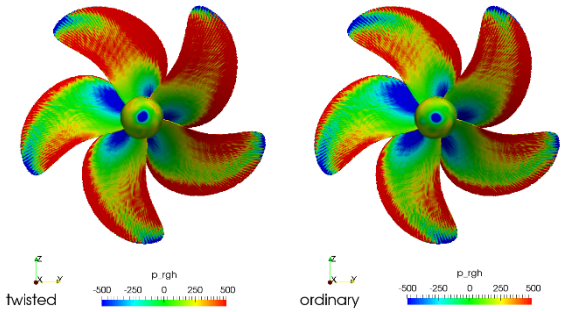


Figure 15 dynamic pressure contours on the pressure sides

Conclusions

In this paper, the energy-saving mechanism of twisted rudder is discussed and studied by numerical methods. All simulations are performed by our solver, naoe-FOAM-SJTU, developed under the framework of the open source OpenFOAM packages. Overset technique is applied to handle with propeller rotation and ship motions. The propeller open-water curves are obtained using the single run methodology, achieving good agreement with experimental data. Computations of the towed bare hull DTC are performed as needed to obtain self-propulsion factors. Resistance coefficients are compared against experimental data showing good results. Two kinds of rudders speed performance is predicted at the self-propulsion point. The results show that twisted rudder boosts hull efficiency by 2.4% through reducing the thrust deduction fraction and raising the hull efficiency. The results of delivered power is decreased by 3.9% in the model scale. The flow fields around the propeller- rudder system show that twisted rudder efficiently transfers y-direction momentum to x-direction and improves the propulsive performance.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information

Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] el Moctar, O., Shigunov, V., and Zorn, T. (2012) Duisburg Test Case: Post-panamax container ship for benchmarking, *Ship Technology Research* **59**(3), 50–64.

[2] Kim, J.-H., Choi, J.-E., Choi, B.-J., and Chung, S.-H. (2014) Twisted rudder for reducing fuel-oil consumption, *International Journal of Naval Architecture and Ocean Engineering* **6**(3), 715–722.

[3] Liu, J., and Hekkenberg, R. (2017) Sixty years of research on ship rudders: effects of design choices on rudder performance, *Ships and Offshore Structures* **12**(4), 495–512.

[4] Sun, Y., Su, Y., and Hu, H. (2015) Experimental Study and Numerical Simulation on Hydrodynamic Performance of a Twisted Rudder, *Marine Technology Society Journal* **49**(5), 58–69.

[5] Shen, Z., Wan, D., and Carrica, P.M. (2015) Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering, *Ocean Engineering* **108**, 287–306.

[6] Wang, J., Zhao, W., and Wan, D. (2016) Self-propulsion Simulation of ONR Tumblehome Using Dynamic Overset Grid Method, *Proceding of 7th International Conference on Computational Methods*, Berkeley, USA, No. ID 1499-5539-1-PB.

[7] Wang, J., Zhao, W., and Wan, D. (2016) Free maneuvering simulation of ONR tumblehome using overset grid method in naoe-FOAM-SJTU [C], *31st Symposium on Naval Hydrodynamics*, Monterey, CA, USA.

[8] Rusche, H. (2003) Computational fluid dynamics of dispersed two-phase flows at high phase fractions, Ph.D. Imperial College London (University of London).

[9] Berberović, E., Hinsberg, N.P. van, Jakirlić, S., Roisman, I.V., and Tropea, C. (2009) Drop impact onto a liquid layer of finite thickness: Dynamics of the cavity evolution, *Physical Review E* **79**(3), 036306.

[10] Xing, T., Carrica, P., and Stern, F. (2008) Computational Towing Tank Procedures for Single Run Curves of Resistance and Propulsion, *Journal of Fluids Engineering* **130**(10), 101102.

[11] Carrica, P.M., Castro, A.M., and Stern, F. (2010) Self-propulsion computations using a speed controller and a discretized propeller with dynamic overset grids, *Journal of Marine Science and Technology* **15**(4), 316–330.

Numerical Study of Riser Vibration Due to Top-End Platform Motions

Bowen Fu[*], Di Deng, Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: fubowen@outlook.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

Abstract

The CFD simulations of vortex-induced vibrations of a flexible riser under a swaying and surging platform have been numerically investigated based on the strip theory. The top end of the flexible riser are forced to oscillate in one or two directions. Three cases are considered, one with only one-direction excitation, one with '∞'-shaped excitation trajectory, and one with parabolic excitation trajectory. When the riser was excited in a parabolic trajectory, the vibrations in both directions are enhanced. However, vibrations can be reduced in the '∞'-shaped trajectory case, in which a 'hat'-shaped trajectory has been observed from a reference frame which moves with the straight riser axis.

Keywords: Vortex-induced vibration; top-end platform motion; flexible riser; computational fluid dynamics; fluid-structure interaction; viv-FOAM-SJTU solver

Introduction

Marine risers can experience vortex-induced vibrations (VIV) when exposed to currents. Furthermore, offshore floating platforms subject to waves, currents or winds may cause risers to reciprocate in the water. The risers are thus exposed to a relatively oscillatory flow with a degree of shear and forced to cross its own wake, rendering the situation more like wake-induced vibrations. The vortex shedding frequencies keep going up and down due to the continuous flow velocity changes. Lock-in or resonance phenomena occur when the vortex shedding frequencies meet one of the risers' natural frequencies.

Vortex-induced vibrations of risers subject to waves or top-end excitations have received more and more attention. Duggal et al.[1] conducted a large-scale experimental study of vibrations of a long flexible cylinder in regular waves. Also some researchers [2–4] conducted experimental and numerical studies on vibrations of a hanging riser subject to regular or irregular top-end excitations. Riveros [5] experimentally and numerically studied a model riser sinusoidally excited at its top end.

Most of previous numerical studies of risers subject top-end excitations mainly concentrate on excitations in one direction. However, the top-end platform actually moves in more than one direction, making it necessary for research on two-directional excitations. In the present work, vibrations of a vertical top-tensioned riser sinusoidally excited at its top end in one or two directions are numerically investigated using a CFD method based on strip theory. The simulations are conducted by the in-house solver viv-FOAM-SJTU, which has been validated in previous studies [6,7]. The present article is organized as follows. Section 2 introduces the concerned problems, followed by the simulation methods in Section 3 for handling the problems in Section 2. And Section 4 gives the simulation results with detailed analyses. Finally, a curt summary is presented in Section 5.

Problem

To simulate the effect of the top-end platform's motions, the top end of the riser is forced to oscillate in one or two directions. The excitation motion of the riser is a periodic function of time, expressed as

$$x_{\mathrm{s}} = A \cdot \sin(2\pi t \cdot T_{\mathrm{w}}^{-1}) , \quad (1)$$

$$u_{\mathrm{s}} = 2\pi A \cdot T_{\mathrm{w}}^{-1} \cdot \cos(2\pi t \cdot T_{\mathrm{w}}^{-1}) , \quad (2)$$

A being the oscillating amplitude and T_{w} the oscillating period, x_{s} the oscillating displacement and u_{s} the oscillating velocity. The maximum excitation reduced velocity $U_{\mathrm{r\,max}}$ can be written as

$$U_{\mathrm{r\,max}} = \frac{u_{\mathrm{s\,max}}}{f_{\mathrm{n1}} D} = \frac{2\pi A}{T_{\mathrm{w}} f_{\mathrm{n1}} D} , \quad (3)$$

where f_{n1} is the first natural frequency of the riser. In the sinusoidal flow, the Keulegan-Carpenter (KC) number can be expressed as

$$KC = u_{\mathrm{s\,max}} T_{\mathrm{w}} \cdot D^{-1} = 2\pi A \cdot D^{-1} , \quad (4)$$

in which $u_{\mathrm{s\,max}}$ is the maximum excitation velocity.

Table 1: Main structural properties of the flexible riser.

	Symbols	Values	Units
Mass ratio	m^*	1.53	−
Diameter	D	0.024	m
Length	L	12	m
Bending stiffness	EI	10.5	$\mathrm{N \cdot m^2}$
Top Tension	T_{w}	500	N
First natural frequency	f_{n1}	1.08	Hz
Second natural frequency	f_{n2}	2.16	Hz
Third natural frequency	f_{n3}	3.25	Hz

Method

In order to compute the vibrations of flexible risers, the hydrodynamic forces acting on them must be obtained. To do this, the transient incompressible Reynolds-averaged Navier–Stokes equations are solved numerically, the SST $k-\omega$ turbulence model is employed to determine the Reynolds stresses. Considering the large scale in the axial direction of the flow domain, two-dimensional flow fields located equidistantly along the span are solved instead of the entire three-dimensional flow field is not quite feasible. In this case, to solve. As Willden and Graham[8] has mentioned, though three-dimensional vortices might be developed when flow past a riser, an effect of lock-in actually maintain the locally two-dimensional property, making it possible for us to compute the fluid dynamics locally in a two-dimensional way. The hydrodynamic forces at other positions along the span can be interpolated accordingly. The PIMPLE algorithm in the OpenFOAM is used to compute the two-dimensional flow fields.

The flexible riser is modeled as a small displacement Bernoulli–Euler bending beam, with two ends set as pinned. In the present work, the top end of the riser oscillates harmonically as prescribed. Thus, the beam's total displacement is referred to as total displacement x_t, i.e. the sum of the support motion x_s, plus the relative displacement x:

$$x_\mathrm{t} = x_\mathrm{s} + x. \tag{8}$$

The equilibrium of forces for this system can be written as

$$f_\mathrm{I} + f_\mathrm{D} + f_\mathrm{S} = f_\mathrm{H}\,, \tag{9}$$

where f_I, f_D, f_S, f_H are the inertial, the damping, the spring, and the hydrodynamic forces, respectively. The force components can be expressed as $f_\mathrm{I} = m\ddot{x}_\mathrm{t}$, $f_\mathrm{D} = c\dot{x}$, $f_\mathrm{S} = kx$, where m, c, k are the mass, the damping and the stiffness of the system. We have

$$m\ddot{x}_\mathrm{t} + c\dot{x} + kx = f_\mathrm{H}\,, \tag{10}$$
$$m\ddot{x} + c\dot{x} + kx = f_\mathrm{H} - m\ddot{x}_\mathrm{s}\,. \tag{11}$$

Thus there will be one additional contribution to the total forces from the point of view of the riser, the additional inertial force. In the finite element method the equations can be discretized as

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{F}_{\mathrm{H}x} - \mathbf{M}\ddot{\mathbf{x}}_\mathbf{s}\,, \tag{12}$$
$$\mathbf{M}\ddot{\mathbf{y}} + \mathbf{C}\dot{\mathbf{y}} + \mathbf{K}\mathbf{y} = \mathbf{F}_{\mathrm{H}y} - \mathbf{M}\ddot{\mathbf{y}}_\mathbf{s}\,, \tag{13}$$

where $\mathbf{x}$, $\mathbf{x_s}$, $\mathbf{y}$, and $\mathbf{y_s}$ are nodal displacement vectors, and $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$ are the mass, the damping and the stiffness matrices. The Rayleigh damping $\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}$ is adopted, where α and β are calculated based on the natural frequencies of two mainly involved modes, with a damping ratio ζ of 0.03. The equation can be written as

$$\begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \frac{2\zeta}{f_{\mathrm{n}i}+f_{\mathrm{n}j}}\begin{bmatrix}2\pi f_{\mathrm{n}i} f_{\mathrm{n}j}\\ 1/(2\pi)\end{bmatrix}. \tag{14}$$

$\mathbf{F}_{\mathrm{H}x}$ and $\mathbf{F}_{\mathrm{H}y}$ are the hydrodynamic force vectors in corresponding directions (including hydrodynamic mass forces). The equations are solved using the Newmark-beta method [9].

At the beginning of each time step the hydrodynamic forces are mapped to the structural model nodes. Then the displacements of the riser are computed. With the displacements obtained, the mesh can be moved or deformed accordingly, thus resulting in new flow fields from which the hydrodynamic forces can be gained. In this way, a time step is advanced. The procedure is shown in Figure 2, based on which the solver viv-FOAM-SJTU is formed. Twenty strips equidistantly located along the span of the riser are plotted in Figure 3. These strips share the same initial flow field mesh, as shown in Figure 4. The motion solver "displacementLaplacian" in OpenFOAM is applied to handle the dynamic mesh [10]. Imposed on the surface of the riser is the no-slip boundary, and no external current is applied. The riser is discretized into 80 elements, with each element imposed of uniformly distributed loads.

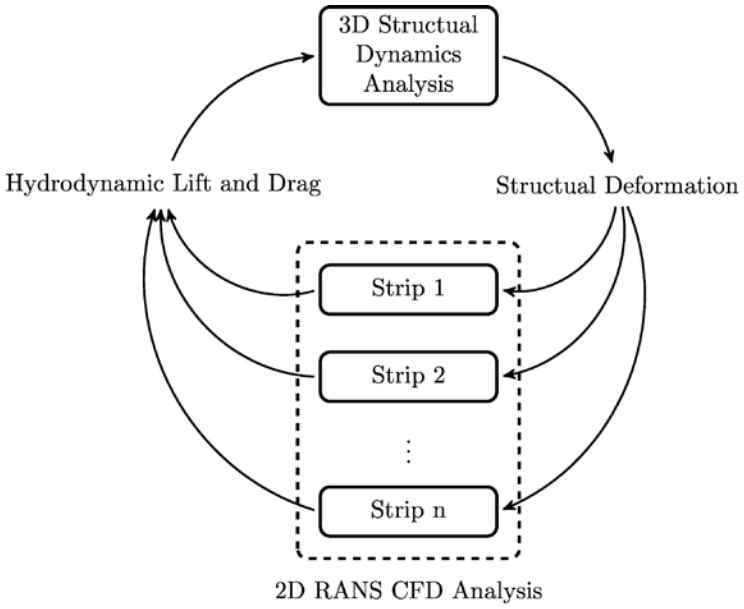


Figure 1: Fluid-structure interaction. The fluid and the structure are coupled by hydrodynamic forces and structural deformations.

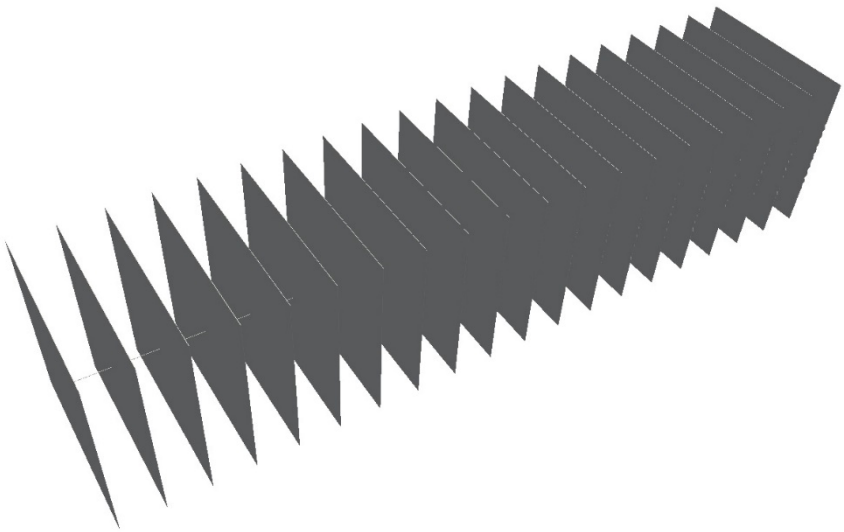


Figure 2: Twenty strips located equidistantly along the span of the flexible riser.

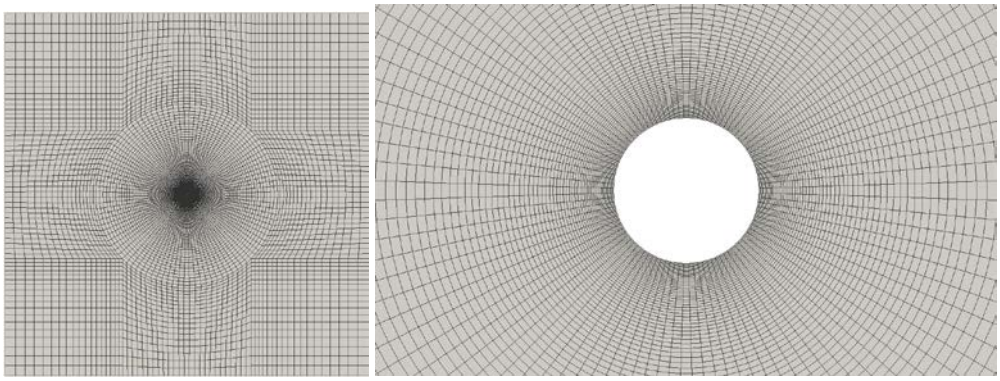


Figure 3: Initial mesh on each strip. The mesh near the riser is magnified. Eighty diameters in the in-line direction x, forty diameters in the cross-flow direction y.

Table 2: Main parameters for simulation cases. Symbols KC_x, KC_y, $U_{\mathrm{r\,max}}$ and $V_{\mathrm{r\,max}}$ denote KC numbers and reduced velocities at the top end of the riser in the in-line and cross-flow directions, ψ being the phase difference between the excitation in two directions.

Case	KC_x	KC_y	$U_{\mathrm{r\,max}}$	$V_{\mathrm{r\,max}}$	ψ
1	84	0	12	0	-
2	84	21	12	6	0
3	84	21	12	6	90

Results

Three cases considered in the present work are listed in Table 2. In Figure 4 are plotted the trajectories of them. When the riser is excited in two directions, the frequency of the cross-flow excitation is set as twice that of the in-line excitation, thus forming '∞'-shaped or parabolic trajectories in Figure 4. An interesting 'hat shaped trajectory (viewed from a reference frame which moves with the straight riser axis) is found in case 2. In case 2, risers move upwards in the cross-flow direction when passing the intersection ('X') parts of the '∞' trajectories. Near the intersections are the high speed periods and consequently large drag forces, causing large deflections in the opposite directions. High speeds also mean more intense vibrations in the locally cross-flow direction, forming the lower 'crab plier' shaped parts. Thus the two 'crab plier'-like parts in Figure 5 correspond to the 'X' parts in Figure 4 while the top-end knots in Figure 5 correspond to two sides in Figure 4, the zero in-line excitation velocity periods. Since at two sides of '∞' the riser is always moving downwards, causing the top-end knots in Figure 5 to move upwards. Though the trajectories in cases 1 and 3 seem similar in Figure 4, i.e. of overlapping forward and backward routes, the curvature actually plays a subtle role in the vibrations, resulting in quite different results in Figs. 5-7. The constantly changing motion directions together with the wake effects render more complex flow situations and also larger deflections in case 3.

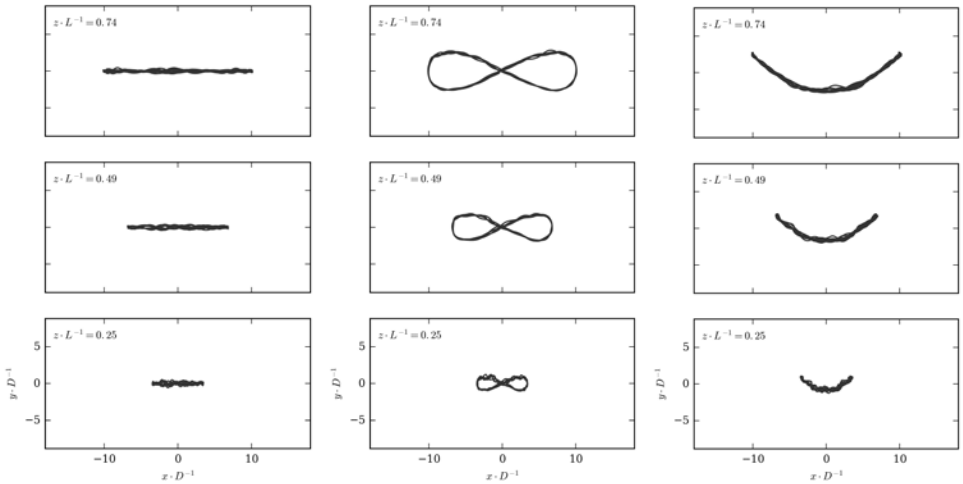


Figure 4: Actual trajectories of the vibrations of the riser in cases 1 (left), 2 (middle), and 3 (right) viewed from a fixed camera.

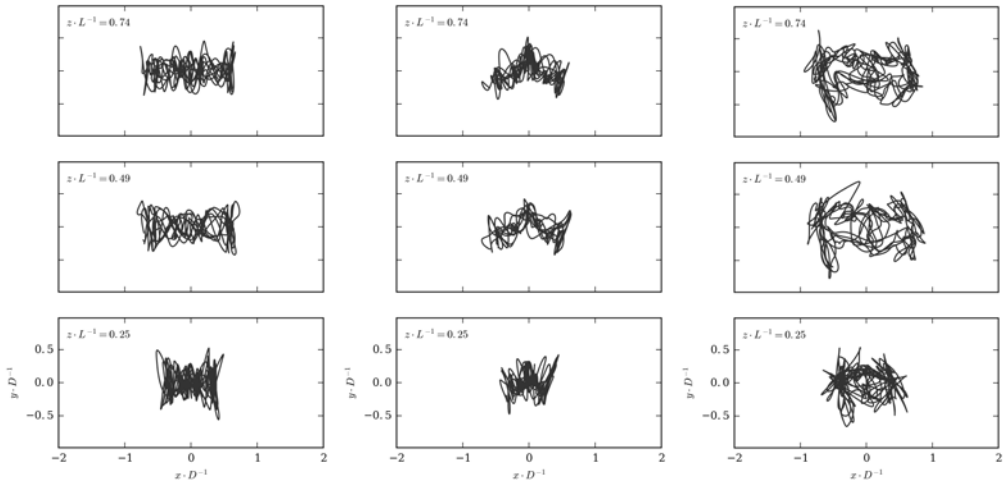


Figure 5: Trajectories of the vibrations of the riser in cases 1 (left), 2 (middle), and 3 (right) viewed from a reference frame which moves with the straight riser axis.

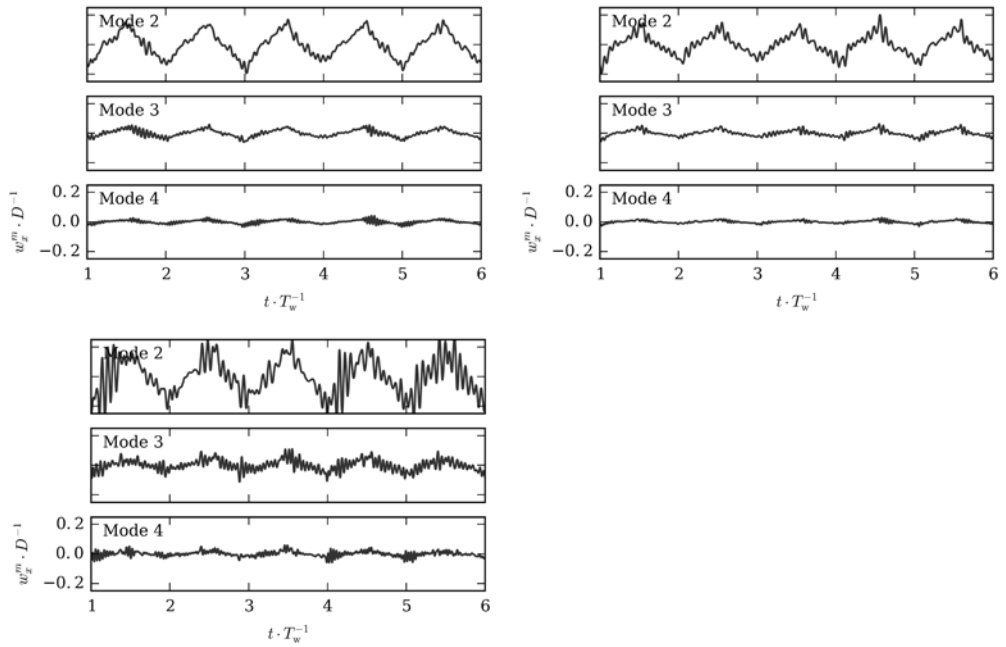


Figure 6: Time series of modal weights of the in-line displacements $x \cdot D^{-1}$ in cases 1 (left), 2 (middle), and 3 (right), $w_x^m \cdot D^{-1}$.

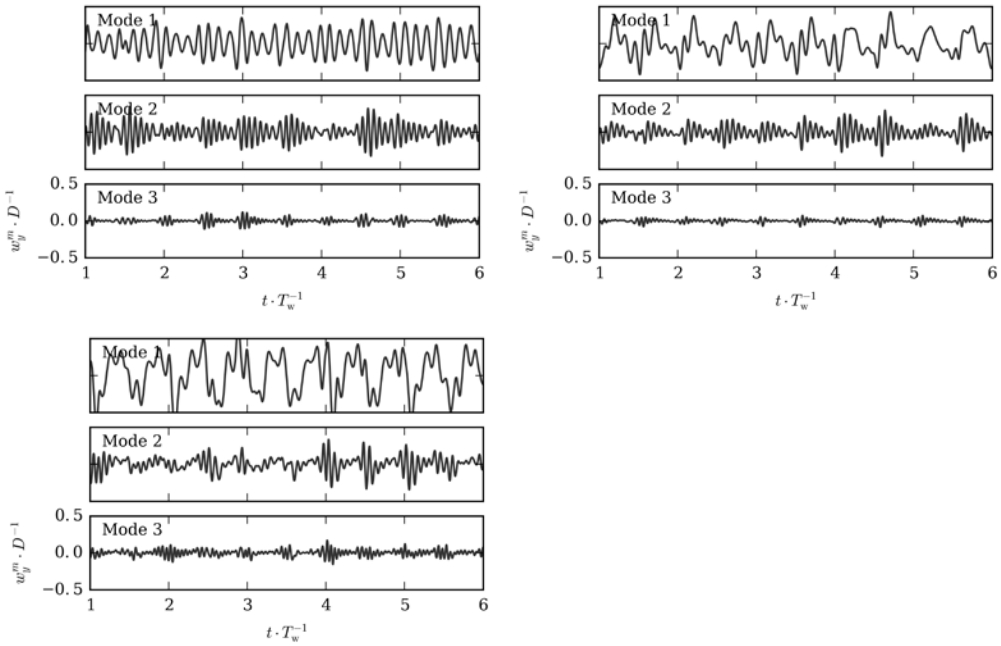


Figure 7: Time series of modal weights of the cross-flow displacements $y \cdot D^{-1}$ in cases 1 (left), 2 (middle), and 3 (right), $w_y^m \cdot D^{-1}$.

Time series of modal weights of the in-line and cross-flow displacements are presented in Figures 6 and 7, respectively. Modal decompositions are conducted in the least-squares sense[11,12]. The low-frequency components due to the current speed variation along the span appear in higher in-line modal weights in all cases. These also occur in the cross-flow direction in case 2 and 3, but the low frequency is twice of the in-line one, the same with the excitation frequency in the corresponding direction. It is clear from Figures 5-7 that the vibrations become more intense in both directions in case 3, especially in the in-line direction, indicating that the excitation in the other direction can enhance the vibrations when the riser has to cross its own wake. The pretty intense second modal weight of in-line displacement in case 2 can be explained as the component of actual 'cross-flow' vibrations in 'x' direction (still referred to as in-line direction). As a result, the second modal weight of cross-flow displacement become smaller compared to that in case 1. For case 2, the riser does not necessarily cross the wake, rendering less intense vibrations compared to case 1, in which the riser is excited in only one direction. Comparing results of case 2 and 3, it is safe to say that the almost same forward and backward trajectory, which ensures that the riser crosses the wake, greatly enhance the vibrations.

Conclusions

The vortex-induced vibrations of a flexible riser excited at the top-end in one and two directions have been numerically simulated. A CFD method based on the strip theory is used in the simulations. Three cases are considered, one with only one-direction excitation, one with '∞'-shaped excitation trajectory, and one with parabolic excitation trajectory. When the riser was excited in a parabolic trajectory, the vibrations in both directions are enhanced. However, vibrations can be reduced in the '∞'-shaped trajectory case, in which a 'hat'-shaped trajectory has been observed（viewed from a reference frame which moves with the straight riser axis）. The key factor is whether the riser would cross its own wake.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1]. Duggal, A.S., Niedzwecki and J.M. (1995) Dynamic response of a single flexible cylinder in waves, *J. Offshore Mech. Arct. Eng.* 117, 99–104.

[2]. Park, H.I., Hong, Y.P., Nakamura, M. and Koterayama, W. (2002) An Experimental Study on Transverse Vibrations of A Highly Flexible Free-Hanging Pipe in Water, in: *Proceedings of The Twelfth (2002) International Offshore and Polar Engineering Conference*. Kitakyushu, Japan, pp. 199–206.

[3]. Park, H.I., Hong, Y.P., Nakamura, M. and Koterayama, W. (2004) Experimental Study on Vortex Induced Vibrations of Highly Flexible Immersed Pipe Subjected to Top End Oscillations, *J. Waterw. PORT, Coast. Ocean Eng.* 130, 207–214.

[4]. Senga, H., Koterayama and W. (2005) An Experimental and Numerical Study on Vortex Induced Vibrations of a Long Flexible Riser Undergoing Irregular Motion at its Top End, *Int. J. Offshore Polar Eng.* 15, 274–281.

[5]. Riveros, C.A., Utsunomiya, T., Maeda, K. and Itoh, K. (2009) Dynamic response of oscillating flexible risers under lock-in events, *Int. J. Offshore Polar Eng.* 19, 23–30.

[6]. Duan, M., Wan, D. and Xue, H. (2016) Prediction of response for vortex-induced vibrations of a flexible riser pipe by using multi-strip method, Rhodes, Greece, pp. 1065–1073.

[7]. Fu, B., Duan, M. and Wan, D. (2016) Effect of mass ratio on the vortex-induced vibrations of a top tensioned riser, in: *The Second Conference of Global Chinese Scholars on Hydrodynamics Effect*, Wuxi, China, pp. 431–435.

[8]. Willden, R.H.J. and Graham, J.M.R. (2004) Multi-modal Vortex-Induced Vibrations of a vertical riser pipe subject to a uniform current profile, *Eur. J. Mech. B/Fluids* 23, 209–218.

[9]. Clough, R.W. and Penzien, J. (2003) Dynamics of Structures, 3rd ed, *Computers & Structures*, Inc., Berkeley.

[10]. Jasak, H. (2009) Dynamic Mesh Handling in OpenFOAM, in: *47th AIAA Aerospace Sciences Meeting Including The New Horizons Forum and Aerospace Exposition*. American Institute of Aeronautics and Astronautics, Inc., Orlando, Florida.

[11]. Lie, H. and Kaasen, K.E. (2006) Modal analysis of measurements from a large-scale VIV model test of a riser in linearly sheared flow, *J. Fluids Struct.* 22, 557–575.

[12]. Trim, A.D., Braaten, H., Lie, H. and Tognarelli, M.A., 2005. Experimental investigation of vortex-induced vibration of long marine risers, *J. Fluids Struct.* 21, 335–361.

The effect of Wave Steepness on Wave Breaking Properties over Submerged Reef

Ke Xia[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: xiakee@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

Wave turning and breaking is a common phenomenon in epicontinental sea which is prominent subject in coastal and marine engineering that deserve investigated in detail. In this paper, a viscous flow solver (naoe-FOAM-SJTU) which is developed and based on the popular open source toolbox OpenFOAM is presented. The solver is adopted to study the effect of wave steepness on wave breaking properties over submerged reef with a 2D simulation. Wave steepness will be considered and the asymmetry properties and cavity properties that induced in the process of wave over-turning will be investigated in detail.

Keywords: submerged reef; wave breaking; wave steepness; asymmetry characteristic

Introduction

Wave breaking as a common phenomenon is an important element in many oceanographic and offshore engineering [1]. It is a two phase flow phenomenon involving air and water, and it strongly influences the air–sea interaction by enhancing mass, momentum and energy transfer between the phases [2]. Due to the nonlinear effect, the wave crests tend to sharpen while the troughs tend to flatten with decreasing of water depth [3]. According to the dispersion relation, the wave velocity and the wave length decrease with the reduction of the water depth in the process of wave propagation shoreward. Hence, the velocity of the front part is smaller than the rear part, which related to the wave energy accumulate upward with the shorter wave length, and the wave height increase which leads to the wave evolution in this process. When the steepness (H/L) is too large and exceed the limitation, the wave breaking occurs.

A relationship reported by Kjeldsen and Myrhaug (1978) [4] firstly that exists between the asymmetric parameters and the breaker type. They introduced steepness and asymmetry parameters to describe the asymmetry of the wave profile: crest front steepness (ε), crest rear steepness (δ), the vertical asymmetry factor (λ) and the horizontal asymmetry factor (μ) as depicted in Figure 1. Parameterization of the skewness and asymmetry of waves is an efficient way to describe the wave evolution process and also meaningful to the application and development of physical model. Hence, these parameters are always chosen to descript the transformation of waves [5-8]. Additionally, the geometric properties of breaking waves can be related to the breaker type, which plays a key role in estimation of breaking wave forces on marine structures [2].

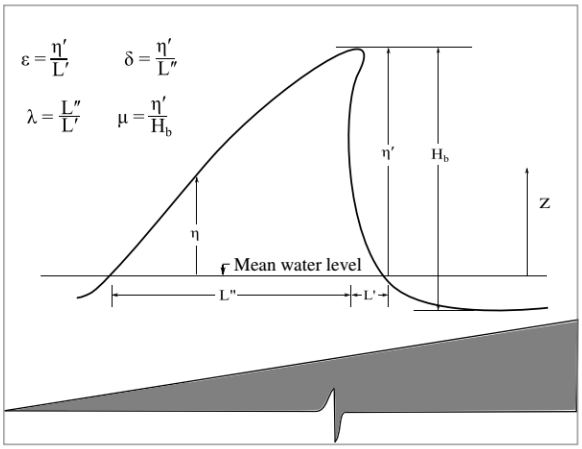


Figure 1. Definition of local wave profile asymmetry (Kjeldsen and Myrhaug, 1978)

Numerous studies have attempted to explain the wave over-turning and breaking process and their characteristics. The detailed literature review on wave breaking in deep and shallow water can be found in Cokelet (1977), Peregrine (1983), Basco (1985), Banner and Peregrine (1993), and Perlin et al. (2013) [9-13]. In the case of submerged terrain and reefs, wave breaking is strongly influenced by the local environmental parameters, such as water depth (d) and sea bed slope (m). This has been studied in laboratory experiments by Blenkinsopp and Chaplin (2008) [14]. In order to research the nonlinear phenomenon, several surface wave theories have been investigated and proposed to resolve the wave breaking issues. Meanwhile, many theories were put forward to describe the wave breaking, and most studies in the field of submerged breakwater structures have only focused on the prediction of the reflection and transmission characteristics of waves for a given environmental condition. Ting and Kim (1994) [15] investigated the wave transformation over a submerged structure and concluded that potential theory cannot be applied to model the flow process such as flow separation and energy dissipation. However, the breaking process and generation and dissipation of vortices are created by rotational flow [16]. Numerical modeling of wave breaking becomes challenging due to the intricacy in describing the physical processes involved such air-sea interaction, vorticity generation, overturning motion and the air entrainment. Hence, a straightforward approach to describing the breaking process numerically is applied to solve the fundamental fluid dynamic equations with CFD (Computational Fluid Dynamics) method. Chella et al. (2015) [2] did the investigation about the characteristics and profile asymmetry properties of wave breaking over an impermeable submerged reef by the CFD method, and the capture of free surface is conducted by the level set method. The numerical result showed great correlation to the experimental results which is just the contribution of consideration of vortex and viscosity.

Despite a considerable number of experimental, numerical and theoretical studies and field observations have been carried out to investigate the process, the wave breaking mechanism is not completely understood. A comprehensive examination of breaking wave properties is inevitable to understand the mechanism of wave breaking and thus the description of the breaking process. Wave breaking over a submerged reef primarily depends on the tidal level and the characteristics of the incident waves. Moreover, an accurate description of waves breaking over submerged structures has always been a central issue in estimation of hydrodynamic loads on marine structures. So that, in this paper, a viscous flow solver (naoe-FOAM-SJTU) which is developed and based on the popular open source toolbox OpenFOAM is presented. The solver is adopted to study the effect of wave steepness on wave breaking properties over submerged reef in 2D simulation. The asymmetry properties and cavity

properties that induced in the process of wave over-turning will be investigated in detail, and the numerical results will be compared with the experimental data and other numerical data so that to validate the accuracy of the solver and simulation. Some regularity about the wave characteristic of the over-turning wave were found in this work.

Numerical Methods

The present solver naoe-FOAM-SJTU [17] adopted for numerical simulation is based on a built-in solver in OpenFOAM named interDyFoam, which can be used to solve two-phase flow which is incompressible, isothermal and immiscible. To deal with common air-sea interaction and wave evolution problems in coastal and offshore engineering, wave generating module was adopted in this work. SST K-ω model is carried out in all the calculations. Mathematical formulae related to the solver are described as follows in detail.

Governing Equations

For transient, incompressible and viscous fluid, flow problems are governed by Navier-Stokes equations:

$$\nabla \cdot U = 0 \tag{1}$$

$$\frac{\partial \rho \mathbf{U}}{\partial t} + \nabla(\rho(\mathbf{U} - \mathbf{U}_g)\mathbf{U}) = -\nabla p_d - \mathbf{g} \cdot \mathbf{x} \nabla \rho + \nabla(\mu \nabla \mathbf{U}) + \mathbf{f}_\sigma \tag{2}$$

Where $\mathbf{U}$ and $\mathbf{U}_g$ represent velocity of flow field and grid nodes separately; $p_d = p - \rho \mathbf{g} \cdot \mathbf{x}$ is dynamic pressure of flow field by subtracting the hydrostatic part from total pressure p; $\mathbf{g}, \rho$ and μ denote the gravity acceleration vector, density and dynamic viscosity of fluid respectively; $\mathbf{f}_\sigma$ is surface tension which only takes effect at the free surface and equals zero elsewhere. The Laminar model means that the Navier-Stokes equation will be solved directly and the turbulence model is not been considered in the calculation.

Wave Generation and Relaxation Zone

Wave generation is a vital part for the investigation of floating offshore structures and wave evolution. The wave generation and wave damping work are implemented by an open-source toolbox for CFD library: waves2foam. The wave was generated by modification of the velocity boundary condition and the phase boundary condition. In this study, Stokes 2nd wave theory was adopted in the g0eneration of the wave according to the calculated wave cases. The equation of Stokes 2nd wave theory was below:

$$\eta = \eta_1 + \eta_2 \tag{3}$$

$$\eta_1 = \frac{H}{2}\cos \omega t \tag{4}$$

$$\eta_2 = \frac{\pi H^2}{8L}\frac{\cosh kd}{\sinh^3 kd}(2 + \cosh 2kd)\cos 2\omega t \tag{5}$$

In which, η is the wave elevation of free surface in certain point, and H is wave height of the generated wave, k is the wave number and d is the water depth at the local position.

In this wave maker module, relaxation zones are implemented to absorb the incident wave that keeps mass conservation and avoids reflection of waves from outlet boundaries at the same time and what else to avoid waves reflected internally in the computational domain to interfere with the floating structure and wave maker boundaries. The former obviously contaminates the results, and the latter is found to create discontinuities in the surface elevation at the wave making boundary, which leads to divergent solutions [18]. A relaxation function:

$$\alpha_R(\chi_R) = 1 - \frac{\exp(\chi_R^{3.5}) - 1}{\exp(1) - 1} \quad \text{for} \quad \chi_R \in [0, 1] \tag{6}$$

is applied inside the relaxation zone in the following way

$$\phi = \alpha_R \phi_{\text{computed}} + (1 - \alpha_R)\phi_{\text{target}} \tag{7}$$

in which ϕ is either velocity of phase indices. The definition of χ_R is such that α_R is always 1 at the interface between the non-relaxed part of the computational domain and the relaxation zone, as illustrated in Figure 2.

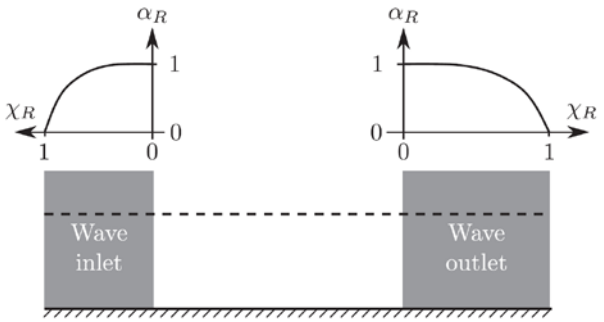


Figure2. A sketch of the variation of χ_R for both inlet and outlet relaxation zones (Jacobsen et al., 2012).

Numerical Examples

The numerical wave tank consists of a submerged reef with a height of 0.618m and a slope of about 1/10 gradient, located 3.8 m from the wave generation zone as shown in Figure 3. The numercial set-up, incident wave parameters and the coordinate system are the same as the experimental conditions that presented in Blenkinsopp and Chaplin (2008) [14], the wave cases that presented in this work is listed in the Table 1. And in the computational domain, several wave gauges were set to measure the wave height at different location.

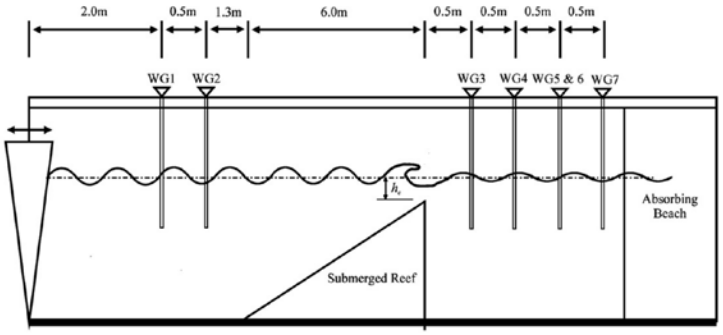


Figure3. Sketch of the computational domain (Blenkinsopp and Chaplin, 2008) [14]

The investigation was set up in a 2D numerical basin. The numerical domain of this work is shown in Figure4 which includes the computational mesh and arrangement of the domain. The total mesh of this work is about 0.82 million, and the mean grid size around the free surface is about 0.5 mm.

Table 1. List of computational cases

Simulation cases	Simulation No.	Wave steepness, H_0/L_0	Reference water depth, d (m)
Based on wave steepness (H_0/L_0)	1	0.02	0.7
	2	0.03	
	3	0.04	
	4	0.05	
	5	0.06	
	6	0.07	

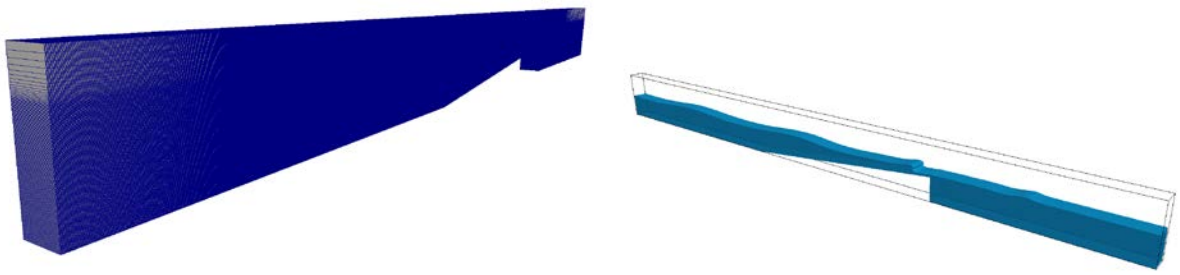


Figure4. Sketch of the mesh and arrangement of the computational domain

The breaking point is assessed in the present study as the point where part of the wave front becomes vertical. Thus, the computed water depth (d_b) and wave height (H_b) at the breaking point are used to calculate the breaker indices. The breaker depth index, γ_b, is the ratio of the breaker height H_b to the water depth at breaking d_b:

$$\gamma_b = \frac{H_b}{d_b} \tag{8}$$

The breaker height index, Ω_b is the ratio of the breaker height H_b to offshore wave height H_0:

$$\Omega_b = \frac{H_b}{H_0} \tag{9}$$

It is well known that the wave profile becomes asymmetric as it approaches the breaking point and cannot be described by the wave steepness (H/L). Hence, additional parameters are required to describe the asymmetric shape of the wave at breaking. Four additional geometric parameters proposed by Kjeldsen and Myrhaug (1978) are used in this study to describe the asymmetry of the wave profile as shown in Figure 1. And these parameter are described in the following:

$$\varepsilon = \frac{\eta'}{L'}, \delta = \frac{\eta'}{L''}, \lambda = \frac{L''}{L'}, \mu = \frac{\eta'}{H_b} \tag{10}$$

In order to examine the effect of the wave steepness and relative reef submergence it was necessary to define some measurable parameters that could be used to quantify the observed variation in wave breaking intensity. As seen in Figure 5. lc and wc are the length and width respectively of the cavity of air enclosed between the overturning jet and the wave face and the x and y axes suitably aligned with the arrows in this figure. From photographs of breaking waves at 23 international surfing breaks, Mead and Black (2001) [19] observed that those on steep reefs broke in a more violent plunging manner, and that there was an inverse linear

relationship between the length to width ratio of the cavity lc/wc and the orthogonal reef gradient (for present purposes the same as the actual seabed gradient m), independent of the incoming wave parameters:

$$\frac{l_c}{w_c} = \frac{0.065}{m} + 0.821 \tag{11}$$

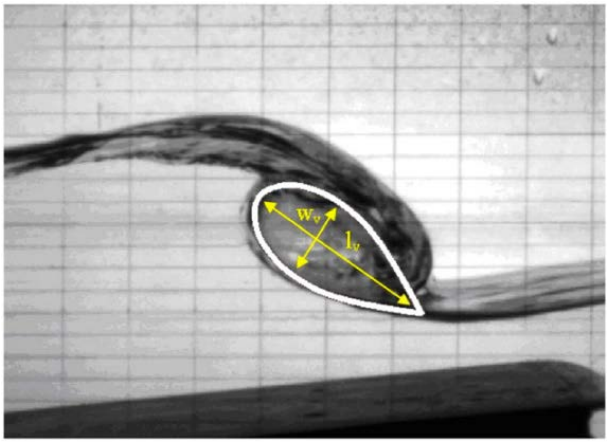


Figure 5. sketch of the l_c and w_c in the cavity (Blenkinsopp and Chaplin, 2008)

Results and Discussions

In the section, wave breaking phenomenon, the asymmetry characteristic and the feature of the cavity in the process of wave over-turning of the wave over submerged reefs will be investigated in detail.

Wave breaking phenomenon over reef

In the practice, wave will be nonlinear in the process that wave propagate from deep sea to the shallow that the asymmetry of the wave will be more clear and when the wave steepness is too large to maintain the waveform of the water particles, the wave breaking will occurs. As shown in Figure 6, which shows the changes in the wave surface profile and the velocity during the breaking process clearly depict that a portion of the wave crest with higher velocity moves forward faster than the rest of the wave. Initially the wave spread to the very shallow position and it is evident that the waveform is very asymmetry which is caused by the accumulation of the water particles from the rear side to the front side and results in the wave height increase at the same time. When the wave crest spread to the breaking position, the wave become to turn, and the velocity of the wave crest become increase and towards the right side which is greater than the wave spread velocity. When the overturned and ejected wave front hits the free surface at the base of the wave, it almost falls over the wave trough of the preceding wave and generates a surface roller. Then water splashes up causing a rise in the water surface with an air cavity inside the water as shown in the Figure 6. An extreme changeover from irrotational flow to rotational flow leads to increased vorticity and turbulence as the wave approaches the beach and eventually violent mixing of air and water occurs. It is worth to notice that the impingement of the rotating plunging vortex is causing a secondary wave with new wave characteristics that propagates shoreward as shown in Figure 6. Moreover, the numerical prediction of the flow pattern and the wave profile changes are very similar to the observation of the flow features of plunging breakers over plane slopes by Basco (1985) [11]. And the return flow phenomenon can be captured clearly in the first figure in the right side of the wave peak which is induced by the shallow water depth and the wave overturning phenomenon.

Figure 6. Wave breaking phenomenon over the reef in half wave period colored by velocity.

Study of wave asymmetry parameter

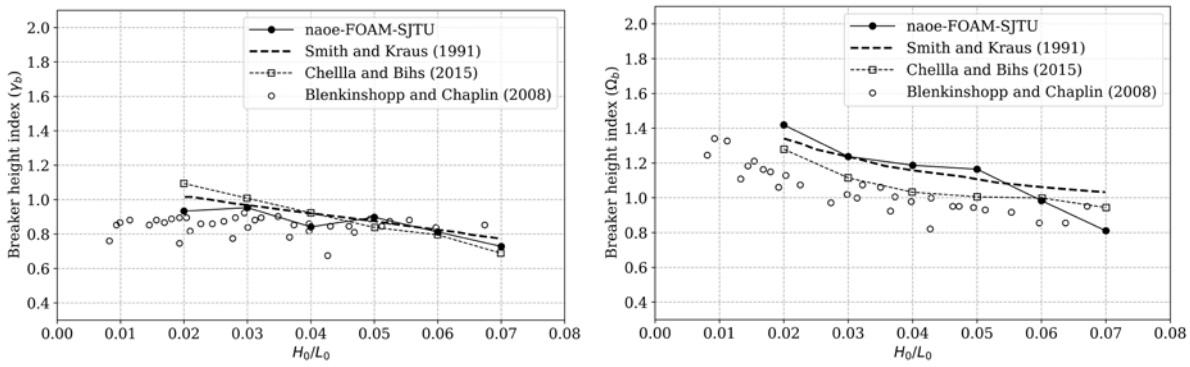


Figure 7. Breaker depth index (left) and breaker height index (right) as a function of offshore wave steepness (H_0/L_0).

Figure 7 presents the comparison of the numerical results and the measured breaker depth index and the breaker height index. It is evident that the γ_b and Ω_b decreases with the increasing of wave steepness. Although the experimental data by Blenkinsopp and Chaplin (2008) [14] do not vary much versus H_0/L_0. However the computed results are in good agreement with the experimental data. At the same time the present numerical results agree the equation solution that given by Smith and Kraus very well. And this work validates the accuracy of the simulation work in the wave overturning calculation and the solver and computational mesh adopted in this work is reliable. Moreover, it is evident that the breaker height and depth index will decrease with the increasing of the wave steepness.

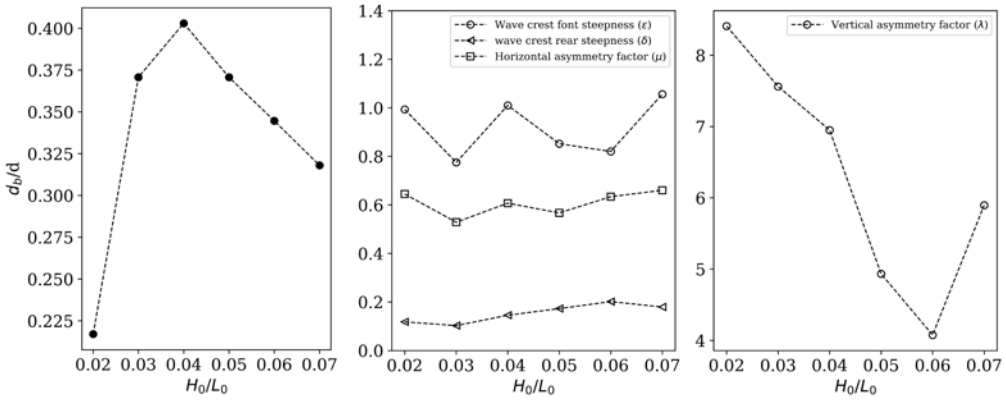


Figure 8. Simulated non-dimensional water depth at breaking (left), wave profile asymmetry parameters as functions of offshore wave steepness (H_0/L_0) (mid and right)

After the validation work, the analysis of the wave asymmetry will be carried out. As shown in the Figure 8, it is evident that, all of these parameters will change with the increasing of the wave steepness. And to analyze the water dimensionless water depth, it is evident that the relative water depth will decrease with the increasing of the wave steepness except the case 1 whose wave steepness is 0.02, and in the author's opinion, it may be caused by the error in the calculation which is apparently conflicting the other cases and can be removed in the analysis of the trends. In the mid figure, it is evident that all these three parameter do not change very clearly in the change trends, and to do a deep analysis it can be found that the variation tendency of wave crest front steepness, wave crest rear steepness and the horizontal asymmetry factor increase with the increasing of the wave steepness. In the right figure, it is

evident that the wave vertical asymmetry factor will decrease with the increasing of the wave steepness.

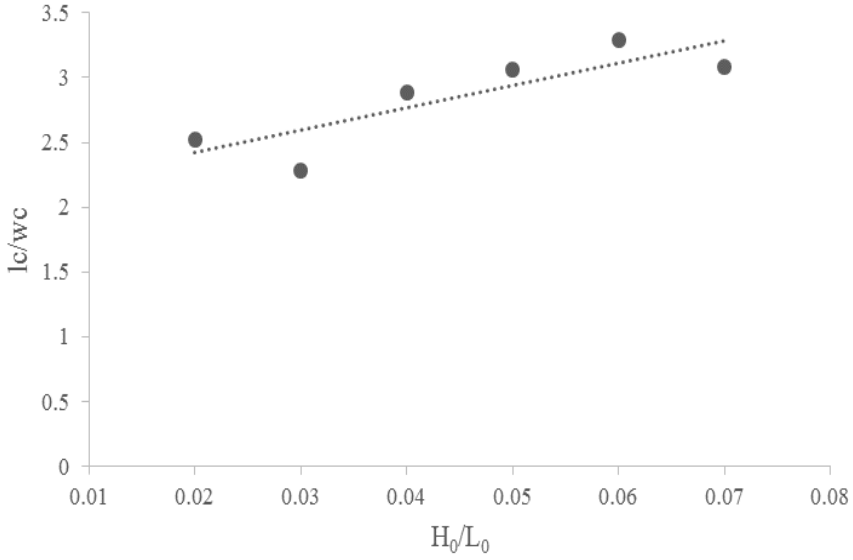


Figure 9. Simulated cavity length to width as a function of wave steepness

In the final section, the characteristic of the cavity in the simulation of the wave overturning process will be analyzed. The parameter of the cavity is related to the wave overturning intensity which is also a significant parameter in the research and quantification of the wave breaking. And as shown in Figure 9, the black point is the calculated results and the dash line is the trend line of the results. From the trend it is obvious that the results of lc/wc is greater with the increasing of the wave steepness and so that it means with the increase of the wave steepness, the cavity in the wave overturning process is more and more narrow and long which means that the wave breaking intensity is greater that the velocity of the wave crest if larger. Finally, the vertical asymmetry decrease with the increase of steepness. For a further analysis of the wave breaking intensity, the characteristic of cavity that induced by the wave overturning was investigated and it is found that the cavity will be more narrow and long with the increasing of the wave steepness which shows that the wave breaking intensity is greater.

Conclusions

In this paper, the wave breaking and overturning phenomenon that affected by the wave steepness over a submerged reef was mainly investigated, and a 2D numerical wave basin was set to do the simulation. All these work was carried out by our in-house two-phase flow CFD solver naoe-FOAM-SJTU. Stokes 2nd wave theory was primarily used in the wave generation work. In these simulations, wave overturning phenomenon can be clearly analyzed and some phenomenon in detail such as cavity, return flow, breaking induced vortex can be captured. Moreover, some regularity can be drawn from the analysis of wave asymmetry characteristic. First of all, the numerical results of the breaker depth and height index show great correlation to the experimental data meanwhile, it is apparent that both these two index decrease with the increasing of the wave steepness. Subsequently, the relative breaking depth decrease with the increasing of the steepness, and it can be found that the variation tendency of wave crest front steepness, wave crest rear steepness and the horizontal asymmetry factor increase with the increasing of the wave steepness.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] K. She, C.A. Greated, W.J. Easson (1994), Experimental study of three-dimensional wave breaking, J. Waterw. Port Coastal Ocean Eng. 120 (1) 20–36.
[2] Chella, M.A., Bihs, H. and Myrhaug, D. (2015). Characteristics and profile asymmetry properties of waves breaking over an impermeable submerged reef. *Coast Eng*, 100 26–36.
[3] M.L. Banner, D.H. Peregrine (1993), Wave breaking in deep water, Annu. Rev. Fluid Mech. 25 (1) 373–397.
[4] Kjeldsen, S.P., Myrhaug, D., 1978. Kinematics and dynamics of breaking waves. Technical Report. River and Harbour Laboratory (NHL), The Norwegian Institute of Technology.
[5] Crawford, A.M. (2000), Field observations of linear transition ripple migration and wave orbital velocity skewnessPhD Thesis Memorial University of Newfound-land,Newfoundland, Canada.
[6] Crawford, A.M., Hay, A.E. (2001), Linear transition ripple migration and wave orbital velocity skewness: observations. J. Geophys. Res. 106 (C7), 14113–14128.
[7] Doering, J.C., Bowen, A.J. (1995), Parameterization of orbital velocity asymmetries of shoaling and breaking waves using bispectral analysis. Coast. Eng. 26, 15–33.
[8] Peng, Z., Zou, Q.P., Reeve, D.E. (2009). Parameterization and transformation of wave asymmetries over a low-crested breakwater. Coast. Eng. 56, 1123–1132.
[9] Cokelet, E (1977). Breaking waves. Nature 267, 769–774.
[10] Peregrine, DH (1983). Breaking waves on beaches. Annu. Rev Fluid Mech, 149–178.
[11] Basco, DR (1985). A qualitative description of wave breaking. J Waterw Port Coast. Ocean Eng, 3, 171–188.
[12] Banner, M, Peregrine, DH (1993). Wave breaking in deep water. Annu Rev Fluid Mech, 373–397.
[13] Perlin, M, Choi, W, Tian, Z (2013). Breaking waves in deep and intermediate waters. Annu Rev Fluid Mech, 115–145.
[14] Blenkinsopp, C, Chaplin, J (2008). The effect of relative crest submergence on wave breaking over submerged slopes. Coast Eng, 55, 967–974.
[15] Ting, FCK, Kim, YK (1994). Vortex generation in water waves propagating over a submerged obstacle. Coast Eng, 24, 23–49.
[16] Takikawa, K, Yamada, F and Matsumoto, K (1997). Internal characteristics and numerical analysis of plunging breaker on a slope. Coast Eng, 31, 143–161.
[17] Shen, Z, Zhao, W, Wang, J and Wan, D. (2014). Manual of CFD solver for ship and ocean engineering flows: naoe-FOAM-SJTU. Technol Rep Solver Man, Shanghai Jiao Tong University.
[18] Jacobsen, NG, Fuhrman DR and Fredsφe (2012). A wave generation toolbox for the open-source CFD library: OpenFoam. Int J Numer Method Fluid. 70:1073–1088.
[19] Mead, S.T., Black, K. (2001), Predicting the breaking intensity of surfing waves. Journal of Coastal Research SI29, 103–130.

CFD Simulation of Flow around a Fixed Paired-Column Semi-Submersible

Weiwen Zhao[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: weiwen.zhao@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

Abstract

The Paired-Column Semi-Submersible (PC Semi) is a concept design by Houston Offshore Engineering (HOE). It distinguished from conventional semi-submersibles from three aspects: eight columns rather than four, rectangular columns rather than square and larger column slenderness. The current study numerically investigate the flow around a fixed PC-Semi at different velocities and current headings. The finite volume CFD solver naoe-FOAM-SJTU is utilized to archive all simulations. Turbulence flow around the semi-submersible is modeled by shear stress transport based delayed detached-eddy simulation (SST-DDES). The present computed drag forces are compared with existing experimental and numerical results. Instantaneous flow visualizations are presented and analyzed. This preliminary study show comprehensive wake interactions between columns and provide a better understanding of Vortex-Induced Motions (VIM) mechanism for multi-column offshore structures.

Keywords: multi-column; wake interference; current heading; naoe-FOAM-SJTU solver

Introduction

Flow past bluff bodies, such as chimneys, bridges, marine cables, risers and offshore platforms may induce flow separation and vortex shedding within a certain range of current speed. The periodic vortex shedding will cause oscillating hydrodynamic forces on bodies in transverse direction. For elastically mounted structures, transverse motions occur under the excitation of these fluctuation forces. The phenomena are commonly observed and gain much attention of engineers and scientists in the field of offshore engineering. It is termed vortex-induced vibrations (VIV) for marine cables and risers and vortex-induced motions (VIM) for large-volume offshore platforms. Despite a considerable number of experimental, numerical studies have been carried out to investigate the process on various kinds of offshore platforms (see for example [1–5]), it remain crucial to understand the flow mechanism behind VIM.

Model tests is the most common research method for VIM. Due to the high cost as well as speed limit of current generation in deepwater offshore basin, VIM model tests are usually conducted in towing tank or water circular channel. The model scale ratio from prototype is limited by the dimensions of towing tank and facilities. According to Fujarra et al. [6], small-scale tests with floating units subjected to VIM are generally employed due to the restriction of available facilities. The scale ratios for offshore platform VIM test vary from 40 to 100. Under such circumstances, it is impossible to ensure Reynolds number ($Re = U_C D / \nu$, where U_C is current velocity, D is characteristic length and ν is kinematic viscosity of fluid) equality between model and prototype. Typical Reynolds numbers for VIM model tests are in the order of 10^3 to 10^5, which fall into the sub-critical range. It is crucial to understand the flow characteristics around a fixed platform prior to an elastically mounted one in the sub-critical range.

Flow around single cylinder have been studied extensively in the past years. Delaney and Sorenson [7] experimentally investigated rounded-corner effect on the drag of an infinite length square cylinder with a wide range of Reynolds number ($Re=10^4$~2×10^6). However, the cylinder in their experiment was infinite. For cylinder with free end, Kawamura et al. [8]

measured surface pressure and Strouhal number around a finite circular cylinder on a flat plane at Re=3.2×10^4. The aspect ratio in their experiments varies from 1 to 8. Okamoto and Uemura [9] experimentally investigated the round-corner effects on aerodynamic forces and turbulent wake of a square column with free-end.

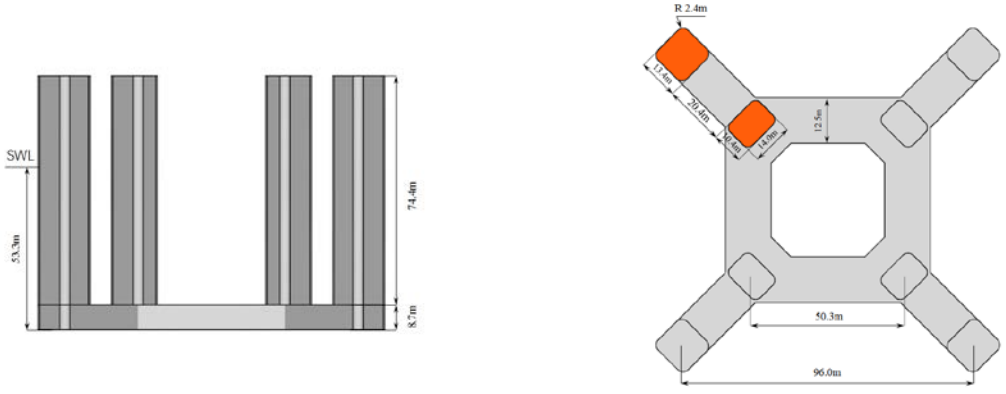
For flow past cylinder array, the wake interference between cylinders is important and has immediate significance in engineering applications. Sumner et al. [10] identified different flow patterns for flow past two cylinders in staggered arrangement with different center-to-center pitch ratios and angles of incidence by conducting experiments at sub-critical Reynolds number (Re=850 to 1900). Sayers [11] conducted experiments of flow past four equispaced cylinders at sub-critical Reynolds number (Re=3×10^4) with a varied staggered angle over the range of 0° to 180°. Liang et al. [12] experimentally and numerically studied flow around four rectangular columns with free end. The Reynolds number in their studies varies from 2.6×10^4 to 4.3×10^4.

In present study we numerically investigate the characteristics of flow around a fixed Paired-Column Semi-Submersible at different velocities and current headings. The detached-eddy simulation is employed for turbulence modeling. This paper is organized as follows: the geometry specifics of the semi-submersible are given first, followed by the introduction of computational domain, boundary conditions and mesh. Furthermore, the solver used in this study is briefly introduced. Then comes the computational results and discussions. Finally, some conclusions are drawn.

Methodologies

Geometry

This study is based on a Paired-Column Semi-Submersible (PC Semi) offshore platform, proposed and designed by the Houston Offshore Engineering (HOE), as shown in Figure 1. PC Semi is designed as an alternative to Spar platform for dry-tree application in the Gulf of Mexico (GoM). It can provide larger payload capability than Spar, while maintain the low dynamic response comparing with conventional semi-submersible. The overall height is 83.1m, in which the column height is 74.4 and pontoon height is 8.7m. The width of pontoon is 12.5m. The designed draft of the platform is 53.3m. PC Semi has eight rectangular columns rather than four squared columns compared with conventional semi-submersibles. The eight columns are divided into four outer columns (OC) and four inner columns (IC) with different dimensions. The OCs are connected to ICs with pontoons at four corner. The dimensions of OC and IC are 14.0m × 13.4m and 14.0m × 10.4m, respectively. The round corner radius for OC and IC are both 2.4m. The center-to-center distances of OC and IC are 96.0m and 50.3m, respectively. The aspect ratios (ratio between immersed length of the column and characteristic length) for OC and IC are 2.75 and 3.06, respectively.



(a) Front view (b) Top view

Figure 1 Geometry of the PC Semi

Prior works to investigate VIM characteristics of the PC Semi have been presented as parts of the Research Partnership to Secure Energy for America (RPSEA) 4405 and 5404 projects. Meanwhile, a large number of experimental and numerical data have been published [13–18]. In these publications, the characteristic length (or effective diameter) of the rectangular column was defined as the diagonal length of the cross section without considering the corner radius. The characteristic lengths for OC and IC are 19.4m and 17.4m, respectively. In the present study, the model is scaled at ratio 1:54 which is the same with that in Antony's work [15].

Computational domain and boundary conditions

The model is displaced in a computational domain consisting of polyhedral cells. The computational domain is illustrated in Figure 2. This is a prior work of investigation for PC Semi VIM, which requires grids moving and deforming. Therefore the overset grids technique is utilized here. Two mesh blocks, namely the background grid and hull grid, are generated individually and then assembled into a single mesh. The domain of the background grid extends to $-14D \le x \le 28D$, $-11D \le y \le 11D$ and $-3H \le z \le 0$, where D is OC's characteristic length and H is the draft of PC Semi.

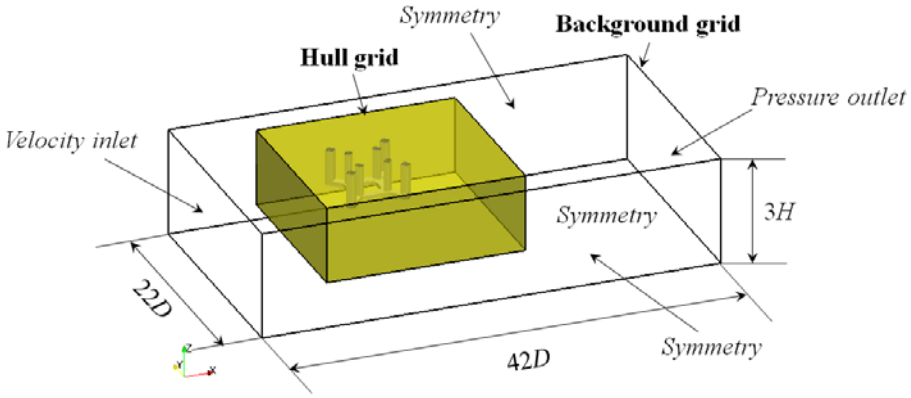


Figure 2 Computational domain and boundary condition

For background grid, Neumann boundary condition for velocity (fixed inlet) and Dirichlet boundary condition for pressure (zero gradient) were used on the upstream inlet patch ($x = -14D$), and vice versa for downstream outlet patch. Symmetry was applied for two sides and bottom of the domain. The free surface effect is neglected due to the small Froude number, thus the top plane at free surface is treated as symmetry. The boundary condition on hull surface is set to no-slip, i.e., zero for velocity and zero normal gradient for pressure.

Meshing strategy

Although the PC Semi in the current study is fixed, we use overset grid technique to perform our simulations as the present work will extends to VIM investigation which requires dynamic mesh. The overset grid approach is proved to be efficient and robust in the current solver [19]. As for stationary problems without grid moving, static overset mesh is applied. The domain connectivity information (DCI) just needed to be calculated only once at the beginning of the simulation. As mentioned before, the computational mesh consists of two mesh blocks, the background and hull grid which are generated individually. Figure 3 illustrates the static overset grid system in the current study. The background grid is

hexahedral and has a uniform grid spacing. The hull grid is based on predominantly Cartesian cut cell approach and refined near hull and wake regions in order to capture the boundary layers and wake structures induced flow separations. Four different levels of refinement zones are utilized to archive higher accuracy in critical regions. In the vicinity of columns and pontoons, four prism cell layers are applied to hull boundary to capture the boundary layer development. For all cases, the non-dimensioned wall distance of the first layer satisfy $y+<5$ which make sure the first layer cells are located in the viscous sublayer.

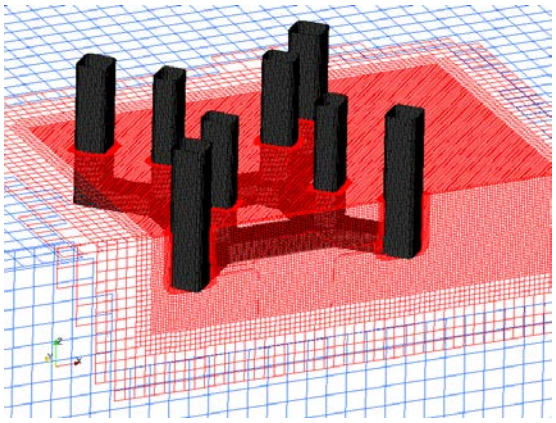


Figure 3 The overset grid distribution for 0° current heading (medium mesh)

Turbulence modeling

Flow past bluff bodies involves unsteady behavior and is dominated by large-scale structures. Therefore, it is not readily to solve these kinds of flow by statistical turbulence models. Essentially, large-eddy simulation (LES) is more suitable as it resolve the large-scale part of the turbulent eddies which has significant impact on the oscillating hydrodynamic forces of bodies. However, LES requires huge computational cost, most of which will be used to resolve the thin boundary layer when dealing with high Reynolds wall-bounded flows. According to Spalart [20], a pure LES simulation for practical engineering flow problem should be possible in approximately 2045. Detached-eddy simulation (DES) was proposed by Spalart [21] to address the challenge of massively separated flow at high Reynolds numbers. It combines the best practice of Reynolds-Averaged Navier-Stokes (RANS) and LES methods by employing unsteady RANS modeling in the near wall region and LES-like manner in the separated flow regions away from wall. In such a way, DES reduces grid resolution at boundary layer while maintaining the ability for accurately predicting eddy structures after flow separation. In the current study, naoe-FOAM-SJTU which is a solver developed based on OpenFOAM toolbox is utilized to perform all the simulations. We choose delayed DES (DDES) approach based on the two-equation Shear Stress Transport SST model for turbulence modeling [22].

Results and discussions

Two different current headings (0° and 22.5 °) are considered in our study. The definition of current heading herein are consistent with the model tests performed by Antony et al. [15], as shown in Figure 4.

Figure 4 Schematic of different current headings

Two different current velocities (U_c=2.0m/s, 2.75m/s) are investigated for each current heading. These velocities are for prototype and not scaled. For the 1:54 scaled model, the corresponding velocities are scaled by $1/\sqrt{54}$ and become 0.272m/s and 0.374m/s, respectively. The Reynolds number defined by OC's characteristic length ranges from 0.86 × 10^5 and 1.1 × 10^5.

The temporal derivatives are discretized using a second-order implicit backward differencing scheme. The convection term in momentum equation is discretized using a second-order upwind scheme, stabilized for transport (linear-upwind stabilized transport, LUST). For turbulent quantities, convection terms are discretized using a second-order TVD limited linear scheme. The merged PISO-SIMPLE (PIMPLE) algorithm is used for pressure-velocity decoupling.

Grid convergence study

Prior to all simulations, the accuracy and reliability of the current CFD approach is assessed by performing grid convergence study. The 0° and 0.272m/s case is selected to perform this study. Three different mesh sizes are considered. All simulations employs a time step of 0.02s. For overset grid system, grid refinement is performed for all mesh blocks. In this case, the background mesh block and the initial hexahedral mesh used generate hull mesh block by cut cell approach are refined by a factor of $\sqrt{2}$. Table 1 shows the results of grid convergence test. The mean drag parameter of numerical simulation are reasonably in good agreement with experiments, suggesting the present numerical simulations are accurate and reliable. The deviations of mean drag parameter $\overline{F_x}/(\rho U^2)$ and RMS lift parameter $F_y'/(\rho U^2)$ shows monotonic convergence, indicating the medium mesh is enough to resolve turbulent eddies around the hull. Therefore, the medium mesh is selected for all the remaining simulations.

Table 1 Results of grid convergence tests

Case	No. of cells			$\overline{F_x}/(\rho U^2)$ [m^2]	$F_y'/(\rho U^2)$ [m^2]
	Total	Background	Hull		
Coarse	1.04M	0.04M	1.00M	0.970	0.0644
Medium	2.53M	0.10M	2.43M	0.920	0.0303
Fine	6.25M	0.29M	5.96M	0.886	0.0280
Experiment[15]	-	-	-	0.912(±3.0%)	-

Forces and flow fields

Figure 5 shows the comparison of drag among the experimental and numerical results by Antony et al. [15] and present CFD results. Overall, all CFD methods agree well with experimental data, except AcuSolve and Fluent at 22.5° current heading. Fluent underestimated drag by 4.9% at 22.5° current heading and 0.374 m/s. However, no obvious deviations are observed in the present results.

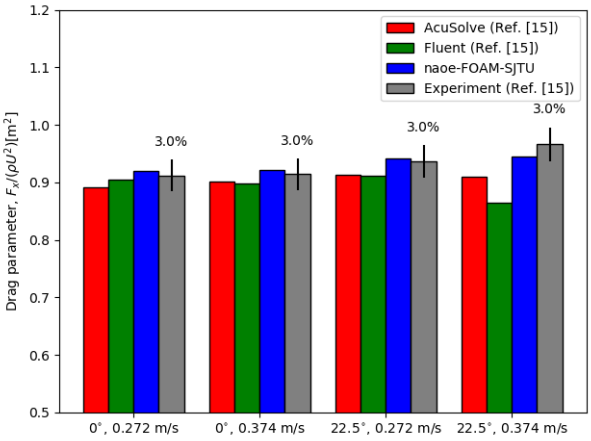


Figure 5 Comparison of drag between ref. [15] and present results

Figure 6 shows the instantaneous flow visualizations presented by streamwise velocity contours and streamlines on the cut-plane at z=-H/2 at U_c=0.272 m/s for different current headings. It can be seen that at 0° current heading, wake interference between side-by-side OCs is insignificant. However, the streamwise velocity increases between two upstream side-by-side ICs due to the narrower gap between ICs. Taking the upper-left OC as example, the streamlines behind it indicate there are two main recirculation bubbles located at position that has a lateral deviation to OC's x-direction centerline. The deviation is caused by the speed up flow between OC and corresponding IC. Looking from the streamwise direction, the wake interference between upstream and downstream OC is trivial because of the large spacing ratio L/D=4.95 (center-to-center distance to characteristic length). While the spacing ratio of ICs L/d=2.89 is small enough that the wake interference is non-trivial and cannot be neglected. Notably, the wakes of the downstream OCs are effected by ICs and are much wider due to the interaction between vortices of near wakes of the downstream OCs and upstream wakes. As for 22.5° current heading, a similar lateral deviation of recirculation region is observed behind the upstream OC. The staggered arrangement weakened the interaction between wakes of upstream and downstream ICs compared with 0° current heading. Nevertheless, the wake of downstream OC is influenced by the front IC and becomes wider.

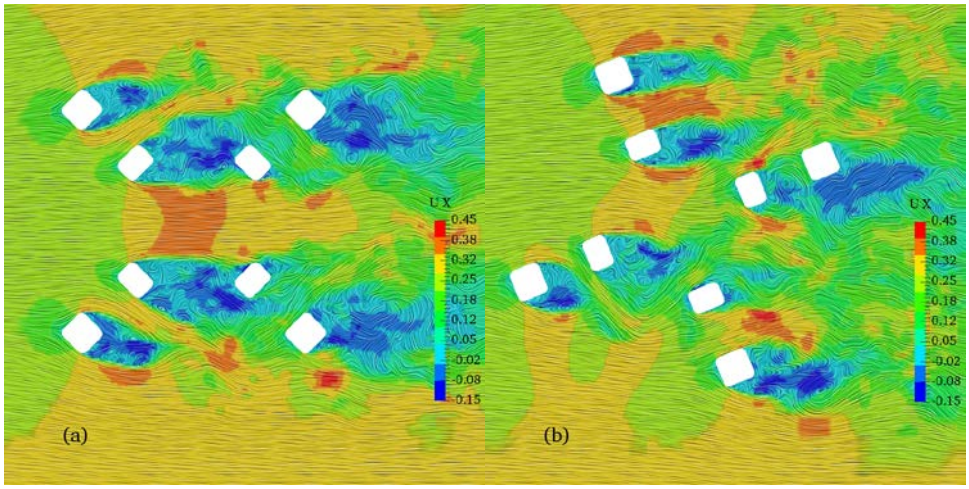


Figure 6 Instantaneous streamwise velocity contours and streamlines on the half-draft plane (z=-H/2) at U_c=0.272 m/s with (a) 0° and (b) 22.5°current headings

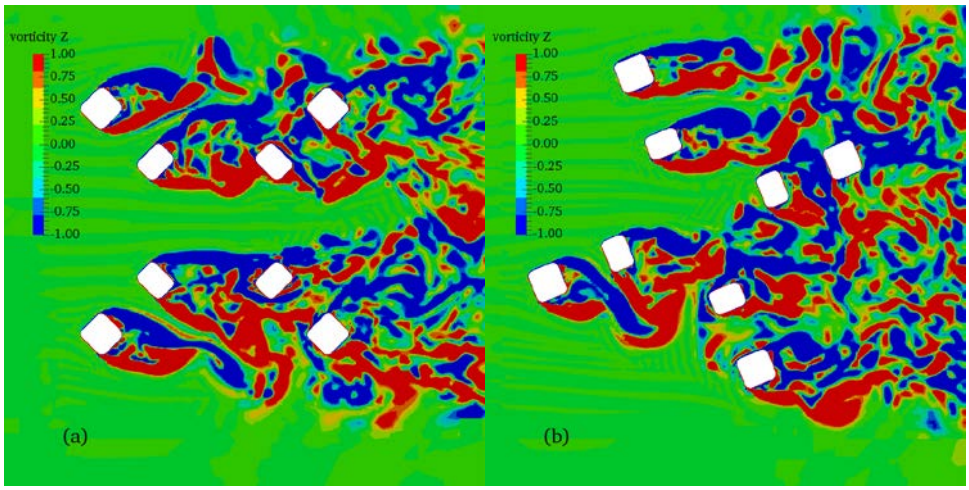


Figure 7 Instantaneous vorticity contours on the half-draft plane (z=-H/2) at U_c=0.272 m/s with (a) 0° and (b) 22.5°current headings

Figure 7 presents the instantaneous vorticity contour on the half-draft plane (z=-H/2) at U_c=0.272 m/s with different current headings. For upstream OCs and ICs at 0° current heading, the flow separation occurs in the vicinity of the rounded corner. No vorticities are found between the side-by-side ICs. The wakes of upstream ICs are strongly effected by upstream OCs, thus become much wider. In contrast, the upstream lateral IC at 22.5° current heading does not interfered by corresponding OC due to the providential orientation to current direction. Overall, the wakes of downstream columns interact with the vortices shed from upstream columns and break into small-scale eddies in the rear of PC Semi.

Figure 8 is the instantaneous pressure contour on the hull surface at U_c=0.272 m/s with different current headings. The view is seen from the upstream. For 0° current heading, the high-pressure region of the hull surface occurs exactly at the rounded-corner of the upstream OCs and ICs. For 22.5° current heading, it also appears at some downstream columns due to the staggered arrangement. It worth noting that in perpendicular surface to the high-pressure region for some columns, there exists some low-pressure regions which, in conjunction with the high-pressure regions, will result higher rotational moments around z-axis in compare with 0° current heading.

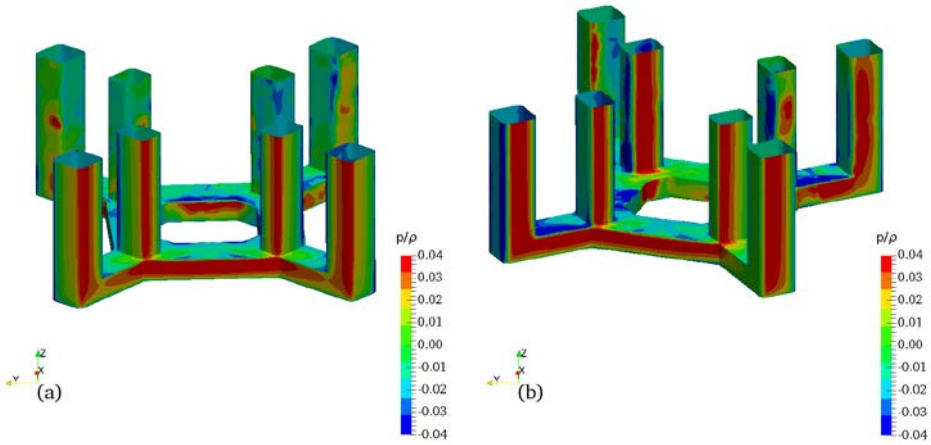


Figure 8 Instantaneous pressure contour on the hull surface at U_c=0.272 m/s with (a) 0° and (b) 22.5°current headings

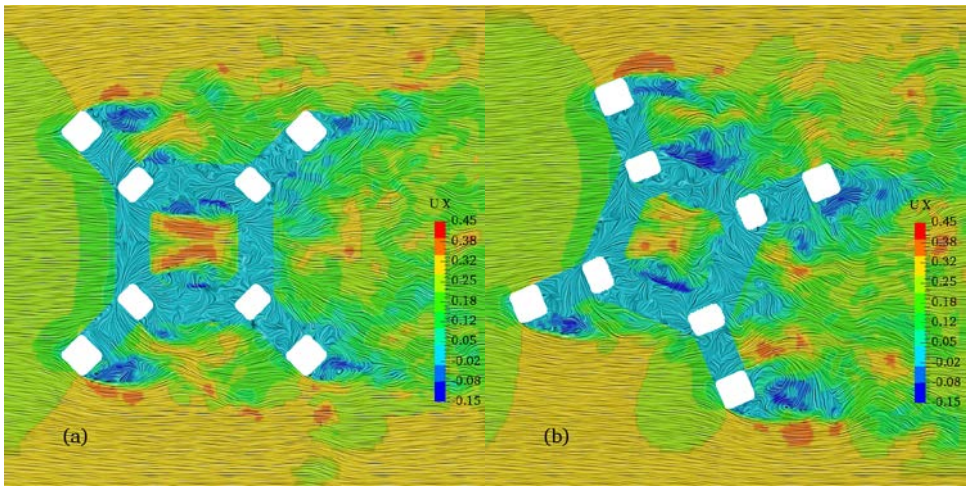


Figure 9 Instantaneous streamwise velocity contours and streamlines on the plane above pontoon at U_c=0.272 m/s with (a) 0° and (b) 22.5°current headings

Figure 9 presents the instantaneous streamwise velocity contours and streamlines on the plane above pontoon with different current headings. It clearly shows that the pontoon suppresses vortex sheds from the columns interior. In the region above moon pool, the high velocity area is smaller than that on the plane at $z=H/2$ in Figure 6. The diminution of streamwise velocity is induced by the large recirculation bubble around pontoon in the moon pool, as illustrated in Figure 10. Figure 10 also reveals that vortex mainly sheds from bottom shape corner of the pontoon. The vortex that sheds from top shape corner of the pontoon is mainly suppressed by the large flow velocities between upstream ICs.

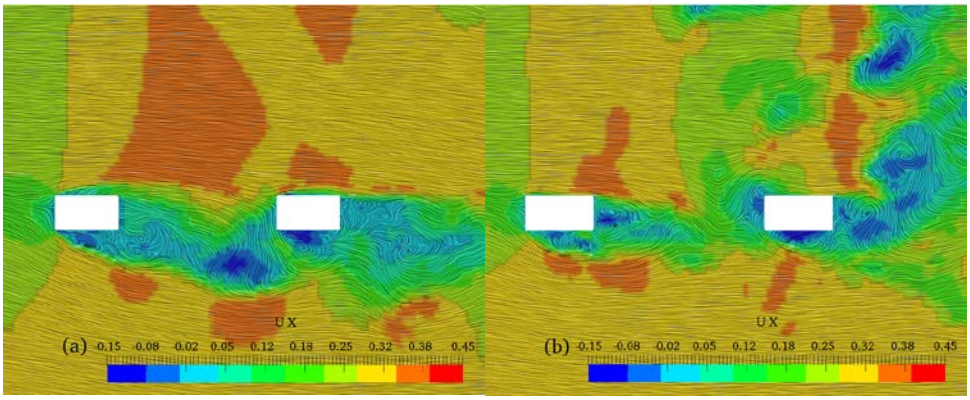


Figure 10 Instantaneous streamwise velocity contours and streamlines on the longitudinal section (y=0) at U_c=0.272 m/s with (a) 0° and (b) 22.5°current headings

Conclusions

In the present study, flow past a fixed Paired-Column Semi-Submersible at model scale 1:54 are numerically investigated at different flow velocities and current headings. The simulations are performed by the finite volume solver naoe-FOAM-SJTU developed on top of the OpenFOAM framework. The following conclusion can be drawn:

1. The drag parameters with 0° current heading vary little between current velocity 0.272m/s and 0.374m/s. However, they increase when the current heading changes from 0° to 22.5°. At 22.5° current heading, larger velocity will result in larger drag parameters.

2. Distinct wake interferences are observed at 0° current heading between ICs. Consequently, the pressures on the hull surface of downstream columns oscillate much stronger than that of upstream. At 22.5° current heading, surface pressure oscillations of downstream columns are much weaker due to the staggered arrangement of ICs.
3. Pontoons suppress vortex sheds from interior of the columns. Recirculation regions are observed in the moon pool behind the upstream pontoon.

The above conclusions can help us better understanding the wake interference in multi-column offshore structures. The current numerical approaches can be easily extended to VIM phenomena of the PC Semi. Investigations of the current headings effects on dynamic VIM response characteristics of PC Semi are ongoing.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Finnigan, T., and Roddier, D. (2007) Spar VIM model tests at supercritical reynolds numbers, *Proceedings of the 26th International Conference on Offshore Mechanics and Arctic Engineering*, San Diego, California, USA, 3, 731–740.

[2] Roddier, D., Finnigan, T., and Liapis, S. (2009) Influence of the Reynolds number on spar Vortex Induced Motions (VIM): Multiple scale model test comparisons, *Proceedings of the ASME 28th International Conference on Ocean, Offshore and Arctic Engineering*, Honolulu, Hawaii, USA, 5, 797–806.

[3] Gonçalves, R.T., Fujarra, A.L.C., Rosetti, G.F., and Nishimoto, K. (2010) Mitigation of vortex-induced motion (VIM) on a monocolumn platform: Forces and movements, *Journal of Offshore Mechanics and Arctic Engineering* **132**(4), 041102.

[4] Waals, O.J., Phadke, A.C., and Bultema, S. (2007) Flow induced motions of multi column floaters, *Proceedings of the 26th International Conference on Offshore Mechanics and Arctic Engineering*, San Diego, California, USA, 1, 669–678.

[5] Gonçalves, R.T., Nishimoto, K., Rosetti, G.F., Fujarra, A.L.C., and Oliveira, A.C. (2011) Experimental study on vortex-induced motions (VIM) of a large-volume semi-submersible platform, *Proceedings of the ASME 2011 30th International Conference on Ocean, Offshore and Arctic Engineering*, Rotterdam, The Netherlands, 7, 1–9.

[6] Fujarra, A.L.C., Rosetti, G.F., De, W., and Gonçalves, R.T. (2012) State-of-art on Vortex-Induced Motion: A comprehensive survey after more than one decade of experimental investigation, *Proceedings of the ASME 2012 31st International Conference on Ocean, Offshore and Arctic Engineering*, 4, 561–582.

[7] Delany, N.K., and Sorensen, N.E. (1953) Low-speed Drag of Cylinders of Various Shapes, *National Advisory Committee for Aeronautics*.

[8] Kawamura, T., Hiwada, M., Hibino, T., Mabuchi, I., and Kumada, M. (1984) Flow around a Finite Circular Cylinder on a Flat Plate : Cylinder height greater than turbulent boundary layer thickness, *Bulletin of JSME* **27**(232), 2142–2151.

[9] Okamoto, S., and Uemura, N. (1991) Effect of rounding side-corners on aerodynamic forces and turbulent wake of a cube placed on a ground plane, *Experiments in Fluids* **11**(1), 58–64.

[10] Sumner, D., Price, S.J., and Païdoussis, M.P. (2000) Flow-pattern identification for two staggered circular cylinders in cross-flow, *Journal of Fluid Mechanics* **411**, 263–303.

[11] Sayers, A.T. (1988) Flow interference between four equispaced cylinders when subjected to a cross flow, *Journal of Wind Engineering and Industrial Aerodynamics* **31**(1), 9–28.

[12] Liang, Y., Tao, L., Xiao, L., and Liu, M. (2016) Experimental and Numerical Study on Flow Past Four Rectangular Columns in Diamond Configuration, V001T01A044.

[13] Zou, J., Poll, P., Roddier, D., Tom, N., and Peiffer, A. (2013) VIM testing of a paired column semi submersible, *Proceedings of the ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering*, Nantes, France, 7, V007T08A001.
[14] Antony, A., Vinayan, V., Halkyard, J., Kim, S.-J., Holmes, S., and Spernjak, D. (2015) A CFD based analysis of the Vortex Induced Motion of deep-draft semisubmersibles, *Proceedings of the Twenty-Fifth (2015) International Ocean and Polar Engineering Conference*, Kona, Big Island, Hawaii, USA, 1048–1055.
[15] Antony, A., Vinayan, V., Holmes, S., Spernjak, D., Kim, S.J., and Halkyard, J. (2015) VIM Study for Deep Draft Column Stabilized Floaters, *Offshore Technology Conference*, Houston, Texas, USA.
[16] Kim, S.J., Spernjak, D., Holmes, S., Vinayan, V., and Antony, A. (2015) Vortex-Induced Motion of Floating Structures: CFD Sensitivity Considerations of Turbulence Model and Mesh Refinement, *Proceedings of the ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering*, St. John's, Newfoundland, Canada, 2, V002T08A057.
[17] Vinayan, V., Antony, A., Halkyard, J., Kim, S.-J., Holmes, S., and Spernjak, D. (2015) Vortex-induced motion of deep-draft semisubmersibles: A CFD-based parametric study, *Proceedings of the ASME 2015 34th International Conference on Ocean, Offshore and Arctic Engineering*, St. John's, Newfoundland, Canada, 2, V002T08A003.
[18] Kara, M., Kaufmann, J., Gordon, R., Sharma, P., and Lu, J. (2016) Application of CFD for Computing VIM of Floating Structures, *Offshore Technology Conference*, Houston, Texas, USA.
[19] Shen, Z., Wan, D., and Carrica, P.M. (2015) Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering, *Ocean Engineering* **108**, 287–306.
[20] Spalart, P.R. (2009) Detached-Eddy Simulation, *Annual Review of Fluid Mechanics* **41**(1), 181–202.
[21] Spalart, P., Jou, W., Strelets, M., and Allmaras, S. (1997) Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach, *Advances in DNS/LES* **1**, 4–8.
[22] Zhao, W., and Wan, D. (2016) Detached-Eddy Simulation of Flow Past Tandem Cylinders, *Applied Mathematics and Mechanics* **37**(12), 1272–1281.

Ship Hull Form Design Using a Kriging-based Global Optimization Algorithm

Aiqin Miao[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: maq046@163.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

The kriging-based global optimization algorithm, Efficient Global Optimization (EGO), is applied to ship hull form design problems. The Kriging-based surrogate model is used to approximate the relationship between the design variables (input) and the objective functions (output) using a stochastic process. The kriging model drastically reduces the computational time required for objective function evaluation in the optimization(optimum searching) process. Expected improvement (EI) is used as a criterion to select additional sample points. This makes it possible not only to improve the accuracy of the surrogate model, but also to explore the global optimum efficiently. The present method is applied to five test functions and a ship hull form design, which makes the optimal hull form with the best resistance performance in calm water. It turns out that the EGO method is a good method for optimization design and can further extend to other ship hull form optimization design problems based on computational fluid dynamics.

Keywords: kriging model; efficient global optimization; uncertainty; ship optimization design; OPTShip-SJTU

Introduction

With the growth in computing power of current computers and the advances in computational techniques, today especially computational fluid dynamics (CFD) has become an invaluable tool for ship hull form optimization design. However, in the process of optimization design, the number of objective function evaluations using high-fidelity numerical analysis solvers, is severely limited by time and cost, even with current supercomputers.

One alternative is to construct a simple surrogate model instead of the complicated numerical analysis. The surrogate model expresses the relationship between the design variables (input) and the objective functions (output) using a stochastic Gaussian process. The model requires very little time to evaluate the objective function. It enables designers to greatly save computational cost. The most widely used surrogate model are the polynomial-based model, the response surface model, the kriging model and so on.

The kriging model (Wu J. W., 2017) has gained popularity for the engineering optimization design, recently. The model predicts the value of the unknown point using stochastic processes. The model has sufficient flexibility to represent the nonlinear and multimodal functions at the expense of computation time. However, the computation time to construct the kriging model is still short compared to that of direct high-fidelity numerical analysis.

However, the traditional optimization algorithms, like genetic algorithm (GA), differential evolution (DE), require many objective function evaluations, which may be impractical if we rely solely on the time-consuming high-fidelity numerical analysis solver. Here, the time-consuming numerical analysis solver in the objective function evaluation process of optimization is replaced with the kriging model. However, that it is possible to miss the global optimum in the searching space if we rely only on the prediction value of the kriging model because the model includes uncertainty at the prediction point. For robust exploration of the global optimum point, both the prediction value and its uncertainty should be considered at the same time. This concept is expressed in the criterion expected improvement (EI). EI includes the probability of a point being optimum in the design space. By the selection of the best EI point as the additional sample point, improvement of the model and robust exploration of the global optimum can be achieved at the same time, which is called Efficient Global Optimization (EGO) (Jones D. R., 1998; Jeong S., 2015).

In this paper, differential evolution (DE) algorithm is adopted as the searching algorithm to find maximum value of the EI function. DE is similar to genetic algorithm. However, compared with other evolutionary algorithms, DE retains the population-based global search strategy, which reduces the complexity of genetic operation by using real number coding, simple mutation based on difference and one-to-one competitive survival strategy. At the same time, DE's unique memory ability makes it possible to dynamically track the current search situation to adjust its search strategy, which has strong global convergence ability and robustness, and does not need to use the characteristic information of the problem. DE is suitable for solving the optimization problem in complex environment that the mathematical programming method may not solve.

In the first half of this paper, the EGO method will be briefly introduced back to the Kriging modelling. In the second of this paper, the method will be applied to the optimization problems of the test functions and ship design of Wigley.

Kriging Model

Kriging model (Simpson et al., 1994, 2004) is developed from mining and geostatistical applications involving spatially and temporally correlated data. This model combines a global model and a local component:

$$y(x) = f(x) + z(x) \tag{1}$$

where $y(x)$ is the unknown function of interest, $f(x)$ is a known approximation function of x, and $z(x)$ is the realization of a stochastic process with mean zero, variance σ^2, and non-zero covariance. With $f(x)$ and $z(x)$, the kriging model can build the surrogate model between the input variables and output variables.
The kriging predictor is given by:

$$\hat{y} = \hat{\beta} + \mathbf{r}^{\mathrm{T}}(x)\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{f}\hat{\beta}\right) \tag{2}$$

where y is an ns-dimensional vector that contains the sample values of the response; $\mathbf{f}$ is a column vector of length n_s that is filled with ones when f is taken as a constant; $\mathbf{r}^{\mathrm{T}}(x)$ is the correlation vector of length n_s between an untried x and the sampled data points $\left\{x^{(1)}, x^{(2)}, \ldots, x^{(n_s)}\right\}$ and is expressed as:

$$\mathbf{r}^{\mathrm{T}}(x) = \left[R\left(x, x^{(1)}\right), R\left(x, x^{(2)}\right), \ldots, R\left(x, x^{(n_s)}\right),\right]^{\mathrm{T}} \tag{3}$$

Additionally, the Gaussian correlation function is employed in this work:

$$R(x^i, x^j) = \exp\left[-\sum_{k=1}^{n_{dv}} \theta_k \left| x_k^i - x_k^j \right|^2 \right] \tag{4}$$

In equation (2), $\hat{\beta}$ is estimated using equation (5):

$$\hat{\beta} = \left(\boldsymbol{f}^{\mathrm{T}} \mathbf{R}^{-1} \boldsymbol{f} \right)^{-1} \boldsymbol{f}^{\mathrm{T}} \mathbf{R}^{-1} \boldsymbol{y} \tag{5}$$

The estimate of the variance $\hat{\sigma}^2$, between the underlying global model $\hat{\beta}$ and y is estimated using equation (6):

$$\hat{\sigma}^2 = \left[\left(y - f\hat{\beta} \right)^T R^{-1} \left(y - f\hat{\beta} \right) \right] / n_s \tag{6}$$

where $f(x)$ is assumed to be the constant $\hat{\beta}$. The maximum likelihood estimates for the θ_k in equation (4) used to fit a kriging model are obtained by solving equation (7):

$$\max_{\theta_k > 0} \Phi(\theta_k) = -\left[n_s \ln\left(\hat{\sigma}^2 \right) + \ln|\mathrm{R}| \right] / 2 \tag{7}$$

where both $\hat{\sigma}^2$ and $|\mathrm{R}|$ are functions of θ_k. While any value for the θ_k create an interpolative kriging model, the "best" kriging model is found by solving the k-dimensional unconstrained, nonlinear, optimization problem given by equation (7).

The accuracy of the prediction value largely depends on the distance from sample points. Intuitively speaking, the closer point x to the sample point, the more accurate is the prediction $\hat{y}$. This intuition is expressed as

$$s^2(x) = \hat{\sigma}^2 \left[1 - r^{'} R^{-1} r + \frac{(1 - 1R^{-1}r)^2}{1^{'} R^{-1} 1} \right] \tag{8}$$

where $s^2(x)$ is the mean squared error of the predictor and it indicates the uncertainty at the estimation point. The root mean squared error (RSME) is expressed as $s = \sqrt{s^2(x)}$.

DE algorithm

DE algorithm (Storn R., 1997) is mainly used to solve the global optimization problem of continuous variables. The main working steps are basically the same as other evolutionary algorithms, including Mutation, Crossover and Selection. The basic idea of the algorithm is to start from a randomly generated initial population and use the difference vector of two individuals randomly selected from the population as the random variation source of the third individual, and weight the difference vector and according to certain rules plus the third individual to produce a variation of individuals, which are called mutations. Then, the mutated individuals are mixed with a predetermined target individual to generate the test one. This process is called crossover. If the fitness value of the test individual is superior to the fitness value of the target individual, the test individual replaces the target individual in the next generation, otherwise the target individual is still saved, which is called the selection. In each evolutionary process, each individual vector is taken as the target individual once. The algorithm keeps the good individual and removes the poor individual through the iterative calculation, and the search process is approached to the global optimal solution. Here we use the DE algorithm to maximize the EI value.

A kriging-based global efficient optimization algorithm

Traditionally, once the surrogate model is constructed, the optimum point can be explored using an arbitrary optimizer on the model. However, it is possible to miss the global optimum because the approximation model includes uncertainty at the predicted point.

In Fig. 1, the solid line is the real shape of objective function. Eight points are selected to construct the kriging model, which is shown as red points. The minimum point on the kriging model is located near x=16, whereas, the real global minimum of the objective function is situated near x=17. Searching for the global minimum using the present kriging model will not result in the real global minimum near x=17. For a robust search of the global optimum, both the predicted value by the kriging model and its uncertainty should be considered at the same time.

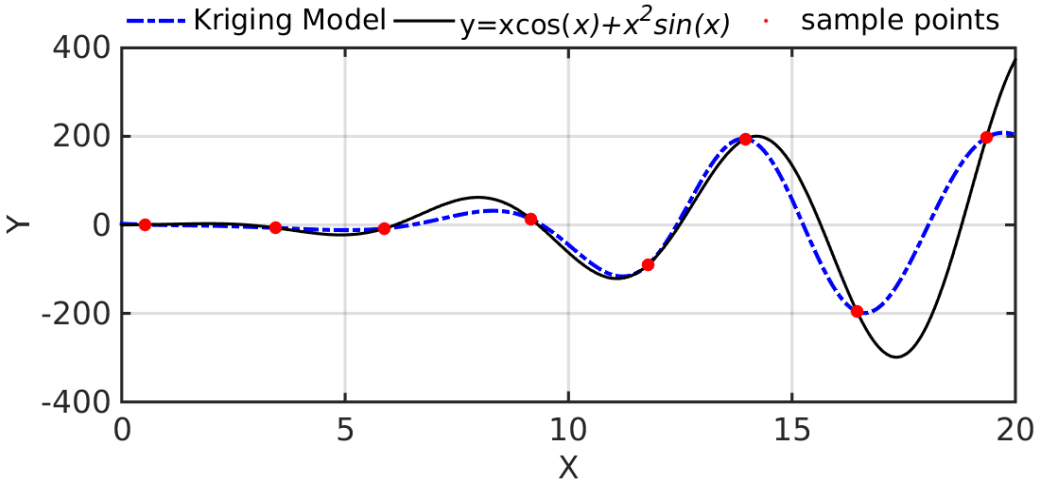


Figure 1. The Kriging model and the real function curve

When Kriging model is built, the mean predicted value and the standard error of the kriging model at any point can be evaluated. Considering the uncertainty of the model, this concept is expressed in the criterion of EI. The EI of minimization problem can be calculated as

$$E[I(x)] = (f_{\min} - \hat{y})\Phi[(f_{\min} - \hat{y})/s] + s\phi[(f_{\min} - \hat{y})/s] \tag{9}$$

Where $f_{\min}$ is the minimum value among n sampled values. Φ and ϕ are the standard distribution and normal density, respectively.
In fact, if we compute the derivative of EI as given in equation (9) with respect to $\hat{y}$ and s, we gets several terms that cancel, resulting in the simple following expressions:

$$\partial EI/\partial\hat{y} < 0, \partial EI/\partial s < 0 \tag{10}$$

It turns out that EI is monotonic in $\hat{y}$ and in s. Thus, we see that the EI is larger the lower is $\hat{y}$ and the higher is s. By selecting the maximum EI point as additional sample point through DE algorithm mentioned above, robust exploration of the global optimum and improvement of the model can be achieved simultaneously.

The overview over the whole efficient global optimization (EGO) procedure mentioned before is shown in Figure 2. First, the initial sample points should be chosen by experiment of design uniformly covering the whole design space. Secondly, an ordinary Kriging model is built and used to predict the objective for each design variable. Thirdly, the expected improvement (EI) balancing between regions of the low mean prediction and of high standard error is constructed to select the next point. The choice of the next sample point is the maximization of the EI value. Next, the objective function in the new point is calculated accurately and used to build a new surrogate model with the initial sample points, thus the

next iteration is initiated. Finally, when the EI value is very small after n iterations, i.e., $\max(EI) < \Delta s \cdot (\max(y) - \min(y))$, where Δs is the relative stopping tolerance, or reach the maximization of iteration steps, the loop should be stopped.

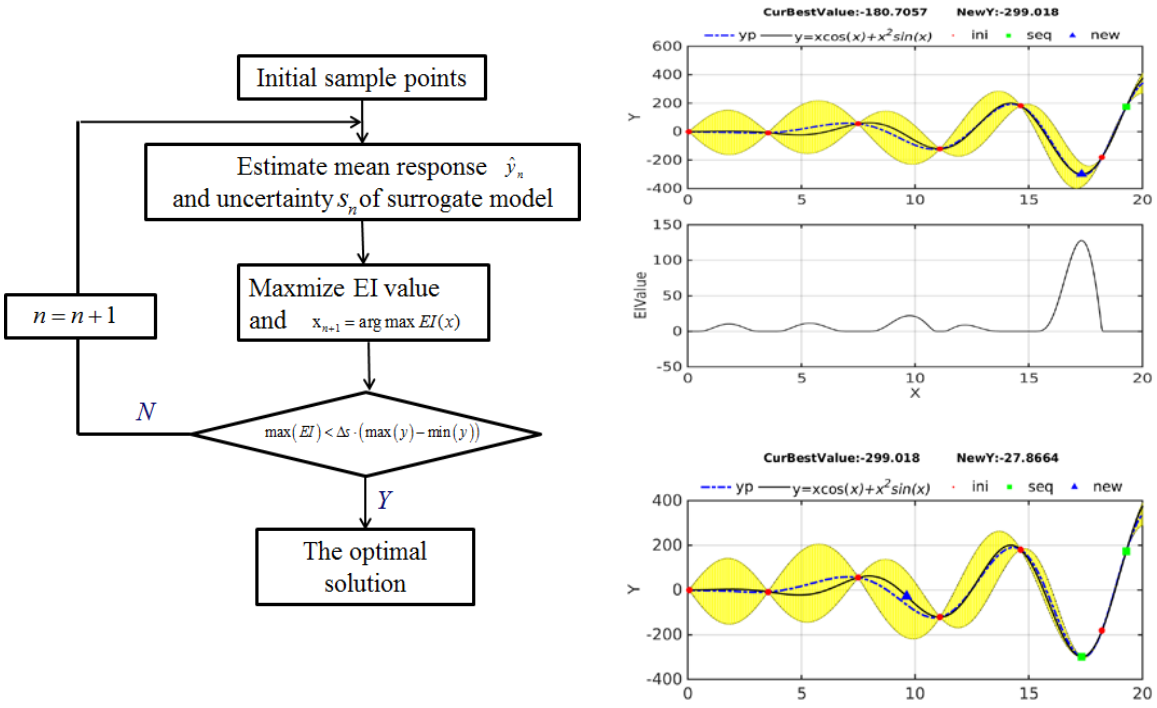


Figure 2. The flow chart of efficient global optimization: on the left, the steps are briefly described; on the right, an example is given (predetermined design points as red dots, the added new points as green squares and the next new point as a blue triangle).

Applications

Application of the test functions

we now apply this method EGO to five test functions, from a literature (Dixon L.C.W. and Szego G. P., 1978): Six-hump camel-back function, the Branin function, the Goldstein-Price function, the Hartman 3 function and the Hartman 6 function. The dimensions of these test functions are 2, 2, 2, 3 and 6, respectively. The results of the test functions by the EGO algorithm is shown in table.

For each function, we report the number of evaluations to meet stopping criterion, and report the actual relative error on convergence.

Table 1. Test function results by the EGO method

Function Name	Dimension	Evaluations to meet stopping criterion	Optimal result by EGO	Real optimal value	Relative error
Six-hump	2	30	-1.0259	-1.03163	-0.56%

Branin	2	30	0.39985	0.398	0.46%
Goldstein-Price	2	31	3.00326	3	0.11%
Hartman 3	3	35	-3.83	-3.86	-0.78%
Hartman 6	6	84	-3.25503	-3.32	-1.96%

As we can see, the results of the former four test functions achieve less than an actual relative error of 1%. For the Hartman 6 function, the actual relative error on convergence is slightly greater. Possibly, it is caused by the unknown but estimated correlation parameters. As a results, the standard error may be a bit too small, causing us to underestimate EI value.

Application of ship optimization design of Wigley

In this study, the EGO method was applied to ship hull form design and the optimization of the resistance performance in calm water.

$$f_{obj} = C_w, Fr = 0.3 \tag{11}$$

The design problem is to minimize the wave-making resistance coefficient of Wigley at the design speed. The Radial Basis Function (RBF) method (De Boer A., 2007) was applied to modify ship hull form. The design variables are parameters closely related to ship hull form modification. The total 7 design variables in table 2 are used to define the geometry of ship hull form. The upper and lower bound of each parameter is determined to avoid unrealistic ship hull geometry.

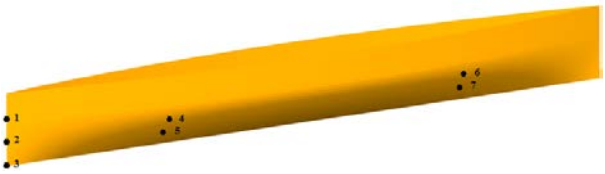


Figure 3. The control points distributed on Wigley hull by RBF method

Table 2. The design variables and their ranges

Design variables	Deformation direction	The bound
1#	x	[-0.0025, 0.0025]
2#	x	[-0.0025, 0.0025]
3#	x	[-0.0025, 0.0025]
4#	y	[-0.0025, 0.0025]
5#	y	[-0.0025, 0.0025]
6#	y	[-0.0025, 0.0025]
7#	y	[-0.0025, 0.0025]

At the early stage of optimization design, 35 sample points are spread over the design space and selected by Optimal Latin Hypercube Sampling (OLHS) to obtain a kriging model (Park

J. S., 1994). In this study, OLHS has the beneficial property of orthogonality and uniformity in the multidimensional design space. The number of sample points is very important to keep the kriging model accurate in the traditional optimization process. however in the present kriging model, additional sample points will be added later in the region where the accuracy is not good enough, based on EI evaluation. The wave-making resistance coefficients of 35 sample hull form and additional sample points are all evaluated using a potential flow theory, Nuemann-Michell method (Noblesse F., 2013; Liu, X. Y., 2016; Wu, J. W., 2016).

After efficient global optimization search, The total number of sample points reached 45, after adding 10 more sample points. The objective function converges to the minimum value, 1.046E-03, a larger reduction of 18.46% than the initial value. the summary of the relative parameters of EGO method is shown in table. And The design values of the optimal hull form is shown in Table 4.

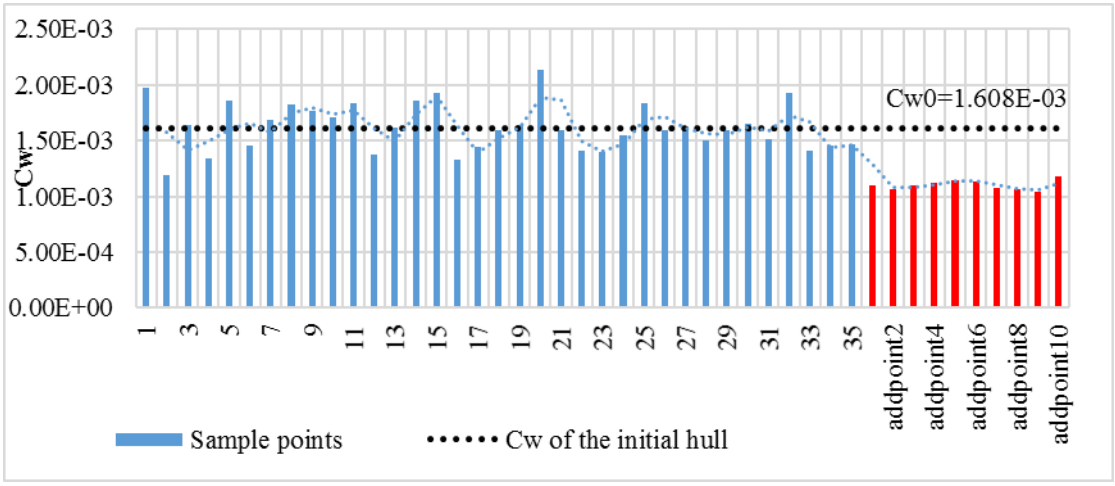


Figure 4. The initial sample hulls and the additional new hulls used in the EGO method

Table 3. The parameters of the EGO method

The initial number of sample points	35	The additional number of sample points	10
The number of iterations	10	Optimization time (h)	1.075h
Node number	1		

Table 4. The values of design variables and the optimal solution of the objective function after optimization procedure

Modification method	Design variables	Convergence value
RBF method	1#	0.002297
	2#	-0.002346
	3#	-0.002450

4#	-0.00250
5#	-0.00250
6#	-0.00250
7#	-0.00250
The initial objective function value	1.618×10^{-3}
The convergent objective function value	1.046×10^{-3}

The following figures depict the comparisons of body lines and shapes of the initial Wigley and the optimal ship. The optimal hull form is thinner than the initial one globally. Another apparently different region is the fore part of the hull. The optimal hull is changed like 's' curve, indent near the bottom and convex close to the top part.

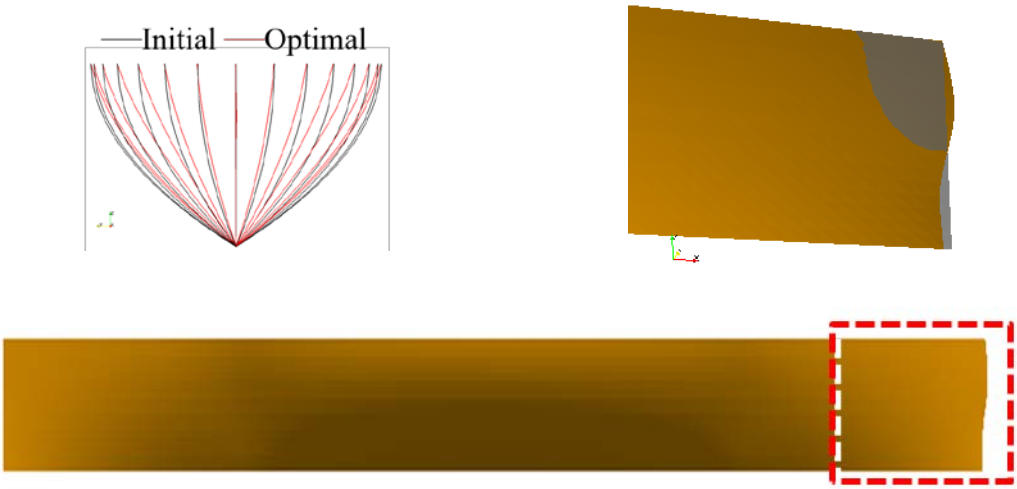


Figure 5. Comparisons of the body lines and shapes between the initial hull and the optimal one

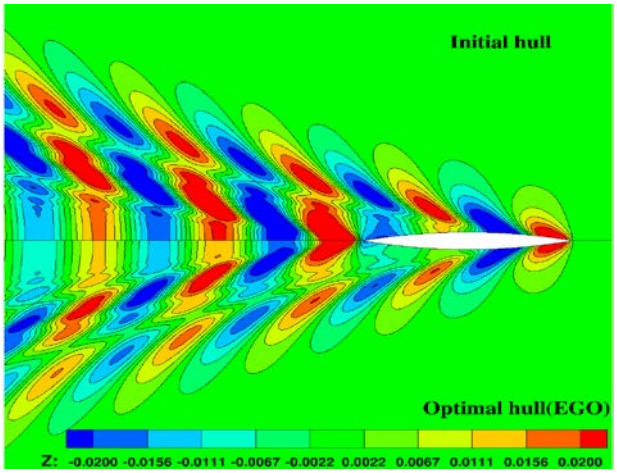


Figure 6. Comparison of free surface elevation between the initial hull and the optimal one

Figure 6 is the wave profile of free surface of the two ships. In contrast, the optimal hull by EGO method generates lower wave than the initial one obviously, thus leads to the reduction of the wave-making resistance of the optimal hull. Also comparing the pressure distribution of the two ships in Fig. 7, the fore part of the optimal hull is with lower pressure than the initial one. Figure 8 shows the wave-making resistance coefficients of the two ships at a series of speeds, which makes it clear that the optimal ship has a lower wave-making resistance coefficient than the initial ship at certain range of speeds. This results validate that our optimization design in this study is robust.

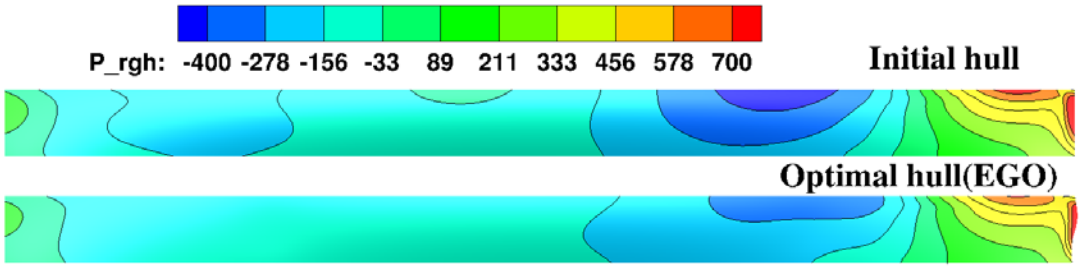


Figure 7. Comparison of pressure distribution between the initial hull and the optimal one

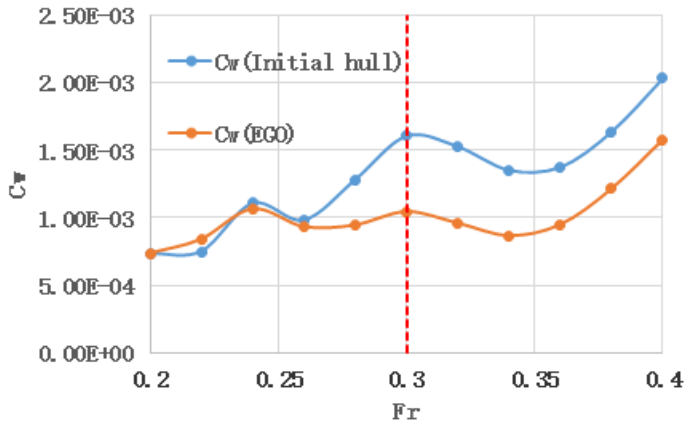


Figure 8. The wave-making resistance coefficients of the best variation at a series of speeds

Conclusions

This paper presents a Kriging-based global optimization method, efficient global optimization (EGO), different from the ordinary optimization method. It combines the surrogate modeling with the optimization algorithm. By this method, not only the accuracy of surrogate model is continuously improved but also the solution of the optimization problem keeps searched in the iterative procedure. The method is successfully applied to the test functions and ship hull form optimization design. The results demonstrate the usability of the method in ship hull form optimization design. In the future, it will be used to the more ship optimization problem, such as the ship hull form design to improve comprehensive hydrodynamic performance, based on entire CFD.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] De Boer A, Van der Schoot M S, Bijl H. (2007) Mesh deformation based on radial basis function interpolation. *Computers and Structures* **85**, 784-795.

[2] Dixon, L.C.W. and Szego, G.P. (1978) The global optimization problem: an introduction. *L.C.W. Dixon and G.P. Szego (eds.), Towards Global Optimisation* **2**, 1-15.

[3] Jones, D. R., Schonlau, M., and Welch, W. J. (1998) Efficient global optimization of expensive black-box functions. *Journal of Global Optimization* **13**, 455-492.

[4] Jeong, S., Murayama, M., and Yamamoto, K. (2015) Efficient optimization design method using kriging model. *Journal of Aircraft* **42**, 413-420.

[5] Liu, X. Y., Wu, J. W., and Wan, D. C. (2016) Multi-objective optimization for a surface combatant using Neumann-Michell theory and approximation model. *Proceedings of 12th International Conference on Hydrodynamics* (ICHD), 18-23.

[6] Noblesse, F., Huang, F. X. and Yang, C. (2013) The Neumann-Michell theory of ship waves. *Journal of Engineering Mathematics* **79**, 51-71.

[7] Park, J. S. (1994) Optimal Latin-hypercube designs for computer experiments. *Journal of Statistical Planning and Inference* **39**, 95-111.

[8] Simpson, T. W. (1998) Comparison of response surface and kriging models in the multidisciplinary design of an aerospike nozzle. *Institute for Computer Applications in Science and Engineering (ICASE)*.

[9] Simpson, T. W., Booker, A. J., Ghosh, D., Giunta, A. A., Koch, P. N. and Yang, R. J. (2004) Approximation methods in multidisciplinary analysis and optimization: a panel discussion. *Structural and multidisciplinary optimization* **27**, 302-313.

[10] Storn, R., and Price, K. (1997) Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Kluwer Academic Publishers*.

[11] Weihs, C., Herbrandt, S., Bauer, N., Friedrichs, K. and Horn, D. (2016) Efficient global optimization: Motivation, variations and applications.

[12] Wu J. W., Liu X. Y., and Wan D. C. (2016) Multi-objective hydrodynamic optimization of ship hull based on approximation model. *Proceedings of 26th International Offshore and Polar Engineering Conference* (ISOPE), Rhodes, Greece, 814-820.

[13] Wu, J. W., Liu, X. Y., Zhao, M., and Wan, D. C. (2017) Neumann-Michell theory-based multi-objective optimization of hull form for a naval surface combatant. *Applied Ocean Research **63***, 129-141.

Numerical investigation of open water performance of hybrid CRP podded propulsion system

Dongya He[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: tjuwjh@126.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

Abstract

Hybrid CRP podded propulsion system inherits merits of contra-rotating propellers (CRPs) and Azipod propulsion system. Firstly it matains good maneuverability of Azipod system and will decrease carbon emission since Azipod system is powered by electricity. Secondly the rear propeller can drastically recover swirl energy wasted by the front propeller in the wake flow field. Thirdly the pressure magnitude suffered by the two propellers is reduced dramatically, therefore the hybrid CRP podded propulsion system can relieve undesired cavitating and vibrating level compared with single propeller . In this paper, the hybrid system is composed of a CRPs a streamlined pod. Based on CFD (Computational Fluid Dynamics) technology with RANS function, k-ω SST turbulence model and sliding mesh method, the hybrid CRP podded propulsion system is simulated at different advance coefficients. Predicting results show that the Azipod propulsion system has little influence on the front propeller while the latter is strongly affected by the former. K_T and $10K_Q$ have 8 small periods when propeller rotates 180 degrees and their fluctuating amplitude is very small under considered axial spacing between the propeller centerlines. The efficiency is 3.5% higher than single propeller, the rear pod shape should be optimized to reduce its drag force. As the spacing increases to some extent, K_T and $10K_Q$ become smaller while efficiency almost keeps constant. Wake flow field information such as velocity distribution and vorticity structure reflect complex interactions between front propeller and Azipod propulsion system.

Keywords: Hybrid CRP podded propulsion system, sliding mesh method, open water performance

Introduction

Nowadays in order to reduce costs of ship operation, cargo carriers such as container ships, are designed to be more bigger and faster, then propeller needs larger radius and suffers more heavier weight and load. This may lead to serious non-uniform wake flow and undesirable cavitation and propeller vibration which will greatly deteriorate propulsion efficiency. In the past decades, majority of vehicles are propelled by traditional screw propeller with one shaft linked to main engine, on one hand, low efficiency of this propulsion way under heavy load condition consumes large quantity of fuel oil, on the other hand, the limitation of engine power cannot supply sufficient thrust to propel larger ships. Every year all kind of ships will discharge large amount of pollution gas, report from IMO points out that in 2009, shipping industry all over the world discharged 96.5 million tons of carbon dioxide and this data will be increased to 153 million tons in 2030. In order to promote development of “Green ship”, IMO puts forward three indexes , namely EEDI (Energy Efficiency Design Index), EEOI (Energy Efficiency Operational Index), SEEMP (Ship Energy Efficiency Management Plan)

to regulate ship design and operation. Ship design with lower energy consuming, higher propulsion efficiency and better cavitation performance has turn to be an international tendency.

Contra-rotating propellers (CRPs) , as figure 1 shows, is composed of two conventional propellers which rotate axially in reverse direction. Compared with single propeller, CRPs has better propulsion efficiency and energy-saving effect. However complexity of shaft system and high costs of installation and maintenance limit its further development. Then hybrid CRP podded propulsion system was put forward and applied widely in recent years. Hybrid CRP podded propulsion system, as figure 2 shows, inherit merits of CRPs and Azipod propulsion system. Firstly it maintains good maneuverability of Azipod system and will decrease carbon emission since Azipod system is powered by electricity. Secondly the rear propeller can drastically recover swirl energy wasted by the front propeller in the wake flow field. Thirdly the pressure magnitude suffered by the front and Azipod system decreases dramatically compared to single propeller, therefore, the hybrid CRP podded propulsion system can relieve undesired cavitation and vibrating level to obtain higher propulsion efficiency.

Figure 1. CRPs configuration

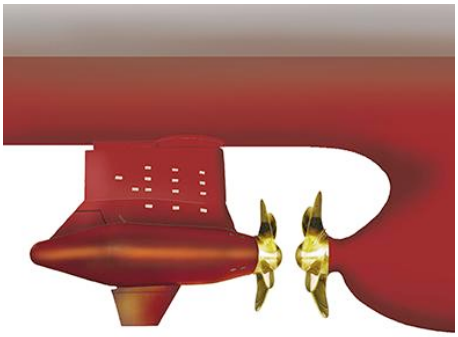


Figure 2. Hybrid CRP podded system

Azipod propulsion system[1] was firstly put forward and applied by ABB with great success in 1999 and then in 2002 hybrid CRP podded propulsion system[2] was developed based on previous work. In 2001, Samsung Heavy Industry Co. Ltd[3]. and ABB agreed to develop a Ultra Large Container Vessel (ULCC) with the hybrid CRP podded propulsion system, later Samsung Ship Model Basin carried model tests for ULCC with single screw, twin screw, and hybrid CRP podded propulsion system, and finally hybrid CRP podded propulsion system was proven to the most efficient propulsion system for carrier with this type and size, power savings to twin screw were about 9% and to single screw were 5%. Sheng et al. (2012)[4] performed open water experiment of hybrid CRP podded propulsion system in cavitation tunnel and the experimental data were compared with results predicted by CFD. Both model test and numerical model were proven to be reliable and credible.

Open water performance and interactions between front propeller and the rear Azipod propulsion system are critical issues in initial design period of hybrid CRP podded propulsion system. Up to now, there three main methods, model test in cavitation tunnel, lifting surface theory and CFD technology, have been widely used to propeller design. Model test has relatively high accuracy, however high cost of model building, experimental devices and long experimental period limit its popular application in ordinary people, meanwhile experimental result is sensitive to accuracy of detection instruments. Lifting surface theory has high computing efficiency, but for ignorance of viscosity, it rely on much experience to rectify model. Yang et al. (1991,1992) [5][6] investigated the steady and unsteady performance of contra-rotating propellers by lifting surface theory. Owing to the great progress in numerical

algorithms and supercomputer, the applications of CFD technology are advancing rapidly in the fields of ship hydrodynamics. Because CFD technology is based on actual fluid control functions (Navier-Stokes equations) which take the viscosity and rotation into account, thus it can correctly model nonlinear wake deformation and flow separation due to heavy loading. Up to now, MRF (Multiple Reference Frame) method, overset mesh method, and sliding mesh method are the three main techniques dealing with propeller' rotation. MRF method can only be used to solve steady problems, in other words, dynamic flow field cannot be obtained. Though this method has high efficiency, its precision is not so good as overset mesh method and sliding mesh method. Overset mesh method has been extensively applied to handle problems that have multiple moving objects with many degrees of freedom. Different grids will exchange their information through an interpolating code named SUGGAR++(Noack et al., 2009）[7]on the overlapping area. At sacrifice of relatively large of computing resources, this method will guarantee high accuracy. Shen, et al (2012)[8] carried out KCS self-propulsion and maneuvering by CFD solver naoe-FOAM-SJTU with overset mesh technique, predicted results agree well with their experimental data. Comparing with MRF and overset mesh method, sliding mesh method keeps equal precision as overset mesh method, but its computing efficiency is greatly improved since this method only needs interpolation between overlapping faces of rotational region and static region. Wu et al (2016) [9]compared accuracy and computing efficiency of those three method applied to numerical prediction of open-water performance of single propeller. Based on sliding mesh method, Zhou (2014) [10]investigated unsteady flow around wind turbines with different blades numbers. Wang, et at (2012)[11] studied two sets of CRPs' open-water performance developed by David W Taylor Naval Ship R&D Center using CFD method, numerical predicted results agree well with their experimental counterparts, furthermore, he investigated CRPs' periodical unsteady thrust and torque in detail. In the present work, CFD technology by solving RANS functions and k-ω model which has been widely used to predict hydrodynamics of propeller with higher computing accuracy and efficiency is used.

Numerical methods

Governing equations

In present work, fluid is assumed to be incompressible, RANS functions including mass and momentum conservation equations are listed as follow:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{1}$$

$$\frac{\partial}{\partial t}(u_i) + \frac{\partial}{\partial x_j}(u_i u_j) = -\frac{\partial p}{\partial x_j} + \frac{1}{\mathrm{Re}}\frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{\partial}{\partial x_j}\left(-\overline{u_i' u_j'}\right) \tag{2}$$

Where u_i are averages of velocity components in three directions; j (=1, 2, 3) means different direction; R_e is Reynolds number; $-\overline{u_i' u_j'}$ are Reynolds stress.

Turbulence model

In order to solve RANS functions, turbulence model is introduced to model Reynolds stress. k-ω SST[5](Shear Stress Transport) was put forward by Menter in 2003, k-ω turbulence model is effective near wall surface while in the far field, k-ε turbulence model become effective. k-ω SST has been widely used in research of hydrodynamics of navel architectures.

Sliding mesh method

Sliding mesh method is typically competent to solve unsteady problems in industries, such as propeller, pump, turbine and so on. It has higher accuracy and computing efficiency because fluid information is transferred only at overlapping area in a simple interpolation way. As can be seen in figure 5, the whole computing zone is divided into three zones, two of them surround the front and rear propellers respectively and they will rotate synchronously with propeller. The residual big zone is kept static. Interpolation will be done at overlapping area based on weight factor. In figure 3, blue surface and red surface are named master surface slave surface respectively. The contribution master cell 1 makes to slave cell 1 is measured by the weight factor which is defined by how much overlapping area master cell 1 accounts .

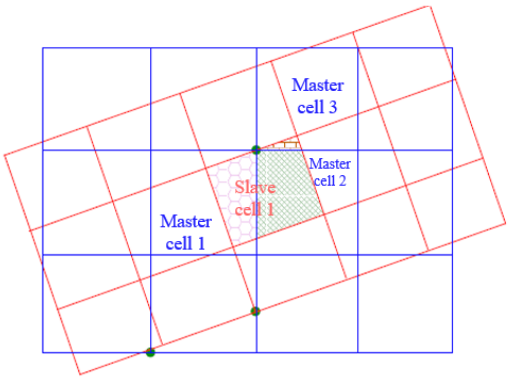


Figure 3. Diagram of overlapping area

Geometry, grid and test conditions

Geometry

In the present work, the hybrid CRP podded propulsion system is composed of a CRPs model and a streamlined pod model. The CRPs model was developed by Miller[1] in David W Taylor Naval Ship R&D Center in 1976. As Table 1 shows, DTMB3686 works as front propeller and DTMB3687A as the rear propeller. Both propellers own 4 blades and diameter of the latter is slightly smaller than the former owing to shrinking effect of wake flow field. Azipod propulsion system is composed of the rear propeller and a streamlined pod who not only stows motor but also works as rudder.

Table 1. Main parameters of CRPs

Items	DTMB3686	DTMB3687A
Diameter/mm	305.2	299.1
Blade number	4	4
$(P/D)_{0.7R}$	1.291	1.326
Disc ratio	0.303	0.324
Direction	Left	right
Airfoil	NACA66mod	NACA66mod

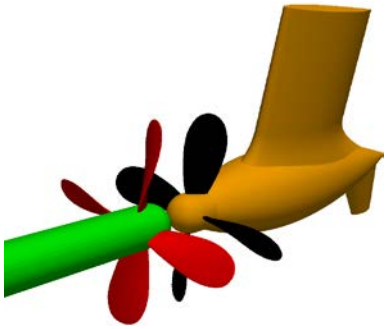


Figure 4. Hybrid CRP podded propulsion system

Grid distribution and test conditions

Grid generation is a crucial but very tough work. In present work, firstly the background grid is generated by ANSYS ICEM CFD , it is quite convenient to generate structured grid by

creating 'O block' . Secondly background grid is imported to OpenFOAM, one of its grid generation application, named snappyHexMesh will delete cells in the bodies, move boundary to the surfaces and add boundary layers sequentially. Thirdly, another application, named topoSet, will be used to create two rotational zones that surround the front and the rear propeller and sliding mesh faces will be constructed by topology technology. The final grid zone and grids are displayed in figure 5 and figure 6. Total grid number is 2.9 million where 2.6 million is gathered around propeller blades in order to capture important flow features. Y^+ is chosen to be 40 which is required to be more than 30 if wall functions is applied.

Open water performance validation of contra-rotating propellers is carried out at different advance coefficients. They are 0.7, 0.8, 0.9, 1.0, 1.1 which can be achieved by altering inflow velocity while keeping rotation rate constant (12 rps). Time step is settled to be 1.1574e-4s so that propeller will rotate with 1 degree, this small time step will improve computing accuracy of unsteady forces suffered by propeller. In order to investigate influence of axial spacing between the propeller centerlines on hydrodynamics of hybrid CRP podded propulsion system, three distances, 0.343 D_F, 0.543 D_F, 0.743 D_F are going to be taken into consideration at advance coefficient 0.9.

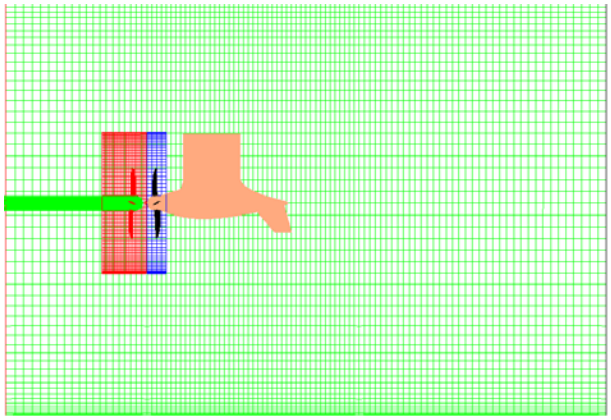


Figure 5. Grids allocation

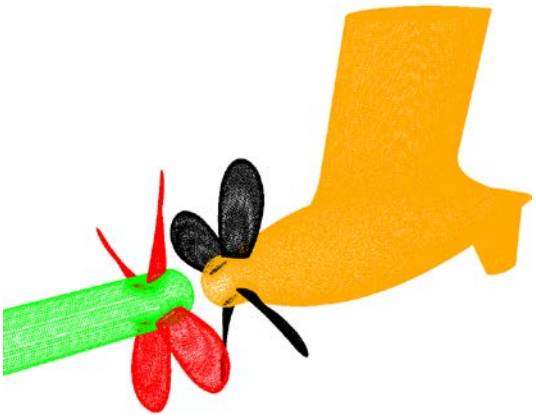


Figure 6. Grids of the model

Simulation results and analysis

Hybrid CRP podded propulsion system is composed of a CRPs and a streamlined pod, model test of this CRPs was carried out by Miller (1976)[13] at David W Taylor Naval Ship R&D Center, open water performance validation of CRPs will help verify whether numerical methods and numerical model in present work is correct. Then open water performance of hybrid CRP podded propulsion system is analyzed in detail including interactions between front propeller and Azipod propulsion system, influence of axial spacing between the propeller centerlines.

Open water performance validation of contra-rotating propellers

Some important hydrodynamic coefficients should be defined to measure open water performance of hybrid CRP podded propulsion system, they are listed as follow:

$$J = \frac{U_0}{nD_F} \tag{3}$$

$$K_T = \frac{T_F + T_P}{\rho n^2 D_F^4} \tag{4}$$

$$K_Q = \frac{Q_F + Q_P}{\rho n^2 D_F^5} \tag{5}$$

$$\eta_0 = \frac{JK_T}{2\pi K_Q} \tag{6}$$

Where U_0 is inflow velocity, T_F, Q_F, T_P, Q_P are thrust coefficient and torque coefficient of the front propeller and Azipod propulsion system respectively, D_F, n are diameter and rotation rate of front propeller.

Figure 7 shows open water performance of CRPs, predicting results agree well with their experimental counterparts, errors of thrust coefficient and torque coefficient are about 2%, 3.6% respectively which are slightly higher than experimental data, errors of efficiency is -1.6% which is slightly lower than experimental data. In general, numerical model and algorithms in present work are reliable and credible. Now, a streamlined pod will be added to CRPs as hybrid CRP podded propulsion system.

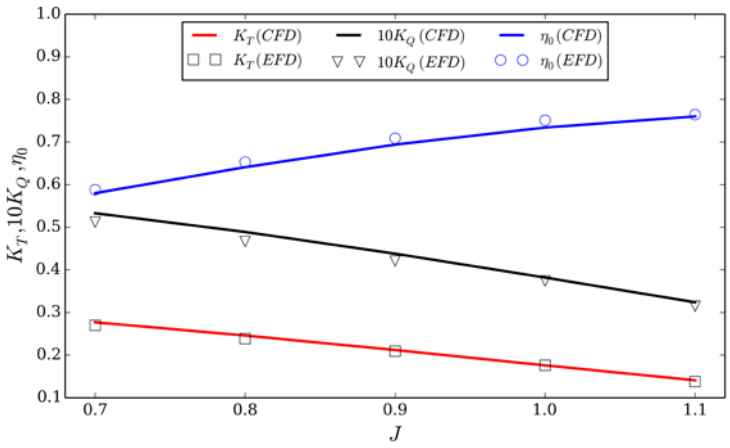


Figure 7. Open water performance of CRPs

Open water performance of hybrid CRP podded propulsion system

Interactions between front propeller and Azipod propulsion system are critical issues of hybrid CRP podded propulsion system, in present work, computing results of this propulsion system are compared with computing open water results of single propeller and Azipod propulsion system. Figure 8 shows that the Azipod propulsion system has little effect on the front propeller. As can be seen in figure 10(a) and figure 10(b), owing to the reason that blockage effect induced by the rear pod counteracts suction effect induced by the rear propeller, flow fields around single propeller and the front propeller show no obvious difference. However Figure 9 shows that hydrodynamic coefficients of the Azipod propulsion system are greatly affected by the front propeller. As can be seen in Figure 10(b) and Figure 10(c), inflow velocity before Azipod propeller has been dramatically boosted by the front

propeller, thrust and torque of the Azipod propulsion system are reduced obviously typically at high advance coefficients.

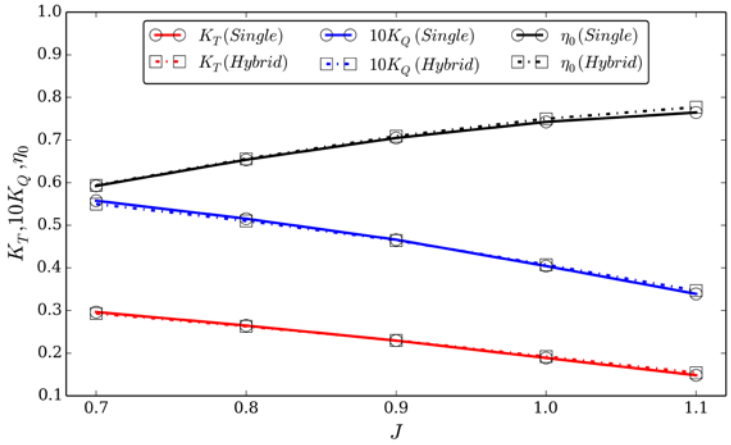


Figure 8. Effect of the podded propeller on the front

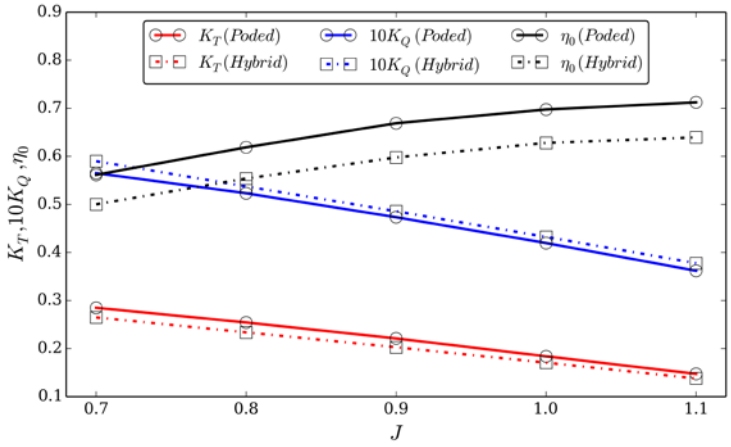
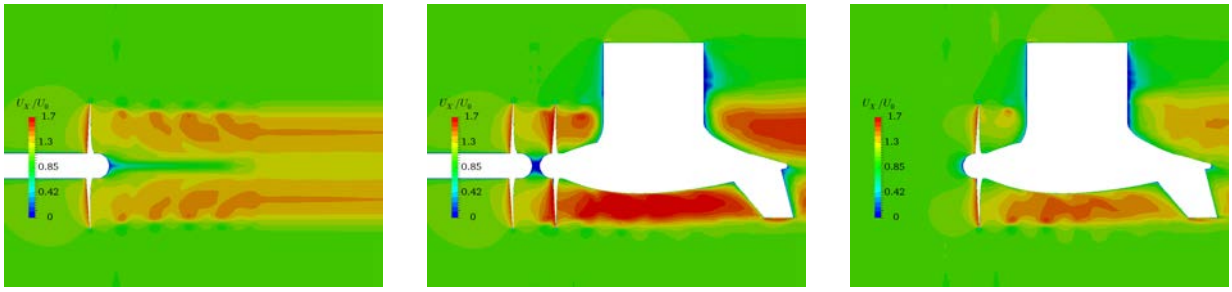


Figure 9. Effect of the front propeller on the podded propeller



(a) Single propeller (b) Hybrid CRP podded propulsion system (c) Azipod propulsion system

Figure 10. Axial velocity distribution

Figure 11 shows that tangential velocity can be obviously utilized by the rear propeller which means hybrid CRP podded propulsion system could obtain better energy-saving effect than single propeller. Meanwhile, Figure 12 shows that the rear propeller will intensify magnitude of axial velocity. At last, rudder performance will be improved with smaller tangential velocity but higher axial velocity, on the other hand, reverse rotating direction of the front and rear propeller lead to minimum unbalanced torque suffered by this hybrid system, it will improve curse-keep performance of torpedo.

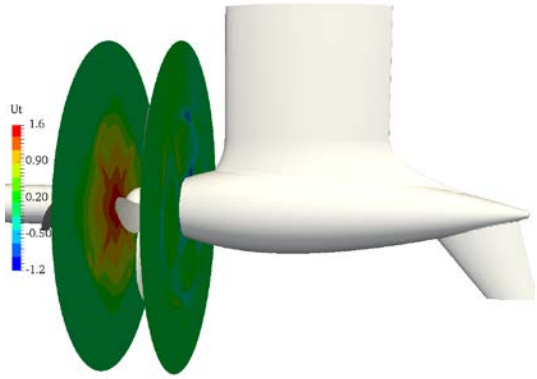


Figure 11. Tangential velocity distribution

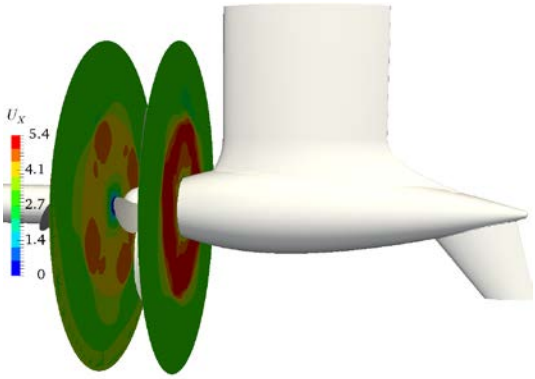
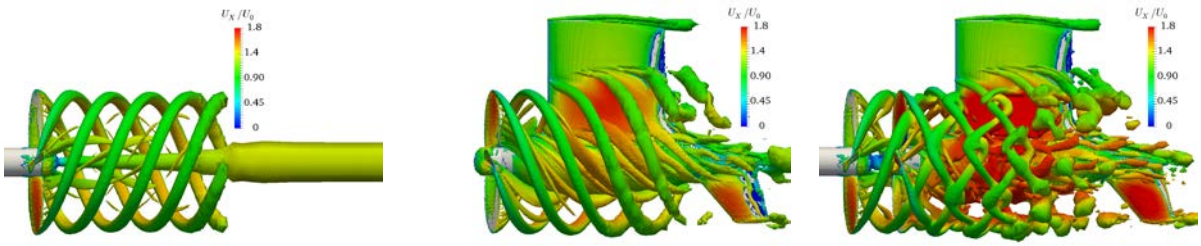


Figure 12. Axial velocity distribution

Figure 13 displays vortex structure of those three types of propulsion when Q is equal to 200. All of them are colored by U_X/U_0 where U_X , U_0 are axial velocity and inflow velocity respectively. Both vortex structures of Azipod propulsion system and hybrid CRP podded propulsion system will climb up when they encounter the rear pod , it is induced by the blockage effect of pod who functions as a rudder. Meanwhile axial velocity will be boosted around convex surface of pod. After comparison of figure 13(b) and figure 13(c), it's easy to find that vortex structure of hybrid CRP podded owes small tangential velocity in wake flow field, in other words, the rear propeller recovers swirl energy induced by the front propeller.



(a) Single propeller **(b) Azipod propulsion system** **(c) Hybrid CRP podded propulsion system**

Figure 13. Vortex structure distribution

In the period of designing hybrid CRP podded propulsion system, much more attention should be paid to unsteady forces, because large amplitude of fluctuation may induce ship vibration and serious noise. Figure 14 shows time history of unsteady forces suffered by front propeller and Azipod propulsion system, where 'Front', 'Azipod' mean the front propeller and Azipod propulsion system of hybrid CRP podded propulsion system respectively. Owing to the interaction between front propeller and the Azipod propulsion system, both K_T and $10K_Q$ have 8 small periods when propeller rotate 180 degrees which is related to blade numbers and blade number ratio, on the other hand, fluctuating amplitude is very small when axial spacing between the propeller centerlines is $0.343D_F$. it means that hybrid CRP podded propulsion system will not suffer obvious unsteady forces under a certain spacing where Azipod propulsion system could rotate normally.

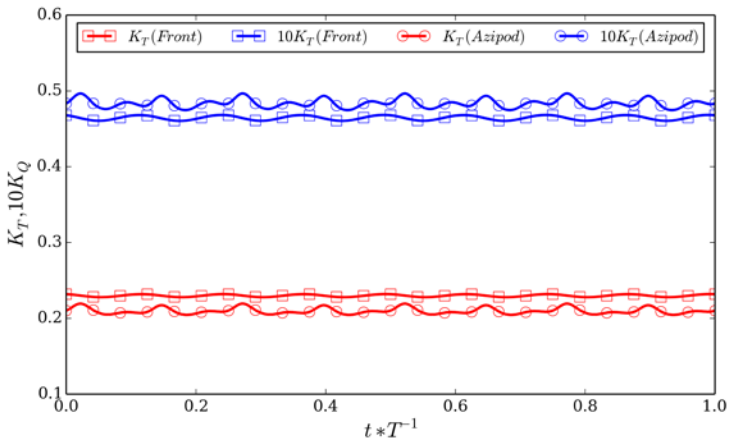


Figure 14. Unsteady forces suffered by front propeller and Azipod system

In table 2, three types of propulsion are compared under the condition that they could approximately produce equal thrust. Efficiency of CRPs is 10% higher than single propeller while efficiency of hybrid CRP podded propulsion system is 3.5% higher than single propeller for the reason that the rear pod unit will produce drag force. In future work, pod unit shape should be optimized to improve propulsion efficiency.

Table 2. Comparison of efficiency among three types of propulsion

Propulsion manner	T(N)	Q(N·M)	η_0
Single propeller	529.12	30.31	0.630
CRPs	529.91	33.41	0.694
Hybrid CRP podded propulsion system	540.00	36.24	0.652

Effects of spacing between propeller centerlines

The question of whether and how the spacing between propeller centerlines affects hydrodynamics of hybrid CRP podded propulsion system is also investigated in detail. Figure 15 shows that with the spacing increasing, K_T, $10K_Q$ will decrease to some extent while η_0 almost keeps constant. Figure 16 displays axial velocity distribution at different spacing of propeller centerlines, where U_0, U_X are inflow velocity and axial velocity in flow field respectively. It could be found that with the spacing becoming larger, inflow velocity near leading edge before the rear propeller will become larger to some extent typically. As Wang et al[14]. (2016) pointed out, the smaller spacing, the better energy-saving effect can be obtained.

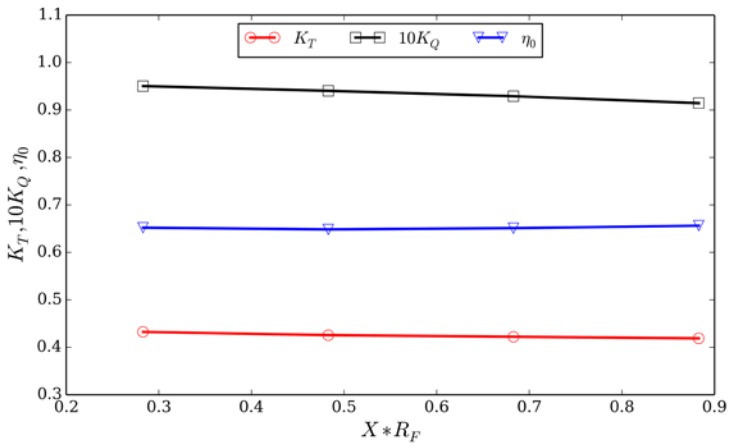
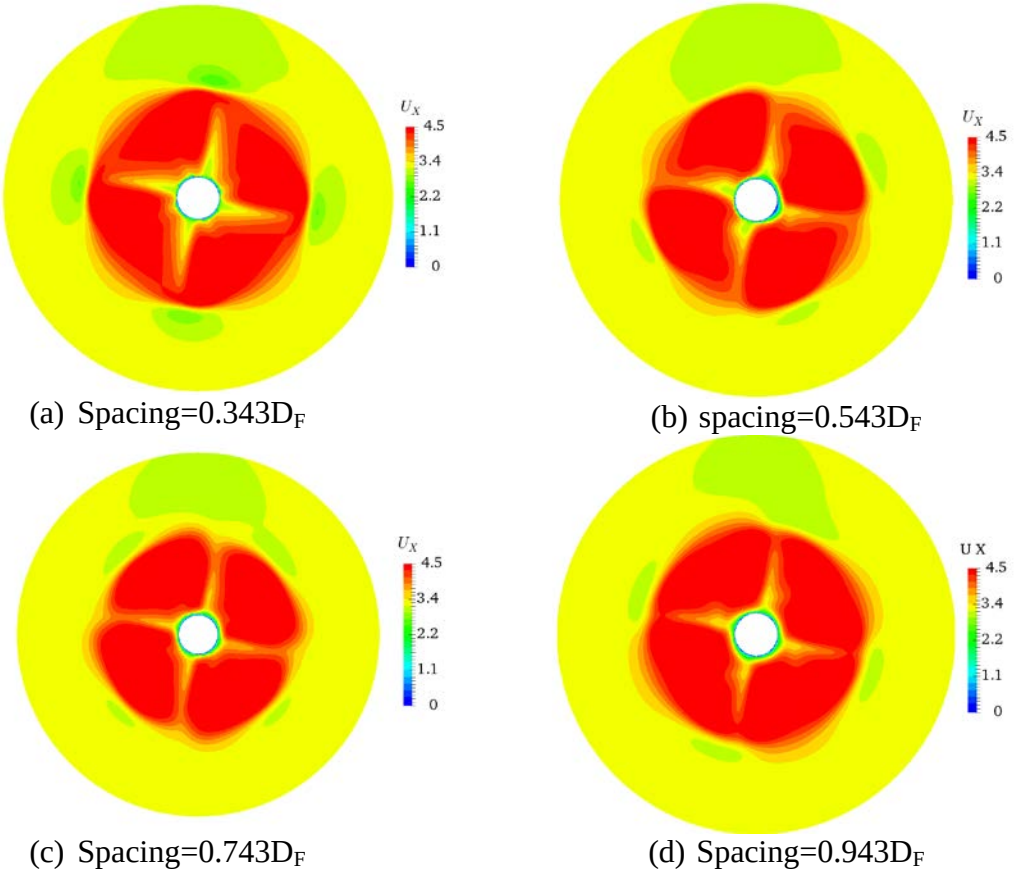


Figure 15. hydrodynamic coefficients at different spacing



(a) Spacing=0.343D_F (b) spacing=0.543D_F

(c) Spacing=0.743D_F (d) Spacing=0.943D_F

Figure 16. Axial velocity distribution before rear propeller at different spacing of propeller centerlines

Conclusions

This paper investigates open water performance of hybrid CRP podded propulsion system by CFD software, OpenFOAM. Based on analysis of hydrodynamics and wake flow field information, some useful conclusions can be draw as follow:

1) Predicting results of open water performance of CRPs agree well with their experimental counterparts, it proves that numerical methods and numerical models in present work are reliable and creditable.
2) The Azipod propulsion system has little influence on the front propeller, because suction effect of the rear propeller is counteracted by blockage effect of the rear pod, while Azipod propulsion system is greatly affected by the front propeller, thrust and torque will decrease because its inflow velocity has been boosted by front propeller by a large scale.
3) Magnitude of tangential velocity in wake flow field is reduced dramatically which means hybrid CRP podded propulsion system could obtain better energy-saving performance than single propeller.
4) Efficiency of hybrid CRP podded propulsion system is 3.5% higher than single propeller, the rear pod shape, which will produce drag force, should be optimized in future work.
5) With the spacing between propeller centerlines increasing, K_T, $10K_Q$ of the front propeller almost keep constant, while for the Azipod propulsion system, they will be reduced gradually, however η_0 of the hybrid CRP podded propulsion system changes slightly.

Future wok will focus on optimization of pod shape and investigation of power ratio influence on hydrodynamics of hybrid CRP podded propulsion system.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Pakaste, R. (1999) Experience with Azipod propulsion system on boat marine vessels, *ABB Review*, **2**, 12-18.
[2] Levander O. (2002) Advanced machinery with CRP propulsion for fast RoPax vessels, *Proceedings of the 24th Motor ship Marine Propulsion Conference*, 132-143,
[3] Kim, S. E., Choi, S. H. (2002) Model tests on propulsion systems for ultra large container vessel, *Proceedings of the Twelfth International Offshore and Polar Engineering Conference*, Kitakyushu, Japan, 520-524.
[4] Sheng, L., Xiong, Y. (2012) Numerical simulation and experimental investigation on hydrodynamics performance of hybrid CRP podded propulsion, *Journal of Nanjing University of Aeronautics and Astronautics*, **44**(2), 184-189.
[5] Yang, C., Tamashima, M., Wang, G. (1991), Prediction of the steady performance of contra-rotating propellers by lifting surface theory, *Transactions of the West-Japan Society of Naval Architects*, **82**, 17-31.
[6] Yang, C., Tamashima, M., Wang, G. (1992), Prediction of the unsteady performance of contra-rotating propellers by lifting surface theory, *Transactions of the West-Japan Society of Naval Architects*, **83**, 47-65.

[7] Noack, R. Boger, D., Kunz, R. Carrica, P.M. (2009), Suggar++: An improved general overset grid assembly capability, *Proc 47th AIAA Aerosp Sci Exhib*, 22-25.
[8] Shen, Z., Wan, D. and Carrica, P.M. (2015), Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering, *Ocean Eng*, **108**, 287-306.
[9] Wu, J., Yin, C., and Wan, D. (2016), Numerical prediction of open-water performance of propeller based on three methods, *Chin J Hydrodyn*, **31**(2), 177-187.
[10] Zhou, H., Wan, D. (2014), Numerical simulation of the unsteady flow around wind turbines with different blades numbers, *J Hydrodyn Ser A*, **29**(4): 444-453.
[11] Wang, Z., Xiong, . and Qi, W. (2012), Numerical prediction of open-water performance of contra-rotating propellers, *J Huazhong Univ of Sci & Tech*, **40**(11), 77-88.
[12] Menter, F. R., Kunz, M., Langtry, R. (2003) Ten years of industrious experiences with SST turbulence model, *Turbulence, Heat and Mass Transfer*, **4**, 625-632.
[13] Miller, M. (1976) Experimental determination of unsteady forces on contra-rotating propellers in uniform flow, Technical Report No. AD-A032337, David W Taylor Naval Ship R&D Center.
[14] Wang, Z. Z., Xiong, Y., Wang, R. (2016) Effect of the main design parameters on the open water performance of hybrid CRP podded propulsion system, *Journal of Harbin Engineering University*, **37**(1), 98-103.

GPU Acceleration of MPS for Three-Dimensional Sloshing

Xiang Chen[*], Youlin Zhang, Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: just_chenxiang@163.com

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

In this paper, the application of a GPU-based particle method to three-dimensional sloshing problem is presented. Moving particle semi-implicit (MPS) method is a Lagrangian method which can be used to simulate nonlinear flow effectively. But one of its drawbacks is the high computation cost with the increase of particle number. Based on modified MPS, the MPS-GPU-SJTU solver is developed to simulate a large sum of particles by using GPU which supports large-scale scientific computations. In addition, one optimization strategy is applied to reduce the storage and computation cost of Poisson equation of pressure (PPE). Then the convergent validation is carried out to verify the accuracy of present solver. And the accuracy and performance of GPU-based solver are investigated by comparing the results with those by CPU. As a summary of results, the GPU-based solver shows a good agreement with CPU solver (MLParticle-SJTU). And the computation efficiency of GPU is much higher than CPU.

Keywords: Moving particle semi-implicit (MPS); GPU acceleration; MPS-GPU-SJTU solver; sloshing

Introduction

With the development of economy, the demand of energy is increasing. Many countries which lack energy import liquefied natural gas, oil and liquefied petroleum gas by the transportation of vessels such as LNG, LPG, VLCC and so on. Because of the variable sea conditions, the liquid in a partially filled tank will be prone to complicated and nonlinear sloshing phenomenon. The instantaneous impact pressure induced by liquid sloshing may destroy the structure of tank walls and the stability of the ship. Therefore, many researchers have devoted themselves to investigating the characters and mechanisms of sloshing.

Faltisnen (1978) developed the boundary element method (BEM) to study sloshing in a two-dimensional (2-D) rectangular tank under translational excitation [1]. Nakayama and Washizu (1980) used the finite element method (FEM) to study 2-D problem of nonlinear liquid sloshing in a rectangular tank under pitch excitation [2]. Arai et al. (1992) used the Marker-and-cell (MAC) method to simulate numerically a 3-D sloshing phenomenon in liquid tank with vertical baffle [3]. Koh et al. (1998) developed a coupled BEM-FEM method for analyzing 3-D liquid sloshing in rectangular tanks [4]. Kim (2001) studied 2-D and 3-D sloshing in a rectangular tank with a large vertical baffle and three horizontal baffles by using the finite difference method (FDM) [5]. Liu and Lin (2009) modeled the vertical baffle in 2-D and 3-D tank by the virtual boundary force (VBF) method [6]. Yin et al. (2012) employed the volume of fluid (VOF) method by utilizing their inhouse solver naoe-FOAM-SJTU to study the effect of filling rate on sloshing [7].

In addition to the above methods, the fully Lagrangian particle methods are used to research the nonlinear sloshing problem. Delorme et al. (2009) applied the smoothed particle hydrodynamics (SPH) method to research the impact pressure of sloshing in shallow filled tank under roll excitation [8]. Shao et al. (2012) simulated 2-D liquid sloshing by using SPH [9]. Koh et al. (2013) used the consistent particle method (CPM) to investigate the effect of a constrained floating baffle in prismatic tank [10]. Kim et al. (2014) developed MPS method to research multiliquid-layer sloshing problems and investigated the elevation of interface under different excited frequencies and amplitudes [11]. Yang et al. (2015) studied the effects of excitation period on 2-D liquid sloshing by MPS Method [12]. Hashimoto et al. (2016) estimated oil overflow from an oil storage tank subjected to a possible Nankai trough earthquake in Osaka bay area based explicit MPS method [13]. Zhou et al. (2016) modified the pressure correction algorithm and simulated the sloshing of multiphase flows with large density ratio [14]. Chen et al. (2016) used modified MPS method to study the effect of the location of horizontal baffle on liquid sloshing in 3-D tank [15].

Though mesh-free methods can easily track free surface and effectively simulate sloshing problem, they still suffer from high computation cost with the increase of particle number. The GPU (Graphics Processing Unit) is a multi-processor designed to optimize for the execution of massive number of threads. Because of the explicit algorithm, it is easy to apply GPU technology to SPH. Harada et al. (2007) applied the acceleration technique of GPU to SPH and the computation speed of GPU is up to 28 times faster than CPU [16]. Hérault et al. (2010) used CUDA to implement SPH on GPU to simulate the problems of dam break and paddle-generated waves [17]. Crespo et al. (2011) developed a CPU-GPU solver DualSPHysics to deal with free-surface flow problems and achieved a speedup of 64 in the simulation of 3-D dam break by using one million particles [18]. Domínguez et al. (2013) improved DualSPHysics by using several optimizations for the GPU implementations and accelerated serial SPH codes with a speedup of 56.2 [19]. Fourtakas and Rogers (2016) applied a two-phase model based on SPH to simulate the problem of two-phase liquid-sediments flows [20]. However, it is difficult to implement a GPU-based MPS calculation due to the semi-implicit algorithm adopted to obtain the pressure field. Hori et al. (2011) developed a GPU-accelerated MPS code by using CUDA language and simulated 2-D elliptical drop evolution and dam break [21]. Kakuda et al. (2012, 2013) presented a GPU-based MPS to calculate 2-D and 3-D dam break problems compared with CPU simulations and the speed-up is 12 times and 17 times respectively [22]. Li et al. (2015) applied GPU acceleration in the solution of PPE and search of neighboring particles and achieved speedups of 10 and 6 respectively [23]. Gou et al. (2016) simulated the isothermal multi-phase fuel-coolant interaction by using MPS method with GPU acceleration [24].

In this study, the MPS-GPU-SJTU solver based on modified MPS method is employed for numerical simulation of 3-D liquid sloshing by using GPU acceleration. In the first section, a brief description of MPS method is presented. In the second section, the flow chart of implementations on GPU follows. Then one optimization strategy to reduce the storage and computation time of PPE is applied and improves the computation efficiency. In addition, the convergent validation is carried out to verify the accuracy of present solver. And the 3-D sloshing problems are simulated by GPU and CPU solvers at the same time. It is shown that the results of GPU solver show a good agreement with CPU and a large amount of computation time is reduced by GPU.

MPS Method

Koshizuka and Oka (1996) have explained the MPS method in detail [25]. In this section, the numerical models adopted in this paper are introduced briefly.

Governing Equations

The governing equations for incompressible and viscous fluid include conservation equations of mass and momentum.

$$\frac{1}{\rho}\frac{D\rho}{Dt} = \nabla \cdot \vec{V} = 0 \tag{1}$$

$$\frac{D\vec{V}}{Dt} = -\frac{1}{\rho}\nabla P + \nu\nabla^2\vec{V} + \vec{g} \tag{2}$$

where ρ is the fluid density, t is the time, $\vec{v}$ is the velocity vector, P is the pressure, ν is the kinematic viscosity and $\vec{g}$ is the gravitational acceleration vector.

Particle Interaction Models

Kernel Function

A kernel function is used for all interaction models to describe the particle interaction in MPS method.

$$W(r) = \begin{cases} \dfrac{r_e}{0.85r + 0.15r_e} - 1 & 0 \le r < r_e \\ 0 & r_e \le r \end{cases} \tag{3}$$

where r is the distance between two particles and r_e is the radius of the particle interaction. The particle number density and gradient model is $r_e=2.1l_0$, while $r_e=4.0l_0$ is used for the Laplacian model, where l_0 is the initial distance between two adjacent particles.

Gradient Model

The gradient operator is modeled as a local weighted average of the gradient vectors between particle i and its neighboring particle j.

$$<\nabla P>_i = \frac{D}{n^0}\sum_{j\neq i}\frac{P_j + P_i}{|\vec{r}_j - \vec{r}_i|^2}(\vec{r}_j - \vec{r}_i)\cdot W(|\vec{r}_j - \vec{r}_i|) \tag{4}$$

where D is the number of space dimension, n^0 is the initial particle number density and $\vec{r}$ is coordinate vector of fluid particle.

Laplacian Model

The Laplacian operator is modeled by weighted average of the distribution of a quantity ϕ from particle i to its neighboring particle j.

$$<\nabla^2\phi>_i = \frac{2D}{n^0\lambda}\sum_{j\neq i}(\phi_j - \phi_i)\cdot W(|\vec{r}_j - \vec{r}_i|) \tag{5}$$

$$\lambda = \frac{\sum_{j\neq i} W(|\vec{r}_j - \vec{r}_i|)\cdot|\vec{r}_j - \vec{r}_i|^2}{\sum_{j\neq i} W(|\vec{r}_j - \vec{r}_i|)} \tag{6}$$

where λ is applied to make sure that the increase of variance is equal to the analytical solution.

Model of Incompressibility

In this paper, PPE is solved by using a mixed source term method which is developed by Lee et al. (2011) [26].

$$< \nabla^2 P^{k+1} >_i = (1-\gamma)\frac{\rho}{\Delta t}\nabla \cdot \vec{V}_i^* - \gamma \frac{\rho}{\Delta t^2}\frac{< n^* >_i - n^0}{n^0} \tag{7}$$

where γ is a blending parameter which varies from 0 to 1, n^* is the temporal particle number density and Δt is the time step. In this paper, γ =0.01 is employed for all numerical simulations.

Free Surface Detection

Zhang (2012) developed a modified surface particle detection method, which is based on the asymmetry arrangement of neighboring particles [27].

$$< \vec{F} >_i = \frac{D}{n^0}\sum_{j \neq i}\frac{1}{|\vec{r}_i - \vec{r}_j|}(\vec{r}_i - \vec{r}_j)W(r_{ij}) \tag{8}$$

$$< |\vec{F}| >_i \; > \; \alpha \tag{9}$$

$$\alpha = 0.9\left|\vec{F}\right|^0 \tag{10}$$

where $\vec{F}$ is a vector which represents the asymmetry of arrangements of neighbor particles, $\left|\vec{F}\right|^0$ is the initial value of $|\vec{F}|$.

Boundary Condition

For MPS, multilayer particles are used to present the wall boundary. The wall particles are arranged at the boundary and the pressures of them are solved by PPE. Two layers of ghost particles are configured to fulfill the particle number density near the boundary so that the particle interaction can be properly simulated near the boundary. The pressure of ghost particle is obtained by interpolation.

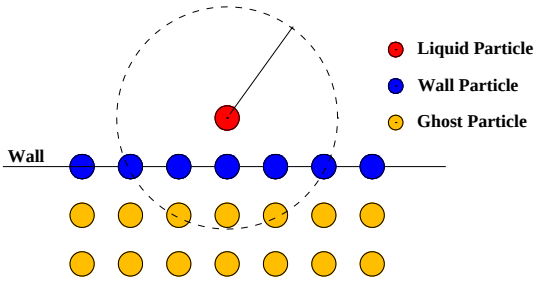


Figure 1. Schematic of boundary particles

Implementations on GPU

It is very important to reduce the CPU cost of MPS with the high-efficient parallelization. As shown in Figure 2, GPU is designed to possess more calculation threads to process data

simultaneously. For example, when one thousand particles are simulated, one thousand threads are launched to calculate. The function of CPU is only to conduct GPU codes to implement and communicate data between GPU. Therefore, GPU is a better choice for high parallel MPS method.

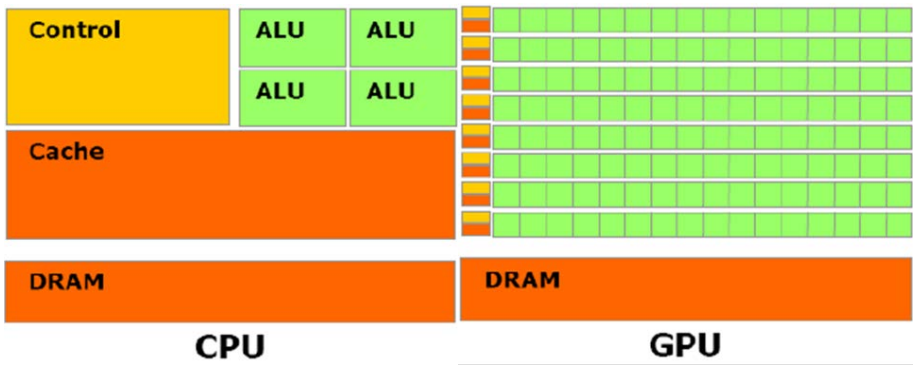


Figure 2. The frameworks of CPU and GPU

For MPS, the time integration is mainly composed of two steps. One step is an explicit calculation considering the gravity and viscosity terms. Another step corresponds to an implicit calculation of PPE. The computational flow chart of MPS on GPU is shown in Figure 3. Except the exchange of data between GPU and CPU, the GPU implementation mainly consists of six steps:

1) Search neighboring particles and create neighbor list
2) Explicit calculation (gravity and viscosity terms)
3) Calculate particle number density and detect free surface
4) Solve PPE to obtain pressure field
5) Calculate pressure gradient
6) Update velocities and positions of particles

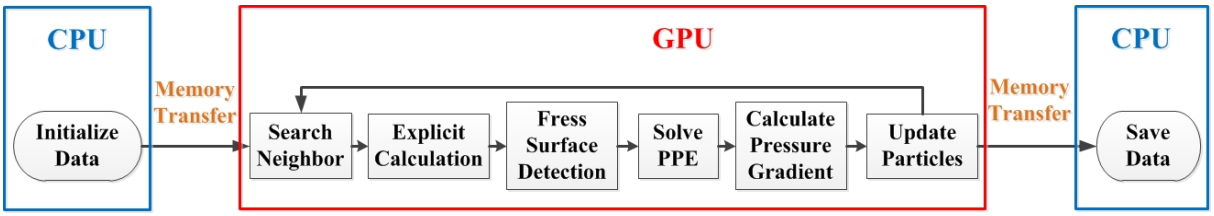


Figure 3. The flow chart of MPS on GPU

Simulation and Result

In this section, a 3-D liquid tank same as the experimental model given by Kim (2001) is selected as the numerical model to simulate. The geometry of liquid tank is shown in Figure 4. The dimensions of tank are 0.8 m (L), 0.35 m (B) and 0.5 m (H), respectively. The depth of water (D) is 0.25 m, corresponding filling level is 50%. The liquid tank is subject to move by the external surge excitation:

$$x = A \cdot \sin(\omega \cdot t) \tag{11}$$

where A is the amplitude of excitation set to be 0.02 m and the excitation frequency ω is 5.39 rad/s which is same as the first natural period of fluid motion in the tank. In addition, two pressure probes are arranged on lateral wall to measure the variation of pressure. The arrangements of pressure probes are listed in Table 1.

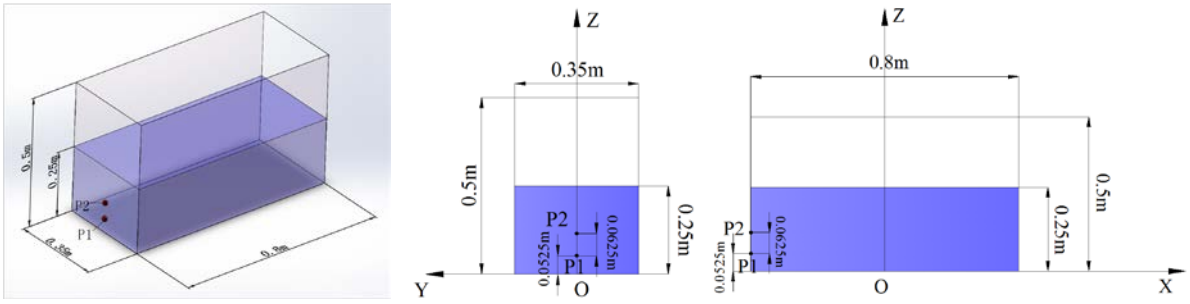


Figure 4. The sketch of numerical model

Table 1. Arrangements of pressure probe

	X/m	Y/m	Z/m
P1	-0.4	0	0.0525
P2	-0.4	0	0.115

All simulations are performed on parallel high performance computing (HPC) with multi cores of Intel(R) Xeon(R) E5-2680 v2, 2.80 GHz. And the GPU device is NVIDIA graphics card Tesla K40M, which has 2880 CUDA cores with 12GB graphics memory. Table 2 shows the parameters of computing devices. In this paper, the double precision floating point computation is only used in both CPU and GPU codes.

Table 2. Computational environment of CPU and CPU

	HPC	GPU
Card	Intel(R) Xeon(R) E5-2680 v2, 2.80 GHz	Tesla K40M
Memory	DDR3 1600, 16GB	12GB
Core	10	2880
Compiler	gcc, MVAPICH	CUDA 7.0, CULA Sparse S6

PPE Optimization

It is well known that the most computation time of MPS is consumed to solve PPE which can be discretized into a linear system **Ax**=**b**. Eq. 12 shows the discretization of linear system. The coefficient matrix of PPE, **A** is a typical sparse matrix. In order to reduce the storage of matrix **A**, the compressed sparse row (CSR) data format is employed. Moreover, the Biconjugate gradient stabilized (BiCGSTAB) method is applied to solve this linear system based on CSR format. Nevertheless, the pressure of ghost particles and free surface particles is no need to solve in PPE and set to zero as the boundary condition. For example, if particle *i* is ghost particle or surface particle, the relevant row of matrix A and array B will be set to constant in Eq. 13. Here, Thrust is a C++ template library for CUDA which provides a rich collection of data parallel primitives such as scan, sort, and reduce [28]. In order to reduce the iteration time and storage of PPE, the relevant rows and columns of ghost and free surface particles are removed from the linear system by using Thrust. Therefore, the matrix dimensions of **A**, **x** and **b** can be reduced in Eq. 14 and the iteration speed of PPE will be faster. The sparse linear algebra library of CULA-Sparse is utilized for accelerating the iteration of PPE [29].

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,i-1} & a_{1,i} & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,i-1} & a_{i,i} & a_{i,i+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,i-1} & a_{i+1,i} & a_{i+1,i+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & a_{n,i} & a_{n,i+1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_{i-1} \\ b_i \\ b_{i+1} \\ \vdots \\ b_n \end{pmatrix} \tag{12}$$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,i-1} & a_{1,i} & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} & a_{i-1,i} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+1,1} & \cdots & a_{i+1,i-1} & a_{i+1,i} & a_{i+1,i+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & a_{n,i} & a_{n,i+1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_i \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_{i-1} \\ 0 \\ b_{i+1} \\ \vdots \\ b_n \end{pmatrix} \tag{13}$$

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,i-1} & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,i-1} & a_{i-1,i+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,i-1} & a_{i+1,i+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & a_{n,i+1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_{i-1} \\ x_{i+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_{i-1} \\ b_{i+1} \\ \vdots \\ b_n \end{pmatrix} \tag{14}$$

The comparison of computation time between original and optimized PPE is conducted in this sub-section. Total 750793 particles are used to model the liquid tank. The simulation model includes 144812 ghost particles and uncertain free surface particles which are filtered after free surface detection at each step. Figure 5 shows the comparison of pressure history between original and optimized PPE. The overall trends of pressure are almost same for two cases. The relative errors of average pressure peak on both P1 and P2 are small. The calculation time of every step is shown in Figure 6. For optimized PPE, the time cost of solving PPE is about 7% smaller than original PPE. With the increase of ghost and free surface particles, the optimized PPE may reduce more computation time. The concrete results are listed in Table 3.

Table 3. The results of original and optimized PPE

PPE Type	Pressure Peak of P1 (Pa)	Relative error	Pressure Peak of P2 (Pa)	Relative error	Computation time of PPE (s)
Original	3166	-	2630	-	3.101
Optimized	3180	0.44%	2639	0.34%	2.877

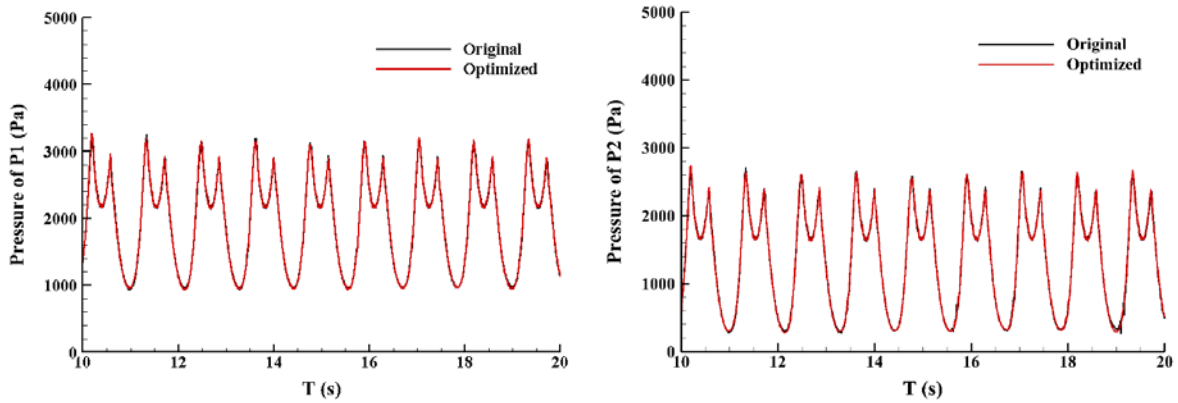


Figure 5. The pressure histories of original and optimized PPE

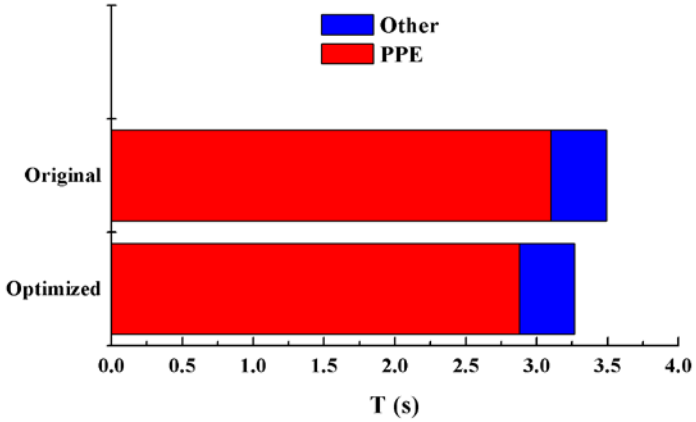


Figure 6. The computation times of original and optimized PPE

Convergence verification

In this sub-section, the convergence verification is conducted to investigate the effect of particle spacing on the numerical results. Three different spatial resolutions (0.006 m, 0.005 m, 0.004 m) are employed to check the convergence of numerical results. The pressure variations of different spatial resolutions are shown in Figure 7. The average values of pressure peak are listed to check the convergence of three cases quantitatively in Table 4. For three spatial resolutions, the variation tendency of pressure is almost same. Furthermore, the relative error is so tiny that the results are convergent with respect to the spatial resolution. In addition, the computation times of different spatial resolution are compared in Figure 8. For every step, the computation time of medium resolution is 1.888 times than coarse resolution. And every step of fine resolution is about 7.086 s which is almost twice than medium resolution. Considering the computation time and accuracy, the medium spatial resolution is selected in following sections.

Table 4. The pressure peaks of different spatial resolutions

Spatial resolution	Particle spacing (m)	Particle number	Pressure Peak of P1 (Pa)	Relative error	Pressure Peak of P2 (Pa)	Relative error
Coarse	0.006	456588	3156	-	2630	-
Medium	0.005	750793	3180	0.76%	2639	0.34%
Fine	0.004	1390732	3224	1.35%	2647	0.30%

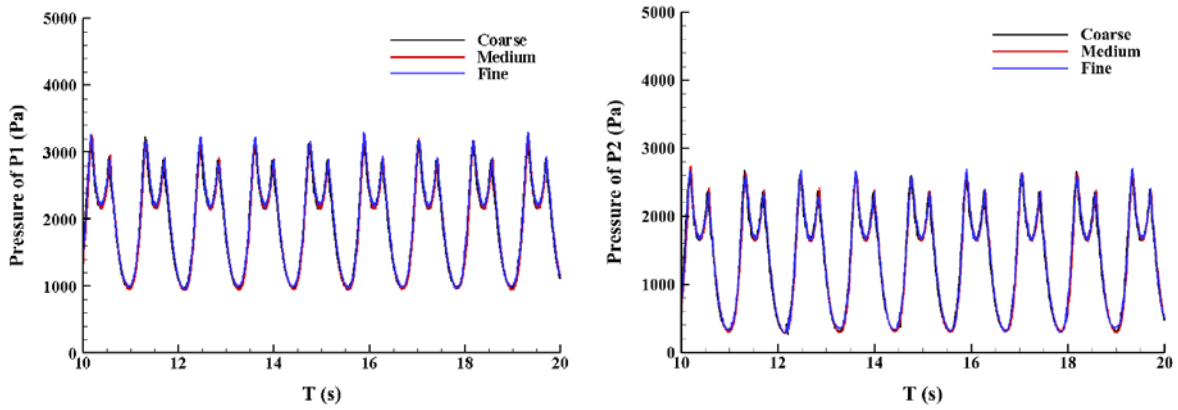


Figure 7. The pressure histories of different spatial resolutions

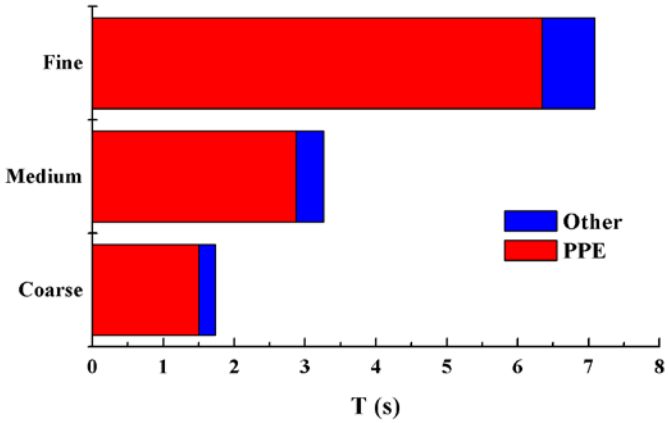
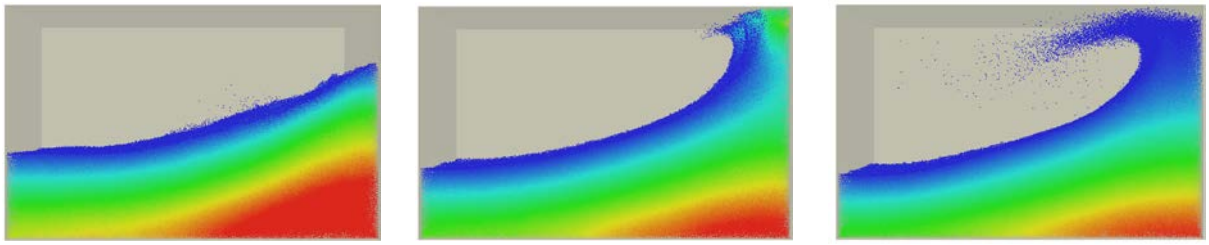


Figure 8. The computation times of different spatial resolutions

3-D Sloshing

In this sub-section, there are some comparisons between CPU and GPU including simulation results, computation time and so on. Figure 9 shows some snapshots of numerical flow field. The nonlinear deformation and large fragmentation of free surface can be observed. The numerical flow field of GPU is in good agreement with CPU simulation. Many details of flow field such as overturning wave, local splashing and travelling wave are clearly captured. The flow of fluid is opposite to the movement of liquid tank. The tank moves to the left lateral wall while fluid flows to right. Therefore, the fluid runs up along lateral wall and impacts the ceiling of the tank. Then fluid spreads along the roof and drops down under the action of gravity. In this process, a part of water even splashes on the right side wall. Then the sloshing wave travels to the left side wall.

CPU 10cores



$t \approx 15.35$ s $t \approx 15.45$ s $t \approx 15.55$ s

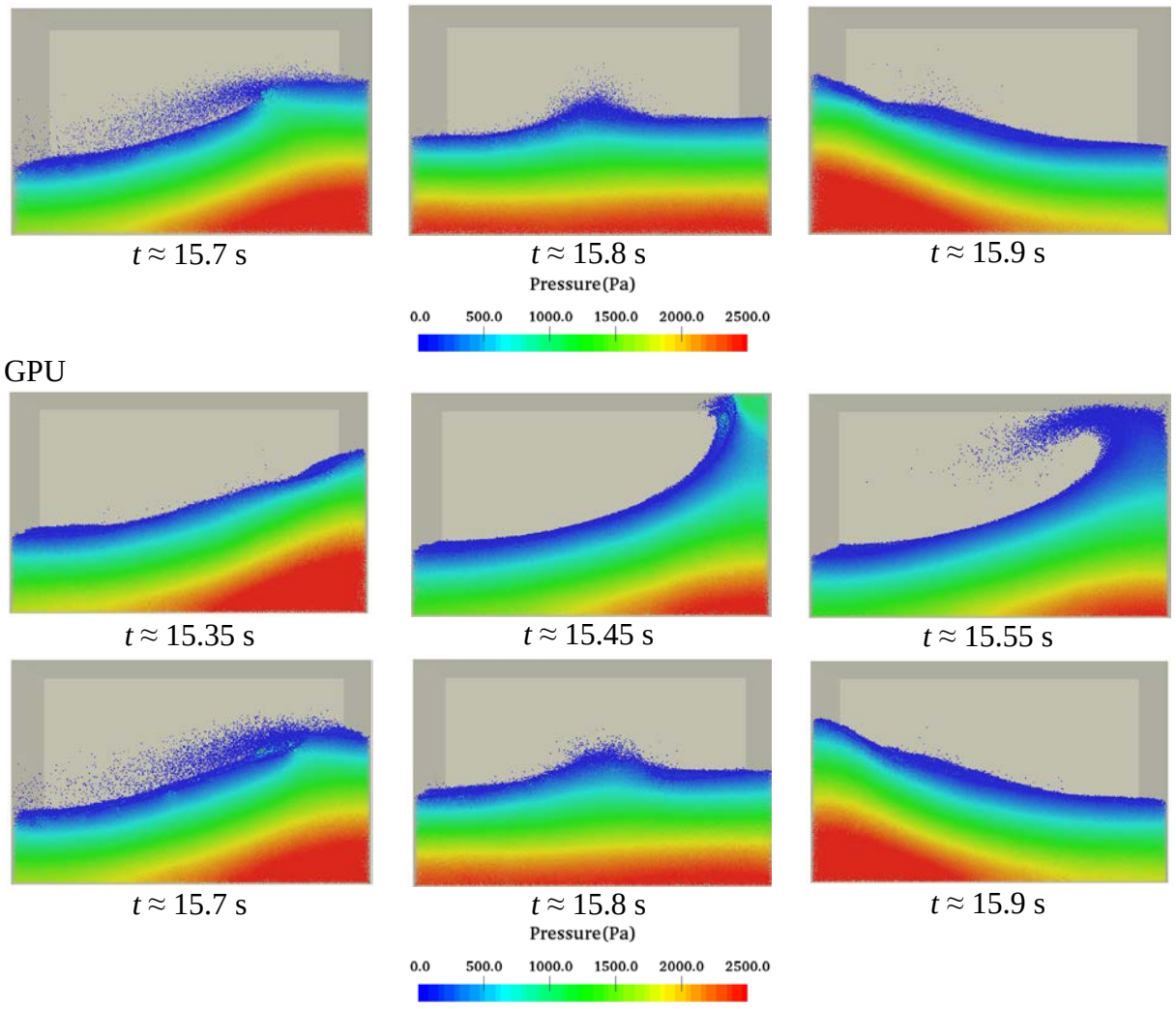


Figure 9. The flow fields of CPU and GPU

In addition, the numerical pressure histories of CPU and GPU are shown in Figure 10. The pressures of GPU simulation on two probes show a good congruency with CPU results. Two successive pressure peaks in each period can be observed. Because of the phase difference between fluid and tank, the sloshing wave that impacts on lateral wall induces the first pressure peak. Then the pressure decreases when fluid runs up along lateral wall. And the second pressure peak results from the fallen water which spreads along the roof and drops down on the free surface. Figure 11 shows the spectrum analysis of GPU results. When frequency is 0 Hz, the spectrum amplitudes of P1 and P2 are similar to hydrostatic pressure. When the frequency is same as excitation frequency, a peak of spectrum amplitude explains the first successive pressure peak. As frequency is twice as excitation frequency, there is the second peak of spectrum amplitude corresponding to the second pressure peak in pressure history.

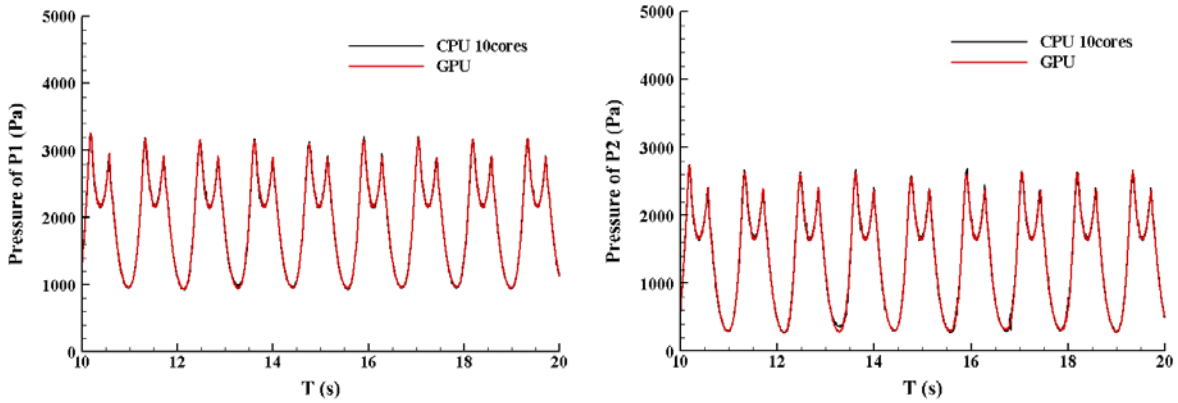


Figure 10. The pressure histories of CPU and GPU

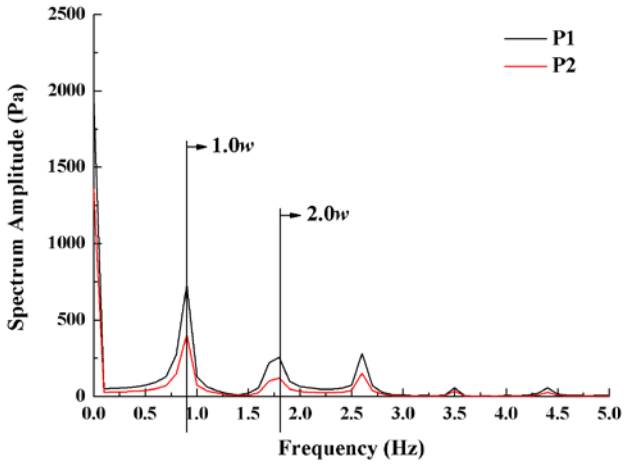


Figure 11. Spectrum of numerical pressures

In this paper, various schemes of multi cores are implemented to run CPU code on HPC. Figure 12 shows the computation times on CPU and GPU devices. The specific computation time of every step is listed in Table 5. No matter CPU and GPU, solving PPE is the most cost for calculation. From Figure 12, the total time decreases with the increase of CPU cores. Comparing GPU and CPU one core, the speedups of PPE and total time are 25.22 and 24.05, respectively. Therefore, how to solve PPE quickly is the greatest problem for researchers.

Table 5. The computation times of CPU and GPU

	CPU 1core	CPU 2cores	CPU 4cores	CPU 6cores	CPU 8cores	CPU 10cores	GPU
PPE	72.562	40.167	25.398	20.667	17.892	16.607	2.877
Other	6.011	3.416	1.872	1.592	1.404	1.424	0.390
Total	78.573	43.584	27.270	22.259	19.295	18.032	3.267

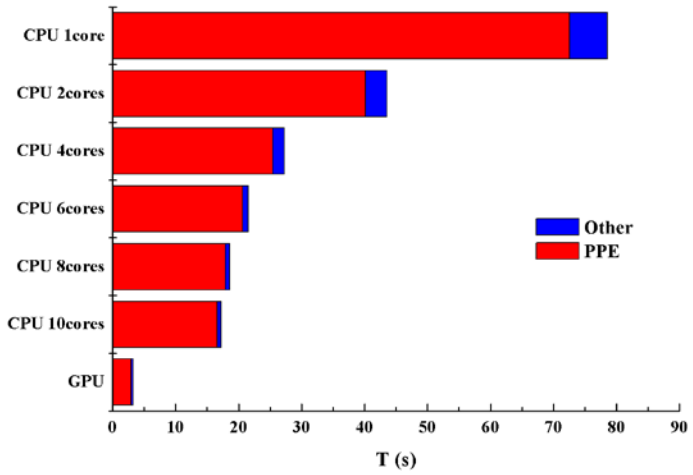


Figure 12. The computation times of CPU and GPU

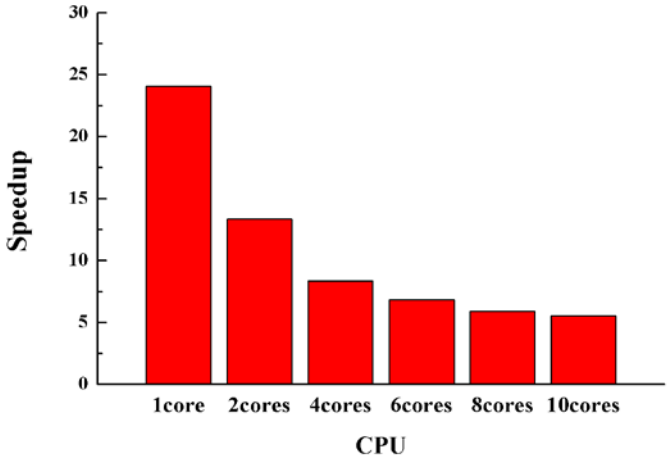


Figure 13. The speedup by GPU

Conclusions

In this paper, the MPS-GPU-SJTU solver based on modified MPS is developed to simulate 3-D sloshing problem by applying GPU acceleration technique. By reducing the dimensions of matric and arrays, the computation efficiency of solving PPE is improved. The convergence verification is conducted to prove the stability of GPU solver. In addition, the results of GPU show a good agreement with CPU. The computation time of GPU solver is much smaller than CPU and the speedup of every time iteration is up to 24.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information

Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Faltinsen, O. M. (1978) A Numerical Nonlinear Method of Sloshing in Tanks with Two Dimensional Flow, *Journal of Ship Research* **22** (3), 193-202.
[2] Nakayama, T., and Washizu, K. (1980) Nonlinear Analysis of Liquid Motion in a Container Subjected to Forced Pitching Oscillation, *International Journal for Numerical Methods in Engineering* **15** (8), 1207-1220.
[3] Arai, M., Cheng, L. Y., and Inoue, Y. (1992) 3D Numerical Simulation of Impact Load due to Liquid Cargo Sloshing, *Journal of the Society of Naval Architects of Japan* **171**, 177-184.
[4] Koh, H. M., Kim, J. K., and Park, J. H. (1998) Fluid-structure Interaction Analysis of 3-D Rectangular Tanks by a Variationally Coupled BEM-FEM and Comparison with Test Eesults, *Earthquake Engineering and Structural Dynamics* **27** (27), 109-124.
[5] Kim, Y. (2001) Numerical Simulation of Sloshing Flow with Impact Load, *Applied Ocean Research* **23**, 53-62.
[6] Liu, D. M., and Lin, P. Z. (2009) Three-dimensional Liquid Sloshing in a Tank with Baffles, *Ocean Engineering* **36**, 202-212.
[7] Yin C. H., Wu J. W., Wan D. C. (2014) *Numerical study on liquid sloshing in three-dimensional rectangular tanks with different filling rates and fixed baffle*, Proceedings of the 24th International Offshore and Polar Engineering Conference, Hawaii, United States: 381-391.
[8] Delorme, L., Colagrossi, A., Iglesias, A. S., Rodriguez, R. Z., and Vera, E. B. (2009) A Set of Canonical Problems in Sloshing, Part I: Pressure Field in Forced Roll-comparison between Experimental Results and SPH, *Ocean Engineering* **36** (2), 168-178.
[9] Shao, J. R., Li, H. Q., Liu, G. R., and Liu, M. B. (2012) An Improved SPH Method for Modeling Liquid Sloshing Dynamics, *Computers and Structures* **100-101**, 18-26.
[10] Koh, C. G., Luo, M., and Bai, W. (2013) Modelling of Liquid Sloshing with Constrained Floating Baffle, *Computers and Structures* **122**, 270-279.
[11] Kim, K. S., Kim M. H., and Park, J. C. (2014) Development of Moving Particle Simulation Method for Multiliquid-Layer Sloshing, *Mathematical Problems in Engineering* **137**(5), 282-290.
[12] Yang, Y. Q., Tang, Z. Y., Zhang, Y. L., and Wan, D. C. (2015) *Investigation of Excitation Period Effects on 2D Liquid Sloshing by MPS Method*, Proceedings of the Twenty-fifth (2015) International Ocean and Polar Engineering Conference, Hawaii, USA, 937-944.
[13] Hashimoto, H., Hata, Y., and Kawamura, K. (2016) Estimation of oil overflow due to sloshing from oil storage tanks subjected to a possible Nankai Trough earthquake in Osaka bay area, *Journal of Loss Prevention in the Process Industries*, 1-10.
[14] Zhou, L., Cai, Z. W., Zong, Z., and Chen, Z. (2016) An SPH pressure correction algorithm for multiphase flows with large density ratio, *Int. J. Numer. Meth. Fluids* **81**, 765-788.
[15] Chen, X., Zhang, Y. L., and Wan, D. C. (2016) *Effects of the location of horizontal baffle on liquid sloshing by MPS method*, Proceedings of the Second Conference of Global Chinese Scholars on Hydrodynamics, Wuxi, China, 448-454
[16] Harada, T., Koshizuka, S., and Kawaguchi, Y. (2007) Smoothed particle hydrodynamics on GPUs, *Structure* **4** (4), 671-691.
[17] Crespo, A. J. C., Dominguez, J. M., Barreiro, A., Gómez-Gesteira, M. and Rogers, B. D. (2011) GPUs, a new tool of acceleration in CFD: efficiency and reliability on smoothed particle hydrodynamics methods, *PLoS One* **6** (6): e20685.
[18] Domínguez, J. M., Crespo. A. J. C., and Gómez-Gesteira M. (2013) Optimization strategies for CPU and GPU implementations of a smoothed particle hydrodynamics method, *Computer Physics Communications*, **184**(3), 617-627.
[19] Fourtakas, G., and Rogers, B. D. (2016) Modelling multi-phase liquid-sediment scour and resuspension induced by rapid flows using Smoothed Particle Hydrodynamics (SPH) accelerated with a Graphics Processing Unit (GPU), *Advances in Water Resources* **92**, 186-199.
[20] Hori, C., Gotoh, H., Ikari, H., and Khayyer, A. (2011) GPU-acceleration for moving particle semi-implicit method, *Computers & Fluids* **51**, 174-183.
[21] Kakuda, K., Nagashima, T., Hayashi, Y., Obara, S., Toyotani, J., Katsurada, N., Higuchi, S., and Matsuda, S. (2012) Particle-based fluid flow simulations on GPGPU using CUDA, *Computer Modeling in Engineering & Sciences* **88**(1), 17-28.
[22] Kakuda, K., Nagashima, T., Hayashi, Y., Obara, S., Toyotani, J., Miura, S., Katsurada, N., Higuchi, S., and Matsuda, S. (2013) Three dimensional fluid flow simulations using GPU-based particle method, *Computer Modeling in Engineering & Sciences* **93**(5), 363-376.
[23] Li, H. Z., Zhang, Y. L., and Wan, D. C. (2015) *GPU Based Acceleration of MPS for 3D Free Surface Flows*, Proceedings of the 9th International Workshop on Ship and Marine Hydrodynamics, Glasgow, UK.

[24] Gou, W., Zhang S., and Zheng Y. (2016) Simulation of isothermal multi-phase fuel-coolant interaction using MPS method with GPU acceleration, *Kerntechnik* **81**(3), 330-336.
[25] Koshizuka, S., and Oka, Y. (1996) Moving-particle semi-implicit method for fragmentation of incompressible fluid, *Nuclear Science and Engineering* **123**(3), 421-434.
[26] Lee, B.H., Park, J.C., Kim, M.H., and Hwang, S.C. (2011). Step-by-step improvement of MPS method in simulating violent free-surface motions and impact loads, *Computer Methods in Applied Mechanics and Engineering* **200**(9-12), 1113-1125.
[27] Zhang, Y. X., and Wan, D. C. (2012) Numerical simulation of liquid sloshing in low-filling tank by MPS, *Journal of Hydrodynamics* **27**(1), 101-107.
[28] CUDA Toolkit Documentation v8.0.61. <http://docs.nvidia.com/cuda/>.
[29] CULA (GPU-Accelerated LAPACK). <http://www.culatools.com/>.

Numerical Simulation of Regular Waves onto a Vertical Circular Cylinder

Zhenghao Liu[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: xiakee@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

Wave run-up phenomenon is of great significance during the design of a fixed or floating structure. In this work, the wave run-up on a truncated surface-piercing circular cylinder is investigated using a numerical wave tank. The numerical simulations are carried out by the in-house CFD solver naoe-FOAM-SJTU which is developed on the open source platform OpenFOAM. The volume of fluid (VOF) method is applied to capture the free surface. The surface elevation around the cylinder is probed by a series of wave gauges and analyzed using the Fourier analysis. The response amplitude operates (RAOs) of surface elevation are presented and compared with experimental data. Reasonable agreement shows the present solver is capable to investigate the wave run-up on a cylinder. The local surface elevation around the cylinder, the wave force and the scattered wave field around the cylinder is also investigated in detail.

Keywords: wave run-up; truncated vertical cylinder; naoe-FOAM-SJTU solver; surface elevation

Introduction

Offshore structures such as Spars, semi-submersibles and tension leg platforms (TLP) have been widely applied in ocean engineering. Wave run-up on the columns of the structures can be relatively large and even causes green water on deck in severe environment. Generally, wave run-up height is defined as the maximum vertical wave elevation to the still water surface. Significant nonlinear wave-structure interactions can be observed during the wave run-up process. The accurate prediction of wave run-up is of great importance for the air-gap design of offshore structures. Numerous researchers have done experimental, theoretical and numerical studies on wave run-up onto piles, sloped beaches and columns of both fixed and floating offshore structures.

Various experimental investigations of wave run-up on vertical cylinders have previously been performed. Galvin and Hallermeier (1972) experimentally studied the wave run-up on a cylindrical column for the first time [1]. A series of wave gauges were mounted near the column to obtain the distribution of free surface around the column. It was found that when the waves pass through the vertical column, two important factors will affect the wave run-up effect: 1) scattering effect due to wave-structure interaction; 2) viscous dissipation effect at the column wake region. Chakrabarti and Tam (1975) conducted a series of model tests to investigate the regular waves onto a large-scale cylinder [2]. The incident wave steepness kA is 0.03~0.19 and the scattering parameter kr is 0.34~1.55, respectively. Their work was focused on the wave run-up phenomenon and the effect of incident wave angle on the distribution of the cylinder surface pressure, while the free surface around the cylinder was

not discussed. Morris-Thomas et al. (2002) investigated the effect of wave steepness and the scattering parameter on the wave run-up on a fixed cylinder in a towing tank [3]. Nielsen (2003) firstly studied the effect of cross-sectional shape on the wave run-up and free surface of cylinders [4]. Compared with circular cylinder, stronger nonlinear interaction can be found for circular-like cylinder. Systematic model tests were conducted in Shanghai Jiaotong University to investigate the effect of aspect ratio, cross-sectional shape, wave parameters and current on wave run-up phenomenon in 2010.

Besides experimental research on wave run-up, the nonlinear wave-structure interaction phenomenon has been studied by numerous researchers through theoretical methods. Generally, the theoretical methods are based on potential theory in which an idealized fluid domain is assumed. The Laplace equation is solved with applied boundary conditions to yield a velocity potential. The free-surface elevation around the column can be obtained by application of the unsteady Bernoulli equation and velocity potential at the free-surface position. In the early days, the approximate results of wave run-up on a single cylinder according to first- and second-order potential flow theory [5][6]. Trulsen and Teigen (2002) developed fully nonlinear numerical wave tanks (NWT) to investigate nonlinear interaction between wave and a truncated cylinder, and compared the predicted results with experimental data [7]. They found the effects of viscosity may take account for the discrepancies between the theoretical method and experiment. Similar researches can be found in [8]. Morris-Thomas and Thiagarajan (2004) adopted the linear diffraction theory and the commercial software WAMIT to predict the wave run-up around a cylinder [9]. They suggested that linear diffraction theory is insufficient for wave run-up estimation. This confirms conclusions made by previous authors. The second-order harmonic components predicted by WAMIT show reasonable estimates of wave run-up when the scattering parameter is small.

With the rapid development of computer technology during the last several decades, computational fluid dynamics (CFD) method has been widely applied in ocean engineering field. Previous numerical simulations based on CFD have also been performed on wave run-up problems. Based on the open source platform OpenFOAM, Cao and Wan (2014, 2015, 2017) simulated the regular and solitary waves on to a circular cylinder, and the comparison of numerical results of wave run-up and experimental data showed the reasonable agreement [10]-[12]. Donald G Danmeier et al. [13] used the ComFLOW software to simulate the wave run-up around a semi-submersible platform. Sun et al. [14] used a both potential flow solver DIFFRACT and a full CFD solver OpenFOAM to investigate nonlinear interactions between regular waves and a single truncated circular column. The predicted free surface elevation around the column and the wave forces were analyzed and compared with experimental data.

To investigate wave run-up, the ITTC committee organized several studies, including experimental and numerical researches in 2013. The results of a series of model tests for a truncated circular column in regular waves at MARINTEK and MOERI are used in the ITTC benchmark study [15]. The time histories of surface elevations and wave forces were provided.

The objective of the present work is to investigate the wave run-up on a surface-piercing circular cylinder. The computations in this paper are performed with the in-house CFD solver naoe-FOAM-SJTU, which is developed based on the open source code OpenFOAM. The numerical results of the free surface elevation around the column and the wave forces are presented and compared with experiments performed at MOERI. The details information of wave run-up around cylinder will be given and discussed. The results show that the present approach can be an alternative tool to deal with nonlinear wave-structure interactions.

Numerical methods

naoe-FOAM-SJTU solver

The CFD solver naoe-FOAM-SJTU is designed for computing viscous flows around ships and ocean structures [16-20] and mainly composed of a dynamic deforming mesh module, a 6DoF motion module, a velocity inlet wave-making module and a mooring system module.

In present study, the incompressible unsteady Reynolds averaged Navier-Stokes (URANS) equations are adopted as the governing equations, and can be written as follows:

$$\nabla \cdot \boldsymbol{U} = 0 \tag{1}$$

$$\frac{\partial \rho \boldsymbol{U}}{\partial t} + \nabla \cdot (\rho(\boldsymbol{U} - \boldsymbol{U}_g)\boldsymbol{U}) = -\nabla p_{\boldsymbol{d}} - \boldsymbol{g} \cdot x \nabla \rho + \nabla \cdot (\mu_{\boldsymbol{eff}} \nabla \boldsymbol{U}) + (\nabla \boldsymbol{U}) \cdot \nabla \mu_{\boldsymbol{eff}} + f_\sigma + f_s \tag{2}$$

where $\boldsymbol{U}$ and $\boldsymbol{U}_\mathbf{g}$ are the velocity field and the velocity of grid nodes, respectively. p_d is the dynamic pressure and p is the total pressure, ρ is the mixed density of the two phases water and air. μ_{eff} is the effective dynamic viscosity, in which ν and ν_t are kinematic viscosity and eddy viscosity, respectively. f_σ is the surface tension, which impacts the free surface. f_s is a source term, added to generate the sponge layer for wave absorbing.

Capture of Free Surface

The free surface of the two-phase flow is captured by the volume of fluid (VOF) method [21] with artificial bounded compression techniques. The method is based on a volume fraction α which can control numerical diffusion and capture the interface with high resolution. The volume fraction function can be determined by solving a transport equation:

$$\frac{\partial \alpha}{\partial t} + \nabla \cdot [(\boldsymbol{U} - \boldsymbol{U}_g)\alpha] + \nabla \cdot [\boldsymbol{U}_{\boldsymbol{r}}(1-\alpha)\alpha] = 0 \tag{3}$$

The first two terms on the left-hand side of Eq. (3) stand for traditional volume of fluid transport equation while the third term represents the artificial compression term. The velocity field $\boldsymbol{U}_{\boldsymbol{r}}$ compressing the interface is computed at cell faces by the maximum velocity magnitude at the interface region:

$$U_{r,f} = n_f \min\left\{ C_\alpha \frac{|\phi|}{|S_f|}, \max\left(\frac{|\phi|}{|S_f|} \right) \right\} \tag{4}$$

where φ is face volume flux; C_α is a compression coefficient controlling the magnitude of compression, in this paper it is chosen to be 1.0. Larger value will increase the compression of the interface, leading to larger detrimental velocity gradients around the interface. The compression term only works on the interface without affecting the numerical computation out of the transition layer due to term $(1-\alpha)\alpha$.

Different phases are marked using volume fraction α which indicates the relative proportion of fluid in each cell, it is defined as Eq. (5). For an interface cell, the value of volume fraction α is between 0 and 1, representing it contains both water and air.

$$\begin{cases} \alpha=0 & \text{air} \\ \alpha=1 & \text{water} \\ 0<\alpha<1 & \text{interface} \end{cases} \tag{5}$$

In physical domain, the density of fluid ρ and the dynamic viscosity μ can be obtained by a weighted value based on the volume fraction α:

$$\rho = \alpha\rho_1 + (1-\alpha)\rho_2 \tag{6}$$

$$\mu = \alpha\mu_1 + (1-\alpha)\mu_2 \tag{7}$$

Where ρ_1 and $\rho_2$denote the density of water and air, μ_1 and μ_2 denote the viscosity coefficient of water and air, respectively.

Wave Generation and Damping

Our naoe-FOAM-SJTU solver includes a wave generation and damping module. To generate the Stokes first deep water wave, the inlet boundary conditions are set as follows:

$$\eta = a\cos(\mathbf{k}\cdot\mathbf{x} - \omega_e t + \delta) \tag{8}$$

$$u = U_0 + a\omega e^{kz}\cos(\mathbf{k}\cdot\mathbf{x} - \omega t + \delta) \tag{9}$$

$$v = a\omega e^{kz}\cos\beta\cos(\mathbf{k}\cdot\mathbf{x} - \omega t + \delta) \tag{10}$$

$$w = a\omega e^{kz}\sin\beta\sin(\mathbf{k}\cdot\mathbf{x} - \omega t + \delta) \tag{11}$$

Where ζ is transient wave elevation, a, ω and $\boldsymbol{k}$ are wave amplitude, wave frequency and wave number, respectively. $\boldsymbol{U_0}$ is hull speed.

To avoid wave reflection, a sponge layer is setup at the outlet of the computational domain. The term of f_s is introduced into Eq. (2) for absorption of waves and defined as:

$$f_s(x) = \begin{cases} -\rho\alpha_s\left(\dfrac{x-x_s}{L_s}\right)^2(U-U_{ref}) \\ 0 \end{cases} \tag{12}$$

Where in which, ρ is the water density. The α_s is an artificial viscosity coefficient controlling the intensity of the sponge layer. The x_s is the coordinate of the start position of the sponge layer and L_s is the length of the sponge layer. The source term f_s has no effects on the domain out of the sponge layer.

Discretization schemes

The RANS and VOF transport equations are discretized by finite volume method (FVM). Van Leer scheme is applied for VOF equation in OpenFOAM. The PIMPLE (merged PISO-SIMPLE) algorithm is used to solve the coupled equation of velocity and pressure. The convection terms are solved by a second-order TVD limited linear scheme, and the diffusion terms are approximated by a second-order central difference scheme.

Computational model and test conditions

Computational model

To simulate wave run-up on a truncated surface-piercing column, the cylinder model at full scale from MOERI is selected in numerical simulation. The radius of the cylinder is R = 8.0 m, and the draft is 24.0 m. A series of wave probes in both experiment and numerical simulation are shown in Figure 1, and the corresponding location are given in Table 1. The distances between inner circle and outer circle are 0.2063 m and 8 m, respectively.

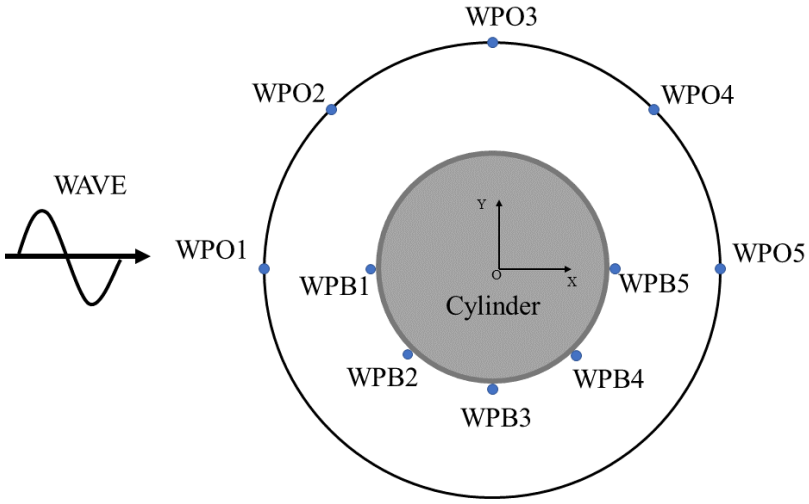


Figure 1 Layout of wave probes

Table 1 Location of wave probes

Inner circle	x (m)	y (m)	Outer circle	x (m)	y (m)
WPB1	-8.2063	0.0000	WPO1	-16.0000	0.0000
WPB2	-5.8027	-5.8027	WPO2	-11.3137	-11.3137
WPB3	0.0000	-8.2063	WPO3	0.0000	16.0000
WPB4	5.8027	-5.8027	WPO4	11.3137	11.3137
WPB5	8.2063	0.0000	WPO5	16.0000	0.0000

Figure 2 shows the arrangement of computational domain. The domain extends to $-2L < x < 3L$, $-L < y < L$, $-L < z < 0.5L$. Where L represents the wave length. The water depth is set as L. The length of sponge length is L, starting from $x = 2L$. The vertical cylinder is fixed in the center of the wave tank.

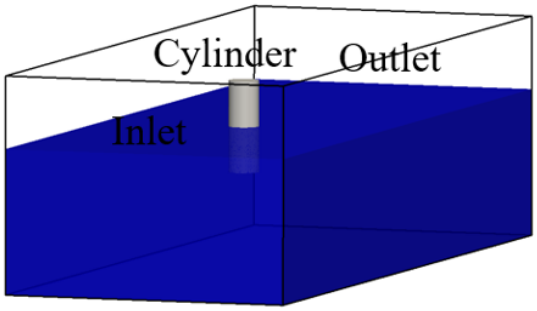


Figure 2 Computational domain

The computational mesh is shown in Figure 3. About 70 grids per wavelength and 20 per wave height are applied. The total grid number is about 1.7 million. To make it easy to converge in each time step, the interface Courant number was controlled to be under 0.3. The

time step is 0.001s in each case. The boundary conditions are as follows: Velocity inlet is adopted. Zero-Gradient condition is applied at the outlet. The no-slip boundary condition is imposed on the cylinder. The symmetry boundary condition is applied to the side walls.

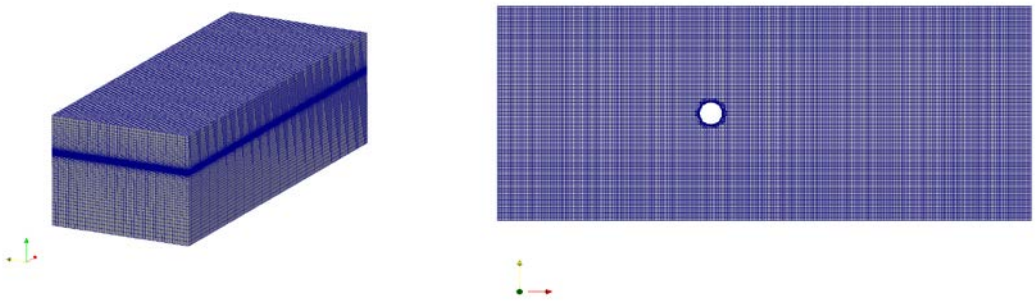


Figure 3 Mesh of computational domain

Test conditions

The incident wave conditions were set up according to the benchmark study conditions proposed by the 27th ITTC committee. The Stokes first order deep water wave is applied in the present work. Two wave periods (T = 7s, 9s) have been investigated. Three steepness parameters (H/L = 1/30, 1/16 and 1/10) were simulated for each wave period, where H is the wave height and L is the wave length. Another parameter in our analysis is the scattering parameter k_0r, where k_0 ($k_0 = 2\pi/L$) is the wave number. Details of the wave conditions are listed in Table2.

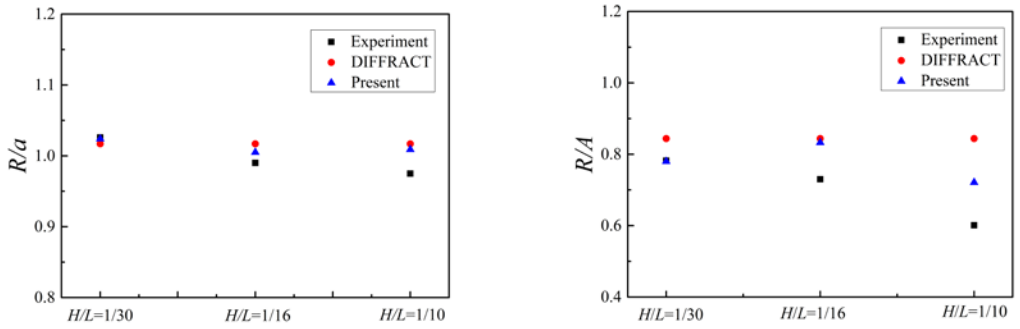
Table 2 Wave conditions

	T = 7 s			T = 9 s		
H/L	L(m)	D/L	H (m)	L(m)	D/L	H (m)
1/30			2.548			4.212
1/16	76.44	0.21	4.777	126.36	0.13	7.898
1/10			7.644			12.636

Results and discussion

RAOs of surface elevations

The response amplitude operates (RAOs) are effectively transfer functions used to determine the effect of wave on the ocean structures. The obtained time histories of free surface elevation are generally analyzed to acquire the RAOs using the Fourier analysis. The acquired RAOs of free surface elevations at ten wave probe locations from our CFD simulation are compared with the experimental data from MOERI. The RAOs of free surface elevations from potential flow solver DIFFRACT [14] are also adopted in this work. Figure 4 shows the comparisons of the RAOs of surface elevation near the cylinder at WPB3 and WPB4 for the condition wave period T = 7s and H/L =1/30, 1/16, 1/10. As shown in Figure 4(a), for small wave steepness condition, both the CFD results and DIFFRACT results agree well with the experimental data at WPB3. As the steepness increases, the CFD solver can give more accurate prediction than the potential flow theory. This is more obvious for WPB4 in Figure 4(b), for the H/L = 1/10 condition, the CFD solver can give a much better agreement with the experimental data. This may be the strong nonlinear interactions at the downstream location WPB4. Reasonable agreement between the CFD results and experimental data implies that the present solver is capable to predict the wave run-up on a truncated cylinder.



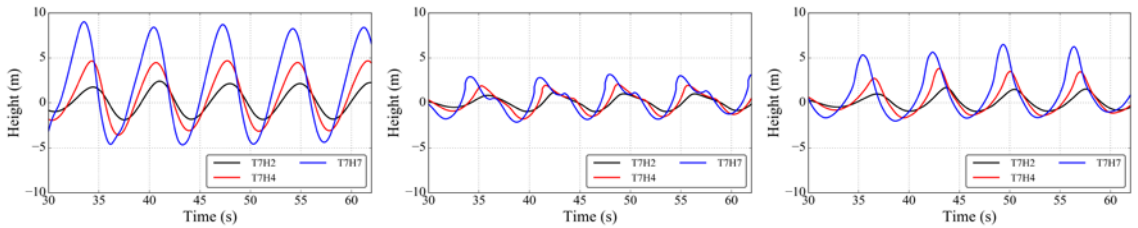
(a)WPB3 **(b)WPB4**

Figure 4 Comparisons of RAOs of surface elevations for T = 7s

Time histories of wave probes

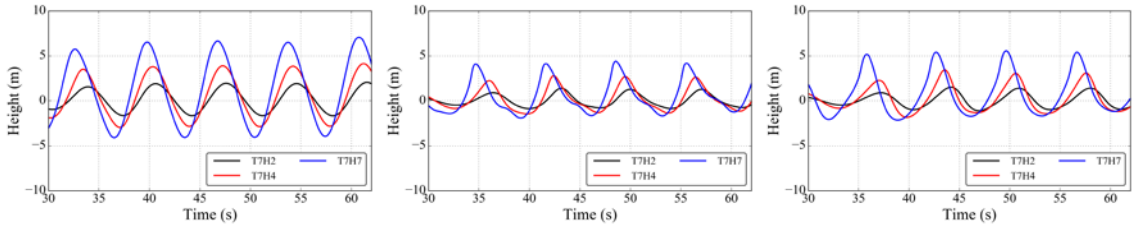
For the wave period T = 7s condition, time histories of the free surface elevation obtained from our CFD simulation are shown in Figure 5 (inner circle probes) and Figure 6 (outer circle probes). As can be seen from Figure 5, the wave probe in front of the cylinder reaches the largest wave amplitude in any wave height condition. This may due to the wave-structure interaction in front of the cylinder. When the wave crest reaches the cylinder, the water runs up vertically along the cylinder. The surface elevation at WPB1 is about two times than the incident wave amplitude. The surface elevation at the downstream quarter point WPB4 is the smallest, and secondary crests can be found for steeper waves (H/L = 16 and 1/10). This may imply strong nonlinear interaction at this location. After passing the sides of the column, the incident wave encounters and interferes with each other, resulting free wave flows backwards in the upstream direction and overlaps with the incident wave. Thus, the secondary crest can be observed. Similar trends can be found for outer circle wave probes, as shown in Figure 6. However, the discrepancies for these wave probes is reduced compared with the inner circle ones. The secondary crests at WPO4 is not so obvious as WPB4 near the cylinder.

For the wave period T = 9s condition, time histories of the free surface elevation are shown in Figure 7 (inner circle probes) and Figure 8 (outer circle probes). The missing of surface elevation data at WPO4 for H/L = 1/16 condition is caused by inappropriate wave probe setup in numerical simulation. Secondary wave crest can also be found at WPB4 in steep waves (H/L = 16 and 1/10).



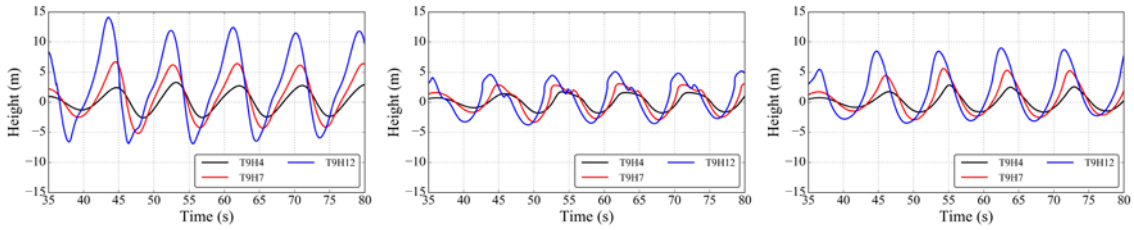
(a)WPB1 **(b)WPB4** **(c)WPB5**

Figure 5 Time series of surface elevation for inner circle wave probes for T = 7s



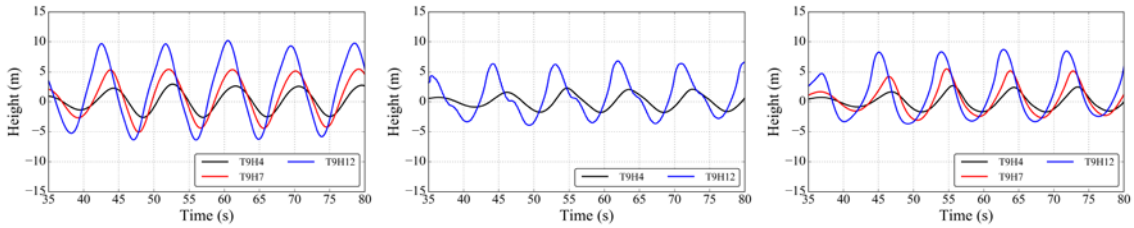
(a)WPO1 **(b)WPO4** **(c)WPO5**

Figure 6 Time series of surface elevation for outer circle wave probes for *T* = 7s



(a)WPB1 **(b)WPB4** **(c)WPB5**

Figure 7 Time series of surface elevation for inner circle wave probes for *T* = 9s



(a)WPO1 **(b)WPO4** **(c)WPO5**

Figure 8 Time series of surface elevation for outer circle wave probes for *T* = 9s

Time histories of wave forces

Figure 9 and Figure 10 show the time histories of horizonal wave force for T = 7s and T = 9s, respectively. The horizonal wave force increases with the wave height. The nonlinearity in wave force is not as strong as surface elevation. This is because the local nonlinear effects are integrated out when computing the force, while the surface elevation show the original local nonlinearity of wave run-up the cylinder. So it is essential to study the local free surface near the cylinder.

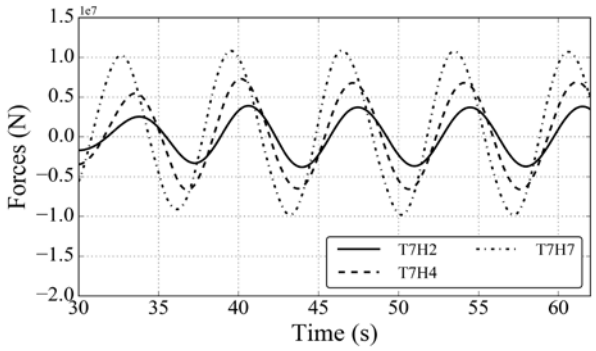


Figure 9 Time series of horizonal wave force for *T* = 7s

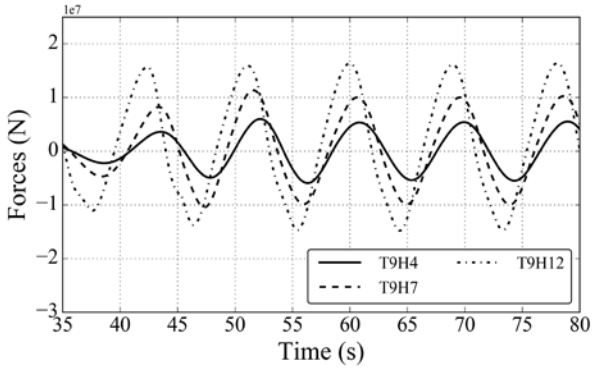


Figure 10 Time series of horizonal wave force for T = 9s

Scattered wave field around the cylinder

Figure 11 shows the local free surface around the cylinder for wave at T = 7s and H/L =1/10. The wave diffraction field around the cylinder can be clearly seen. When the wave crest approaches the cylinder, the water is blocked in front of the cylinder and concentric wave field (type1) can be observed. This is specified by Swan and Sheikh (2015) [22]. Then the water bypasses the side of the cylinder, the non-concentric wave field (type2) is developed at the downstream shoulders. This may induce strong nonlinearity at the shoulders of the cylinder. When the water encounters and overlaps at the rear side of cylinder, the run-up phenomenon can also be observed.

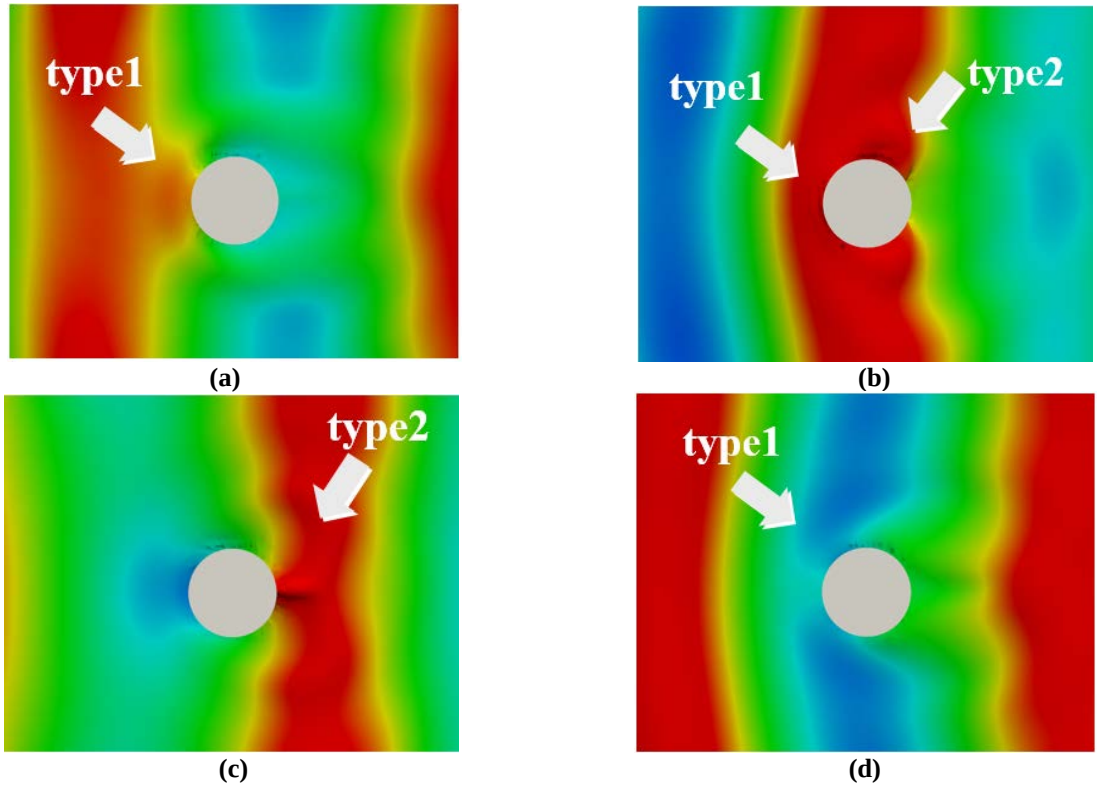
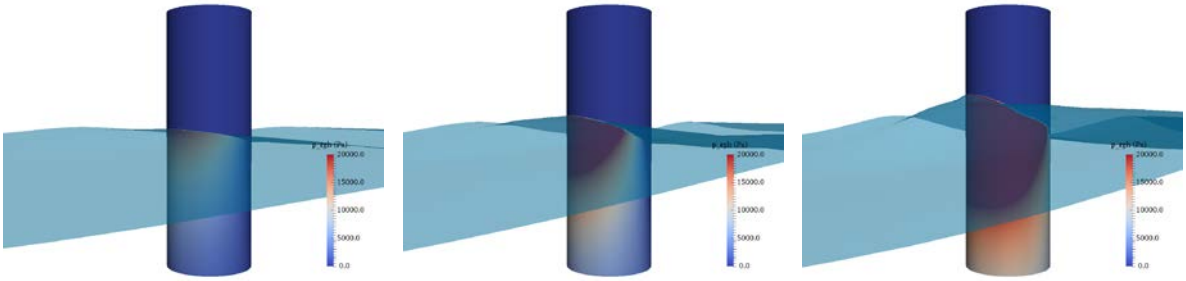


Figure 11 Local free surface around the cylinder

Figure 12 shows the dynamic pressure contours on the cylinder surface at the moment that the wave crest encounters the cylinder surface for wave at T = 7s and H/L =1/30, 1/16, 1/10. The dynamic pressure increases with the increasing wave height. For wave at T = 7s and H/L =

1/10 condition, the maximum dynamic pressure in the front of the cylinder can been found. For each case, the maximum dynamic pressure of the cylinder is close to the free surface.



(a)*H*/*L* = 1/30 **(b) *H*/*L* = 1/16** **(c) *H*/*L* = 1/10**

Figure 12 Dynamic pressure the cylinder

Conclusions

In this paper, numerical simulation of the wave run-up on a fixed surface-piercing cylinder in regular waves are performed by the in-house naoe-FOAM-SJTU solver. Two wave periods (T = 7s and 9s) and three wave heights (H/L =1/30, 1/16, 1/10) are conducted to investigate the wave run-up phenomenon. The predicted RAOs of local surface elevation are compared with the experimental data and good agreement can be acquired even at the strong nonlinear interaction location, where secondary crests can be observed for steeper waves. This can be caused by the overlap between the backward wave from rear part and the incident wave. Concentric and non-concentric wave fields around the cylinder can be captured by the present CFD solver. The surface elevation around the cylinder and the dynamic pressure increases with the increasing wave height. This study shows the capability of the present solver to investigate the wave run-up on a fixed cylinder. Further work should be focused on high-order harmonic and detailed flow information around the cylinder.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Galvin, C. J., Hallermeier, R. J. (1972) Wave runup on vertical cylinders, *Coastal Engineering* **1**, 1955-1974.

[2] Chakrabarti, S. K., Tam, W. A. (1975) Wave height distribution around a vertical cylinder, *Journal of Waterways, Harbours and Coastal Engineering* **101**, 225-230.

[3] Morris-Thomas, M. T., Thiagarajan, K., Krokstad, J. R. (2002) An experimental investigation of wave steepness and cylinder slenderness effects on wave run-up, In *Proceedings of 21st International Conference on Offshore Mechanics and Arctic Engineering*, OMAE2002-28050.

[4] Nielsen, F. G. (2003) Comparative study on airgap under floating platforms and run-up along platform columns, *Marine Structures* **16**, 97-134.

[5] Havelock, T. H. (1940) The pressure of water waves upon a fixed obstacle, In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **175**, 409-421.

[6] Hunt, J. N., Baddour, R. E. (1981) The diffraction of a nonlinear progressive wave by a vertical cylinder, *Quarterly Journal of Mechanics and Applied Machematics* **34**, 69-88.

[7] Trulsen, K., Teigen, P. (2002) Wave scattering around a vertical cylinder: fully nonlinear potential flow calculations compared with low order perturbation results and experiment, In *Proceedings of 21st International Conference on Offshore Mechanics and Arctic Engineering*, OMAE2002-28173.

[8] Teigen, P., Niedzwecki, J. M. (2003) Wave diffraction effects and runup around multicolumn structure, In *Proceedings of 13th International Offshore and Polar Engineering Conference*, 137-144.

[9] Morris-Thomas, M. T., Thiagarajan, k. (2004) The run-up on a cylinder in progressive surface gravity waves: harmonic components, *Applied Ocean Research* **26**, 98-113.

[10] Cao, H. J., Wan, D. C. (2014) Development of multidirectional nonlinear numerical wave tank by naoe-FOAM-SJTU solver. *International Journal of Ocean System Engineering* **4**, 52-59.

[11] Cao, H. J., Wan, D. C. (2017) Benchmark computations of wave run-up on single cylinder and four cylinders by naoe-FOAM-SJTU solver, *Applied Ocean Research* **65**, 327-337.

[12] Cao, H. J., Wan, D. C. (2015) RANS-VOF solver for solitary wave run-up on a circular cylinder, *China Ocean Engineering* **29**, 183-196.

[13] Danmeier, D. G., Seah, R. K., Finnigan, T., Roddier, D., Aubault, A., Vache, M., Imamura, J. T. (2008) Validation of wave run-up calculation methods for a gravity based structure. In *Proceedings of the ASME 27th International Conference on Offshore Mechanics and Arctic Engineering*, OMAE2008-57625.

[14] Sun, L., Zang, J., Chen, L., Taylor, R. E., Taylor, P. H. (2016) Regular waves onto a truncated circular column: A comparison of experiments and simulations, *Applied Ocean Research* **59**, 650-662.

[15] Kristiansen, T., Baarholm, R., & Stansberg, C. T. (2004) Validation of second-order analysis in predicting diffracted wave elevation around a vertical circular cylinder, In *Proceedings of 14th International Offshore and Polar Engineering Conference*, 342-349.

[16] Shen, Z. R., Wan, D. C., Carrica, P. M. (2015) Dynamic overset grids in OpenFOAM with application to KCS self-propulsion and maneuvering. *Ocean Engineering* **108**, 287-306.

[17] Wang, J. H., Wan, D. C. (2016) Numerical simulation of pure yaw motion using dynamic overset grid technology. *Chinese Journal of Hydrodynamics* **31**, 567-574.

[18] Ye, H. X., Wan, D. C. (2017) Benchmark computations for flows around a stationary cylinder with high Reynolds numbers by RANS-overset grid approach, *Applied Ocean Research* **65**, 315-326.

[19] Shen, Z. R., Wan, D. C. (2016) An irregular wave generating approach based on naoe-FOAM-SJTU solver, *China Ocean Engineering* **30**, 177-192.

[20] Liu, Z. H., Wan, D. C. (2016) Numerical simulation of green water on S-175 containership, In *Proceedings of the Second Conference of Global Chinese Scholars on Hydrodynamics*, 815-821.

[21] Hirt, C. W., Nichols, B. D. (1981) Volume of fluid (VOF) method for the dynamics of free boundaries. *Journal of Computational Physics* **39**, 201-225.

[22] Swan, C., Sheikh, R. (2015) The interaction between steep waves and a surface-piercing column, *Philosophical Transactions of the Royal Society of London A* **373**, 20140114.

Viscous modeling of liquefaction-induced settlement of existing structures using dynamic mesh

***†Wuwei Mao[1], Yu Huang[1], and Rouzbeh Rasouli[2]**

[1]Department of Geotechnical Engineering, Tongji University, Shanghai, China.
[2]Amec Foster Wheeler, Scarborough, Ontario, Canada
†Corresponding author: 09_maowuwei@tongji.edu.cn

Abstract

Seismic liquefaction is commonly observed in most of the large earthquakes. Settlement of structures, especially private houses resting on liquefiable soils has drawn continuous public concerns. The present study deals with the problem of structure subsidence through the framework of computational fluid dynamics. The liquefied ground was simulated using a viscous fluid model. The structure was modelled as a rigid body moving through the computational domain. Dynamic mesh method was used to reconstruct the mesh grid in order to avoid convergence problem due to invalidated meshes caused by structure movement. The volume of fluid method was used for free surface tracking. Furthermore, 1-G shaking table test was performed to obtain the subsidence behavior of a simulated structure resting a liquefied ground. The results showed that settlement of structure obtained from the proposed numerical model was comparable with the measured settlement from shaking table test. It is therefore suggested that the proposed method could be used for assessment of house settlement problems in liquefaction prone regions.

Keywords: Computational Fluid Dynamics, viscous, dynamic mesh, shaking table test, settlement

Introduction

From the most recent major earthquakes, soil liquefaction was substantially observed and its impact to the infrastructures was further recognized, such as the 2001 Arequipa earthquake[1], the 2008 Wenchuan earthquake[2] and the 2011 Tohoku earthquake[3]. Among many liquefaction related damages, e.g. sand boils, ground subsidence, ground cracking and lateral spreading, etc., the subsidence of the existing structures is one of the most intractable problems. Many structures nowadays are built on soft sandy deposits and are therefore potentially prone to liquefaction induced settlement. During the 2011 Tohoku earthquake of Japan, numerous buildings were observed subsided, especially for those structures with shallow foundations resting on the young reclaimed landfills near the Tokyo Bay area[4][5]. Although soil improvements for liquefaction mitigation are often employed in public or industrial projects, it is still not a common practice in residential developments. Damage assessment of such buildings from seismic liquefaction has become an important issue.

During liquefaction, the excess pore water pressure reaches the same value as the effective stress within the soil. The shear resistance of the post-liquefaction soil is substantially reduced to a negligible level. Many laboratory tests have been carried out to investigate the viscous fluid characteristics of liquefied soils[6]-[8]. The viscous description of the liquefied soils has been successfully applied to solve the liquefaction related problems such as lateral spreading[9][10] and large flow deformation[9][10]. For numerical analysis toward the problem of liquefaction-induced structure settlement, however, the studies are limited.

This study concerns with the subsidence of a uniformly weighed structure resting on a post-liquefied sandy deposit. To solve the problem, the liquefied ground was considered as a

viscous fluid and the finite volume method was used to realize the division of the computing domain. In previous studies reported by Uzuoka et al. (1998)[9], Hadush et al. (2000)[10] and Huang et al. (2012)[12], only meshes representing the moving fluid deforms. However, during the process of structure subsidence, it becomes the problem of a rigid body moving inside a fluid field. The Rigid mesh of moving structure will compress or drag the meshes within the fluid field and cause negative volume problems. Therefore, the previous fixed meshing method does not apply anymore. To overcome such limitations, a dynamic meshing method is used in the current study. The computing area is re-meshed at certain iteration steps when the preset criteria are satisfied. To verify the proposed method, a 1-g shaking table test was conducted to obtain the structure settlement behavior subjected to ground liquefaction.

Theory for viscous modeling

Viscous modeling of liquefied soil was newly developed in the recent years. The essence of the method is to regard the post-liquefied soils as fluid and use a viscous fluid model (e.g. the Bingham model) to simulate its behavior. The essence of the numerical model is described below.

Governing equation

The process of viscous fluid modeling solves the governing equations through domain partition and mesh generation. In fluid dynamic analysis, the conservation laws are applied to finite control volumes. The continuity equation and the momentum conservation equation can be written as:

$$\frac{\partial \rho}{\partial \boldsymbol{t}} + \nabla(\rho \boldsymbol{u}) = 0 \tag{1}$$

$$\frac{\partial (\rho \boldsymbol{u})}{\partial \boldsymbol{t}} + \nabla(\rho \boldsymbol{u}\boldsymbol{u}) = \nabla \boldsymbol{T} + \rho \boldsymbol{f}_v \tag{2}$$

For Newtonian fluid, the stress tensor $\boldsymbol{T}$ can be written as:

$$\boldsymbol{T} = -(\boldsymbol{p} + \frac{2}{3}\eta \nabla \boldsymbol{u})\boldsymbol{I} + 2\eta \boldsymbol{S} \tag{3}$$

and the Navier-Stokes equation can be written as:

$$\frac{\partial}{\partial \boldsymbol{t}}(\rho \boldsymbol{u}) + \nabla(\rho \boldsymbol{u}\boldsymbol{u}) = -\nabla \boldsymbol{p} + \rho \boldsymbol{f}_v + \nabla \boldsymbol{\tau} \tag{4}$$

where

$$\nabla \boldsymbol{\tau} = \eta \Delta \boldsymbol{u} + \frac{1}{3}\eta \nabla(\nabla \boldsymbol{u}) \tag{5}$$

In the above equations, ρ is the density, t is the time, $\boldsymbol{I}$ is the unit tensor, $\boldsymbol{T}$ is the stress tensor, $\boldsymbol{S}$ is the strain rate tensor, $\boldsymbol{u}$ is the velocity tensor, $\boldsymbol{f}_v$ is the volume force, η is the viscosity, p is the static pressure, $\boldsymbol{\tau}$ is the shear stress tensor.

Dynamic mesh

In the current study, the existing structure lies above the liquefied ground is modeled as a rigid body. During the process of its subsidence, the surrounding ground soils are subjected to deform and move to new positions, resulting in a twisted mesh grid within the computation domain. Convergence problems occur under the most unfavorable conditions when negative volume appears due to invalidated meshes. To overcome such limitation, a dynamic mesh method is adopted in this study by updating the mesh at each time step (or every several time steps) during structure motion. The mesh is updated once the mesh properties meet one of the

following criteria: (a) the skewness of a cell is greater than a specified maximum value; (b) the length of an edge is smaller/larger than a specified minimum/maximum length value.

Interface tracking

The volume of fluid (VOF) method developed by Hirt and Nichols (1981)[13] is used for interface tracking. As is mentioned above, the liquefied soil is modeled as a viscous liquid. It is therefore the whole computational domain is filled by two distinct phases: air and liquid. A volume fraction is assigned to each phase within a specific control volume, and the sum of the fractions in the same control volume should be unity (one). In such case, the free surface between the air and the liquid can be constructed based on the volume fraction of each phase. That is, if the liquid's volume fraction in the cell is α_{liq} , then the cell is empty (filled with air) when α_{liq}=0; the cell is full of liquid when α_{liq}=1; and the cell contains the air-liquid interface when0< α_{liq}<1. For cell contains air-liquid interface, the piecewise-linear approach is further used to reconstruct the interface geometry[14].

1-G shaking table test

The model test was conduct on a 1-G shaking table. A rectangular soil container with inside dimensions of 2.0m in length, 0.4m in width and 0.6m in height was mounted on the shaking table. The front side of the soil container was installed with transparent windows to directly observe sand deformation during shaking. The colored sand bars were placed on the side window in both vertical and horizontal direction to construct 10cm×10cm squares, which were helpful for observing sand formation. Fig.1 shows layout of experimental setting.

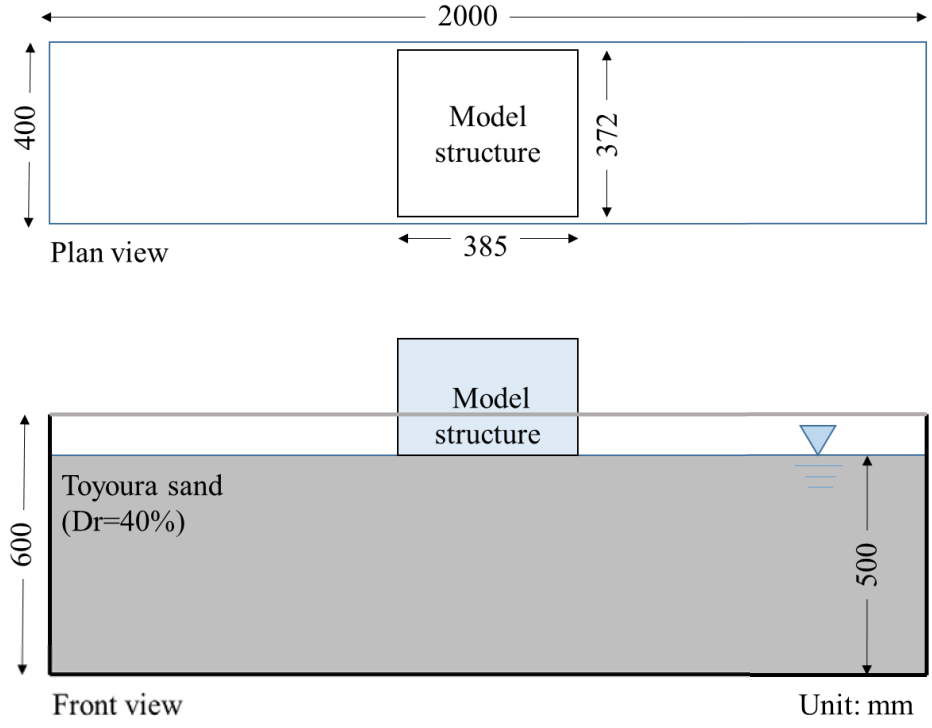


Fig.1 Layout of the shaking table test

The model ground was prepared with water pluviation method. Firstly, water level was risen up to 10cm over the previous layer and then the sand was poured uniformly at a specific height. Toyoura sand was used in the test and the basic properties of the sand are shown in Table 1. A wooden box filled with sand was used to model the surface structure. The base dimensions of the model structure were 38.5cm×37.2cm, which produced a surface pressure

of 2.05kPa when the model building was set to 30kg. During the test, the time history of structure settlement was recorded accurately by means of a laser displacement transducer. After finishing the model preparation, a sinusoidal motion of 350 Gal, 10 Hz, and 30 sec was applied to the model.

Table 1 Basic properties of the Toyoura sand

Property	Value
Specific gravity, G_s	2.65
Maximum void ratio, e_{max}	0.98
Minimum void ratio, e_{min}	0.61
D_{50}, mm	0.28
Coefficient of uniformity, Cu	1.30

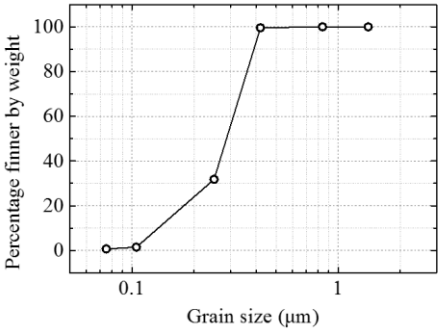
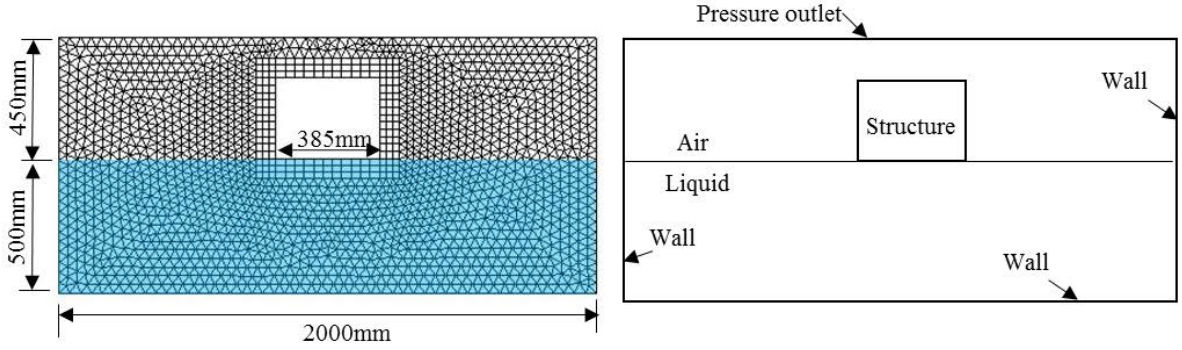


Fig.2 Grain size distribution of the sand

Computational model

The experimental case was further evaluated using the above proposed viscous modelling method. The simulation was implemented in FLUENT software. The schematic arrangements of the model details are shown in Fig.3. Same dimension was adopted with reference to the experimental setup. The existing structure lies one the model ground was modelled as a moving wall and six degree of freedom solver is used for buoyancy and drag force computing. Quadrilateral grids were used as boundary layer for region around the structure, while triangular grids were used for the other regions. When the structure subsidence occurs, the surrounding quadrilateral grids deform synchronously, leading to the deformation of outer triangular grids. Re-meshing is carried out when the specified criteria are met. In total, 1807 nodes, generating 168 quadrilateral cells and 3084 triangular cells, were used at the initial step. The upper edge of the model is set as pressure outlet, with a value of 1.013×10^5Pa. The left, right and bottom edge were set as wall (Fig.3b).



(a) mesh grid (b) boundary conditions

Fig.3 Details of the computational model

Results and discussions

Mesh configuration

Fig.4 show the mesh configurations at different computing time with an interval of 5s. Overall, the updated mesh maintained good quality throughout the whole modelling. The viaration of main paramters of the mesh condition is summerized in Table 2. It is seen that the numbers of toal nodes and cells were approximately the same.

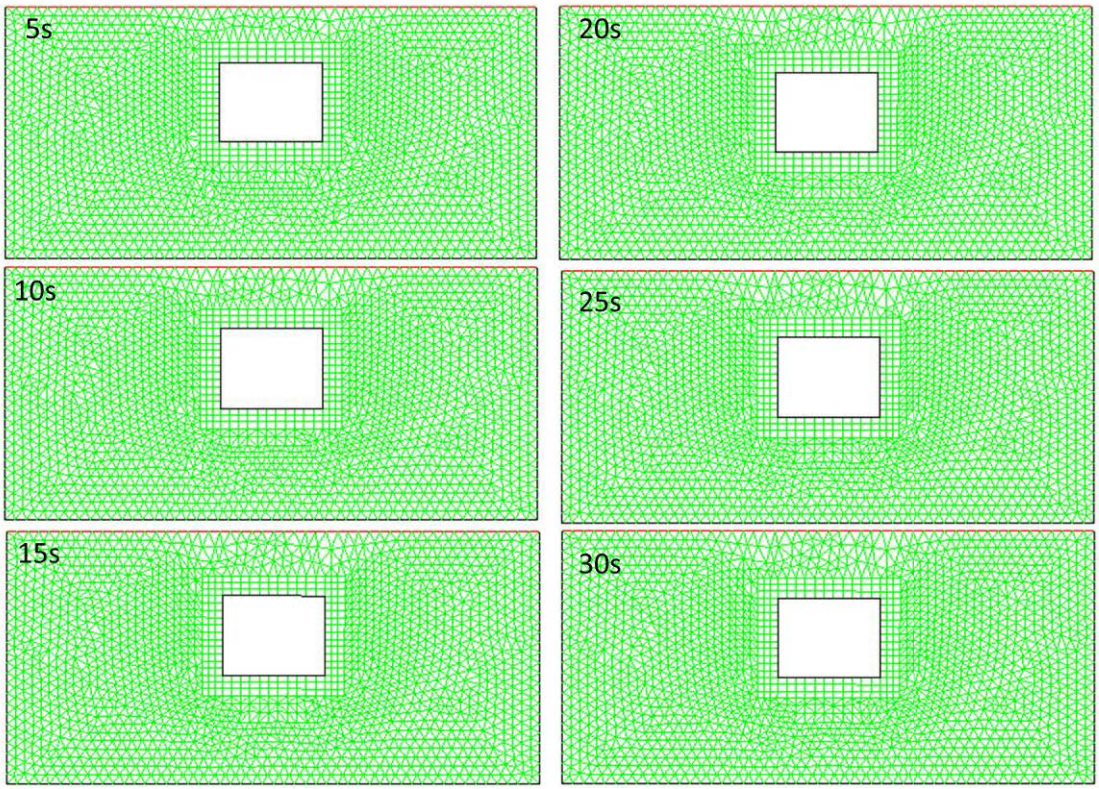


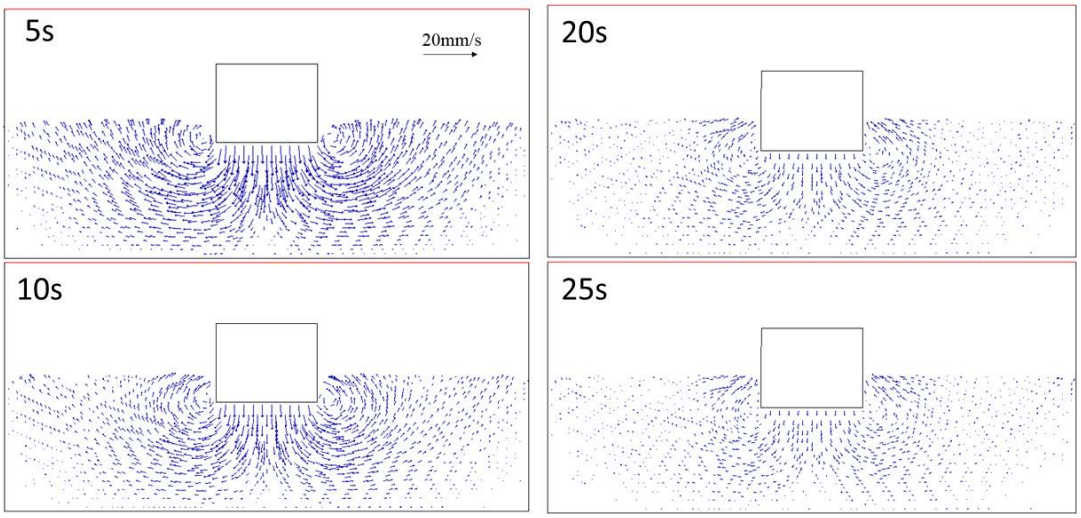
Fig.4 Mesh configurations at different time

Table 2 parameters of mesh grid at different time

Property	Initial	5s	10s	15s	20s	25s	30s
Quadrilateral cells	168	168	168	168	168	168	168
Triangular cells	3084	2928	2952	2942	2944	2948	2950
Nodes	1807	1729	1741	1736	1737	1739	1740

Velocity field

Fig.5 shows the velocity field of ground deformation. The maximum velocity is found to be near the edge of structure, and decreases towards the far field. The soils below the structure were pushed downward and flow to the two sides. In general, lager velocity was found at the first 10seconds, after which it significantly decreased, and it almost became zero at the end of simulation when the buoyancy and viscous drag force were balanced by the structure weight.



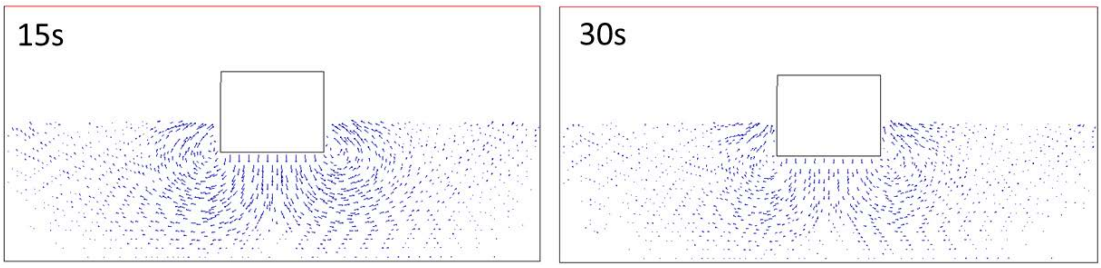


Fig.5 Development of velocity field during structure subsidence

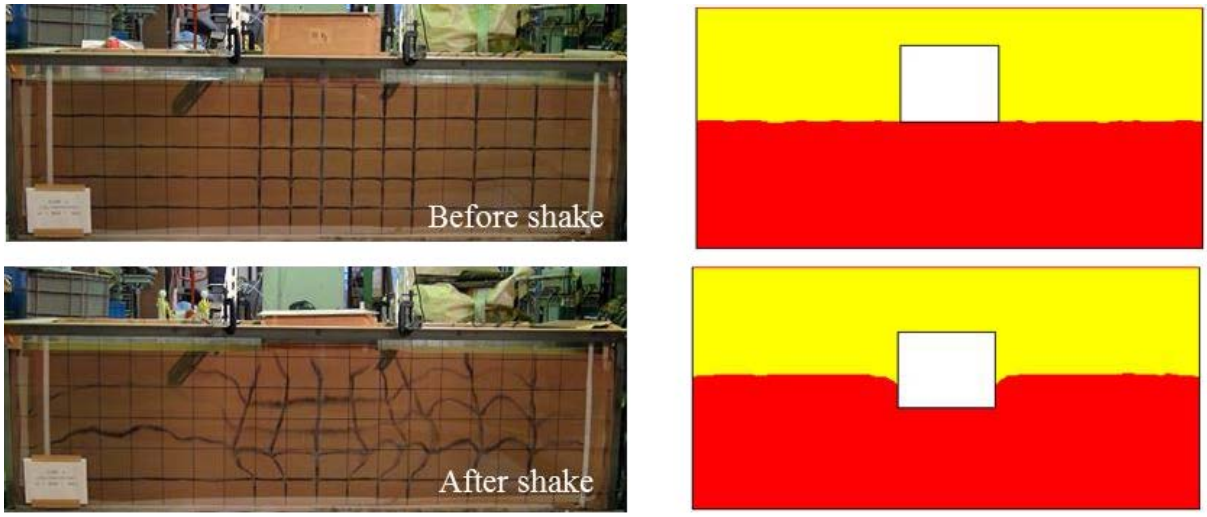


Fig.6 Comparison of experimental and simulation configuration before and after shake

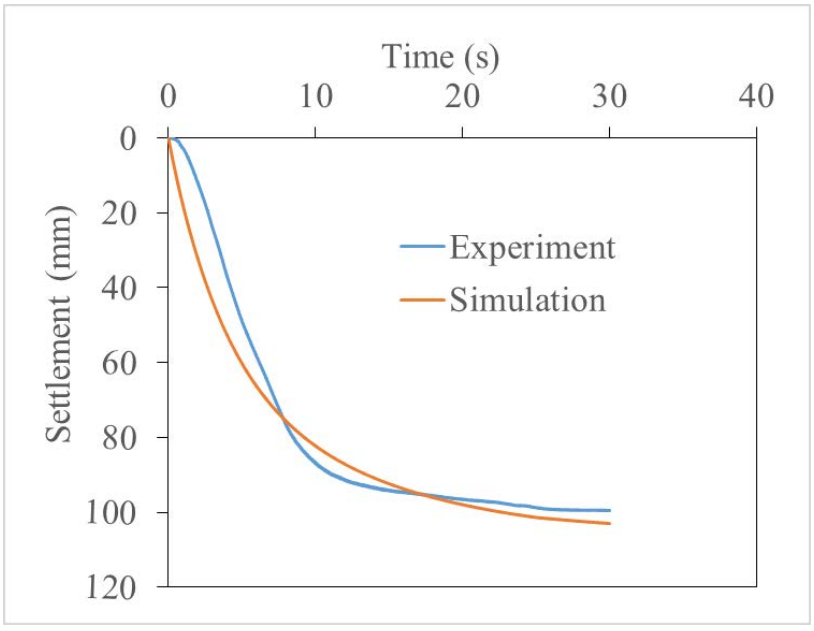


Fig.7 Comparion of total structrue settlement

Structure subsidence

The results of structure settlement before and after shake during shaking table test, together with the results obtained from simulation are shown in Fig.6. Notable settlement can be obervated both from experiment and numerical simulaition, which demonstrates the harmfulness of seimic liquefaction. Fig. 7 show the compariosn of settlement development during the whole process. The overall tendency as well as the ultimate settlement values are

in good agreement. The settlement was developed most rapidly within the first 10 seconds of test, and then it turned to become stable. This is also comfirmed by the moving velocity of the structure as shown in Fig. 8. It is found that the settlement increased significnatly at very beginning of liqufaction. The velocity climbed to an peak value within 0.5s and then decreased gradually. It is suggested that the bouyancy force plays a important role after the structure sinked into the liquifed soil.

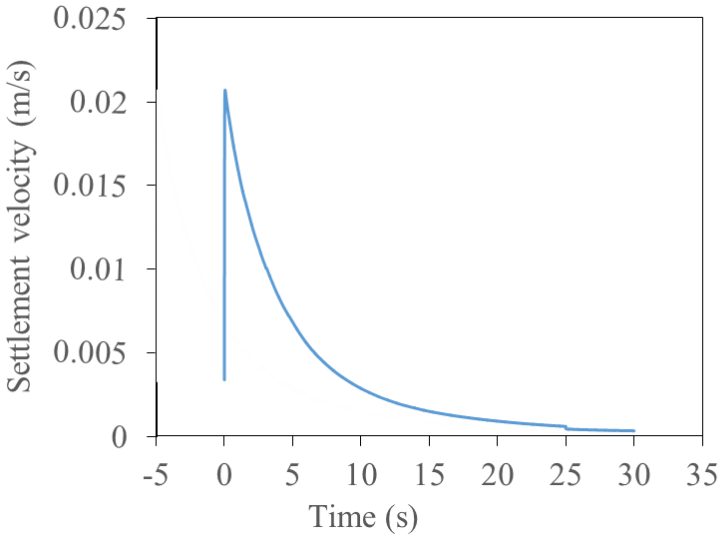


Fig.8 Developtment of settlment velocity of the structrue

Conclusions

Structure settlement is one of main damages observed in liquefaction prone areas. To analysis the settlement behavior of the existing structure. A viscous fluid modelling method with dynamic mesh was used to simulate the process of structure subsidence on liquefied ground. The liquefied ground was modelled as a viscous fluid, the structure was modelled as a moving rigid and the deformed mesh was updated using dynamic mesh.

The proposed method was examined by simulating and comparing with the results obtained from a 1-G shaking table test. It is found that the settlement of structure occurred most rapidly at the very beginning of liquefaction. The settlement behavior from viscous fluid modelling was consist with the measured settlement from shaking table test. It is therefore verified the validity of the proposed method.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grants Nos. 41372355 and 41625011) and the Program of Shanghai Academic/Technology Research Leader (Grant No.17XD1403700).

References

[1] Audemard FA, Gomez JC, Tavera HJ, Orihuela N (2005) Soil liquefaction during the Arequipa Mw 8.4, June 23, 2001 earthquake, southern coastal Peru. Eng Geol 78(3–4):237–255.
[2] Huang Y, Jiang XM (2010) Field-observed phenomena of seismic liquefaction and subsidence during the 2008 Wenchuan earthquake in China. Nat Hazards 54(3):839–850
[3] Towhata, I., Maruyama, S., Kasuda, K. I., Koseki, J., Wakamatsu, K., Kiku, H., and Hayashida, T. (2014). Liquefaction in the Kanto region during the 2011 off the pacific coast of Tohoku earthquake. Soils and Foundations, 54(4): 859-873.

[4] Ishihara, K., Araki, K. & Bradley, B., (2011) Characteristics of liquefaction-induced damage in the 2011 great east Japan earthquake. Hanoi, Proc. of International conference on geotechnics for sustainable development, Geotec, Hanoi.
[5] Tokimatsu, K., Tamura, S., Suzuki, H., and Katsumata, K. (2012). Building damage associated with geotechnical problems in the 2011 Tohoku Pacific Earthquake. Soils and Foundations, 52(5): 956-974.
[6] Towhata I, Vargas-Monge W, Orense RP, Yao M (1999) Shaking table tests on subgrade reaction of pipe embedded in sandy liquefied subsoil. Soil Dyn Earthq Eng 18:347–361
[7] Hwang JI, Kim CY, Chung CK, Kim MM (2006) Viscous fluid characteristics of liquefied soils and behavior of piles subjected to flow of liquefied soils. Soil Dyn Earthq Eng 26:313–323.
[8] Dashti, S., Bray, J.D., Pestana, J.M., Riemer, M. and Wilson, D. (2010). Mechanisms of seismically induced settlement of buildings with shallow foundations on liquefiable soil. Journal of Geotechnical and Geoenvironmental Engineering, 136(1):151 – 164.
[9] Uzuoka, R., Yashima, A., Kawakami, T., and Konrad, J. M. (1998). Fluid dynamics based prediction of liquefaction induced lateral spreading. Computers and Geotechnics, 22(3): 243-282.
[10] Hadush, S., Yashima, A., and Uzuoka, R. (2000). Importance of viscous fluid characteristics in liquefaction induced lateral spreading analysis. Computers and Geotechnics, 27(3): 199-224.
[11] Tamate, S., and Towhata, I. (1999). Numerical simulation of ground flow caused by seismic liquefaction. Soil Dynamics and Earthquake Engineering, 18(7): 473-485.
[12] Huang, Y., Mao, W., Zheng, H., and Li, G. (2012). Computational fluid dynamics modeling of post-liquefaction soil flow using the volume of fluid method. Bulletin of Engineering Geology and the Environment, 71(2): 359-366.
[13] Hirt, C. W., & Nichols, B. D. (1981). Volume of fluid (VOF) method for the dynamics of free boundaries. Journal of computational physics, 39(1), 201-225.
[14] Youngs, D. L. (1982). Time-dependent multi-material flow with large fluid distortion. Numerical methods for fluid dynamics, 24(2), 273-285.

Consistent meshfree method for phase-field model of brittle fracture

Yulong Shao*, †Qinglin Duan, Shuhui Li, Xikui Li, Hongwu Zhang

The State key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian, Liaoning 116024, P.R.China.

*Presenting author: yulongshaow@163.com
†Corresponding author: qinglinduan@dlut.edu.cn

Abstract

Numerical solution of the phase-field model of brittle fracture involves the capture of high gradients at cracks. Meshfree methods are convenient to construct high order approximation and to implement local refinement. This facilitates the capture of local high gradients to a large extent. Due to this reason, this paper introduces the element-free Galerkin (EFG) method, which is one of the major meshfree methods, to the phase-field modeling of crack growth. The meshfree discretization of the governing equations is described. To further improve the computational efficiency, the consistent integration scheme for quadratic EFG method is employed. The capability of the proposed consistent meshfree method for phase-field model of brittle fracture in predicting the load-displacement response and crack patterns is demonstrated by numerical examples.

Keywords: Meshfree; Phase-field model; Brittle fracture; EFG; Crack propagation

Introduction

The prediction of failure mechanisms due to crack initiation and propagation in solids is of great importance for engineering applications. The Griffith theory provides a criterion for crack propagation, but it is insufficient to determine curvilinear crack paths, crack kinking and branching angles. In particular, such a theory is unable to predict crack initiation. These defects of the classical Griffith-type theory of brittle fracture can be overcome by variational methods based on energy minimization as suggested by Francfort & Marigo [1], see also Bourdin, Francfort and Marigo [2]-[3]. The approximation regularizes a sharp crack surface topology in the solid by diffusive crack zones governed by a scalar auxiliary variable. Kuhn and Müller[4], using the framework of Gurtin [5] on the thermodynamics of order parameter based models, reformulated the energy minimization problem as the system of the stress equilibrium equation and a Ginzburg–Landau type evolution equation for phase field. However, the phase-field formulation does not distinguish between fracture behavior in tension and compression [3]. To avoid such situations, and, additionally, to prevent the interpenetration of the crack faces under compression, two modified regularized formulation were proposed by Amor et al. [6] and Miehe et al. [7]-[8] using an additive decomposition of the elastic energy density. Borden [9] presented a higher-order phase-field model formulation which could gain more regular and faster converging solutions of the variational problem of brittle fracture. Ambati et al. [10] and Nguyen [11] et al. modified the phase-field model in [8] in computational efficiency and accuracy, respectively. In order to avoid split the fourth order differential equation into two second order differential equations, Amiri et al. [12] applied a fourth order phase-field model for fracture based on local maximum entropy (LME) approximants. Now the phase field approaches to fracture have been applied in many fields such as the fracture in biological tissues, poroelastic medium and so on. They offer important new perspectives towards the theoretical and computational modeling of complex crack topologies. But these models have the high gradient of phase field at cracks which need high smoothness and accuracy and result in much more computational cost.

EFG method as a representative of Meshfree methods [13]-[16], has the advantages of high smoothness and offers substantial potential in many fields. However, their relatively low computational efficiency seriously hinders their applications. One main reason is that a plenty of integration points cost much more CPU time. Many efforts have been devoted to developing stable and efficient integration methods with reduced number of integration points [17]-[20]. Among them, the CEFG method by Duan [19] not only dramatically reduces the number of quadrature points in domain integration but also accurately passes the linear and quadratic patch tests, and remarkably improves the computational efficiency, accuracy and convergence of the standard EFG methods. Now it has been applied in many fields such as heat conduction [21], dynamics [22], elastic-plastic [23], nearly-incompressible elasticity [24]-[25] and so on. It is very promising.
In this paper, we will apply CEFG method to the phase-field model of brittle fracture, deduce the meshfree discretization of phase-field model and demonstrate its validity.

Phase-field model of brittle fracture

The variational approach to fracture mechanics provided by Francfort and Marigo [1] introduces the following energy functional for cracked body:

$$E(\mathbf{u},\Gamma) = E_u(\mathbf{u},\Gamma) + E_s(\Gamma) = \int_\Omega W_u(\boldsymbol{\varepsilon}(\mathbf{u}))d\Omega + \int_\Gamma g_c d\Gamma \tag{1}$$

where $E_u(\mathbf{u},\Gamma)$ represents the elastic energy stored in the cracked body, $E_S(\Gamma)$ is the energy required to create the crack according to the Griffith criterion, $\boldsymbol{\varepsilon}$ is the strain field, $\mathbf{u}$ is the displacement field, Γ is the geometry of the crack, W_u is the energy density function and g_c is the fracture toughness. In a regularized framework (phase field method), the above functional is substituted by the functional:

$$E(\mathbf{u},\Gamma) = \int_\Omega W_u(\boldsymbol{\varepsilon}(\mathbf{u}),d)d\Omega + g_c\int_\Omega \gamma(d)d\Omega \tag{2}$$

where

$$\gamma(d) = \frac{1}{2l}d^2 + \frac{l}{2}\nabla d\bullet\nabla d \tag{3}$$

Bounded phase field $d\in(0,1)$, $d = 0$ and $d = 1$ correspond to the unbroken and fully broken states, respectively. l is the regularization parameter which is related to the diffusive approximation of the sharp crack. The total energy is then rewritten as $E = \int_\Omega W d\Omega$ in which

$$W = W_u(\boldsymbol{\varepsilon}(\mathbf{u}),d) + g_c\gamma(d) \tag{4}$$

can be identified as the free energy.
In order to distinguish between fracture behavior in tension and compression, assuming isotropic elastic behavior of the body accounting for damage induced by traction only, we choose the following form for W_u

$$W_u(\boldsymbol{\varepsilon}(\mathbf{u}),d) = \Psi^+(\boldsymbol{\varepsilon})[g(d)+k] + \Psi^-(\boldsymbol{\varepsilon}) \tag{5}$$

k is a numerical parameter. The strain field is decomposed into tensile and compressive sections as

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^+ + \boldsymbol{\varepsilon}^- \tag{6}$$

and

$$\Psi^\pm(\boldsymbol{\varepsilon}) = \frac{1}{2}\lambda\langle tr(\boldsymbol{\varepsilon})\rangle_\pm^2 + \mu tr[(\boldsymbol{\varepsilon}^\pm)^2] \tag{7}$$

where λ, μ are the elastic bulk modulus and shear modulus.

$$\boldsymbol{\varepsilon}^\pm = \sum_{I=1}^{D}\langle\varepsilon_I\rangle_\pm \mathbf{n}_I\otimes\mathbf{n}_I \tag{8}$$

ε_I and $\mathbf{n}_I$ are the eigenvalues and eigenvectors of $\boldsymbol{\varepsilon}$. In (8) $\langle x\rangle_{\pm}=\frac{1}{2}(x\pm|x|)$. The degradation function $g(d)=(1-d)^2$.

To handle the irreversibility of the crack phase-field evolution in a general and possibly cyclic, loading/unloading scenario. Miehe et al. [8] introduced the strain history functional:

$$H(\mathbf{x},t)=\max_{\tau\in[0,t]}\Psi^{+}(\boldsymbol{\varepsilon}(\mathbf{x},\tau)) \tag{9}$$

For a discussion and justification of the use of this function, the reader is invited to refer to the mentioned reference. Then we can obtain the governing equations for phase field and displacement field as follows

$$\begin{cases} div\boldsymbol{\sigma}(\mathbf{u},d)=\mathbf{0} \\ 2(1-d)H+\dfrac{g_c}{l}\left(d-l^2\Delta d\right)=0 \end{cases} \tag{10}$$

in the domain Ω along with boundary conditions:

$$\mathbf{n}\cdot\boldsymbol{\sigma}=\bar{\mathbf{t}} \quad \text{on} \quad \partial\Omega_{\sigma} \tag{11}$$

$$\mathbf{u}=\bar{\mathbf{u}} \quad \text{on} \quad \partial\Omega_{u} \tag{12}$$

$$d(\mathbf{x})=1 \quad \text{on} \quad \Gamma \tag{13}$$

$$\nabla d(\mathbf{x})\cdot\mathbf{n}=0 \quad \text{on} \quad \partial\Omega \tag{14}$$

where

$$\boldsymbol{\sigma}(\mathbf{u},d)=\left((1-d)^2+k\right)\left\{\lambda Tr\boldsymbol{\varepsilon}_{+}1+2\mu\boldsymbol{\varepsilon}^{+}\right\}+\lambda Tr\boldsymbol{\varepsilon}_{-}1+2\mu\boldsymbol{\varepsilon}^{-} \tag{15}$$

Meshfree discretization

The phase field and phase field gradient are approximated

$$d(\mathbf{x})=\sum_I \mathbf{N}_d^I(\mathbf{x})d_I \quad \nabla d(\mathbf{x})=\sum_I \mathbf{B}_d^I(\mathbf{x})d_I \tag{16}$$

$\mathbf{N}_d(\mathbf{x})$ and $\mathbf{B}_d(\mathbf{x})$ are vectors and matrices of meshfree nodal shape functions and their derivatives for phase field, respectively. By using the standard Galerkin procedure with Penalty method enforcing the essential boundary condition for initial crack induced by phase-field, the final discretized equation is

$$(\beta\mathbf{K}_d^p+\mathbf{K}_d)\mathbf{d}=\beta\mathbf{F}_d^p+\mathbf{F}_d \tag{17}$$

where β is the penalty parameter.

$$\mathbf{K}_d^p=\int_{\Gamma}\mathbf{N}_d^{\mathrm{T}}\mathbf{N}_d d\Gamma,\ \mathbf{F}_d^p=\int_{\Gamma}\mathbf{N}_d^{\mathrm{T}}dd\Gamma \tag{18}$$

$$\mathbf{K}_d=\int_{\Omega}\left\{\left(2H+\frac{g_c}{l}\right)\mathbf{N}_d^{\mathrm{T}}\mathbf{N}_d+g_c l\mathbf{B}_d^{\mathrm{T}}\mathbf{B}_d\right\}d\Omega \tag{19}$$

$$\mathbf{F}_d=\int_{\Omega}2\mathbf{N}_d^{\mathrm{T}}Hd\Omega \tag{20}$$

The displacement field is approximated

$$\mathbf{u}(\mathbf{x})=\sum_I \mathbf{N}_I(\mathbf{x})\mathbf{u}_I \tag{21}$$

$\mathbf{N}(\mathbf{x})$ is the matrices of meshfree nodal shape functions of displacement field. Introduce two shifted strain matrix

$$\boldsymbol{\varepsilon}^{+}=\mathbf{P}^{+}\boldsymbol{\varepsilon},\ \boldsymbol{\varepsilon}^{-}=\mathbf{P}^{-}\boldsymbol{\varepsilon} \tag{22}$$

By using the standard Galerkin procedure with Nitsche's Method [26] enforcing the essential boundary condition, the final discretized equation is

$$\left\{\mathbf{K}_u - \mathbf{K}^{\partial\Omega_u} + \beta\mathbf{K}_u^p\right\}\mathbf{u} = \mathbf{F}_u - \mathbf{F}^{\partial\Omega_u} + \beta\mathbf{F}_u^p \tag{23}$$

where

$$\mathbf{K}_u = \int_\Omega \mathbf{B}^{\mathrm{T}}\mathbf{D}(d)\mathbf{B}\mathrm{d}\Omega \qquad \mathbf{F}_u = \int_\Omega \mathbf{N}^{\mathrm{T}}\mathbf{f}d\Omega + \int_{\partial\Omega_\sigma} \mathbf{N}^{\mathrm{T}}\bar{\mathbf{t}}d\partial\Omega_\sigma \qquad \mathbf{B} = \mathbf{L}\mathbf{N} \tag{24}$$

$$\mathbf{K}^{\partial\Omega_u} = \int_{\partial\Omega_u}\left[\mathbf{N}^{\mathrm{T}}\left(\mathbf{L}_n^{\mathrm{T}}\mathbf{D}(d)\mathbf{B}\right) + \left(\mathbf{L}_n^{\mathrm{T}}\mathbf{D}(d)\mathbf{B}\right)^{\mathrm{T}}\mathbf{N}\right]\mathrm{d}\partial\Omega_u \tag{25}$$

$$\mathbf{F}^{\partial\Omega_u} = \int_{\partial\Omega_u}\left(\mathbf{L}_n^{\mathrm{T}}\mathbf{D}(d)\mathbf{B}\right)^{\mathrm{T}}\bar{\mathbf{u}}\mathrm{d}\partial\Omega_u \tag{26}$$

$$\mathbf{K}_u^p = \int_{\partial\Omega_u}\mathbf{N}^{\mathrm{T}}\mathbf{N}\mathrm{d}\partial\Omega_u, \quad \mathbf{F}_u^p = \int_{\partial\Omega_u}\mathbf{N}^{\mathrm{T}}\bar{\mathbf{u}}\mathrm{d}\partial\Omega_u \tag{27}$$

$$\mathbf{L} = \begin{bmatrix}\partial/\partial x & 0 & \partial/\partial y \\ 0 & \partial/\partial y & \partial/\partial x\end{bmatrix}^{\mathrm{T}} \qquad \mathbf{L} = \begin{bmatrix}n_x & 0 & n_y \\ 0 & n_y & n_x\end{bmatrix}^{\mathrm{T}} \tag{28}$$

$$\mathbf{D}(d) = \left((1-d)^2 + k\right)\left\{\lambda \mathrm{R}^+\mathbf{I}\mathbf{I}^{\mathrm{T}} + 2\mu\mathbf{P}^+\right\} + \lambda \mathrm{R}^-\mathbf{I}\mathbf{I}^{\mathrm{T}} + 2\mu\mathbf{P}^- \tag{29}$$

$$\mathbf{I} = [1,1,0]^{\mathrm{T}}, R^+ = \frac{1}{2}\left(sign(tr(\boldsymbol{\varepsilon})) + 1\right), \; R^- = \frac{1}{2}\left(sign(-tr(\boldsymbol{\varepsilon})) + 1\right) \tag{30}$$

In this paper, the meshfree nodal shape functions and their derivatives for phase field and displacement field are identical. The solution of them can be obtained in [13] and [19]. We adopted the modified derivatives of meshfree nodal shape functions in [19], namely CEFG.

Results

Single edge notched pure shear test

Consider a squared plate containing horizontal notch located at middle height from the left edge with a length of 0.5 mm. The geometric setup is depicted in Fig. 1. In order to capture the crack pattern properly, the mesh is refined in areas where the crack is expected to propagate. An effective nodal distance h≈0.001 mm in the critical zone is obtained. The regularization parameter $\mathrm{l} = 3\mathrm{h}$. The elastic bulk modulus is chosen to λ =121.15 kN/mm^2, the shear modulus to μ=80.77 kN/mm^2 and the critical energy release rate to g_c=2.7×10^{-3} kN/mm.

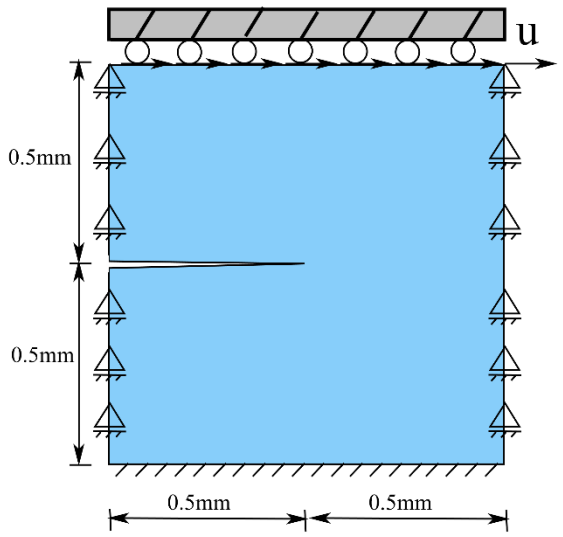


Fig. 1.Geometry and boundary conditions of single edge notched pure shear test

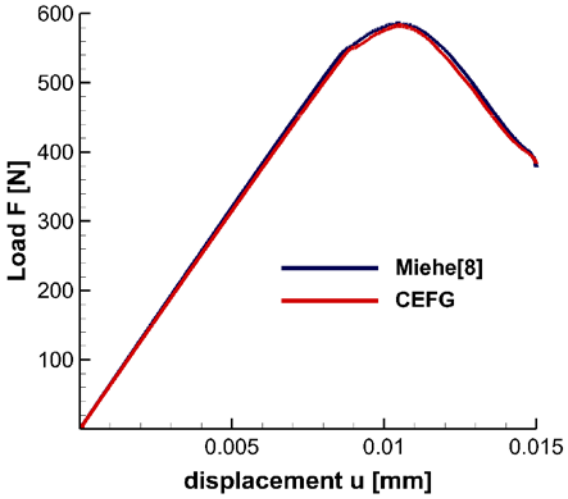


Fig. 2. Load-deflection curves of single edge notched pure shear test

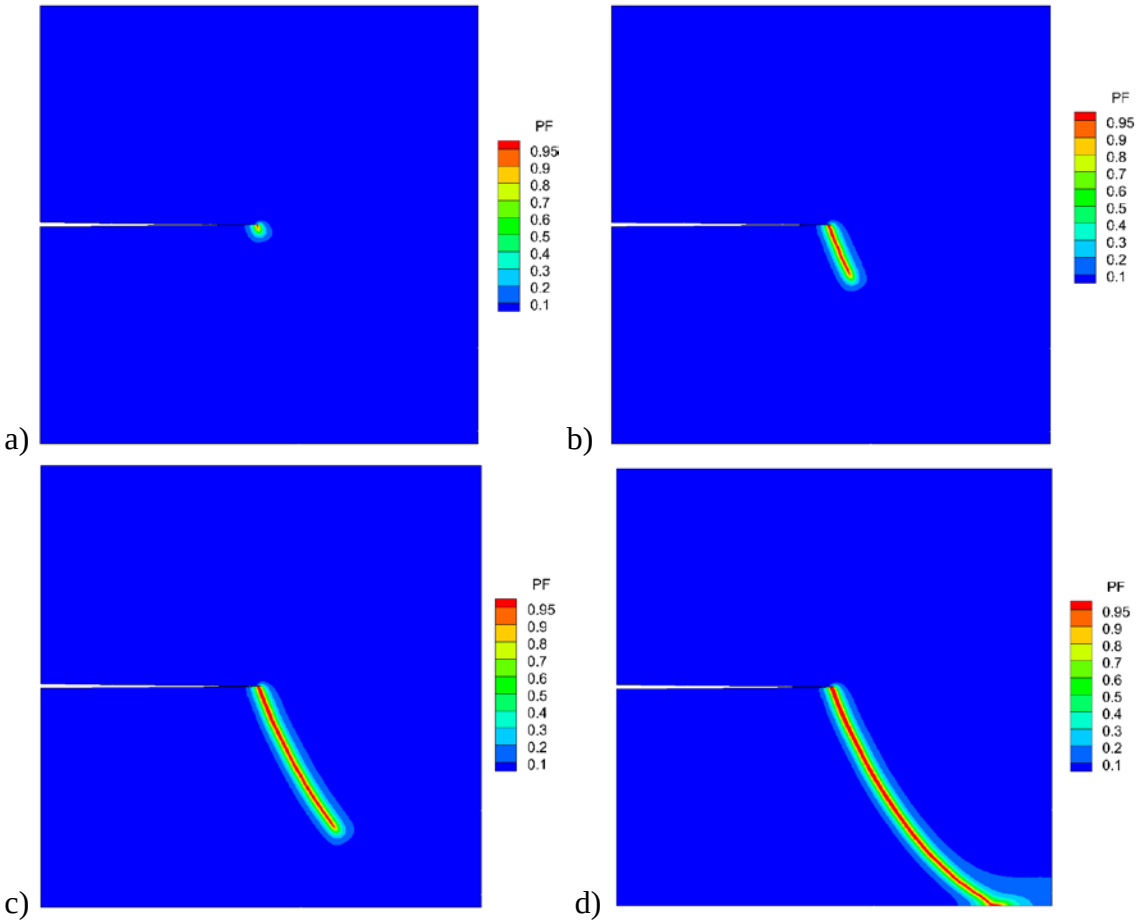


Fig. 3. Crack pattern of single edge notched pure shear test at a displacement of a) u=9.0×10^{-3} mm, b) u=11.0×10^{-3} mm, c) u=13×10^{-3} mm, d) u=15×10^{-3} mm

The computation is performed in a monotonic displacement driven context with constant increment $\Delta u=1\times10^{-5}$ mm. The load-deflection curves are depicted in Fig. 2. Obviously, it agrees well with the result in Miehe [8]. Figure 3 shows the crack patterns at several stages of loading. Here the red and blue colors indicate the damaged and undamaged material, respectively. From the Fig.3, we can find the phase field at crack is very smooth and CEFG method can accurately capture the crack pattern of single edge notched pure shear test.

Symmetric three point bending test

This test is a simply supported notched beam. The geometric setup as well as the loading conditions are illustrated in Fig. 4. The discretization is refined in the expected crack propagation zone. And an effective nodal distance is h≈0.008 mm in the critical zone. The bulk modulus is chosen to λ=12.00 kN/mm^2, the shear modulus to μ=8.0 kN/mm^2 and the critical energy release rate to $g_c=5.0\times10^{-4}$ kN/mm.

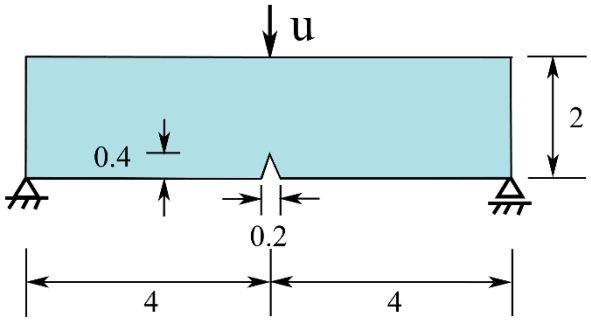


Fig. 4.Geometry and boundary conditions of simply supported notched beam

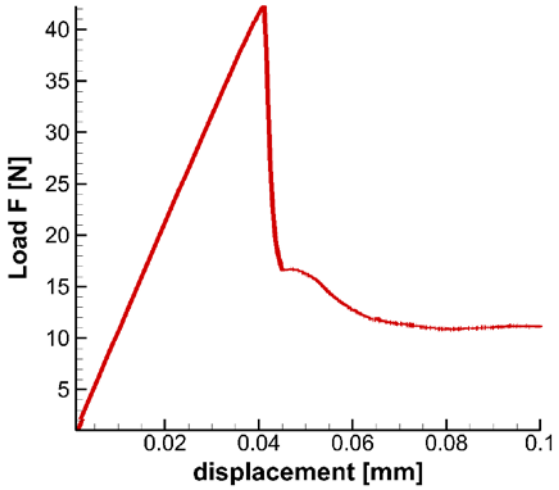


Fig. 5. Load-deflection curves of simply supported notched beam

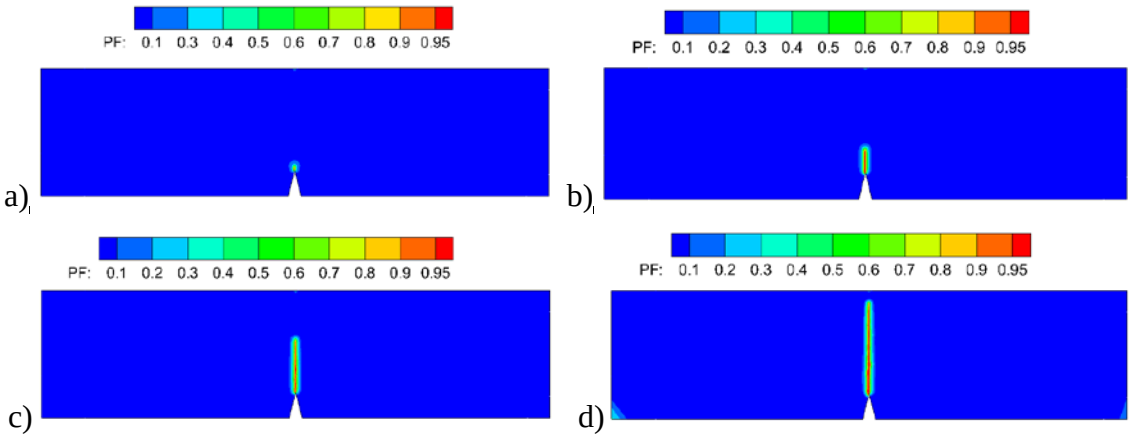


Fig. 6. Crack pattern of simply supported notched beam at a displacement of a) u=3.9×10^{-2} mm, b) u=4.131×10^{-2} mm, c) u=4.5×10^{-2} mm, d) u=1×10^{-1} mm

The computation is performed in a monotonic displacement driven context with constant increment $\Delta u=1\times10^{-3}$ mm in the first 39 loading steps. Continuing crack propagation then

demands for an adjustment of the displacement increment to $\Delta u=1\times10^{-5}$ mm for the subsequent 600 loading steps and $\Delta u=1\times10^{-4}$ for the last 550 loading steps. The load-deflection curves obtained are depicted in Fig. 5. It shows CEFG method can catch the basic characteristics of brittle fracture. The resulting contour plots of the crack topology are given in Fig. 6. We can reach the same conclusion like single edge notched pure shear test.

Conclusions

The quadratic CEFG method for the phase-field model of brittle fracture is implemented. The feasibility of the proposed method for modeling crack extension is investigated by two numerical examples, i.e. the single edge notched pure shear test and the simply supported notched beam. It is demonstrated that the load-displacement response and crack patterns are predicted correctly by the proposed consistent meshfree method.

Acknowledgement

The authors are pleased to acknowledge the support of this work by Science Challenge Project (No.JCKY2016212A502), the National Key Research and Development Program of China (2016YFC1402705, 2016YFC1402706), the National Natural Science Foundation of China (11372066, 11232003), the Fundamental Research Funds for the Central Universities (DUT15LK07), the open funds of the state key laboratory of water resources and hydropower engineering science (2015SGG03), the open funds of the state key laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP2016K007).

References

[1] G.A. Francfort, J.J. Marigo. (1998) Revisiting brittle fracture as an energy minimization problem, *Journal of the Mechanics and Physics of Solids* **46**, 1319–1342.
[2] B. Bourdin, G.A. Francfort, J.J. Marigo. (2008) *The Variational Approach to Fracture*, Springer Verlag, Berlin.
[3] B. Bourdin, G.A. Francfort, J.J. Marigo. (2000) Numerical experiments in revisited brittle fracture, *Journal of the Mechanics and Physics of Solids* **48**, 797–826.
[4] Kuhn C, Müller R. (2008) A phase field model for fracture, *Proc Appl Math Mech* **8**, 10223–10224.
[5] Gurtin M E. (1996) Generalized Ginzburg–Landau and Cahn–Hilliard equations based on a microforce balance, *Phys D* **92** (3–4), 178–192
[6] Amor H, Marigo JJ, Maurini C. (2009) Regularized formulation of the variational brittle fracture with unilateral contact: numerical experiments, *J Mech Phys Solids* **57**, 1209–1229.
[7] Miehe C, Welschinger F, Hofacker M. (2010) Thermodynamically consistent phase-field models of fracture: variational principles and multi-field FE implementations, *Int J Numer Methods Eng* **83**, 1273–1311.
[8] Miehe C, Hofacker M, Welschinger F. (2010) A phase field model for rate-independent crack propagation: robust algorithmic implementation based on operator splits, *Comput Methods Appl Mech Eng* **199**, 2765–2778.
[9] Borden MJ, Hughes TJR, Landis CM, Verhoosel CV. (2014) A higher-order phase-field model for brittle fracture: formulation and analysis within the isogeometric analysis framework, *Comput Methods Appl Mech Eng* **273**, 100–118.
[10] Ambati M, Gerasimov T, De Lorenzis L. (2015) A review on phase-field models of brittle fracture and a new fast hybrid formulation, *Computational Mechanics* **55**, 383-405.
[11] Nguyen T T, Yvonnet J, Zhu Q Z, et al. (2015) A phase field method to simulate crack nucleation and propagation in strongly heterogeneous materials from direct imaging of their microstructure, *Engineering Fracture Mechanics*, **139**, 18-39.
[12] Amiri F, Millán D, Arroyo M, et al. (2016) Fourth order phase-field model for local max-ent approximants applied to crack propagation, *Computer Methods in Applied Mechanics and Engineering* 2016.
[13] Belytschko T, Lu YY, Gu L. (1994) Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering* **37**, 229–256.
[14] Liu W.K., Jun S, Zhang Y.F. (1994) Reproducing kernel particle methods. *International Journal for Numerical Methods in Fluids* **20**, 1081–1106.
[15] Zhang X. (2001) Least-Squares collocation meshless method. *International Journal for Numerical Methods in Engineering* **51**, 1089–1100.

[16] Liu GR, Gu YT and Dai K. (2004) Assessment and applications of point interpolation methods for computational mechanics, *International Journal for Numerical Methods in Engineering* **59**, 1373-1397.
[17] Chen JS, Wu CT, Yoon S, You Y. (2001) A stabilized conforming nodal integration for Galerkin mesh-free methods. *International Journal for Numerical Methods in Engineering* **50**, 435–466.
[18] G. R. Liu, J. Zhang, K. Y. Lam, H. Li, G. Xu, Z. H. Zhong, G. Y. Li and X. Han. (2008) A gradient smoothing method (GSM) with directional correction for solid mechanics problems, *Computational Mechanics* **41**, 457-472.
[19] Duan Q L, Li X K, Zhang H W, Belytschko T. (2012) Second-order accurate derivatives and integration schemes for meshfree methods, *International Journal for Numerical Methods in Engineering* **92**, 399-424.
[20] Wang Dongdong,Wu Junchao. (2016) An efficient nesting sub-domain gradient smoothing integration algorithm with quadratic exactness for Galerkin meshfree methods, *Comput. Methods Appl. Mech. Engrg.* **298**, 485–519.
[21] Wang Bingbing, Gao Xin, Duan Qinglin. (2014) Meshfree method for heat conduction using quadratic consistent 1-point integration scheme, *Engineering Mechanics* **31**(9), 1-6.
[22] Wang Bingbing, Chen Songtao, Duan Qinglin. (2014) Quadratically consistent meshfree method for dynamics, *Chinese journal of applied mechanics* **31**(3), 353-358.
[23] Duan Qinglin, Gao Xin, Wang Bingbing,et al. (2014) A four-point integration scheme with quadratic exactness for three-dimensional element-free Galerkin method based on variationally consistent formulation, *Comput. Methods Appl. Mech. Engrg* **280**(10), 84–116.
[24] Ortiz-Bernardin A, Hale J.S.,Cyron C.J.. (2015) Volume-averaged nodal projection method for nearly-incompressible elasticity using meshfree and bubble basis functions, *Comput. Methods Appl. Mech. Engrg*, **285**, 427–451.
[25] Ortiz-Bernardin A, Puso M A, Sukumar N. (2015) Improved robustness for nearly-incompressible large deformation meshfree simulations on Delaunay tessellations, *Comput. Methods Appl. Mech. Engrg* **293**,348–374.
[26] Fernández-Méndez S, Huerta A. (2004) Imposing essential boundary conditions in mesh-free methods, *Computer Methods in Applied Mechanics and Engineering* **193**(12), 1257–1275.

Meshfree modeling of heat transfer in selective laser melting process

Songtao Chen*, †Qinglin Duan, Shuhui Li, Xikui Li and Hongwu Zhang

The State key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian, Liaoning 116024, P.R.China.

*Presenting author: chensongtao1221@163.com
†Corresponding author: qinglinduan@dlut.edu.cn

Abstract

Selective laser melting (SLM) process produces 'near net shape' parts from a stream of metal powders in a layer-by-layer manner for use in medical, aerospace and other industries. The thermal history generated by SLM is significant in determining the microstructure, mechanical properties, residual stress, and distortion of manufactured components. This study employs three-dimensional element-free Galerkin (EFG) method to investigate the temperature evolution in SLM. The moving heat source and the dependence of material properties on temperature are considered. It is found that the developed meshfree method is able to capture the high gradients of temperature in heat affected zoon. The effects of scan speed and laser power are also investigated.

Keywords: Meshfree; Selective laser melting; Additive manufacturing; EFG; Heat transfer

1. Introduction

Selective laser melting (SLM) is a promising additive manufacturing (AM) technique in which successive layers of metal powders are heated via laser in order to build a part. It has advantages for direct fabrication of three-dimensional (3D) parts with complex structures [1][2]. A typical process of layer by layer SLM fabrication is showed in Figure 1:

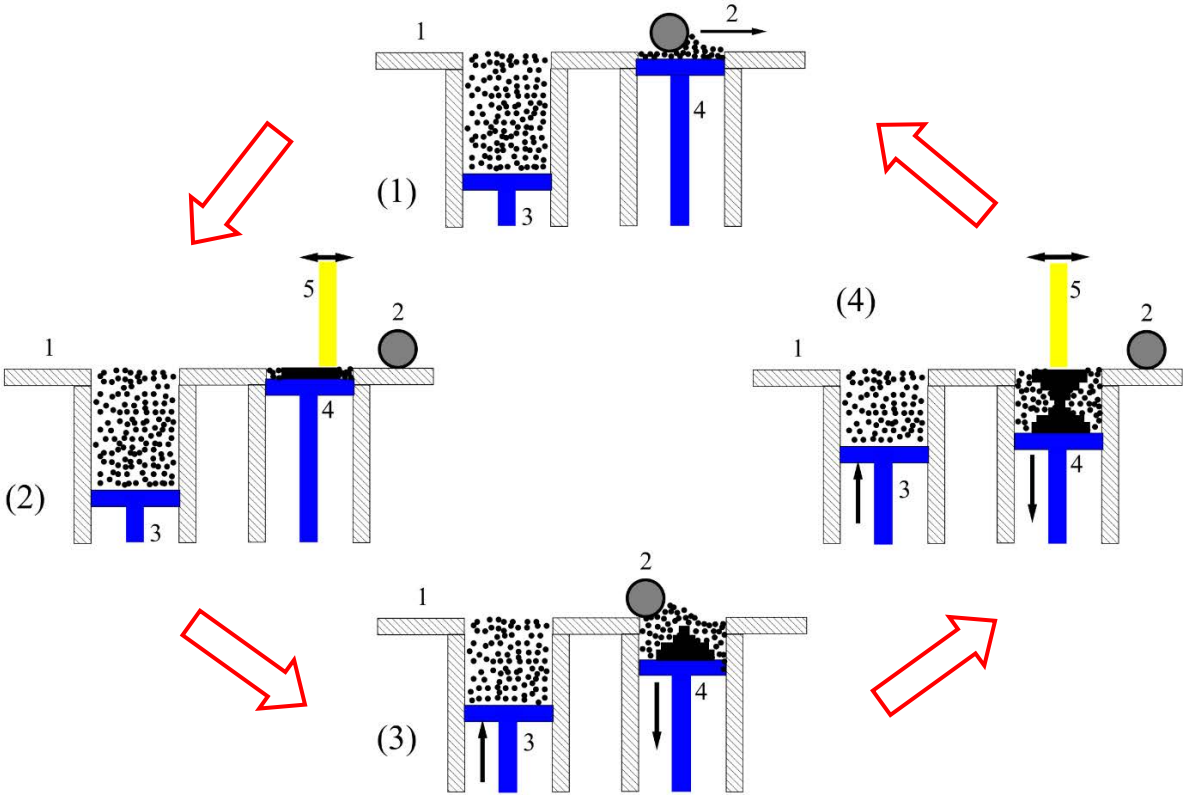


Figure 1.Scheme of the process (1) deposition of a powder layer, (2) scanning of the first layer, (3) deposition of a powder layer in the interval, (4) scanning of the last layer, and preparing for next part.

The SLM machine is composed of a fabrication plate 1 with two holes to which two containers are attached. A roller 2 drives powder from the left container and deposits a thin powder layer in the right container. The thickness of the layer is controlled by pistons moving up 3 and down 4. Then a laser beam 5 scans on the deposited powder under inert-gas protection. The absorbed laser energy melts the powder and makes a molten pool locally. As the molten pool freezes, this part will be fixed on the substrate. The product will be fabricated by repeating this procedure layer-by-layer. After that, the product will be cooled and removed out.

All the manufacturing course of SLM is in an inert-gas protection system, which prevents metal powders from assimilating oxygen, nitrogen, and hydrogen from the atmosphere. But as SLM is a complex physic metallurgy process, the structure always has uncontrollable defects such as balling, distortion, and heat fatigue cracking. Therefore, it is significant to study the thermal behavior during SLM. However, experimental measurements of the temperature during SLM are considered to be almost impossible because of the localized heating and superfast melting and solidification involved. Numerical simulation is a useful measure to solve these problems.

Recently, many FEM models have been established to investigate thermal behavior during AM. Gusarov *et al* [3]-[5] numerically analyzed the temperature distribution of a 316L powder bed during SLM. They discovered that the stability of the process is highly dependent on the scan speed, powder layer thickness, and the thermal properties of the materials. Li [6] studied the scan speed and the power's influence on the size of molten pool, and carried out many experiments of SLM process under different laser processing conditions to demonstrate the reliability of the physical model and simulation results. Hussein *et al* [7] used ANSYS parametric design language (APDL) to develop a non-linear transient model based on sequentially coupled thermo-mechanical field analysis code and simulated the temperature and stress fields in single 316L stainless steel layers built on the powder bed without support in SLM. They found that the predicted length of the molten pool increases at higher scan speed while both width and depth of the melt pool decrease. The cyclic melting and cooling rates in the scanned tracks result in high VonMises stresses in the consolidated tracks of the layer.

Considering that the temperature gradient around the molten pool is very high and fluctuates violently, traditional FEM shape function is difficult to simulate accurately. The meshfree methods [8]-[13] that have the apparent advantage of building highly smooth field is more suitable for this process. Meshfree methods such as the element-free Galerkin (EFG) method developed in recent twenty years have become a formidable competitor to the traditional finite element method (FEM) which dominates engineering analysis for decades. But the basis functions in Galerkin-based meshfree method are rational (nonpolynomial) functions. This is the central issue that introduce numerical errors when using standard Gauss quadrature for numerical integration. The errors can be reduced by using a large number of Integral points per cell; however, this substantially increases the computational costs in the numerical integration.

Many efforts have been devoted to develop stable and efficient integration methods with reduced number of sampling points such as the nodal integration [14]-[16], the stress-point integration [17]-[18], the support domain integration [19].et al. Among these, the nodal integration initialed by Bessial and Belytschko [15] can dramatically improve the efficiency since it use the minimum evaluating points (the nodes) as integration points. However, direct nodal integration is not stable and can't pass the patch tests, lots of works have been done to relieve this issue. Among these, Chen *et al.* [20] developed a stabilized conforming nodal integration (SCNI) by using a strain smoothing method at the nodal representative domain, which is stable and provides even better accuracy than Gauss integration. It can pass the linear patch test whereas Gauss integration fails. One outstanding advantage of this method is that no additional term or stabilization parameter is involved. So far, SCNI has developed to be a major integration scheme in meshfree method and the strain smoothing technique in SCNI has been extend into FEM analysis [21].

This paper uses three-dimensional SCNI to simulate the temperature fields in single 316L stainless steel layers built on the substrate in SLM. The material properties are obtained from [7]. The remained of the paper is organized as follows. Section 2 presents a summary of the Mathematic model for SLM process. The stabilized conforming nodal integration (SCNI) and numeral examples are next presented in Section 3. Conclusions are then summarized in Section 4.

2. Mathematic model for selective laser melting (SLM) process

2.1 Thermal modelling

SLM is a swift remelting process. When the laser beam scans the surface of a powder bed, a portion of the laser energy is reflected and the remainder is absorbed by the powder. The absorbed laser energy melts the upper part of powder, thereby yielding a molten pool. Metallurgical bonds form between adjacent tracks and neighboring layers as solidification occurs.The thermal equilibrium equation satisfies the following classical 3D heat conduction equation [7]:

$$\rho c\frac{\partial T}{\partial t}=\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial \mathrm{y}}\left(k\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial \mathrm{z}}\left(k\frac{\partial T}{\partial z}\right)+Q \tag{1}$$

where ρ is the material density (kg/m^3); c is the specific heat capacity(J/kg K); T is the temperature; t is the interaction time; k is thermal conductivity (W/m K); and Q=(x, y, z, t) is the volumetric heat generation (W/m^3).

The effective thermal conductivity is a function of the powder's porosity [1]. The thermal conductivity of the powder can be expressed as [22],

$$k_p=k_b\left(1-\phi\right) \tag{2}$$

where ϕ is the porosity of the powder, it varies from 0.4 for powder state to 0 at solid state. k_p and k_b are thermal conductivities of powder and bulk materials.

Considering the melting and solidification phenomena that happen during additive manufacturing, the latent heat for phase change cannot be neglected. To define the latent heat of fusion, enthalpy is expressed as a function of temperature, T, density, p, and specific heat, c, according to

$$H=\int \rho c\left(T\right)dT \tag{3}$$

The material properties are obtained from [7]

2.1.1. Initial and boundary conditions

The initial temperature distribution throughout the powder bed at $t=0$ can be defined as

$$T\left(x,y,z,t\right)\big|_{t=0}=T_0,\quad \left(x,y,z\right)\in\Omega \tag{4}$$

where T_0 is the ambient temperature and is taken as 25 C.Since the layers are built on substrate with large thickness, the temperature at the bottom of the structure can be assumed to stay at T_0.

$$T\left(x,y,z,t\right)\big|_{z=0}=T_0 \tag{5}$$

During SLM, besides thermal conduction, heat losses due to convection should also be taken into account for proper description of the thermal behavior.The natural boundary condition can be expressed as [23]

$$k\frac{\partial T}{\partial n}+q_c=0,\quad \left(x,y,z\right)\in\Gamma_c \tag{6}$$

where Γ_c represents the surfaces attached to imposed heat fluxes(convection), n is the normal vector of Γ_c, and q_c is heat convection defined as

$$q_c = h\left(T - T_0\right) \tag{7}$$

where h is the coefficient for heat convection.

2.1.2. Heat source model

Heat input model: The Gaussian cylindrical source heat input model is used to simulate the heat input on to the part. The heat source distribution Q is given by

$$Q = \frac{AP}{\pi r_0^2 h}\exp\left(-\frac{r^2}{r_0^2}\right) \tag{8}$$

r is the distance from the laser center to the point, r_0 is the laser beam waist radius, P is the laser power, A is the absorptivity of the powder material which can be calculated if the reflectivity of the material k is known. (a reflectivity of iron=0.7 is considered for 316L stainless steel).

3. Numerical solution: meshfree method with stabilized conforming nodal integration (SCNI)

3.1 meshfree method

EFG uses the moving-least squares (MLS) approximation to construct the nodal shape functions. Given a set of nodes $\mathbf{X}_I$ in the domain $\Omega \subset \mathbf{R}^3$, the MLS approximation of the displacement is

$$\mathbf{u}^h(\mathbf{x}) = \mathbf{N}(\mathbf{x})\mathbf{U} = \sum_I \mathbf{N}_I(\mathbf{x})\mathbf{U}_I \tag{9}$$

where $\mathbf{U}$ is the vector of nodal displacement parameters. $\mathbf{N}(\mathbf{x})$ is the matrix of nodal shape functions which can be written as

$$N_I(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{X}_I)\, w_I(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) \tag{10}$$

where $\mathbf{p}(\mathbf{x})$ is a vector of base functions which usually includes a complete basis of the polynomials to a given order, $w_I(\mathbf{x})$ a weight function and $\boldsymbol{\alpha}(\mathbf{x})$ the unknown vector. The unknown vector $\boldsymbol{\alpha}(\mathbf{x})$ can be determined by the so called reproducibility condition, i.e. the consistency condition

$$\mathbf{p}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I) N_I(\mathbf{x}) \tag{11}$$

Substitution of Eq.(10) into Eq.(11) leads to

$$\mathbf{A}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{p}(\mathbf{x}) \tag{12}$$

where

$$\mathbf{A}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I)\mathbf{p}^{\mathrm{T}}(\mathbf{X}_I)\, w_I(\mathbf{x}) \tag{13}$$

The nodal MLS shape functions $N_I(\mathbf{x})$ can be obtained from Eq.(11) after the unknown vector $\boldsymbol{\alpha}(\mathbf{x})$ is solved from Eq.(12). Computation of the derivatives of the MLS shape functions is by taking the derivative of Eq.(10)

$$N_{I,i}(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{X}_I)\left[w_{I,i}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) + w_I(\mathbf{x})\boldsymbol{\alpha}_{,i}(\mathbf{x})\right] \tag{14}$$

where subscripts preceded by commas denote partial derivatives with respect to spatial coordinates. The unknown $\boldsymbol{\alpha}_{,i}(\mathbf{x})$ in Eq.(7) can be solved from the derivative of Eq.(13)

$$\mathbf{A}(\mathbf{x})\boldsymbol{\alpha}_{,i}(\mathbf{x}) = \mathbf{p}_{,i}(\mathbf{x}) - \mathbf{A}_{,i}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) \tag{15}$$

with

$$\mathbf{A}_{,i}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I)\mathbf{p}^{\mathrm{T}}(\mathbf{X}_I)w_{I,i}(\mathbf{x}) \tag{16}$$

3.2 Consistent integration schemes

In order to develop a stabilized nodal integration for the Galerkin meshfree method to achieve higher efficiency with desired accuracy and convergent properties, Chen *et al.* [20] developed a stabilized conforming nodal integration (SCNI). The Strain smoothing method used in SCNI at the nodal representative domain depends on

$$\tilde{u}^h_{,i}(\mathbf{X}_L) = \frac{1}{A_L}\int_{\Omega_L} u^h_{,i}\,d\Omega, \quad A_L = \int_{\Omega_L} d\Omega \tag{17}$$

where Ω_L is a nodal representative domain-the yellow cube covered by the Γ_L in Figure 2(a). Applying divergence theorem to Equation (17), the smoothed gradient operator can be rewritten as

$$\tilde{u}^h_{,i}(\mathbf{X}_L) = \frac{1}{A_L}\int_{\Gamma_L} \mathbf{n}u^h_i\,d\Gamma \tag{18}$$

where Γ_L is the boundary of the representative domain as shown in Figure 2. Introducing MLS shape functions into (17) yields

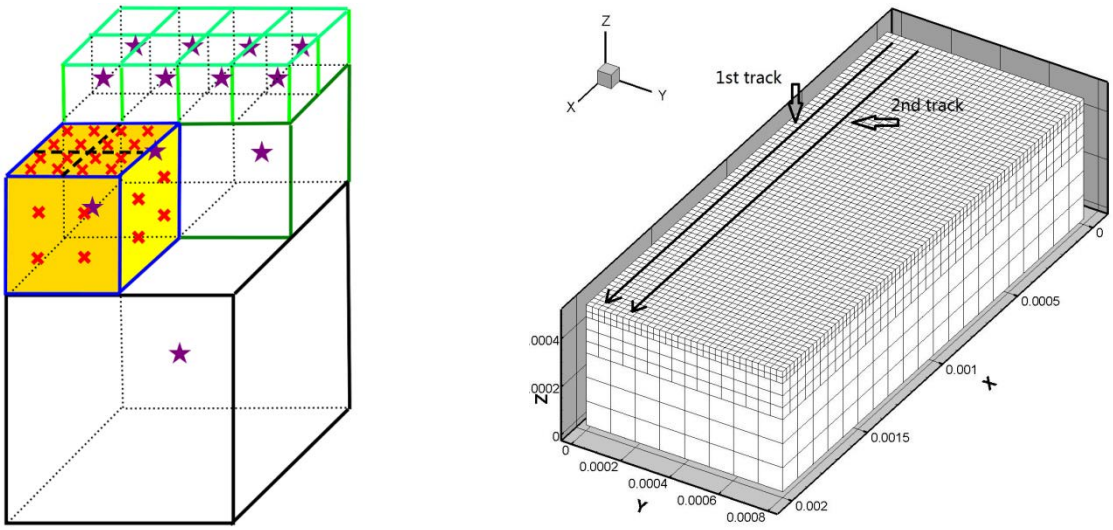
$$\tilde{N}_{I,i}(\mathbf{X}_L) = \frac{1}{A_L}\int_{\Gamma_L} \mathbf{n}N_{I,i}\,d\Gamma \tag{19}$$

The discrete form of (19) is

$$\tilde{N}_{I,i}(\mathbf{X}_L) = \frac{1}{A_L}\int_{\Gamma_L} \mathbf{n}N_{I,i}\,d\Gamma = \frac{1}{A_L}\sum_{L=1}^{NS}\sum_{G=1}^{NG} N_I(\mathbf{X}_L)n_i^L w_G \tag{20}$$

where *NS* is the number of the sides of the representative domain, *NG* the number of the gauss points per side. n_i^L the unit normal vector of the boundary. w_G the weight of the gauss points.

The whole computational domain is partitioned by background eight-node hexahedral element and each element represents an integration sub-domain (cell), as shown in Figure 2(a). The purple star is not only approximation node but cubature point, thus it is a one-point node integration scheme. The blue ridges denote the edge of background integration mesh. The four red crosses on each surface of the hexahedral cell are the quadrature points for boundary integration.



(a) (b)

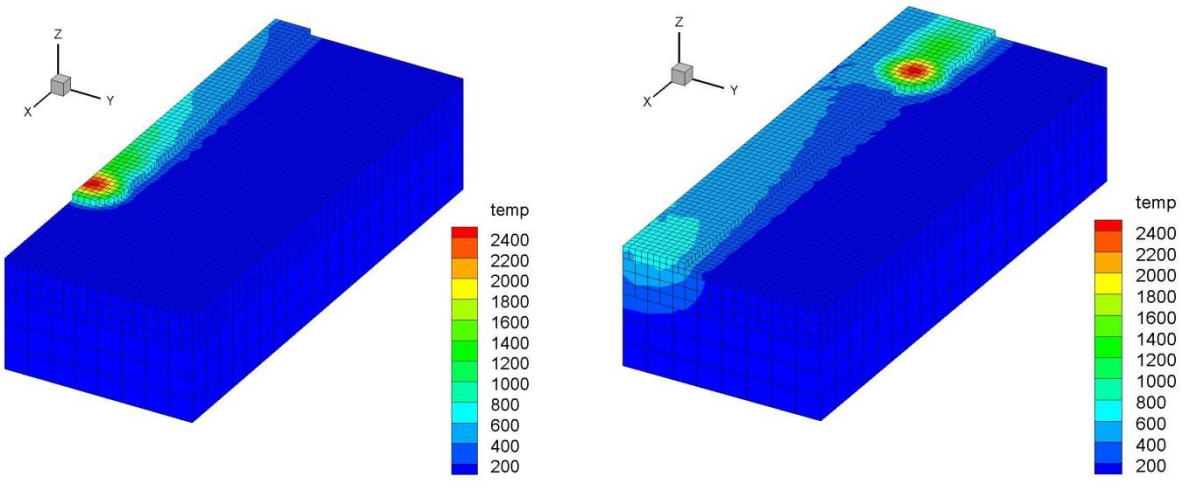
Figure 2. (a) Schematic diagram of integration schemes for SCNI. (b) Three dimensional model of SLM.

3.3 Numerical example 1

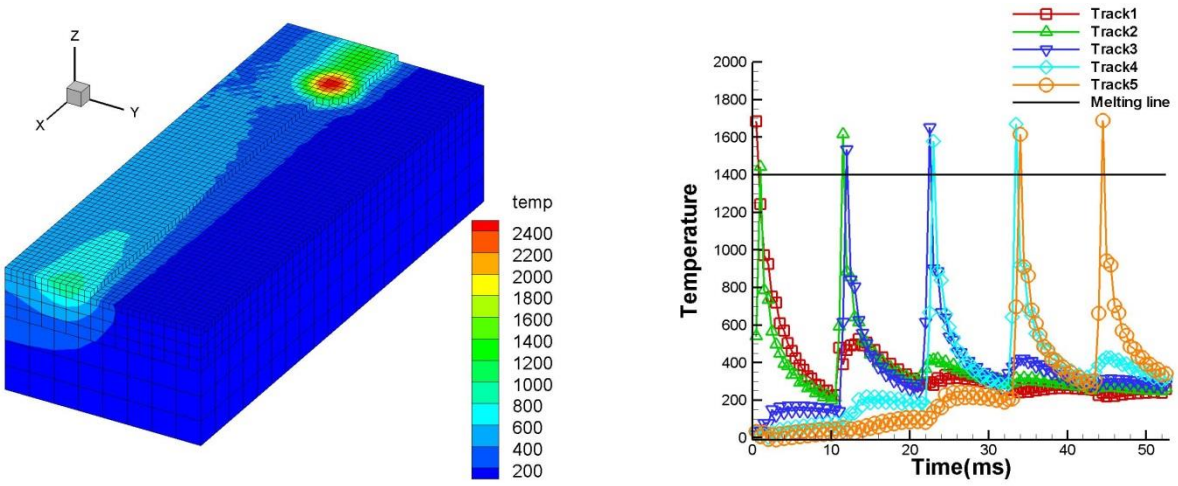
In this paper, we constructed a self-adaption coarsening model as shown in Figure 2(b) .The size of it is $2.0mm \times 0.8mm \times 0.5mm$. The upper part of the structure is divided into fine mesh. All the topmost elements are given the powder's thermophysical parameter at first. If the laser scans on the element, it will be activated in the post process, and thermophysical parameter will change. When the heat affected zone moved away from the elements, a grid coarsening operation will be done to accelerate computing. The Table 1 presents the parameters in SLM:

Table 1. Parameters for SCNI simulation

Parameter	Value
Powder layer thickness, d	0.025mm
Laser spot size, D	0.1mm
Hatch spacing, s	0.0mm
Laser power, P	100,150,200W
Scan speed, v	50,100,200mm/s



(a) (b)



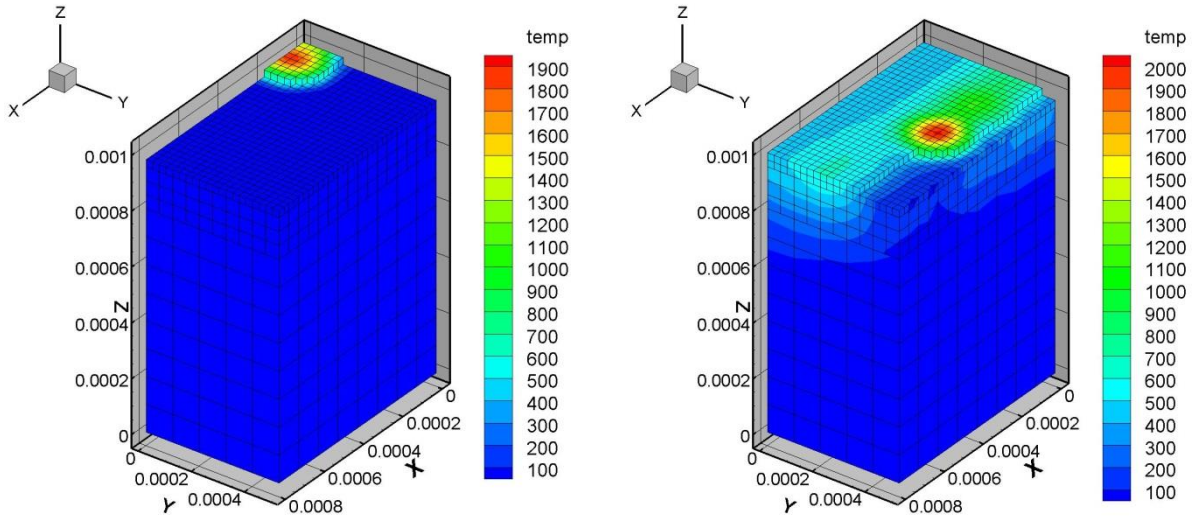
(c) (d)

Figure 3. (a) Temperature distribution at 1st track in the 20's layer. (b) Temperature distribution at 2nd track in the 20's layer. (c) Temperature distribution at 3rd track in the 20's layer. (d) Temperature history of the points in different track.

Figure 3 (a)-(c) show the temperature distribution as the laser beam reaches different positions during SLM for a scan speed of 100 mm/s and laser power of 150 W. When the first track was irradiated, the predicted maximum temperature of the powder bed was 2546 C in the region under direct irradiation, exceeding the melting point of 316L stainless steel powder(1398 C). The minimum temperature is only 25 C in the majority of the area of the substrate, At the beginning of the 2ed scan track, the maximum temperature of the structure soars to 2620 C. At the 3rd scan track, it increases to 2638 C. From the Figure 3 (d), we can also get the conclusion above. Every track melts twice during the scan, and after the 5th scan, almost the whole surface's temperature increases to 200 C.

3.4 Numerical example 2

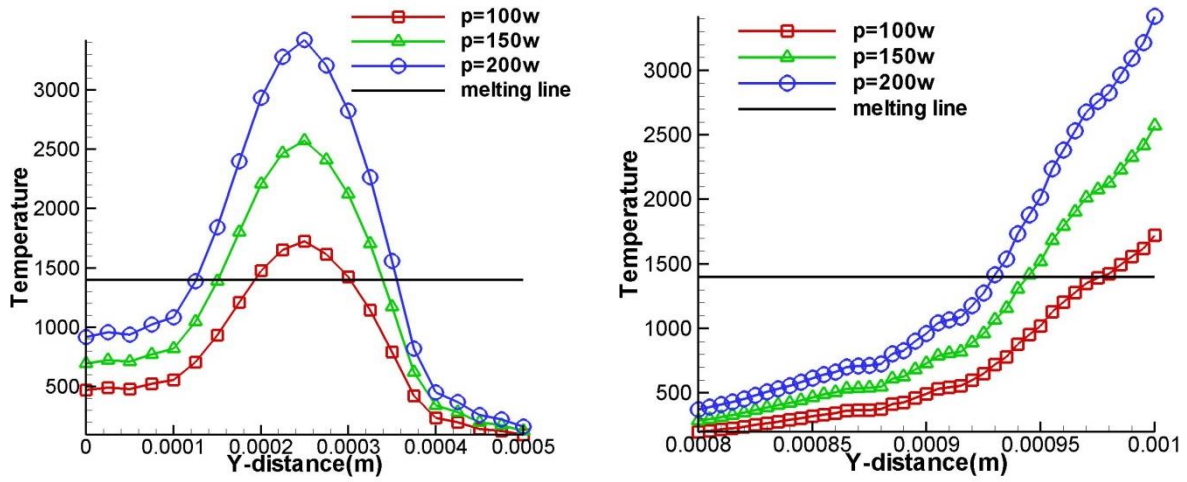
We use another model to study the scan speed and power's affect on the molten pool, the physical dimension is $0.8mm \times 0.5mm \times 1.0mm$, other parameters are the same as example1.



(a) (b)

Figure 4. (a) Scanning 1st track in the 40th layer. (b) Scanning 3rd track in the 40th layer.

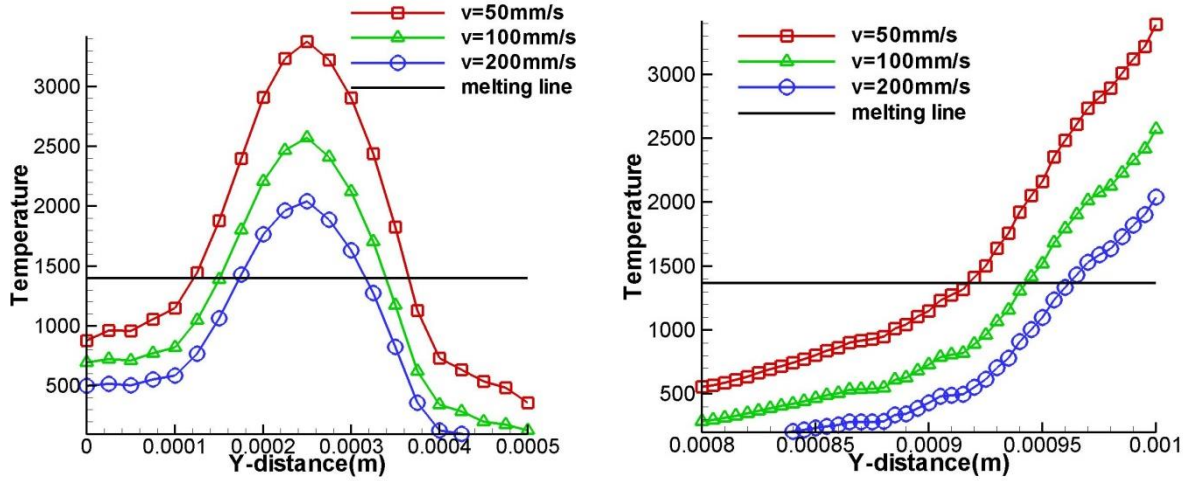
Figure 4 (a)(b) demonstrate the temperature distribution when the laser beam reaches different track in manufacturing.



(a) (b)

Figure 5. (a) Distribution of temperature along Y-axis at the molten pool of 3rd layer for different laser power. (b) Distribution of temperature along Z-axis at the molten pool of 3rd layer for different laser power.

Figure 5 exhibits the molten pool size for a given scan speed of 100mm/s as the laser power changes. If the laser power is 100W, the width and the depth are only 0.1mm and 0.025mm. When the laser power changes to 200W the width and the depth observably increase to 0.23mm and 0.072mm, respectively. It indicates that laser power has a significant influence on the molten pool.



(a) (b)

Figure 6. (a) Distribution of temperature along Y-axis at the molten pool of 3rd layer for different scan speed. (b) Distribution of temperature along Z-axis at the molten pool of 3rd layer for different scan speed.

Figure 6 shows the width and depth of molten pool for a given laser power of 150W, when the scan speed changes from 50mm/s to 200mm/s. The width and the depth are 0.15mm and 0.035mm, since the scan speed is 200mm/s. When the scan speed decreases to 50mm/s, the width and the depth markedly increase to 0.24mm and 0.073mm. Just like the laser power, Scan speed impacts the pool size obviously.

4. Conclusion

Three dimensional meshfree method is developed for predicting the temperature fields within a single metallic layer formed on the substrate in SLM process. The major findings are:

1. The proposed meshfree method is able to simulate the heat transfer in SLM.
2. Due to heat accumulation, the maximum and average temperatures gradually increase during the SLM process. And this is demonstrated in this paper.
3. Both the scan speed and power have impact on the molten pool and its size can be predicted by the developed method.

5. Acknowledgement

The authors are pleased to acknowledge the support of this work by Science Challenge Project (No. JCKY2016212A502), the National Key Research and Development Program of China (2016YFC1402705, 2016YFC1402706), the National Natural Science Foundation of China (11372066, 11232003), the Fundamental Research Funds for the Central Universities(DUT15LK07), the open funds of the state key laboratory of water resources and hydropower engineering science (2015SGG03), the open funds of the state key laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP2016K007).

References

[1] Kruth JP, Levy G, Klocke F, Child THC. (2007) Consolidation phenomena in laser and powder-bed based layered manufacturing, *CIRP Ann Manuf Technol* **56**, 730–59.
[2] Gu DD, Meiners W, Wissenbach K, Poprawe R. (2012) Laser additive manufacturing of metallic components: materials, processes and mechanisms. *IntMater Rev* **57**, 133–64.
[3] Gusarov AV, Yadroitsev I, Bertrand P, Smurov I. (2007) Heat transfer modelling and stability analysis of selective laser melting, *Appl Surf Sci* **254**, 975–9.
[4] Gusarov AV, Smurov I. (2009) Two-dimensional numerical modelling of radiation transfer in powder beds at selective laser melting, *Appl Surf Sci* **255**, 5595-5509.
[5] Gusarov AV, Smurov I. (2010) Modeling the interaction of laser radiation with powder bed at selective laser melting, *Phys Proc* **5**, 381–94.
[6] Y. Li, D. Gu. (2014) Thermal behavior during selective laser melting of commercially pure titanium powder: numerical simulation and experimental study, *Addit. Manuf*. **1**, 99–109.
[7] A. Hussein, L. Hao, C. Yan, R. Everson, (2013) Finite element simulation of the temperature and stress fields in single layers built without-support inselective laser melting, *Mater. Des*. **52**, 638–647.
[8] Nayroles B, Touzot G, Villon P. (1992) Generalizing the finite element methods: Diffuse approximation and diffuse elements, *Computational Mechanics* **10**, 307-318.
[9] Belytschko T, Lu YY, Gu L. (1994) Element-free Galerkin methods, *International Journal for Numerical Methods in Engineering* **37**, 229–256.
[10] Liu W.K., Jun S, Zhang Y.F. (1994) Reproducing kernel particle methods, *International Journal for Numerical Methods in Fluids* **20**, 1081–1106.
[11] Duarte CA, Oden JT. (1996) H-p clouds—an h-p meshless method, *Numer. Methods Partial Differ. Equations* **12**, 673-705.
[12] Zhang X. (2001) Least-Squares collocation meshless method, *International Journal for Numerical Methods in Engineering* **51**, 1089–1100.
[13] Liu GR, Gu YT and Dai K. (2004) Assessment and applications of point interpolation methods for computational mechanics, *International Journal for Numerical Methods in Engineering* 59, 1373-1397.
[14] Dolbow J, Belytschko T. (1999) Numerical integration of the Galerkin weak form in meshfree methods, *Computational Mechanics* **23**, 219–230.
[15] Beissel S, Belytschko T. (1996) Nodal integration of the element-free Galerkin method, *Computer Methods in Applied Mechanics and Engineering* **139**, 49–74.
[16] Liu GR, Zhang GY, Wang YY, Zhong ZH, Li GY, Han X. (2007) A nodal integration technique for meshfree radial point interpolation method (NI-RPIM) , *International Journal of Solids and Structures* **44**, 3840–3860.
[17] Dyka CT, Randles PW, Ingel RP. (1997) Stress points for tension instability in SPH, *International Journal for Numerical Methods in Engineering* **40**, 2325–2341.
[18] Fries TP, Belytschko T. (2008) Convergence and stabilization of stress-point integration in mesh-free and particle methods, *International Journal for Numerical Methods in Engineering* **74**, 1067–1087.

[19] Liu Y, Belytschko T. (2010) A new support integration scheme for the weak form in mesh-free methods, *International Journal for Numerical Methods in Engineering* **82**, 699–725.
[20] Chen JS, Wu CT, Yoon S, You Y. (2001) A stabilized conforming nodal integration for Galerkin mesh-free methods, *International Journal for Numerical Methods in Engineering* **50**, 435–466.
[21] Liu G, Dai K, Nguyen T. (2007) A smoothed finite element method for mechanics problems, *Computational Mechanics* **39**, 859-877.
[22] Thummler F, Oberacker R. (1993) An introduction to powder metallurgy, Cambridge (London): The University Press.
[23] Dai K, Shaw L. (2005) Finite element analysis of the effect of volume shrinkage during laser densification, *Acta Mater* **53**, 4743–54.

Numerical simulation of turbulent flows in a channel with a series of groynes by ZDES

***†J.X. Zhang[1,2]**

[1]MOE Key Laboratory of Hydrodynamics, Shanghai Jiao Tong University, China
[2]School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, China

*Presenting author: zhangjingxin@sjtu.edu.cn
†Corresponding author: zhangjingxin@sjtu.edu.cn

Abstract

In the present work, the hydrodynamics in a straight open channel constructing a multiple-embayment groyne field on one of its sides is investigated by numerical simulations. The open channel is composed of a flat bottom in the main channel and a gradual slope toward one of the sides with a ratio 1:3. A series of five embayment are constructed on the slope, and the ratio of the embayment's length to width is specified as 3. A thorough analysis of the three-dimensional flow characteristics was carried out by Zonal detached-eddy simulation (ZDES) model. The numerical simulation highlights the turbulent coherent structures induced by groynes, including local horseshoe vortex (HV) around the groyne tip, shedding vortex (SV) from the groyne and the mixing layer flow at the channel-embayment interface. Both the instantaneous and the mean flows show a spatial evolution of turbulence in the main channel and embayment, which reveals different coherent structures from the upstream embayment to the downstream embayment. The analysis of the large-scale eddies, which populates the mixing between the channel and embayment, are contributed to investigate the mechanisms of fluid mixing, mass transportation and exchange, and some issues about hydraulic engineering.

Keywords: Hydrodynamics, channel flow, groynes, Zonal-DES, free surface flow

Introduction

Groynes are commonly built with an angle to the main flow to fulfill multiple objectives. One of the most usual targets is to maintain the channel navigability by keeping the flow away from the banks and increasing the velocity in the channel as well as increasing the energy of sediment transportation. There are also other objectives, for example, restoring fish habitat by degrading the river current [1]. Because the embayment region between successive groynes acts as a dead water zone, the fluid and dissolved mass exchange ratio decreases and the residence times of matter are much larger, which raises an eco-hydraulic problem [2].

The turbulence coherent structures are dominated by the channel geometry and groyne shapes, including the groyne orientation, space of the embayment, emerged, submerged, permeable and the groyne tip shape [3]-[4]. On the other hand, the large-scale eddies populate the fluid mixing between the main channel and embayment. Most of the knowledge about the dynamics of turbulent flows in groyne flows comes from experimental scale-model researches by means of experimental techniques to capture the mean velocity and instantaneous turbulence quantities [5]-[6].

Besides of experimental researches, the numerical model is much more efficient to capture some quantitative hydrodynamic factors of the groyne field flows. The early numerical simulation works were based on the RANS models [7]-[8], which was contributed to the mean recirculation simulations. Large eddy simulation (LES) method is prior to the RANS model in predicting large-scale eddies, which is considered as the key dynamic force

dominating the mass transportation and exchange in the groyne area [9]-[11]. Much more hydrodynamic mechanisms have been highlighted by LES, but the high computational cost hinders the practical applications in large-scale channel flows, for example the natural river flows. A family of hybrid RANS/LES model, named as detached-eddy simulation (DES) is practical to promote the high-resolution numerical simulations in practical hydraulic engineering simulation.

In the present paper, a Zonal detached-eddy simulation (ZDES) model is developed to simulate free surface flows [12], and then is employed to study the turbulent flows in a channel composed of a flat bottom in the main channel and a slope bottom along one side with a series of emergent groynes constructed on the slop bottom with a constant ratio of the embayment length to width. It is feasible using ZDES to investigate the 3D flow features referring to turbulent fluctuations, distributions of turbulent energy, and the bed shear stress distribution.

Numerical method and simulation setup

Governing Equations

In the present 3D numerical model, the total pressure is split into the hydrostatic component $p_h = \rho g(\zeta - z)$ and non-hydrostatic component p_n. The 3D governing equations are written in a conversation form:

$$\frac{\partial \zeta}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_\sigma}{\partial \sigma} = 0 \tag{1}$$

$$\frac{\partial q_x}{\partial t} + \frac{\partial q_x u}{\partial x} + \frac{\partial q_x v}{\partial y} + \frac{\partial q_x \tilde{\omega}}{\partial \sigma} = -gD\frac{\partial \zeta}{\partial x} - \frac{D}{\rho_0}\frac{\partial p_n}{\partial x} + \frac{\partial}{\partial x}\left(\upsilon_t \frac{\partial q_x}{\partial x}\right) + \frac{\partial}{\partial y}\left(\upsilon_t \frac{\partial q_x}{\partial y}\right) + \frac{1}{D}\frac{\partial}{\partial \sigma}\left(\frac{\upsilon_t}{D}\frac{\partial q_x}{\partial \sigma}\right) \tag{2}$$

$$\frac{\partial q_y}{\partial t} + \frac{\partial q_y u}{\partial x} + \frac{\partial q_y v}{\partial y} + \frac{\partial q_y \tilde{\omega}}{\partial \sigma} = -gD\frac{\partial \zeta}{\partial y} - \frac{D}{\rho_0}\frac{\partial p_n}{\partial y} + \frac{\partial}{\partial x}\left(\upsilon_t \frac{\partial q_y}{\partial x}\right) + \frac{\partial}{\partial y}\left(\upsilon_t \frac{\partial q_y}{\partial y}\right) + \frac{1}{D}\frac{\partial}{\partial \sigma}\left(\frac{\upsilon_t}{D}\frac{\partial q_y}{\partial \sigma}\right) \tag{3}$$

$$\frac{\partial q_z}{\partial t} + \frac{\partial q_z u}{\partial x} + \frac{\partial q_z v}{\partial y} + \frac{\partial q_z \tilde{\omega}}{\partial \sigma} = -\frac{1}{\rho_0}\frac{\partial p_n}{\partial \sigma} + \frac{\partial}{\partial x}\left(\upsilon_t \frac{\partial q_z}{\partial x}\right) + \frac{\partial}{\partial y}\left(\upsilon_t \frac{\partial q_z}{\partial y}\right) + \frac{1}{D}\frac{\partial}{\partial \sigma}\left(\frac{\upsilon_t}{D}\frac{\partial q_z}{\partial \sigma}\right) \tag{4}$$

where g is gravitational acceleration, ζ is the free surface elevation. When the non-hydrostatic component p_n is ignored, the vertical momentum equation is also neglected, and the system degenerates to the common shallow water equations with only two horizontal momentum equations and one continuity equation, which is also named as hydrostatic model. In order to fit the free surface and the uneven bottom boundary, the vertical coordinate z is transformed to the σ coordinate. In the governing equations. The new variables $q_x = Du, q_y = Dv, q_z = Dw, q_\sigma = D\tilde{\omega}$ are introduced instead of the velocity u, v and w, and the vertical velocity in σ coordinate is calculated by the following formula:

$$q_\sigma = \frac{q_z}{D} - \frac{q_x}{D}\left(\sigma\frac{\partial D}{\partial x} + \frac{\partial \zeta}{\partial x}\right) - \frac{q_y}{D}\left(\sigma\frac{\partial D}{\partial y} + \frac{\partial \zeta}{\partial y}\right) - \left(\sigma\frac{\partial D}{\partial t} + \frac{\partial \zeta}{\partial t}\right) \tag{5}$$

Turbulence model

The eddy viscosity coefficient υ_t is determined based on the one-equation S-A model. The working variable $\tilde{\upsilon}$ obeys the following transport equation

$$\frac{D\tilde{\upsilon}}{Dt} = c_{b1}\tilde{S}\tilde{\upsilon} - c_{w1}f_w\left[\frac{\tilde{\upsilon}}{d}\right]^2 + \frac{1}{\tilde{\sigma}}\left\{\nabla\cdot\left[\left(\upsilon+\tilde{\upsilon}\right)\nabla\tilde{\upsilon}\right] + c_{b2}\left(\nabla\tilde{\upsilon}\right)^2\right\} \tag{6}$$

The eddy viscosity υ_t is given by

$$\upsilon_t = \tilde{\upsilon} f_{v1} \tag{7}$$

where $f_{v1} = \frac{\chi^3}{\chi^3 + c_{v1}^3}$, $\chi \equiv \frac{\tilde{\upsilon}}{\upsilon}$. υ is the molecular viscosity. Here $|\bar{S}|$ is the magnitude of the strain rate tensor, $\tilde{S} = |\bar{S}| + \frac{\tilde{\upsilon}}{\kappa^2 d^2} f_{v2}$, $f_{v2} = 1 - \frac{\chi}{1+\chi f_{v1}}$.The function f_w is calculated as

$f_w = g\left[\frac{1+c_{w3}^6}{g^6 + c_{w3}^6}\right]^{1/6}$, $g = r + c_{w2}\left(r^6 - r\right)$, $r \equiv \frac{\tilde{\upsilon}}{\tilde{S}\kappa^2 d^2}$. The constants in the above equations are $c_{b1} = 0.1355$, $\tilde{\sigma} = 2/3$, $c_{b2} = 0.622$, $\kappa = 0.41$, $c_{w1} = c_{b1}/\kappa^2 + \left(1+c_{b2}\right)/\tilde{\sigma}$, $c_{w2} = 0.3$, $c_{w3} = 2.0$, $c_{v1} = 7.1$.

The DES model concerns the destruction term in the transport Equation (6), and the distance d to the solid wall is replaced by a modified length $\tilde{d}$. In the present model, the unstructured grid is used in the horizontal plane and a structured sigma-grid is used in the vertical direction, and the $\tilde{d}$ is calculated by

$$\tilde{d} = \min\left(d, C_{DES}\Delta\right), \text{with } \Delta = \max\left(\sqrt{4A/\pi}, \Delta z\right) \tag{8}$$

where A is the horizontal mesh area; C_{DES} is 0.65. There is a layer near the solid wall where $\tilde{d} = d$, which is referred to as the traditional "RANS" zone, and the simulation is turned into "LES" zone above this layer where $\tilde{d} = C_{DES}\Delta$.

The model-stress depletion (MSD) is one severe problem of the RANS/LES hybrid model, which is pronounced for the DES when the switch from RANS to LES takes place in the boundary layer, so Spalart [13] proposed a new version of DES, i.e. DDES, to delay the switch from RANS to LES. Shur et al. [14] developed a new model, named IDDES, for the simulation with an ambiguous grid scale, which can solve the MSD problem by modifying the turbulent length scale, initial condition and inflow condition. Keating et al. [15] presented a dynamic stochastic forcing method, which significantly speeds up the transition resulting in more accurate predictions of the velocity fluctuation. The Zonal-DES approach is well adopted to handle separated flows in which strong instabilities develop rapidly, thus overwhelming the turbulence inherited from upstream boundary layers [16]-[17]. In the present model, a Zonal-DES model is developed by means of modifying the functions f_{v1}, f_{v2} and f_w as shown in (9) and adopting a new sub-grid length scale $\Delta = \left(A\Delta z\right)^{1/3}$ in the LES computational domain.

$$f_{v1} = 0,\ f_{v2} = 0,\ f_w = 1 \tag{9}$$

Numerical scheme

A predictor-corrector scheme is used in the numerical method [18]. In summary, the flow driven by hydrostatic pressure is calculated as the predictor step, and then the flow driven by non-hydrostatic pressure is further resolved in the corrector step. The grid system is composed of unstructured horizontal grids and multi-layers in the vertical direction. The governing equations are discretized based on finite volume method (FVM). The 2nd-order

total variation diminishing (TVD) scheme, sometimes referred to as the modified-flux approach. In the present numerical model, the OSHER scheme is adopted. In the corrector step, the Poisson-type equation for the non-hydrostatic pressure is numerically obtained by the pre-conditioned BI-CGSTAB approach. The in-house codes are paralleled by OpenMP library.

Simulation setup

The sketch of the computational domain is shown in Fig.1, which is composed of a flat main channel bottom and a slope bottom along one side. The slope bottom extends from the bank into the main channel with a constant slope 1 in 3, which is also revealed in Fig.2. The groynes are emerged and perpendicular to the main flow direction. The depth in the main channel $D = 0.2m$ was chosen as the length scale. The mean velocity $U = 0.22m/s$ in the main channel was chosen as the velocity scale. The Reynolds number was $\mathrm{Re} = UD/\upsilon = 44,000$, and Froude number $Fr = U/\sqrt{gD} = 0.2$. In the computational domain, x = stream wise direction, z = vertical direction (originating on the still water level), and in the transversal direction, $y = 0$ corresponds to the sidewall. The domain extends upstream 20D of the first embayment and 50D downstream of the last embayment. The thickness of the groynes is 0.25D. The width (3D) over length (9D) ratio of each embayment is 1/3. The depth in the embayment area increases from 0.15D at the sidewall to D with a slope of 1:3, and the horizontal length of the slope in the embayment is 2.5D, which is shorter than the groyne length 3D (Fig.2). The length scale of the embayment guarantees the applicability of the SWE (shallow water equation) in hydrostatic mode, but the groyne tip and the presence of the change of depth between the main channel and the embayment are likely to induce strong 3D effects. Consequently, the present non-hydrostatic model is needed for the flow simulations.

In ZDES by means of S-A model as it stands in the above section, the grid scale is strictly limited in the wall units: $\Delta_\perp^+ = O(1)$ at the wall. We maintained $\Delta_\perp^+ \approx 1$ between the first grid point to the nearest wall, and stretched the grid size with a ratio of 1.15 away the solid wall until the grid size meets the criterion for LES simulations. For open channel flows, Hinterberger et al.[10] proposed a valuable reference grid size in LES. Because the large-scale eddies in open channel flow are limited by the characteristic length, i.e. the water depth h, the grid size was properly designed in the range of $(\frac{1}{20} \square \frac{1}{10})h$. The total number of the meshes is about 9 million.

The free surface boundary can be captured by the stretching σ grid, which is high efficient in the simulation of free surface motion. The sidewall and the bottom were considered as non-slip boundaries, and the lateral wall opposite the groynes was considered as slip boundary. The grid scale was designed to ensure the local LES domain only covering the embayment, and beyond the LES area, the RANS domain extending to the inlet and outlet boundaries. A steady inflow boundary condition was used, and the flow in the inlet section was fully turbulent obtained by means of a precalculated fully-developed open channel flow simulation. In the outlet section, a fixed water level was specified and a convective flux condition was used in the discretizing the momentum equations.

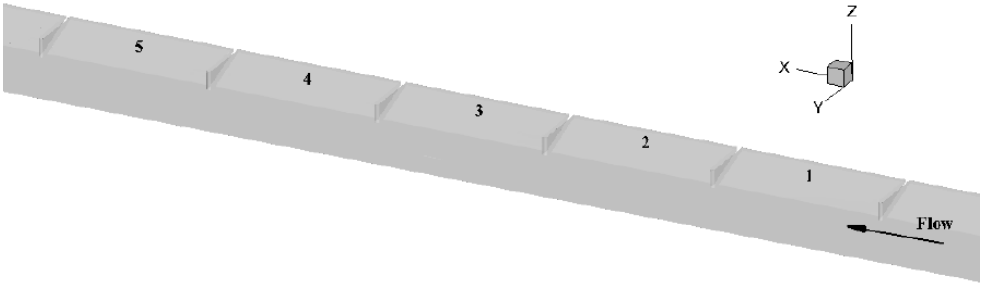
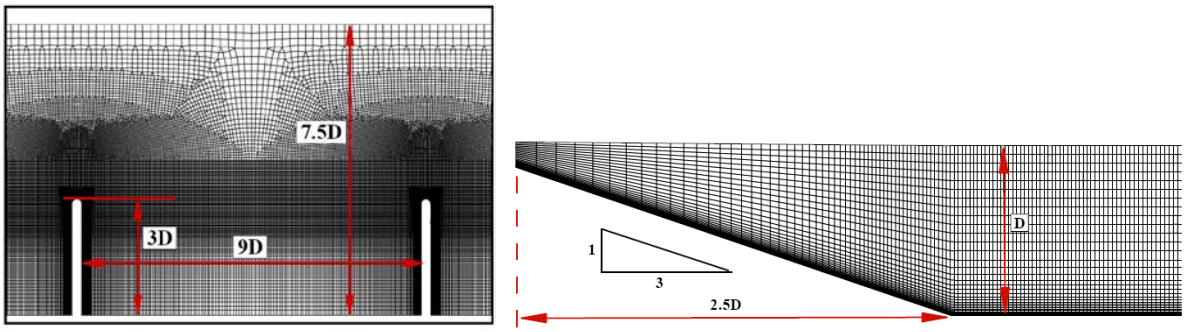


Fig. 1 Sketch showing computational domain with five shallow embayment.



(a) (b)

Fig.2 Computational mesh: (a) mesh in a horizontal plane; (b) mesh in a vertical cross section

Results and Discussion

The simulation was firstly run until the transients were eliminated, i.e. being fully turbulent. Statistics were further calculated using the instantaneous flow fields over the next $120D/U$. The present work was focused on the evolution of the turbulent flows from the first groyne to downstream. And the instantaneous and mean flow were analyzed to investigate the hydrodynamics.

Instantaneous flow

The instantaneous free surface reveals the evolution of the turbulent flows induced by the groynes on the uneven bottom. Fig.3 shows a snapshot of the free surface. The vortical structures beneath the free surface deform the geometry of free surface. Surface patterns, such as upwelling, downdraft, and ripples were predicted in the present simulations. A series of regular surface downdraft are observed justly downstream of the first groyne, and these downdraft surface motion are gradually weakened in the following embayment. This surface pattern corresponds to the shedding vortex originated from the tip of the groynes. At the channel-embayment interface, a large-scale ripple structure is observed, which dominates the fluid exchange between the main

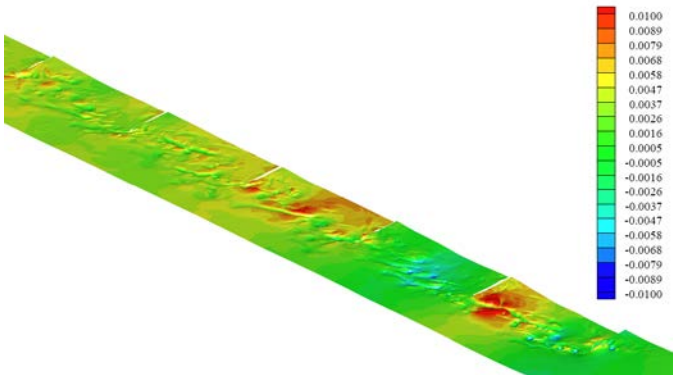


Fig.3 The instantaneous free surface

channel and the embayment. The present simulations show no identical free surface pattern until the last groyne.

It is expected that small-scale vortical structures are predicted by the ZDES model. A positive isovalue of criterion Q was used to show the turbulent structures, which defines as the vortex tubes in the regions where the second invariant of the velocity gradient tensor is positive. Fig.4 shows the instantaneous Q (= 6) in the groyne area, in which the small-scale vortical structures develop very quickly originated from the first groyne. In order to investigate the details of the coherent structures, the local vortical structures identified by isovalue of Q in embayment 1 and embayment 3 are presented in Fig.5. At instantaneous $t = t_0$, there are mainly three distinct vortical structures around the fire groyne. The necklace-like vortices (NV) originate from the upstream groyne face near the free surface. The organized shedding vortex (SV) is observed beneath the free surface, and the horseshoe vortex (HV) is formed at the basement of the groyne. The SV gradually decays downstream of the first groyne. The near surface vortical streak becomes oriented relatively parallel to the free surface, which corresponds to the mixing layer induced by the upstream groyne. At instantaneous $t = t_0 + 3D/U$, the NV degenerates, but the HV and the SV become intensity and develop gradually downstream. At instantaneous $t = t_0 + 5D/U$, the SV continually develop downstream along with merging and decaying with ambient vortices . The large-scale HV breaks down during the interaction with small-scale vortices. In embayment 3 area, the similar vortical structures are observed in the snapshot of the flow field shown in Fig.5, but the SV and HV at the basement are not obviously formed, which is likely because of the attenuation of the incident flows. However, a large-scale vortical structure penetrating into the embayment is observed. The instantaneous coherent structures of the turbulent flow are corresponded to the flow separating at the tip of the groynes and the mixing layer flow formed by the high speed channel flows and the low speed embayment flows at the interface. Conversely, the large-scale vortices predominate the mass exchange between the main channel and the embayment, the local scouring and the suspended sediment transportation.

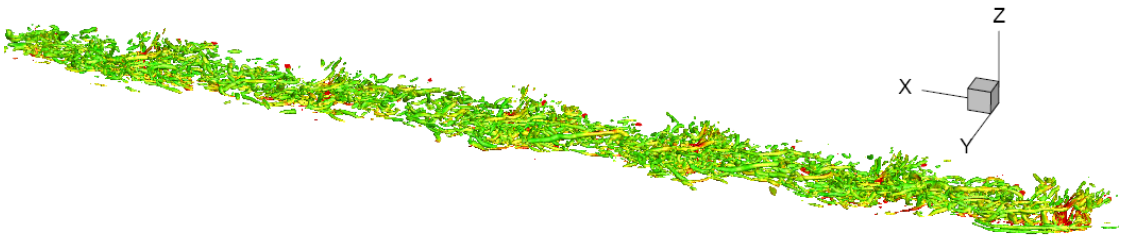
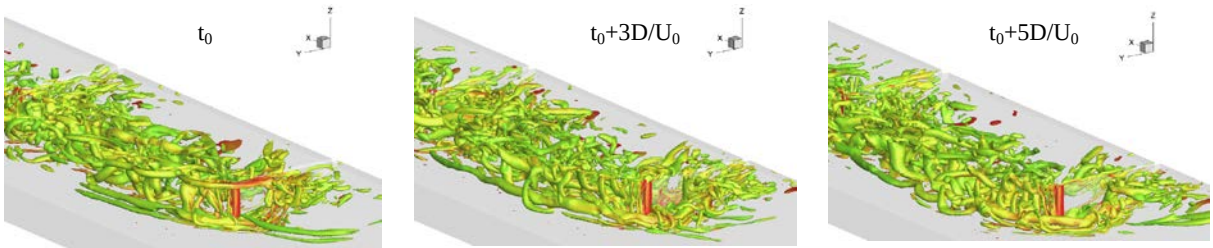
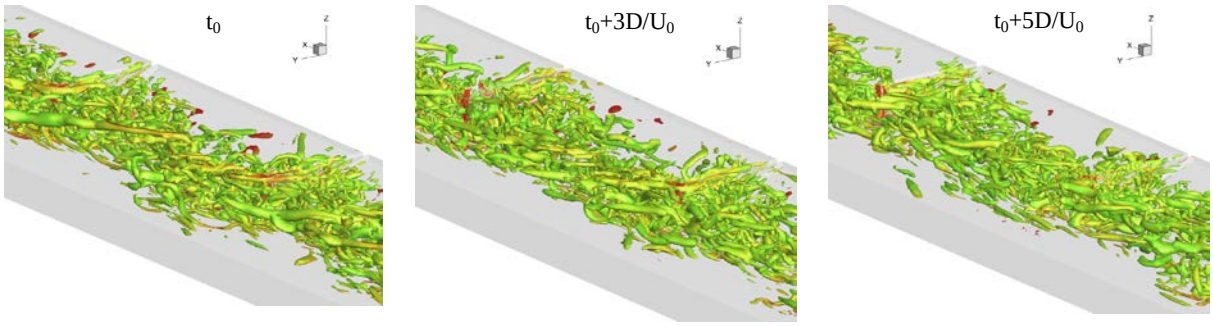


Fig.4 The coherent structures of turbulent flows identified by Q flooded by vorticity.



(a) Embayment 1



(b) Embayment 3

Fig. 5 The coherent structures of the turbulent flow identified by Q: (a) Embayment 1, and (b) Embayment 3.

Time-averaged mean flow

The mean flows were calculated using the instantaneous flow fields over $120D/U$. The random small-scale vortices are disappear in the mean velocity field, and the related steady flow patterns contribute to analyze the long-term hydrodynamics in the channel flows with groynes.

Fig.6 shows the rough mean velocity streamlines in the computational domain. A one-gyre circulation is observed located in the corner in front of the first groyne, and the local streak is highly three dimensional. A two-gyre circulation pattern is distinctly observed in embayment 1, embayment 3, and embayment 5, but the similar recirculation is not clearly observed in the other embayment. Fig.6 reveals that the identical flow pattern in the embayment has not fully developed.

The mean velocity streamlines present the long-term fluid particle's kinematic path, and can be qualitatively used to analyze the mean mass exchange between different reaches. Fig.7 (a) clearly shows a two-gyre recirculation in embayment 3. Affected by the change bottom topography, the mean recirculation reveals a distinct 3D pattern rather than a 2D pattern commonly observed in shallow water channel flows. One single streamline originated from the upstream is elaborately selected to show the path of one fluid particle motion (Fig.7 (b)), which roughly presents one possible mass transport route. The scale of the secondary gyre at the corner is commonly smaller than that observed in groyne flow with the similar ratio of length to width but on flat bottom.

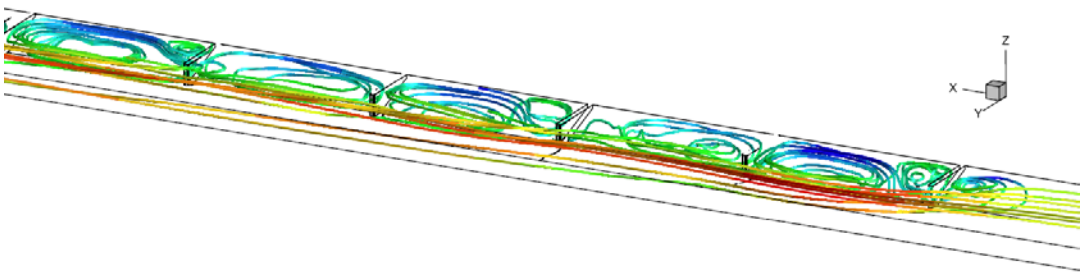
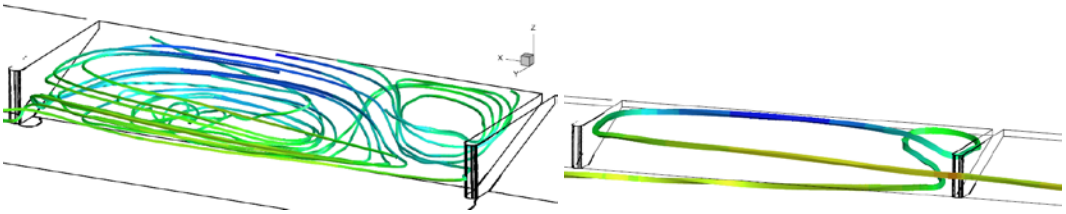


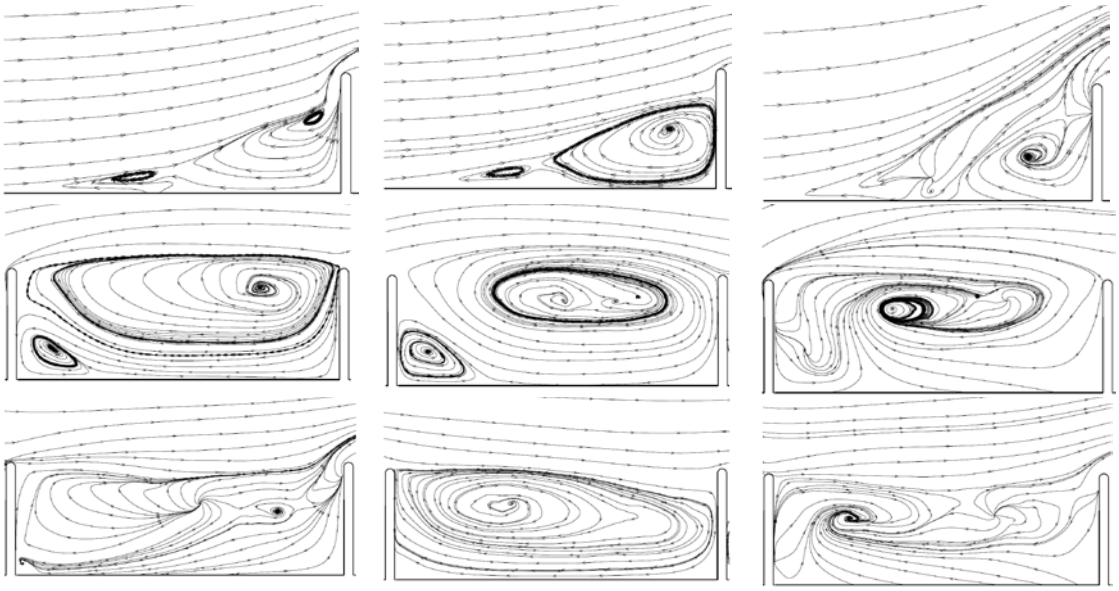
Fig. 6 Mean velocity streamlines flooded by stream wise velocity.

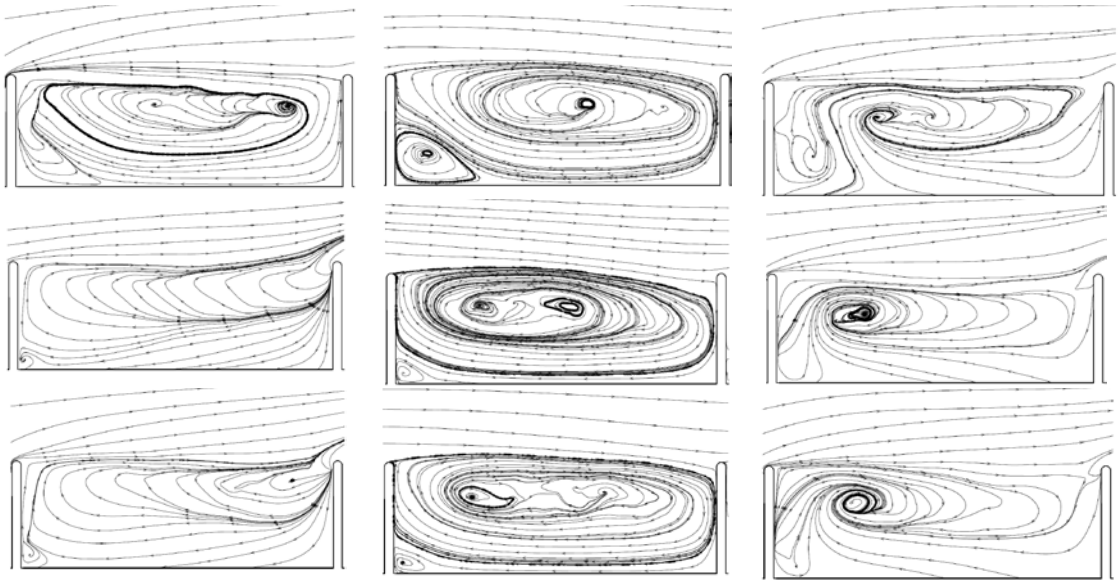


a) mean velocity streamlines b) one selected velocity streamline

Fig.7 Local mean velocity streamlines showing the recirculation in embayment 3.

Fig.8 reveals the mean velocity streamlines in the horizontal plane at different vertical position, i.e. near the surface ($\sigma = -0.05$), at the mid-depth ($\sigma = -0.5$), and near the bottom ($\sigma = -0.95$). $\sigma \in [-1, 0]$ presents the relative depth. The panels from top to bottom in Fig.8 correspond to the embayment in the stream wise. The mean velocity streamlines at the mid-depth in front of the first groyne reveal a large-scale clockwise recirculation and a small-scale unclockwise recirculation upstream of it. The recirculation pattern near the free surface is a bit different, especially in the area very near to the tip of the groyne. The surface ripples likely effect the local flows shown in Fig.8. Near the bottom, some complex structures presented by saddle points, nodes and spiral or focus points are observed. The details of the characteristic streamlines or singular points clearly describe the flow structures, for example, the separation of the boundary flows. In the present simulation, an identical flow pattern have not developed until the last groyne, which can be observed from the streamlines not only in the near surface plane, but also the mid-depth plane or the adjacent bottom plane. Justly analyzing the recirculation in the mid-depth plane shown in the middle column panels in Fig.8, the size of a two-gyre is different among the five embayment. The relative size of the secondary gyre to the main one is about identical in embayment 4 and 5, which is roughly considered as an identical flow pattern. The two-gyre flow pattern is affected by the incident flow condition and the geometry of the computational domain. A periodic open boundary condition is commonly used in numerical simulations of flow passing one embayment, in which the computational cost decreases because of the smaller computational domain. The simulation of the flow passing series of groynes give rise to the selection of the periodic boundary position.

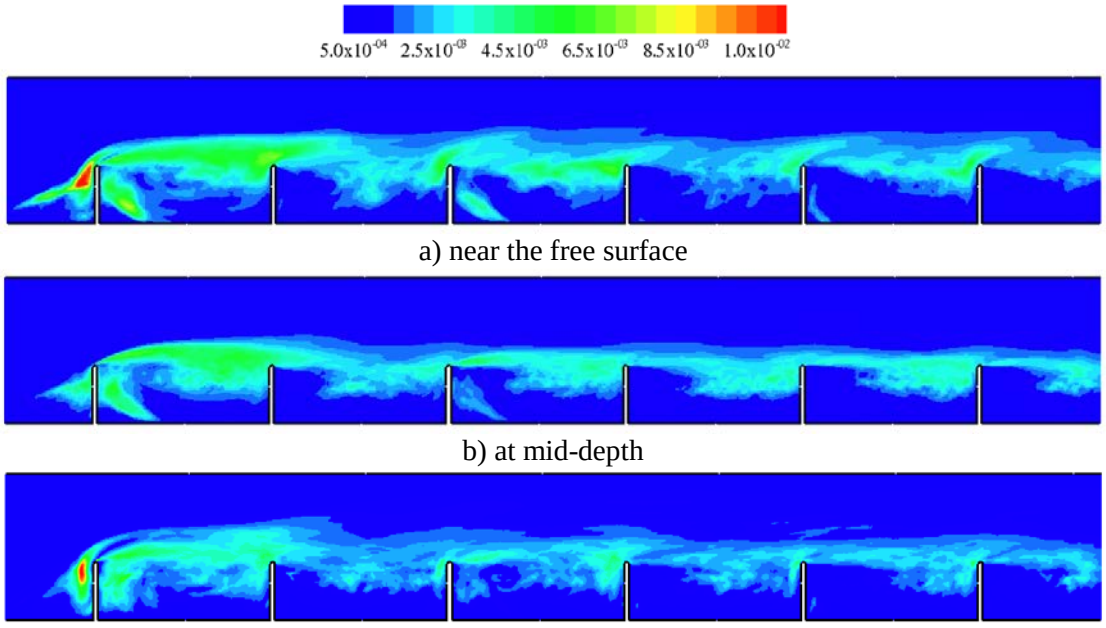




a) near the free surface b) at mid-depth c) near the bottom

Fig. 8. Mean velocity streamlines in horizontal planes along the groynes.

Fig.9 shows the distribution of the mean turbulent kinetic energy (TKE) in the horizontal planes from the surface to the bottom. The maximum value occurs near the tip of the first groyne near the free surface and near the basement, which reveals relative stronger flow fluctuations compared to the flow in the mid-depth. The higher level of TKE in embayment 1 corresponds the stronger incident flow and the sharp distortion of the local streamlines around the frontal groyne tip, and the next groynes are sheltered in the downstream flows. Because of the weakened incident flow and the little distortion of the local streamlines around the groyne tip, the level of TKE decreases downstream.



a) near the free surface

b) at mid-depth

c) near the bottom

Fig. 9 The mean turbulent kinetic energy in horizontal planes.

Conclusions

A 3D non-hydrostatic numerical model was developed and used to investigate the channel flow with a series of groynes constructed on a slope along one side. The ZDES was used to simulate the turbulence by aid of its superiority in predicting small-scale vortical structures compared to RANS model. The simulations deeply investigate the turbulent coherent structures induced by the groynes and the composited topography. The attention were focused on the spatial evolution of the vortices not only in the streamwise, but also in the vertical direction. The analysis of the coherent structures is benefit to investigate the mechanism of the hydrodynamic forces and the mass advection or exchange across the chaneel-embayment interface.

Acknowledgements

This work was jointly sponsored by the National Natural Science Foundation of China (No. 11572196) and the National Basic Research Program of China (973 Program, No. 2014CB046200).

References

[1] Shields, F.D., Cooper, C.M., and Knight, S.S. (1995). Experiments in stream restoration. J. Hydrauli.Eng., 121, 494-502.
[2] Engelhardt, C., Kruger, A., Sukhodolov,A., and Nicklisch, A. (2004). A study of phytoplankton spatial distributions, flow structure and characteristics of mixing in a river reach with groynes. J. Plankton Res., 26, 1351-1366.
[3] Sukhodolov A., Uijttewaal W.S.J., and Engelhardt C. On the correspondence between morphological and hydrodynamical patterns of groyne fields. Earth Surface Processes and Landforms, 2002, 27: 289-305.
[4] Weitbrecht V., Scocolofsky S.A., and Jirka G.H. Experiments on mass exchange between groin fields and main stream in rivers. Journal of Hydraulic Engineering, 2008, Vol. 134, No. 2, pp. 173-183.
[5] Uijttewaal W.S.J., Lehmann D., and van Mazijk A. Exchange processes between a river and its groyne fields: model experiments. Journal of Hydraulic Engineering, 2001, Vol.127, No.11, pp. 928-936.
[6] Uijttewaal W.S.J. Effects of groyne layout on the flow in groyne fields: laboratory experiments. Journal of Hydraulic Engineering, 2005, Vol.131, No.9, pp.782-791.
[7] Peng, J., Kawahara, Y., and Huang, G. (1999a). Evaluation of modified k-ε models in simulating 3D flows over submerged spur dikes. Proc., Turbulence and Shear Flow-1, First Int. Symp., S. Banerjee and J.K. Eaton,eds., Santa Barbara, Calif.
[8] Peng, J., Tamai, N., Kawahara, Y., and Huang, G. (1999b). Numerical modeling of local scour around spur dikes. Proc., 28th IAHR Congress, Graz, Austria.
[9] Uijttewall, W., and van Schijndel, S.A.H. (2004). The complex flow in groyne fields: Numerical modeling compared with experiments. Proc. River Flow 2004, Naples, Italy, 1331-3838.
[10] Hinterberger, C. (2004). Three-dimensional and depth-average large eddy simulation of shallow water flows. Ph.D. thesis, karlsruhe Univ., Karlsruhe, Germany.
[11] McCoy, A., Constantinescu G., and Weber L.J. Numerical investigation of flow hydrodynamics in a channel with a series of groynes. Journal of Hydraulic Engineering, 2008, Vol. 134, No.2, pp. 157-172.
[12] Zhang, J.X., Wang X.K., Liang D.F., and Liu H. (2015). Application of detached-eddy simulation to free surface flow over dunes. Engineering Applications of Computational Fluid Mechanics, 2015, Vol.9, No.1, 556-566.
[13] Spalart, P.R., Deck, S., Shur, M.L., Squires, K.D., Strelets, M.K., and Travin, A., (2006). A new version of detached-eddy simulation, resistant to ambiguous grid densities. Theoretical Computational Fluid Dynamics, 20: 181-195.
[14] Shur K.L., Spalart P.R., Strelets M.K., and Travin A.K., (2008). A hybrid RANS-LES approach with delayed-DES and wall-modelled LES capabilities. International Journal of Heat and Fluid Flow, 29: 1638-1649.
[15] Keating A., and Piomelli U., (2006). A dynamic stochastic forcing method as a wall-layer model for large-eddy simulation. Journal of Turbulence, 7(12): 1-24.
[16] Deck S.,(2005a). Numerical simulation of transonic buffet over a supercritical airfoil. AIAA Journal, 43(7): 1556-1566.
[17] Deck S., (2005b). Zonal-detached eddy simulation of the flow around a high-lift configuration. AIAA Journal, 43(11): 2372-2384.
[18] Zhang, J.X., Sukhodolov, A.N., and Liu, H.(2014). Fully hydrodynamic versus hydrostatic modeling for shallow environmental flows, J.of Hydrodynamics, 26(4): 840-847.

The fully coupled effects of FPSO with different filling ratio tanks in CFD method

Yuan Zhuang[*], Decheng Wan[‡]

State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

[*]Presenting author: nana2_0@sjtu.edu.cn

[‡]Corresponding author: dcwan@sjtu.edu.cn

http://dcwan.sjtu.edu.cn

Abstract

With the increasing demand of oil and natural gas, the FPSO with LNG tanks was proposed as a popular type LNG platform. Equipped with liquid natural gas so that the tanks has free surface during the operation, FPSO will suffer from sloshing forces and moments from inner tanks. At the meantime, the sloshing flow in tanks is influenced by the ship motion in return. This paper applied numerical methods to study the coupling effects between FPSO and sloshing tanks. The coupling effects were solved in CFD method, using our in house solver naoe-FOAM-SJTU, based on open package toolbox OpenFOAM. The outer field and inner free surface were both simulated by solving Navier-Stocks equation thus the ship motion and sloshing tanks were fully coupled. The two LNG tanks were in different filling conditions, and the sloshing forces and moments were considered in current study.

Keywords: FPSO LNG; sloshing; naoe-FOAM-SJTU solver; fully coupled analysis

Introduction

For the oil and natural gas occupied the main position during the production in recent years, the demand of these energy sources calls for the need of deep sea platform, which has the ability to exploit gas field and process the crude oil in deep water. As a result, FPSO (Floating Production Storage and Offloading) was proposed as a new type of floating LNG (Liquid Natural Gas) platform. With tanks equipped with FPSO, the existence of inner free surface in tanks would alter the motion of FPSO, and the external excitation such as wave force would change the inner free surface in return. Therefore, the coupling effect between ship motion and tank sloshing become an attractive research.

In recent years, many studies have been done on the sloshing tanks coupled with ship motion. Rognebakke and Faltinsen [1] conducted two-dimensional experiments about hull section with two tanks in waves. Later on, they adopted nonlinear theory to simulate the sloshing tanks and the external fluid separately. Newman [2] analyzed ship motion and sloshing tank in a unified approach – he extended panel code WAMIT and defined inner surface as an extension of external fluid surface. Seo et al [3] applied Rankine panel method to solve both ship motion and inner tanks. Since the viscous flow has advantages over linear potential flow in solving violent sloshing, Lee et al [3] developed a potential-viscous hybrid method to solve ship motion with sloshing tanks. Since then, the violent sloshing tanks were solved in CFD methods, such as fluent or OpenFOAM, and the external flow fields were solved in potential theory[4]-[8]. Shen and Wan [9] applied CFD methods both in external fluid and inner fluid to achieve fully coupled analysis. Zhuang and Wan [10] used fully coupled CFD method to simulate FPSO with two LNG tanks in low-filling ratios, and consider the wave breaking and

overturning in tanks. Its results showed well agreement with experiments and the ability to simulate the violent flow in tanks.

In this paper, the model scale of FPSO [11] was chosen. The main focus is on the different filling conditions in fore tank and aft tank. The results were compared to the existing experimental results. In order to discuss the function of sloshing fluid in fore and aft tank, the same filling conditions is also considered. The same filling conditions in fore and aft tank is discussed before [10], yet the sloshing forces and moments were not included. This time we added new function into our CFD solver to calculate the sloshing forces and moments, to discover the coupling effects between ship motion and two different filling ratio tanks.

Numerical Methods

Governing Equations

The incompressible Reynolds-Averaged Navier-Stocks equations are adopted in this paper to investigate the viscous flow. Using dynamic deformation mesh, the governing equations are:

$$\nabla\cdot\mathbf{U} = 0 \tag{1}$$

$$\frac{\partial\rho\mathbf{U}}{\partial t} + \nabla\cdot(\rho(\mathbf{U}-\mathbf{U}_g)\mathbf{U}) = -\nabla p_d - \mathbf{g}\cdot\mathbf{x}\nabla\rho + \nabla\cdot(\mu_{eff}\nabla\mathbf{U}) + (\nabla\mathbf{U})\cdot\nabla\mu_{eff} + f_\sigma + f_s \tag{2}$$

Where $\mathbf{U}$ is velocity field, $\mathbf{U}_g$ is velocity of grid nodes; $p_d = p - \rho\mathbf{g}\cdot\mathbf{x}$ is dynamic pressure; $\mu_{eff} = \rho(\nu + \nu_t)$ is effective dynamic viscosity, in which ν and ν_t are kinematic viscosity and eddy viscosity respectively. ν_t is obtained by k-ω SST turbulence model[12]. f_σ is the surface tension term in two phases model.

The solution of momentum and continuity equations is implemented by using the pressure-implicit spit operator (PISO) algorithm [13]. PISO algorithm applies mass conservation into pressure equation, thus when pressure equation converges, continuity error decreases. This method uses a predictor-corrector on solving pressure-velocity coupling, and utilizes a collocated grid method [14].

VOF Method

The Volume of fluid (VOF) method with bounded compression techniques is applied to control numerical diffusion and capture the two-phase interface efficiently. The VOF transport equation is described below:

$$\frac{\partial\alpha}{\partial t} + \nabla\cdot[(\mathbf{U}-\mathbf{U}_g)\alpha] = 0 \tag{3}$$

Where α is volume of fraction, indicating the relative proportion of fluid in each cell and its value is always between zero and one:

$$\begin{cases} \alpha = 0 & \text{air} \\ \alpha = 1 & \text{water} \\ 0 < \alpha < 1 & \text{interface} \end{cases} \tag{4}$$

6DOF Motions and Patch forces

A fully 6DOF module with bodies is implemented. Two coordinate systems are used to solve 6DOF equation. We describe $(x_1, x_2) = (x, y, z, \varphi, \theta, \psi)$ as the translation and rotation angles of

the ship, representing motions of surge, sway, heave, roll, pitch and yaw, respectively. $(v_1, v_2) = (u, v, w, p, q, r)$ are the velocities in the earth-fixed coordinate system, which can be transformed to the body-fixed coordinate system by equations given below:

$$v_1 = J_1^{-1} \cdot x_1 \qquad v_2 = J_2^{-1} \cdot x_2 \tag{5}$$

where J_1, J_2 are transformation matrices based on Euler angle. The forces and moments are projected into the earth-fixed system in following way:

$$F = (X, Y, Z) = J_1^{-1} \cdot F_e \qquad M = (K, M, N) = J_1^{-1} \cdot M_e \tag{6}$$

In order to compute the sloshing forces and moments, the single body is divided into several patches. In present study, the single body is divided into three patches: the hull of FLNG which can reveal external forces and moments, the left side of inner tank and the right side of inner tank. The forces are integrated on each patch separately, and accumulated as a whole to compute for next motion. The process is shown in Fig.1.

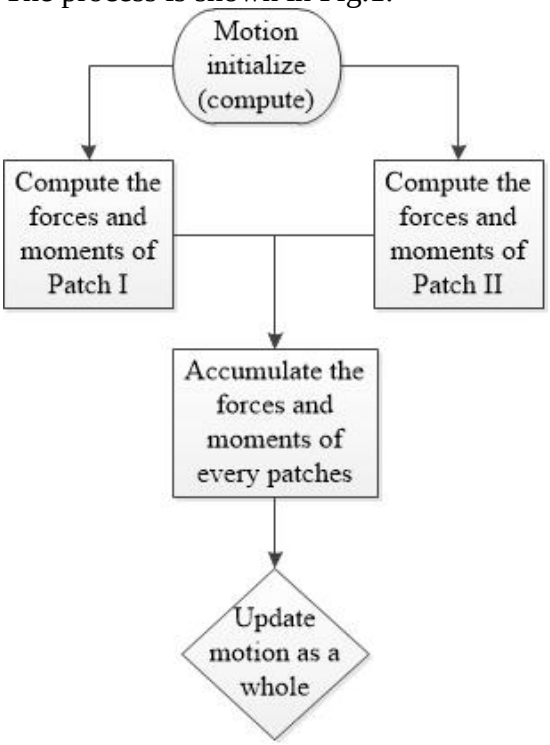


Figure 1. The algorithm for calculating the forces and moments of each patch

Wave Generation and Damping

The incoming regular wave is generated by imposing the boundary conditions of α and **U** at the inlet. The linear Stokes wave in deep water is applied for the wave generation.

$$\xi(x,t) = a\cos(kx - \omega_e t) \tag{7}$$

$$u(x,y,z,t) = a\omega e^{kz}\cos(kx - \omega_e t) \tag{8}$$

$$w(x,y,z,t) = a\omega e^{kz}\sin(kx - \omega_e t) \tag{9}$$

where ξ is the wave elevation; a is the wave amplitude; k is the wave number; U_0 is the ship velocity; ω is the natural frequency of wave; ω_e is the encounter frequency, in this condition, $\omega_e = \omega$.

Numerical Results

Geometry and mesh generation

The chosen model is simplified LNG FPSO [11]. The main particulars of ship-shaped FPSO vessel in both full scale and model scale are illustrated in Table 1.

Table 1. Main particulars of LNG FPSO.

Main particulars		Full Scale	Model
Scale factor	—	1	1/100
Length between perpendiculars	L_{PP} (m)	285	2.85
Maximum beam of waterline	B_{WL} (m)	63	0.63
Draft	T (m)	13	0.13
Displacement	Δ (m^3)	220017.6	220.0176
Natural period of roll	T_{ϕ} (s)	13	1.3
Vertical Center of Gravity (from keel)	KG (m)	16.5	0.165
Radius of gyration	K_{xx}	19.45	0.1945
	K_{yy}	71.25	0.7125

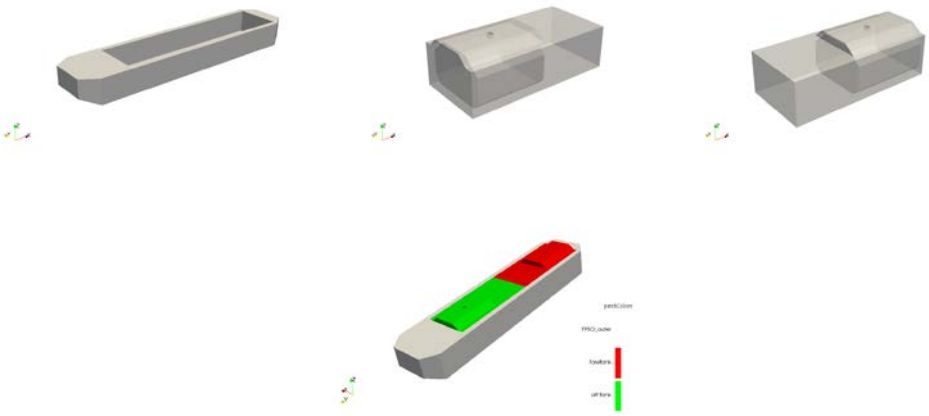


Figure 2. The geometry of each part from FPSO and inner tanks

The model scaled LNG FPSO and computational domain setup is shown in Fig.3. The filling condition is also included, shows that 82.6% filling ratio in fore tank, and 23.5% filling ratio in aft tank. The computational domain was set as $-1.0L_{pp}<x<2.0L_{pp}$, $-1.5L_{pp}<y<1.5L_{pp}$, $-1.0L_{pp}<z<1.0L_{pp}$. The sponge layer was included to avoid the wave reflection and keep mass conservation in computational domain. The FPSO was restrained to heave, roll and pitch motion.

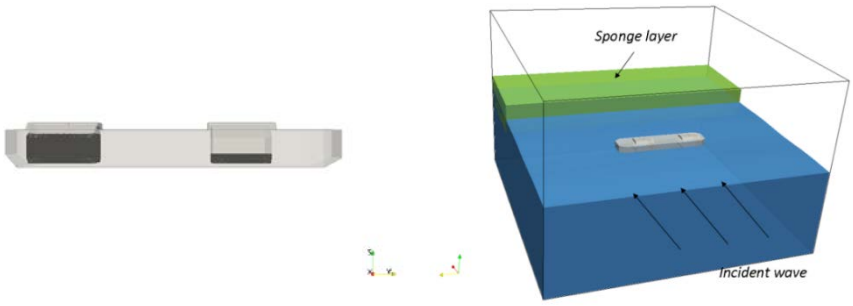


Figure 3. The geometry and computational domain setup

The meshes are generated by snappyHexMesh, an auto mesh generation utility provided by OpenFOAM. The total cell numbers are around 2.1M, and the LNG tanks require 0.5M cells. The mesh details are shown in Fig. 4. The right side of picture in Fig.4 shows two small tunnels connected the LNG tanks to the external region, which can keep the pressure inside the tanks same to the external region and simplify the computations. Meanwhile, the connection tunnel makes the computation fully coupled.

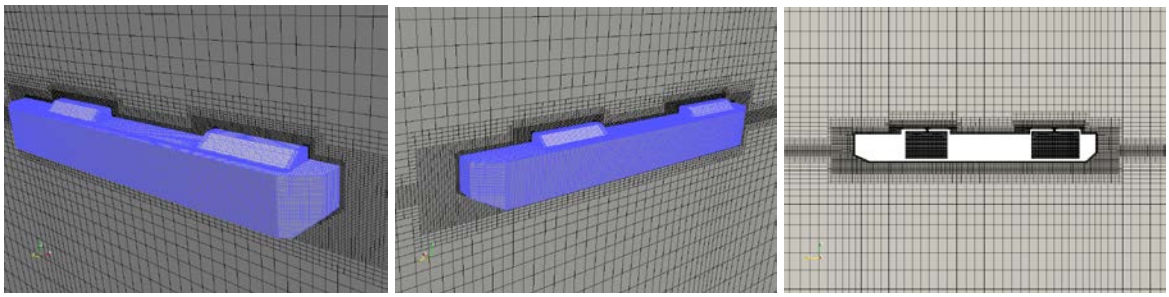


Figure 4. Demonstrations of meshes

Numerical verification

The roll motion response in frequency domain is shown in Fig.5. The normalized motion amplitude and natural frequency are considered to compare with experimental data. The normalized roll motion is given as: $\theta B/2A$, in which θ is maximum degree of roll motion, B is beam of ship and A is wave amplitude. The normalized natural frequency is given as: $\omega(L/g)^{1/2}$, in which ω is natural frequency of water, L represents length of ship.

In order to compare with the existing experimental data, the simulations are all under the same incident wave height, which is 0.025m. Two different incident wave frequencies are considered, and the normalized wave frequencies were chosen to be 2.5 and 3.5 separately. Fig.5 indicates the comparison of roll motion between present work and experiment. Besides, the results of other numerical results were also included. The potential theory [15] adopted in 20% filling conditions cannot simulate the 82.6% filling ratio, for the high filling condition reach the top corner of the LNG tank and the linear potential theory did not have the ability to simulate it. The results of circumstances were better than other simulation results when compared to the experimental data, showing an advantage in simulating the hybrid of high-filling condition and low-filling condition. Some of the results about 20% filling condition were from the work done before [10], and the case of normalized wave frequency equals to 3.5 was new added in this paper. The results of current study show well agreement with experimental results, which proves that our methods are reliable.

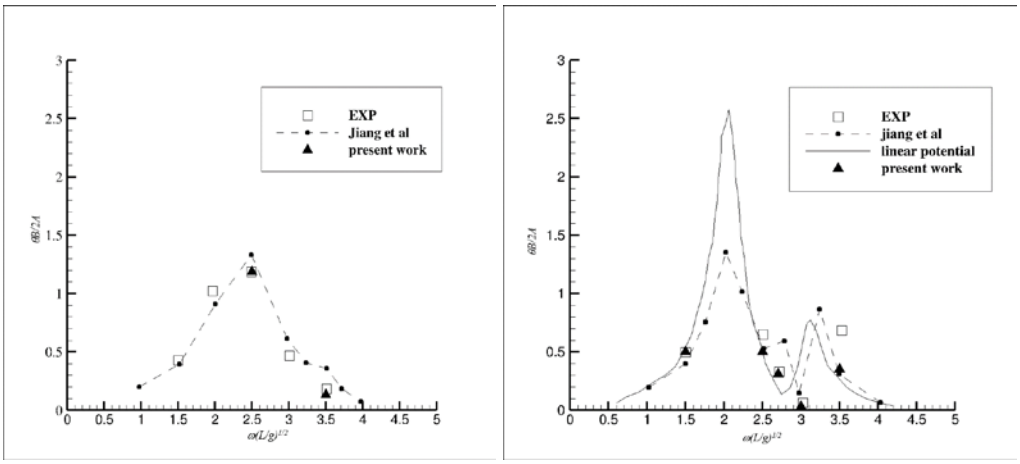


Figure 5. Comparison among current studies and experiments and other numerical methods.(left 82.6%-23.5%, right 20%-20%)

It can be observed that the 20% filling conditions had two peaks of roll motion response, while the 82.6%-23.5%filling condition only had one peak. To better understand the coupling effects between ship motion and sloshing tanks, and the coupling effects of two sloshing tanks as well, the inner sloshing moments and external exciting moments were discussed.

Sloshing moment and external exciting moment

The sloshing moments of 82.6%-23.5% filling condition were shown in Fig.6. The sloshing moments of fore tank and aft tank were compared in normalized moment. It can be seen that under these two wave frequencies, the external wave force produce larger moment on the low filling ratio tank. Under different wave frequencies, the time history of moment from each tank shows different phenomenon. In the normalized wave frequency of 2.5, the moment curve of fore tank (82.6% filling ratio) and the aft tank (23.5%) did not show phase difference. However, in the normalized wave frequency equals to 3.5, the time history of moment in each tank shows the phase difference, almost near 180 degree. The phase difference means under this wave frequency, the sloshing moments in aft tank and fore tank influence each other, and reduce the total sloshing moment which was acted on FPSO.

Considering the external exciting moment, shown in Fig.7, it can be observed that the time history of fore tank moment have phase difference with external exciting moments under both wave conditions. That's the reason why the sloshing fluid can reduce the ship motion. Besides, the number of sloshing moment and exciting moment under the case of normalized wave frequency equals to 3.5 is much smaller than that in the case of 2.5, therefore the roll motion of 3.5 case was smaller than that of 2.5 case.

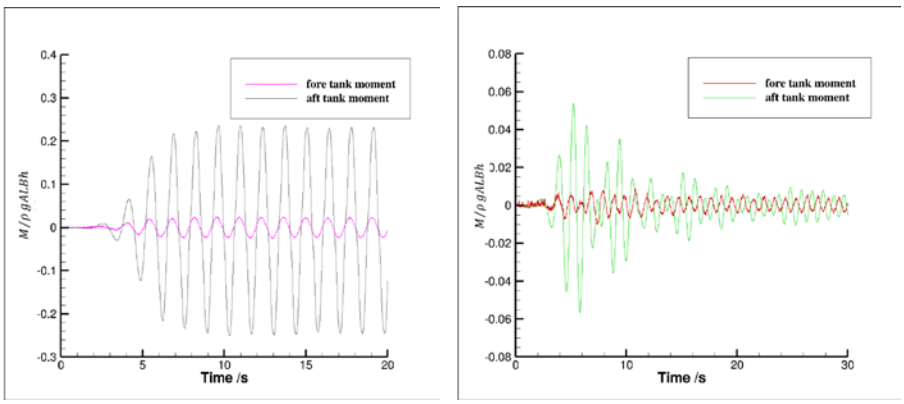


Figure 6. Time history of inner sloshing moment in aft and fore tank in different wave frequencies. (filling ratio 82.6%-23.5%, from left to right: normalized wave frequency 2.5, 3.5.)

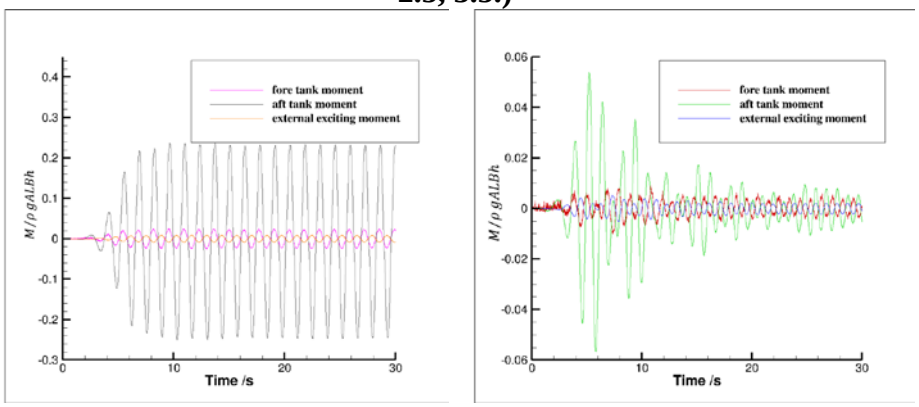


Figure 7. Time history of inner sloshing moment in aft and fore tank and external exciting moment in different wave frequencies. (filling ratio 82.6%-23.5%, from left to right: normalized wave frequency 2.5, 3.5.)

Fig.8 illustrated the sloshing moment of 20%-20% filling condition. For the filling ratios were the same in fore and aft tanks, the sloshing moments were also the same. Fig.9.included the external exciting moments. It can be observed that in the case of 2.5, the external exciting moments showed phase difference with the sloshing moment. The degree of phase difference were almost 180 degree. The external exciting moment excited the sloshing moments, while the sloshing moment reduced the external exciting moment in return. However, the case that normalized wave frequency equals to 3.5 were totally different. The phase of sloshing moment and exiting moment were the same. Although the external exiting moment in case of 3.5 was smaller than that in case of 2.5, the induced sloshing moment increase the total moment acted on ship. Therefore, the value of roll ship motion didn't show difference between these two cases.

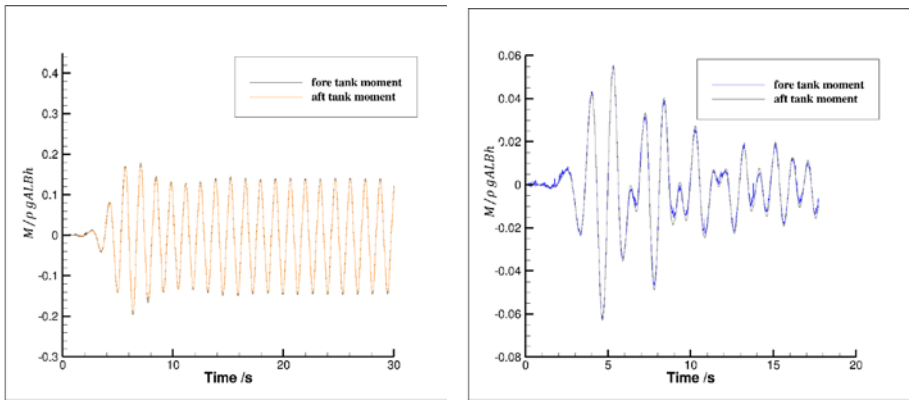


Figure 8. Time history of inner sloshing moment in aft and fore tank and external exciting moment in different wave frequencies. (filling ratio 20%-20%, from left to right: normalized wave frequency 2.5, 3.5.)

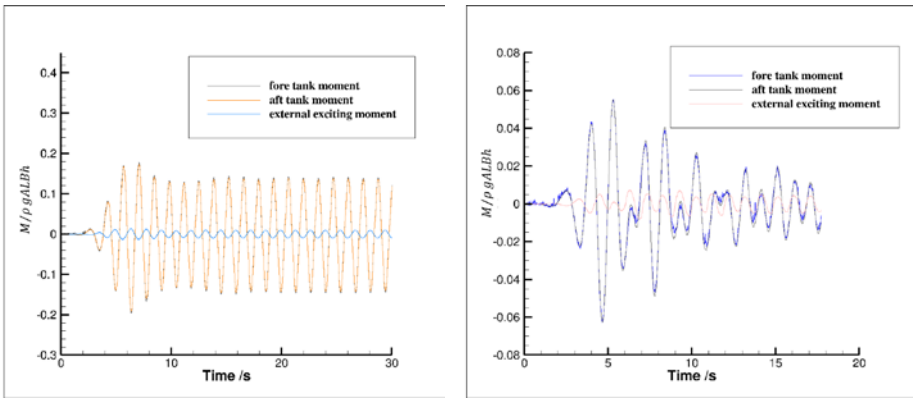


Figure 9. Time history of inner sloshing moment in aft and fore tank and external exciting moment in different wave frequencies. (filling ratio 20%-20%, from left to right: normalized wave frequency 2.5, 3.5.)

Sloshing flow

To observe the coupling effects, the internal tank flow were included. Fig.10 shows three snapshots of sloshing flow in fore tank and aft tank. The images of fore tank and aft tank were all in the same time. To better observe the detail of sloshing flow, the snapshots of fore tank had mirror reversal. Thus the phase of sloshing flow in fore tank and aft tank were the same, just accord with the sloshing moment shown before.

The sloshing flow in fore tank is almost linear, for the filling ratio is high. While in aft tank appeared a wave crest among the sloshing flow. This peak moved along with ship, probably induced by ship motion. To better observe the sloshing flow, we simulated a large wave height case, shown in Fig.11, where the wave height equals to 0.1m. It can be seen that the sloshing flow in high filling condition tank was still linear, and it had a phase difference with external flow. Due to the low-filling condition, the sloshing flow in aft tank showed a great nonlinearity. The inner fluid broke on the bulkhead, and formed a propulsion wave that moved with ship motion. And then the propulsion wave reached the other bulkhead, and broke on the bulkhead. In order to observe the sloshing flow in aft tank, the velocity vector of the wave crest was included and shown in Fig.12. The velocity distribution along the tank length was not the same because of the pitch motion. The direction of velocity in wave crest changed when the propulsion wave on its way to the bulkhead, thus wave rolled over during the propulsion. This kind of sloshing flow showed a phenomenon like that of shallow water wave.

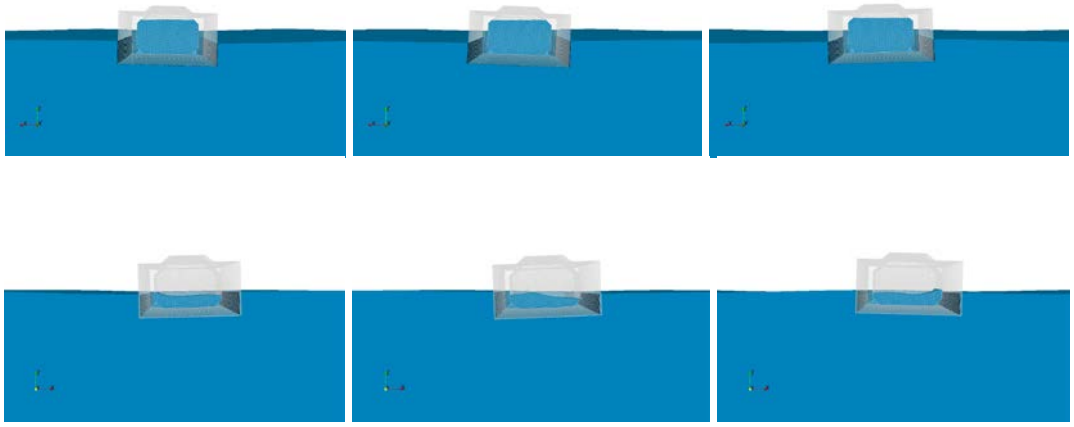


Figure 10. Snapshots of inner tank flow and outer flow (above were fore tank filling 82.6%, bottom were aft tank filling 23.5%; normalized wave frequency is 2.5)

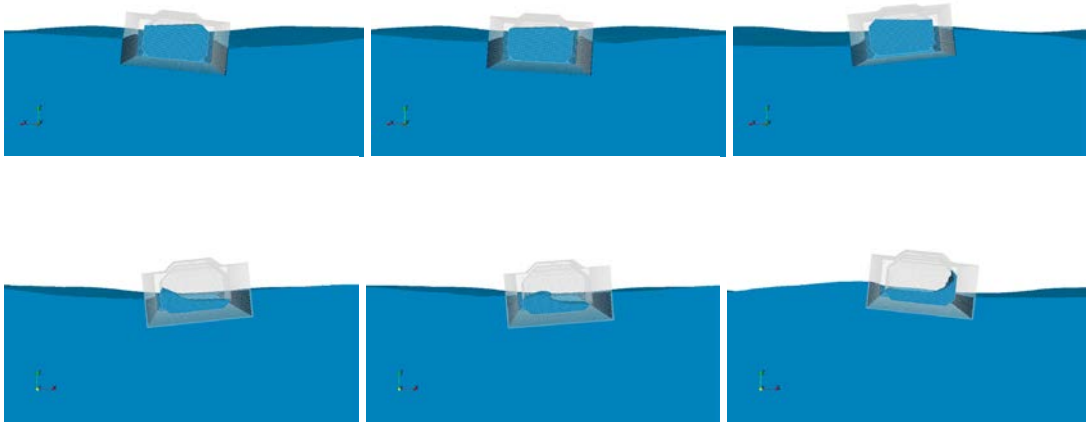


Figure 11. Snapshots of inner tank flow and outer flow in wave height equals 0.1m. (above were fore tank filling 82.6%, bottom were aft tank filling 23.5%; normalized wave frequency is 2.5)

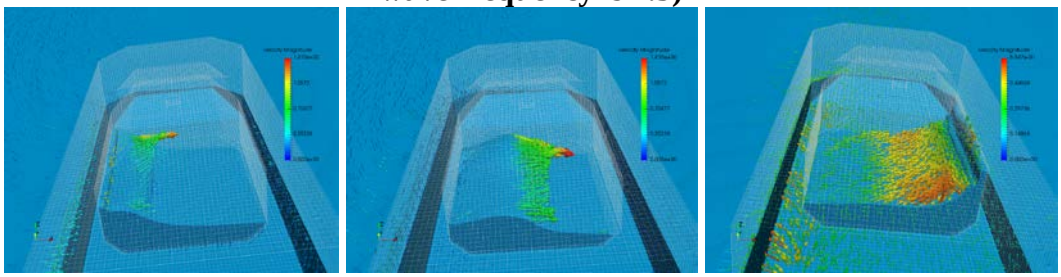


Figure 12. Velocity vector of inner tank flow and outer flow in wave height equals 0.1m. (above were fore tank filling 82.6%, bottom were aft tank filling 23.5%; normalized wave frequency is 2.5)

Conclusions

In this paper, the coupling effects were discussed using our in house solver naoe-FOAM-SJTU. The LNG FPSO was equipped with two different filling ratio tanks, and both the ship motion and sloshing moments were considered. The roll motion in different wave frequencies of FPSO were compared with experimental results, and showed well agreement, which revealed our solver had the advantages to solve both the high filling condition and low filling condition. Through observing the results, we found that the roll motion of 82.6%-23.5% filling condition reduced at the case that normalized wave frequency equals to 3.5, while that of 20%-20% was almost the same. In order to figure it out, the study of sloshing moments was carried out.

Our in house CFD solver naoe-FOAM-SJTU has added new functions to solve the problem. The model was departed into inner and outer parts to obtain the sloshing moments and external exciting moments. The results were analyzed to figure out how the sloshing moments influence the ship motion. The sloshing moments in tanks have a phase difference with external exciting moments when the normalized wave frequency equals to 2.5, thus the existence of sloshing tank would reduce the ship motion. When the normalized wave frequency equals to 3.5, In the case of 82.6%-23.5%, not only the sloshing moment in fore tank and aft tank showed phase difference, but external exciting moments showed the phase

difference as well; whereas in the case of 20%-20%, both the sloshing moments in fore tank and aft tank and external exciting moments were in the same phase, therefore the ship roll motion didn't reduce as that in case of 82.6%-23.5%.

At last, the sloshing flow and velocity were included to observe the details of sloshing flow. For the high filling condition, the sloshing flow is almost linear even under large wave height. However, the low filling condition shows the nonlinearity both under small wave height and large wave height. The break and roll over were observed under large wave height condition.

In the future, the wave height in the tank would be observed to deeper understand the relationship between external wave excitation and internal flow. Different wave direction would be considered.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (51379125, 51490675, 11432009, 51579145), Chang Jiang Scholars Program (T2014099), Shanghai Excellent Academic Leaders Program (17XD1402300), Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (2013022), Innovative Special Project of Numerical Tank of Ministry of Industry and Information Technology of China (2016-23/09) and Lloyd's Register Foundation for doctoral student, to which the authors are most grateful.

References

[1] Rognebakke, O. F., Faltinsen, O. M. (2003) Coupling of sloshing and ship motions, *Journal of Ship Research* **47**, 208-221.

[2] Newman, J. N. (2005) Wave effects on vessels with internal tanks. *In: Proceedings of the 20th Workshop on Water and Floating Bodies*. Sptisbergen, Massachusetts.

[3] Seo, M. G., Kim, Y., and Park, D. M. (2017). Effect of internal sloshing on added resistance of ship. *Journal of Hydrodynamics*, Ser. B, **29**, 13-26.

[4] Lee, S. J., Kim, M. H., Lee, D. H. and Shin, Y. S. (2007) The effects of tank sloshing on LNG vessel response. *In: Proceedings. of 26th International Conference on Offshore Mechanics and Arctic Engineering*, San Diego, California, OMAE 2007-29665.

[5] Li, Y. L., Zhu, R. C., Miao, G. P. and Fan, J. (2012) Simulation of tank sloshing based on OpenFOAM and coupling with ship motions in time domain. *Journal of Hydrodynamics*, Ser. B, **24**, 450-457.

[6] Kim, J. W., Kim, K. (2007) Response-based evaluation of design Sloshing loads for membrane-type LNG carriers. *In: Proceedings of 26th International Conference on Offshore Mechanics and Arctic Engineering*. San Diego, USA, OMAE2007-29746.

[7] Jiang, S. C., Teng, B., Bai W., Guo, Y. (2015) Numerical simulation of coupling effect between ship motion and liquid sloshing under wave action. *Ocean Engineering*, **108**, 140-154

[8] Li, X., Zhang, T., Zhang, Y. O., and Wang, Y. X. (2014, April). Numerical analysis of ship motion coupled with tank sloshing. *In OCEANS 2014-TAIPEI* IEEE, 1-10.

[9] Shen, Z. R. and Wan, D. C. (2012) Numerical Simulations of Large-Amplitude Motions of KVLCC2 with Tank Liquid Sloshing in Waves. *In Proceedings 2nd Int Conf Violent Flows*. Ecole Centrale Nantes, Nantes, France.

[10] Zhuang, Y. and Wan, D. C. (2016) Numerical Study on Coupling Effects of FPSO Ship Motion and LNG Tank Sloshing in Low-Filling Conditions. *Applied Mathematics & Mechanics*, **37**, 1378-1393.

[11] Nam, B. W, Kim, Y., Kim, D. W., Kim, Y. S. Experimental and Numerical Studies on Ship Motion Responses Coupled with Sloshing in Waves. *Journal of Ship Research*, 2009, **53**, 68-82.

[12] Issa, R. I. (1986). Solution of the implicitly discretized fluid flow equations by operator-splitting. *Journal of computational physics*, **62**, 40-65.

[13] Rhie, C. M. and Chow, W. L. (1983). Numerical study of the turbulent flow past an airfoil with trailing edge separation. *AIAA journal*, **21**, 1525-1532.

[14] Dhakal, T. P. and Walters, D. K. (2009, January). Curvature and rotation sensitive variants of the K-Omega SST turbulence model, *In ASME 2009 Fluids Engineering Division Summer Meeting*. American Society of Mechanical Engineers. 2221-2229.

[15] Guo, Y., Kim, Y., Kim, T. Y. (2011) A numerical study on coupling between ship motions and sloshing in frequency and time domain. *Proceedings of the 21th International Offshore and Polar Engineering Conference*, Maui, Hawaii, USA, 158-164.

Computational Hemodynamics for a Pair of Intracranial Aneurysms Treated with Flow Diverting Stents

Xudong Liu[1], Zhuangyuan Meng[1], *Shengzhang Wang[1], and Xiaolong Zhang[2]

[1]Institute of Biomechanics, Department of Aeronautics and Astronautics, Fudan University, 220 Handan Road, Shanghai 200433, China

[2]Department of Radiology, Huashan Hospital affiliated to Fudan University, 12 Middle Urumuqi Road, Shanghai 200040, China

*Presenting and corresponding author: szwang@fudan.edu.cn

Abstract

Flow diverting stent is a stent-like device with higher metal coverage rate to occlude the intracranial aneurysms. Comparing to coils, flow diverting stent is considered more effective and lower recanalization rate. In this article, a pair of intracranial aneurysms, both located at the two sides of the parent artery and one aneurysm sac involving a branch artery, were treated virtually with two kinds of flow diverting stents: LVIS and Pipeline by rapid deployment method. Blood flows in the pair of aneurysms before and after treatment were simulated by computational fluid dynamics and the hemodynamic parameters such as velocity, wall shear stress were calculated. Analyzing the numerical results, it was observed that velocity reduction of the treated aneurysms with Pipeline was much larger than with LVIS, wall shear stress on the treated aneurysm with Pipeline is much smaller than with LVIS, and the increment of flow rate at the branch artery when the aneurysm treated with Pipeline was larger than that with LVIS. The following conclusions can be summarized that Pipeline is more effective to occlude the aneurysm than LVIS, but the aneurysm treated with Pipeline has more rupture risk than with LVIS. The branch artery located at the aneurysm sac has suction effect when the aneurysm is treated with flow diverting stent, which reduces velocity further inside the aneurysm sac and benefits embolization formation.

Keywords: Hemodynamics, Intracranial aneurysm, Flow diverting stent, Computational fluid dynamics

Introduction

Intracranial aneurysms are pathologic dilations of the intracranial arteries, generally located at the circle of Willis. Rupture of intracranial aneurysms causes subarachnoid hemorrhage with an associated high mortality and morbidity rate [1]. Endovascular treatment is a minimally invasive surgery which releases coils or stent-assisted coils to occlude aneurysms in order to prevent blood flowing into the aneurysmal sacs and promote thrombus formation. However the aneurysms occluded with coils have high recanalization rate comparing to surgery clipping [2]. Flow diverting stent is a new endovascular device to occlude the aneurysm and it is a braided stent with higher metal coverage rate. An intracranial aneurysm occluded by a flow diverting stent is considered much lower recanalization risk than coils embolization [2]. Many investigators have studies the relationship between hemodynamics and intracranial aneurysms and it is widely agreed that hemodynamics plays a very important role in the initiation, growth and rupture of intracranial aneurysms. The change of hemodynamics when flow diverting stents deployed to treat intracranial aneurysms has been studies by many research groups [3-7]. However, when intracranial aneurysms involving branch arteries, the aneurysms treated with flow diverting stents may occlude branch arteries to induce cerebral ischemia. Hence change of hemodynamics when the aneurysm involved branch vessel treated with flow diverting stent is very complicated but has rarely been discussed. In this study, a pair of intracranial aneurysms involving a branch artery at one aneurysm sac will be treated virtually with two kinds of flow diverting stents by Rapid deployment technique. Blood flows in the aneurysms before and after treatment will be simulated by computational fluid dynamics, and the hemodynamic parameters such as velocity, flow rate, and wall shear stress will be calculated. By analyzing the changes of these parameters after the aneurysms treated

with flow diverting stents, the effects of the different stents will be discussed and compared from the hemodynamic viewpoint.

Methodology

A 53 year-old female patient was suffering from a pair of aneurysm sacs located two sides of the left internal carotid artery. The patient-specific geometrical model was constructed from the medical image data. The two kinds of flow diverting stents were deployed into the parent artery across the orifices of both aneurysms by rapid virtual deployment technique. Blood flows in the aneurysms before and after treatment were simulated by computational fluid dynamics. This section introduces the details of these methods.

Vascular Models

A pair of intracranial aneurysm sacs located at the left internal carotid artery were selected to study. The two aneurysm sacs located at the two sides of the parent artery and a branch artery locates at the upper sac of the aneurysm model, see the left figure of Figure 1. The 3D geometry of the artery was constructed from the patient undergoing clinically indicated conventional angiography with rotational data acquisition. The reconstructed 3D geometry of the artery was exported into STL (Stereo Lithography format) format file. The geometry was imported into the reverse engineering software Geomagic studio 11.0 to segment, repair and smooth. After these stages, the geometrical model of the aneurysm was constructed. The diameters of the inlet and the outlet at the artery are 4.8 mm and 4.4 mm, respectively, and the diameter of branch artery is 1.8 mm.

Virtual Deployment

To treat a pair of aneurysm sacs located at two sides of the same position of the parent artery, comparing to coils embolization, flow diverting stent has lower money cost and simpler procedure. Assuming that the pair of aneurysm sacs are treated with two kinds of flow diverting stents: LVIS and Pipeline, both are widely employed in clinic. The difficulty of the virtual treatments is to deploy flow diverting stent into the aneurysm model reasonably. Finite element method is capable to deploy flow diverting stents virtually but it is very complicated for modeling and time consuming for simulation [8-9]. Rapid virtual deployment technique based on simplex mesh method or spring theory has been applied to deploy flow diverting stents virtually by many groups and it has been proved acceptable from the geometry and the artery wall attaching [10-12]. Both flow diverting stents are deployed virtually into the parent artery across the aneurysms orbits by rapid virtual deployment technique. Actually, both flow diverting stents have the different geometric parameters such as the number of struts, the diameter of the strut and the metal coverage rate, presented in Table 1.

Table 1. Geometric parameters of LVIS and Pipeline

	LVIS	Pipeline
Number of the struts	16	48
Diameter of the strut (μm)	61	48
Metal coverage rate	20%	29%

Figure 1: The models of a pair of aneurysm sacs before (Left) and after treated with LVIS (Middle) and Pipeline (Right)

Numerical Simulations

The computational meshes of the aneurysm models were generated by Ansys ICEM-CFD 15.0, and the unstructured meshes composed of tetra elements were specified as follows: the max element size in the artery was 0.3 mm, the max element sizes near the metal struts of LVIS and Pipeline were 0.03 mm and 0.02 mm, respectively, in order to capture the strut wires. Blood was assumed as Newtonian flow with density 1050 kg/m^3 and dynamic viscosity 0.0035 Pa•s. The steady laminar flows in the models were simulated by the commercial CFD package ANSYS CFX 15.0. A parabolic velocity profile was imposed at the inlet and the average velocity at the inlet was calculated by Poiseuille law to guarantee the average wall shear stress at the inlet artery to be 1.5 Pa [4-7]. The boundary condition was traction-free with 10000 Pa reference pressure at the outlets. The walls of the artery and the struts of the flow diverting stents had no-slip condition.

Results and Discussion

Figure 2 demonstrates velocity contours on the center plane cross the pair of aneurysm sacs before and after treated with LVIS and Pipeline. It is observed that the velocity in the pair of aneurysm sacs, especially in the upper sac, is reduced remarkably after treated with Pipeline, but LVIS is not so effective to reduce the velocity inside the aneurysm sacs. Figure 3 demonstrates wall shear stress on the pair of aneurysms before and after treated with LVIS and Pipeline. It is observed that wall shear stress on the aneurysm sacs is reduced remarkably after treated with Pipeline, but the reduction on the aneurysm sacs after treated with LVIS is small. Moreover, no obvious change can be observed on the parent artery after treated with flow diverting stents. In order to occlude the aneurysm by flow diverting stents, the more velocity reduction is expected and therefore Pipeline is more effective to occlude aneurysms than LVIS. On the other hand, lower wall shear stress is considered to be a high risk factor related to the aneurysm rupture [3-4], therefore the aneurysm treated with Pipeline has higher risk to rupture than treated with LVIS. Table 2 demonstrates the variation of flow rat at the inlet, outlet and small branch artery before and after treatment. It is observed that the flow rate at the branch artery increases when Pipeline or LVIS is deployed to occlude the aneurysm sacs and the increment with Pipeline is larger than that with LVIS. Actually, Pipeline has higher metal coverage rate than LVIS, therefore it is understood easily that Pipeline is more effective to reduce the velocity inside the aneurysm sacs than LVIS. However, the branch artery located at the upper aneurysm sac has the suction effect when a flow diverting stent is deployed to occlude the aneurysms, that is why the flow rate at the branch artery is increased

after treatment, and the suction effect of branch artery can reduce the velocity inside the aneurysm sac further.

Table 2: Flow rate at the branch vessel before and after the aneurysms treated with LVIS and Pipeline

Flow rate (ml/s)	No treatment	LVIS	Pipeline
inlet	4.88	4.88	4.88
outlet	4.65	4.54	4.52
branch	0.23	0.34	0.36

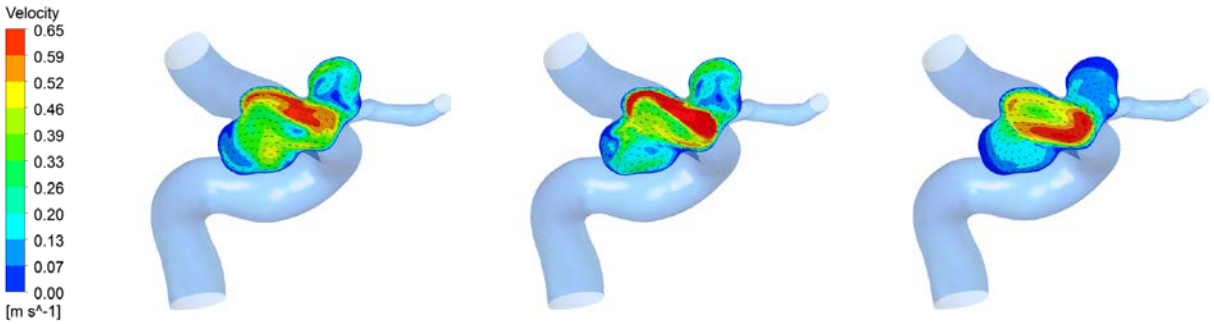


Figure 2: Velocity contours on the cross section of the pair of aneurysms before (Left) and after treated with LVIS (Middle) and Pipeline (Right)

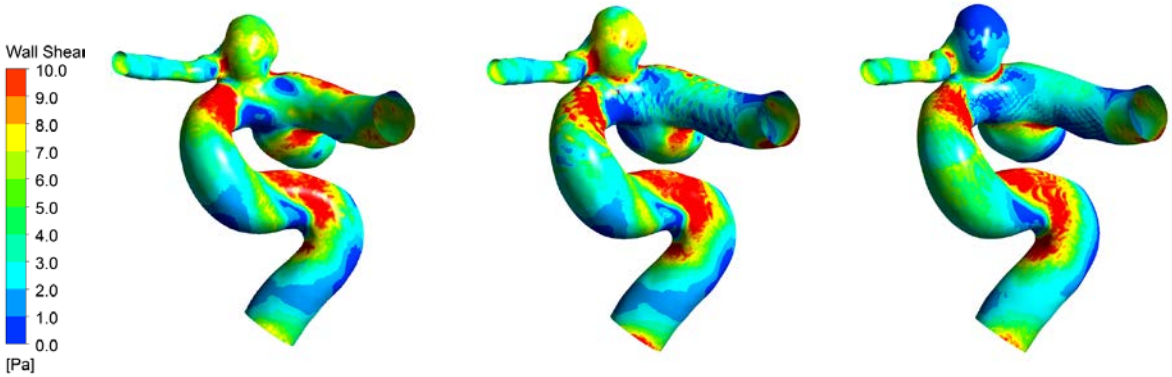


Figure 3: Wall shear stress contours on the aneurysm before (Left) and after treated with LVIS (Middle) and Pipeline (Right)

Conclusions

The following conclusions can be drawn from the numerical results: Pipeline is more effective to occlude the aneurysm than LVIS, but the aneurysm treated with Pipeline has more rupture risk than with LVIS. The branch artery located at the aneurysm sac has the suction effect when the aneurysm sacs are occluded with flow diverting stents, which reduces the velocity inside the aneurysm sac further and benefits embolization formation.

Acknowledgements

This study was supported by National Natural Science Foundation of China (Nos.81571128, 81371308, 81371315).

References

[1] Schievink W I. Intracranial aneurysms. New England J Medicine, 336: 28–40, 1997.
[2] Seibert B, Tummala RP, Chow R, Faridar A, Mousavi SA and Divani AA. Intracranial Aneurysms: Review of Current Treatment Options and Outcomes. Frontiers in Neurology, 2:1-10, 2011.
[3] Sforza M, Putman CM and Cebral JR. Hemodynamics of cerebral aneurysms, Annual Review of Fluid Mechanics, 41:91-107, 2009.
[4] Chung B, Cebral JR. CFD for evaluation and treatment planning of aneurysms: review of proposed clinical uses and their challenges. Annals of biomedical engineering, 43:122-138, 2015.
[5] Cebral JR, Mut F, Raschi M, Scrivano E, Ceratto R, Lylyk P and Putman CM. Aneurysm Rupture Following Treatment with Flow-Diverting Stents: Computational Hemodynamics Analysis of Treatment. American Journal of Neuroradiology, 32:27-33, 2011.
[6] Hassan T, Ahmed YM and Hassan AA. The adverse effects of flow-diverter stent-like devices on the flow pattern of saccular intracranial aneurysm models: computational fluid dynamics study. Acta Neurochirurgica, 153:1633-40, 2011.
[7] Mut F, Scrivano E, Bleise C, Lylyk P and Cebral JR. Hemodynamics in two tandem aneurysms with flow diverters. Int J Numer Meth Biomed Engng, 30:517-524, 2014.
[8] Damiano RJ, Ma D, Xiang J, Siddiqui AH, Snyder KV, Meng H. Finite element modeling of endovascular coiling and flow diversion enables hemodynamic prediction of complex treatment strategies for intracranial aneurysm. Journal of biomechanics, 48:3332-3340, 2015.
[9] Ma D, Xiang J, Choi H, et al. Enhanced aneurysmal flow diversion using a dynamic push-pull technique: an experimental and modeling study. AJNR American journal of neuroradiology, 35:1779-1785, 2014.
[10] Bernardini A, Larrabide I, Petrini L, Pennati G, Flore E, Kim M. Deployment of self-expandable stents in aneurysmatic cerebral vessels: comparison of different computational approaches for interventional planning. Computer Methods in Biomechanics and Biomedical Engineering, 15:303-11, 2012.
[11] Larrabide I, Kim M, Augsburger L, Villa-Uriol MC, Rufenacht D, Frangi AF. Fast virtual deployment of self-expandable stents: method and in vitro evaluation for intracranial aneurysmal stenting. Medical image analysis 2012;16:721-730.
[12] Peach TW, Ngoepe M, Spranger K, Zajarias-Fainsod D and Ventikos Y. Personalizing flow-diverter intervention for cerebral aneurysms: from computational hemodynamics to biochemical modeling. Int J Numer Meth Biomed Engng, 30:1387-1407, 2014.

Consistent high order meshfree Galerkin methods and applications

†Qinglin Duan*, Bingbing Wang, Xin Gao, Xikui Li and Hongwu Zhang

The State key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian, Liaoning 116024, P.R.China.

*Presenting author: qinglinduan@dlut.edu.cn
†Corresponding author: qinglinduan@dlut.edu.cn

Abstract

Meshfree methods such as the element-free Galerkin (EFG) method have been developed to be a formidable competitor and also a beneficial complement to the traditional finite element method (FEM) which dominates engineering analysis for decades. One attractive advantage of meshfree methods is that constructing high order approximation is much more convenient than that in the finite element method (FEM). However, high order meshfree methods are computationally inefficient since a large number of integration points are required. On the other hand, the stabilized conforming nodal integration method based on strain smoothing is very efficient for linear meshfree Galerkin methods, but it cannot exploit the high convergence and accuracy of meshfree methods with high order approximation. In this work, the number of quadrature points for high order meshfree methods is remarkably reduced by correcting the nodal derivatives. Such correction is rationally developed based on the Hu-Washizu three-field variational principle. The proposed method is able to exactly pass patch tests in a consistent manner and is therefore, named as consistent high order meshfree Galerkin methods. In contrast, the traditional meshfree methods cannot exactly pass patch tests. Numerical results of elastostatic problems show that the proposed technique remarkably improves the numerical performance of high order meshfree methods in terms of accuracy, convergence, efficiency and stability. Applications of the proposed methods to thin plates and shells as well as crack problems are also presented.

Keywords: Meshfree; EFG; Numerical integration; Plate; Shell; Crack

Introduction

Meshfree methods[1-6] such as the element-free Galerkin (EFG) method developed in recent twenty years have become a formidable competitor to the traditional finite element method (FEM) which dominates engineering analysis for decades. The common feature shared by the so called “meshfree” or “meshless” methods is that they are based on scattered data approximation which do not need explicit nodal connectivity(element). So far, EFG is one of the most popular and successful meshfree methods. Its fundamental advantage against traditional FEM is the smoothness of the approximation function. With careful choice of the weight function for the MLS process, the continuity of the approximation function can be C^{∞} which provides superiority for solving high order partial differential equation such as thin plates and shells whereas the continuity of the FEM approximation is only C^{0}. Furthermore, high order approximation can be conveniently obtained in EFG. The required change in the input data is minimum. In contrast, in FEM, high order elements such as 6-node or 10-node triangle elements have to be constructed and this changes the input data a lot. Finally, EFG is easy to achieve h-adaptive computation[7] since its approximation is only based on nodes (not elements).

However, a main issue of the EFG method (or more generally the meshfree methods) is the efficient numerical integration of the weak form. Background meshes with Gauss integration points are commonly used in EFG. Due to the non-polynomial character of the MLS approximants, high order Gauss integration has to be employed to result a stable method. Clearly, the large number of integration points consumes more CPU time and thus severely impairs the computational efficiency. What’s worse is that even the high order integration cannot integrate the weak form accurately enough to make the method exactly pass the patch test.

Many efforts have been devoted to develop stable and efficient integration methods with reduced number of sampling points such as the nodal integration [8-10], the stress-point

integration [11-12], the support domain integration [13], etc. Among these, the nodal integration initiated by Bessial and Belytschko [8] can dramatically improve the efficiency since it uses the minimum evaluating points (the nodes) as integration points. However, direct nodal integration is not stable and can't pass the patch tests, some works have been done to relieve this issue [10-15]. Among these, Chen *et al.* [14] developed a stabilized conforming nodal integration (SCNI) which is stable and provides even better accuracy than Gauss integration. They showed that SCNI can pass the linear patch test whereas Gauss integration fails. One outstanding merit of this method is that no additional term or stabilization parameter is involved. So far, SCNI has developed to be a major integration scheme in meshfree method and the strain smoothing technique in SCNI has been extend into FEM analysis [16].

However, Puso *et.al.* [17] reported SCNI may still cause oscillation near the boundary of the solution domain. What is more, Duan *et.al.* [18] reported that SCNI is only linear exactness and is not adequate for quadratic meshfree approximation. They further presented a consistency framework guiding the correction of the nodal derivatives based on the divergence theorem between a nodal shape function and its derivatives to remedy this issue. Particularly, a three-point integration scheme with second order accuracy named quadratically consistent three-point (QC3) integration method for second order meshfree method is developed in such framework. QC3 employs triangular background integration cells. In each integration cell, the nodal shape functions on six boundary sampling points and three domain sampling points are used to determinate the nodal corrected derivatives on the three domain sampling points. Later, by further reformulating the framework of nodal derivative correction based on the Hu-Washizu three-field variational principle, Duan *et.al.* [19] proposed the consistent element-free Galerkin (CEFG) method and showed its much better numerical performance in terms of accuracy, convergence, efficiency and stability than the standard EFG method. It should be stressed that the proposed EFG method is based on the Hu-Washizu three-field variational principle and can pass the patch tests in a consistent manner, i.e. EFG with linear, quadratic and cubic bases can, respectively, pass the linear, quadratic and cubic patch tests. Therefore, the method is named as consistent element-free Galerkin (CEFG) method. In contrast, the traditional EFG method is based on the classical one-field variational principle and cannot pass the patch tests.

The paper is structured as follows. The standard EFG method is first reviewed in section 2. The proposed consistent EFG method is then described in section 3. Applications of the proposed method to thin-plates and shells as well as crack problems are presented in section 4 followed by the conclusions in section 5.

Element-free Galerkin (EFG) method: approximation and discretization

EFG was invented by Belytschko *et al.* [2] about twenty years ago and so far it has already developed into one of the most popular and successful meshfree Galerkin methods. Consider a two dimensional elastostatic problem in the domain $\Omega \subset \mathbf{R}^2$ with a set of nodes $\mathbf{X}_I$, the displacement $\mathbf{u}(\mathbf{x})$ at an arbitrary point $\mathbf{x}$ is approximated in a form similar to that in FEM

$$\mathbf{u}^h(\mathbf{x}) = \mathbf{N}(\mathbf{x})\mathbf{U} = \sum_I \mathbf{N}_I(\mathbf{x})\mathbf{U}_I \tag{1}$$

where $\mathbf{U}$ is the unknown vector of nodal displacement parameters. $\mathbf{N}(\mathbf{x})$ is the matrix of nodal shape functions

$$\mathbf{u}^h(\mathbf{x}) = \mathbf{N}(\mathbf{x})\mathbf{U} = \sum_I \mathbf{N}_I(\mathbf{x})\mathbf{U}_I \tag{2}$$

The nodal shape function $N_I(\mathbf{x})$ is constructed by MLS and can be written as

$$N_I(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{X}_I) w_I(\mathbf{x}) \boldsymbol{\alpha}(\mathbf{x}) \tag{3}$$

where $\mathbf{p}(\mathbf{x})$ is a vector of base functions which usually includes a complete basis of the polynomials to a given order, $w_I(\mathbf{x})$ a weight function and $\boldsymbol{\alpha}(\mathbf{x})$ the unknown vector. The unknown vector $\boldsymbol{\alpha}(\mathbf{x})$ can be determined by the so called reproducibility condition, i.e. the consistency condition

$$\mathbf{p}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I) N_I(\mathbf{x}) \tag{4}$$

Substitution of Eq.(3) into Eq.(4) leads to

$$\mathbf{A}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) = \mathbf{p}(\mathbf{x}) \tag{5}$$

where

$$\mathbf{A}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I)\mathbf{p}^{\mathrm{T}}(\mathbf{X}_I) w_I(\mathbf{x}) \tag{6}$$

The nodal MLS shape functions $N_I(\mathbf{x})$ can be obtained from Eq.(3) after the unknown vector $\boldsymbol{\alpha}(\mathbf{x})$ is solved from Eq.(5). Computation of the derivatives of the MLS shape functions is by taking the derivative of Eq.(3)

$$N_{I,i}(\mathbf{x}) = \mathbf{p}^{\mathrm{T}}(\mathbf{X}_I)\left[w_{I,i}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) + w_I(\mathbf{x})\boldsymbol{\alpha}_{,i}(\mathbf{x})\right] \tag{7}$$

where subscripts preceded by commas denote partial derivatives with respect to spatial coordinates. The unknown $\boldsymbol{\alpha}_{,i}(\mathbf{x})$ in Eq.(7) can be solved from the derivative of Eq.(6)

$$\mathbf{A}(\mathbf{x})\boldsymbol{\alpha}_{,i}(\mathbf{x}) = \mathbf{p}_{,i}(\mathbf{x}) - \mathbf{A}_{,i}(\mathbf{x})\boldsymbol{\alpha}(\mathbf{x}) \tag{8}$$

with

$$\mathbf{A}_{,i}(\mathbf{x}) = \sum_I \mathbf{p}(\mathbf{X}_I)\mathbf{p}^{\mathrm{T}}(\mathbf{X}_I) w_{I,i}(\mathbf{x}) \tag{9}$$

For a elastostatic problem on a 2D domain Ω bounded by Γ, EFG uses the classical displacement variational principle to construct the weak form $\delta\Pi(\mathbf{u})$, i.e

$$\delta\Pi(\mathbf{u}) = \int_\Omega \delta\boldsymbol{\varepsilon}^{\mathrm{T}}\mathbf{D}\boldsymbol{\varepsilon}\,\mathrm{d}\Omega - \int_{\Gamma_t} \delta\mathbf{u}^{\mathrm{T}}\bar{\mathbf{t}}\,\mathrm{d}\Gamma - \int_\Omega \delta\mathbf{u}^{\mathrm{T}}\mathbf{b}\,\mathrm{d}\Omega \tag{10}$$

where $\mathbf{D}$ is the material modulus, prefix δ denotes a variation and the strain is

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_{xx} & \varepsilon_{yy} & \gamma_{xy} \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial v}{\partial y} & \dfrac{\partial u}{\partial y} + \dfrac{\partial v}{\partial x} \end{bmatrix}^{\mathrm{T}} = \mathbf{B}\mathbf{U} = \sum_I \mathbf{B}_I \mathbf{U}_I \tag{11}$$

with

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_n \end{bmatrix} \quad \text{and} \quad \mathbf{B}_I = \begin{bmatrix} \dfrac{\partial N_I}{\partial x} & 0 \\ 0 & \dfrac{\partial N_I}{\partial y} \\ \dfrac{\partial N_I}{\partial y} & \dfrac{\partial N_I}{\partial x} \end{bmatrix} \tag{12}$$

By taking the variation, the following discretized equation can be obtained

$$\mathbf{K}\mathbf{U} = \mathbf{f} \tag{13}$$

where

$$\mathbf{K} = \int_\Omega \mathbf{B}^{\mathrm{T}}\mathbf{D}\mathbf{B}\,\mathrm{d}\Omega \qquad \mathbf{f} = \int_\Omega \mathbf{N}^{\mathrm{T}}\mathbf{b}\,\mathrm{d}\Omega + \int_{\Gamma_t} \mathbf{N}^{\mathrm{T}}\bar{\mathbf{t}}\,\mathrm{d}\Gamma \tag{14}$$

Consistent integration schemes for high order approximation

Consistent element-free Galerkin methods (CEFG) proposed by Duan *et.al.* [19] can pass patch tests exactly and consistently. This is due to special integration schemes with corrected

nodal derivatives are developed for EFG methods. The computation of the corrected nodal derivatives is based on the divergence theorem and it starts form the following equation:

$$\int_{\Omega_S} \tilde{N}_{I,i}(\mathbf{x})\mathbf{q}(\mathbf{x})\mathrm{d}\Omega = \int_{\Gamma_S} N_I(\mathbf{x})\mathbf{q}(\mathbf{x})n_i\mathrm{d}\Gamma - \int_{\Omega_S} N_I(\mathbf{x})\mathbf{q}_{,i}(\mathbf{x})\mathrm{d}\Omega \tag{15}$$

where $\tilde{N}_{I,i}(\mathbf{x})$ is the corrective derivatives, Ω_S bounded by Γ_S the cell for domain integration, $\mathbf{q}(\mathbf{x})$ the base obtained by

$$\mathbf{q}(\mathbf{x}) = \mathbf{p}_{,x}(\mathbf{x}) \cup \mathbf{p}_{,y}(\mathbf{x}) = \begin{bmatrix} q_1(\mathbf{x}) & q_2(\mathbf{x}) & \cdots & q_m(\mathbf{x}) \end{bmatrix}^{\mathrm{T}} \tag{16}$$

Using the divergence theorem to the right term of Eq.(15) leads to

$$\int_{\Omega_S} \tilde{N}_{I,i}(\mathbf{x})\mathbf{q}(\mathbf{x})\mathrm{d}\Omega = \int_{\Omega_S} N_{I,i}(\mathbf{x})\mathbf{q}(\mathbf{x})\mathrm{d}\Omega \tag{17}$$

Considering the $\mathbf{q}(\mathbf{x})$ is the base of the assumed Cauchy stress $\hat{\boldsymbol{\sigma}}$ space, Eq.(17) can be further rewritten as

$$\int_{\Omega_S} \hat{\boldsymbol{\sigma}}^{\mathrm{T}} \left(N_{I,i} - \tilde{N}_{I,i} \right) \mathrm{d}\Omega = \mathbf{0} \tag{18}$$

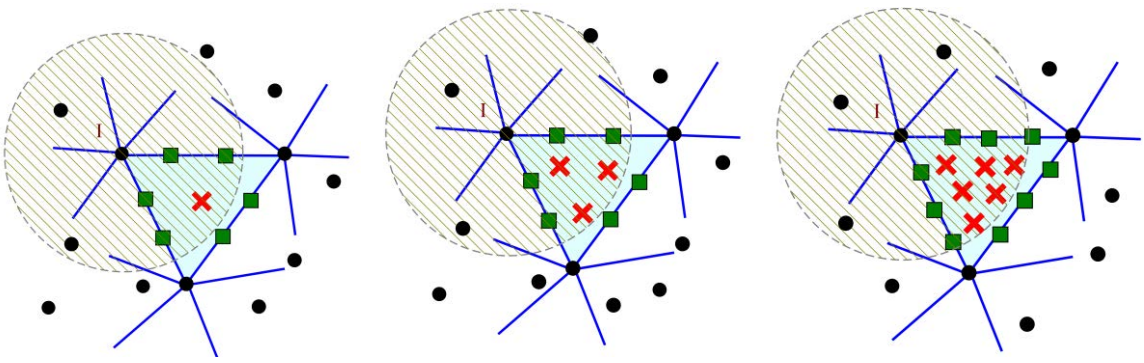
Since the corrective derivatives used for domain integrations represent the assumed strain, Eq.(18) can be replaced by:

$$\int_{\Omega_S} \hat{\boldsymbol{\sigma}}^{\mathrm{T}} \left(\boldsymbol{\varepsilon} - \tilde{\boldsymbol{\varepsilon}} \right) \mathrm{d}\Omega = \mathbf{0} \tag{19}$$

The above equation shows that the nodal corrected derivatives satisfy the orthogonality condition which means the nodal corrected derivative formulation can be derived from the Hu-Washizu three-field variational principle rational. The according consistent integration schemes are designed based on the numerical integration form of Eq.(15) with triangular integration cells as

$$\sum_{D=1}^{n_\Omega} W_D \frac{\partial \tilde{N}_I(\mathbf{x}_D)}{\partial x_i} \mathbf{q}(\mathbf{x}_D) = \sum_{e=1}^{3} \sum_{b=1}^{n_\Gamma} w_b N_I(\mathbf{x}_b)\mathbf{q}(\mathbf{x}_b) n_i^e - \sum_{D=1}^{n_\Omega} W_D N_I(\mathbf{x}_D)\mathbf{q}_{,i}(\mathbf{x}_D) \tag{20}$$

where W_D and w_b are, respectively, the integration weights of the evaluation points $\mathbf{x}_D$ in cells and $\mathbf{x}_b$ on edges, n_i^e the unit normal to the edge e. For linear, quadratic and cubic meshfree approximation, the following one, three and six integration schemes are designed [19] based on Eq.(20) as showed in Figure 1.



(a) (b) (c)

Figure 1. Schematic diagram of integration schemes for (a) Linear CEFG; (b) Quadratic CEFG and (c) Cubic CEFG; Dark dots denote approximation nodes, red crosses denote evaluation points for domain integration and green squares denote evaluation points for contour integration.

Once nodal corrected derivatives at the domain quadrature points are obtained, the discretized equation based on the Hu-Washizu variational principle is:

$$\tilde{\mathbf{K}}\mathbf{U} = \mathbf{f} \tag{21}$$

where

$$\tilde{\mathbf{K}} = \int_\Omega \tilde{\mathbf{B}}^{\mathrm{T}}\mathbf{D}\tilde{\mathbf{B}}\mathrm{d}\Omega \qquad \mathbf{f} = \int_\Omega \mathbf{N}^{\mathrm{T}}\mathbf{b}\mathrm{d}\Omega + \int_{\Gamma_t} \mathbf{N}^{\mathrm{T}}\bar{\mathbf{t}}\mathrm{d}\Gamma \tag{22}$$

$$\tilde{\mathbf{B}} = \begin{bmatrix} \tilde{\mathbf{B}}_1 & \tilde{\mathbf{B}}_2 & \cdots & \tilde{\mathbf{B}}_n \end{bmatrix} \quad \text{and} \quad \tilde{\mathbf{B}}_I^{\mathrm{T}} = \begin{bmatrix} \dfrac{\partial \tilde{N}_I}{\partial x} & 0 & \dfrac{\partial \tilde{N}_I}{\partial y} \\ 0 & \dfrac{\partial \tilde{N}_I}{\partial y} & \dfrac{\partial \tilde{N}_I}{\partial x} \end{bmatrix} \tag{23}$$

Numerical examples

Patch tests

Patch tests [19] are first investigated on a 2×2 domain with 5×5 nodes with irregular nodal distribution. Table 1, Table 2 and Table 3, respectively, compare the numerical results obtained by linear, quadratic and cubic approximations. It is observed that the proposed CEFG methods can pass patch tests in a consistent manner, i.e. linear CEFG is able to pass linear patch test, quadratic CEFG is able to pass up to quadratic patch test and cubic CEFG is able to pass up to cubic patch test. In contrast, the EFG methods fail to pass any patch test both in displacement and in energy.

Table 1 Patch test results: Linear methods

		Linear Patch Test	Quadratic Patch Test	Cubic Patch test
Linear EFG	E^{disp}	0.37E-06	0.75E-01	0.21E+00
	E^{eng}	0.33E-05	0.23E+00	0.53E+00
Linear CEFG	E^{disp}	0.35E-12	0.36E-01	0.13E+00
	E^{eng}	0.24E-11	0.16E+00	0.44E+00

Table 2 Patch test results: Quadratic methods

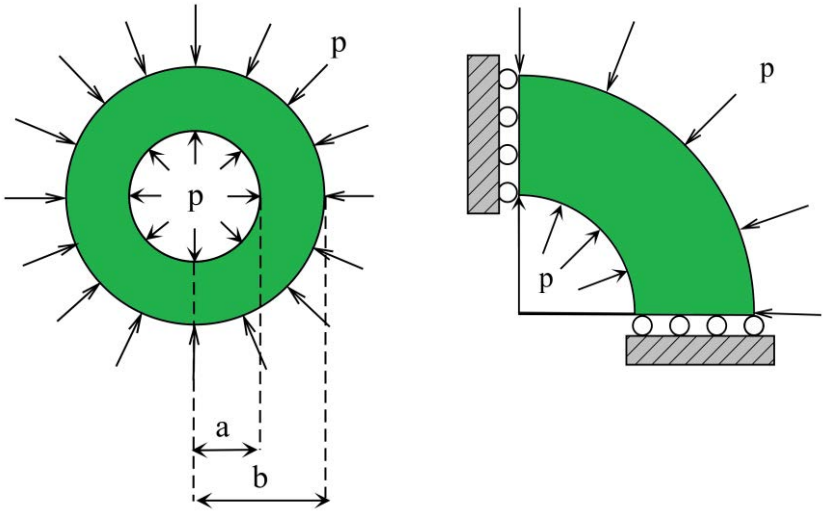
		Linear Patch Test	Quadratic Patch Test	Cubic Patch test
Quadratic EFG	E^{disp}	0.48E-05	0.91E-05	0.26E-01
	E^{eng}	0.36E-04	0.42E-04	0.97E-01
Quadratic CEFG	E^{disp}	0.27E-12	0.43E-12	0.19E-01
	E^{eng}	0.23E-11	0.16E-11	0.78E-01

Table 3 Patch test results: Cubic methods

		Linear Patch Test	Quadratic Patch Test	Cubic Patch test
Cubic EFG	E^{disp}	0.23E-04	0.99E-05	0.21E-05
	E^{eng}	0.15E-03	0.35E-04	0.59E-05
Cubic CEFG	E^{disp}	0.14E-11	0.85E-12	0.44E-12
	E^{eng}	0.72E-11	0.27E-11	0.13E-11

Pressurized hollow cylinder

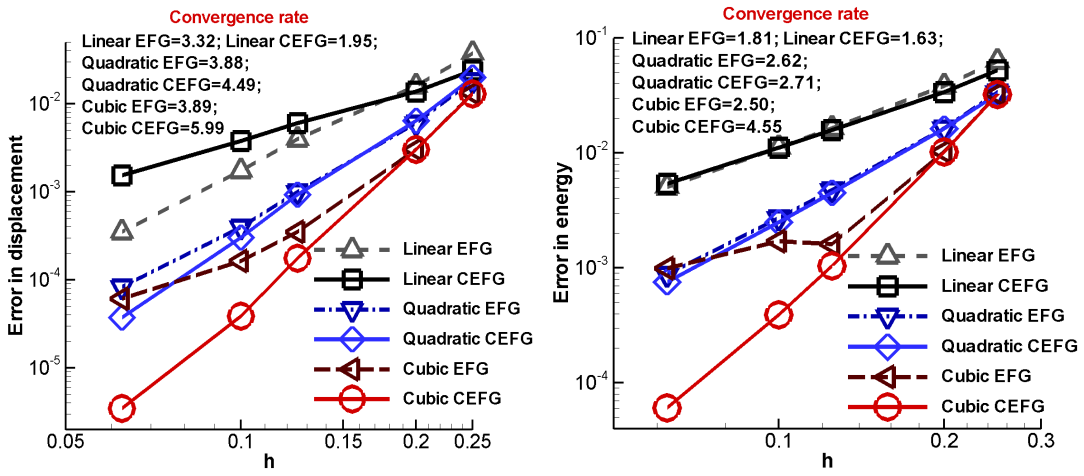
A hollow cylinder subjected to internal and external pressure as shown in Figure 2a is examined. As shown in Figure 2b, due to two-fold symmetry, only the first quadrant is modeled.



(a) (b)

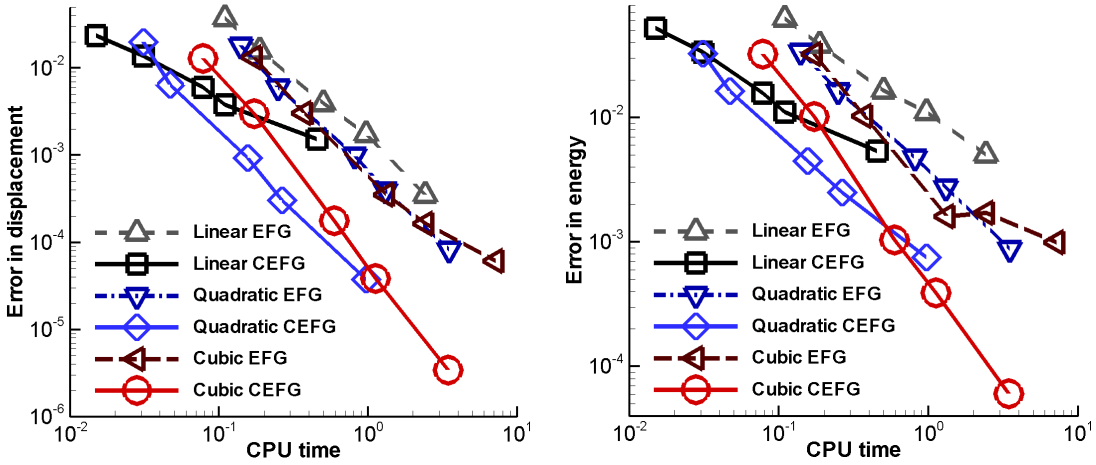
Figure 2. Plate with a hole problem: (a) schematic diagram; (b) solution domain

Five regular grids with the typical size of the discretization $h = 0.25, 0.2, 0.125, 0.1, 0.0625$ are employed in convergence study and the results are plotted in Figure 3. The proposed CEFG shows better accuracy and convergence rate than EFG both in displacement and in energy. Figure 4 compares the computational efficiency of displacement and energy. The proposed CEFG is much more efficient than standard EFG since less integration points are used. Figure 5 shows the $\sigma_{yy,y}$ fields obtained by the six methods. Clearly, the proposed cubic CEFG method obtains the best $\sigma_{yy,y}$ field which is very smooth. In contrast, considerable oscillations present in the result of the cubic EFG method. The $\sigma_{yy,y}$ fields given by quadratic CEFG and quadratic EFG methods are similar. Same observation is applied to linear CEFG and linear EFG methods.



(a) (b)

Figure 3. Convergence of the plate with a hole problem: (a) displacement; (b) energy



(a) (b)

Figure 4. Computational efficiency of the plate with a hole problem: (a) displacement; (b) energy

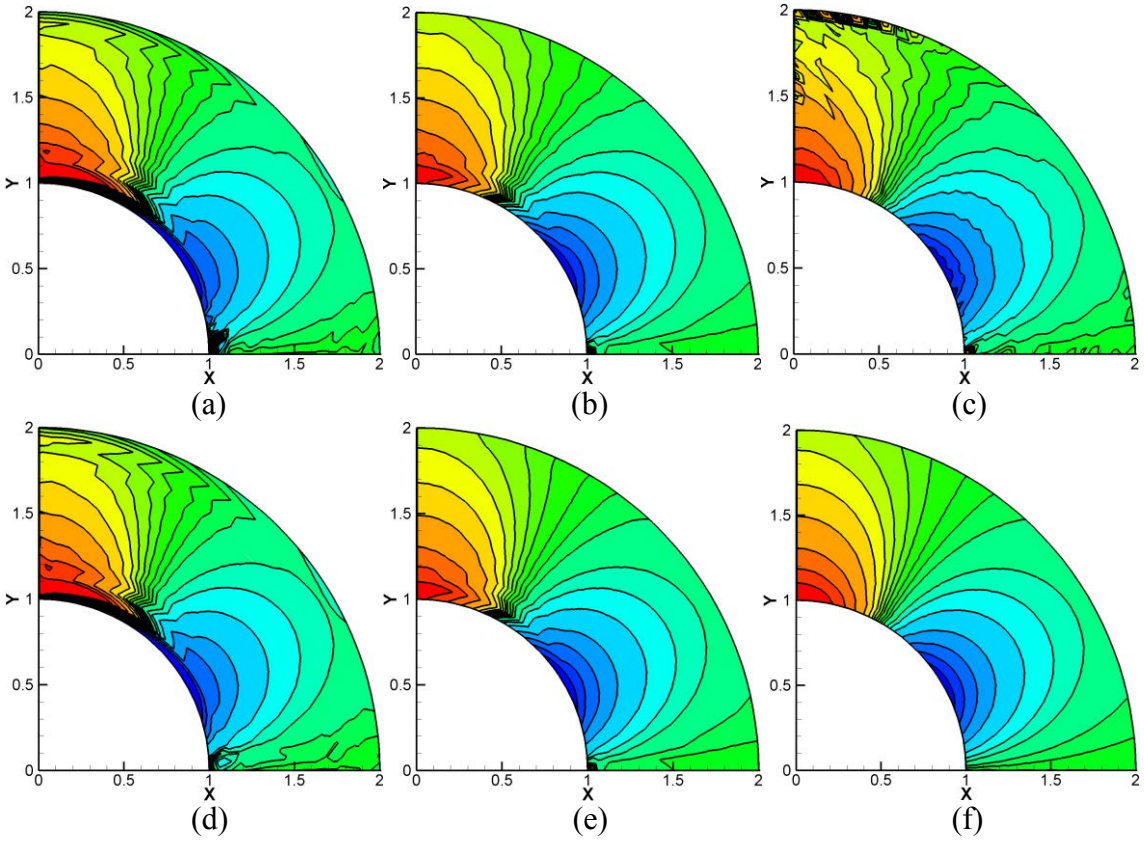


Figure 5. Comparison of $\sigma_{yy,y}$ fields of the pressurized hollow cylinder problem obtained by: (a) Linear EFG; (b) Quadratic EFG; (c) Cubic EFG;(d) Linear CEFG; (e) Quadratic CEFG; (f) Cubic CEFG.

Application to thin plates and shells

Due to the high smoothness of the meshfree approximations, various works for thin-plate and –shell problems [20-26] have been investigated by meshfree methods. However, high order quadratures are commonly employed to evaluate the Galerkin weak form, which is computationally expensive. In this section, consistent element-free Galerkin method is further

extended into thin-plates and thin-shells problems. For such high-order differential equations, cubic approximation is employed in this section.

Thin plates

Since the government equation of thin plates problem is a fourth-order partial differential equation, the kernel idea is to correct the nodal second order derivatives which leads to a curvature smoothing(CS) formulation. The consistent curvature smoothing formulation in each integration cell Ω_k is as follows:

$$\begin{aligned}\int_{\Omega_k}\tilde{N}_{I,\alpha\beta}\mathbf{q}(\mathbf{x})\mathrm{d}\Omega=&\int_{\Gamma_k}\frac{1}{2}\left[N_{I,\alpha}\mathbf{q}(\mathbf{x})n_\beta+N_{I,\beta}\mathbf{q}(\mathbf{x})n_\alpha\right]\mathrm{d}\Gamma\\&-\int_{\Gamma_k}\frac{1}{2}N_I\left[\mathbf{q}_{,\beta}(\mathbf{x})n_\alpha+N\mathbf{q}_{,\alpha}(\mathbf{x})n_\beta\right]\mathrm{d}\Gamma\\&-\int_{\Omega_k}N_I\mathbf{q}_{,\alpha\beta}(\mathbf{x})\mathrm{d}\Omega\end{aligned}\tag{24}$$

For cubic meshfree approximation which leads to a linear curvature smoothing(LCS) formulation, the above equation reduces to:

$$\begin{aligned}&\int_{\Omega_k}\tilde{N}_{I,\alpha\beta}\mathrm{d}\Omega=\int_{\Gamma_k}\frac{1}{2}\left(N_{I,\alpha}n_\beta+N_{I,\beta}n_\alpha\right)\mathrm{d}\Gamma\\&\int_{\Omega_k}\tilde{N}_{I,\alpha\beta}x\mathrm{d}\Omega=\int_{\Gamma_k}\frac{1}{2}x\left(N_{I,\alpha}n_\beta+N_{I,\beta}n_\alpha\right)\mathrm{d}\Gamma-\int_{\Gamma_k}\frac{1}{2}N_I\left(\delta_1^\beta n_\alpha+\delta_1^\alpha n_\beta\right)\mathrm{d}\Gamma\\&\int_{\Omega_k}\tilde{N}_{I,\alpha\beta}y\mathrm{d}\Omega=\int_{\Gamma_k}\frac{1}{2}y\left(N_{I,\alpha}n_\beta+N_{I,\beta}n_\alpha\right)\mathrm{d}\Gamma-\int_{\Gamma_k}\frac{1}{2}N_I\left(\delta_2^\beta n_\alpha+\delta_2^\alpha n_\beta\right)\mathrm{d}\Gamma\end{aligned}\tag{25}$$

Eq.(25) are used to correct the nodal second order derivatives.

Numerical examples

A square with four simply supported edges under uniform load

As showed in Figure 6, a 1×1 square subjected to uniform load with the magnitude $q=1$ is tested. The four edges of the square are simply supported. The Young's modules and the Poission ratio are $E=1\times10^{10}$ and $\nu=0.3$. The thickness of the plate is $t=0.001$.

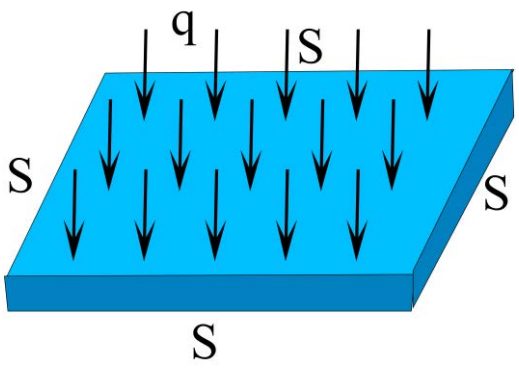
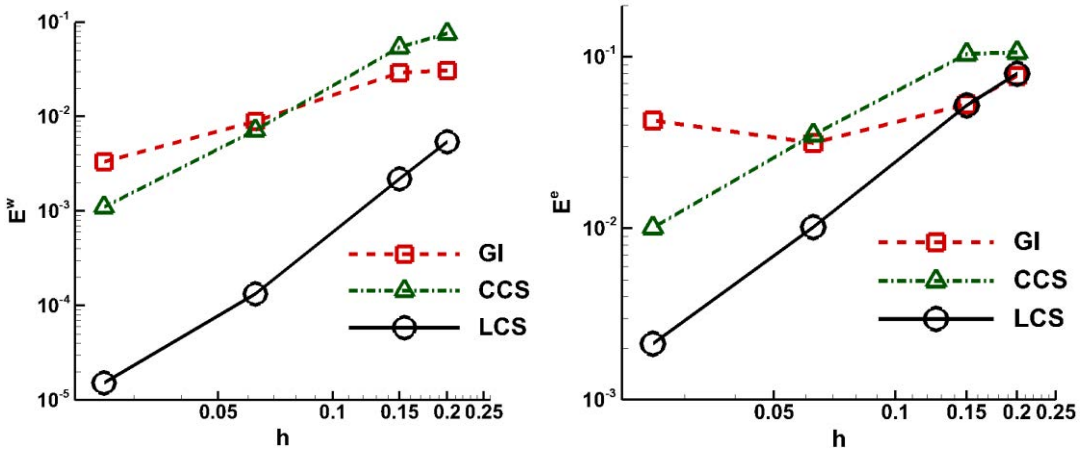


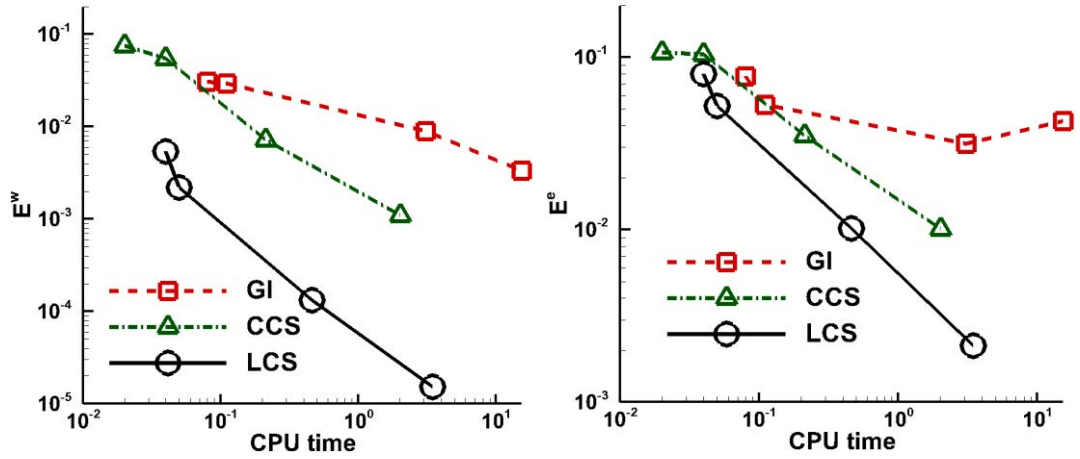
Figure 6. A square with four simply supported edges under uniform load

Four irregular grids are used for convergence study in this problem and the results are showed in Figure 7. The proposed LCS achieves the highest accuracy for both deflection and energy. Figure 8 shows the efficiency results of the deflection and energy. It is observed that the proposed LCS is the most efficiency method for both deflection and energy.



(a) (b)

Figure 7. Convergence of the square with four simply supported edges under uniform load problem: (a) deflection; (b) energy.



(a) (b)

Figure 8. Efficiency of a square with four simply supported edges under uniform load problem: (a) deflection; (b) energy

Simply supported circular plate under uniform load

A circular simply supported plate under uniform load as shown in Figure 9 is next investigated. The radius of the plate is $R=2$ and the thickness of the plate is $t=0.001$.Material parameters are the Young's modules $E=1\times10^{10}$ and Poission ratio $v=0.3$. The magnitude of applied uniform load is $q=0.1$.

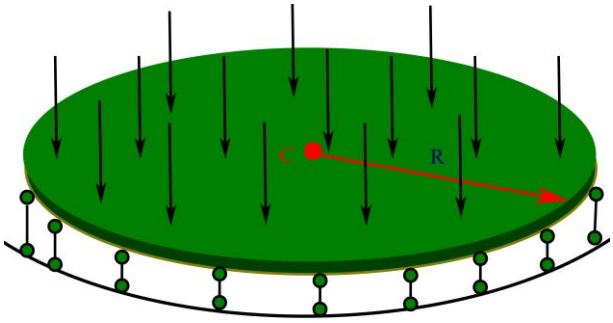


Figure 9. Simply supported circular plate under uniform load

The full model is used for analysis with three grids containing 168, 414 and 1547 nodes, respectively. The convergence of the deflection of the center point w_c is showed in figure 10. It is clearly that the proposed LCS agrees best with the exact solution.

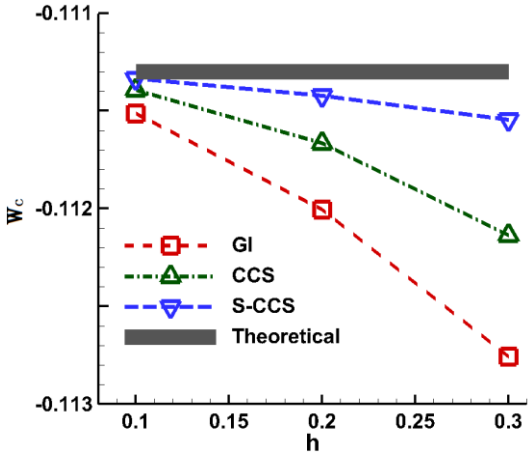


Figure 10. The convergence result of the deflection of the center point W_C

Thin shells problem

The geometrically exact thin shell model [26] is considered in this paper. A linear strain smoothing formulation is proposed in parametric space for thin shell analysis which contains the membrane strain smoothing and curvature smoothing:

$$\int_{\mathscr{A}_S} \tilde{N}_{I,\alpha}(\xi)\mathbf{q}(\xi)d\mathscr{A}_S = \int_{\partial\mathscr{A}_S} N_I(\xi)\mathbf{q}(\xi)n_\alpha dl_\xi - \int_{\mathscr{A}_S} N_I(\xi)\mathbf{q}_{,\alpha}(\xi)d\mathscr{A}_S \tag{26}$$

$$\int_{\mathscr{A}_S} \tilde{N}_{I,\alpha\beta}(\xi)\mathbf{q}(\xi)d\mathscr{A}_S = \int_{\partial\mathscr{A}_S} N_{I,\alpha}(\xi)\mathbf{q}(\xi)n_\beta dl_\xi - \int_{\mathscr{A}_S} N_{I,\alpha}(\xi)\mathbf{q}_{,\beta}(\xi)d\mathscr{A}_S \tag{27}$$

where $\mathscr{A}_S$ is the background integration cell in parametric space. $\mathbf{q}(\xi)$ is chosen as a linear base of the parametric space $\mathscr{A}$:

$$\mathbf{q}(\xi) = \begin{pmatrix} 1 \\ \xi^1 \\ \xi^2 \end{pmatrix} \tag{28}$$

Numerical examples

Pinched hemispherical shell with 18° *hole*

Hemispherical shell problem is tested. The hemispherical has an 18° hole at the top. The thickness of the shell is $h = 0.04$ and the radius $R = 10$. The material parameters contain Young's modulus $E = 6.825e7$ and Poisson's ratio $v = 0.3$. The shell is subjected to two pairs of load which are equal and opposite along the X- and Y- axes.

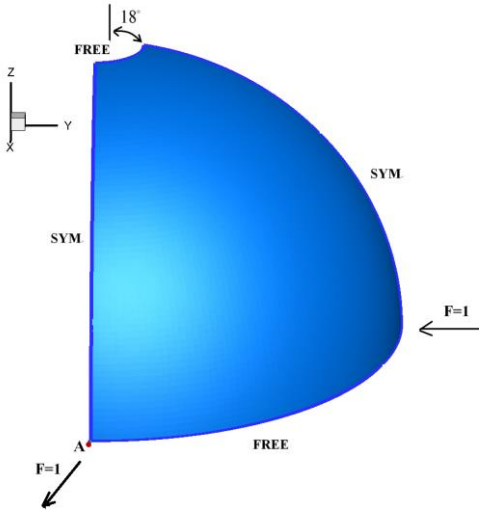


Figure 11. Schematic diagram of hemispherical shell with 18° hole problem

Due to the symmetry, only a quarter of the geometry is modelled, see figure 11. The convergence result is shown in figure 12. It can be seen that the proposed method agrees well with the reference solution and performs better than GI-16 and CCS with the densest node configuration.

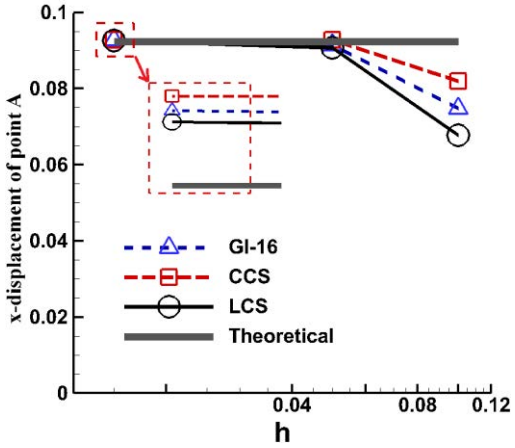


Figure 12. Convergence of the X-displacement of point A

Scordelis-Lo roof

The Scordelis-Lo roof problem [26] is a benchmark problem for a curved shell analysis. The geometrical parameter of the roof are: the length $L=50$, the radius $R=25$, the thickness $h=0.25$ and the span angle $\theta=80^{\circ}$. The material properties are: Young's modulus $E=4.32e8$ and the Poisson ratio $v=0$. The roof is loaded by a self-weight $q=90$. The model is fixed by two opposite rigid diaphragms and the other two edges are free. Four irregular grids with 36, 121, 441 and 1684 nodes are employed for the convergence study and the convergence curves are presented in Figure 13.

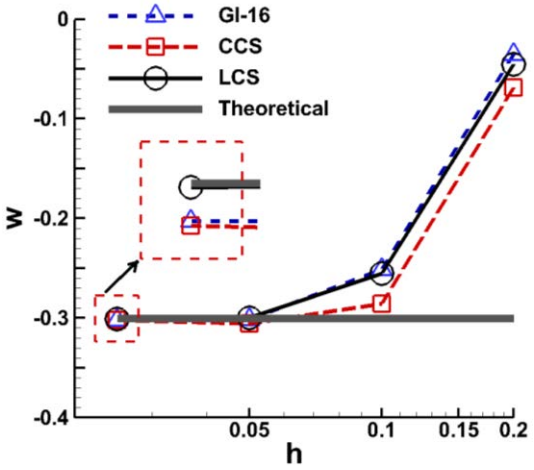
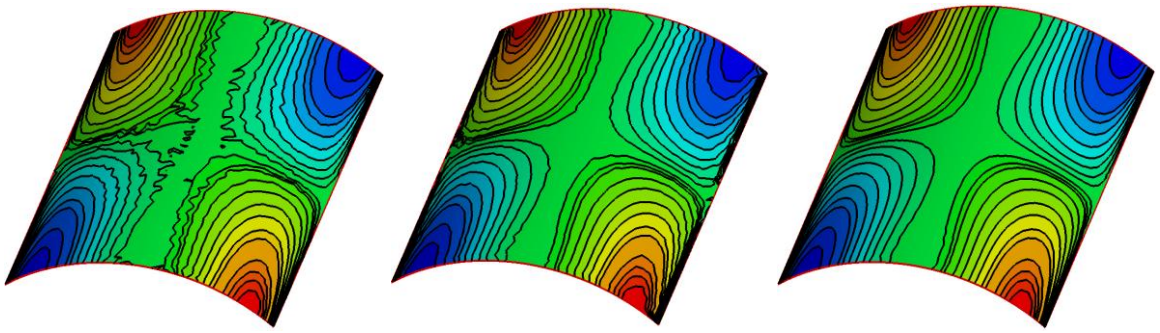


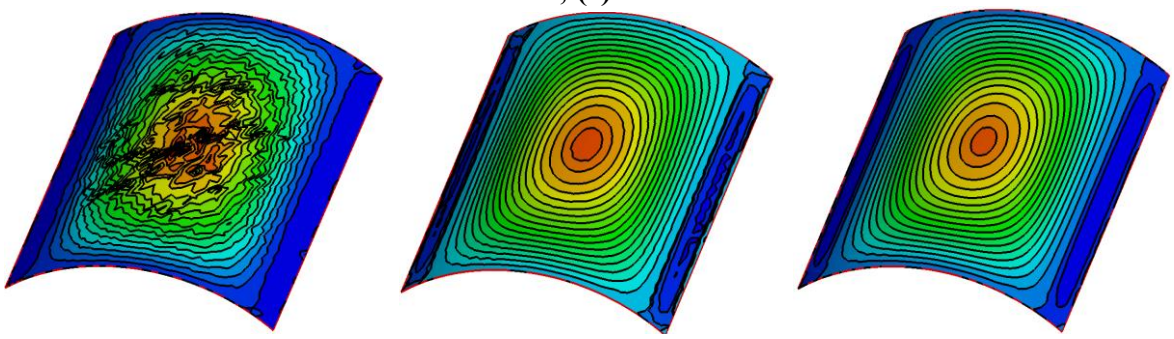
Figure 13. Convergence of the mid-side vertical displacements of Scordelis-Lo roof problem

Figure 14 shows the contours of membrane stress n_{12} by GI-16, CCS and LCS. The proposed LCS achieves smoothed contour while mildly spurious oscillations are observed by GI-16 and CCS. Figure 15 shows the contours of bending stress m_{11} by GI-16, CCS and LCS. GI-16 results severely spurious oscillations although it uses the most quadrature points. LCS achieves much better bending stress m_{11} contour, but mild oscillations still appear. Only the proposed LCS results smoothed bending stress m_{11} contour. This shows the superiority of the proposed LCS.



(a) (b) (c)

Figure 14. Membrane stress n_{12} contour of Scordelis-Lo roof problem by: (a) GI-16; (b) CCS; (c) LCS



(a) (b) (c)

Figure 15. Bending stress m_{11} contour of Scordelis-Lo roof problem by: (a) GI-16; (b) CCS; (c) LCS

Application to crack problems

The method has been developed to deal with a problem of linear elastic fracture mechanics. There is no enrichment function for the discontinuous displacement field.

Crack description

A phantom-node method[27] is developed to describe cracks. We start with the discontinuous displacement field in an element

$$u(\mathbf{x})=\sum_{I=1}N_I(\mathbf{x})\left\{\mathbf{u}_I+\mathbf{q}_I\left[H\left(f(\mathbf{x})\right)-H\left(f(\mathbf{x}_I)\right)\right]\right\} \tag{29}$$

where $H(\cdot)$ is the Heaviside step function given by

$$H(x)=\begin{cases}1 & \forall x>0\\ 0 & \forall x\le 0\end{cases} \tag{30}$$

and $f(\mathbf{x})=0$ represents the position of the crack. So, Eq.(29) are subdivided each term into parts that are associated with $f(\mathbf{x})<0$ and $f(\mathbf{x})>0$, we have

$$\mathbf{u}=\sum_{I=1}\left[\left(\mathbf{u}_I-\mathbf{q}_I\right)H_I^+N_I(1-H)+\mathbf{u}_IH_I^-N_I(1-H)+\mathbf{u}_IH_I^+N_IH+\left(\mathbf{u}_I+\mathbf{q}_I\right)H_I^-N_IH\right] \tag{31}$$

If we then let

$$\mathbf{u}_I^1=\begin{cases}\mathbf{u}_I & f(\mathbf{x}_I)<0\\ \mathbf{u}_I-\mathbf{q}_I & f(\mathbf{x}_I)>0\end{cases} \tag{32}$$

$$\mathbf{u}_I^2 = \begin{cases} \mathbf{u}_I + \mathbf{q}_I & f(\mathbf{x}_I) < 0 \\ \mathbf{u}_I & f(\mathbf{x}_I) > 0 \end{cases} \tag{33}$$

We can write the displacement field as

$$\mathbf{u}(\mathbf{x}) = \sum_{I \in S_1} \mathbf{u}_I^1 N_I(\mathbf{x}) H(-f(\mathbf{x})) + \sum_{I \in S_2} \mathbf{u}_I^2 N_I(\mathbf{x}) H(f(\mathbf{x})) \tag{34}$$

where S_1 and S_2 are the index sets of the relative nodes of superposed element 1 and 2, respectively. As can be seen from Figure 16, each element contains original real nodes and phantom nodes.

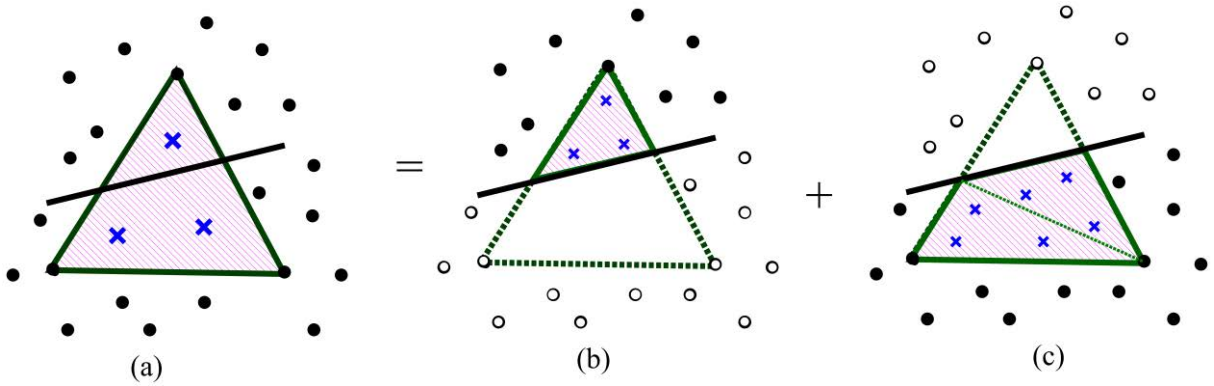


Figure 16 The decomposition of a cracked element into two elements

Therefore, elements are overlapped on the position of the crack, and they are partially integrated to implement the discontinuous displacement across the crack.

Numerical examples

The discontinuous patch test

This example was devised by Dolbow and Devan[28]. Shown in Figure.17 is a 3×3 square domain that is subjected to a discontinuous horizontal traction with $t=10$. It is assumed that the domain is completely bisected into two regions by the failure surface(i.e. red line).

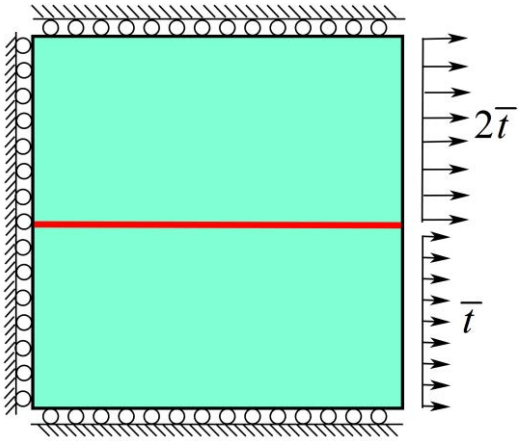


Figure.17 Schematic diagram of the discontinuous patch test configuration and loading

A 4×4 grid is used for analysis. The distributions of σ_{xx} are shown in Figure.18 for compressible case and incompressible case, respectively. The results suggest that the method can exactly pass the test in the both cases.

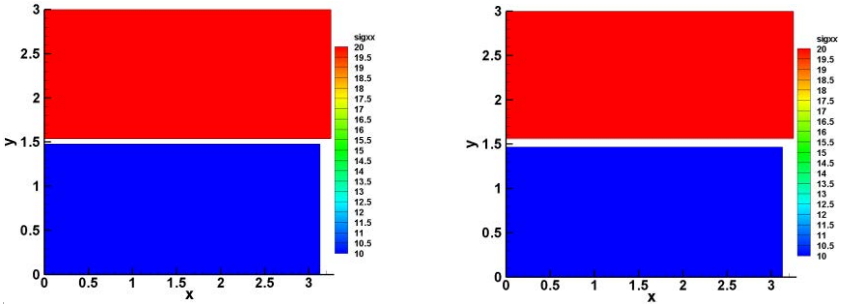


Figure. 18 σ_{xx} distribution of: (a) compressible case ;(b) incompressible case

Edge cracked plates under tension or shear

Consider a plane stress plate of width $b=7$ and height $l=16$ with an edge crack length of $a=b/2=3.5$. The material properties are $E=10^7$, and $\nu=0.3$.The plate is subjected to a tension $\sigma=1$ at the top(see Figure.19a) or sustains a shear $\tau=1$ on the top edge(see Figure.19b).

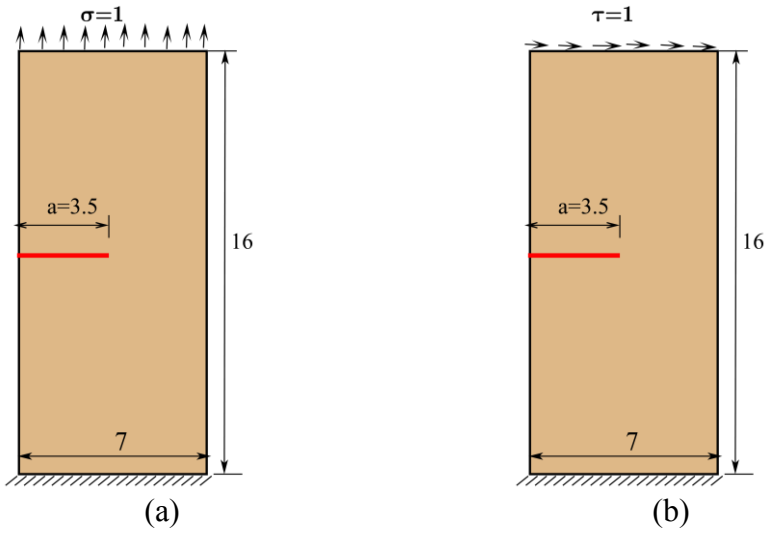


Figure.19 Schematic diagram of the edge cracked plate under:(a) tension; (b) shear

The analytical solution for the plate under tension is given by

$$K_I = F(\lambda)\sigma\sqrt{\pi a} \tag{35}$$

with $F(\lambda)=1.12-0.231\lambda+10.55\lambda^2-21.72\lambda^3+30.39\lambda^4$.The values of K_I and K_{II} for the shear case in [29] are used as the reference solution:

$$K_I = 34.0,\ K_{II} = 4.55 \tag{36}$$

Table.4 The results of normalized K_I and K_{II}

Cases	EFG		CEFG	
	K_I	K_{II}	K_I	K_{II}
Tension case	1.067	--	1.008	--
Shear case	1.062	0.991	1.005	0.998

A 30×71 regular grid is used for the evaluating of stress intensity factors. Table.4 lists the evaluating results of normalized K_I and K_{II} with two methods for tension case and shear case, respectively. The results show that CEFG method is able to give more accurate stress intensity factors than the standard EFG method.

Crack growth from a fillet

This example shows the growth of a crack from a fillet in a structural member. The experiment to be modelled is shown in Figure.20, with the computational domain outlined with a dashed line. The material properties are $E = 200Gpa$, and $\nu = 0.3$, respectively. The applied load is $P = 1.0N$. The initial crack length is $a_0 = 5mm$. To model a very thick beam, the displacement is fixed along the entire bottom of the computational domain, that is, rigid constraint. A flexible constraint is idealized to model a very thin beam.

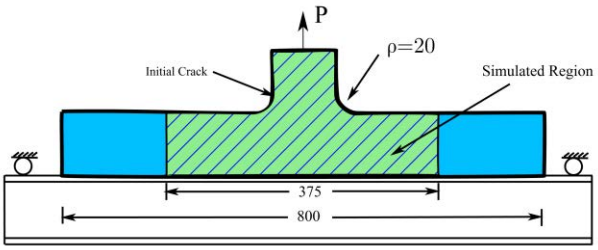


Figure.20 Experimental configuration and simulated region for a fillet problem

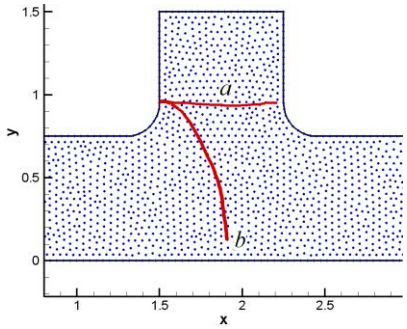


Figure.21 Crack growth paths of two different constraints: a. Flexible constraint; b. Rigid constraint

This experiment is presented in Reference [30] to investigate the effect of the thickness of the lower I-beam on crack growth. Figure.21 shows the crack paths predicted by the proposed method for a thick I-beam and a thin I-beam, respectively. These results agree well with the experimental results [30].

Conclusions

Consistent element-free Galerkin method and its applications are presented. For elastostatic solids, CEFG performs much better than standard EFG in terms of accuracy, convergence and efficiency, according to the numerical results. Application of the high order CEFG method to thin plates and shells as well as crack problems is also presented. It is demonstrated that the CEFG method is really promising in these applications.

Acknowledgement

The authors are pleased to acknowledge the support of this work by Science Challenge Project (No. JCKY2016212A502), the National Key Research and Development Program of China (2016YFC1402705, 2016YFC1402706), the National Natural Science Foundation of China (11372066, 11232003), the Fundamental Research Funds for the Central Universities (DUT15LK07), the open funds of the state key laboratory of water resources and hydropower engineering science (2015SGG03), the open funds of the state key laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP2016K007).

References

[1] Nayroles B, Touzot G, Villon P. (1992) Generalizing the finite element methods: Diffuse approximation and diffuse elements. *Computational Mechanics* **10**, 307-318.

[2] Belytschko T, Lu YY, Gu L. (1994) Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering* **37**, 229–256.
[3] Liu W.K., Jun S, Zhang Y.F. (1994) Reproducing kernel particle methods. *International Journal for Numerical Methods in Fluids* **20**, 1081–1106.
[4] Duarte CA, Oden JT. (1996) H-p clouds—an h-p meshless method. *Numer. Methods Partial Differ. Equations* **12**, 673-705.
[5] Zhang X. (2001) Least-Squares collocation meshless method. *International Journal for Numerical Methods in Engineering* **51**, 1089–1100.
[6] Liu GR, Gu YT and Dai K. (2004) Assessment and applications of point interpolation methods for computational mechanics, *International Journal for Numerical Methods in Engineering* 59, 1373-1397.
[7] Rabczuk T, Belytschko T. (2005) Adaptivity for structured meshfree particle methods in 2D and 3D. *International Journal for Numerical Methods in Engineering* **63**, 1559–1582.
[8] Dolbow J, Belytschko T. (1999) Numerical integration of the Galerkin weak form in meshfree methods. *Computational Mechanics* **23**, 219–230.
[9] Beissel S, Belytschko T. (1996) Nodal integration of the element-free Galerkin method. *Computer Methods in Applied Mechanics and Engineering* **139**, 49–74.
[10] Liu GR, Zhang GY, Wang YY, Zhong ZH, Li GY, Han X. (2007) A nodal integration technique for meshfree radial point interpolation method (NI-RPIM). *International Journal of Solids and Structures* **44**, 3840–3860.
[11] Dyka CT, Randles PW, Ingel RP. (1997) Stress points for tension instability in SPH. *International Journal for Numerical Methods in Engineering* **40**, 2325–2341.
[12] Fries TP, Belytschko T. (2008) Convergence and stabilization of stress-point integration in mesh-free and particle methods. *International Journal for Numerical Methods in Engineering* **74**, 1067–1087.
[13] Liu Y, Belytschko T. (2010) A new support integration scheme for the weak form in mesh-free methods. *International Journal for Numerical Methods in Engineering* **82**, 699–725.
[14] Chen JS, Wu CT, Yoon S, You Y. (2001) A stabilized conforming nodal integration for Galerkin mesh-free methods. *International Journal for Numerical Methods in Engineering* **50**, 435–466.
[15] G. Y. Zhang, G. R. Liu, Y. Y. Wang H. T. Huang, Z. H. Zhong, G. Y. Li and X. Han. (2007) A linearly conforming point interpolation method (LC-PIM) for three-dimensional elasticity problems. *International Journal for Numerical Methods in Engineering* **72**, 1524-1543.
[16] Liu G, Dai K, Nguyen T. (2007) A smoothed finite element method for mechanics problems. *Computational Mechanics* **39**, 859-877.
[17] Puso MA, Chen JS, Zywicz E, Elmer W. (2008) Meshfree and finite element nodal integration methods. *International Journal for Numerical Methods in Engineering* **74**, 416–446.
[18] Duan QL, Li XK, Zhang HW, Belytschko T. (2012) Second-order accurate derivatives and integration schemes for meshfree methods. *International Journal for Numerical Methods in Engineering* **92**, 399–424.
[19] Duan QL, Gao X, Wang BB et.al. (2014) Consistent element-free Galerkin method. *International Journal for Numerical Methods in Engineering* **99**, 79–101.
[20] Krysl P, Belytschko T. (1995) Analysis of thin plates by the element-free Galerkin method. *Computational Mechanics* **17**, 26–35.
[21] Cui X, Liu G, Li G et.al. (2011) A thin plate formulation without rotation DOFs based on the radial point interpolation method and triangular cells. *International Journal for Numerical Methods in Engineering* **85**, 958–986.
[22] Chen JS, Wang D. (2006) A constrained reproducing kernel particle formulation for shear deformable shell in Cartesian coordinate. *International Journal for Numerical Methods in Engineering* **68**, 151–172.
[23] Wang D, Chen JS. (2008) A Hermite reproducing kernel approximation for thin-plate analysis with sub-domain stabilized conforming integration. *International Journal for Numerical Methods in Engineering* **74**, 368–390.
[24] Krysl P, Belytschko T. (1996) Analysis of thin shells by the element-free Galerkin method. *International Journal of Solids and Structures* **33**, 3057-3080.
[25] Rabczuk T, Areias P, Belytschko T. (2007) A meshfree thin shell method for non-linear dynamic fracture. *International Journal for Numerical Methods in Engineering* **72**, 524–548.
[26] Millan D, Rosolen A, Arroyo M. (2011) Thin shell analysis from scattered points with maximum-entropy approximations. *International Journal for Numerical Methods in Engineering* **85**, 723–751.
[27] Song J.H., Areias P. M. A. and Belytschko T.(2006) A method for dynamic crack and shear band propagation with phantom nodes, *Int. J. Numer. Meth. Engng* **67**, 868–893.
[28] Dolbow J, Devan A.(2004) Enrichment of enhanced assumed strain approximations for representing strong discontinuities: addressing volumetric incompressibility and the discontinuous patch test. *Int J Numer Methods Eng*, **59**:47–67.
[29] Yau J.F., Wang S.S., Corten H.T. (1980) A mixed-mode crack analysis of isotropic solids using conservation laws of elasticity. *J Appl Mech ASME* **47**,335-341.
[30] Sumi Y., Yang C., Wang Z. N. (1996) Morphological aspects of fatigue crack propagation. Part II—effects of stress biaxiality and welding residual stress. *International Journal of Fracture* **82**(3), 221-235.

Applying virtual stent deployment to study flow-diversion treatment for intracranial aneurysms: the effect of stent compaction on post-treatment wire configuration

†*M. Zhang[1, 2], Y. Li[1, 2], D.I. Verrelli[1], W. Chong[3, 4], M. Ohta[5], and Y. Qian[1]

[1] Faculty of Medicine and Health Sciences, Macquarie University, Australia.
[2] Graduate School of Engineering, Tohoku University, Japan.
[3] Neuroradiology Department, Monash Medical Centre, Australia.
[4] Department of Surgery, Faculty of Medicine, Monash University, Australia
[5] Institute of Fluid Science, Tohoku University, Japan

[*] Presenting author: mingzi.zhang@biofluid.ifs.tohoku.ac.jp
[†] Corresponding author: mingzi.zhang@biofluid.ifs.tohoku.ac.jp

Abstract

Flow-diverting (FD) stent implantation has become a popular treatment mode for intracranial aneurysms (IAs). The stent wire configurations post-treatment can greatly affect the treatment outcomes. However, it remains a challenge to predict the stent wire configurations prior to a treatment. In this study, we propose to compare the FD stent structures and wire configurations between treatments with FD stents of different diameters being deployed at different compaction levels.

We adopted a recently reported spring–mass model to virtually implant FD stents of three diameters — 4.0, 4.5, and 5.0 mm, with each diameter modelled at three compaction levels (no compaction, a low compaction, and a high compaction) — into two clinically observed IAs: one successfully treated and the other unsuccessfully treated previously with a single FD stent. We then examined the morphological differences in stent wire configurations across different treatment scenarios, and quantified the porosity achieved in each such scenario.

This led us to two main findings. Firstly, at the same compaction level, the porosity differences attributable to device diameter were limited (SD < 2%). Secondly, stent deployment with some compaction could effectively reduce the FD wire porosity within the aneurysm ostium — a low compaction reduced the porosity by around 10%, and a high compaction reduced the porosity by around 30%.

The FD stent structures observed from virtual deployment can be used in the subsequent aneurysmal haemodynamic simulations. Thus, the virtual stent deployment strategy, together with the stent compaction measurement and simulation technique presented in this study, may contribute to research into flow-diversion treatment planning.

Keywords: Virtual stent deployment, Spring–mass model, Flow-diverting (FD) stents, Porosity, Treatment planning.

Introduction

Intracranial aneurysms (IAs) are common cerebrovascular disorders that can be identified as balloon-like dilations from a segment of brain artery [1, 2]. Untreated IAs can be dangerous, since subarachnoid haemorrhage could be induced due to aneurysm rupture. As an endovascular

therapy, flow-diverting (FD) stent implantation is intended to reduce the blood velocity within the aneurysm sac, thereby promoting thrombotic occlusion of the aneurysm [2–4].

A favourable FD treatment outcome depends heavily upon sufficient flow diversion produced by the implanted device, whilst flow diversion is closely associated with the stent size and deployment procedure determined prior to a treatment [5]. Recently, stent deployment with a "push and pull" manoeuvre was introduced as an effective approach for a single stent to achieve lower porosity within the aneurysm ostium [6–8]. Along with other factors that may change the stent structure post-deployment (*e.g.* stent diameter), it remains a challenge to predict the stent wire configuration, so as to predict the treatment outcomes.

In this study, we propose to investigate the procedure of stent deployment when a compaction technique is applied, and to quantify the stent wire porosity post-deployment across various treatment scenarios. To realise this objective, we employed a numerical method for virtual stent deployment simulation. Furthermore, we defined a variable to characterise the level of stent compaction, to enable quantitative assessments to be made. Adopting these approaches, we examined a total of 18 deployment procedures for two patient-specific aneurysms, considering different combinations of stent diameter and compaction ratio.

Methods

Patient Aneurysm Geometries and Model FD Stent

We accessed the patient aneurysm geometries after institutional ethics approval was granted. Two patient aneurysms were studied — one successfully treated case in which complete aneurysm occlusion was confirmed 6 months after treatment, and the other an unsuccessfully treated case in which an aneurysm lobule was identified by digital subtraction angiography (DSA) scans. Both patients were treated with the commercially available *Silk+* stent (Balt Extrusion, Montmorency, France).

The *Silk+* stents comprise a total of 48 braided Nitinol-alloy filaments, with each filament having an average diameter of 35 μm [9]. We employed a fast geometric deployment approach — the spring–mass analogue [10–13] — in the simulation of FD stent deployment. In this spring–mass analogue, a series of assumptions were adopted:

1) the FD wire filaments were represented by a group of centreline trajectories without an explicit thickness;
2) the intersections of FD wire filaments were assumed to be mass points (nodes), which were connected with each other by fictitious springs based upon the topology of the FD stent being modelled;
3) the stiffness of each spring was related to the length and thickness of the FD wire filament being represented.

With these simplifications applied, the internal force of a *Silk+* stent could be calculated following the 3D Hooke's law:

$$F_i = \sum_{j=1}^{n_i} k_{ij}(\delta_j - \delta_i), \tag{1}$$

in which δ_i and δ_j refer to the displacements of nodes i and j; k_{ij} denotes the stiffness of the fictitious spring connecting node i and its neighbour j; and n_i is the number of nodes directly connected to node i.

Virtual Stent Deployment

The *Silk+* stent was assumed to be initially crimped in alignment with the centreline of the parent artery, and then expanded to its unloaded condition. The expansion process was driven by the internal restoring forces from the mass points, and was constrained by vascular geometries — the arterial wall, *etc*. To constrain the nodal movements within the boundaries of the parent artery, a wall-touch detection algorithm was invoked after each step of nodal movement. The benchmark of simulation convergence was set as the displacement of each mass point being less than 1 μm after a step of nodal movement.

Measurement of the Compaction Ratio

The stent compaction ratio was measured in accordance with the maximum compaction length $\Delta L_{\max}$ of a FD stent, a parameter defined as the maximum longitudinal distance that an additional length of FD stent could possibly be compacted into the aneurysm neck segment. The parameter $\Delta L_{\max}$ was calculated as

$$\Delta L_{\max} = \left(\frac{p_1}{p_{\min}} - 1 \right) \cdot L_{\text{neck}}, \tag{2}$$

where L_{neck} denotes the longitudinal length of the aneurysm neck segment, p_1 is the helical filament pitch of a FD stent in the neutrally expanded condition, prior to compaction, and $p_{\min}$ is the filament pitch in the fully compacted condition, *i.e.* in a condition where no space exists between any two neighbouring wires and the metal coverage ratio (MCR) therefore equals 100%. By letting MCR = 100%, $p_{\min}$ could be calculated from

$$\text{MCR}\ (\%) = \frac{nd\sqrt{p^2 + \pi^2 D^2}}{\pi p D} \times 100, \tag{3}$$

where n is the number of FD wire filaments; d is wire thickness; and D denotes the nominal diameter of a FD stent. The stent compaction ratio could therefore be computed as

$$\text{CPT} = \frac{\Delta L}{\Delta L_{\max}} \tag{4}$$

We investigated FD stent deployment under three conditions: non-compacted (CPT = 0%), low-compacted (CPT = 20%), and high-compacted (CPT = 70%) conditions.

Results

Virtual Stent Deployment Procedures

Figures 1 and 2 demonstrate the representative deployment procedures for the successful and the unsuccessful case. As can be observed from the stent deployment procedures, FD stents expand progressively from the initial fully crimped condition to the final fully expanded condition, controlled by the boundaries of the aneurysm's parent artery.

FD Stent Structures and Wire Configurations

Figures 3 and 4 depict the FD stent structures and wire configurations after treatments with FD stents of three diameters deployed at three compaction ratios. Both figures show that: (i) the length of a FD stent becomes shorter when the stent is subjected to compaction during deployment; and (ii) the mesh density within the aneurysm ostium becomes higher when a higher compaction is applied during FD stent deployment.

Figure 5 presents the values of stent porosity within the aneurysm ostium measured from treatments using the default deployment technique (CPT = 0%), and following the deployments with a low compaction (CPT = 20%) and a high compaction (CPT = 70%). Regardless of compaction levels, the porosities obtained in the unsuccessfully treated case were higher than

those obtained in the successful one. However the differences attributable to the device's diameter were modest, with a standard deviation (SD) less than 2%. Compared to the default stent deployment, deployment with a low compaction reduced the average porosity from 72.5% to 64.2 % for the successful case, and from 76.3% to 69.3% for the unsuccessful case. In addition, stent deployment with a high compaction further reduced the porosities by 20% (on average), compared to the low compaction conditions in both cases.

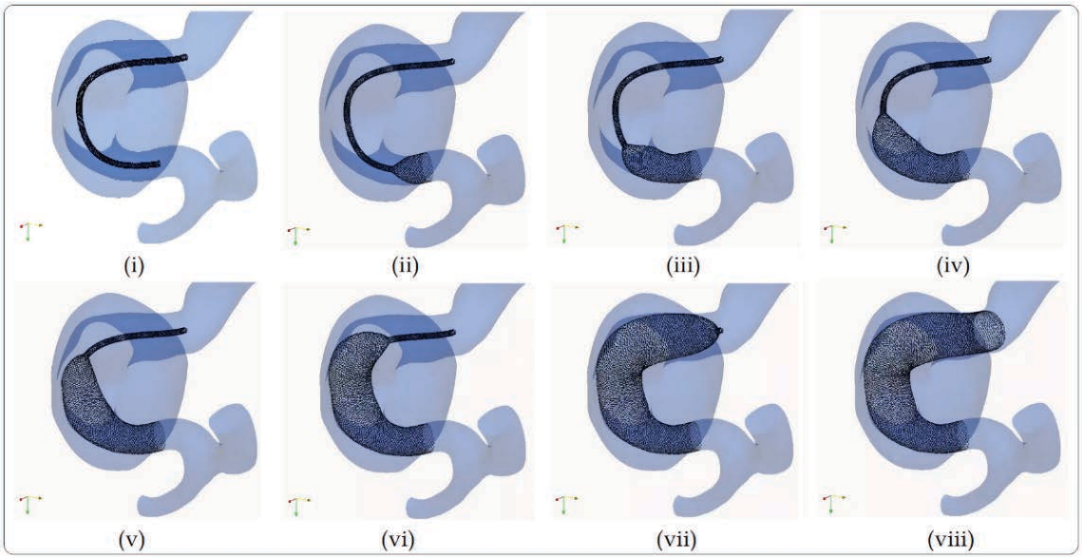


Figure 1. A representative virtual stent deployment procedure for the successfully treated aneurysm. (FD stent size: 4.5 mm, under non-compacted condition; subfigures i to viii: the deployment procedure from stent fully crimped to stent fully expanded.)

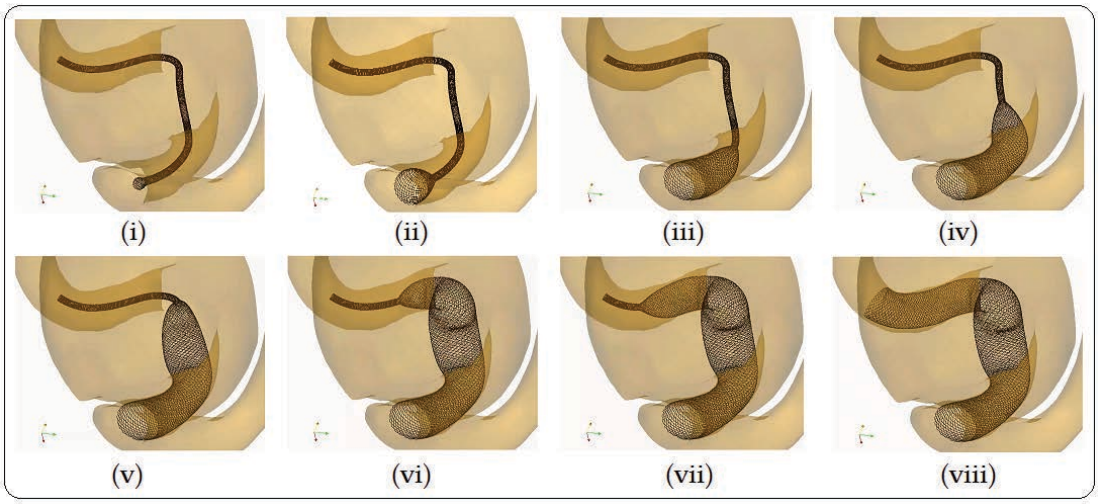


Figure 2. A representative virtual stent deployment procedure for the unsuccessfully treated aneurysm. (FD stent size: 4.5 mm, under non-compacted condition; subfigures i to viii: the deployment procedure from stent fully crimped to stent fully expanded.)

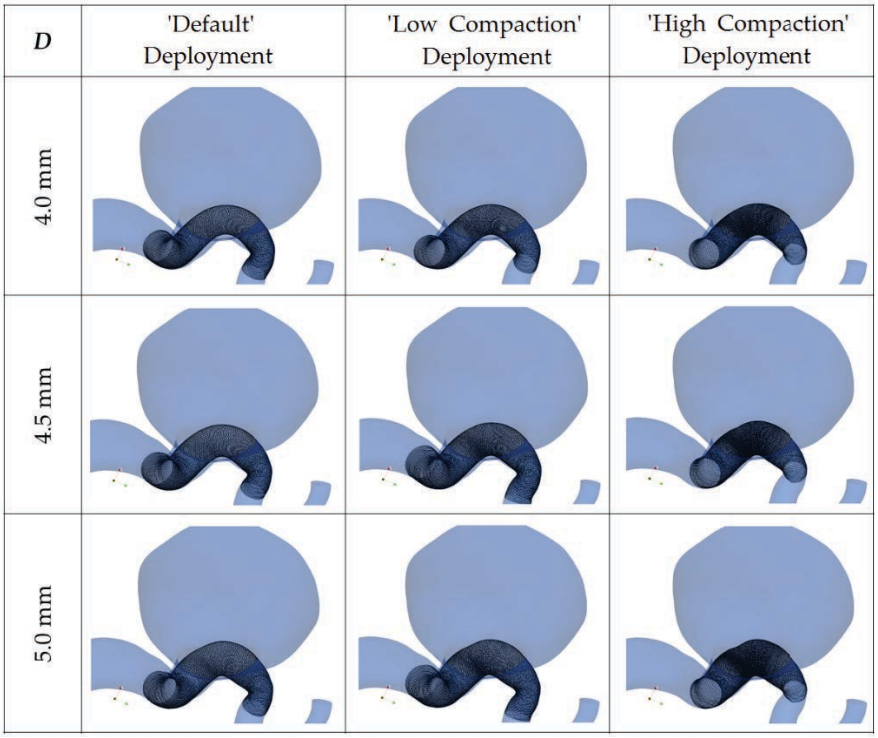


Figure 3. The stent wire configurations of treatments with FD stents of three diameters: 4.0, 4.5, and 5.0 mm, at three compaction levels: 'default' deployment (CPT = 0%), 'low compaction' deployment (CPT = 20%), and 'high compaction' deployment (CPT = 70%) in the successfully treated case.

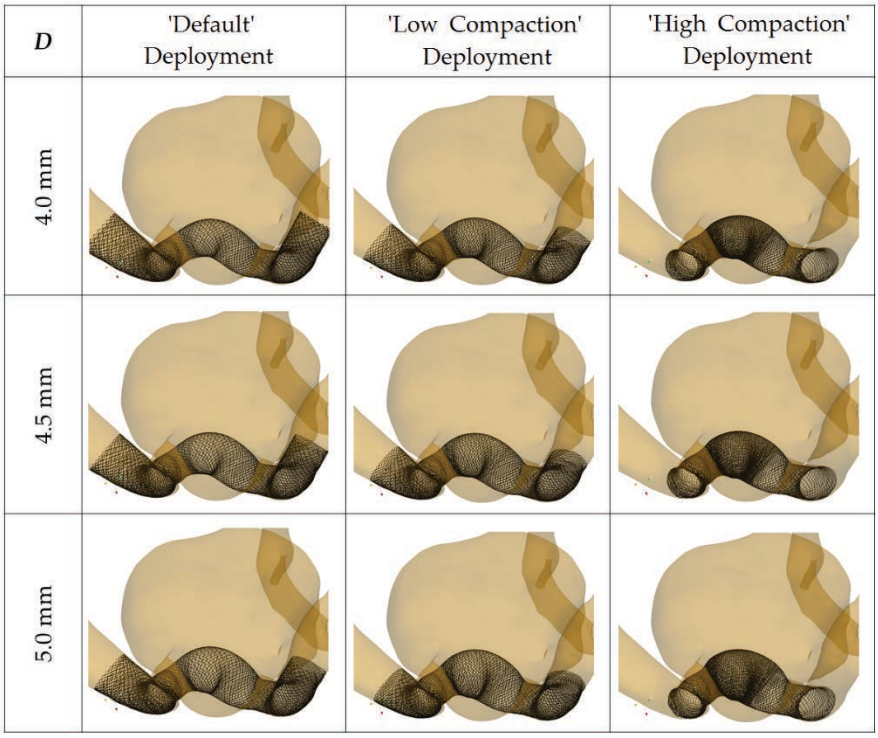


Figure 4. The stent wire configurations of treatments with FD stents of three diameters: 4.0, 4.5, and 5.0 mm, at three compaction levels: 'default' deployment (CPT = 0%), 'low compaction' deployment (CPT = 20%), and 'high compaction' deployment (CPT = 70%) in the unsuccessfully treated case.

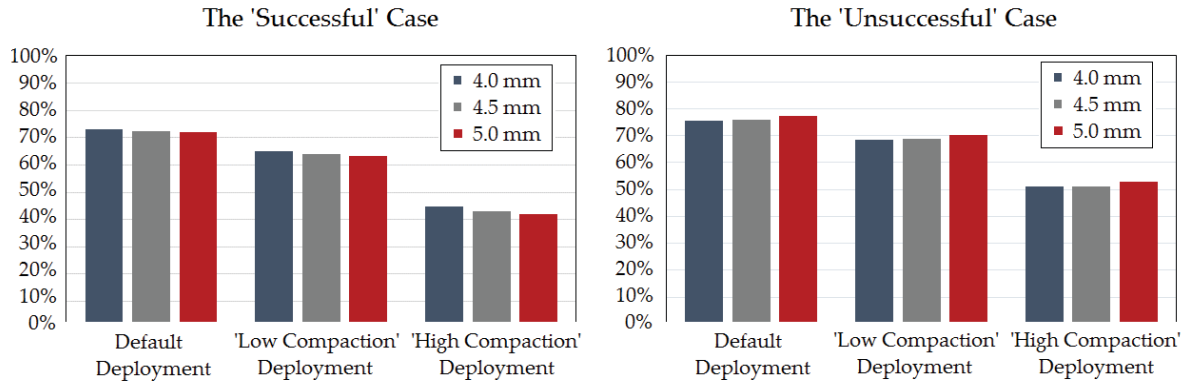


Figure 5. The porosities of the FD stent within the aneurysm ostium measured from treatments with FD stents of three diameters: 4.0, 4.5, and 5.0 mm, at three compaction levels: 'default' deployment (CPT = 0%), 'low compaction' deployment (CPT = 20%), and 'high compaction' deployment (CPT = 70%) in the two cases (successfully treated case at left, unsuccessfully treated case at right).

Discussion

Wire Configurations and Porosities after Deployment with Compaction

We confirmed that stent compaction could decrease the wire porosity across the aneurysm ostium, regardless of the stent size or the recipient aneurysm geometry. The difference between porosities achieved from a low compaction and a high compaction was measured to be 20%. Although the determination of FD stent diameter shows limited effects on stent wire porosity, substantial effects of stent diameter on post-stenting stent shape can be observed.

Different Wire Configurations Obtained in the Successful and the Unsuccessful Case

Observing the post-treatment wire configurations, we found that the stent wires dilate further into the aneurysm ostium in the unsuccessfully treated case. This phenomenon may most likely be due to the morphological characteristics: the unsuccessfully treated case manifests a highly-curved parent artery, compared to the successful one (see Figure 6). Furthermore, the aneurysm in the unsuccessful case was located at the 'apex' of the parent artery curve; treatment with compaction applied may therefore push the FD wires deeper into the aneurysm sac, causing a larger discrepancy between the expected porosity and the obtained one.

In addition, we found a gap between the vascular boundary and the FD stent wires in all treatment scenarios for the unsuccessful aneurysm, whereas no obvious gap was seen in the successful case. This might be an important clue in exploring the reasons why that treatment was ultimately unsuccessful *in vivo*: when blood is flowing through the parent artery, its main stream has a chance to enter the aneurysm sac through the 'gap' (whether or not such gap-induced strong aneurysm inflow eventuates). This point merits further computational fluid dynamics simulations and follow-up DSA scans to gain more specific evidence on such anomalous local haemodynamic behaviour.

If the treating clinicians had have had access to predictive modelling results illustrating the wire configurations and the shape of a virtually deployed FD stent, it is possible that they may have opted to adjust their proposed treatment plan.

Conclusions

Following the definition of compaction ratio developed in this study, we have virtually deployed FD stents of three diameters, with each diameter modelled at three compaction levels, into two patient-specific IAs. Our results indicate that: (i) the porosity differences attributable to device diameter were limited (SD < 2%); and (ii) stent deployment with compaction applied could markedly reduce the wire porosity within the aneurysm ostium (reduction by around 10–30%, depending on the degree of compaction). We observed that the stent wires prolapsed further into the aneurysm sac in the unsuccessful case, most likely due to its geometric characteristics: the unsuccessful case has a highly-curved parent artery.

Treatment rehearsal with a virtual stent deployment simulation would help predict the wire configurations following treatment with a specific device diameter using a given deployment strategy prior to the real treatment, which may be useful in assisting the treating clinicians in determining an optimal treatment plan.

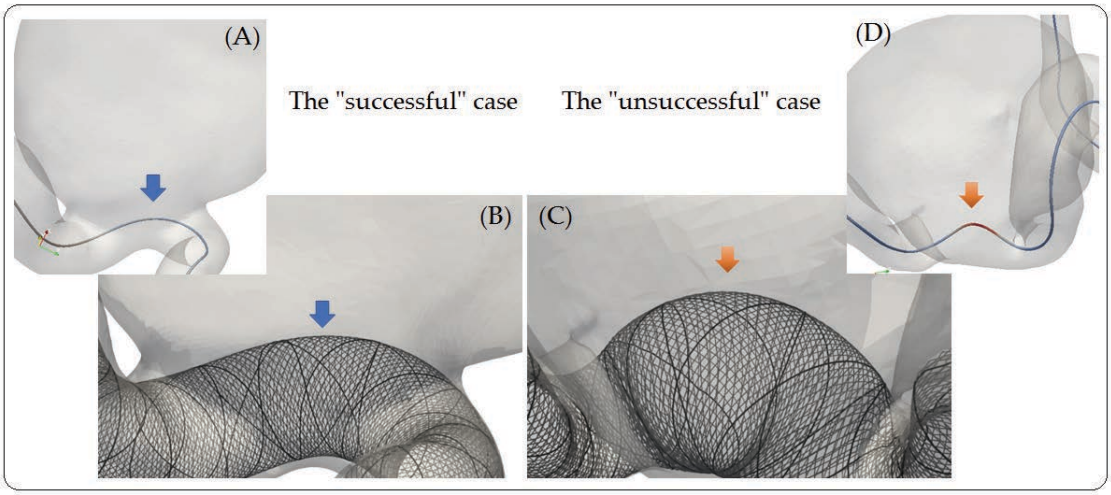


Figure 6. Magnified view of FD wire dilation into the aneurysm lumen. Subfigures A and B show the FD wires after deployment in respectively the successful and the unsuccessful cases; subfigures C and D depict the centrelines extracted from the respective aneurysm geometry for comparison.

Acknowledgements

This study was supported by the ImPACT Program of Council for Science, Technology and Innovation (Cabinet Office, Government of Japan), the Macquarie University Research Excellence Scholarship (iMQRES, No. 2015257), and the Australian Research Council Linkage Projects (Grant ID: LP130100423). The authors would like to thank Dr. Xi ‘George’ Zhao for many fruitful discussions and valuable suggestions regarding the programming of wall-touch detection.

References

[1] Brisman, J.L., Song, J.K., and Newell, D.W. (2006) Cerebral aneurysms. *New England Journal of Medicine* **355** (9), 928–939.

[2] Sforza, D.M., Putman, C.M., and Cebral, J.R. (2009) Hemodynamics of cerebral aneurysms. *Annual Review of Fluid Mechanics* **41** (1), 91–107.

[3] Zanaty, M., Chalouhi, N., Tjoumakaris, S.I., Rosenwasser, R.H., Gonzalez, L.F., and Jabbour, P. (2014) Flow-diversion panacea or poison? *Frontiers in Neurology* 5.

[4] Qureshi, A.I., Janardhan, V., Hanel, R.A., and Lanzino, G. (2007) Comparison of endovascular and surgical treatments for intracranial aneurysms: an evidence-based review. *The Lancet Neurology* **6** (9), 816–825.

[5] Cebral, J.R., Mut, F., Raschi, M., Hodis, S., Ding, Y.-H., Erickson, B.J., et al. (2014) Analysis of hemodynamics and aneurysm occlusion after flow-diverting treatment in rabbit models. *American Journal of Neuroradiology* **35** (8), 1567–1573.

[6] Ma, D., Xiang, J., Choi, H., Dumont, T.M., Natarajan, S.K., Siddiqui, A.H., et al. (2014) Enhanced aneurysmal flow diversion using a dynamic push-pull technique: an experimental and modeling study. *American Journal of Neuroradiology*.

[7] Makoyeva, A., Bing, F., Darsaut, T.E., Salazkin, I., and Raymond, J. (2013) The varying porosity of braided self-expanding stents and flow diverters: an experimental study. *American Journal of Neuroradiology* **34** (3), 596–602.

[8] Gentric, J.-C., Salazkin, I., Gevry, G., Raymond, J., and Darsaut, T. (2015) Compaction of flow diverters improves occlusion of experimental wide-necked aneurysms. *Journal of NeuroInterventional Surgery* neurintsurg-2015-012016.

[9] Lubicz, B., Collignon, L., Raphaeli, G., Pruvo, J.-P., Bruneau, M., De Witte, O., et al. (2010) Flow-diverter stent for the endovascular treatment of intracranial aneurysms: a prospective study in 29 patients with 34 aneurysms. *Stroke* **41** (10), 2247–2253.

[10] Spranger, K. and Ventikos, Y. (2014) Which Spring is the best? comparison of methods for virtual stenting. *IEEE Transactions on Biomedical Engineering* **61** (7), 1998–2010.

[11] Spranger, K., Capelli, C., Bosi, G.M., Schievano, S., and Ventikos, Y. (2015) Comparison and calibration of a real-time virtual stenting algorithm using finite element analysis and genetic algorithms. *Computer Methods in Applied Mechanics and Engineering* **293**, 462–480.

[12] Peach, T.W., Spranger, K., and Ventikos, Y. (2015) Towards predicting patient-specific flow-diverter treatment outcomes for bifurcation aneurysms: from implantation rehearsal to virtual angiograms. *Annals of Biomedical Engineering*.

[13] Zhang, M., Li, Y., Zhao, X., Verrelli, D.I., Chong, W., Ohta, M., et al. (2017) Haemodynamic effects of stent diameter and compaction ratio on flow-diversion treatment of intracranial aneurysms: a numerical study of a successful and an unsuccessful case. *Journal of Biomechanics* **58**, 179–186.

Numerical Predictions of Hydrodynamic Forces and Squat of Ships in Confined Waters

***Y. Liu[1], †L. Zou[1,2], Z. J. Zou[1,2], T.C. Lu[1] and J.X. Liu[3,4]**

[1]School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiao Tong University, Shanghai, China.
[2]Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai, China
[3]School of Navigation, Wuhan University of Technology, Hubei, China
[4]Hubei Key Laboratory of Inland Shipping Technology, Wuhan University of Technology, Wuhan, Hubei, China

*Presenting author: lucy261@sjtu.edu.cn
†Corresponding author: luzou@sjtu.edu.cn

Abstract

Due to the blockage effects in the flow, hydrodynamic performances of a ship in confined waterways are significantly different from those in deep waters. In particular, the hydrodynamic interactions between ship hull and sea bottom or bank wall in the vicinity tend to be more complicated. This gives rise to notable increases in hydrodynamic forces on the hull, along with more pronounced dynamic sinkage and trim where the ship squat phenomenon occurs. The predictions of hydrodynamic forces and ship squat are of great importance from the safe navigation point of view and are also challenging because of the remarkable viscous effects and flow separations in confined waters.

In this paper, an unsteady Reynolds-Averaged Navier Stokes (URANS) solver is applied to simulate the viscous flows around a tanker and a container ship in a confined tank, which is characterized by both shallow sea bottom and close side bank. In each case, the ship is moving along a straight course. The free surface elevation caused by the ship motion is captured by the Volume of Fluid method. In all simulations, the ship position during the motion is updated at each time step according to the computed hydrodynamic forces acting on the hull, from which the dynamic sinkage and trim of the ship is determined. A grid dependency study is performed so as to estimate the numerical error resulted from the grid discretization. The influences of water depth, ship-to-bank distance, ship speed and ship hull form on the hydrodynamic forces and squat of the ship are investigated through systematic computations. Numerical results are evaluated in combination with available experimental data. For the squat, additional data from a mathematical model are used for comparison. The hydrodynamic performances of the ships are generally in good agreement with the data. Furthermore, the mechanisms of the shallow water effects and bank effects involved in the confined waters are analyzed from the simulated flow field around the hulls.

Keywords: Hydrodynamic forces, squat, URANS simulation, shallow water, side bank

Introduction

A ship manoeuvring in confined waters usually experiences much larger hydrodynamic forces than in unrestricted waters due to the hydrodynamic interaction between the ship and the bottom/bank of the waterway. This hydrodynamic interaction has detrimental influence on ship manoeuvrability and may result in marine accidents such as collision or grounding. The ship undergoes dynamic sinkage and trim, notably at very small water depths, due to the hydrodynamic forces acting on the hull. This phenomenon is commonly termed as squat [1]. A decrease of the net Under Keel Clearance (UKC) resulted from squat may affect the ship's manoeuvrability dramatically, giving rise to the loss of ship control. In the latest report of International Towing Tank (ITTC) Manoeuvring Committee [2], it is pointed out that the knowledge of ship motions in confined waters remains a complex and challenging issue, which can be influenced by many different factors such as: free surface elevation, bank blockage, water depth, ship speed, etc. Therefore, to ensure a safe navigation, it is of great

importance to accurately predict the hydrodynamic forces acting on the ship manoeuvring in confined waters by taking the so-called shallow-water and bank effects into account.

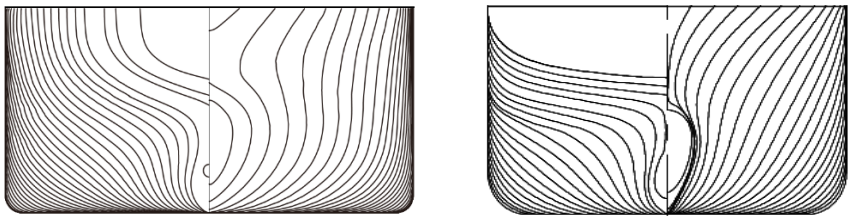
Ship manoeuvring in confined waters has been studied in many ways for a long time. In general, most of the investigations rely on Experimental Fluid Dynamics (EFD) or model tests, theoretical and semi-theoretical methods [3, 4]. In Kazerooni [5], the squat of several ship models were measured and a regression formula was established from model tests results. The squat measurements in model tests of ships moving into and out of a lock chamber were conducted by Xu et al. [6]. Dand and Ferguson [7] established a semi-empirical method (D&F method) for the calculation of ship squat. Based on this theory, Latarie et al. [8] proposed a new mathematical model to predict the squat in rectangular cross sections. Gourlay et al. [9] calculated the dynamic sinkage and trim of modern container ships in confined waters by using slender-body method and Rankine-source method. Most recently, Mucha et al. [10] carried out model tests in confined waters, which was contributed as benchmark tests to the PreSquat-Workshop on Numerical Prediction of Ship Squat in Restricted Waters [11], aimed to benchmark the capabilities of available numerical methods for squat prediction and to increase the safety of manoeuvring ships.

Nowadays, with the rapid development of computer technique and Computational Fluid Dynamics (CFD) method, CFD-based numerical prediction of the hydrodynamic forces has become possible [12]. Especially, the viscous method, typically the Reynolds Averaged Navier-Stokes (RANS), is shown to be able to produce promising and comprehensive predictions of ship manoeuvrability. Many numerical investigations on the squat of ships advancing in confined waterways have been conducted, such as Toxopeus et al. [13], Kaidi et al. [14], Tezdogan et al. [15] and Linde et al. [16]. Moreover, Zou and Larsson [17] utilized a RANS solver to investigate the bank effects on a tanker hull in two canals. A similar simulation for a bulk carrier passing a lock was conducted by Wang and Zou [18]. Liu et al. [19] assessed the ship manoeuvring stability in Planar Motion Mechanism (PMM) tests taking the ship-bank distance into account. Wang et al. [20] numerically simulated the berthing manoeuvre of a ship in the prescribed translational motion.

In this paper, a numerical investigation is performed aiming to find an efficient numerical method for the ship squat and hydrodynamic forces predictions in confined waters combing the comparisons with available experimental data.

Ship geometry and computational cases

Two benchmark ship models are considered in the present study, namely, KVLCC2 tanker and KCS container ship. Fig. 1 presents their line plans without the scale drawn. Geometries are available at the website of SIMMAN2014 [21]. Main particulars of the two ships are given in Table 1.



a) KVLCC2 b) KCS

Figure 1. Line plans of ships

Table 1. Main particulars of the three ships

Particulars	KVLCC2	KCS
Length between perpendiculars L_{pp} (m)	320.0	230.0
Breadth B (m)	58.0	32.2
Design Draft T (m)	20.8	10.8
Block coefficient C_B (-)	0.8098	0.651
Displacement ∇ (m^3)	312,622	52,022
Moment of Inertia K_{xx}/B	0.4	0.4
Moment of Inertia K_{yy}/L_{pp}, K_{zz}/L_{pp}	0.25	0.25
Wet surface area S (m^2)	27194	9424
Vertical Center of Gravity (from keel) KG(m)	18.6	7.28
LCB (%L_{pp}), fwd+	3.48	-1.48

In the numerical computations, the same test conditions are used as in the experiments to ensure direct comparisons. Benchmark tests of KVLCC2 hull in shallow water in model scale (scale factor λ=75) are obtained from Flanders Hydraulics Research (FHR) [22]. The ship model was towed at a constant forward speed U along different lateral positions y_{wall} in a rectangular tank (See Fig. 2). The width W of the tank section is varied from 2.5B up to the entire width of the towing tank (9.05B, 7m). Also, a range of water depths varied from 1.1T to 3T was tested. The blockage factor m_b=(B×T)/(h×W) is defined as the ratio of the ship's underwater cross-sectional area at mid-ship to the cross-sectional area of the waterway. The water depth h, tank width W and the forward speed U are summarized in Table 2.

The computations of KCS model (λ=40) are conducted for h=1.3T, 1.6T at 0.73m/s forward speed and for 1.2T water depth at 0.82m/s respectively. These conditions are identical to the experiments carried out at the Development Center for Ship Technology and Transport Systems (DST) in Duisburg, Germany [23, 24]. The bank effect is neglected for this ship model. It will be mainly used for comparing the ship resistance and squat predictions with KVLCC2 model at the same Froude depth number Fr_h (=$U/\sqrt{gh}$).

Table 2. Overview of the test conditions for KVLCC2 model

U (full scale)	U	W	h	m_b	Fr_h	y	y_{wall}
Knot	m/s	-	-	-	-	m	m
7	0.416	*9.05B*	*1.2T*	0.092	0.230	0	0
			1.5T	0.074	0.203		
			3T	0.037	0.146		
15.5	0.921	*9.05B*	*1.2T*	0.092	0.510	0	0
			1.5T	0.074	0.456		
			3T	0.037	0.323		
8	0.475	5B	1.1T	0.182	0.275	0	0
			1.35T	0.148	0.248		
			1.5T	0.133	0.235		
8	0.475	5B	1.5T	0.133	0.235	1.526	0.02
						1.352	B/4
						0.773	B

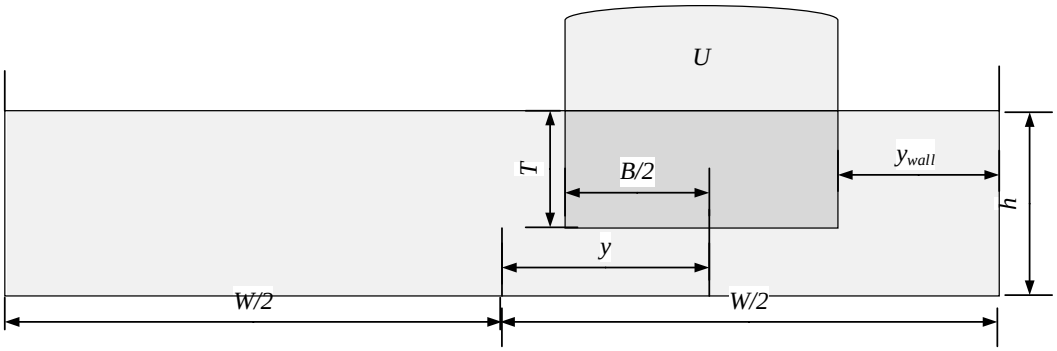


Figure 2. The variables in a cross section of the tank

Numerical Models

Governing equations

The viscous flow involved in the manoeuvring tests can be treated as incompressible, and is governed by the Navier-Stokes (N-S) equations. After averaging the N-S equations over time, the unsteady Reynolds Averaged Navier-Stokes (RANS) equations can be written as follows:

$$\begin{cases} \dfrac{\partial u_i}{\partial x_i} = 0 \\ \rho\dfrac{\partial u_i}{\partial t} + \rho u_j \dfrac{\partial u_i}{\partial x_j} = -\dfrac{\partial p}{\partial x_i} + \dfrac{\partial}{\partial x_j}\left(\mu\dfrac{\partial u_i}{\partial x_j} - \rho\overline{u_i' u_j'}\right) \end{cases} \quad (i, j = 1, 2, 3) \tag{1}$$

where $x_{i,j}$ (i, j=1, 2, 3, $i \neq j$) is the ith or jth component of the fixed coordinate system, ρ is the density of fluid, $u_{i,j}$ (i, j=1, 2, 3, $i \neq j$) is ith or jth mean velocity component, p is the mean pressure, μ is the viscosity and $-\rho\overline{u_i' u_j'}$ is the Reynolds stress which needs to be solved through turbulence modelling.

Numerical Methods

To simulate the viscous flow, the CFD software STAR-CCM+ [25] is applied in the present study to solve the RANS equations. A Finite Volume Method (FVM) is used to discretize the flow domain into a finite number of control volumes (CVs). The temporal discretization is based on a first-order Euler difference, and the spatial discretization is performed with second-order upwind scheme for the convection term and secondary gradient contribution for the diffusion term.

The air-water interface is captured using the Volume of Fluid (VOF) method [26]. VOF assumes a common velocity and pressure field for both phases within a single CV, and monitors the phase fraction. The governing equations for mass and momentum continuity in a single-phase flow are thus solved for an equivalent fluid, whose physical properties (density and laminar viscosity) are functions of the constituent phase's properties and volume fractions. The transport of volume fraction is described by an additional conservation equation:

$$\frac{\partial \alpha}{\partial t} + \frac{\partial (u_i \alpha)}{\partial x_i} = 0 \tag{2}$$

where α represents the volume fraction, indicating the relative proportion of fluid in each cell; its value is always between 0 and 1. The High Resolution Interface Capturing (HRIC) convection discretization scheme [27] is used to improve the VOF interface tracking capabilities.

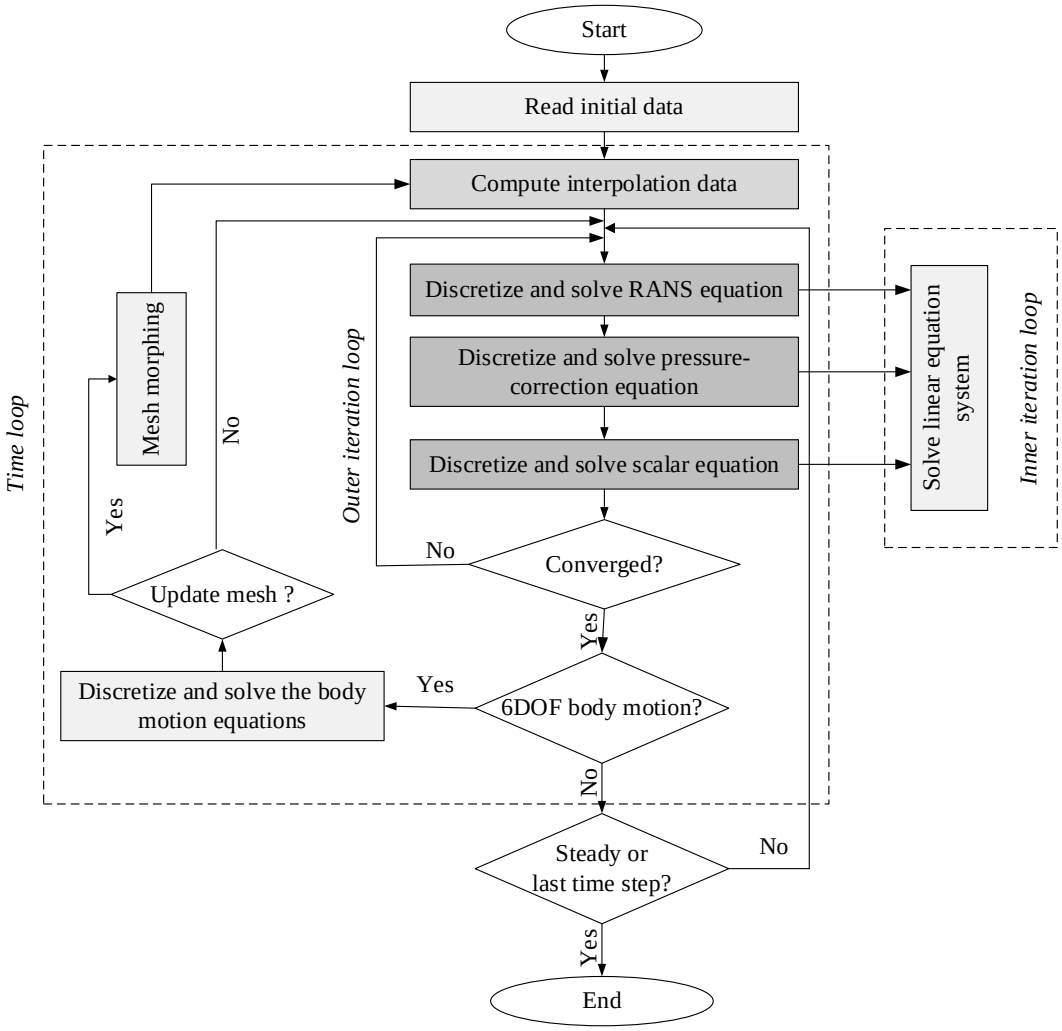


Figure 3. Main flow chart in the STAR-CCM+ solver

Mean flow quantities near the solid wall are simulated according to an all y^+ wall treatment where blended wall function is adopted. This approach is flexible as it's capable of handling a range of local grid refinement levels near the wall. If the grid is fine enough ($y^+<1$), the viscous sublayer is resolved and and thus the wall shear stress is computed as it would be in a laminar flow. If the grid is coarse ($y^+> 30$), the wall law is equivalent to a logarithmic profile. The SIMPLE algorithm [28] is employed for pressure-velocity coupling. Moreover, the RANS equations are closed with *k*-*ε* turbulence model [29]. Particularly, the sinkage and trim motion are specified with the "Dynamic Fluid body Interaction (DFBI) morphing" module. It involves actual displacement of the grid vertices, and can use control points and their associated displacements to generate an interpolation field throughout the region, which can model the 6-DOF motion of a rigid body within the fluid system. Then the resultant force and

moment acting on the body due to all influences are calculated, and the flow field is updated to find the new position and orientation of the ship. More details about the DFBI formulation can be found in Ohmori [30]. For resistance computations in calm water, the time step is determined by 0.05-0.01L_{pp}/U in accordance with the related guidelines of ITTC[31]. A much smaller time step (<0.002L_{pp}/U) is used in the present work since the flow in the confined water condition is more unstable. A general flow chart of this solver is indicated in Fig. 3.

Computational setup and grid generation

The computational domain and general grid distributions for the case of KVLCC2 model at h/T=1.5 and W=5B are given in Fig. 4. The bottom position of the domain is determined according to the corresponding water depth, and the tank width is related to the test conditions in Table 2. The computational domain in the numerical tests is made by eight boundaries: inlet plane, outlet plane, hull surface, top plane, tank bottom, as well as two side walls representing the bank of the tank. It should be noted that in the W=9.05B case, only a half of the domain is modeled since this case represents the hull located in the center of the tank, indicating a symmetry flow configuration. The fluid domain extends 1.5L_{pp} from the bow to the inlet plane, 3.5L_{pp} from the aft-perpendicular to the outlet plane and 0.33 L_{pp} from the free surface to the top plane. The water depth and tank width in the domain vary in the following systematic computations.

As to the adopted boundary conditions in the computations, the velocity inlet condition is set on the inlet and top plane where the boundary pressure is extrapolated by using reconstruction gradients. The pressure outlet condition is used on outlet plane where the boundary pressure is governed by a field function which monitors the instantaneous pressure on, above and below the free surface at the boundary. The symmetry condition is set on the symmetry plane in the half domain case where the shear stress is zero. A no slip condition is satisfied on the hull and slip wall condition on the side walls. It should be noted that in shallow water cases, the effect of the boundary layer on the tank bottom greatly influences the flow in the gap between the ship and bottom, so a moving no-slip condition is used on the bottom. The two side bank walls are also set as moving no-slip wall in the narrow bank case (W<9.05B). The pressure resistance fluctuation is usually found to be caused by the wave reflection at the non-physical side boundaries. Therefore, a numerical damping method with a damping length of 10m is applied on the inlet and outlet boundaries to remove the fluctuation. Furthermore, the release and ramp time are up to 20s to allow enough time for the fluid flow to be stable.

Fig. 5 provides a closer look at the grid cells around the ship hull. As shown in Fig. 5, the grid is refined near the free surface, tank bottom, hull surface, bank sides and in the wake region to ensure that the complex flow features are appropriately captured. The grid refinements in these zones are achieved using volumetric controls. Orthogonal prismatic cells are generated next to the hull, tank bottom surfaces and side walls to improve the accuracy of the flow resolution. The prism layer is not used on the two side surfaces in the wide tank condition (W=9.05B), where the blockage effect is minor. To avoid numerical difficulty related to the DFBI morphing approach as much as possible for shallow water cases, y^+ is larger than 30 so as to use wall function on the tank bottom and the close side tank wall, while y^+ is smaller than 1 at the hull surface to get a more precise flow field simulation near the ship.

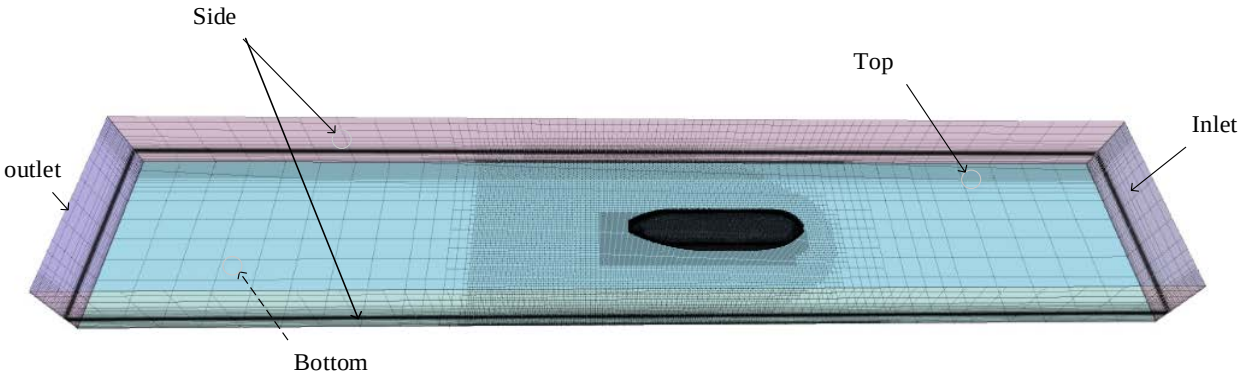
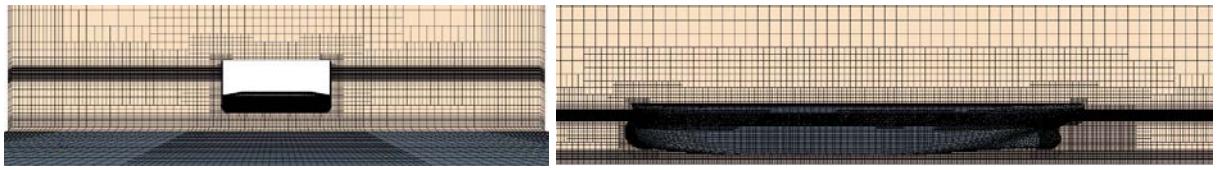


Figure 4. Overview of computational domain



(a) (b)

Figure 5. Grid structure around ship and bottom in shallow water
(a) cross section at mid-ship (b) longitudinal section at *y*=0

Results and Discussion

Grid dependency study for KVLCC2

The purpose of grid dependency study is to estimate the numerical error and uncertainty resulted from the grid discretization. In this paper, this study is conducted following the Grid Convergence Index(GCI) method [32]. It is applicable for unstructured grid and only requires the grid refinement to be done systematically. Therefore, all grid quantities are given as percentages in terms of a base size, in order to refine the grid in a more systematic way. The case of KVLCC2 at h=1.2T, 9.05B bank width and the 0.921m/s forward speedis chosen for the study. Only half of the computational domain is used to reduce the calculation cost. Three grid sets (coarse, medium and fine) are adopted in the study and the grid refinement is achieved by applying a refinement factor r_G=√2 to the base size. The fine grid (No.1) consists of approximately 2.9M cells; the medium grid (No.2) contains about 1.3M; and about 0.69M in the coarse grid (No.3). The changes in solutions between two successive grids are defined as: $\varepsilon_{32} = \phi_3 - \phi_2$, $\varepsilon_{21} = \phi_2 - \phi_1$. The apparent order p of the method is expressed by:

$$p = \frac{1}{\ln(r_G)} \left| \ln \left| \varepsilon_{32} / \varepsilon_{21} \right| \right| \tag{3}$$

The extrapolated values ϕ_{ext}^{21} can be calculated by:

$$\phi_{ext}^{21} = (r_G \phi_1 - \phi_2)/(r_G^P - 1) \tag{4}$$

The approximate relative error between medium-fine solutions e_a^{21} and extrapolated relative error e_{ext}^{21} can be computed as follows:

$$e_a^{21} = \left| \frac{\phi_1 - \phi_2}{\phi_1} \right| \tag{5}$$

$$e_{ext}^{21} = \left| \frac{\phi_{ext}^{21} - \phi_1}{\phi_{ext}^{21}} \right| \tag{6}$$

The fine-grid convergence index GCI_{fine}^{21} is calculated by:

$$GCI_{fine}^{21} = \frac{1.25 e_a^{21}}{r_G^P - 1} \tag{8}$$

The total resistance (R_t) of a ship is mainly composite of two components: the pressure resistance (R_p) and the frictional resistance (R_f). The computed values of R_p, R_f, R_t and the ship sinkage σ of the coarse, medium and fine grids are given in Table 3. The e_a^{21} results show that all the resistances have small approximate relative errors. Moreover, the values of GCI in all coefficients are no more than 5%. From the grid dependency study, the observed errors in all grid sets are not large, thus a medium grid density is chosen to maintain an affordable computation cost.

Table 3. Grid convergence parameters

	Pressure resistance R_p (N)	Frictional resistance R_f (N)	Total Resistance R_t(N)	Sinkage σ (mm)
ϕ_3	14.23	7.76	21.99	24.59
ϕ_2	14.44	8.18	22.62	24.11
ϕ_1	14.81	8.20	23.02	23.76
p	1.72	9.05	1.37	1.00
ϕ_{ext}^{21}	15.27	8.20	23.66	22.94
e_a^{21}%	2.53	0.23	1.71	1.43
e_{ext}^{21}%	3.01	0.01	2.73	3.58
GCI_{fine}^{21}%	3.89	0.01	3.51	4.32

For more clear comparison, the resistance R_p, R_f and R_t and sinkage σ are expressed in non-dimensional form in the next section. The resistance coefficients C_p, C_f and C_t are achieved by dividing each term by $0.5\rho U^2 S$ and the sinkage σ' by L_{pp}, respectively.

KVLCC2

1. W=9.05B-centred ship

Firstly, the predicted total resistance and sinkage of KVLCC2 model in a wide tank (W=9.05B) at three different water depths: 1.1T, 1.35T and 1.5T are compared to the experimental data [22] and the results computed from D&F mathematical formula, which is a semi-empirical method based on the Bernoulli Equation[7]. As shown in Fig. 6, the overall computed resistances at different water depths agree well with the measurements and the sinkage also indicates satisfactory agreement.

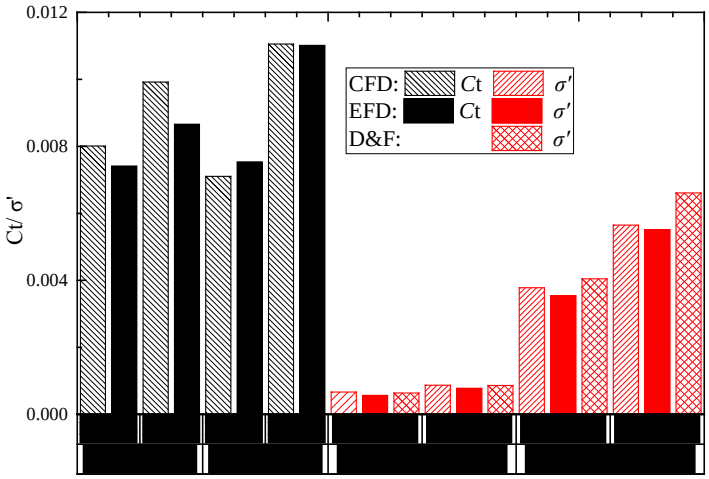


Figure 6. Comparison with experimental data [22] and D&F mathematical model

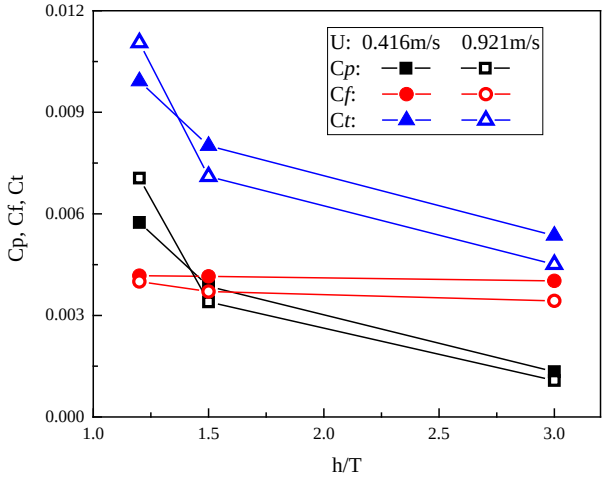


Figure 7. Resistance coefficients at different water depths and forward speeds

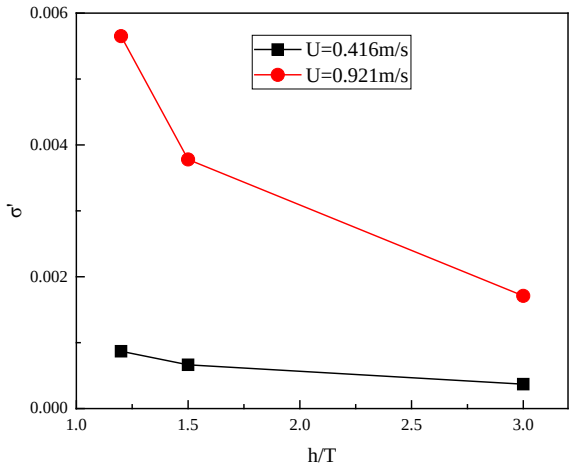


Figure 8. Sinkage coefficients at different water depths and forward speeds

The numerical results of resistance and sinkage with the same tank width but different Fr_h are shown in Fig. 7-8. It can be seen that all the resistance coefficients-C_p, C_f and C_t increase as the water depth decreases at the same forward speed, among which the increase in C_f is less. The sinkage coefficient increases rapidly as Fr_h becomes larger.

2. *W*=5*B*-centred ship

The results of resistance and sinkage coefficients of KVLCC2 with a narrow tank width (*W*=5*B*) are given in Fig. 9 and Fig. 10, respectively. The results of sinkage are also compared with the experimental data from Lataire et al. [8] and the computed results from D&F mathematical formula [7]. The computed sinkage coefficients qualitatively follow the same trend as the measurements with slight under-predictions. Results from the D&F method also indicate a similar trend, but there are large deviations from either computation or experiment. The difference between computed and measured sinkage is due to the fact that the sinkage values are small, which are difficult to capture precisely. It might be also caused by the propulsion effects. In the experiments, the ship model is tested with a rotating propeller, while in the numerical simulations no appendage is equipped.

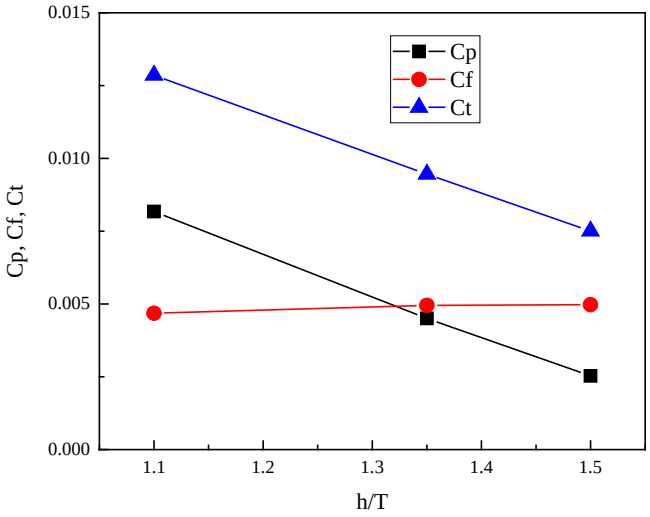


Figure 9. Resistance coefficients in the 5*B* tank width case

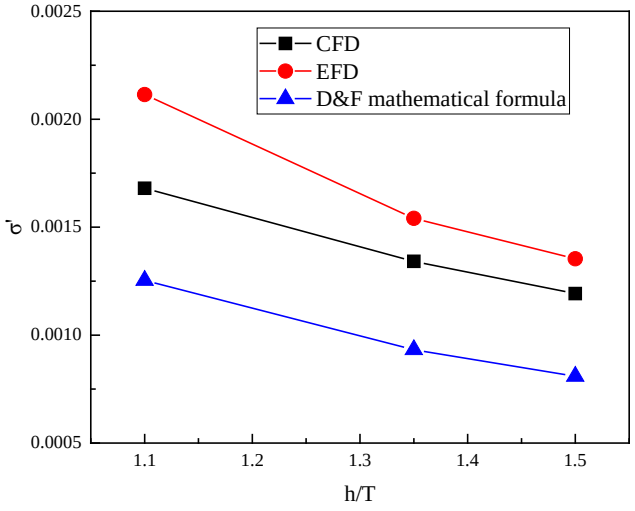


Figure 10. Sinkage coefficient in the 5*B* tank width case

3. *W*=5*B*-Offset ship longitudinal central position

To investigate the bank effects, three configurations with different ship-bank distances are simulated. The distance y_{wall} (see Fig. 3 and Table 2) is defined as:

$$y_{wall} = \frac{W}{2} - \frac{B}{2} - y \tag{9}$$

The lateral position y specifies the distance between the centerlines of the ship and the tank.

The resistance coefficients are given in Fig. 11. In this investigation, two more influential quantities are examined, namely, sway force Y and yaw moment N (see Fig. 11). The two hydrodynamic forces are nondimensionlized by $0.5\rho U^2 L_{pp} T$ and $0.5\rho U^2 L_{pp}^2 T$, respectively. The results show that hydrodynamic forces are affected by the y_{wall}. These forces are larger when the ship is getting closer to the bank. The sinkage coefficients are shown in Fig. 12 and compared with the mathematical model from Lataire et al. [8]. The sinkage predicted by CFD method is lower than that the experimental data in [8], since the mathematical model considers the propeller effects.

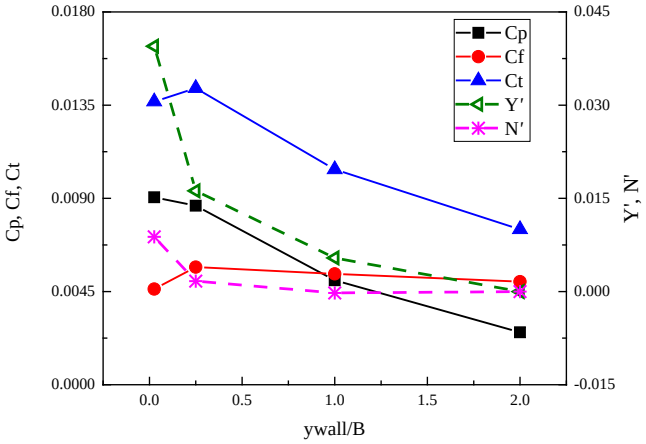


Figure 11. Resistance, sway force, yaw moment coefficients vs. y_{wall}/B

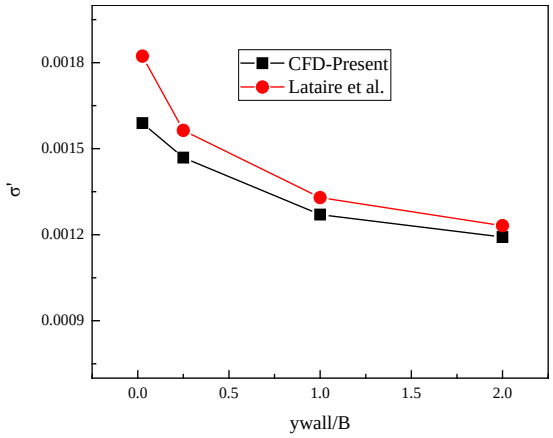
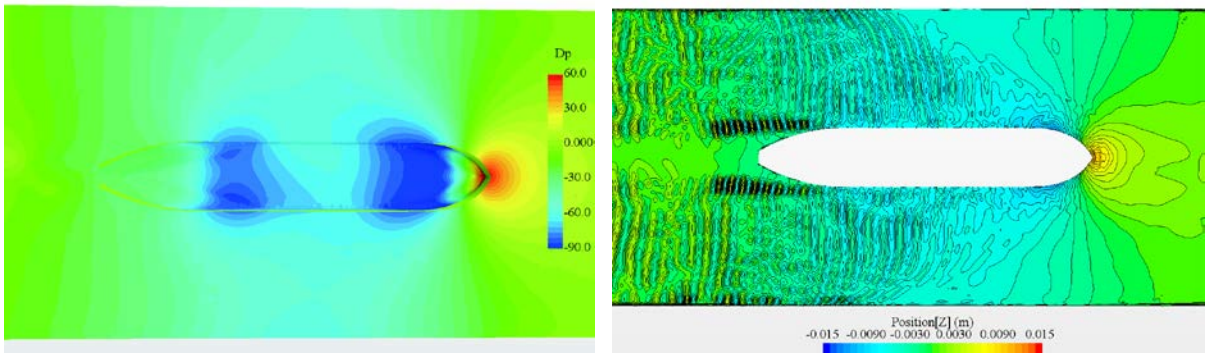


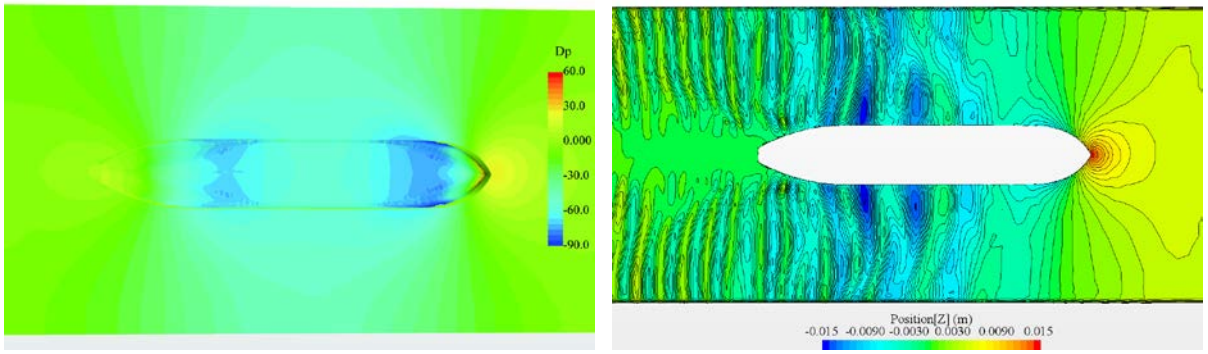
Figure 12. Sinkage coefficients vs. y_{wall}/B

To gain a deeper insight into the physical mechanism in the hydrodynamic forces on the ship, pressure distributions on the hull and the tank bottom surface, along with the free surface elevation are discussed in Figs. 13-14. As shown in Fig. 13, there are significant differences between the pressure distributions at different water depths. For the shallowest case, a distribution of significant suction pressures can be observed on the aft and fore part of the hull. The Kelvin wave pattern can be seen for the three cases.

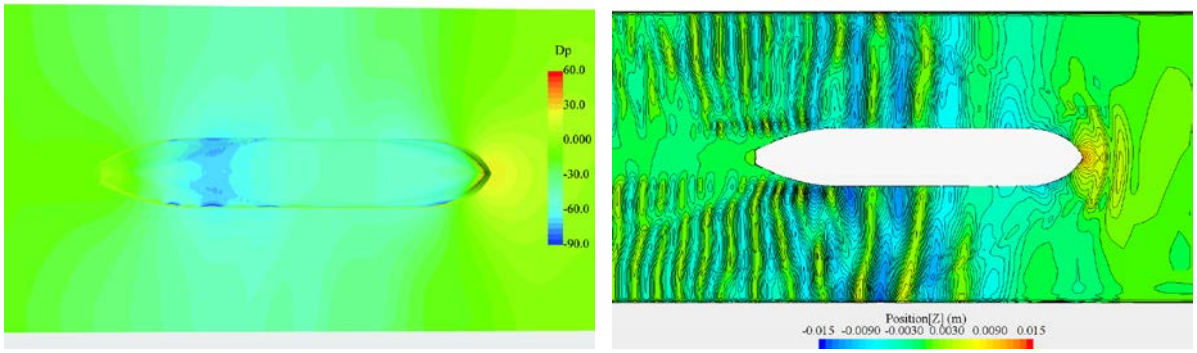
The bank effects are clearly indicated in Fig. 14. The hull surface facing the near bank side has a larger suction pressure and it is clearly noticed for the case with y_{wall}=0.02. The pressure increases as positive at the bow which pushes the ship to the tank centre.



a) h/T=1.1

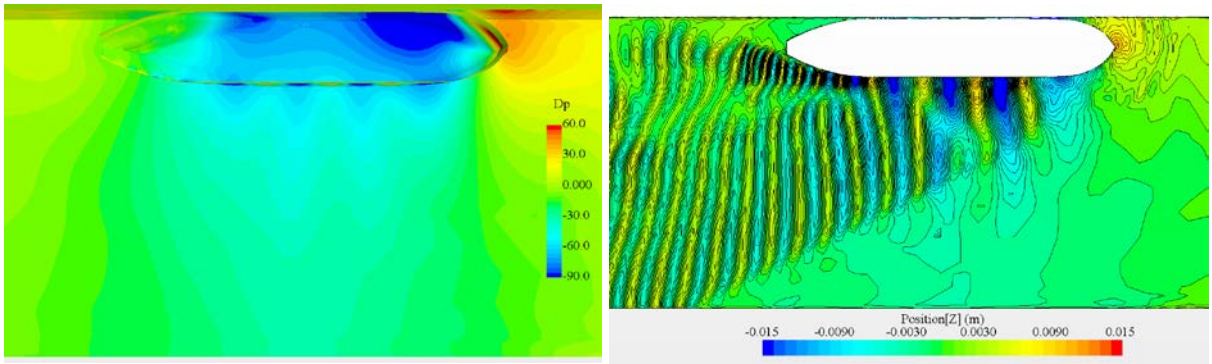


b) h/T=1.35

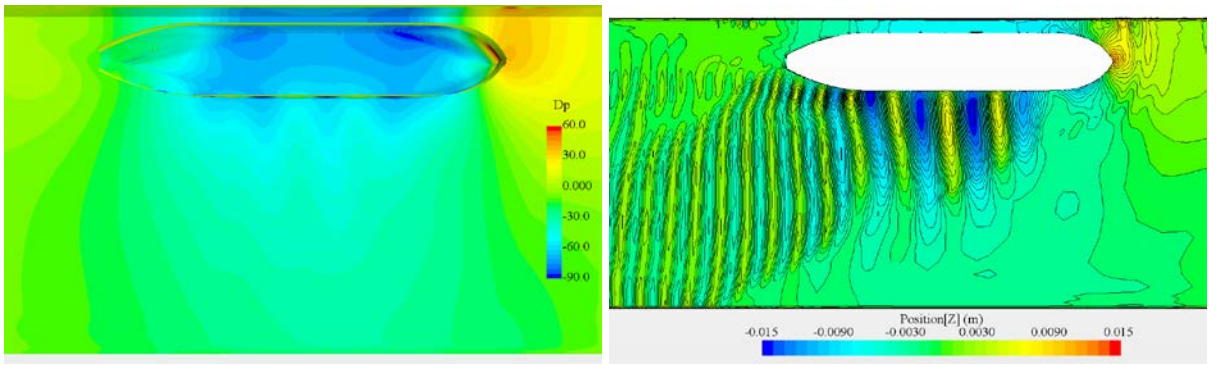


c) h/T=1.5

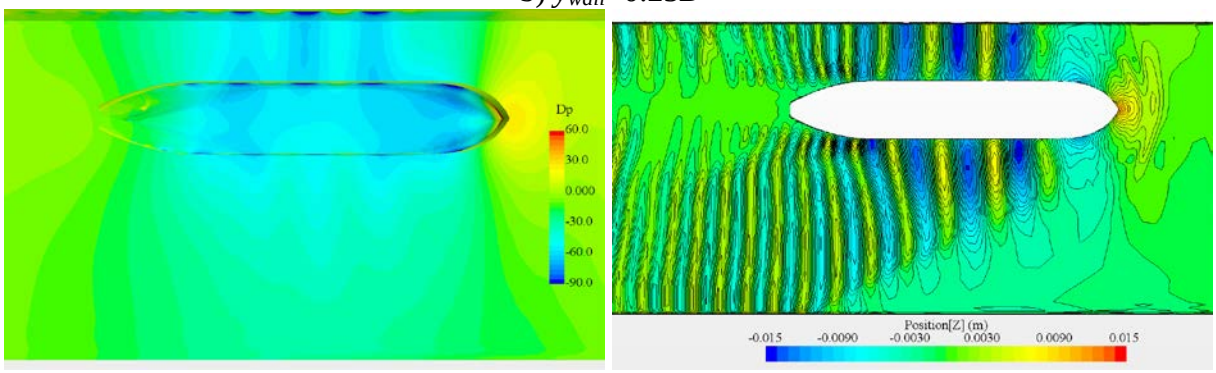
Figure 13. Dynamic pressures and wave elevation at different water depths (Left: dynamic pressure; right: wave elevation)



a) y_{wall}=0.02B



b) y_{wall}=0.25B



c) y_{wall}=B

Figure 14. Dynamic pressure and wave elevation under different y_{wall} (Left: dynamic pressure; right: wave elevation)

KCS

Similarly, the predicted total resistance and sinkage coefficients of KCS model using the RANS method are compared to the experimental data [24]. As shown in Fig. 15, the overall agreement between EFD and CFD is good, except for the C_t at h/T=1.2.

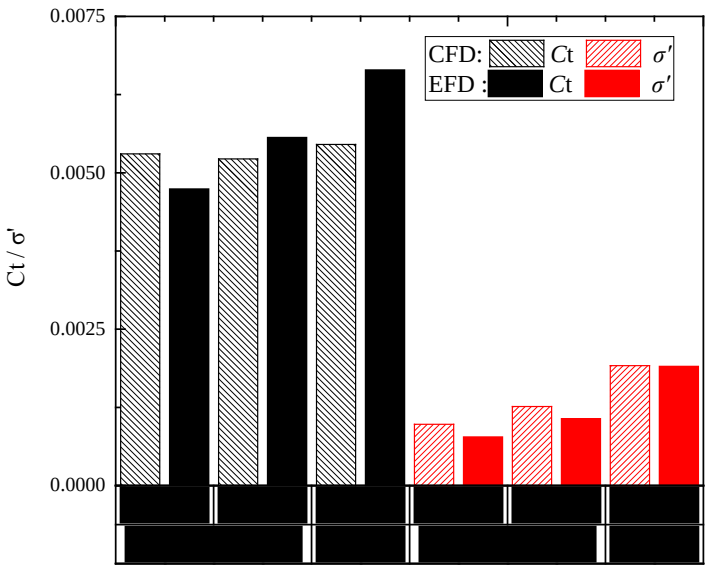


Figure 15. Comparison with experimental data [24]

Then the frictional, pressure, total resistance coefficients and the sinkage of the KCS model are compared with those of KVLCC2 model at the same Fr_h. That is to say, the motion and

the resistance of KCS model at three different water depth ratio *h/T*: 1.2, 1.5 and 3.0 at the forward speed 0.874m/s in the range of Froude depth numbers Fr_h=0.51, 0.456, 0.323 are simulated corresponding to the KVLCC2 model at the 0.921m/s forward speed at the same three water depth ratios. The comparisons are shown in Fig. 16-17. It can be seen that all the resistance coefficients and the sinkage coefficients of KVLCC2 are larger than those of KCS at the same Fr_h. Given that in model tests the KVLCC2 is shorter but has larger C_B than KCS, it seems that the blunt hulls will be more affected by the shallow water effects.

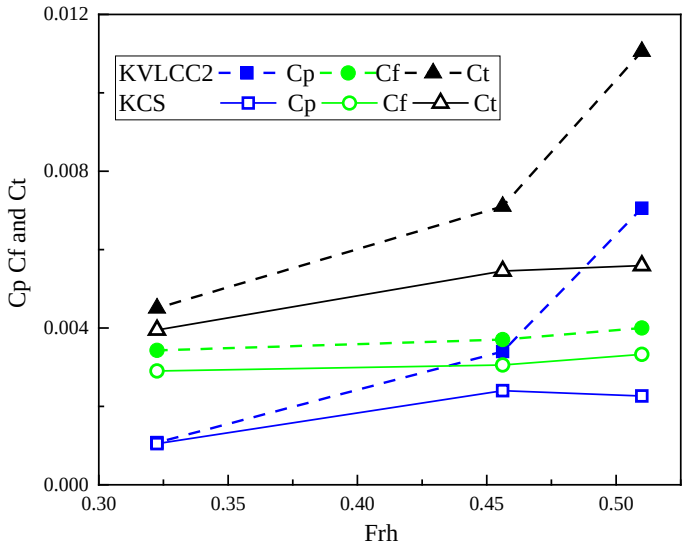


Figure 16. Resistance coefficients comparison between KCS and KVLCC2

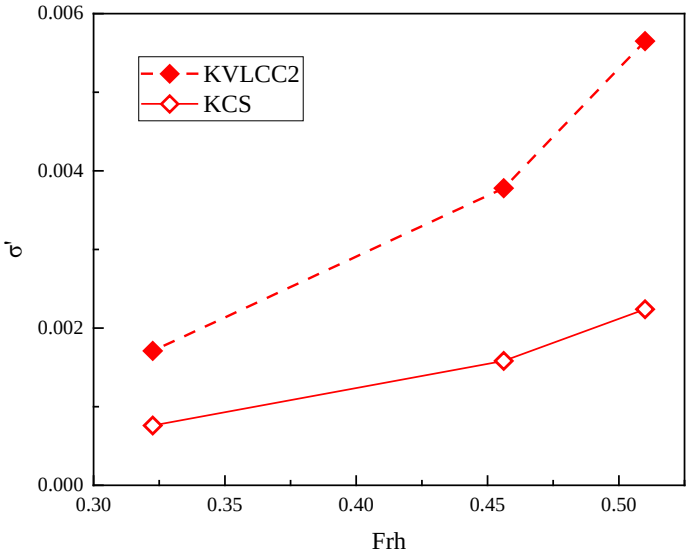


Figure 17. Sinkage coefficients comparison between KCS and KVLCC2

Conclusions

In this paper, a numerical method based on RANS equation is proposed to study the influence of tank width and water depth in terms of the shallow-water and bank effects. The sinkage and resistance results of the three benchmark ships obtained from the CFD simulations are presented in a range of ship speeds at different water depths. The numerical results of squat were also compared to those from available experiments. The general agreement between the resistance predictions by experimental measurements and numerical simulations is satisfactory. For KVLCC2, the pressure resistance is dominant in the wide tank width case.

However, the resistance coefficients of KCS are all smaller than those of KVLCC2 at the same Fr_h. The agreement in the squat predictions for both hulls is also promising. The mid-ship sinkage is dominated by the local pressure along the parallel middle body which can be captured well by the present CFD method. Therefore, a better understanding of the hydrodynamic performances for ships travelling in confined waters is obtained.

Acknowledgments

The work described in this paper was supported by grants from the National Natural Science Foundation of China (NSFC) (No.: 51479156 and 51309152). And the authors would like to thank the reviewers for their constructive and valuable comments which lead to the better presentation of this paper.

References

[1] Briggs, M. J., Vantorre, M., Uliczka, K. and Debaillon, P. (2010) Prediction of squat for underkeel clearance, *Handbook of Coastal and Ocean Engineering World Scientific Publishing Company*, 723-774.

[2] International Towing Tank Conference Maneuvering Committee.(2014) *Final report and recommendations to the 27th ITTC, 27th International Towing Tank Conference(ITTC)*, Copenhagen, Denmark.

[3] Ankudinov, V., Daggett, L., Huval, C. and Hewlett, C.(1996) *Squat predictions for manoeuvring applications, the 7th International Conference on Marine Simulation and Manoeuvrability (MARSIM)*, Copenhagen, Denmark.

[4] Briggs, M. J. (2006). Ship squat predictions for ship/tow simulator.

[5] Kazerooni, M. F. and Seif, M. S. (2014) Experimental evaluation of ship squat in shallow waters, *Journal of the Brazilian Society of Mechanical Sciences and Engineering* **36**, 559-569.

[6] Xu, J. C., Xuan, G. X., Li, Y., Li, Z. H., Hu, Y., Jin, Y. and Huang, Y. (2016) Study on the squat of extra-large scale ship in the Three Gorges ship lock, *Ocean Engineering* **123**, 65-74.

[7] Dand, I. and Ferguson, A. (1973) Estimatinq the bow and stern sinkage of a ship underway in shallow water, *The Naval Architect*.238.

[8] Lataire, E., Vantorre, M. and Delefortrie, G. (2012) A prediction method for squat in restricted and unrestricted rectangular fairways, *Ocean Engineering* **55**, 71-80.

[9] Gourlay, T. P., Ha, J. H., Mucha, P. and Uliczka, K.(2015) *Sinkage and trim of modern container ships in shallow water, the 22nd Australasian Coastal and Ocean Engineering Conference*, Auckland, New Zealand.

[10] Mucha, P., Moctar, O. E. and Böttner, C. U. (2014) Technical note: PreSquat–Workshop on numerical prediction of ship squat in restricted waters, *Ship Technology Research* **61**, 162-165.

[11] https://www.uni-due.de/IST/ismt_presquat_test.shtml. (2014) *Introduction-Dynamic Squat Test Parameter, Presquat.*

[12] Stern, F., Yang, J. M., Wang, Z. Y., Sadat-Hosseini, H., Mousaviraad, M., Bhushan, S. and Xing, T. (2013) Computational ship hydrodynamics: Nowadays and way forward, *International Shipbuilding Progress* **60**, 3-105.

[13] Toxopeus, S. L., Simonsen, C. D., Guilmineau, E., Visonneau, M., Xing, T. and Stern, F. (2013) Investigation of water depth and basin wall effects on KVLCC2 in manoeuvring motion using viscous-flow calculations, *Journal of Marine Science and Technology* **18**, 471-496.

[14] Kaidi, S., Smaoui, H. and Sergent, P. (2017) Numerical estimation of bank-propeller-hull interaction effect on ship manoeuvring using CFD method, *Journal of Hydrodynamics, Ser B* **29**, 154-167.

[15] Tezdogan, T., Incecik, A. and Turan, O. (2016) A numerical investigation of the squat and resistance of ships advancing through a canal using CFD, *Journal of Marine Science and Technology* **21**, 86-101.

[16] Linde, F., Ouahsine, A., Huybrechts, N. and Sergent, P. (2016) Three-Dimensional Numerical Simulation of Ship Resistance in Restricted Waterways: Effect of Ship Sinkage and Channel Restriction, *Journal of Waterway, Port, Coastal, and Ocean Engineering* **143**.

[17] Zou, L. and Larsson, L. (2013) Computational fluid dynamics (CFD) prediction of bank effects including verification and validation, *Journal of Marine Science and Technology* **18**, 310-323.

[18] Wang, H. Z. and Zou, Z. J. (2014) Numerical prediction of hydrodynamic forces on a ship passing through a lock, *China Ocean Engineering* **28**, 421-432.

[19] Liu, H., Ma, N. and Gu, X. C. (2016) Numerical simulation of PMM tests for a ship in close proximity to sidewall and maneuvering stability analysis, *China Ocean Engineering* **30**, 884-897.

[20] Wang, H. M., Li, X. H., Chen, L. and Sun, X. J. (2016) Numerical study on the hydrodynamic forces on a ship berthing to quay by taking free-surface effect into account, *Journal of Marine Science and Technology* **21**, 601-610.
[21] http://www.simman2014.dk/.(2014) *Workshop on Verification and Validation of Ship Manoeuvring Simulation Methods (SIMMAN), Copenhagen, Denmark.*
[22] Delefortrie, G., Eloot, K. and Mostaert, F. (2012) *SIMMAN 2013: Execution of model tests with KCS and KVLCC2*. Antwerp, Belgium.
[23] Mucha, P., Deng, G., Gourlay, T. and Moctar, O. E.(2016) *Validation Studies On Numerical Prediction Of Ship Squat And Resistance In Shallow Water*, the *4th International Conference on Ship Manoeuvring in Shallow and Confined Water (MASHCON): Ship - Bottom Interaction*, Hamburg, Germany.
[24] Mucha, P. and Moctar, O. E. (2014) *Numerical Prediction of Resistance and Squat for a Containership in Shallow Water, the 17th Numerical Towing Tank Symposium*, Marstrand, Sweden.
[25] STAR-CCM+ User Guide. (2014) CD-Adapco.
[26] Hirt, C. W. and Nichols, B. D. (1981) Volume of fluid (VOF) method for the dynamics of free boundaries, *Journal of Computational Physics* **39**, 201-225.
[27] Muzaferija, S., Peric, M., Sames, P. and Schellin, T.(1998) *A two-fluid Navier-Stokes solver to simulate water entry, the 22nd symposium on naval hydrodynamics(ONR),* Washington, DC., USA.
[28] Patankar, S. V. and Spalding, D. B. (1972) A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows, *International Journal of Heat and Mass Transfer* **15**, 1787-1806.
[29] Launder, B. E. and Spalding, D. B. (1974) The numerical computation of turbulent flows, *Computer Methods in Applied Mechanics and Engineering* **3**, 269-289.
[30] Ohmori, T. (1998) Finite-volume simulation of flows about a ship in maneuvering motion, *Journal of Marine Science and Technology* **3**, 82-93.
[31] ITTC-Recommended Procedures and Guidelines. (2011) Practical Guidelines for Ship CFD Applications.
[32] Celik, I. B., Ghia, U., Roache, P. J. and Freitas, C. J. (2008) Procedure for estimation and reporting of uncertainty due to discretization in CFD applications, *Journal of Fluids Engineering-Transactions of the ASME* **130**.

An Optimal Control Obtained by Finite Dimensional Approximation for a Flexible Robot Arm

Xuezhang Hou
Mathematics Department, Towson University, Baltimore, Maryland 21252-0001, USA

Abstract
In this paper, we are concerned with a flexible robot arm formulated by partial differential equations with initial and boundary conditions. An optimal energy control of the robot arm is investigated after the system has been transformed to an abstract evolution system in an appropriate Hilbert space. An optimal energy control problem is discussed, and evantually, it is shown that an optimal energy control can be obtained by a finite dimensional approximation.

Keywords
Flexible Robot Arm, Optimal Control, Finite Dimensional Approximations.

1 Introduction

The vibration suppression of flexible Euler-Bernoulli beams has been studied extensively in many articles [1]−[7] due to its wide applications. We have studied control theory for flexible robot arms in the articles [8−11]. In the present paper, we are going to discuss a kind of optimal energy control problems for a flexible robot system.

We are now concerned with a flexible robot arm in the x-y plane. Since any motion of the robot arm in the $x-y$ plane can be decomposed into its x and y components, the vibration in the x-direction and y-direction can be considered independently which can be formulated by the following partial differential equations with the initial and the boundary conditions [9]:

$$\rho\ddot{w}(t,r) + EIw''''(t,r) = -\rho\ddot{x}(t), \qquad 0 < r < t, \tag{1.1}$$
$$w(t,0) = w'(t,0) = 0, \tag{1.2}$$
$$M[\ddot{w}(t,l) + \ddot{x}(t)] - EIw'''(t,l) = 0, \tag{1.3}$$
$$J\ddot{w}'(t,l) + EIw''(t,l) = 0, \tag{1.4}$$
$$w(0,r) = w_0(r), \quad \dot{w}(0,r) = w_1(r), \tag{1.5}$$

where m is the mass of a moving body driven by a control motor, M is the mass of the payload of the flexible robot arm that is attached to this moving body, $w(t,r)$ represents the amplitude of vibration of the flexible beam at time t and position r, $\ddot{w}(t)$ denotes the acceleration of the moving body in the x-direction, "$\cdot$" denotes the time derivative, and "$'$" denotes the spatial derivative, ρ denotes the line density of mass for the arm, EI denotes the bending rigid degree of the flexible beam, l denotes the length of the arm, J denotes the turning inertia, $w_0(r)$ and $w_1(r)$ denote the initial displacement and initial velocity of the arm, respectively.

For the motor system, we shall establish the following control equation:

$$m\dot{x}(t) = u(t) - EIw'''(t,0), \tag{1.6}$$

where the sliding friction was neglected and $u(t)$ is a control.
Let $y(t,r)$ be the total displacement in the x-direction of the flexible beam. Thus, we have

$$y(t,r) = w(t,r) + x(t). \tag{1.7}$$

Substituting (1.7) into the system (1.1)-(1.5) yields the following controlled closed-loop system equation about state $y(t,r)$:

$$\begin{aligned}
\rho\ddot{y}(t,r) + EIy''''(t,r) &= 0, \quad 0 < r < l, t > 0, && (1.8)\\
y'(t,0) &= 0, && (1.9)\\
m\ddot{y}(t,0) + EIy'''(t,0) &= u(t), && (1.10)\\
M\ddot{y}(t,l) - EIy'''(t,l) &= 0, && (1.11)\\
J\ddot{y}'(t,l) + EIy''(t,l) &= 0. && (1.12)
\end{aligned}$$

In order to investigate the system (1.8)-(1.12) under the abstract frame, we now consider a real Hilbert space $H = R^3 \times L^2_\rho(0,l)$ equipped with the inner product as

$$(\Phi_1, \Phi_2)_H = m\xi_1\xi_2 + M\eta_1\eta_2 + J\zeta_1\zeta_2 + \langle\varphi_1, \varphi_2\rangle_\rho$$

where $\Phi_i = [\xi_i, \eta_i, \zeta_i, \varphi_i]^\tau \in H$, $i = 1, 2$, $\langle\varphi_1, \varphi_2\rangle_\rho = \int_0^l \rho\varphi_1(x)\overline{\varphi_2}(x)dx$, and τ means the transpose. We define a linear operator A with domain $D(A)$ in H as follows:

$$A\tilde{\varphi} = \begin{bmatrix} \frac{EI}{m}\varphi'''(0) \\ -\frac{EI}{M}\varphi'''(l) \\ \frac{EI}{J}\varphi''(l) \\ \frac{EI}{\rho}\varphi''''(\cdot) \end{bmatrix}, \quad \text{for} \quad \tilde{\varphi} = \begin{bmatrix} \varphi(0) \\ \varphi(l) \\ \varphi'(l) \\ \varphi(\cdot) \end{bmatrix} \in D(A)$$

where $D(A) = \{\tilde{\varphi} \in H : \varphi, \varphi', \varphi'', \varphi''', \varphi'''' \in L^2_\rho(0,l), \varphi' = 0\}$.
Using the operator A, the system (1.8)-(1.12) becomes the following second-order abstract evolution equation in H:

$$\frac{d^2\tilde{y}(t)}{dt^2} + A\tilde{y}(t) = bu(t), \tag{1.13}$$

where $\tilde{y}(t) = [y(t,0), y(t,l), y'(t,l), y(t,\cdot)]^\tau, b = [1/m, 0, 0, 0]^\tau$.
First, we are going to propose and prove the following theorems[9]

Theorem 1.1 *$A : D(A) \to H$ is a nonnegative self-adjoint operator.*

Theorem 1.2 *The resolvent of A is a compact operator.*

Theorem 1.3 *The spectrum of A consists of only nonnegative eigenvalues with single multiplicity.*

We now define a Hilbert space $\mathcal{H} = H \times H$, and a linear operator $\mathcal{A}$ on $\mathcal{H}$ as follows

$$\mathcal{A} = \begin{bmatrix} 0 & A_r^{1/2} \\ -A_r^{1/2} + rA_r^{-1/2} & 0 \end{bmatrix}$$

where r is in $\rho(A)$, the resolvent set of A, and $D(\mathcal{A}) = D(A_r^{1/2}) \times D(A_r^{1/2})$. We also denote $[0, b]^\tau$ by $\mathcal{B}$, where b is defined in (1.13). Let us consider a subspace $\mathcal{T}$ of $\mathcal{H}$ consisting of $z = [z_1, z_2]^\tau$, where $z_1 = A_r^{1/2}\tilde{y}$, and $z_2 = \dot{\tilde{y}}$, and $\tilde{y}$ is defined in (1.13). In these notations, the equation (1.13) with initial conditions

$$\begin{aligned}
\tilde{y}(0) &= \tilde{y}_0 && (1.14)\\
\dot{\tilde{y}} &= \tilde{y}_1 && (1.15)
\end{aligned}$$

becomes a first-order evolution equation in $\mathcal{T}$ with initial conditions as follows:

$$\begin{aligned}
\frac{dz}{dt} &= \mathcal{A}z + \mathcal{B}u, && (1.16)\\
z_0 &= [A_r^{1/2}\tilde{y}_0, \tilde{y}_1]^\tau. && (1.17)
\end{aligned}$$

The semigroup property of the linear operator $\mathcal{A}$ has been proved in the article [9] as follows:

Theorem 1.4 *The operator $\mathcal{A}$ is the infinitesimal generator of a $C_0 - semigroup$ $T(t)$, $t \geq 0$.*

2 A Minimum Energy Problem

Since $T(t)$ is the semigroup of linear operators generated by the operator $\mathcal{A}$ (See the Theorem 1.4), it follows from the theory of semigroup of linear operators that the system (1.16)- (1.17) has an unique mild solution [12] given by

$$z(t) = T(t)z_0 + \int_0^t T(t-s)\mathcal{B}u(s)\, ds \tag{2.18}$$

Let $\varphi(\cdot)$ be an arbitrary element in $C([0,T];\mathcal{H})$, and

$$\rho = inf_{u\in L^2([0,T];\mathcal{H})}\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u(s)\, ds\|\,, \tag{2.19}$$

define the admissible control set of the system (1.16)-(1.17)as follows

$$\begin{aligned} U_{ad} = \{u \in L^2([0,T];\mathcal{H}) : \|\varphi(t) - T(t)z_0 \\ - \int_0^t T(t-s)\mathcal{B}u(s)\, ds\| \le \rho + \epsilon\} \end{aligned} \tag{2.20}$$

where ϵ is any positive number.
It can be seen from (2.2) that U_{ad} is not empty and contains infinitely many elements related to φ and ϵ. The minimum energy control problem is actually to find the element u, satisfying

$$\|u_0\| = min\{\|u\| : u \in U_{ad}\} \tag{2.21}$$

where u_0 is said to be a minimum energy control element.

Lemma 2.1 *The admissible control set U_{ad} defined by (2.3) is a closed convex set in Hilbert space $L^2([0,T];\mathcal{H})$.*

Proof. Convexity. For any u_1, $u_2 \in U_{ad}$ and a real number λ, $0 < \lambda < 1$, it is easy to see from (2.2) that

$$\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u_i(s)\| \quad \le \quad \rho + \epsilon \quad i = 1, 2$$

and hence

$$\begin{aligned} & \|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}[\lambda u_1(s) \\ & \qquad +(1-\lambda)u_2(s)]\, ds\|x \\ = \quad & \|\varphi(t) - T(t)z_0 - \int_0^T T(t-s)[\lambda\mathcal{B}u_1(s) \\ & \qquad +(1-\lambda)\mathcal{B}u_2(s)]\, ds\| \\ \le \quad & \lambda\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u_1(s)\, ds\| \\ & +(1-\lambda)\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u_2(s)\, ds\| \\ \le \quad & \lambda(\rho+\epsilon) + (1-\lambda)(\rho+\epsilon) = \rho + \epsilon. \end{aligned}$$

Since $\lambda u_1 + (1-\lambda)u_2 \in L^2([0,T];H)$, it follows that $\lambda u_1 + (1-\lambda)u_2 \in U_{ad}$, this implies that U_{ad} is a convex subset of $L^2([0,T];H)$.
Closedness. Suppose $\{z_n\} \subset U_{ad}$, and $\lim_{n\to\infty}\|u_n - u^*\| = 0$. It can be shown that $u^* \in U_{ad}$. In fact, we see from the definition of U_{ad} that

$$\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u_n(s)\, ds\| \le \rho + \epsilon, \quad n = 1, 2, \cdots$$

Since $T(t),\ t \geq 0$ is a C_0-semigroup in Hilbert space $\mathcal{H}$, there is a constant $M > 0$ such that $\sup_{0 \leq t \leq T} \|T(t)\| \leq M$. On the other hand, since $u(s)$ is differentiable on $[0,T]$, it is continuous on $[0,T]$, and hence $\{u(s) : s \in [0,T]\}$ is a bounded set in $L^2([0,T];Y)$. Thus there is a constant $N > 0$ such that $\|Bu(s)\| \leq N\ \ (0 \leq s \leq T)$ and

$$\begin{aligned} & \|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u^*(s)\ ds\| \\ \leq\ & \|\varphi(t) - T(t)z_0 - \int_0^t \mathcal{B}u_n(s)\ ds\| \\ & \qquad + \|\int_0^t T(t-s)\mathcal{B}[u_n(s) - u^*(s)]\ ds\| \\ \leq\ & \rho + \epsilon + M\|u_n - u^*\| \cdot NT \end{aligned}$$

Letting $n \to \infty$ leads to

$$\|\varphi(t) - T(t)z_0 - \int_0^t T(t-s)\mathcal{B}u^*(s)\ ds\| \leq \rho + \epsilon\,.$$

Thus, $u^* \in U_{ad}$, and U_{ad} is a closed set and the proof is complete.

3 Optimal Energy Control of the System

In this section, we will discuss the existence and uniqueness of optimal energy control for the flexible robot system (1.16)-(1.17). Let us begin with the following theorem.

Theorem 3.1 *There exists an unique minimum energy control element in the admissible control set U_{ad} corresponding to the system (1.16) and (1.17).*

Proof. Since $L^2([0,T],Y)$ is a Hilbert space, it is naturally a strict convex Banach Space. From the preceding Lemma 2.1, we have seen that U_{ad} is a closed convex set in $L^2([0,T],Y)$, and it follows from [13] that there is an unique element $u_0 \in U_{ad}$ such that

$$\|u_0\| = min\ \ \{\|u\| : u \in U_{ad}\}$$

According to the definition (2.4), u_0 is just the desired minimum energy control element of the system (1.16)-(1.17). The proof is complete.
Finally, we shall show that the minimum energy control element can be approached.

Theorem 3.2 *Suppose that u_0 is the minimum energy control element of the system (1.16)-(1.17), then there exists a sequence $\{u_n\} \subset U_{ad}$ such that $\{u_n\}$ converges strongly to u_0 in $L^2([0,T];Y)$, namely,*

$$\lim_{n\to\infty} \|u_n - u_0\| = 0$$

Proof. Let $\{u_n\}$ be a minimized sequence in the admissible control set U_{ad}, then it follows that

$$\|u_{n+1}\| \leq \|u_n\|, \quad n = 1, 2, \cdots \tag{3.22}$$

and

$$\lim_{n\to\infty} \|u_n\| = inf\{\|u\| : u \in U_{ad}\} \tag{3.23}$$

It is obvious that $\{u_n\}$ is a bounded sequence in $L^2([0,T];Y)$, and so there is a subsequence $\{u_{n_k}\}$ of $\{u_n\}$ such that $\{u_{n_k}\}$ weakly converges to an element $\widetilde{u}$ in $L^2([0,T];Y)$[14].
Since U_{ad} is a closed convex set in $L^2([0,T];Y)$ based on the Lemma 2.1, we see from Mazur's Theorem that U_{ad} is a weakly closed set in $L^2([0,T];Y)$, thus $\tilde{u} \in U_{ad}$. Combining (3.2) and employing the properties of limits of weakly convergent sequence on norm yield

$$\begin{aligned} inf\{\|u\| : u \in U_{ad}\} \leq \|\tilde{u}\| \ &\leq\ \varliminf_{k\to\infty} \|u_{n_k}\| \\ = \lim_{n_k\to\infty} \|u_{n_k}\| = \lim_{n\to\infty} \|u_n\| \ &=\ \inf\{\|u\|; u \in U_{ad}\}. \end{aligned}$$

Thus, we have

$$\lim_{n\to\infty} \|u_n\| = \|\tilde{u}\| \tag{3.24}$$

and

$$\|\tilde{u}\| = inf\{\|u\| u \in U_{ad}\}. \tag{3.25}$$

Since $\{u_{n_k}\}$ is weakly convergent to $\tilde{u}$, it follows from (3.3) that $\{u_{n_k}\}$ converges to $\tilde{u}$. Therefore, we see from the Theorem 3.1 and (3.4) that $\tilde{u} = u_0$, namely, $\tilde{u}$ is the minimum energy control element. Thus, $\{u_{n_k}\}$ strongly converges to the minimum energy control element in $L^2([0,T];Y)$. Without loss of generality, we can rewrite $\{u_{n_k}\}$ by $\{u_n\}$, and the conclusion of theorem is now obtained.

4 Coclusion

In this paper, we have investigated a kind of optimal energy control for a flexible robot arm formulated by partial differential equations with initial and boundary conditions. After a discussion of minimum energy problem for the beam system, we have proposed and proved the existence and uniqueness Theorem 3.1 of the optimal energy control in terms of semigroup approach of linear operators. Finally, we gave an approximation result Theorem 3.2 that points out that the minimum energy control element can be approached by a weakly convergent sequence in the control space, and provides the theoretical basis of approximate computation for finding an optimal energy control element.

References

[1] G. Chen, M.C. Delfour, A.M. Krall, G. Payre, *Modeling Stabilization and Control of Serially Connected Beam*, SIAM J. Control Optim. 25(1987), pp526-546.

[2] F. Conrad, *Stabilization of Beams by Pointwise Feedback Control*, SIAM J. Control Optim. 28 (1990), pp. 423 - 437.

[3] Z.H.Luo., *Direct Strain Feedback Control of Flexible Robot Arms: New Theoretical and Experimental Results*, IEEE Trans. Automat. Control, 38 (1993), pp. 1610 - 1622.

[4] B.Z.Guo, *Riesz Basis Approach to Stabilization of a Flexible Beam with a Tip Mass*, SIAM J. Control Optim. 39(2001), pp. 1736-1747.

[5] B.Z. Guo, *Riesz Basis Property and Exponential Stability of Controlled Euler-Bernoulli Beam Equation with Variable Coefficients*, SIAM J. Control & Optim., 40(6)(2002), 1905-1923.

[6] R.F. Curtain, H.J. Zwart, *An Introduction to Infinite Dimensional Linear System Theory*, Springer-Verlag, New York, 1995.

[7] Z.H.Luo, B.Z.Guo and O.Morgul, *Stability and Stabilization of Infinite-dimensional System with Applications*, Springer-Verlag, London (1999).

[8] Bao-Zhu Guo, Yu Xie, Xuezhang Hou, *On Spectrum of a General Petrovsky Type Equation and Riesz Basis of N-Connected Beams with Linear Feedback at Joints*, Journal of Dynamical and Control Systems, Vol. 10, No.2, 187-211(2004).

[9] Xuezhang Hou and Sze-Kai Tsui, *A Control Theory for Cartesian Flexible Robot Arms*, Journal of Mathematical Analysis and Applications; **225**(1998), 265-288.

[10] Xuezhang Hou and Sze-Kai Tsui, *Control and Stability of a Torsional Elastic Robot Arm*, Journal of Mathematical Analysis and Applications; **243**(2000), 140-162.

[11] Xuezhang Hou and Sze-Kai Tsui, *A Feedback Control and a Simulation of a Torsional Flexible Robot Arm*, Applied Mathematics and Computations, Vol. 142(2003), pp. 389-407.

[12] A Pazy, *Semigroup of Linear Operators and Applications to Partial Differetnail Equations*, Springer-Verlag, Berlim, 1983.

[13] Barbu, V. and Precupana, Th.,*C*onvexity and Optimization in banach Space, Ed. Acad. Rep. Soe Romania, Bucuresti, 1978.

[14] A. E. Taylor. D. C. Lay, *Introduction for Functional Analysis.* 2nd ed. New York, John Wiley and Sons, (1980).

Finite element based micromechanical model for elastic materials containing nanoscale inhomogeneities

***S. Chaisuwannakorn [1] and †Y. Sapsathiarn[1]**

[1] Department of Civil and Environmental Engineering, Faculty of Engineering, Mahidol University, Thailand

*Presenting author: salisa.chi@student.mahidol.edu
†Corresponding author: yasothorn.sap@mahidol.edu

Abstract

Nano-structured materials (e.g. nanocomposites, nanoporous materials, nanocrystalline materials, etc.) and nano-scale structural elements (e.g. nanotubes, nanofilms, nanobeams, etc.) have unique mechanical and physical properties. For nano-structured materials containing inhomogeneities (e.g. voids and particles) in the nanoscale dimensions such as nanoporous materials and nanocomposites, the size effect due to surface energy due to nanoscale inhomogeneities can play an important role on their mechanical properties and responses. In this paper, the finite element based micromechanical model for analysis of materials containing nanoscale inhomogeneities incorporating Gurtin-Murdoch surface theory is presented. The proposed micromechanical model is applied to examine the responses and properties of nano-structured materials, i.e., nanoporous and nanocomposite materials. Selected numerical results are presented to portray the features of the elastic field responses and properties of elastic materials with nanoscale inhomogeneities. The finite element-based micromechanical model presented in this paper is an efficient tool to analyze the response and predict the mechanical properties of nano-structured materials.

Keywords: Nanotechnology, micromechanics theory, inhomogeneity, nanocomposites, nanoporous materials

Introduction

Nanomechanics is a study of responses and properties of materials and structures at the nanoscale. Steritz et al. [1] and Dingreville et al. [2] have clearly explained that atoms adjacent to the free surface have a different local environment than do atoms in the bulk of a material. The nanoscale materials or structures contain larger fraction of energy associated with surface atoms when compared to those in the bulk material, and as a result, structures at the nanoscale are known to exhibit size-dependent behavior.

Within the context of modelling nanoscale behavior of materials and structures, two predominant mathematical approaches have been commonly employed in the literature, one known as the molecular or atomistic simulations and the other corresponding to the modified continuum-based models. The molecular-based models, while providing more direct response prediction, generally consume tremendous computational resources because billions of atoms at the nanoscale is needed to include in the simulation models. The continuum-based models, in contrast to the molecular-based models, are less complicated and much more computationally efficient.

Gurtin-Murdoch model, proposed by Gurtin and Murdoch [3, 4], is a mathematical model that incorporates the effects of surface and interfacial energy into continuum mechanics. A good agreement between solutions based on the Gurtin–Murdoch model and atomistic simulations for nano-scale structures has been reported by various researchers (e.g., [5-7]). Sapsathiarn

and Rajapakse [7] shown that the Gurtin-Murdoch nanoscale beam model is capable of simulating the experimental results of chromium cantilever beams loaded by an atomic force microscope. Mogilevskaya et al. [8] considered the multiple interaction of circular nano-inclusions in unbounded domain by using a complex variables formulation. Fang et al. [9] studied the elastic interaction between screw dislocations and an embedded coated circular nanowire with interface stresses based on Gurtin and Murdoch theory and explained that the effect of the interface stress on the motion and the equilibrium position of the dislocation near the nanowire is significant when the radius of the nanowire is reduced to nanometer dimensions. A finite-element formulation for static and dynamic modeling of circular nanoplates based on Gurtin-Murdoch theory has been presented by Sapsathiarn and Rajapakse [10]. Mi and Kouris [11] examined the stress concentration Stress concentration in the vicinity of a nanovoid near the free surface of an elastic half-space and its dependence on surface properties.

In this paper, a finite element based micromechanical model for elastic materials with nanoscale inhomogeneities incorporating Gurtin-Murdoch surface stress effects is developed. Selected numerical results for the elastic fields and properties of elastic composites containing nanoscale inhomogeneities, i.e., nanoporous and nanocomposite materials, are presented. The finite element-based micromechanical model of nanoparticle-reinforced composites developed in the present study is an efficient tool to investigate the response and properties of nano-structured materials with practically useful arbitrary shaped nanoscale inhomogeneities, multiple voids/particles, non-symmetric loading, etc.

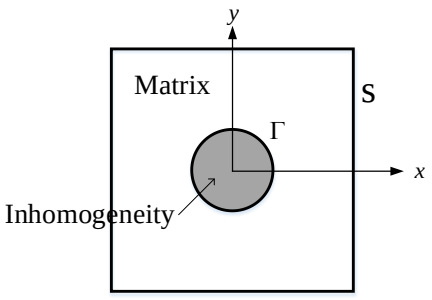


Figure 1. A representative volume element for composite materials containing a nanoscale inhomogeneity, e.g., nanovoid or nanoparticle.

Theoretical consideration

Consider a two-dimensional material plane containing a nanoscale inhomogeneity, e.g., nanovoid or nanoparticle, as shown in Fig. 1. In the case of material with a nanoparticle, the matrix and inhomogeneity are considered as linearly orthotropic materials and the matrix-inhomogeneity bonding is assumed to be perfect. The Cartesian coordinates (x,y) is used in the formulation. The research methodology, procedures and fundamental theories to be employed in the analysis are summarized in the subsequent sections.

Governing equations and surface elasticity model

Regarding linear elasticity theory, the equilibrium equation in the absence of body forces and the constitutive relation of the bulk material can be written using the standard indicial notation as

$$\sigma_{ij,j}^{B} = 0;\ \sigma_{kl}^{B} = C_{ijkl}\varepsilon_{ij} \tag{1}$$

where $\sigma_{ij,j}^{B}$ denotes the components of stress tensor for a bulk material (inhomogeneity and matrix); ε_{ij} and σ_{kl}^{B} denote the second-rank tensors of strain and stress respectively; and the elastic matrix C_{ijkl} is the fourth-rank tensor.

The elastic matrices **D** in the Voigt notation for plane stress and plane strain deformations can be written as

$$\mathbf{D}=\left[C_{ijkl}\right]=\begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{22} & 0 \\ 0 & 0 & C_{66} \end{bmatrix}; \quad \mathbf{D}=\begin{bmatrix} C_{11}-(D_{13}^{\ 2}/C_{33}) & C_{12}-(C_{13}C_{23}/C_{33}) & 0 \\ C_{12}-(C_{13}C_{23}/C_{33}) & C_{22}-(C_{23}^{\ 2}/C_{33}) & 0 \\ 0 & 0 & C_{66} \end{bmatrix} \tag{2}$$

(plane stress) (plane strain)

where C_{ij} is the component of elastic compliance matrix in the Voigt notation.

The incorporation of surface stress effects is needed due to the fact that the inhomogeneity in the composite is in a nanoscale size. Gurtin and Murdoch [3, 4] proposed a surface stress model to account for the surface effects at the nanoscale. Models of nanoscale structures based on the Gurtin-Murdoch continuum theory have an elastic surface, mathematically zero thickness, perfectly bonded to the bulk material. The elastic surface has distinct material properties and accounts for the surface energy effects [6]. The generalized Young-Laplace equation [12], surface constitutive relations and strain-displacement relationship of the surface can be expressed as [3, 4]

$$\sigma_{\beta\alpha,\beta}^{s}+\langle\sigma_{\beta\alpha}^{B}n_{\beta}\rangle=0;\ \langle\sigma_{ji}^{B}n_{i}n_{j}\rangle=\sigma_{\beta\alpha}^{s}k_{\beta\alpha} \tag{3a}$$

$$\sigma_{\beta\alpha}^{s}=\tau_{0}\delta_{\beta\alpha}+2\ (\mu^{S}-\tau_{0})\ \varepsilon_{\beta\alpha}+\ (\lambda^{S}+\tau_{0})\ \varepsilon_{\gamma\gamma}\delta_{\beta\alpha}+\tau_{0}u_{\beta,\alpha}^{s} \tag{3b}$$

$$\varepsilon_{\alpha\beta}^{s}=\frac{1}{2}\ u_{\alpha,\beta}^{s}+u_{\beta,\alpha}^{s} \tag{3c}$$

where superscripts *B* and *S* denote the quantities corresponding to the bulk and the surface respectively; $\langle * \rangle=(*)_{M}-(*)_{I}$ denotes the jump of the field quantity across the inhomogeneity and matrix interface where the subscripts *M* and *I* are used to identify quantities associated with the matrix and the inhomogeneity respectively; μ^{S} and λ^{S} are surface Lamé constants; τ_{0} is the residual surface tension under unstrained conditions; n_{i} denotes the components of the unit normal vector of the surface; and $k_{\beta\alpha}$ is the curvature tensor of the surface. It should be noted that the surface material properties, μ^{S}, λ^{S} and τ_{0} can be determined from atomistic simulations [13].

Tian and Rajapakse [14] presented a finite element formulation for the analysis of a two-dimensional elastic material plane containing a nanoscale inhomogeneity by employing the energy method. The potential energy ($\prod$) of the system in Fig. 1 consists of the elastic strain energies of the bulk inhomogeneity (U^{BI}) and matrix (U^{BM}) materials, the surface elastic strain energy (U^{S}) due to the surface effects and the potential energy (*W*) due to the application of external loads, and can be written as

$$\prod=U^{BI}+U^{BM}+U^{S}+W \tag{4}$$

in which the superscript *BI* and *BM* denotes quantities corresponding to the bulk inhomogeneity and matrix materials respectively.

The elastic strain energies for the bulk inhomogeneity (U^{BI}) and for the matrix material (U^{BM}) can be expressed as

$$U^{BI} = \int_{V_I} \int_0^{\varepsilon_{ij}} \sigma_{ij}^B d\varepsilon_{ij} dV \; ; \; U^{BM} = \int_{V_M} \int_0^{\varepsilon_{ij}} \sigma_{ij}^B d\varepsilon_{ij} dV \tag{5}$$

The potential energy (W) due to the application of external loads can be expressed as

$$W = -\int_{V_I} \{u\}^{\mathrm{T}} \{T\} dS \tag{6}$$

where T and u denote the vectors of surface traction and surface displacement respectively; and the superscript T denotes the transpose of a vector or matrix.

Based on the Gurtin-Murdoch surface stress model expressed in Eq. (3), the surface elastic strain energy (U^S) can be obtained as

$$U^S = \int_{\mathbf{\Gamma}+S} \int_0^{\varepsilon_{\alpha\beta}} \sigma_{\alpha\beta}^S d\varepsilon_{\alpha\beta} d\mathbf{\Gamma} \tag{7}$$

Introducing the element shape function $N(x,y)$ to interpolate the field variables u within an element by

$$\{u\} = [N]\{\bar{u}\} \tag{8}$$

where u denotes nodal displacement vector.

The element strain vector {ε} can be determined from Eq. (8) using the classical strain-displacement relation as

$$\{\varepsilon\} = [B]\{\bar{u}\} \tag{9}$$

where $[B]$ is a strain-displacement matrix in which the elements are the derivatives of the element shape functions, $[B] = \partial[N] / \partial x_i$.

Substitution of Eq. (8) and (9) into Eq. (5) - (7) together with the constitutive relations for bulk (matrix and inhomogeneity) and surface materials, Eq. (4) becomes

$$\begin{aligned} \prod = & \int_{V_M} \frac{1}{2} \{\bar{u}\}^{\mathrm{T}} [B]^{\mathrm{T}} [D]_M [B] \{\bar{u}\} dV + \int_{V_I} \frac{1}{2} \{\bar{u}\}^{\mathrm{T}} [B]^{\mathrm{T}} [D]_I [B] \{\bar{u}\} dV \\ & - \int_S [N] \{\bar{u}\}^{\mathrm{T}} \{T\} dS + \int_{\mathbf{\Gamma}+S} \frac{1}{2} \{\bar{u}\}^{\mathrm{T}} [B]^{\mathrm{T}} [D]_S [B] \{\bar{u}\} d\mathbf{\Gamma} \end{aligned} \tag{10}$$

Applying the stationary condition of $\prod$, i.e., $\delta \prod = 0$, with respect to the nodal displacement components, the equilibrium equations for system in Fig. 1 can be obtained as

$$[K]\{\bar{u}\} = \{f\} \tag{11}$$

where

$$[K]=\int_{V_M}[B]^{\mathrm{T}}[D]_M[B]dV+\int_{V_I}[B]^{\mathrm{T}}[D]_I[B]dV+\int_{\Gamma+S}[B]^{\mathrm{T}}[D]_S[B]d\Gamma \tag{12a}$$

$$\{f\}=\int_S[N]\{T\}dS \quad f \quad =\int_S N \quad T \quad dS \tag{12b}$$

Micromechanical model of nanocomposites

A micromechanical model based on finite element formulation given in the preceding section is developed in the present study for analysis of elastic materials containing nanoscale inhomogeneities. The micromechanical analysis is performed by using a micromechanics theory which relates mechanics between two different length scale problems, i.e., (1) the macroscopic level in which the material is conceptually represented as a homogeneous material and (2) the level of the constituents in which the material properties are always heterogeneous and consist of distinguishable phases such as the main matrix material, inclusions and cavities or voids. Properties of elastic materials containing nanoscale inhomogeneities can be determined by the analysis in the level of the constituents. The analysis might be performed on a "representative volume element" or a "unit cell" which can be isolated from the composite material and is in a state of equilibrium. The unit cell for materials considered in the present work is schematically presented in Fig. 1.

The macroscopic constitutive relation of the materials with nanoscale inhomogeneities can be expressed in terms of the macro stress and the macro strain as

$$\{\bar{\sigma}\}=\left[C^{eff}\right]\{\bar{\varepsilon}\} \tag{13}$$

where

$$\{\bar{\sigma}\}=\begin{matrix}\bar{\sigma}_{xx} & \bar{\sigma}_{yy} & \bar{\sigma}_{xy}\end{matrix}^{\mathrm{T}};\ \{\bar{\varepsilon}\}=\begin{matrix}\bar{\varepsilon}_{xx} & \bar{\varepsilon}_{yy} & \bar{\varepsilon}_{xy}\end{matrix}^{\mathrm{T}} \tag{14a}$$

$$\left[C^{eff}\right]=\begin{bmatrix}C_{11}^{eff} & C_{12}^{eff} & 0\\ C_{12}^{eff} & C_{22}^{eff} & 0\\ 0 & 0 & C_{66}^{eff}\end{bmatrix} \tag{14b}$$

According to the micromechanics theory, the macro stress, $\bar{\sigma}_{ij}$, and macro strain, $\bar{\varepsilon}_{ij}$, can be defined as the volume average stress in a RVE as

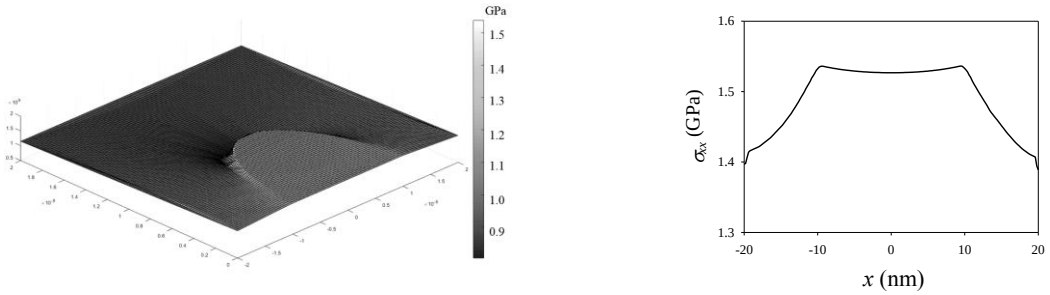
$$\bar{\sigma}_{ij}=\frac{1}{V}\int_\Omega \sigma_{ij}\,d\Omega;\ \ \bar{\varepsilon}_{ij}=\frac{1}{V}\int_\Omega \varepsilon_{ij}\,d\Omega \tag{15}$$

Numerical results and discussion

A selected set of numerical solutions is presented in this section for the plane strain case of elastic material with nanoscale inhomogeneity to portray the response within nano-structured materials and investigate the influence of inhomogeneity volume fraction to the mechanical properties of the nano-structured materials. Two types of materials are considered in the numerical simulation, i.e. (1) nanocomposite materials (i.e., materials containing nanoscale particles) and (2) nanoporous materials (i.e., materials containing nanoscale cavities or voids). The matrix and particle inhomogeneity materials are considered to be linearly elastic and isotropic in the numerical study with Lame′ constants, E_M = 40 GPa, ν_M = 0.20 GPa for the

matrix material and E_I = 80 GPa, ν_I = 0.25 GPa for the particle inhomogeneity material. The surface parameter $K^S = 2\mu^S + \lambda^S - \tau_0 = 10$ N/m is considered in the numerical example.

The unit cell subjected to a prescribed displacement in the *x*-direction over the positive *x* face (a surface that is perpendicular to the *x*-axis and on the positive *x* side) is considered in the numerical example. The other faces are constrained in such a way that only the movement in the *x*-direction is allowed and the displacements in all other directions are prevented. The properties of nanocomposite and nanoporous materials can be determined from the fields within the unit cell being considered by using Eqs. (13)-(15).



(a) σ_{xx} distribution (b) σ_{xx} profile (across the inhomogeneity)

Figure 2. (a) Distribution and (b) profile of the stress σ_{xx} for nanocomposites with circular nano-particles

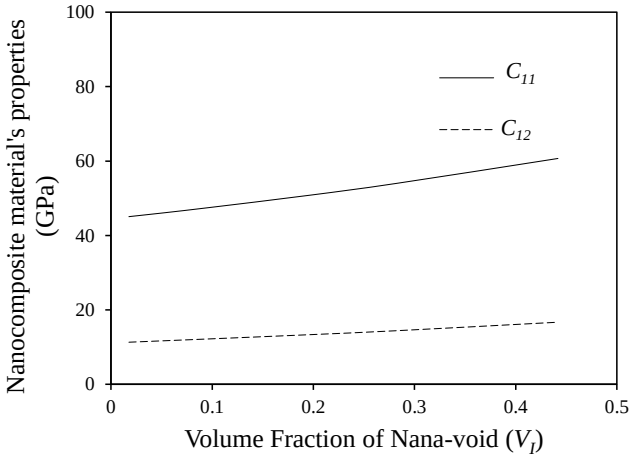


Figure 3. Variation of material properties of nanocomposites with circular nano-particles versus the volume fraction of the inhomogeneity (V_I).

The numerical results for the case of nanocomposites containing circular nano-particles are presented in Figs. 2 and 3. The distribution of stress σ_{xx} over the half-domain of the unit cell is presented in Fig. 2(a) for a nanocomposite material with volume fraction of the inhomogeneity V_I = 0.2. The corresponding profile of stress σ_{xx} along the *x*-axis across the center of the nano-particle inhomogeneity is presented in Fig. 2(b). The unit cell is stretched and the tensile stress σ_{xx} is generated all over the unit cell. The stress σ_{xx} in the domain of inhomogeneity is generally higher compared to those in the matrix domain showing the stress

disturbance in a composite material due to the presence of the nanoscale inhomogeneity. The influence of volume fraction of the inhomogeneity (V_I) to the mechanical properties C_{11}^{eff} and C_{12}^{eff} of the nanocomposite can be observed in Fig. 3. It is shown in Fig. 3 that the properties of nanoinclusion material are increasing as the volume fraction V_I increases. It should be observed that the relations between the material coefficients and V_I are non-linear.

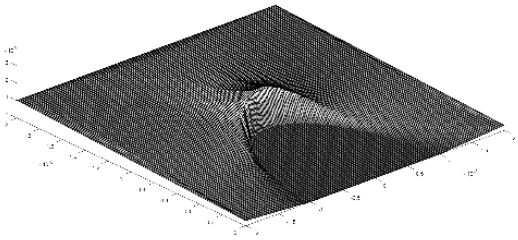


Figure 4. Distribution of the stress σ_{xx} for nanoporous material's properties with circular nano-voids.

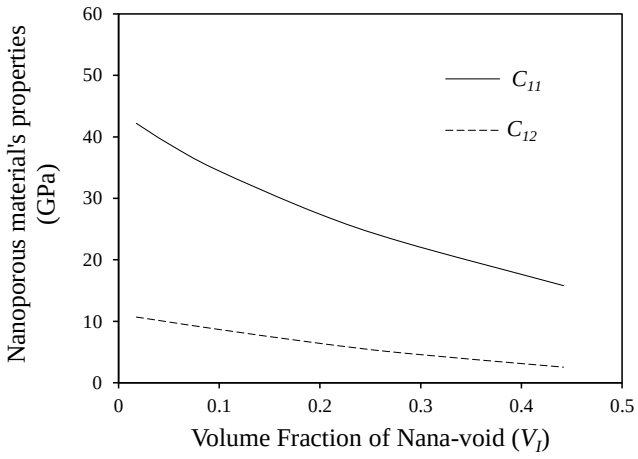


Figure 5. Variation of nanoporous material's properties (circular nano-voids) versus the volume fraction of the inhomogeneity (V_I).

The stress distribution and properties of nanoporous material with circular nano-voids considered in the numerical study are presented in Figs. 4 and 5 respectively. The volume fraction of the inhomogeneity (nanoscale voids) considered in Fig. 4 is V_I = 0.2. Similar behavior is observed for the case of a nanoporous material, i.e., the unit cell is stretched and the tensile stress σ_{xx} is generated all over the domain. Based on the results shown in Fig. 4, the stress concentration in the vicinity of a nanovoid is noted. Similar to the case of nanocomposite, the dependence of nanoporous material's properties on the volume fraction of the inhomogeneity (*VI*) is non-linear (see Figs. 3 and 5). As expected, the properties of nanoporous material are decreasing as the volume fraction of nanovoid V_I increases.

Conclusions

In this paper, a finite element-based micromechanics model for analysis of elastic materials containing nanoscale inhomogeneities incorporating surface stress effects has been developed. The Gurtin-Murdoch surface elasticity is employed in the micromechanical model to incorporate the interface energy effects of nanoscale inhomogeneity. Selected numerical results are presented to portray the features of the elastic field responses and properties of elastic materials with nanoscale inhomogeneities. Two types of materials are considered in the numerical simulation, i.e. nanocomposite materials (i.e., materials containing nanoscale particles) and nanoporous materials (i.e., materials containing nanoscale cavities or voids). Numerical results of stress and material properties for nanocomposite and nanoporous materials show considerable dependence on volume fraction of inhomogeneity. The finite element-based micromechanical model provides an efficient tool to analyze and predict the mechanical response of nano-inhomogeneities with arbitrary-shaped nanoscale particles, multiple particles, nanovoid, multiple nanovoid, non-symmetric loading, etc.

Acknowledgements

The work presented in this paper is supported by MRG research grant for New Scholar from Thailand Research Fund and partially supported by Graduate Studies of Mahidol University Alumni Association, Mahidol University.

References

[1] Streitz, F. H., Cammarata, R. C. and Sieradzki, K. (1994) Surface-stress effects on elastic properties. I. Thin metal films, *Physical Review B* **49**, 10699–706.
[2] Dingreville, R., Qu, J. and Mohammed, C. (2005) Surface free energy and its effect on the elastic behavior of nano-sized particles, wires and films, *Journal of the Mechanics and Physics of Solids* **53**, 1827–54.
[3] Gurtin, M. E. and Murdoch, A. I. (1975) A continuum theory of elastic material surfaces, *Archive for Rational Mechanics and Analysis* **57**, 291–323
[4] Gurtin, M. E. and Murdoch, A. I. (1978) Surface stress in solids, *International Journal of Solids and Structures* **14**, 431–440.
[5] Lee, B. and Rudd, R. E. (2007) First-principles study of the Young's modulus of Si ⟨001⟩ nanowires, *Physical Review B* **75**, 041305(R).
[6] Miller, R. E. and Shenoy V.B. (2000) Size-dependent elastic properties of nanosized structural elements, *Nanotechnology* **11**, 139–147.
[7] Sapsathiarn, Y. and Rajapakse, R. K. N. D. (2012) A model for large deflections of nanobeams and experimental comparison, *IEEE Transactions on Nanotechnology* **11**, 247–254.
[8] Mogilevskaya, S. G., Crouch, S. L. and Stolarski, H. K. (2008) Multiple interacting circular nano-inhomogeneities with surface/interface effects, *Journal of the Mechanics and Physics of Solids* **56**, 2298–327.
[9] Fang, Q. H., Liu, Y. W., Jin, B. and Wen, P. H. (2009) Interaction between a dislocation and a core–shell nanowire with interface effects, *International Journal of Solids and Structures* **46**, 1539–1546,
[10] Sapsathiarn, Y. and Rajapakse, R. K. N. D. (2013) Finite-element modeling of circular nanoplates, *Journal of Nanomechanics and Micromechanics* **3**, 59–66.
[11] Mi, C. and Kouris, D. (2013) Stress concentration around a nanovoid near the surface of an elastic half-space, *International Journal of Solids and Structures* **50**, 2737–2748.
[12] Povstenko, Y. Z. (1993) Theoretical investigation of phenomena caused by heterogeneous surface tension in solids, *Journal of the Mechanics & Physics of Solids* **41**, 1499–1514.
[13] Shenoy, V. B. (2005) Atomistic calculations of elastic properties of metallic fcc crystal surfaces. *Physical Review B* **71**, 094104-1–094104-11.
[14] Tian, L. and Rajapakse R. K. N. D. (2007) Finite element modelling of nanoscale inhomogeneities in an elastic matrix, *Computational Materials Science* **41**, 44–53.

Rectangle clamped at one end: Exact solution

Alexander P. Kerzhaev

Laboratory of Geodynamics, Institute of Earthquake Prediction Theory and Mathematical Geophysics, Russian Academy of Sciences, Russia

Abstract

In the paper the exact solution of a boundary value problem for a rectangle clamped at one end is constructed. The solution is given in the form of explicit expansions in Papkovich–Fadle eigenfunctions. The coefficients of the expansions are clearly determined by means of functions biorthogonal to Papkovich–Fadle eigenfunctions.

Keywords: Plate; clamp; Papkovich–Fadle eigenfunctions; exact solution

Introduction

Numerous publications are devoted to the approximate and numerical solutions of boundary value problems of the theory of elasticity for a rectangle with a clamped end (ends). The main reason for special interest in these problems is, partly, that sometimes the obtained solutions of the same problem could noticeably differ in different authors' works, depending on the way which method (approach) was used for their construction. For example, at the angular point some solutions had a singularity that is characteristic of an infinite rectangular wedge, one face of which is rigidly clamped and to the other one is applied an external load. In other solutions this singularity was absent.

In this paper is constructed the exact solution of a boundary value problem of the theory of elasticity for a rectangle, the left end of which is clamped and on the right end is applied a normal load (even-symmetric and odd-symmetric deformations).

Statement of the Problem

Let us consider a rectangle $\{P:|y|\le 1,\ 0\le x\le d\}$. We will assume that the long sides $y=\pm 1$ are free, i.e.

$$\sigma_y(x,\pm 1)=\tau_{xy}(x,\pm 1)=0, \tag{1}$$

the left end $x=0$ is clamped, and a normal load is applied on the right end $x=d$, i.e.

$$\begin{aligned} u(0,y)=v(0,y)=0,\\ \sigma_x(d,y)=\sigma(y),\tau_{xy}(d,y)=0. \end{aligned} \tag{2}$$

Then the solution in the rectangle written as expansions in Papkovich–Fadle eigenfunctions has the following form:

$$\begin{aligned} U(x,y)=\sum_{k=1}^{\infty} A_k\xi(\lambda_k,y)\sinh\lambda_k x+B_k\xi(\lambda_k,y)\cosh\lambda_k x+\\ +\overline{A_k}\xi\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k}x+\overline{B_k}\xi\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k}x, \end{aligned}$$

$$\begin{aligned} V(x,y)=\sum_{k=1}^{\infty} A_k\chi(\lambda_k,y)\cosh\lambda_k x+B_k\chi(\lambda_k,y)\sinh\lambda_k x+\\ +\overline{A_k}\chi\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k}x+\overline{B_k}\chi\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k}x, \end{aligned}$$

$$\begin{aligned}
\sigma_x(x,y) &= \sum_{k=1}^{\infty} A_k s_x(\lambda_k,y)\cosh\lambda_k x + B_k s_x(\lambda_k,y)\sinh\lambda_k x + \\
&\quad + \overline{A_k} s_x\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k} x + \overline{B_k} s_x\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k} x, \\
\sigma_y(x,y) &= \sum_{k=1}^{\infty} A_k s_y(\lambda_k,y)\cosh\lambda_k x + B_k s_y(\lambda_k,y)\sinh\lambda_k x + \\
&\quad + \overline{A_k} s_y\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k} x + \overline{B_k} s_y\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k} x, \\
\tau_{xy}(x,y) &= \sum_{k=1}^{\infty} A_k t_{xy}(\lambda_k,y)\sinh\lambda_k x + B_k t_{xy}(\lambda_k,y)\cosh\lambda_k x + \\
&\quad + \overline{A_k} t_{xy}\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k} x + \overline{B_k} t_{xy}\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k} x.
\end{aligned} \tag{3}$$

Here $U(x,y) = Gu(x,y)$, $V(x,y) = Gv(x,y)$; $u(x,y)$ and $v(x,y)$ are displacements along the *x*-axis (longitudinal) and along the *y*-axis (transverse) respectively; *G* is the shear modulus; ν is Poisson's ratio.

Assume that the elementary part of the solution is already known. Satisfying the boundary conditions of (2) on the ends of the rectangle, we come to the problem of determining the coefficients a_k from the expansions

$$\begin{aligned}
0 &= \sum_{k=1}^{\infty} B_k \xi(\lambda_k,y) + \overline{B_k}\,\xi\left(\overline{\lambda_k},y\right), \\
0 &= \sum_{k=1}^{\infty} A_k \chi(\lambda_k,y) + \overline{A_k}\,\chi\left(\overline{\lambda_k},y\right), \\
\sigma(y) &= \sum_{k=1}^{\infty} A_k s_x(\lambda_k,y)\cosh\lambda_k d + B_k s_x(\lambda_k,y)\sinh\lambda_k d + \\
&\quad + \overline{A_k} s_x\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k} d + \overline{B_k} s_x\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k} d, \\
0 &= \sum_{k=1}^{\infty} A_k t_{xy}(\lambda_k,y)\sinh\lambda_k d + B_k t_{xy}(\lambda_k,y)\cosh\lambda_k d + \\
&\quad + \overline{A_k} t_{xy}\left(\overline{\lambda_k},y\right)\sinh\overline{\lambda_k} d + \overline{B_k} t_{xy}\left(\overline{\lambda_k},y\right)\cosh\overline{\lambda_k} d.
\end{aligned} \tag{4}$$

Following the general scheme of solving a boundary value problem for a half-strip [1, 2], with the help of the functions $u_k(y)$, $v_k(y)$, $x_k(y)$, $t_k(y)$ biorthogonal to the Papkovich–Fadle eigenfunctions $\xi(\lambda_k,y)$, $\chi(\lambda_k,y)$, $s_x(\lambda_k,y)$, $s_y(\lambda_k,y)$, we obtain the system of algebraic equations for each $k \geq 1$:

$$\begin{aligned}
0 &= B_k\lambda_k M_k + \overline{B_k\lambda_k M_k}, \\
0 &= A_k M_k + \overline{A_k M_k}, \\
\sigma_k^* &= A_k M_k \cosh\lambda_k d + B_k M_k \sinh\lambda_k d + \\
&\quad + \overline{A_k M_k \cosh\lambda_k d} + \overline{B_k M_k \sinh\lambda_k d}, \\
0 &= A_k\lambda_k M_k \sinh\lambda_k d + B_k\lambda_k M_k \cosh\lambda_k d + \\
&\quad + \overline{A_k\lambda_k M_k \sinh\lambda_k d} + \overline{B_k\lambda_k M_k \cosh\lambda_k d}.
\end{aligned} \tag{5}$$

Solving (5), we find

$$A_k = \frac{\sigma_k^*}{M_k}\frac{\lambda_k\overline{\lambda_k}\left(\cosh\lambda_k d - \overline{\cosh\lambda_k d}\right)}{\Delta_k},$$
$$B_k = -\frac{\sigma_k^*\overline{\lambda_k}}{M_k}\frac{\left(\lambda_k\sinh\lambda_k d - \overline{\lambda_k\sinh\lambda_k d}\right)}{\Delta_k}, \tag{6}$$

where

$$\sigma_k^* = \sigma_k + \overline{\sigma_k},\ \ \sigma_k = \int_{-1}^{1}\sigma(y)x_k(y)dy,\ \ M_k = L'(\lambda_k)/2\lambda_k,$$
$$\Delta_k = \lambda_k\overline{\lambda_k}\left(\cosh\lambda_k d - \overline{\cosh\lambda_k d}\right)^2 - \left(\lambda_k\sinh\lambda_k d - \overline{\lambda_k\sinh\lambda_k d}\right)\left(\overline{\lambda_k}\sinh\lambda_k d - \lambda_k\overline{\sinh\lambda_k d}\right).$$

It is obvious that Δ_k is real.

Substituting the coefficients A_k and B_k in formulae for displacements and stresses and isolating null-series [1], we obtain

$$U(x,y) = \sum_{k=1}^{\infty}2\,\mathrm{Re}\left\{\sigma_k\frac{\xi(\lambda_k,y)}{\lambda_k M_k}\frac{\mathrm{Re}\{\lambda_k S(\lambda_k,x)\}}{\Delta_k}\right\},$$
$$V(x,y) = \sum_{k=1}^{\infty}2\,\mathrm{Re}\left\{\sigma_k\frac{\chi(\lambda_k,y)}{M_k}\frac{\mathrm{Re}\,C(\lambda_k,x)}{\Delta_k}\right\},$$
$$\sigma_x(x,y) = \sum_{k=1}^{\infty}2\,\mathrm{Re}\left\{\sigma_k\frac{s_x(\lambda_k,y)}{M_k}\frac{\mathrm{Re}\,C(\lambda_k,x)}{\Delta_k}\right\}, \tag{7}$$
$$\sigma_y(x,y) = \sum_{k=1}^{\infty}2\,\mathrm{Re}\left\{\sigma_k\frac{s_y(\lambda_k,y)}{\lambda_k^2 M_k}\frac{\mathrm{Re}\{\lambda_k^2 C(\lambda_k,x)\}}{\Delta_k}\right\},$$
$$\tau_{xy}(x,y) = \sum_{k=1}^{\infty}2\,\mathrm{Re}\left\{\sigma_k\frac{t_{xy}(\lambda_k,y)}{\lambda_k M_k}\frac{\mathrm{Re}\{\lambda_k S(\lambda_k,x)\}}{\Delta_k}\right\},$$

where

$$C(\lambda_k,x) = \lambda_k\overline{\lambda_k}\left(\cosh\lambda_k d - \overline{\cosh\lambda_k d}\right)\cosh\lambda_k x - \overline{\lambda_k}\left(\lambda_k\sinh\lambda_k d - \overline{\lambda_k\sinh\lambda_k d}\right)\sinh\lambda_k x,$$
$$S(\lambda_k,x) = \lambda_k\overline{\lambda_k}\left(\cosh\lambda_k d - \overline{\cosh\lambda_k d}\right)\sinh\lambda_k x - \overline{\lambda_k}\left(\lambda_k\sinh\lambda_k d - \overline{\lambda_k\sinh\lambda_k d}\right)\cosh\lambda_k x.$$

Examples of Solving a Boundary Value Problem

Example 1. Even-Symmetric Deformation

Let the self-equilibrated normal load act on the right end of the rectangle (Fig. 1):

$$\sigma(y) = \begin{cases} y^4 - \dfrac{6\alpha^2}{5}y^2 + \dfrac{\alpha^4}{5} & (|y| \le \alpha), \\ 0 & (|y| \le \alpha). \end{cases} \tag{8}$$

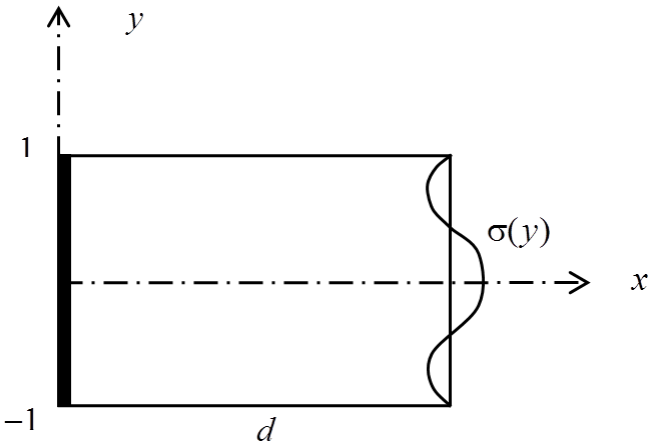


Figure 1. The scheme of the boundary value problem

The Papkovich–Fadle eigenfunctions $\xi(\lambda_k,y)$, $\chi(\lambda_k,y)$, $s_x(\lambda_k,y)$, $s_y(\lambda_k,y)$, $t_{xy}(\lambda_k,y)$ have the form

$$\begin{gathered}\xi(\lambda_k,y)=\left(\frac{1-\nu}{2}\sin\lambda_k-\frac{1+\nu}{2}\lambda_k\cos\lambda_k\right)\cos\lambda_k y-\frac{1+\nu}{2}\lambda_k y\sin\lambda_k\sin\lambda_k y,\\ \chi(\lambda_k,y)=\left(\frac{1+\nu}{2}\lambda_k\cos\lambda_k+\sin\lambda_k\right)\sin\lambda_k y-\frac{1+\nu}{2}\lambda_k y\sin\lambda_k\cos\lambda_k y,\\ s_x(\lambda_k,y)=(1+\nu)\lambda_k\left\{(\sin\lambda_k-\lambda_k\cos\lambda_k)\cos\lambda_k y-\lambda_k y\sin\lambda_k\sin\lambda_k y\right\},\\ s_y(\lambda_k,y)=(1+\nu)\lambda_k\left\{(\sin\lambda_k+\lambda_k\cos\lambda_k)\cos\lambda_k y+\lambda_k y\sin\lambda_k\sin\lambda_k y\right\},\\ t_{xy}(\lambda_k,y)=(1+\nu)\lambda_k^2\left\{\cos\lambda_k\sin\lambda_k y-y\sin\lambda_k\cos\lambda_k y\right\},\end{gathered}\tag{9}$$

where the numbers λ_k, $\overline{\lambda_k}$ $(\operatorname{Re}\lambda_k<0)$ form the set $\left\{\pm\lambda_k;\pm\overline{\lambda_k}\right\}_{k=1}^{\infty}=\Lambda$ of all the complex zeros of the entire function $L(\lambda)=\lambda(\lambda+\sin\lambda\cos\lambda)$.

The functions biorthogonal to the Papkovich–Fadle eigenfunctions $\xi(\lambda_k,y)$, $\chi(\lambda_k,y)$, $s_x(\lambda_k,y)$, $s_y(\lambda_k,y)$ have the form [3]

$$\begin{gathered}u_k(y)=\frac{1}{(1+\nu)}\left[\frac{\lambda_k\cos\lambda_k y}{\sin\lambda_k}-\left(\delta(y-1)+\delta(y+1)\right)\right],\\ v_k(y)=-\frac{\sin\lambda_k y}{(1+\nu)\sin\lambda_k},\ x_k(y)=\frac{\cos\lambda_k y}{2(1+\nu)\lambda_k\sin\lambda_k},\ t_k(y)=-\frac{\sin\lambda_k y}{2(1+\nu)\sin\lambda_k},\end{gathered}\tag{10}$$

where δ is the Dirac delta function.

The numbers M_k and σ_k are equal to

$$M_k=\cos^2\lambda_k,\ \sigma_k=\frac{8\left(\left(15-6\alpha^2\lambda_k^2\right)\sin\alpha\lambda_k+\left(\alpha^2\lambda_k^2-15\right)\alpha\lambda_k\cos\alpha\lambda_k\right)}{5(1+\nu)\lambda_k^6\sin\lambda_k}.$$

Substituting the found coefficients in formulae (7), we obtain the solution of the boundary value problem.

In Fig. 2 the distribution curves of the normal stresses $\sigma_x(1,y)$, $\sigma_x(0.9,y)$ and normal load $\sigma(y)$ are shown (it was assumed that $d=1$, $\alpha=0.5$, $\nu=\frac{1}{3}$).

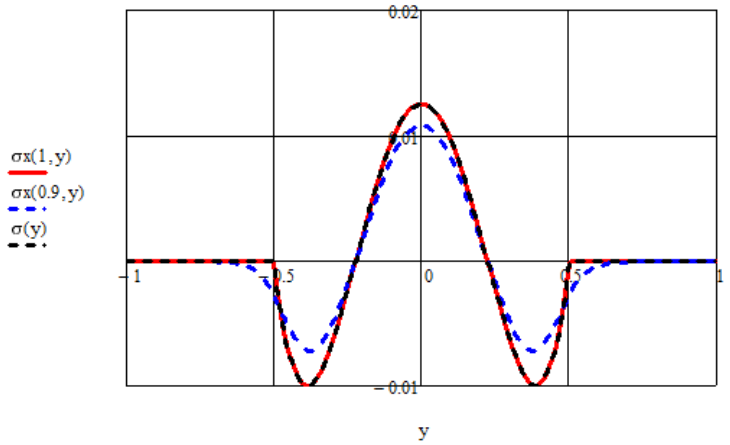


Figure 2. The distributions of the normal stresses $\sigma_x(1,y)$, $\sigma_x(0.9,y)$ and normal load $\sigma(y)$

Example 2. Odd-Symmetric deformation

Let the normal load that is self-equilibrated in moment act on the right end of the rectangle (Fig. 3):

$$\sigma(y)=y^5-\frac{10\alpha^3}{7}y^2+\frac{3}{7}y. \tag{11}$$

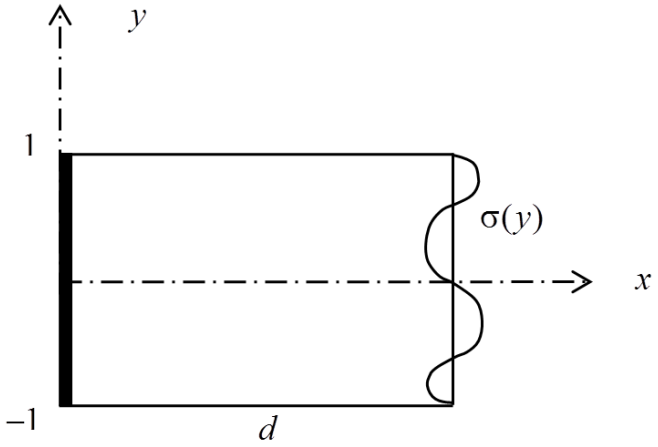


Figure 3. The scheme of the boundary value problem

In this case the Papkovich–Fadle eigenfunctions have the following form:

$$\begin{aligned}
\xi(\lambda_k,y)&=\left(\sin\lambda_k-\frac{1+\nu}{2}\lambda_k\cos\lambda_k\right)\sin\lambda_k y+\frac{1+\nu}{2}\lambda_k y\sin\lambda_k\cos\lambda_k y,\\
\chi(\lambda_k,y)&=-\left(\frac{1-\nu}{2}\sin\lambda_k+\frac{1+\nu}{2}\lambda_k\cos\lambda_k\right)\cos\lambda_k y-\frac{1+\nu}{2}\lambda_k y\sin\lambda_k\sin\lambda_k y,\\
s_x(\lambda_k,y)&=(1+\nu)\lambda_k\left\{(2\sin\lambda_k-\lambda_k\cos\lambda_k)\sin\lambda_k y+\lambda_k y\sin\lambda_k\cos\lambda_k y\right\},
\end{aligned} \tag{12}$$

$$s_y(\lambda_k, y) = (1+\nu)\lambda_k^2\{\cos\lambda_k \sin\lambda_k y - y\sin\lambda_k \cos\lambda_k y\},$$
$$t_{xy}(\lambda_k, y) = (1+\nu)\lambda_k\{(\sin\lambda_k - \lambda_k\cos\lambda_k)\cos\lambda_k y - \lambda_k y\sin\lambda_k \sin\lambda_k y\},$$

and the numbers λ_k, $\overline{\lambda_k}$ $(\mathrm{Re}\,\lambda_k < 0)$ form the set $\left\{\pm\lambda_k;\pm\overline{\lambda_k}\right\}_{k=1}^{\infty} = \Lambda$ of all the complex zeros of the entire function $L(\lambda) = \lambda - \sin\lambda\cos\lambda$.

The functions biorthogonal to the Papkovich–Fadle eigenfunctions $\xi(\lambda_k, y)$, $\chi(\lambda_k, y)$, $s_x(\lambda_k, y)$, $s_y(\lambda_k, y)$ have the form [4]

$$u_k(y) = \frac{1}{(1+\nu)}\frac{\sin\lambda_k y}{\sin\lambda_k},\ v_k(y) = \frac{1}{(1+\nu)}\frac{\cos\lambda_k y}{\lambda_k \sin\lambda_k},$$
$$x_k(y) = \frac{1}{2(1+\nu)\lambda_k^2}\left(\frac{\sin\lambda_k y}{\sin\lambda_k} - y\right),\ t_k(y) = \frac{1}{2(1+\nu)\lambda_k}\frac{\cos\lambda_k y}{\sin\lambda_k}. \tag{13}$$

The numbers M_k and σ_k are equal to

$$M_k = \frac{\sin^2\lambda_k}{\lambda_k},\ \sigma_k = \frac{8\left(\left(105 - 45\lambda_k^2 + \lambda_k^4\right)\sin\lambda_k + 5\left(2\lambda_k^2 - 21\right)\lambda_k\cos\lambda_k\right)}{7(1+\nu)\lambda_k^8 \sin\lambda_k}.$$

Substituting the found coefficients in formulae (7), we obtain the solution of the boundary value problem.

In Fig. 4 the distribution curves of the normal stresses $\sigma_x(1, y)$, $\sigma_x(0.9, y)$ and normal load $\sigma(y)$ are shown (it was assumed that $d = 1$, $\nu = \frac{1}{3}$).

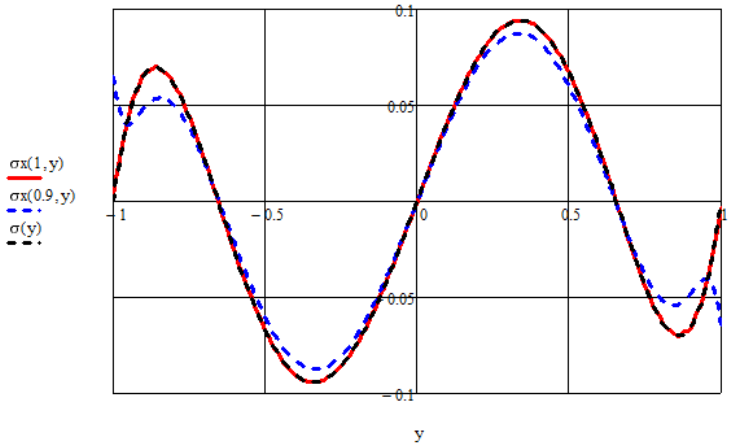


Figure 4. The distributions of the normal stresses $\sigma_x(1, y)$, $\sigma_x(0.9, y)$ and normal load $\sigma(y)$.

Conclusions

In both examples, on the end, the expanded functions $u = v = 0$, and we continue them by zero. Therefore, there will be no singularity at the angles if it is not introduced artificially by choosing a non-zero continuation outside the segment (end) [-1,1] in this or that way.

Acknowledgment

This research was supported by the Russian Foundation for Basic Research, research project no. 16-31-60028 mol_a_dk.

References

[1] Kovalenko, M. D. and Shulyakovskaya, T. D. (2011) Expansions in Fadle–Papkovich functions in a strip. Theory foundations, *Izv. Akad. Nauk. Mekh. Tverd. Tela* 5, 78–98 [*Mech. Solids* (Engl. Transl.) **46**(5), 721–738].

[2] Kovalenko, M. D., Menshova, I. V. and Shulyakovskaya, T. D. (2013) Expansions in Fadle–Papkovich functions: Examples of solutions in a half-strip, *Izv. Akad. Nauk. Mekh. Tverd. Tela* 5, 121–144 [*Mech. Solids* (Engl. Transl.) **48**(5), 584–602].

[3] Kovalenko, M. D. and Menshova, I. V. (2014) *Analytical Solutions of Two-Dimensional Boundary Value Problems of Elasticity Theory in Finite Domains with Angular Points of a Boundary*, Cheboksary: Chuvash. gos. ped. un-t [in Russian].

[4] Sebryakov, G. G., Kovalenko, M. D., Menshova I. V. and Semenova, I. A. (2015) An odd-symmetric boundary-value problem of elasticity theory for a semi-strip: Exact solution, *Dokl. Akad. Nauk* **462**(6), 662–665 [*Dokl. Phys.* (Engl. Transl.) **60**(6), 274–277].

Higher order extension of PDS-FEM and simulating brittle cracks

***M.L.L. Wijerathne[1], M.K. Pal[2], and M. Hori[1]**

[1]Earthquake Research Institute, The University of Tokyo, Japan.
[2]National Research Institute for Earth Science and Disaster Resilience, Miki city, Hyogo, Japan

Presenting and corresponding author: *lalith@eri.u-tokyo.ac.jp

Abstract

Higher order extension of Particle Discretization Scheme (PDS) and its implementation in FEM framework (HO-PDS-FEM) are presented in this short paper. PDS uses the conjugate tessellation pair Voronoi and Delaunay to approximate functions and their derivatives, respectively. In HO-PDS, a function and its derivatives are approximated as the union of the local polynomial expansions. The support of the base polynomial functions being confined to the domain of each tessellation element, the PDS approximations of function and the derivatives are inherently discontinuous along the boundaries of tessellation elements. Higher order PDS-FEM utilizes these discontinuities in function approximation to model discontinuities like cracks numerically efficiently. Higher order PDS is implemented in FEM framework to solve boundary value problem of elastic solids with mode-I cracks. The verification tests show that the higher order PDS-FEM has higher accuracy and convergence rate, compared to the original 0^{th}-orderPDS-FEM[1] proposed by Hori *et al.* Several benchmark problems are presented to demonstrate the improvement in accuracy. J-integral about a mode-I crack tip field is estimated to demonstrate the improvement in accuracy of crack tip stress fields. It is shown that the singular crack tip stress field also has higher order accuracy and convergence rates, in addition to improved crack surface traction.

Keywords: Particle Discretization Scheme, higher order extension, brittle cracks, J-integral

Introduction

Real materials are far from ideal and contain numerous microscopic cracks, flaws, etc. While these have negligible effect on the ordinary deformations of materials, movements of crack tips, which are moving stress singularities, are very sensitive to these minor heterogeneities and make the crack surfaces to bend, kink and branch. Due to this high sensitivity to minor heterogeneities, even nearly identical and homogeneous samples under same loading conditions do not produce the identical crack pattern. Thus, what is required for practical applications is probability density distribution of possible crack paths, instead of theoretical crack configuration under ideal condition. Generation of probability density distribution with Monte-Carlo simulations requires efficient numerical technique to model propagating cracks.

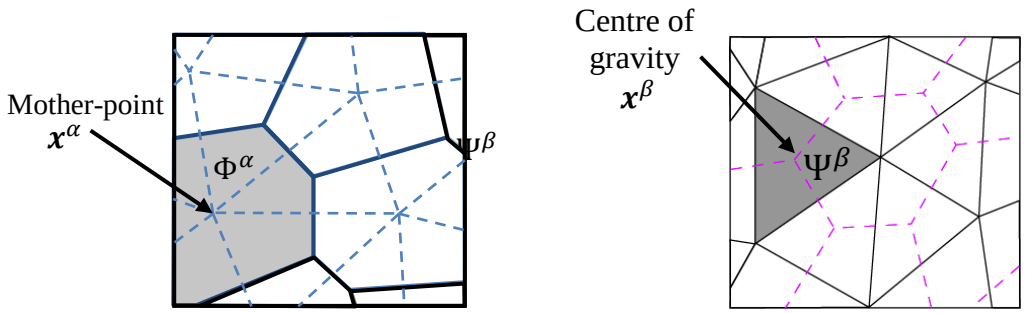
There exists a number of numerical methods, with their own different advantages, for simulating crack propagation. Most of these methods either belong to the family of particle methods or FEM. Recent enhancements of FEM [2] enable accurate modeling of theoretically predicted crack paths. However, most FEM based methods involve significant numerical overhead (e.g. tracking crack front, especial treatment for crack branching; introducing new degrees of freedoms, etc.). Analysis of large deformation and subsequently simulating the complex cracks are easily handled in particle methods [3, 4]. Although particle methods have low computational overhead, those lacks the mathematical rigorousness. Hori *et al.*[1, 5] proposed PDS-FEM as a numerical technique which combine mathematical rigorousness of FEM and simple crack treatment of particle methods. The crack treatment of PDS is fairly simple and involves negligible numerical overhead, making it ideal for probabilistic studies of crack path variability.

The original proposal of PDS-FEM [1, 5], which we refer in this paper as 0^{th}-order PDS-FEM, has only first order accuracy. PDS-FEM is based on Particle Discretization Scheme (PDS), which utilizes the characteristic functions of conjugate domain tessellations to approximate function and its derivatives. This paper presents an overview of higher order extension of PDS and PDS-FEM [6, 7], numerical treatments for modeling cracks and verification, and other improvements.

This short paper consists of five sections. Section two and three provide brief descriptions of higher order PDS and its implementation in FEM framework (PDS-FEM). The latter part of third section presents PDS-FEM's efficient treatment to model cracks. Some numerical results and discussions are included in the fourth section, while the fifth section presents some concluding remarks.

Higher order PDS

A unique feature of PDS is that it uses conjugate domain tessellations for approximation of functions and its derivatives. Though, any pair of tessellation could be utilized, authors have followed the former work and used Voronoi and Delaunay tessellations to approximate function and its derivatives, respectively.



(a) Voronoi (b) Delaunay

Figure 1 Voronoi and Delaunay tessellation in 2D

Assume $f(x)$ to be a target function in a given domain S; its Voronoi and Delaunay tessellations are denoted by $\{\Phi^\alpha\}$ & $\{\Psi^\beta\}$; the set of Voronoi mother points is $\{\boldsymbol{x}^\alpha\}$ and the set of center of gravity of Delaunay tessellation is $\{\boldsymbol{x}^\beta\}$ (see Fig. 1). φ^α and ψ^β are the characteristics function of Φ^α and Ψ^β, respectively. Higher order PDS approximates $f(x)$ and its derivatives $\nabla f(x)$ as

$$f(x) \approx f^d(\boldsymbol{x}) = \sum_{\alpha}^{N^\alpha} \sum_{n}^{|P^\alpha|} f^{\alpha n} P^{\alpha n} \tag{1}$$

$$\nabla f(x) \approx \boldsymbol{g}^d(\boldsymbol{x}) = \sum_{\beta}^{N^\beta} \sum_{n}^{|Q^\beta|} \boldsymbol{g}^{\beta n} Q^{\beta n}$$

where $P^{\alpha n}$ and $Q^{\beta n}$ are sets of base functions. Though any suitable set of functions for modeling the problem under consideration can be used, inspired by the Taylor series we use polynomial base functions such as

$$P^{\alpha n} \in P^\alpha = \{1, (\boldsymbol{x} - \boldsymbol{x}^\alpha), \dots, (\boldsymbol{x} - \boldsymbol{x}^\alpha)^r, \dots\}\phi^\alpha(\boldsymbol{x}),$$
$$Q^{\beta m} \in P^\beta = \{1, (\boldsymbol{x} - \boldsymbol{x}^\beta), \dots, (\boldsymbol{x} - \boldsymbol{x}^\beta)^r, \dots\}\psi^\beta(\boldsymbol{x}).$$

Here, N^{α} and N^{β} are the total number of Voronoi and Delaunay tessellation elements, respectively. $|P^{\alpha}|$ and $|Q^{\beta}|$ denote the number of base functions in each of the sets.

The unknown coefficients of approximations $f^{\alpha n}$ and $\boldsymbol{g}^{\beta n}$ can be found by minimizing the errors where $E^{f} = \int (f - f^{d})\mathrm{d}s$ and $E^{g} = \int (\boldsymbol{\nabla} f - \boldsymbol{g}^{d})\mathrm{d}s$. Minimization of these errors leads to the solving of the following linear system of equations.

$$\sum_{n=0}^{|P^{\alpha}|} f^{\alpha n} I^{\alpha m n} = \int f(\boldsymbol{x}) P^{\alpha m}\mathrm{d}s \tag{2}$$

$$\sum_{n=0}^{|P^{\alpha}|} g_{i}^{\alpha n} I^{\beta m n} = \sum_{\alpha,l}^{N^{\alpha}} f^{\alpha l} \int \left(P^{\alpha l}(\boldsymbol{x})\right)_{,i} Q^{\beta m}\mathrm{d}s \tag{3}$$

Here, $I^{\alpha m n} = \int P^{\alpha n} P^{\alpha m}\mathrm{d}s$ and $I^{\beta m n} = \int Q^{\beta n} Q^{\beta m}\mathrm{d}s$. While there is no restriction that only polynomial bases should be included in P^{α} and Q^{β}, inspired by Taylor expansion, we prefer to include polynomial bases. However, it is best to use suitable set of base functions, according to the nature of the problem.

Multiplying with the characteristic functions of each tessellation elements, the support of base functions are confined to the domain of each tessellation element. Hence the function and derivative approximations of HO-PDS have numerous discontinuities along the boundaries of respective tessellation elements.

Implementation of higher order PDS in FEM framework

The use of higher order PDS to approximate the field variables and their derivatives in FEM framework is referred as higher order PDS-FEM (HO-PDS-FEM). Consider a boundary value problem (BVP) with infinitesimal deformation of a linear elastic domain. Body forces are ignored for the sake of brevity. The standard Lagrange for linear elasticity BVP is stated as follows

$$L[\boldsymbol{\varepsilon}(\boldsymbol{u})] = \frac{1}{2}\int \boldsymbol{\varepsilon} : \boldsymbol{c} : \boldsymbol{\varepsilon}\ \mathrm{d}s \tag{4}$$

Here, $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ is the strain tensor and $\boldsymbol{c}$ is fourth order elasticity tensor defining linear stress-strain relation. Setting the first variation $\delta L = 0$, we can obtain the strong form of the governing equations and essential boundary conditions.

Following the HO-PDS the unknown displacement $\boldsymbol{u}$ is approximated as

$$u_{i}(\boldsymbol{x}) \approx u_{i}^{d}(\boldsymbol{x}) = \sum_{\alpha}^{N^{\alpha}} \sum_{n}^{|P^{\alpha}|} u_{i}^{\alpha n} P^{\alpha n}. \tag{5}$$

Further, following the definition of derivative approximation of HO-PDS, the derivatives of displacement are approximated as $u_{i,j} \approx \sum_{\beta,m} u_{ij}^{\beta m} Q^{\beta m}$. Based on Eq. 2, $u_{ij}^{\beta m}$ can be expressed as

$$u_{ij}^{\beta m}(\boldsymbol{x}) = \sum_{m'}^{|Q^{\beta}|} w^{\beta m m'} \sum_{\alpha}^{N^{\alpha}} \sum_{n}^{|P^{\alpha}|} u_{i}^{\alpha n} \int_{\Psi^{\beta}} Q^{\beta m'} (P^{\alpha n})_{,j}\ \mathrm{d}s$$

$$= \sum_{m'}^{|Q^\beta|} w^{\beta m m'} \sum_{\alpha}^{N^\alpha} \sum_{n}^{|P^\alpha|} u_i^{\alpha n} h_j^{\beta \alpha m' n} \tag{6}$$

where $\left[w^{\beta mn}\right]^{-1} = \left[I^{\beta mn}\right] = \left[\int_{\Psi^\beta} Q^{\beta n} Q^{\beta m} \mathrm{d}v\right]$. Now the strain field can be approximated as

$$\varepsilon_{ij}(\boldsymbol{x}) \approx \sum_{\beta}^{N^\beta} \sum_{n}^{|Q^\beta|} \varepsilon_{ij}^{\beta n} Q^{\beta n},$$

where $\varepsilon_{ij}^{\beta n}$ can be expressed using Eq. 6 as

$$\varepsilon_{ij}^{\beta n} = \sum_{m'}^{|Q^\beta|} w^{\beta m m'} \sum_{\alpha}^{N^\alpha} \sum_{n}^{|P^\alpha|} \frac{1}{2} (h_j^{\beta \alpha m' n} u_i^{\alpha n} + h_i^{\beta \alpha m' n} u_j^{\alpha n}). \tag{7}$$

For the sake of brevity, let's express Eq. 7 in tensor form as

$$\boldsymbol{\varepsilon}^{\beta m} = symm\left(\boldsymbol{B}^{\beta m \alpha n} \otimes \boldsymbol{u}^{\alpha n}\right), \tag{8}$$

where $B_i^{\beta m \alpha n} = \sum_{m'}^{|Q^\beta|} w^{\beta m m'} h_i^{\beta \alpha m' n}$.

Similarly stress tensor $\boldsymbol{\sigma}$ can be approximated as $\sigma_{ij}(\boldsymbol{x}) \approx \sum_{\beta}^{N^\beta} \sum_{n}^{|Q^\beta|} \sigma_{ij}^{\beta n} Q^{\beta n}$. The elasticity tensor $\boldsymbol{c}$ is also approximated using the characteristic functions of Delaunay tessellation as $c_{ijkl}(\boldsymbol{x}) \approx \sum_{\beta}^{N^\beta} c_{ijkl}^{\beta} \psi^{\beta}(\boldsymbol{x})$. It is straight forward to obtain $\sigma_{ij}^{\beta n} = c_{ijkl}^{\beta} \varepsilon_{kl}^{\beta n}$.

Substituting the Eq. 8 into L in Eq. 4 and setting its first variation to zero (i.e. $\delta L = 0$), the following governing matrix equation of HO-PDS-FEM can be obtained.

$$\sum_{\alpha', n, n', m'} \boldsymbol{w}^{\beta n n'} \cdot \left(\boldsymbol{B}^{\beta n \alpha m} \cdot \boldsymbol{c}^{\beta} \cdot \boldsymbol{B}^{\beta n \alpha' m'}\right) \cdot \boldsymbol{u}^{\alpha' m'} = \boldsymbol{0} \tag{9}$$

According to the above equation, the element stiffness matrix of HO-PDS-FEM is

$$\boldsymbol{K}^{\boldsymbol{\beta}} = \boldsymbol{w}^{\beta} \cdot \boldsymbol{B}^{\beta n \alpha m} \cdot \boldsymbol{c}^{\beta} \cdot \boldsymbol{B}^{\beta n \alpha' m'} \tag{10}$$

Size of this element stiffness matrix depends on the space dimensions and number of basis functions used in analyzing the target problem. For example, the size is $(6 \times |P^\alpha|) \times (6 \times |P^\alpha|)$ for a 2D Delaunay triangle, and $(12 \times |P^\alpha|) \times (12 \times |P^\alpha|)$ for 3D Delaunay tetrahedral.

Modeling brittle crack in HO-PDS-FEM

Major advantage of PDS-FEM is its simple and efficient treatment for modeling propagating discontinuities like cracks. The displacement field approximation $u_i^d(\boldsymbol{x})$ is inherent with discontinuities along each boundary of Voronoi elements Φ^α , $\partial\Phi^\alpha$, as a consequence of

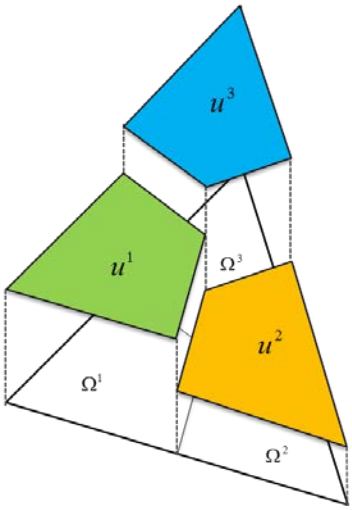
limiting the support of each polynomial base $P^{\alpha n}$ to the domain of the corresponding tessellation element Φ^{α}. Figure 2(a) shows an exaggerated illustration of a displacement component approximated with PDS over a Delaunay tessellation. As explained in this sub-section, HO-PDS-FEM utilizes these existing discontinuities along $\partial\Phi^{\alpha}$'s to numerically efficiently model moving discontinuities in BVPs.

The contribution to the strain from the above mentioned discontinuities can be isolated by expressing base functions with compact support within each Voronoi as $P^{\alpha n} = F^{\alpha n}(\boldsymbol{x})\,\phi^{\alpha}(\boldsymbol{x})$ and substituting to $h_j^{\beta\alpha m' n}$ in Eq. 6.

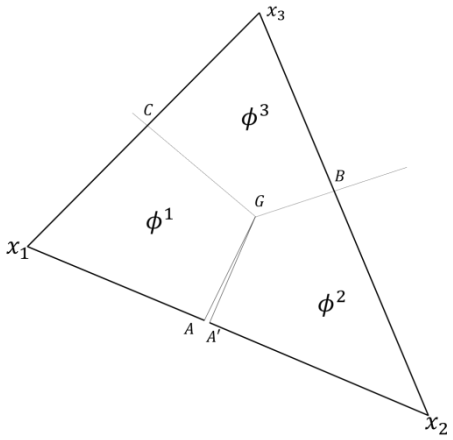
$$\begin{aligned} h_j^{\beta\alpha m' n} &= \int_{\Psi\beta} Q^{\beta m'} (P^{\alpha n})_{,j}\, \mathrm{d}s \\ &= \int_{\Psi\beta} Q^{\beta m'} \left((F^{\alpha n})_{,j}\, \phi^{\alpha} + F^{\alpha n} \phi^{,\alpha}_{,j} \right) \mathrm{d}s \\ &= \int_{\Psi\beta} Q^{\beta m'} (F^{\alpha n})_{,j}\, \phi^{\alpha}\, \mathrm{d}s + \int_{\partial\Phi^{\alpha}} Q^{\beta m'} F^{\alpha n} n_j\, \mathrm{d}l \end{aligned} \tag{11}$$

In the above equation, the surface integration $\int_{\partial\Phi^{\alpha}} Q^{\beta m'} F^{\alpha n} n_j\, \mathrm{d}l$ carries the contribution to strain ε_{ij} from the above mentioned discontinuities along boundaries $\partial\Phi^{\alpha}$; note that we have used the Gauss divergence theorem. Eliminating this contribution is equivalent to introducing a discontinuity to the physical problem by removing the contribution from an infinitesimally thin neighborhood along $\partial\Phi^{\alpha}$.

As an example, an opening crack AGA' as shown in Fig. 2(b), along the common boundary of Voronoi elements Φ^1 and Φ^2 in Fig. 2(a), can be modelled by simply dropping the contributions $\int_{AG} Q^{\beta m'} F^{\alpha n} n_j\, \mathrm{d}l$ and $\int_{GA'} Q^{\beta m'} F^{\alpha n} n_j\, \mathrm{d}l$ while evaluating element stiffness matrix $\boldsymbol{K}^{\beta}$ of the Delaunay element encompassing the crack surface.



(a) Discontinuities in the approximated displacement field obtained with PDS-FEM



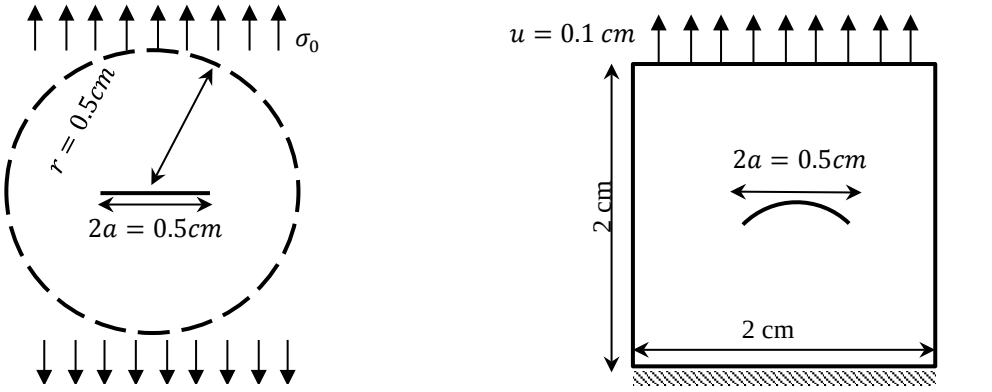
(b) Mode-I crack AGA' is modeled by dropping contribution from a thin neighborhood of boundary between Φ^1 and Φ^2.

Figure 2. Modeling a mode-I crack

Most of the existing numerical tools require introduction of additional nodes, enrichment functions, etc., subsequently adding substantial numerical overhead and/or complex process. On the other hand, higher order PDS-FEM only requires only re-calculation of an element stiffness matrix eliminating the contributions $\int_{\partial\Phi^\alpha} Q^{\beta m'} F^{\alpha n} n_j \, \mathrm{d}l$ along the required Voronoi boundaries $\partial\Phi^\alpha$. This very low numerical overhead in modeling cracks is a notable feature which makes PDS-FEM one of the most numerically efficient numerical treatment for modeling cracks. This is especially useful in simulation of 3D crack propagation in large scale models.

Numerical examples

As explained above, a function or vector field approximated with PDS consists of numerous discontinuities along each Voronoi boundary. The use of such discontinuous approximations in solving BVP is rare and one may doubt about the quality of the solution. In this section some numerical examples are presented to demonstrate that PDS-FEM is accurate and higher order versions have the expected higher accuracy and convergence rates. A major advantage of PDS-FEM being the numerical efficient crack treatment, majority of the examples given in this sections are to demonstrate the accuracy stationary crack modelled with HO-PDS-FEM.



(a) Centre crack in an infinite domain (b) Square plate with a circular crack

Figure 3 Considered numerical examples with stationary model-I cracks

Problem setting

A classical mode-I finite crack in an infinite domain (see Fig. 3(a)) is chosen as a numerical example to verify crack tip singularity modelled with HO-PDS-FEM. An arch shape crack shown in Fig. 3(b) is analyzed to demonstrate that it can model nearly traction free crack surfaces. For all the problems provided in this section, Young's modulus of 1GPa and Poison's ratio 0f 0.33 are assumed.

Results with two different pairs of base functions sets are compared to demonstrate the improvement in accuracy. The first case is with lowest order base functions $P^\alpha = \{1\}$ and $Q^\beta = \{1\}$, which is referred as 0th-order PDS-FEM. The other case is with the polynomial bases $P^\alpha = \{1, (x - x^\alpha), (y - y^\alpha)\}$ and $Q^\beta = \{1, \ (x - x^\beta), \ (y - y^\beta), \ (x - x^\beta)^2, \ (y - y^\beta)^2, (x - x^\beta)(y - y^\beta)\}$, which is referred as 1st-order PDS-FEM.

Results and discussion

Figure 4 shows the stress component along the right half of the crack surface. As seen, in the neighborhood of crack tip, HO-PDS-FEM's solution has a large deviation from the analytical

solution. However, σ_{yy} is in good agreement with analytic solution elsewhere. Moreover, higher order PDS-FEM reproduces the traction free crack-surface, which is a significant improvement compared to 0[th]-order PDS-FEM. Analytical solution of this boundary value problem is of nature of $\sqrt{r}Sin(\theta/2)$ and cannot be accurately approximated as a linear combination of polynomials. Hence, this significant disagreement in the crack tip neighborhood is not unexpected.

The best way to eliminate this large deviation in the crack tip neighborhood is to utilize the analytic solution for crack tip stress field as the basis functions of higher order PDS-FEM. Another less precise technique is to adjust the point of inflection of polynomial bases in Q^{β} only for the Delaunay elements encompassing crack tips. It is found that choosing the mid-point of Delaunay edge, through which crack enters (see Fig. 5), as the point of inflection of polynomial bases in Q^{β} improves crack tip stress field.

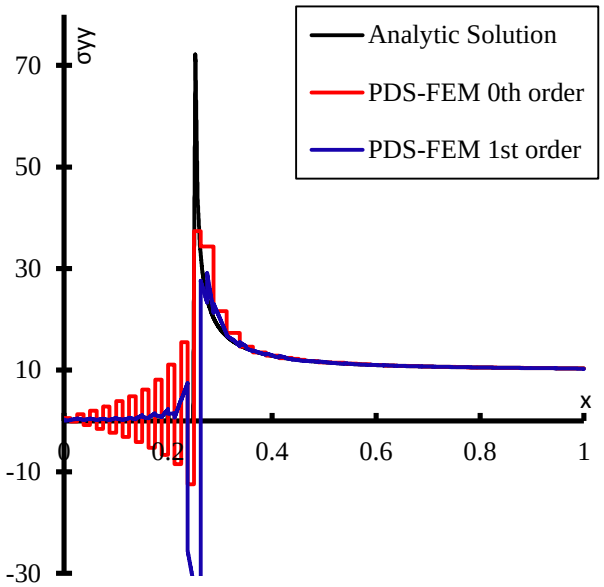


Figure 4 σ_{yy} along the right half of crack line

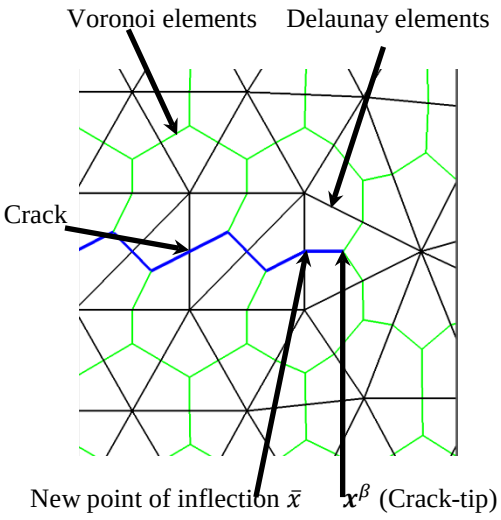


Figure 5 Point of inflexion

Figure 6 compares the results of 0[th]-order and 1[st]-order PDS-FEM with analytic solution, when point of inflection is moved to the entry point of the crack. As is seen, the results are in good agreements with the analytic solution. Also, the crack surface remains nearly traction free, which is noteworthy improvement compared to 0[th]-order (see Fig. 6). Further, Fig. 7 shows the J-integral with different number of degrees of freedoms. As is seen, both the accuracy and rate of convergence have improved with 1[st]-order PDS-FEM.

Although the accuracy of crack tip stress filed can be improved using analytical solutions of crack tip stress field, the above presented less precise approach by moving the point of inflexion is attractive in large scale simulations since it does not increase the numerical overhead. Though the use of analytic solution of crack tip stress field as the basis functions improves the accuracy, it increases the numerical overhead, leading to load imbalance in parallel computing and lower scalability. On the contrary, moving of point of inflexion does not involve any additional numerical overhead.

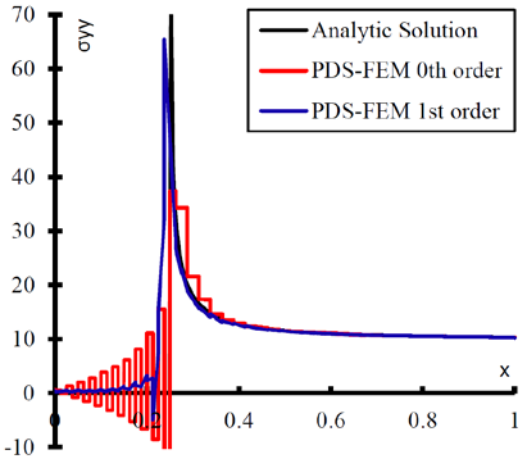


Figure 6 σ_{yy} along the right half of crack

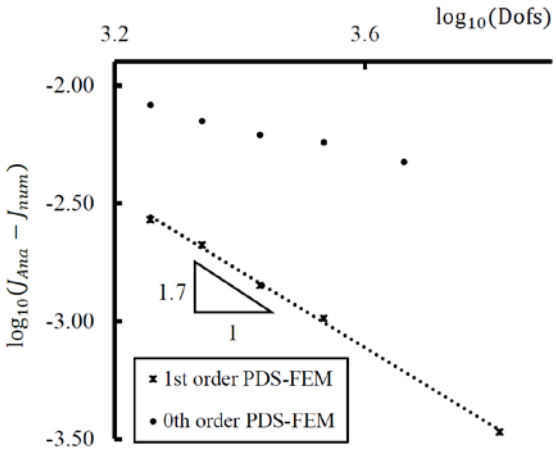
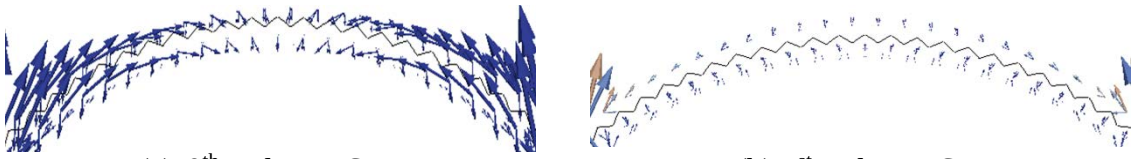


Figure 7 Convergence rate of J-integral

Figure 8 compares the tractions, obtained with 0^{th} -order and 1^{st}-order PDS-FEM, along the curved crack surfaces in the rectangular plate. It can be clearly seen that the 1^{st}-order PDS-FEM can reproduce nearly traction free crack surfaces, and it is a significant improvement compares to 0^{th} -order.



(a) 0^{th}-order PDS-FEM (b) 1^{st}-order PDS-FEM

Figure 8 Traction normal to crack surfaces

Unlike most other FEM based crack treatments, PDS-FEM does not emphasizes modeling the crack tip and crack surface precisely. Instead PDS-FEM tries model crack tip and surface to a sufficient degree of accuracy for practical problems focusing on lower numerical overhead so that large scale 3D crack propagation problems can be solved efficiently. However, as shown above the crack tip and surfaces modeled with HO-PDS-FEM have a fairly high accuracy and convergence rates. The crack surfaces modeled with PDS-FEM.

Uniformly pressured thick hollow cylinder

A thick cylinder subjected to internal and external pressure is considered to verify the 3D implementation of HO-PDS-FEM. Figure 8 illustrates the problem settings. The boundary conditions over the top and bottom surfaces are set to reproduce plain strain conditions. In this 3D problem, the set of polynomial bases used are $P^{\alpha} = \{1, (x - x^{\alpha}), (y - y^{\alpha}), (z - z^{\alpha})\}$ and $Q^{\beta} = \{1, (x - x^{\beta}), (y - y^{\beta}), (x - x^{\beta})^2, (y - y^{\beta})^2, (z - z^{\beta})^2, (x - x^{\beta})(y - y^{\beta}), (x - x^{\beta})(z - z^{\beta}), (y - y^{\beta})(z - z^{\beta})\}$.

Figure 10 compares the analytic solutions and numerical results of radial displacement u_r and strain component ε_{rr} along a radial line. A quick comparison advocates the improvement in solution with the mesh refinement, and that the numerical solutions are in good agreement with analytic solution. For this specific setting, u_r reaches its maximum at $r = 0.2$. Figure 11 shows

the error of u_r at $r = 0.2$ for several tessellations with different element sizes. As is seen, the error diminishes at second order rate with respect to the number of degrees of freedoms, which is the expected.

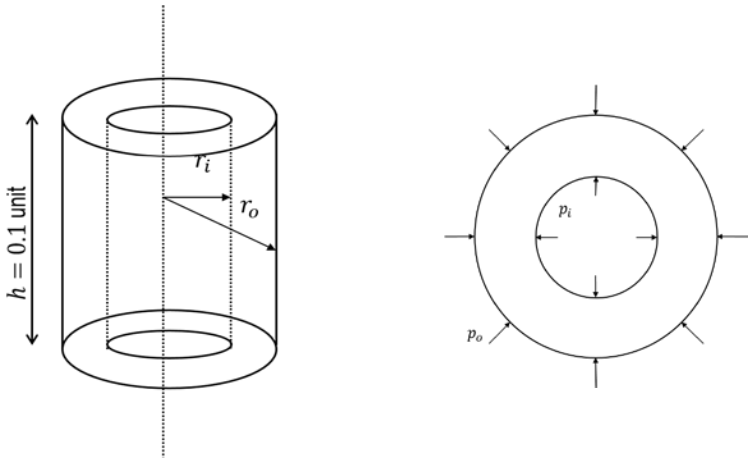


Figure 9 Thick hollow cylinder applied with uniform internal and external pressure

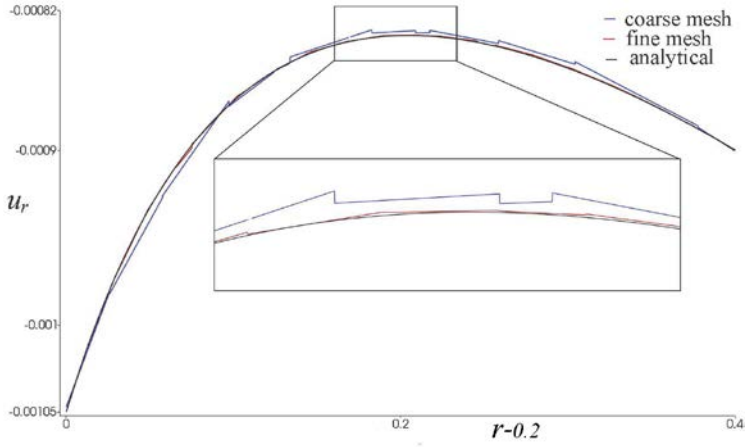


Figure 10 Comparison of displacement along radial direction u_r

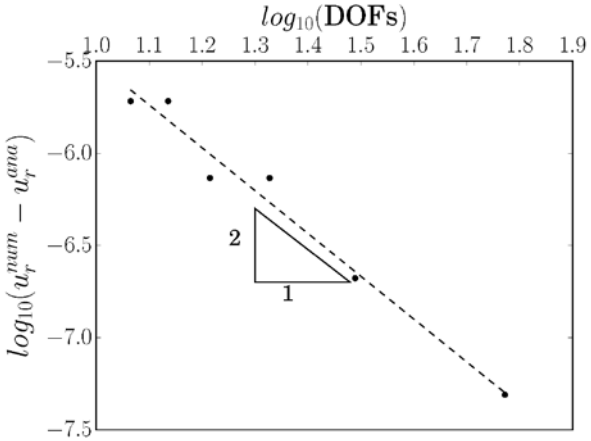


Figure 11 Convergence rate of the displacement along radial direction

Concluding remarks

Implementation of higher order PDS-FEM and its application to simulated brittle stationary cracks are presented. With numerical example, it is demonstrated that HO-PDS-FEM provides higher accuracy and theoretically expected convergence rates. Two major advantages of HO-PDS-FEM over former 0^{th}-order PDS-FEM are the improvement in the accuracy of crack tip stress field and significant reduction in traction along the model-I crack surfaces. Like in the former 0^{th}-order implementation, the numerical treatment for modeling cracks with HO-PDS-FEM also numerically efficient to simulate crack propagation in large scale 3D models.

References

[1] Hori, M., Oguni, K., Sakaguchi, H., (2005) Proposal of FEM implemented with particle discretization scheme for analysis of failure phenomena, *J. of Mech. and Phys. of Solids*, **53** 681–703.

[2] Moes, N., Dolbow, J., Belytschko, T., (1999) A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, **46**(1) 131-150.

[3] Gingold, R.A., Monaghan, J.J., (1977) Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Monthly Notices of the Royal Astronomical Society*, **181** 375–389.

[4] Schlangen E., Garboczi E.J., (1997) Fracture simulations of concrete using lattice models: computational aspects. *Engrg. Frac. Mechanics*, **57**(2) 319–332.

[5] Wijerathne, M.L.L. , Oguni, K., Hori, M. (2009) Numerical analysis of growing crack problem using particle discretization scheme, *Int. J. for Numerical Methods in Engineering*, **80** 46-73.

[6] Pal, M.K., Wijerathne, L,, Hori, M., Ichimura, T., Tanaka, S. (2014) Implementation of Finite Element Method with higher order Particle Discretization Scheme, *J. of Japan Society of Civil Engineers, Ser.A2* **70(2)** 297-305.

[7] Pal, M.K., Wijerathne, L, Hori, M., Ichimura, T. (2015), Simulation of cracks in linear elastic solids using higher order Particle Discretization Scheme-FEM, *J. of Japan Society of Civil Engineers, Ser.A2* **71(2)** 327-337

Young's modulus determination of the collagen molecule via steered molecular dynamics simulations

***H.B. Ge, †C.Q. Zhang, Y Song, Q Liu**

Tianjin Key Laboratory of the Design and Intelligent Control of the Advanced Mechatronical System, School of Mechanical Engineering, Tianjin University of Technology, Tianjin 300384, China

*Presenting author: ghb_zw@126.com

†Corresponding author: zhang_chunqiu@126.com

Abstract

Articular cartilage is a special kind of connective tissue and unique zonal architecture. In the different layers of cartilage, the functions of collagen fiber cannot be ignored. The mechanical properties of the cartilage directly related to its microstructure, and depended on the function of collagen fibers to some extent. The purpose of this paper is to use GROMACS 2016 and steered molecular dynamics (SMD) simulation method to directly measure the Young's modulus of collagen molecule. And under the condition of stretching rate of v=0.01nm/ps, we obtained the Young's modulus of collagen molecule is 6.1±0.3 GPa.

Keywords: Articular cartilage, Collagen molecule, GROMACS 2016, Steered Molecular Dynamics

Introduction

The primary role of articular cartilage is to provide a low friction joint surface that resists wear, and distributing stresses in a demanding joint environment [1]. Articular cartilage is a special kind of connective tissue and have the unique zonal architecture (Figure 1).

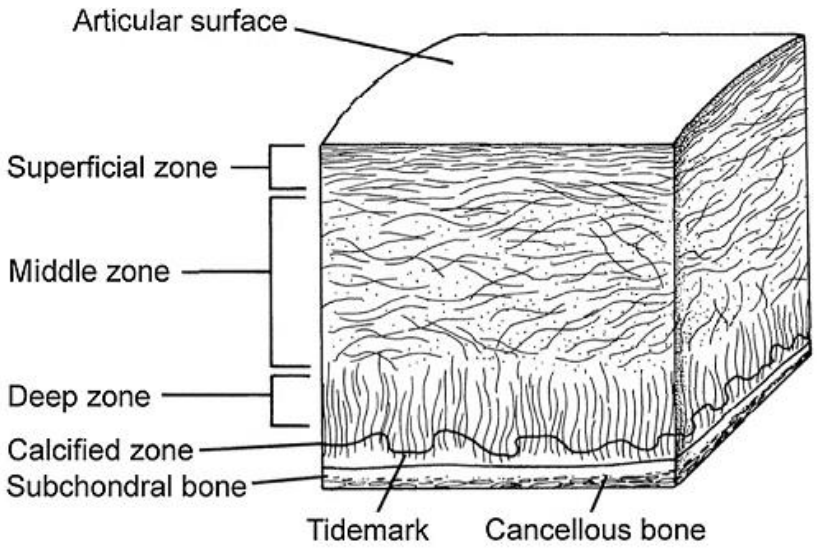


Figure 1. Schematic of the architecture of collagen fibers of human adult articular cartilage (Republished with permission from J. Am. Acad. Orthop. Surg. 2(4):193, 1994.)

The articular cartilage function is dependent on the molecular composition of the extracellular matrix—the interactions between interstitial fluid, the proteoglycan and Type II collagen [2]. Articular cartilage is composed of chondrocytes and extracellular matrix (ECM). Collagen is the most primary component of the extracellular matrix, which is approximately 60%-70% of dry weight of adult articular cartilage. Type II collagen, in articular cartilage, is the principal constituents of collagen.

The functions of collagen fiber cannot be ignored, in the different layers of cartilage [3]. The superficial zone provides a smooth, gliding surface with minimal friction [4], and the collagen fibers, which constrains the tensile stress, are arranged parallel to the joint surface. In middle zone, collagen fibers are mutually crosswise arranged, which plays a primary role in resisting the shear stress and compression stress. In deep zone, collagen fibers are perpendicular to the articular surface, which contributes to resisting the compression stress.

The mechanical properties of the cartilage directly related to its microstructure, and depend on the function of collagen fibers to some extent. In order to research the mechanical properties of the collagen molecule, Harley [5] and co-workers measured collagen Young's modulus to be 9.0 GPa, in 1977; Cusack [6] and Miller estimated 5.1GPa, in 1979; Sasaki [7] and Odajima obtained 2.9 GPa, in 1996; Sun [8] and co-workers reported the Young's modulus between 0.35 and 12.2 GPa, in 2002. Although the mechanical properties of collagen molecules can be determined by the experimental techniques, the results are largely dependent on the biological samples, and the nanoscale phenomena cannot be observed during the experiment. In order to overcome these disadvantages, molecular dynamics method is used to research the collagen molecules. Using molecular dynamics method, Lorenzo [9] and Caffarena obtained the Young's modulus of 4.8 GPa, in 2005; Buehler [10] assessed the Young's modulus ranging from 6.99 to 18.82 GPa, in 2006; Alfonso [11] assessed 4.6 GPa, in 2008; Andrzej [12] obtained 7.4 GPa, in 2015.

Due to the continuous improvement of the simulation software and the parameters of the biomolecular force field, the simulation results of the mechanical properties of collagen may be different. The purpose of this paper is to use GROMACS 2016 to directly research the mechanical properties of collagen molecule when submitted to the virtual traction along its principal axis with steered molecular dynamics (SMD) simulation method. The Young's modulus of collagen molecule was obtained, and compared with previous results.

Method

In order to study the mechanical properties of the collagen molecule, the SMD method is used to carry out the stretching along the principal axis. Due to the length of type II collagen molecules is too long (~300nm), it is time-consuming that use this structure to carry out full-atomistic simulation. Therefore, the collagen-like peptide: $(Pro\text{-}Hyp\text{-}Gly)_4$-Glu-Lys-Gly-$(Pro\text{-}Hyp\text{-}Gly)_5$, which widely used as a model for collagen, substituted for type II collagen.

The GROMACS 2016 simulation package is used to perform the SMD simulation, which GROMOS96 54a7 Force Field [13] and the extended single point charge (SPC/E) [14] water

model for solvent were considered. For the solvation, a 24nm×3.2nm×3.2nm cubic box containing 6719 water molecules, 20 $Na^+$and 20 Cl^- ions are implemented (Figure 2). This solution is equivalent to the normal saline.

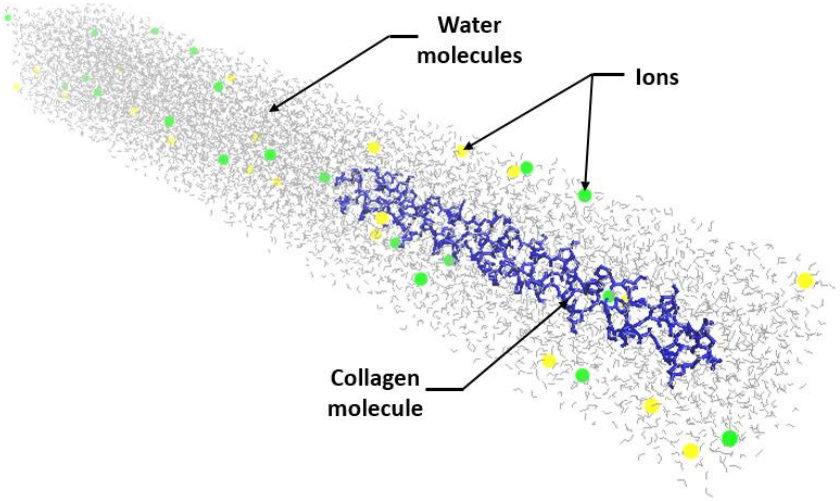


Figure 2. Schematic representation of the model

Since the simulation is to be carried out under isothermal and isobaric conditions, the system must be carried out the equilibrium simulation firstly. In the equilibrium process, Figure 3 indicates the fluctuant curve of the system temperature, and Figure 4 indicates the fluctuant curve of the system density. The system temperature is approximately equal to 300 K; the system density approximately approach to 1022 kg/m^3, equaling to the density of normal saline (1033 kg/m^3) basically. It is express that the system reaches equilibrium.

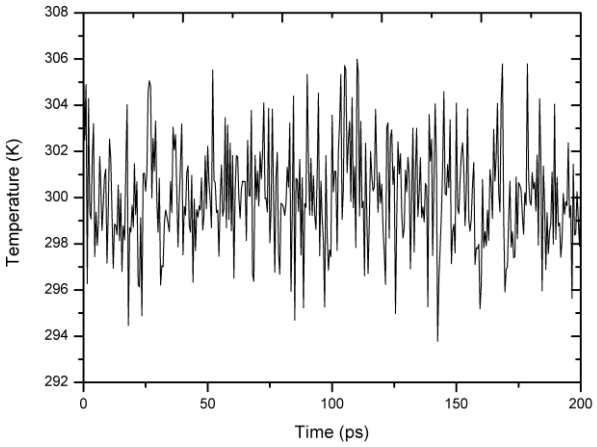


Figure 3. The fluctuant curve of the system temperature

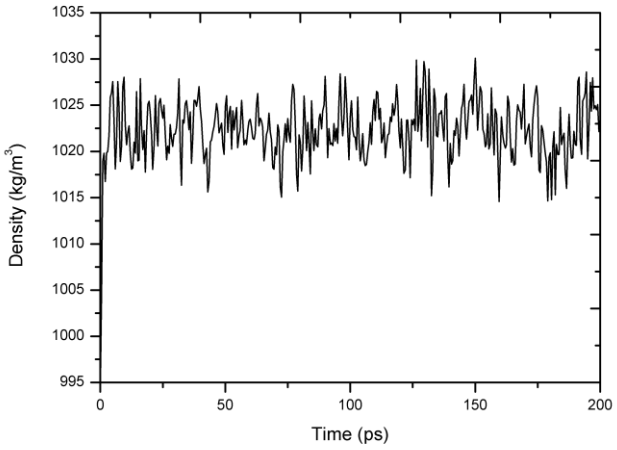


Figure 4. The fluctuant curve of the system density

To beginning the SMD simulations, the pull-group of this molecular is linked to a spring with an elastic constant k_{spring} =3800kJ/mol/nm^2, which is moved at a velocity of $\mathrm{v} = 0.01\mathrm{nm/pN}$ along the molecular axis. The values of velocity and the constant elastic are chosen based on previous works [10]-[11]. The fix-group is kept fixed with a strong position restraint (Figure 5).

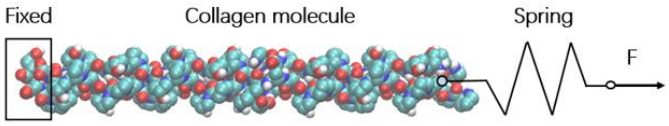


Figure 5. Schematic representation of the whole system of molecule plus virtual spring

The force applied by the virtual spring is

$$\mathrm{F(t)} = k_{spring}(x_{spring}(t) - x_{pill}(t))$$

Where x_{spring} represent the position of spring and x_{pill} represent the position of the pull-group.

On the assumption that the collagen molecule represents an elastic response when performed to stretching along the principal axis. Therefore, the whole system, molecule and spring, will respond as two springs connected in series of elastic constant k_{system}, given by the function:

$$\frac{1}{k_{system}} = \frac{1}{k_{collagen}} + \frac{1}{k_{spring}}$$

The value of k_{system} is calculated from the slot of $\mathrm{F(t)}$ versus ΔL_{system}.

We determined the structure of molecular is a cylindrical shape, and estimated a molecular radius and length, which we used to calculate the Young's modulus:

$$\mathrm{Y} = \frac{\sigma}{\varepsilon} = \frac{F/A}{\Delta L/L_0} = \frac{F}{\Delta L} \times \frac{L_0}{A} = k_{collagen} \times \frac{L_0}{A}$$

Where L_0 is the initial length of the molecule and A is the area of cross-section.

Results

We calculate the Young's modulus of the collagen molecule in terms of the elastic constant with SMD simulation method. As an example of the molecular response to the linear traction, Figure 6 indicates the Force, putting on the center of mass of the pull-group, versus collagen molecular elongation ΔL, and Table 1 shows the main features of collagen molecule. The Young's modulus of the collagen molecule is 6.1 ± 0.3 GPa.

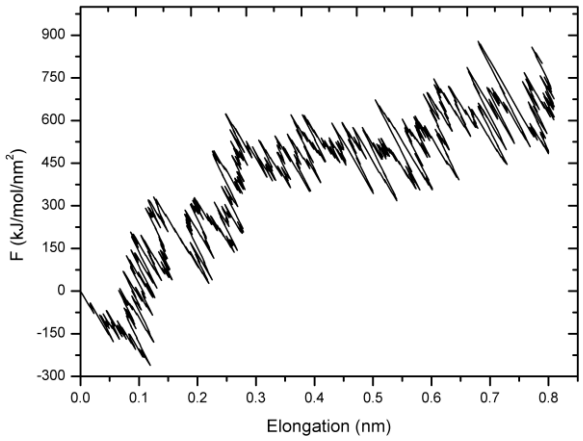


Figure 6. The force versus collagen molecular elongation ΔL

Table 1. The main features of collagen molecule

	(Pro-Hyp-Gly)$_4$-Glu-Lys-Gly-(Pro-Hyp-Gly)$_5$
$k_{collagen}$	1552 ± 59.2 pN/nm
L_0	84 Å
A	214.34 Å^2
Y	6.1 ± 0.3 GPa

Conclusions

In the present study, we build the aqueous environment of normal saline. Using GROMACS 2016 and SMD simulation method, we analyzed the Young's modulus of the collagen molecule in terms of the elastic constant. Using this model, we assessed the value of Young's modulus about 6.1 ± 0.3 GPa. The simulation results of the author are slightly larger than the simulation data of Lorenzo and Alfonso, and are slightly smaller than Buehler and Andrzej. Two reasons could be cause the difference of the result. For one thing, the reason is that the continuous improvement of the simulation software and the parameters of the biomolecular force field. For another thing, we use a kind of stretch ratio without considering other situations, in simulation process. This could be make our result less comprehensive. Further studies should focus on these situations. And we believe that these study will be published in our subsequent articles.

References

[1] Isaac E. Erickson, Sydney R. Kestle, Kilief H. Zellars, George R. Dodge, Jason A. Burdick. (2012) Improved Cartilage Repair via *in vitro* Pre-maturation of MSC Seeded Hyaluronic Acid Hydrogels, *Biomedical Materials* **7(2)**, 024110.

[2] Zhang L. H, Saeed Miramini, David W. Smith, Bruce S. Gardiner, and Alan J. Grodzinsky. (2015) Time evolution of deformation in a human cartilage under cyclic loading. *Annals of Biomedical Engineering* **43(5)**, 1166–1177.

[3] Hosseini S.M, Wu Y, Ito K, van Donkelaar C.C. (2014) The importance of superficial collagen fibrils for the function of articular cartilage. *Biomechanics and Modeling in Mechanobiology* **13(1)**, 41-51.

[4] Correa D, Lietman S.A. (2016) Articular cartilage repair: Current needs, methods and research directions. *Seminars in Cell & Developmental Biology* **62**, 67-77.

[5] Harley R, James D, Miller A, White J.W. (1977) Phonons and the elastic moduli of collagen and muscle. *Nature* **267(5608)**, 285–287.

[6] Cusack S, Miller A. (1979) Determination of the elastic constants of collagen by Brillouin light scattering. *Journal of Molecular Biology* **135(1)**, 39–51.

[7] Sasaki N, Odajima S. (1996) Stress–strain curve and Young's modulus of a collagen molecule as determined by the X-ray diffraction technique. *Journal of Biomechanics* **29(5)**, 655–658.

[8] Sun Y.L, Luo Z.P, Fertala A, An K.N. (2002) Direct quantification of the flexibility of type I collagen monomer. *Biochemical and Biophysical Research Communications* **295(2)**, 382–386.

[9] Lorenzo A.C, Caffarena E.R. (2005) Elastic properties, Young's modulus determination and structural stability of the tropocollagen molecule: a computational study by steered molecular dynamics. *Journal of Biomechanics* **38(7)**, 1527–1533.

[10] Buehler M.J. (2006) Atomistic and continuum modelling of mechanical properties of collagen: elasticity, fracture and self-assembly. *Journal of Material Research* **21(8)**, 1947–1961.

[11] Alfonso Gautieri, Simone Vesentini, Franco M. Montevecchi, Alberto Redaelli. (2008) Mechanical properties of physiological and pathological models of collagen peptides investigated via steered molecular dynamics simulations. *Journal of biomechanics* **41(14)**, 3073-3077

[12] Andrzej Mlyniec, Lukasz mazur, Krzysztof A. Tomaszewsk, Tadeusz Uhl. (2015) Viscoelasticity and failure of collagen nanofibrils: 3D Coarse-Grained Simulation Studies. *Soft Materials* **13(1)**, 47–58.

[13] Schmid N, Eichenberger AP, Choutko A, Riniker S, Winger M, Mark AE, Van Gunsteren WF. (2011) Definition and testing of the GROMOS force-field versions 54A7 and 54B7. *Biophysics of Structure & Mechanism* **40(7)**, 843–856.

[14] H. J. C. Berendsen, J. R. Grigera. T and T. P. Straatsma. (1987) The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry* **91(24)**, 6269-6271

Plate/Shell Topology Optimization with Buckling and Frequency Constraints Based on Independent Continuous Mapping Method

***†H.L. Ye[1], W.W. Wang[1], Y.K. Sui[1]**

[1]College of Mechanical Engineering and Applied Electronics Technology, Beijing University of Technology, China

*Presenting author: yehongl@bjut.edu.cn

†Corresponding author: yehongl@bjut.edu.cn

Abstract

In this paper, the topological optimization problem considering the stability and vibration characteristics of structure is studied. Firstly, the topology optimization model with the minimum structure volume as object is established based on independent, continuous and mapping (ICM) method, which subjects to the critical buckling load and frequencies as constraints. The filter functions with composite exponential function of elemental mass matrix, elemental stiffness matrix and elemental geometric matrix would be introduced, by which the three matrixes are updated in iteration putted into the topology optimization of differential equation to analyse the design sensitivity and optimize the structure. Secondly, the optimal model is conversed into the quadratic programming with the introducing of filter functions, Taylor expansion and sensitivity analysis. And then, the mathematical model is solved by dual sequence quadratic programming (DSQP) algorithm. Finally, the bisection method is applied to reduce searching region of threshold space to find the optimal mapping from "continuous" to "discrete". Numerical examples are given to illustrate the feasibility and efficiency of the proposed method.

Keywords: Topology optimization, Critical buckling load, Frequency constraints, ICM method, Plate/shell structure

1 Introduction

The topology optimization of continuum structure can reduce the cost of structure with little waste of material, owing to the ability of finding an optimum path transferring load on the base structure with given constraints [1][2]. Topology optimization has widely prospect in the application of automobile, machinery, aerospace, civil engineering and etc. The representative topology optimization methods for continuous structure include homogenization method [3], variable density method (including SIMP and RAMP interpolation model) [4]-[5], evolutionary structural optimization (ESO) [6], level set method [7], moving morphable components(MMC) [8], phase field method [9] etc.The plate/shell structures become more and more popular among the engineering designers for simple structure form, and the stability and dynamics characteristics of structure are considered as the significant factors for assessing the plate/shell structures design [10]-[14]. Therefore, considering buckling and frequency constraints is important for the plate/shell topology optimization.

In this paper, the plate/shell topology optimization model, which takes the critical buckling load and natural frequencies as constraints, is established based on independent, continuous and mapping (ICM) method. The discrete topology optimization model is translated into continuous model with the introduction of composite exponential filter function. The sensitivity analysis and first-order Taylor expansion are applied to explicit the buckling and frequency constraints, and the optimal model is solved by conversing into the

quadratic programming model. Bisection inversion strategy is used to realize the inversion of topology variables from “continuous” to “discrete”.

2 Mathematical model of multi-constraints topology optimization

2.1 ICM method and CEF filter function

The filter functions are the key technology of ICM, which realize the mapping of topological variables from “discrete” to “continuous” and establish the relationship between the optimal model and the topological variables and the physical properties of the element. Here, the composite exponential function (CEF) is selected as filter functions to realize the approximate transformation of the topological variables from 0/1 to (0, 1].

$$f(t_i) = (e^{t_i/\alpha} - 1) / (e^{1/\alpha} - 1) , \tag{2}$$

where α is the given positive constant.

The element volume, stiffness matrix, geometric stiffness matrix, and mass matrix v_i, $\boldsymbol{k}_i$, $\boldsymbol{g}_i$, $\boldsymbol{m}_i$ of i-th element in the optimal process are recognized by the filter functions as follows

$$v_i = f_v(t_i) v_i^0,\ \boldsymbol{k}_i = f_k(t_i) \boldsymbol{k}_i^0,\ \boldsymbol{g}_i = f_g(t_i) \boldsymbol{g}_i^0,\ \boldsymbol{m}_i = f_m(t_i) \boldsymbol{m}_i^0 , \tag{3}$$

where v_i^0, $\boldsymbol{k}_i^0$, $\boldsymbol{g}_i^0$, $\boldsymbol{m}_i^0$ represent the initial element volume, stiffness matrix, geometric stiffness matrix, mass matrix of i-th element, respectively, and $f_v(t_i)$, $f_k(t_i)$, $f_g(t_i)$, $f_m(t_i)$ are the corresponding filter functions.

2.2 Establishment of the topological optimization model

In order to guarantee the optimal structure meeting the requirements of stability and dynamic characteristics, the topological optimization model with the minimum structure volume as object , which subjects to the buckling load and frequencies constraints, is established.

$$\begin{cases} \text{find} \quad \boldsymbol{t} \in E^N , \\ \text{make} \quad V \to \min , \\ \text{s.t.} \quad P_{\text{cr1}} \geq \underline{P}_1 , \\ \qquad \gamma_l \geq \underline{\gamma}_l \quad (l = 1, \cdots, L) , \\ \qquad t_{\min} \leq t_i \leq 1 \quad (i = 1, \cdots, N) , \end{cases} \tag{1}$$

where $\boldsymbol{t}$ is the vector of topological variables and E^N denotes N-dimensional Euclidean space, V is the total volume of the structure, P_{cr1} and $\underline{P}_1$, respectively, present the 1-th critical buckling load and lower limit of buckling critical load, γ_l and $\underline{\gamma}_l$ are the l-th natural frequency and lower limit of l-th natural frequency, respectively, L is the total number of the frequency constraints and N is the total number of elements.

By taking advantage of the relations between the filter functions and physical properties of the element, the optimal model (1) can be rewritten as

$$\begin{cases} \text{find} \quad \boldsymbol{t} \in E^N, \\ \text{make} \quad V = \sum_{i=1}^{N} f_v(t_i) v_i^0 \to \min, \\ \text{s.t.} \quad P_{\text{cr1}}\left(f_k(t_i), f_g(t_i)\right) \geq \underline{P}_1, \\ \qquad \gamma_l(f_k(t_i), f_m(t_i)) \geq \underline{\gamma}_l \quad (l = 1, \cdots, L), \\ \qquad t_{\min} \leq t_i \leq 1 \quad (i = 1, \cdots, N). \end{cases} \tag{4}$$

In the process of optimization, the reciprocal variable of the stiffness filter function $x_i = 1/f_k(t_i)$ is introduced as design variable.

3 Standardization and solution of the optimization model

3.1 Explicit approximation of buckling constraints

The finite element method is applied for the mechanical analysis of plate/shell structure. The buckling characteristic equation for the linear elastic plate/shell structure is expressed as

$$(\boldsymbol{K} + \lambda_1 \boldsymbol{G})\boldsymbol{u}_1 = \boldsymbol{0}, \tag{5}$$

where $\boldsymbol{K}$ and $\boldsymbol{G}$ denote the structural stiffness matrix and geometric stiffness matrix respectively, λ_1 is the 1-th buckling critical load factor and $\boldsymbol{u}_1$ represents the corresponding eigenvector of λ_1. In the linear buckling finite element analysis, the λ_1 is an important index to evaluate the structural buckling performance, which is linearly related to P_{cr1}.

$$P_{\text{cr1}} = \lambda_1 \times P, \quad \underline{P}_1 = \underline{\lambda}_1 \times P, \tag{6}$$

where P is the given external mechanical load and $\underline{\lambda}_1$ is the lower limit of 1-th buckling critical load factor. Therefore, the critical buckling constraint can be simplified as the constraint of buckling critical load factor.

$$\lambda_1 \geq \underline{\lambda}_1. \tag{7}$$

Equation (5) shows that the buckling critical load factor is associated to the structural stiffness matrix and geometric stiffness matrix. Then the sensitivity analysis and first-order Taylor expansion are used to get the explicit equations of the buckling constraint.

$$\lambda_1(\boldsymbol{x}) \approx \lambda_1(\boldsymbol{x}^{(v)}) + \sum_{i=1}^{N} A_{i1}^{(v)} \frac{1}{x_i^{(v)}} x_i - \sum_{i=1}^{N} A_{i1}^{(v)}, \tag{8}$$

where $A_{i1} = \dfrac{U_{i1} + \lambda_1 \beta(x_i) V_{i1}}{V_{\sum 1}}$, $\beta(x_i) = \dfrac{f_g'(t_i) f_k(t_i)}{f_g(t_i) f_k'(t_i)} = \dfrac{\gamma_k (e^{t_i/\gamma_k} - 1)}{\gamma_g (e^{t_i/\gamma_g} - 1)} e^{t_i(\frac{1}{\gamma_g} - \frac{1}{\gamma_k})}$. $U_{i1} = 0.5\boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{k}_i \boldsymbol{u}_1$ and $V_{i1} = 0.5\boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{g}_i \boldsymbol{u}_1$ are the 1-th mode strain energy and geometric strain energy for i-th element, respectively, $V_{\sum 1} = \boldsymbol{u}_1^{\mathrm{T}} \boldsymbol{G} \boldsymbol{u}_1$ is the total geometric strain energy of structure for the 1-th

buckling mode, which can be obtain from the results of finite element buckling analysis. Then the buckling constraints in optimal model (4) can be written as

$$\sum_{i=1}^{N} c_{i1} x_i \geq d_1 \ , \tag{9}$$

where $c_{i1} = A_{i1}^{(v)} \dfrac{1}{x_i^{(v)}}$, $d_1 = \underline{\lambda}_1 - \lambda_1(\boldsymbol{x}^{(v)}) + \sum_{i=1}^{N} A_{i1}^{(v)}$.

3.2 Explicit approximation of frequencies constraints

For linear multi-degree of freedom system without considering damping and additional dynamic load, the frequency equation of structure can be expressed as follows

$$(\boldsymbol{K} - \xi_l \boldsymbol{M})\boldsymbol{p}_l = \boldsymbol{0} \ , \tag{10}$$

where $\boldsymbol{M}$ is the mass matrix of structure. ξ_l is the l-th frequency eigenvalue and $\boldsymbol{p}_l$ is the eigenvector corresponding ξ_l. Considering the relationship between natural frequency γ_l and frequency eigenvalue ξ_l meets $\xi_l = (2\pi\gamma_l)^2$, and $\gamma_l > 0$, $\underline{\gamma}_l > 0$, the frequency constraints equation can be transformed into

$$\xi_l(f_k(t_i), f_m(t_i)) \geq \underline{\xi}_l \quad (l = 1, \cdots, L) \ , \tag{11}$$

where $\underline{\xi}_l = (2\pi\underline{\gamma}_l)^2$ is the corresponding lower limit of frequency eigenvalue constraint.

The explicit method of critical buckling load factor is applied to get the explicit function of frequency eigenvalue. When the eigenvector is satisfied $\boldsymbol{p}_l^{\mathrm{T}} \boldsymbol{M} \boldsymbol{p}_l = 1$, the frequency eigenvalue can be expressed as follows

$$\xi_l(\boldsymbol{x}) \approx \xi_l(\boldsymbol{x}^{(v)}) + \sum_{i=1}^{N} B_{il}^{(v)} \frac{1}{x_i^{(v)}} x_i - \sum_{i=1}^{N} B_{il}^{(v)} \ \ (l = 1, \cdots, L) \ . \tag{12}$$

where $B_{il} = \dfrac{2}{x_i^{(v)}}(\phi(x_i^{(v)}) D_{il} - U_{il})$, $\phi(x_i) = \dfrac{f_m'(t_i) f_k(t_i)}{f_m(t_i) f_k'(t_i)} = \dfrac{\gamma_k (e^{t_i/\gamma_k} - 1)}{\gamma_m (e^{t_i/\gamma_m} - 1)} e^{t_i(\frac{1}{\gamma_m} - \frac{1}{\gamma_k})}$. $U_{il} = \dfrac{1}{2} \boldsymbol{p}_l^{\mathrm{T}} \boldsymbol{k}_i \boldsymbol{p}_l$ and $D_{il} = \dfrac{1}{2} \boldsymbol{p}_l^{\mathrm{T}} \boldsymbol{m}_i \boldsymbol{p}_l$ represent the strain energy and the kinetic energy of i-th element corresponding to the j-th frequency, respectively.

Let $B_{il} = \dfrac{2}{x_i^{(v)}}(\phi(x_i^{(v)}) D_{il} - U_{il})$, $g_l = \underline{\xi}_l - \xi_l(\boldsymbol{x}^{(v)}) - \sum_{i=1}^{N} 2(\phi(x_i^{(v)}) D_{il} - U_{il})$, and the frequency constraints can simplified as follows

$$\sum_{i=1}^{N} B_{il} x_i \geq g_l \ , \tag{13}$$

3.3 The standardization of the optimal model

The second-order Taylor expansion is applied to obtain the approximate explicit function of structure volume. Then the model (4) can be translated into the standard quadratic programming model.

$$\begin{cases} \text{find} \quad \boldsymbol{x} \in E^N \ , \\ \text{make} \quad V = \sum_{i=1}^{N} \left(a_i x_i^2 + b_i x_i \right) \to \min \ , \\ \text{s.t.} \quad \sum_{i=1}^{N} c_{i1} x_i \geq d_1 \quad , \\ \qquad \sum_{i=1}^{N} B_{il} x_i \geq g_l \quad (l = 1, ..., L) \ , \\ \qquad 1 \leq x_i \leq \bar{x}_i \quad (i = 1, ..., N) \ . \end{cases} \tag{14}$$

where

$$\begin{aligned} a_i &= \frac{1}{2} \frac{\gamma_k}{\gamma_m} \frac{m_i^0}{(x_i^{(v)})^4} \frac{e^{1/\lambda_k - 1}}{e^{1/\lambda_m - 1}} \left(\frac{e^{1/\gamma_k} - 1}{x_i^{(v)}} + 1 \right)^{\frac{\gamma_k}{\gamma_m} - 2} \times \left[\left(\frac{\lambda_k}{\lambda_m} + 1 \right) (e^{1/\gamma_k} - 1) + 2 x_i^{(v)} \right] , \\ b_i &= \frac{1}{2} \frac{\lambda_k}{\lambda_m} \frac{m_i^0}{(x_i^{(v)})^3} \frac{e^{1/\gamma_k - 1}}{e^{1/\gamma_m - 1}} \left(\frac{e^{1/\gamma_k} - 1}{x_i^{(v)}} + 1 \right) \times \left[\left(\frac{\gamma_k}{\gamma_m} + 1 \right) (e^{1/\gamma_k} - 1) + 3 x_i^{(v)} \right] . \end{aligned} \tag{15}$$

For the above mathematical model established, the number of design variables is much bigger than that of constraints. In order to reduce the number of design variables and improve the computational efficiency, the duality theory is introduced to convert the Eq.(15) into dual optimization model and dual sequence quadratic programming (DSQP) algorithm is applied to solve this optimal model.

3.4 Bisection inversion strategy

The optimal results obtained from Eq.(15) have some "continuous" topological variables, which affect the mechanical properties of the optimal structure to some degree. Here the bisection inversion strategy is applied to bisect threshold space to get a suitable filter threshold value to realize the mapping from "continuous" to "discrete", with the purpose of obtaining the optimal discrete structure.

The function $g(T)$ is defined to indicate the minimum difference between values of constraints and those of discrete structure's mechanical properties when given the filter threshold value T. Here the concrete format of equation of $g(T)$ can be written as follows

$$g(T) = \min[(P_{\text{cr1}}(T) - \underline{P}_1) / \underline{P}_1 \times 100\% \ , \ (\gamma_l(T) - \underline{\gamma}_l) / \underline{\gamma}_l \times 100\% \ \ (l = 1, \cdots, L) \ ] \tag{16}$$

The flowchart of the bisection inversion strategy is illustrated as Figure 1 and the convergence condition of $g(T)$ is given as

$$0 < g(T) <= \varepsilon \tag{17}$$

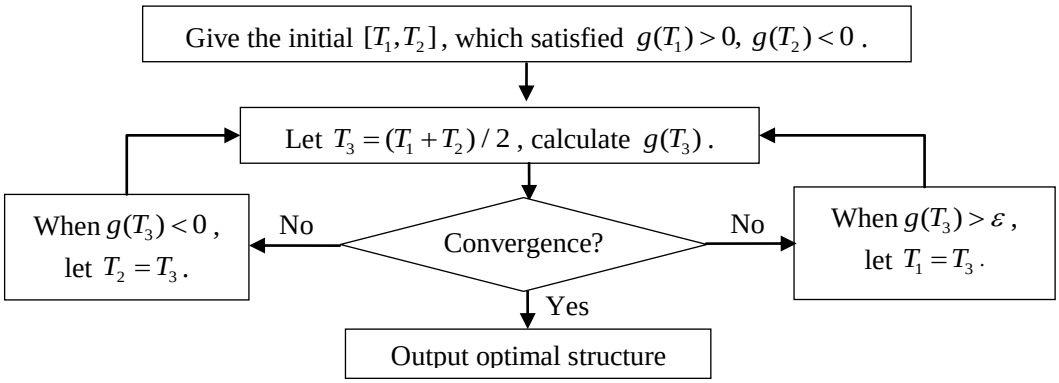


Figure 1. Flowchart of the bisection inversion strategy

4 Numerical example

The design region is a plane elastic body with size $20\times40\times1\text{mm}^3$ as shown in Figure 2. The Young's modulus $E = 68890\text{Mpa}$, Poisson's ratio μ=0.3, and the density ρ=1000kg/m^3. A concentrated mass M=2E-3kg is attached in the midpoint of the top boundary. The bottom boundary is fixed and the concentrated force 1000N is applied at the same position of concentrated mass. The basic structure is divided into 60×120 CQUAD4 elements. After finite element analysis, the first-order buckling load of the basic structures is 180.76N and The first three order frequencies of the basic structure are $\gamma_1 = 236.6\text{Hz}, \gamma_2 = 2213.8\text{Hz}, \gamma_3 = 2332.5\text{Hz}$. $\underline{P}_1 = 130\text{N}$ is defined as buckling constraint and $\underline{\gamma}_l = 0.9\times\gamma_l$ $(l = 1,2,3)$ is defined as frequency constraint.

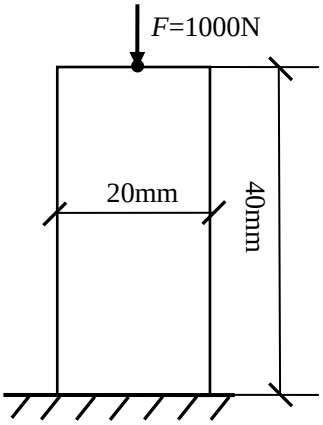


Figure 2. Design region

Table 1 gives the optimal results with different kinds of multi-constraints. The intermediate results and optimal topology configuration are illustrated in Figure 3. Figures 4 and 5 show the iteration history curves of volume and buckling load.

Comparing the optimal results in table 1, it is obvious found that the all the optimal structures satisfy the buckling and frequency constraints when different order of frequency constraint is given. And Figures 4 and 5 indicate that the processes of iteration are stable convergence. Those indicate that the optimization method is effective for solving the multi-constraints problem. From Figure 3, it can be found that the optimal structures with 2-th frequency and 3-th frequency constraints are similar, but all are different from that with 1-th frequency constraint. The main reason for this phenomenon is that the magnitudes of 2-th frequency and

3-th frequency of basic structure are approximately equivalent, meanwhile, the corresponding frequency constraint values are also almost same.

Table 1. Optimal topology process with different kinds of multi-constraints a Buckling & 1-th frequency b Buckling & 2-th frequency c Buckling & 3-th frequency

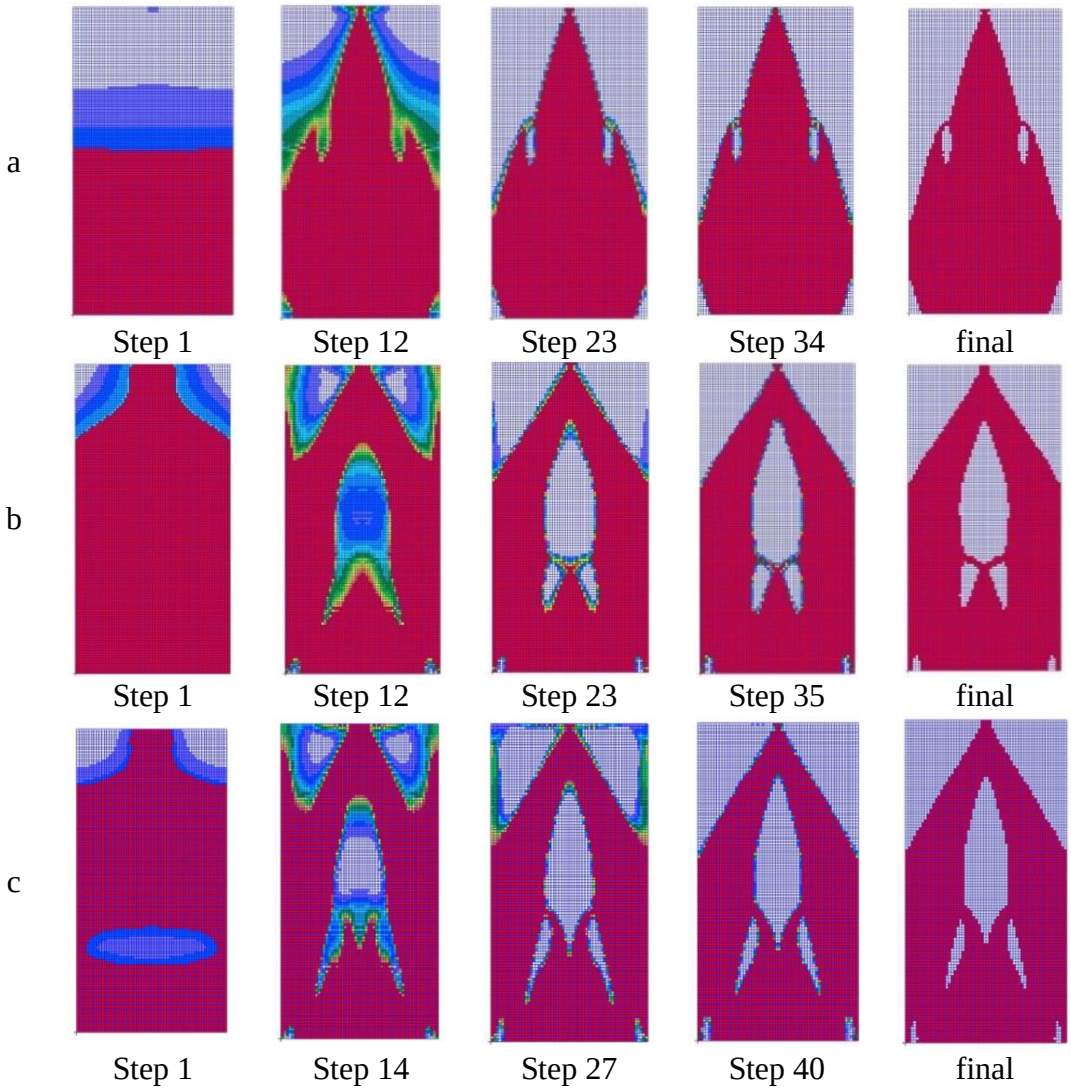


Table 2. Results of topology optimization with different kinds of multi-constraints a Buckling & 1-th frequency b Buckling & 2-th frequency c Buckling & 3-th frequency

Constraint type	a	b	c
Iteration number	34	35	40
Volume/mm^3	515.5	531.5	529.2
Buckling load/N	130.8	130.7	130.8
1-th frequency /Hz	229.3	223.2	223.2
2-th frequency /Hz	1643.2	2169.9	2085.1
3-th frequency /Hz	1647.4	2195.2	2174.3

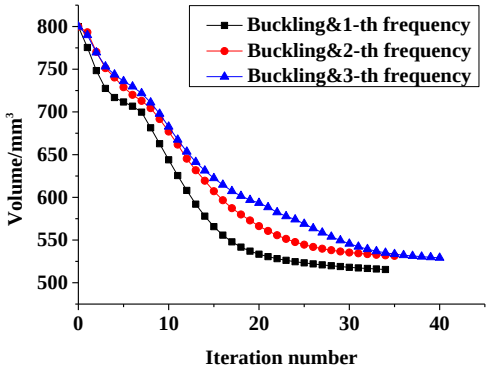


Figure 4. Iterative history curves of volume with different kinds of multi-constraints

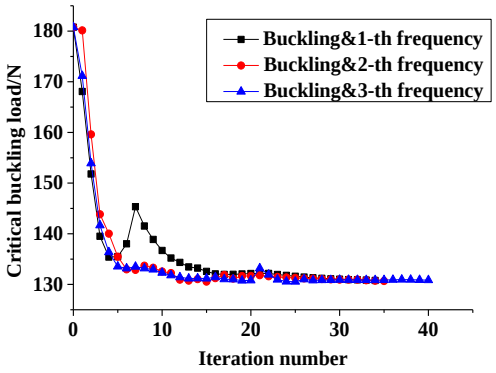


Figure 5. Iterative history curves of buckling load different kinds of multi-constraints

5 Conclusions

The plate/shell topology optimization model with minimum structural volume as objective, subjected to the critical buckling load and natural frequencies, is investigated in this paper. The composite exponential filter functions, Taylor expansion and sensitivity analysis are used to make the optimal model explicit and establish the standard quadratic programming model. And bisection inversion strategy is applied to realize the intelligent mapping of topological variables. The results of example demonstrate that proposed method for plate/shell topology optimization with buckling and frequencies constraints is feasible and valid.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (11072009, 11172013)

References

[1] Sui Y.K., Ye H.L. (2013) *Continuum Topology Optimization Methods ICM*. Science Press, Beijing.

[2] Joshua, D., Grandhi, R.V. (2014) A survey of structural and multidisciplinary continuum topology optimization: post 2000, *Structural & Multidisciplinary Optimization* **49**, 1-38.

[3] Bendsøe, M.P., Kikuchi, N. (1988) Generating optimal topologies in structural design using a homogenization method, *Computer Methods Appllied in Mechanics & Engineering* 71, 197-224.

[4] Sigmund, O. (2001) A 99 line topology optimization code written in Matlab, *Strucural & Multidisciplinary Optimization* **21**, 120-127.

[5] Bendsøe, M. P., Sigmund, O. (1999) Material interpolation schemes in topology optimization, *Archive of Appied Mechanics* **69**, 635–654.

[6] Xie, Y.M., Steven, G.P. (1993) A simple evolutionary procedure for structural optimization, *Computers & Structures* **49**, 885–896.

[7] Daróczy, L., Jármai, K. (2015) From a quasi-static fluid-based evolutionary topology optimization to a generalization of BESO, *Engineering Optimization* **47**, 689–705.

[8] Guo, X., Zhang, W.S., Zhang, J., Yuan, J. (2016) Explicit structural topology optimization based on moving morphable components (MMC) with curved skeletons. *Computer Methods Appllied in Mechanics & Engineering* **310**, 711-748.

[9] Garcke, H., Hecht, C. (2016) Shape and Topology Optimization in Stokes Flow with a Phase Field Approach, Applied Mathematics & Optimization, **73**, 23-70.

[10] Munk, D.J., Vio, G.A., Steven, G.P. (2017) A simple alternative formulation for structural optimisation with dynamic and buckling objectives, *Structural & Multidisciplinary Optimization*, 2017, **55**, 969-986.

[11] Gao, X.J., Ma, H.T. (2015) Topology optimization of continuum structures under buckling constraints, Computers & Structures **157**, 142-152.

[12] Ye, H.L., Wang, W.W., Chen, N., Sui, Y.K. (2016) Plate/shell topological optimization subjected to linear buckling constraints by adopting composite exponential filtering function, *Acta Mechanica Sinica* **32**, 649-658.

[13] Rong, J.H., Xie, Y.M., Yang, X.Y. (2000) Topology optimization of structures under dynamic response constraints. *Journal of Sound & Vibration* **234**, 177-189.

[14] Ye, H.L., Chen, N., Sui, Y.K., Tie, J. (2015) Three-Dimensional dynamic topology optimization with frequency constraints using composite exponential function and ICM method, *Mathematical Problems in Engineering* **2015**, 1-10.

Accurate viscoelastic large deformation analysis using F-bar aided edge-based smoothed finite element method for 4-node tetrahedral meshes (F-barES-FEM-T4)

Yuki Onishi[1,a)], Ryoya Iida[1] and Kenji Amaya[1]

[1] Department of Systems and Control Engineering, Tokyo Institute of Technology, Japan

[a)]Corresponding and Presenting author: yonishi@a.sc.e.titech.ac.jp

ABSTRACT

A state-of-the-art tetrahedral smoothed finite element method, F-barES-FEM-T4, is demonstrated on viscoelastic large deformation problems. The stress relaxation of viscoelastic materials brings near incompressibility when the long-term Poisson's ratio is close to 0.5. The conventional hybrid 4-node tetrahedral (T4) elements cannot avoid the shear locking and pressure checkerboarding issues due to the incompressibility, meanwhile F-barES-FEM-T4 can suppress these issues successfully. A few examples of analyses verify that F-barES-FEM-T4 is locking-free and pressure oscillation-free in viscoelastic analyses as well as in nearly incompressible hyperelastic or elastoplastic analyses.

Keywords: Smoothed finite element method, Tetrahedral element, Large deformation, Viscoelasticity, Volumetric locking, Pressure checkerboarding.

Introduction

In a practical finite element solid analysis handling objects of arbitrary shapes, hexahedral mesh generation is impossible in many cases, and thus tetrahedral mesh generation is frequently used. However, the standard 1st-order 4-node tetrahedral (T4) elements cause the shear/volumetric locking, pressure checkerboarding, etc., which lead to low accuracy solutions. Although the standard 2nd-order 10-node tetrahedral elements avoid the shear locking, it does not improve the accuracy when the material have the incompressibility. In addition, it is well known that the volumetric locking and pressure checkerboarding issues are hardly improved even with finer meshes. Therefore, researching for high accuracy tetrahedral element development to handle material incompressibility is still being carried out actively.

The most popular method in high accuracy analysis with tetrahedral meshes is the hybrid elements based on the mixed variational principle. The standard displacement-based finite elements only have the nodal displacement as the unknown variables, whereas the hybrid elements additionally have the pressure and/or volumetric strain to avoid the volumetric locking and pressure checkerboarding. The 2nd-order 10-node tetrahedral hybrid element [2, 6] is known to have sufficient accuracy in relatively mild large deformation analysis and is widely used. However, this element has some disadvantages: reduced accuracy and convergence in severe large deformation analysis due to the influence of intermediate nodes; unavailability in dynamic explicit analysis due to the influence of hybrid formulation. The 1st-order 4-node tetrahedral hybrid element [2] is considered to be one of the current best options for severe large deformation analysis; yet, the shear locking issue, pressure checkerboarding issue, and disadvantages of hybridization have not been resolved. Some other advanced hybridization techniques for tetrahedral meshes [1, 15, 7] have been proposed but are still in research stage to date.

On the other hand, the smoothed finite element method (S-FEM) [5, 16, 9, 11] has recently attracted attention as a high accuracy analysis technique for tetrahedral meshes without using the mixed variational principle. S-FEM is a displacement-based finite element method that performs strain smoothing between adjacent elements and stress integration in smoothing domains around nodes, element edges or faces. Various S-FEM formulations have been proposed so far; in particular, the authors have proposed the F-bar aided edge-based S-FEM for T4 elements (F-barES-FEM-T4) [12] for high accuracy severe large deformation analysis. F-barES-FEM-T4 is a method that incorporates the F-bar method [3] into the tetrahedral edge-based S-FEM (ES-FEM-T4) so that it resolves the volumetric locking and pressure checkerboarding issues in nearly incompressible cases. The authors validated the effectiveness of F-barES-FEM-T4 for hyperelastic bodies (such as rubber materials) whose Poisson's ratio is close to 0.5 [12, 10, 14], and elastoplastic bodies (such as metal and plastic materials)

whose plasticity coefficient is much smaller than the elastic coefficient [13, 8]. It is thought that F-barES-FEM-T4 is equally effective for viscoelastic bodies (such as thermoplastic resin and glass materials) whose long-term Poisson's ratio is close to 0.5, but demonstration thereof has not yet been carried out.

In this paper, we apply F-barES-FEM-T4 to viscoelastic bodies with nearly incompressible long-term Poisson's ratio and demonstrate its performance in severe large deformation analysis. We handle a Hencky viscoelastic model based on the generalized Maxwell model described by a Prony series [4] and perform quasi-static analysis in which stress relaxation occurs. As the stress relaxation progresses, the incompressibility gradually increases; nevertheless, F-barES-FEM-T4 gives a highly accurate solution without locking or pressure checkerboarding. Analysis results of the conventional methods are also shown to demonstrate the effectiveness of F-barES-FEM-T4.

Methods

The method proposed in this paper is basically the same as F-barES-FEM-T4 [12, 10, 14, 13, 8] already proposed in the case of hyperelastic and elastoplastic body. The difference from the previous methods is only the calculation of the stress via material constitutive models and the state variables to be held. Therefore, in this paper, the calculation method of viscous strain and stress in the case of viscoelastic body is explained in detail, and the formulation of F-barES-FEM-T4 is described in brief. For details of F-barES-FEM-T4, please refer to the reference [12] etc..

Calculation of Deformation Gradient

Let us denote the deformation gradient at edge h in ES-FEM-T4 [5] by ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}$, then the isovolumetric part of the deformation gradient at edge h in F-barES-FEM-T4, ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}^{\mathrm{iso}}$, is described as

$$ {}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}^{\mathrm{iso}} = \left(\frac{1}{{}^{\mathrm{Edge}}_{h}\widetilde{J}} \right)^{1/3} {}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}; \tag{1} $$

$$ {}^{\mathrm{Edge}}_{h}\widetilde{J} = \det({}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}), \tag{2} $$

On the other hand, the volumetric part of the deformation gradient at edge h in F-barES-FEM-T4, ${}^{\mathrm{Edge}}\overline{\boldsymbol{F}}^{\mathrm{vol}}$, is given by the cyclic smoothing of J among elements and nodes followed by the edge-based smoothing (see reference [12] etc. for detail). Finally, the deformation gradient at edge h in F-barES-FEM-T4, ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}$, is obtained by combining ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}^{\mathrm{iso}}$ and ${}^{\mathrm{Edge}}\overline{\boldsymbol{F}}^{\mathrm{vol}}$ with the F-bar method.

$$ {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}} = {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{vol}} \cdot {}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}^{\mathrm{iso}}. \tag{3} $$

The number of cyclic smoothing, c, is the tuning parameter of F-barES-FEM-T4. F-barES-FEM-T4 with c-time cyclic smoothings is referred to as "F-barES-FEM-T4(c)" hereafter in this paper.

Calculation of Stress

The Cauchy stress at edge h, ${}^{\mathrm{Edge}}_{h}\boldsymbol{T}$, is then derived in the standard way with ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}$. The following shows the derivation in case of the Hencky viscoelastic body based on the generalized Maxwell model described with the Prony series [4]. Hereafter in this subsection, ${}^{\mathrm{Edge}}_{h}\square$ is omitted.

The hydrostatic part of the Cauchy stress, $\boldsymbol{T}^{\mathrm{hyd}}$, is calculated in the same way as the Hencky hyperelastic body:

$$ \boldsymbol{T}^{\mathrm{hyd}} = K \, \mathrm{tr}(\boldsymbol{H}) \, \boldsymbol{I}, \tag{4} $$

where K is the bulk modulus (constant), $\mathrm{tr}(\square)$ is the operator to return the trace, and $\boldsymbol{H}$ is the Hencky (logarithmic) strain derived by ${}^{\mathrm{Edge}}\overline{\boldsymbol{F}}$. On the other hand, the deviatoric part of the Cauchy stress, $\boldsymbol{T}^{\mathrm{dev}}$, is given by

$$ \boldsymbol{T}^{\mathrm{dev}} = 2G_0 \left(\boldsymbol{H}^{\mathrm{dev}} - \sum_i g_i \boldsymbol{H}_i^{\mathrm{v}} \right), \tag{5} $$

where G_0 is the instantaneous shear modulus, g_i is the ith non-dimensional shear modulus in the Prony series, and $\boldsymbol{H}_i^{\mathrm{v}}$ is the ith Hencky (logarithmic) viscous strain. The logarithmic viscous strain after a time increment, $\boldsymbol{H}_i^{\mathrm{v}+}$, is calculated with the following time advancing equation.

$$ \boldsymbol{H}_i^{\mathrm{v}+} = \boldsymbol{R} \cdot \boldsymbol{H}_i^{\mathrm{v}} \cdot \boldsymbol{R}^T + \Delta \boldsymbol{H}_i^{\mathrm{v}}; \tag{6} $$

if $\Delta t \ll \tau_i$, then

$$\Delta \boldsymbol{H}_i^{\mathrm{v}} = \frac{\Delta t}{\tau_i}\left(\frac{1}{2}\Delta \boldsymbol{H}^{\mathrm{dev}} + \boldsymbol{R}\cdot(\boldsymbol{H}^{\mathrm{dev}} - \boldsymbol{H}_i^{\mathrm{v}})\cdot\boldsymbol{R}^T\right), \tag{7}$$

else

$$\Delta \boldsymbol{H}_i^{\mathrm{v}} = \left(1 - \frac{\tau_i}{\Delta t}\left(1 - \exp(-\frac{\Delta t}{\tau_i})\right)\right)\Delta \boldsymbol{H}^{\mathrm{dev}} + \left(1 - \exp(-\frac{\Delta t}{\tau_i})\right)\boldsymbol{R}\cdot(\boldsymbol{H}^{\mathrm{dev}} - \boldsymbol{H}_i^{\mathrm{v}})\cdot\boldsymbol{R}, \tag{8}$$

where Δt is the time increment, $\boldsymbol{R}$ is the rigid rotation in the time increment, and τ_i is the ith relaxation time in the Prony series. Conclusively, $\boldsymbol{T} = \boldsymbol{T}^{\mathrm{hyd}} + \boldsymbol{T}^{\mathrm{dev}}$ gives the Cauchy stress at each edge.

Calculation of Nodal Internal Force

The contribution of each edge to the nodal internal force, $\{{}^{\mathrm{Edge}}\boldsymbol{f}^{\mathrm{int}}\}$, is calculated in manner of the F-bar method:

$${}^{\mathrm{Edge}}_{h}f^{\mathrm{int}}_{P:p} = \frac{\partial {}^{\mathrm{Edge}}_{h}\widetilde{D}_{ij}}{\partial \dot{u}_{P:p}}\,{}^{\mathrm{Edge}}_{h}T_{pl}\,{}^{\mathrm{Edge}}_{h}V, \tag{9}$$

where ${}^{\mathrm{Edge}}\widetilde{\boldsymbol{D}}$ is the rate of stretching, u is the displacement, and V is the corresponding volume.

Results

Tensile Suspension of Viscoelastic Block

A large deformation analysis of a viscoelastic block subjected to horizontal tension and vertical gravity is performed. An outline of the analysis is shown in Fig. 1. The size of the block is $3 \times 2 \times 1$ m^3; its left end face is perfectly constrained; the right end face is quickly displaced 3 m in $+x$ direction for 10 s at a constant velocity and is kept the displacement for 1000 s so that the body is suspended. The material model is the Hencky viscoelastic body based on the generalized Maxwell model [4] with 1 MPa instantaneous Young's modulus, 0.3 instantaneous Poisson ratio, 0.96 non-dimensional shear modulus for the 1st (and only) term of the Prony series (i.e., the long-term Young's modulus is about 0.046 MPa and the long-term Poisson's ratio is about 0.49) and its relaxation time constant is 10 s. The mass density is 1000 kg / m^3, and the gravitational acceleration is 9.8 m/s^2 in $-z$ direction. An unstructured 4-node tetrahedral mesh with 0.2 m seed size (1405 nodes and 6183 elements) is prepared and is analyzed by F-barES-FEM-T4(c) for $c = 1, 2, 3$ and also by the conventional 4-node tetrahedral element (ABAQUS C3D4) and its hybrid element (ABAQUS C3D4H). In addition, an analysis is also performed with the 8-node hexahedral selective reduced integration element (ABAQUS C3D8) using a structured mesh in the same seed size (1122 nodes and 800 elements) as a reference solution.

Figure. 2–5 show the deformations and stress distributions of the analysis results. Since the viscoelastic body at the initial stage of analysis shows behavior close to a compressible elastic body with the instantaneous modulus, relatively minor errors are only observed in the results at the end of the tension ($t = 10$ s). However, since it shifts to behavior close to an incompressible elastic body with the long-term modulus over time, relatively major errors are observed in the results at the end of the analysis ($t = 1010$ s). ABAQUS C3D4 clearly suffered from shear/volumetric locking and pressure checkerboarding. Although ABAQUS C3D4H shows a relatively smooth Mises stress distribution, it can be seen that the pressure distribution is partly in the checkerboard pattern. On the other hand, every F-barES-FEM-T4 shows smooth distributions not only in the Mises stress but also in the pressure.

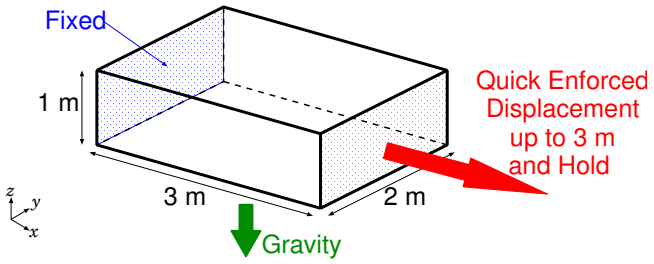
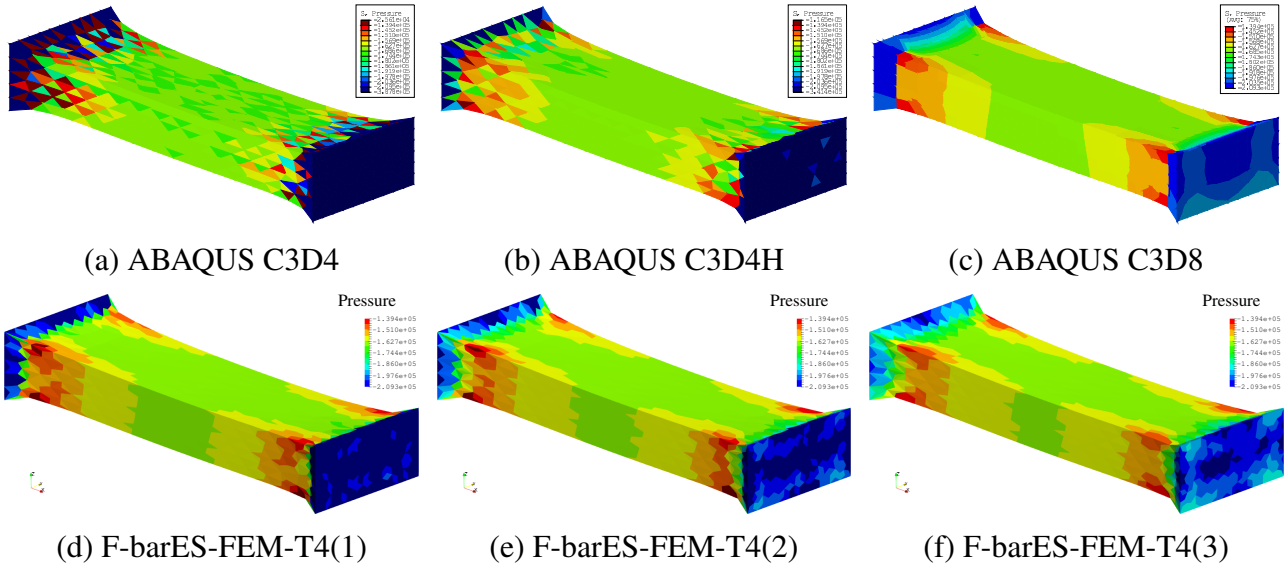


Figure 1. Outline of the tensile suspension analysis of a viscoelastic block.

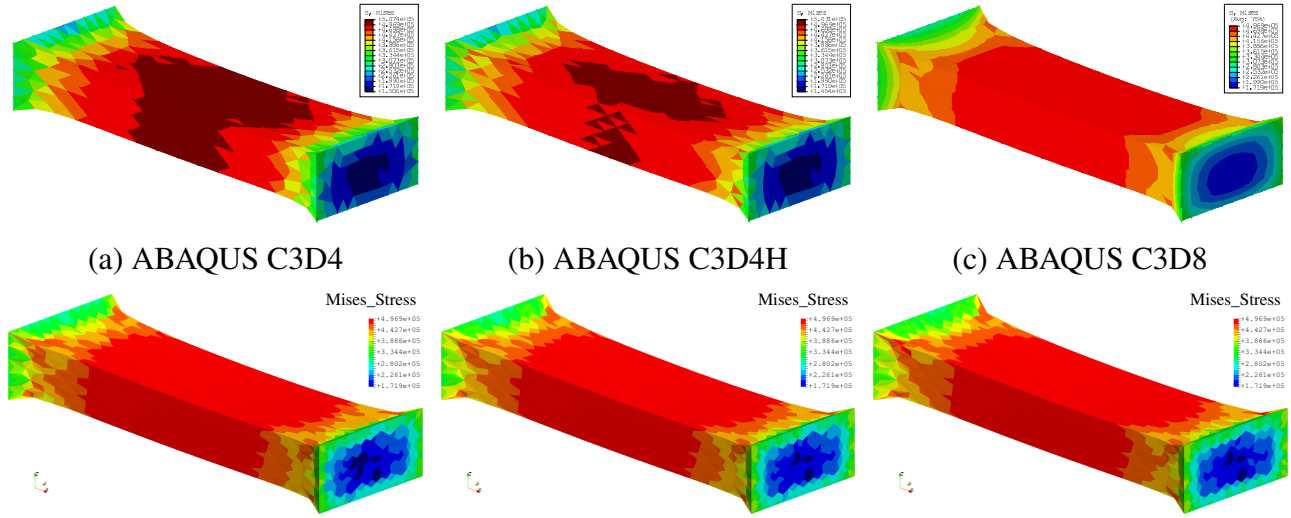
Figure. 6 compares the displacement time histories of the node at the center of the bottom face in z direction. The node is displaced in $+z$ direction up to 10 s because of the decrease in the cross-sectional area due to the quick tension of the block in a compressible state. Meanwhile, the node hangs down in $-z$ direction after 10 s because of the stress relaxation and gravity and approaches to the long-term stable position gradually. Because ABAQUS C3D4 can not avoid both shear and volumetric locking, a much harder solution is obtained. Although ABAQUS C3D4H avoids volumetric locking, sheer locking can not be avoided [12] and thus a slightly hard solution is obtained. On the other hand, F-barES-FEM-T4s show the almost same accuracy solutions as ABAQUS C3D8 regardless of the number of cyclic smoothings c, which confirms that our method avoids both shear and volumetric locking.



(a) ABAQUS C3D4 (b) ABAQUS C3D4H (c) ABAQUS C3D8

(d) F-barES-FEM-T4(1) (e) F-barES-FEM-T4(2) (f) F-barES-FEM-T4(3)

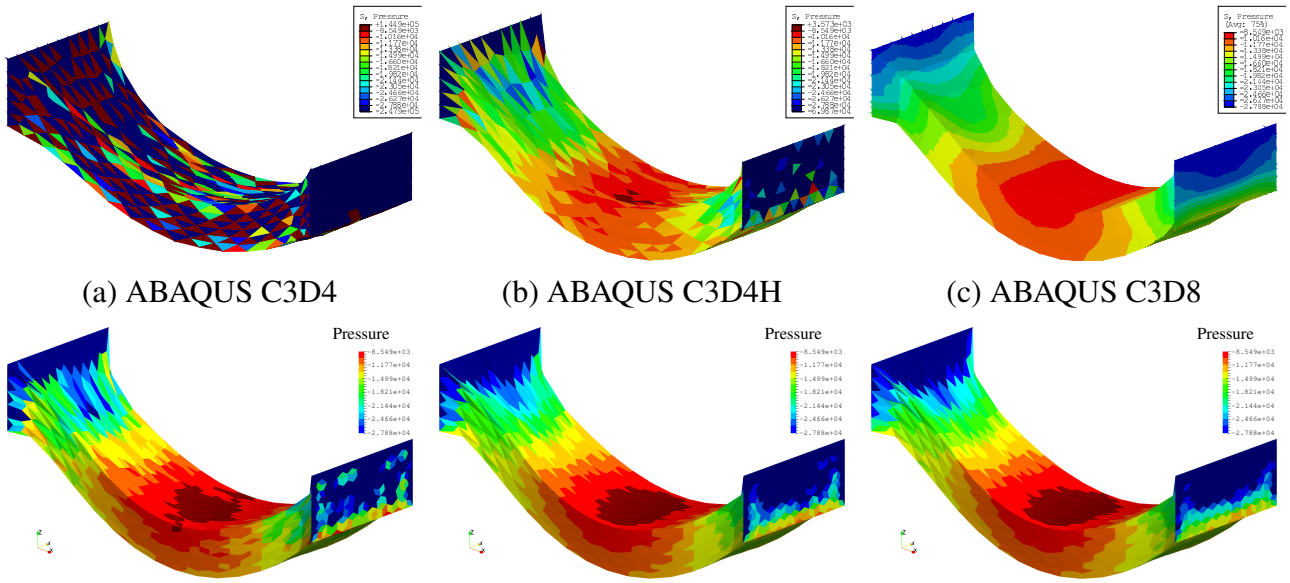
Figure 2. Comparison of pressure distributions at the end of the tension (t = 10 s) in the tensile suspension analysis. The contour range is [−209.3 kPa, −139.4 kPa] and is in common with each other.



(a) ABAQUS C3D4 (b) ABAQUS C3D4H (c) ABAQUS C3D8

(d) F-barES-FEM-T4(1) (e) F-barES-FEM-T4(2) (f) F-barES-FEM-T4(3)

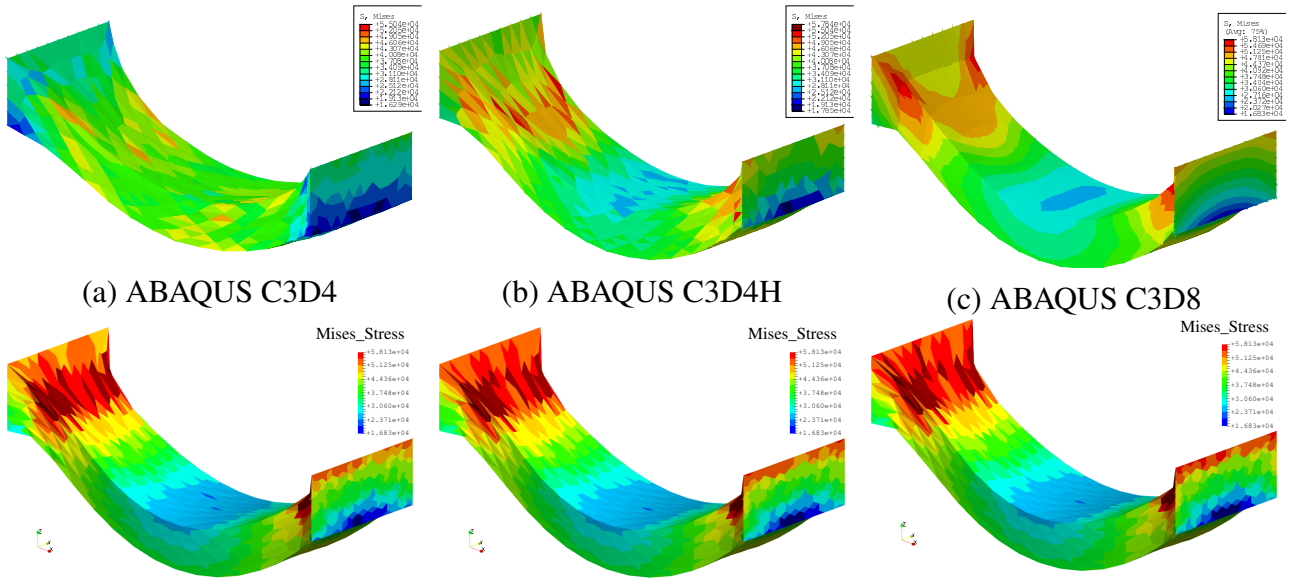
Figure 3. Comparison of Mises stress distributions at the end of the tension (t = 10 s) in the tensile suspension analysis. The contour range is [171.9 kPa, 496.9 kPa] and is in common with each other.



(a) ABAQUS C3D4 (b) ABAQUS C3D4H (c) ABAQUS C3D8

(d) F-barES-FEM-T4(1) (e) F-barES-FEM-T4(2) (f) F-barES-FEM-T4(3)

Figure 4. Comparison of pressure distributions at the end of the analysis (t = 1010 s) in the tensile suspension analysis. The contour range is [−27.88 kPa, −8.549 kPa] and is in common with each other.



(a) ABAQUS C3D4 (b) ABAQUS C3D4H (c) ABAQUS C3D8

(d) F-barES-FEM-T4(1) (e) F-barES-FEM-T4(2) (f) F-barES-FEM-T4(3)

Figure 5. Comparison of Mises stress distributions at the end of the analysis (t = 1010 s) in the tensile suspension analysis. The contour range is [16.83 kPa, 58.13 kPa] and is in common with each other.

Tensile Drooping of Viscoelastic Twisted Prism

A large deformation analysis of a viscoelastic twisted prism subjected to vertical tension and gravity is performed. An outline of the analysis is shown in Fig. 7. The body to be analyzed is an object whose height is 10 m and whose cross-section is a right triangle with 3, 4, and 5 m side lengths twisted 180 degrees around the center of cross-sections, which is difficult to be meshed into regular hexahedral elements. The bottom face is perfectly constrained; the top face is quickly displaced 5 m in $+z$ direction for 10 s at a constant velocity and is kept the displacement for 1500 s so that the body is drooped. The material model is almost the same as the previous example, except that the instantaneous Young's modulus

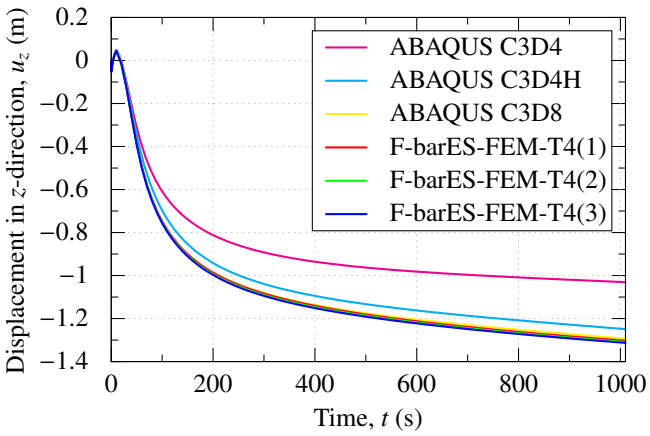


Figure 6. Comparison of displacement time histories of the node at the center of the bottom face in z direction in the tensile suspension analysis.

is 2 MPa. An unstructured 4-node tetrahedral mesh with 0.5 m seed size (960 nodes and 3514 elements) is prepared and is analyzed by F-barES-FEM-T4(2) and ABAQUS C3D4H.

Figure 8–10 compare the deformations and pressure distributions of the analysis results. As in the previous example, there are only minor differences between F-barES-FEM-T4(2) and ABAQUS C3D4H at an earlier stage (t = 10 s) although a little pressure oscillations are observed in the result of ABAQUS C3D4H. However, there are major difference between them at an later stage (t = 1510 s), where a sufficient time has passed compared to the relaxation time constant. The pressure sign distribution of ABAQUS C3D4H shows clear pressure checkerboarding and thus its stress distribution must be inaccurate. Also, the drooping deformation of ABAQUS C3D4H is slightly smaller than that of F-barES-FEM-T4(2), which is probably because of the tendency of ABAQUS C3D4H giving a harder solution. Since F-barES-FEM-T4(2) shows the pressure distributions without any oscillation, it seems to give a more appropriate displacement/stress solution.

Conclusion

A state-of-the-art tetrahedral smoothed finite element method, F-barES-FEM-T4, was demonstrated in quasi-static viscoelastic large deformation analyses. Comparison with the conventional hybrid tetrahedral element (ABAQUS C3D4H) applicable to severe large deformation problems, the proposed method was locking-free and pressure checkerboarding-free and thus gave far better solutions in displacement and stress as well as hyperelastic and elastoplastic cases with material incompressibility. When analyzing the relaxation behavior for a time sufficiently longer than the relaxation time

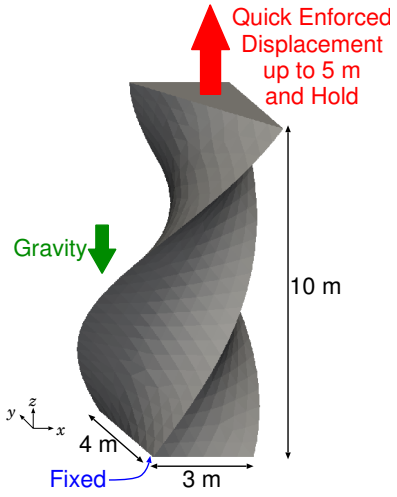
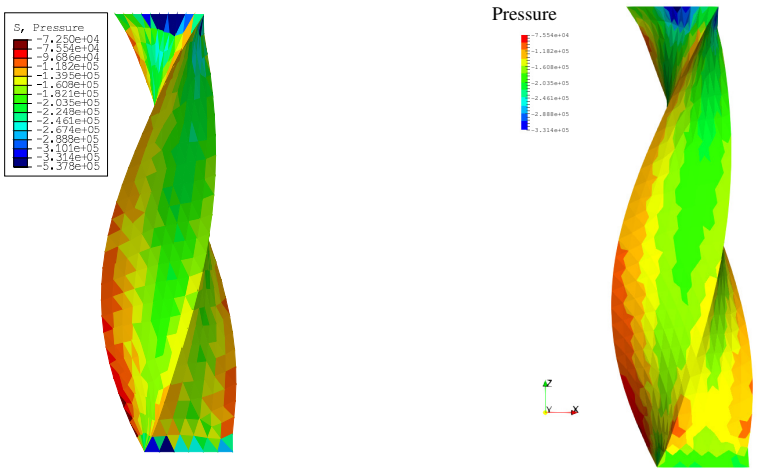
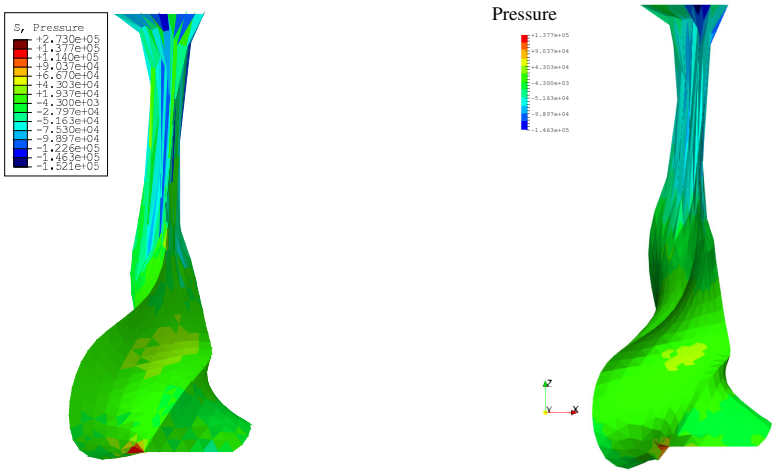


Figure 7. Outline of the tensile drooping analysis of a viscoelastic twisted prism.



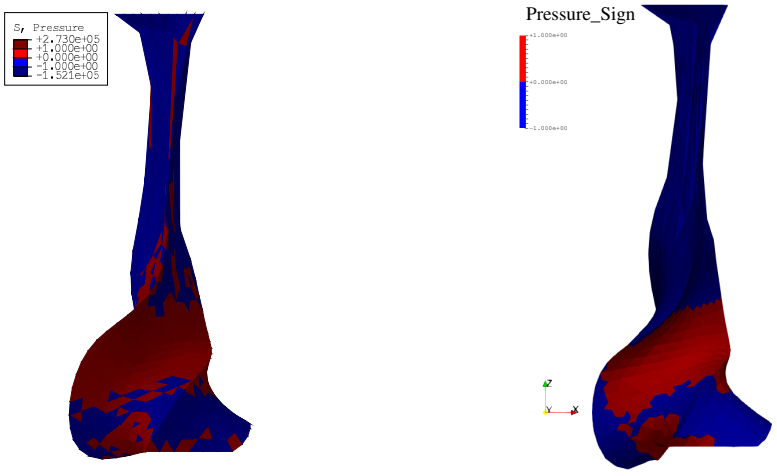
(a) ABAQUS C3D4H (b) F-barES-FEM-T4(2)

Figure 8. Comparison of pressure distributions at the end of the tension (t = 10 s) in the tensile drooping analysis. The contour range is [−331.4 kPa, −75.54 kPa] and is in common.



(a) ABAQUS C3D4H (b) F-barES-FEM-T4(2)

Figure 9. Comparison of pressure distributions at the end of the analysis (t = 1510 s) in the tensile drooping analysis. The contour range is [−146.3 kPa, +137.7 kPa] and is in common.



(a) ABAQUS C3D4H (b) F-barES-FEM-T4(2)

Figure 10. Comparison of pressure sign distributions at the end of the analysis (t = 1510 s) in the tensile drooping analysis. The red and blue colors represent the compressed and tensile parts, respectively.

with respect to the viscoelastic material, it is inevitable to treat the near incompressibility because the long-term material constants decide the phenomenon. Therefore, the proposed method is thought to be effective for resin molding simulations in complex shapes around the glass transition temperature.

F-barES-FEM-T4 is one of the best tetrahedral FE formulations superior in all rubbers, plastic metals and viscoelastic resins in which incompressibility appears. A wide practical use of F-barES-FEM-T4 for more complex body shapes is expected in the future.

References

[1] Bonet, J., Gil, A. J., Lee, C. H., Aguirre, M., and Ortigosa, R. (2015). A first order hyperbolic framework for large strain computational solid dynamics. part i: Total lagrangian isothermal elasticity. *Computer Methods in Applied Mechanics and Engineering*, 283:689 – 732.

[2] Corp., D. S. S. (2013). *ABAQUS 6.13 Theory Guide*. Dassault Systèmes Simulia Corp., Providence, RI, USA.

[3] de Souza Neto, E., Peric, D., Dutko, M., and Owen, D. (1996). Design of simple low order finite elements for large strain analysis of nearly incompressible solids. *International Journal of Solids and Structures*, 33(20-22):3277–3296.

[4] de Souza Neto, E. A., Peric, D., and Owen, D. R. J. (2008). *Computational Methods for Plasticity: Theory and Applications*. Wiley, Hoboken, NJ, USA.

[5] Liu, G. R. and Nguyen-Thoi, T. (2010). *Smoothed Finite Element Methods*. CRC Press, Boca Raton, FL, USA.

[6] MSC Software Corp. (2013). *MARC 2013.1 Volume A: Theory and User Information*. MSC Software Corp., Santa Ana, CA, USA.

[7] Ong, T. H., Heaney, C. E., Lee, C.-K., Liu, G., and Nguyen-Xuan, H. (2015). On stability, convergence and accuracy of bES-FEM and bFS-FEM for nearly incompressible elasticity. *Computer Methods in Applied Mechanics and Engineering*, 285(0):315–345.

[8] Onishi, Y. (2017). F-bar aided edge-based smoothed finite element method with 4-node tetrahedral elements for static large deformation elastoplastic problems. *International Journal of Computational Methods*, under review.

[9] Onishi, Y. and Amaya, K. (2014). A locking-free selective smoothed finite element method using tetrahedral and triangular elements with adaptive mesh rezoning for large deformation problems. *International Journal for Numerical Methods in Engineering*, 99(5):354–371.

[10] Onishi, Y. and Amaya, K. (2015a). F-bar aided edge-based smoothed finite element method with tetrahedral elements for large deformation analysis of nearly incompressible materials. *The 6th International Conference on Computational Methods (ICCM2015)*.

[11] Onishi, Y. and Amaya, K. (2015b). Performance evaluation of the selective smoothed finite element methods using tetrahedral elements with deviatoric/hydrostatic split in large deformation analysis. *Theoretical and Applied Mechanics Japan*, 63:55–65.

[12] Onishi, Y., Iida, R., and Amaya, K. (2017). F-bar aided edge-based smoothed finite element method using tetrahedral elements for finite deformation analysisof nearly incompressible solids. *International Journal for Numerical Methods in Engineering*, 109(11):1582–1606.

[13] Onishi, Y., Ryoya, I., and Amaya, K. (2016). F-bar aided edge-based smoothed finite element methods with 4-node tetrahedral elements for static large deformation hyperelastic and elastoplastic problems. *The 7th International Conference on Computational Methods (ICCM2016)*.

[14] Ryoya, I., Onishi, Y., and Amaya, K. (2016). Performance evaluation of various smoothed finite element methods with tetrahedral elements in large deformation dynamic analysis. *The 7th International Conference on Computational Methods (ICCM2016)*.

[15] Scovazzi, G., Carnes, B., Zeng, X., and Rossi, S. (2016). A simple, stable, and accurate linear tetrahedral finite element for transient, nearly, and fully incompressible solid dynamics: a dynamic variational multiscale approach. *International Journal for Numerical Methods in Engineering*, 106(10):799–839. nme.5138.

[16] Zeng, W. and Liu, G. R. (2016). Smoothed finite element methods (S-FEM): An overview and recent developments. *Archives of Computational Methods in Engineering*, pages 1–39.

Pointwise Gauge Field and Relativistic Structure

Wang Yi-ping [1,2]

[1] Zhejiang Quzhou Association of Senior Scientists and Technicians, Zhejiang Quzhou 324000, China
[2] Qianjiang Mathematics and Power Engineering Institute, Zhejiang Quzhou 324000, China
† † Corresponding author and Presenting author: wyp3025419@163.com

Abstract
The quantization element concept of which infinite pointwise is mechanics element and space-time geometry possesses random "asymmetry and nonuniformity" and other factors is proposed; combination of gauge field and relativistic structure (circle logarithm) is proposed, which includes gravitational field, electromagnetic field, nuclear field, thermodynamic field and photon field to constitute pointwise quantum eleven-dimensional space and the first and second gauge invariance and build relativistic structure (circle logarithm), realizing "exact solution within [0-1/2-1] $^{[0/1/2-1]}$ under the topological variation rule without specific contents". It is provided with a superiority of conciseness, self-consistency and zero error, and is widely applicable to physics, astronomy and mechanics fields.
Keywords: Gauge field Pointwise quantization General relativity Relativistic structure (circle logarithm) Limit topological phase transition

1. Introduction

In 1954, Yang Zhenning and R. L. Mills proposed the theory with positioned isospin invariance and directly extended it to other non-Abel gauge transformation group .which is called quantum gauge field[1].
In 1967-1968, S. Weinberg and A. Salam proposed that the corresponding gauge field quantum is a massless photon through vacuum spontaneous symmetry breaking mechanism (Higgs mechanism) that makes non-Abel gauge field obtain quality proposed by Higgs et al., which is called Weak Quantum Electronic Dynamics (QED).
In 1964, M. Gail-Man and G. Zwick established the invariant strong interaction theory under transformation of the SU(3) domain after proposing the image of hadron constituted by quark, and there are 8 species of gluon as corresponding gauge field quantum which is called Quantum Chromodynamics (QCD).
In 1905-1915, Einstein established the general relativity and special relativity .revealed the macroscopic cosmic gravitational field[2].
Sharp contradiction has been generated between Quantum theory and relativity as well as macroscope (non-uniformity) and microscope (uniformity) [3]. It is a very attractive idea that whether the various interactions — electromagnetism, weak force, strong force, gravitational force, thermal power and photon field have been known as six kinds of mechanical interactions, and constituted a complete and unified new pointwise gauge field, which satisfies the derivation of symmetry principle with gauge transformation and the derivation that gauge field is free of mass element. As well as with the infinite program, infinite set of reciprocity[4] and symmetry expansion[5].
In 2000, the United States Kerry Institute of Mathematics to the standard field as one of the seven mathematical problems in the 21st century.
Quantization unit body which proposed infinite pointwise quantization (including randomized uniform and non-uniform, symmetric and asymmetric, continuous and discontinuous, qualitative and non-qualitative quantization) can be taken as the equivalent replacement of 11-dimensional (massless photons) geometric space-time concept. achieve "topological variation

rules without quality- heat- space-time and accurate solution in the range of $[0\sim1/2\sim1]^{[0\sim1/2\sim1]}$", which is called "pointwise gauge field (circle logarithm, super symmetric cell matrix)".

2. Pointwise quantization and combination

2.1 Basic definition

Definition 1 Pointwise and Pointwise Equation: the pointwise is infinite quantization element with various interaction (including photodynamic field, gravitational field, electromagnetic field, quark field, gluon field and thermodynamic field) and quantized 11-dimensional geometric space-time, including unlimited program novel pointwise mechanics equation $\{L[\psi(x_\beta), A_\mu(x_\beta)]\}^Z$ constituted by randomized "uniform and non-uniform, symmetric and asymmetric, continuous and discontinuous, qualitative and non-qualitative quantization".

Definition 2 Combination Coefficient of Pointwise. The pointwise quantum has a variety of combinations from "1-1" to "11-11" in 11-dimensional space.

$$(C_{1+p})=(S!)/(P!)=(S-0)(S-1)\cdots(S-p)\ !\ /(p-0)(p-1)\cdots(1)\ !;$$

Definition 3 Sample Space of Pointwise

$$\{r_H^{11}\}^Z=\{L[\psi(x_\beta), A_\mu(x_\beta)]_h\}^Z=\sum\sum\{r_h^{11}\}^Z=\sum\prod\{r_h^{11}\}^Z$$

Definition 4 Average Sample Space of Pointwise ≈ refers to equivalent

$$\{r_{0H}^{11}\}^Z=\sum[(C_{1+h}^{-1})^k\sum\{L[\psi(x_\beta), A_\mu(x_\beta)]_h\}^Z$$
$$=\sum[(C_{1+h}^{-1})^k\prod\{L[\psi(x_\beta), A_\mu(x_\beta)]_h\}^Z=\sum\sum\{L[\psi(x_0), A_\mu(x_0)]_0\}^Z$$
$$=\sum\prod\{L[\psi(x_0), A_\mu(x_0)]_0\}^Z=\sum\sum\{r_0^{11}\}^Z=\sum\prod\{r_0^{11}\}^Z;$$

Definition 5 11-dimensional geometric space element $\{r^{11}\}Z$ with positive, middle and negative properties is composed of the three forms of neutrino oscillation (u_e, u_ψ, u_φ), the 8th order matrix of Gell-mann and the Einstein gravitational three-dimensional space. Definition of regularization coefficients in various combinations of pointwise 11-dimensional geometric sample spaces is as follows: Including

$$\{r_{0H}^{11}\}^Z=\sum_{i=h}(C_{1+h}^{-1})\prod_{i=h}L[\psi(x_\beta), A_\mu(x_\beta)]=\sum_{i=h}(C_{1+h}^{-1})\sum_{i=h}L[\psi(x_\beta), A_\mu(x_\beta)]$$
$$\approx\sum\prod L[\psi(x_0), A_\mu(x_0)]\approx\sum\sum L[\psi(x_0), A_\mu(x_0)]$$
$$=(1/1)\{r_0\}^{K(Z\pm0)}+(1/11)\{r_1\}^{K(Z\pm1)}+(1/55)\{r_2\}^{K(Z\pm2)}+(1/165)\{r_3\}^{K(Z\pm3)}$$
$$+(1/330)\{r_4\}^{K(Z\pm4)}+1/462)\{r_5\}^{K(Z\pm5)}+(1/462)\{r_6\}^{K(Z\pm6)}+(1/330)\{r_7\}^{K(Z\pm7)}$$
$$+(1/165)\{r_8\}^{K(Z\pm8)}+(1/55)\{r_9\}^{K(Z\pm9)}+(1/11)\{r_{10}\}^{K(Z\pm10)}+(1/1)\{r_{11}\}^{K(Z\pm11)}$$
$$=\{2\}^{11}\{r_0\}^Z=2048\{r_0\}^Z=1024\{C^2\}^{[Z-1]}$$

Converted to 2048 pointwise energy particles (or 1024 quantum bits) in 11-dimensional $\{r_0^{11}\}^Z$,

Definition 6 Power Function Equation

$$Z=Z/T,t=K(S\pm N\pm N\pm p)/T,t=K(S\pm N\pm N\pm p\pm0)+K(S\pm N\pm N\pm p\pm1)+\cdots$$
$$+K(S\pm N\pm N\pm p\pm p)+\cdots+K(S\pm N\pm N\pm p\pm q)/T,t;$$

Abbreviated as: $=K\{(Z\pm0),(Z\pm1),\ldots,(Z\pm p),\ldots,(Z\pm q)\}/T,t;$

Where K= (+1,0,±0,-1); (S±N) refers to infinite dimension; (±N) refers to finite dimensional order; (±N) refers to calculus order; (±P) refers to polynomial order; T (temperature) t (time). For time and thermodynamic functions are synchronized with geometric space, they are no longer described except it is necessary.

Where $L[\psi(x_\beta), A\mu(x_\beta)]$ (particle, scalar); $L[\psi(x^2{}_\beta), A\mu(x^2{}_\beta)]$ (wave, vector); (C_{1+p}) refers to regularization coefficient of pointwise combination; { } refers to combination set of pointwise; (S!) refers to factorial of number of combinations, and (P!) refers to factorial of the order.

Where {} refers to set; the lower footnotes of $_{h,\ H,\ 0h,\ 0H}$ refer to set of sample space and average sample space of total items and sub-items.

In particular, the values can be zero, defect, error in the combination of regularization coefficients; pointwise position cannot be vacant so as to ensure the stability of "error" automatically eliminating process of "defective point quantum" and regularization.

3. Derivation of pointwise gauge field

3.1 Combination of pointwise gauge field

according to mechanical equation of Yang-Mills (Equation 11.3.17)

$L[\psi(x),A_\mu(x)] = -\psi^-(\gamma_\mu(d/dx_\mu)+m_\beta)\psi - (1/4)F_{\mu\nu}F_{\mu\upsilon} + I e\psi^-\gamma_\mu\psi A_\mu$

expanding to pointwise gauge field

$$L[\psi(x_\beta),A_\mu(x_\beta)]^Z = [\psi^-(\gamma_\mu(d^Z/dx^Z{}_{\mu\beta})+m_\beta)\psi + (1/4)F_{\mu\nu\beta}F_{\mu\upsilon\beta} + I\, m_\beta e\psi^-\gamma_{\mu\beta}\psi A_{\mu\beta}]^Z \quad (1)$$

$M_\beta = (m_p + m_n)$ mass-to-charge ratio: β = (proton + neutron) / electron mass = 3674.836363; the ratio of quantum pointwise combinations constitutes four parts of mechanical equations:

***Item 1*:** combination of Dirac mechanical equation {M} (gravitation-quark)
$\{r_3,r_6,r_9\}^{+Z}=\{r^3\}^{+Z}$; coefficient ($C_{1+3}$, C_{1+6}, C_{1+9});

***Item 2*:** combination of Maxwell electromagnetic equation {Q} (electromagnetism-gluon)
$\{r_2,r_4,r_8\}^{-Z}=\{r^4\}^{-Z}$; coefficient ($C_{1+2}$, C_{1+4}, C_{1+8});

***Item 3*:** $\{r_5,r_{10}\}^{\pm Z}=\{r^5\}^{\pm Z}$ of thermal equation {R}; coefficient (C_{1+5}, C_{1+10});

***Item 4*:** combination of geometric sample space in photon equation {C}
$\{r_0 \sim r_{11}\}^{0Z}=\{r^2\}^{0[Z-1}=\{C^2\}^{0[Z-1]}=C^2$; coefficient ($C_{1+0}\sim C_{1+11}$)

3.2 Interaction of pointwise gauge field

Traditional gauge field belongs to the contribution of quantum equilibrium state, and the gauge invariance is determined by $\{I\ Qe\psi^-\gamma_\mu\psi A_\mu\}\ \alpha$. Uniform and non-uniform quantization pointwise is introduced to make $\{M_g\}$ (gravitational mass), {Q} (electromagnetic charge), {C} (photon) and {R} (thermal particle) a self-consistent whole through the equivalent replacement of 11-dimension $\{r_1\sim r_{11}\}^Z$) sample space, called pointwise gauge field.

(1) Gauge field of interaction between gravitation

The mechanical characteristics of normal gravitational field: entangled state converges towards the boundary: the central force is strong while the boundary force is weak, and the center of sphere is collected with the gravisphere of each level and becomes the sub-unit of the gravitation quantum. and the gravitational constant: G_N =**6.6726**$\times10^{-11}N\cdot m^2/kg^2$;

$$\{I\ Mg\ e\psi^-\gamma_\mu\psi D_\mu\}G_N \approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z = M\{r^3\}^{+Z};\quad (K=+1);$$

(2) Gauge field of electromagnetism and electromagnetic field

Mechanical characteristics of normal electromagnetism: (entangled state diverges towards the boundary), the central force is weak while the boundary force is strong, and the energy (is like donut). Therefore, the planet boundary and the spherical surface are collected with ionized layer and become the electromagnetic unit state. and the electromagnetic interaction coupling constant

k =**1.380658**$\times10^{-23}J\cdot m^2\cdot k^{-1}$;

$$\{I\ Qe\psi^-\gamma_\mu\psi C_\mu\}k \approx L[\psi(x_k),A_\mu(x_k)]^Z = Q\{r^4\}^{-Z};\quad (K=-1);$$

(3) Gauge field of interaction between gravitation and electromagnetic field

$$\{I\ e\ \{M_\beta\psi^-\gamma_\mu\psi G_\mu\}G_k \approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z = M_\beta\{r^{3+4}\}^{0Z};\quad (K=0);$$

(4) Gauge field of interaction between electromagnetism and quark field

$$\{I\ e\ \{M_w\psi^-\gamma_\mu\psi A_\mu\}\alpha_w \approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z = M_w\{r^4+r^2+r\}^{0Z};\quad (K=0);$$

(5) Gauge field of interaction between gluon and quark field

$$\{I\ e\ \{M_\beta\psi^-\gamma_\mu\psi A_\mu\}\alpha_s \approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z = M_p\{\ r^4+(r^2+r)\ \}^{0Z};\quad (K=0);$$

(6) Gauge field of interaction between quark and quark field

The mechanical characteristics of normal strong force field: entangled state diverges towards the boundary, the central force is weal while the boundary force is strong, (spinning+ radiation + vibration) and the surface is collected with quark layer, constituting a strong force unit of "quark confinement". The force coupling constant $\alpha_s=G_s^2/4\pi\hbar c$;

$$\{I\ e\ \{M_s\psi^-\gamma_\mu\psi A_\mu\}\alpha_s \approx L[\psi(x_s),A_\mu(x_s)]^Z = M_s\{2\times(r^2+r)\ \}^{-Z};\quad (K=-1);$$

(7) Gauge field of interaction between gluon and gluon field

The mechanical characteristics of normal gluon field: entangled state diverges towards the boundary, the central force is strong while the boundary force is weak, the vibrations of two gluons are perpendicular to each other and the gluon field spins, vibrates and radiates as per the bi-directional plane, and the weak interaction coupling constant includes: $\alpha_w=g_s^2/4\pi\,\mathbf{hc}\,\sin^2Q_w$; upper and lower spinning $\{r^{\pm2}\}$;

$$\{I\,e\,\{Q_w\psi^-\gamma_\mu\psi A_\mu\}\alpha_w\approx L[\psi(x_w),A_\mu(x_w)]^Z=Q_w\{2\times(r^2+r^2)\}^{+Z};\quad (K=+1);$$

(8) Gauge field of interaction between gravitation and gluon field

$$\{ie\psi_q^{-\{}Q\psi^-\gamma_\mu\psi_qB_\mu\}G_g\approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z=Q\{r^3+(r^2+r^2)\}^{0Z};\quad (K=0);$$

(9) Gauge field of interaction between gravitation and quark field

$$\{ie\psi_q^-Q\gamma_\mu\psi_qE_\mu\}G_g\approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z=Q\{r^3+(r+r^2)\}+^Z;\quad (K=+1);$$

(10) Gauge field of interaction between electromagnetism and gluon field

$$\{I\,e\,\{M_w\psi^-\gamma_\mu\psi A_\mu\}\alpha_w\approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z=M_w\{r^4+2\times r^2\}^{-Z};\quad (K=-1);$$

(11) Gauge field of interaction between thermodynamic fields

The eight-dimension (stage) form with three-dimension (stage) neutrino proved in the experiment [p282] means that the optical particle may be 11-dimension combination or decomposition and interacts with the thermal force ion. The mechanical characteristics: the central force and the boundary force in the discrete state are pairing, and Z =±(S±N±N±p±5)

Thermal force constant $\boldsymbol{\sigma}=\mathbf{5.6703}\times10^{-8}\cdot m^{-2}\cdot K^{-4}$;

$$\{IR\psi_q^-Q\gamma_\mu\psi_qR_\mu\}\boldsymbol{\sigma}\approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z=R\{r^5,r^{10}\}^{\pm0Z};\quad (K=\pm0);$$

(12) Optical force field of quantum

That neutral photon and neutrino can achieve 11-dimensioanl diversified form has been proved by the experiment [p282]. The mechanical characteristics: the central force and the boundary force in the discrete state are pairing, andZ =±(S±N±N±p±11);

Optical force constant: to be determined;

$$\{I\,e\,Q\psi^-\gamma_\mu\psi F_\mu\}\mathbf{C_g}\approx L[\psi(x_\beta),A_\mu(x_\beta)]^Z=\{r^1\sim r^{11}\}^{\pm0Z};\quad (K=0);$$

$$\{r_0^{\,2}\}^{\pm(Z\pm11)}=\{MC^2\}=hu;$$

The sample spaces of interaction of the above mechanical field form pointwise gauge field $L\{\psi(x),A_\mu(x)\}^Z$ and introduce the combination coefficient and force coupling constant of each level:

$$\begin{aligned}L\{\psi(x_\beta),A_\mu(x_\beta)\}^Z&=\{\psi^-(\gamma_\mu(d^Z/dx^Z{}_{\mu\beta})+m_\beta)\psi-(1/4)F_{\mu\nu\beta}F_{\mu\upsilon\beta}\pm IM_\beta\psi\gamma_\mu\psi^-\\&\bullet[A_\mu,B_\mu,C_\mu,D_\mu,E_\mu,F_\mu,G_\mu,H_\mu]\}^Z\bullet[G_N,\,k,\alpha,\,\alpha_s,\alpha_w,\,G_g,\,\alpha_w,\,\boldsymbol{\sigma},\mathbf{C}_g]\\&=\sum[(C_{1+h}^{-1})^k\sum\{r_h^{11}\}^k]^Z=\sum[(C_{1+h}^{-1})^k\prod\{r_h^{11}\}^k]^Z=\{r_{0h}^{11}\}^Z;\end{aligned}\qquad (2)$$

There are no mechanical elements based on circle logarithm-relativistic structure. Whether force coupling constants (G_N, k,α,α_s,α_w, G_g, α_w, $\boldsymbol{\sigma}$,$\mathbf{C}_g$) and interaction constants [A_μ,B_μ,C_μ,D_μ,E_μ,F_μ,H_μ,R_μ] exist does not affect the calculation process.

Note: "d" refers to partial derivative mark, which can be solved by integral equation.

Derivation 1: **homology circle logarithm (the first gauge invariance)**

Homology circle logarithm= sub-item field/total item field=1

$$(1-\eta_h^{\,2})^Z=L[\psi(x_\beta),A_\mu(x_\beta)]_h/L[\psi(x_\beta),A_\mu(x_\beta)]_H=\sum\sum(1-\eta_{h[ijk]}^{\,2})^Z=\sum\prod(1-\eta_{h[ijk]}^{\,2})^Z$$

$$=\begin{vmatrix}(1-\eta_{h1}^{\,2})^{K(Z\pm0)} & 0 & 0\ldots0\ldots0\\ 0 & (1-\eta_{h2}^{\,2})^{K(Z\pm1)} & 0\ldots0\ \ldots0\\ \cdots\cdots & & \\ 0 & 0\ldots & (1-\eta_{hp}^{\,2})^{K(Z\pm p)}\ \ldots0\\ 0 & 0\ldots\ 0\ldots & (1-\eta_{hq}^{\,2})^{K(Z\pm q)}\end{vmatrix}=\begin{vmatrix}\{L[\psi(x),A\mu(x)]_h/L[\psi(x),A\mu(x)]_H\}^{K(Z\pm0)}\\ \{L[\psi(x),A\mu(x)]_h/L[\psi(x),A\mu(x)]_H\}^{K(Z\pm1)}\\ \{\cdots\}\\ \{L[\psi(x),A_\mu(x)]_h/L[\psi(x),A_\mu(x)]_H\}^{K(Z\pm p)}\\ \{L[\psi(x),A_\mu(x)]_h/L[\psi(x),A_\mu(x)]_H\}^{K(Z\pm q)}\end{vmatrix}$$

$$=(1-\eta_{h1[i,jk]}^{\,2})^Z+(1-\eta_{h2[i,jk]}^{\,2})^Z+\ldots+(1-\eta_{hp[i,jk]}^{\,2})^Z+\ldots+(1-\eta_{hq[i,jk]}^{\,2})^Z$$

$$=\{0\sim(1/3)^2,(2/3)^2,(2/3)^2\}^{\{0\sim1\}Z}=\{0\sim1\}^{\{0\sim1\}Z};\qquad (3)$$

It reflects quantization of various combinations of pointwise and their distribution rules.

Where: $(1-\eta_{h[ijk]}^{2})^{Z}$ refers to the mapping (projection) of circle logarithm in the three-dimensional [ijk] coordinates.

Derivation 2: **isomorphism circle logarithm (the second gauge invariance)**

Isomorphism circle logarithm: reciprocal (k=-1) gauge mean field/ positive number (k=+1) gauge mean field=$\{0,1/2,1\}^{Z}$;

$$(1-\eta_{[ijk]}^{2})^{Z}=\{L[\psi(x_0)^{-1},A_\mu(x_0)^{-1}]^{K[Z-1]}_{[ijk]}\cdot L[[\psi(x_0)^{+1},A_\mu(x_0)^{+1}]^{K[Z+1]}_{[ijk]}\}$$

$$=\begin{vmatrix}(1-\eta^2)^{K(Z\pm0)} & 0 & 0\ldots0\ldots0\\ 0 & (1-\eta^2)^{K(Z\pm1)} & 0\ldots0\ \ldots0\\ \ldots\ldots & & \\ 0 & 0\ldots & (1-\eta^2)^{K(Z\pm p)}\ \ldots0\\ 0 & 0\ldots & 0\ \ldots\ (1-\eta^2)^{K(Z\pm q)}\end{vmatrix}=\begin{vmatrix}\{L[\psi(x),A\mu(x)]_{0h}/L[\psi(x),A\mu(x)]_0\}^{K(Z\pm0)}\\ \{L[\psi(x),A\mu(x)]_{0h}/L[\psi(x),A\mu(x)]_0\}^{K(Z\pm1)}\\ \{\cdots\}\\ \{L[\psi(x),A_\mu(x)]_{0h}/L[\psi(x),A_\mu(x)]_0\}^{K(Z\pm p)}\\ \{L[\psi(x),A_\mu(x)]_{0h}/L[\psi(x),A_\mu(x)]_0\}^{K(Z\pm q)}\end{vmatrix}$$

$$=\sum(1-\eta_{[ijk]}^{2})^{Z}=\prod(1-\eta_{[ijk]}^{2})^{Z}$$

$$=(1-\eta_{[ijk]}^{2})^{+Z}+(1-\eta_{[ijk]}^{2})^{0Z}+(1-\eta_{[ijk]}^{2})^{\pm0Z}+(1-\eta_{[ijk]}^{2})^{-Z}=\{0\sim1/2\sim1\}^{\{0\sim1\}};\qquad(4)$$

It reflects isomorphism topological property of various combinations of pointwise.

Derivation 3: **limit (topological phase transition point) valve.**

It can be obtained from Definition 3, 4 and 6 that

$$(1-\eta^2)^{Z}=\sum[(C_{1+h}^{-1})^{k}\sum\{L[\psi(x),A_\mu(x)]\}^{k}]^{Z}/\sum\{L[\psi(x_0),A_\mu(x_0)]\}^{k}]^{Z}$$

$$\approx\sum[(C_{1+h}^{-1})^{k}\prod\{L[\psi(x),A_\mu(x)]\}^{k}]^{Z}/\prod\{L[\psi(x),A_\mu(x)]\}^{k}]^{Z};$$

$$=\sum(1-\eta^2)^{+Z}+\sum(1-\eta^2)^{0}+\sum(1-\eta^2)^{-Z}\qquad(5.1)$$

Including:

$$(1-\eta^2)^{0Z}=\sum(1-\eta^2)^{+Z}+\sum(1-\eta^2)^{-Z}=\{0,1\};$$

$$\sum(1-\eta^2)^{+Z}\cdot\sum(1-\eta^2)^{-Z}=\{0,1\};\qquad(5.2)$$

It can be obtained from simultaneous equations of Equation (5.2) that

$$(1-\eta^2)^{Z}=\{(0,1/2,1)^{(0,1/2,1)}\}^{Z};\qquad(5.3)$$

3.4 Relativistic structure—pointwise gauge field

Including:

$$L[\psi(x_\beta),A_\mu(x_\beta)]^{Z}=(1-\boldsymbol{\eta}^2)^{Z}L[\psi(x_0),A_\mu(x_0)]^{Z};\qquad(6)$$

$$L[\psi(x_0),A_\mu(x_0)]^{Z}=\{MC^2\}^{Z}=MC^2=\{h\upsilon\};\qquad(7)$$

Unified: $$E=(1-\boldsymbol{\eta}^2)^{Z}MC^2\qquad(8)$$

4. Solution to and application of pointwise mechanical equation

To sum up, 11-dimension equation of pointwise quantum mechanics consists of four infinite dimension integral-differential equation or polynomial equation, which can be respectively used for solution. Boundary conditions

$$D=\{^{KS}\sqrt{D}\}^{Z}=L[\psi(D),A_\mu(D)]^{Z}$$

can form combination of four sub-items, which are respectively:

(parallel combination) $D=D_A+D_B+D_C+\cdots=(^{KS}\sqrt{D})_A+(^{KS}\sqrt{D})_B+(^{KS}\sqrt{D})_C+\cdots;$

Or: (serial combination) $D=D_A\cdot D_B\cdot D_C\cdot\cdots=(^{KS}\sqrt{D})_A\cdot(^{KS}\sqrt{D})_B\cdot(^{KS}\sqrt{D})_C\cdot\cdots;$

Boundary conditions constituting pointwise gauge field

（Entangled state）$(^{KS}\sqrt{D})^{Z}=\{r^{11}\}^{Z}$: （Discrete state） $(D_0)^{Z}=\{r_0^{11}\}^{Z}$

$\{r^3\}^{+Z}$ (gravitation equation)

$+\{r^4\}^{-Z}$ (electromagnetic force equation)

$+\{r^2\}^{+Z}$ (quantum wave equation and gauge field)

$+\{r^5\}^{[0,\pm0]Z}$ (thermodynamic equation); (9)

(1) *Integral-differential equation of all sub-items:*

$$F\{ X \pm (^{KS}\sqrt{D})\}^{[Z]} = L[\psi(x_\beta),A_\mu(x_\beta)] \pm L[\psi(D),A_\mu(D)]^Z$$
$$= Ax^{K(Z\pm0)} + Bx^{K(Z-1)}D_0^{K(Z+1)}+\ldots+Px^{K(Z-p)}D_0^{K(Z+p)}+\ldots+ Qx^{K(Z-q)}D_0^{K(Z+q)}\pm(^{KS}\sqrt{D});$$
$$= \{(1-\eta^2)^{K(Z\pm0)}+ (1-\eta^2)^{K(Z\pm1)}+\ldots+(1-\eta^2)^{K(Z\pm p)}+\ldots+(1-\eta^2)^{K(Z\pm q)}\}$$
$$= (1-\eta^2)^Z \{0,2\}^Z \{D_0\}^Z \quad ; \tag{10}$$

(2) *In integral-differential equation,*
we can convert it into primitive function through : $\{2\}^{[N+\Delta N]}$

$$F\{ X\pm(^{KS}\sqrt{D})\}^{[Z]} = F\{ X\pm(^{KS}\sqrt{D})\}^{[Z\pm\Delta]}\cdot\{2\}^{[N\pm\Delta N]}; \tag{11}$$

(3) *Entangled state or discrete state* $(1-\eta^2)^Z=(^{KS}\sqrt{D})/\{D_0\};$ (12)
When: $(1-\eta^2)^Z=1$; it indicates discrete state.
(4) *Sum of combination coefficients*: $(1-\eta^2)^Z=\{2\}^Z$; (13)
4.1 Solution to high-dimension integral-differential equation
Including: N-order integral-differential equation $L[\psi(x),A_\mu(x)]^{[Z-N]}$, boundary conditions $D=(^{KS}\sqrt{D})= L[\psi(D),A_\mu(D)]^Z$, Z, and the number of elements can constitute balance mechanics equation, such as Equation (10).
4.2 Discriminant: to judge the possibility of solution of Equation (10).
(1) Calculate elements of average state $\{D_0\}$ (combination of arbitrary p);

$$\{D_0\}^Z= [^{K(S\pm p)}\sqrt{(P/C_{1+p})}]^{[Z\pm p]}; \tag{14}$$

(2) Meet : $0\le(1-\eta^2) =(^{KS}\sqrt{D}) / D_0]^{[Z\pm p]}\le1$;
(3) The coefficient (A, B, C, D…) is adjusted into $\{D_0\}^{[Z\pm p]}$ and thus conforms to the sum of regularization coefficients: $\sum C_{1+p} = \{2\}^{[Z\pm p]}$;
Where, the numerical value of elements combination can be incomplete but the coefficient can not be null.
4.3 Solution:
Solve $\{D_0\}$ through the simplest sub-item in integral-differential Equation (16):
Including: $\{D_0\}^{[Z-1]}=(B/C_{1+1})$,

$$(1-\eta^2) = (^{KS}\sqrt{D})\}^{[Z-1]} / \{D_0\}^{[Z-1]};$$
$$(1-\eta_h^2)^{[Z-1]} =\sum(1-\eta_h^2)^{[Z-1]} =1; \tag{15}$$

General equation: $\{D_0\}^{[Z-p]}=(P/C_{1+p})$,

$$(1-\eta^2)^{[Z-p]} = (^{KS}\sqrt{D})\}^{[Z-p]} / \{D_0\}^{[Z-p]}$$
$$(1-\eta_h^2)^{[Z-p]} =\sum(1-\eta_h^2)^{[Z-p]} =1; \tag{16}$$
$$\{X_{h1}\}^{[Z-1]}=[(1-\eta^2) / (1-\eta_{h1}^2)]^{[Z-1]} (B/C_{1+1}); \tag{17}$$
$$\{X_{hp}\}^{[Z-p]}= [(1-\eta^2) / (1-\eta_{hp}^2)]^{[Z-p]}(p/C_{1+1}); \tag{18}$$

By virtue of Equation (17) and (18), we can easily know the exact solution of every element or several combined elements in pointwise equation.

5. Engineering application

5.1 Universe cyclic evolution:
Process of $(1-\eta^2)^{+Z}$ (convergence of black hole)→$(1-\eta^2)^{0Z}$ thermal wormhole (thermal topological phase transition, thermal nuclear fission and cosmic bump)"→$(1-\eta^2)^{-Z}$ (expansion of white hole and newborn baby)→$(1-\eta^2)^{0Z}$ cold wormhole (cold topological phase transition and cold nuclear fission)→$(1-\eta^2)^{+Z}$ (extinction of black hole and celestial body) ;
5.2 Description of energy asymmetry:
The universe vacuum will excite and produce extremely asymmetry energy, which can be mathematically proved by polynomial equation. According to $(1-\eta^2)^{0Z}$, topological phase transition, vacuum excitation and Higgs boson, we can know:

$$(1-\eta^2)^{+Z}\{MC^2\} \rightarrow (1-\eta^2)^{-Z}\{MC^2\};.$$

Two parallel polynomials of 11-dimension energy particle $\{x\}^{[Z-11]}=\{x\}^{[Z-6]}+\{x\}^{[Z-5]}$ which are

respectively substituted into (6 prime numbers: 1, 3, 5, 7, 11 and 13) to know entangled state; (5 natural numbers: 1, 2, 3, 4 and 5) are particles in discrete state (the counting process is omitted), and thus we can know:
Mass-energy ratio: (4.758845%: 95.241155%);
Energy ratio: (1: 40.027004);
Surprisingly, the results of above-mentioned calculation are consistent with astronomy observation and test data of high-energy particle collision.

5.3 Application of vacuum excitation

According to principles of universe evolution, we can acquire a patent of super high-energy engine of "vacuum excitation". Project name: *Cold and Negative-pressure Bi-directional Vortex Blades Aero-Engine in Bi-directional Vortex*. (**ZL201410055227.0**), date of authorization is May 2016. It is named man-made mini universe heat engine.

6. Conclusion and prospect

The fusion of traditional quantum gauge field and relativity is aimed to reveal the distinction between uniformity and non-uniformity. Through relativity structure (circle logarithm), the interference of quality and other elements is eliminated, thus realizing abstract four-arithmetic operation without specific content.

Geometric sample space of pointwise quantum {r} and $\{r^2\}$ as well as power and thermodynamic system generate various accelerated speed and energies which are reflected as duality of "wave and particle" and "speed and energy". Through MC^2= M{C…c}= h{U…u}, integrate them into a broader entirety — relativity structure. This algorithm has advantages of conciseness, self-consistency, accuracy, and zero error. It can be generally used in fields like physics, astronomy, and mechanics.

Existing problems: coupling constants in mechanics like gravitation, heating force, nuclear force and electromagnetic force as well as phase transformation point in topology remain to be confirmed and studied.

Of course, there will be inevitable defectiveness in relativity structure; criticizing and improvement suggestions are favorably received. Exchange, popularization, and application by more experts and scholars are expected so as to innovate and develop together.

Gratitude is shown to 2016 Chinese Conference on Computational Mechanics in conjunction with International Symposium on Computational Mechanics in Hangzhou (CCCM-ISCM 2016) for making the statement that *Research and Application of Super-symmetric Element Matrix (circle logarithm)* a novel "abstract analysis" .

I would like to thank Association for Senior Scientists and Technicians in Zhejiang Quzhou for its long-term support, encouragement and blog: explore free sky (LKX0570) (more than 640 open blogs since 2009), which reflects the hard process of establishing circle logarithm-relativity structure.

I would also like to thank American academic journal, Matter Regularity, for publishing *On Light, Polynomial, Relativity, Circle Network Isomorphic Mapping -- One Hundredth Anniversary of Einstein's General Relativity and 2015 International Light Year* and other 6 papers which formed a system from 2014 to 2016. (The End)

References

[1] **Yang Guilin**, Jiang Xingfang, Ke Shanzhe. *Modern Physics*. Beijing: Science Press, third printing in July of 2006.

[2] **John Stachel, J..** editor-in-chief, translated by Fan Dainian and Xu Liangyin, *Einstein's Miraculous Year: Five Papers That Changed the Face of Physics*, Shanghai: Shanghai Science and Technology Education Publishing house, second printing in August of 2003.

[3] **Din Guangbi**. *Difficulties in Coordination Relativity Theory and Quantum Theory, 100 Scientific Puzzles of the 21st Century*. Jilin: Jilin People's Publishing House, third printing in January of 2000, P14-21.

[4] **Kline, M**. *Mathematical Thought from Ancient to Modern Times*, (Volume I, II and III) (page listed in Chinese such as 3-p353 refers to page 287-307 in Volume III), Shanghai: Shanghai Science and Technology Press, second printing in August of 2014.

[5] **Xu Lizhi**. *Berkeley's Paradox, Pointwise Continuity Concept, and Related Questions*, *Advanced Mathematics*, fifth edition of 2013, P33-35.

[6] **Einstein**, translated by Fan Dainian and Xu Liangying. *Selected works of Einstein*, Shanghai: Shanghai Science and Technology Education Publishing house, third printing in April of 1979.

[7] **Wang Yiping**. *On Light, Polynomial, Relativity, Circle Network Isomorphic Mapping — One Hundredth Anniversary of Einstein's General Relativity and 2015 International Light Year* (Chinese and English version) [America], Matter Regularity, 2015/3, P46-61.

Two-Dimensional Unsteady Flow Calculations of a Five Bladed Vertical Axis Wind Turbine

†*Nishant Mishra[1], Praveen Laws[1], Rajagopal Vinod Bethi[1], Santanu Mitra[1]

[1]Department of Mechanical Engineering, Shiv Nadar University, India

*Presenting author, †Corresponding author: nishant.mishra@snu.edu.in

Abstract

With simple construction and mechanism, Vertical Axis Wind Turbine (VAWT) is the most promising renewable energy resource in the modern era. A two-dimensional analysis on the aerodynamic behavior and performance of five bladed VAWT is investigated. The study is systematically focused mainly on parameters such as turbulence modeling, temporal discretization and mesh refinement to figure out the certainty of the unsteady flow behavior and stability of the scheme. All these parameters have a direct impact on the exactness of the simulation result, which is analyzed and presented. Later on, CFD validation is done by comparing with experimental five bladed data based on the aerodynamic behavior and performance. Lastly, it is observed from the numerical analysis that turbulence modeling, time step and mesh refinement has a high impact on the accuracy of the simulation results in understating the unsteady flow phenomena around and through the vertical axis wind turbine. It is very important to develop a fluid-solid integrated model towards understanding of energy production by varying different parameters of wind machine. It is believed that the present 2D work can be extended to 3D without much difficulty and can produce useful results before developing a prototype and making it viable for green energy production.

Keywords: Vertical axis wind turbine, Turbulence Modeling, Time Stepping, Mesh Refinement, performance.

Introduction

To comprehend unsteady flow physics of Darrieus turbine the free vortex approach turn out to give a better picture to figure out the wake formulation around the rotor. These wake formations plays a vital role for calculating the force and moment on the blade. Forces on the turbine are examined by the pressure acts on the blades and the moment using wake vortices impulse. Further this method is suitable only for rotor with single straight bladed with no inclusion of stall effect and extravagant [1]. Later on a model which incorporates dynamic stall effect in free vortex wake has been modeled to validate three dimensional vertical axis wind turbine. A time marching schemes is developed which computes the strength of the bound vortex and wake using Kutta-Joukowski law. In order to probe the effect of dynamic stall an indicial stall model is used which derives the stall effect with a physical approximation [2]. To figure the improved efficiency of vertical axis wind turbine a high fidelity model is developed to examine unsteady, 3D, viscous flow around the rotor. Solving RANS equation using spectral discretization, the model tend to study about solidity and tip speed. The exactness of the system relay on Fourier representation in time because unsteady flow always remains periodic thus gaining computational time by just extracting few number of modes [3].

Also, the complex flow behavior around the turbine due to unsteady flow field and varying angle of attack causes dynamic stall on the blades at low tip speed ratio tends to the formation of eddies. This hysteresis eddies traced from the blade surfaces are due to the impact of vortex shedding, main reason behind the lift reduction in the trailing edge and strong increase of drag

[4]. Further, to simply the physics of flow around the turbine using discrete vortex method, with an assumption that the boundary layer separation is not considered and applicable only for higher tip speed ratio. The method is built upon conformal maps, later solved using Fourier transformation with kutta condition to get the final solution. The model can be further developed to calculate additions effects such as skin friction drag, number blade effect and unsteady calculation of force [5].

The rotation of the turbine in CFD is imitated using the sliding mesh technique, data transfers transpire between the two interfaces. Suitable coupling through interface is made between the steady fixed domain and the unsteady rotating domain. Theses coupling data will be updated at each time step and stable exchange of fluxes between the two domains. When the turbine rotates the blades experience different angle of attack, so at each time step new set of transient aerodynamic parameters will be updated and calculated at each time step [6].The authentication of CFD data are validated against the flow field observed from a Particle image velocimetry (PIV), the system measure the upwind and downwind stall to predict the power coefficient [7].

CFD simulation on the effect of increasing the number of blades show that five bladed vertical axis wind turbine provides higher power coefficient at lower tip speed ratio which is of 20% greater than that of a three and four bladed turbine. The frequency of oscillation torque get reduced leads to an increase in torque coefficient at lower tip speed ratio. To add on increase in blades achieves a marginal reduction of peak radial force (Fr) which may cause structural damage [8].

Another major problem which hits the vertical axis wind turbine is its no self-starting aerodynamics, 2D weak coupling model is generated to analyze the interplay between the rotor and the inducing air. Discussion is made on parameters such as solidity and fixed pitch angle which reveals solidity in the range of $\sigma = 0.6 - 1.0$ provides a minimum self-start at t=0.8s. Delay in the start is influenced by the separated flow vortex with minimum lift [9]. Coupled pressure based solvers with k-ε turbulence modeling turn out to be comparably efficient solver setting which favors with accurate result in modeling periodic flow field. So considering all the research fact above understanding and characterizing the unsteady flow nature around the turbine is complex [10]. The fluid dynamics around the blade surface falls on the effect of tip speed ratio, affected by flow separation and wake created by its own and also from other rotating blades. All these unsteady physics get coupled and affects the performance of the turbine. Understanding all these above discussed details research in this works is spotlighted on culling most advisable time step, mesh refinement and turbulence modeling to harvest an exceptional performance of the turbine. Finally all these desired properties are conjoined to estimate the self-starting behavior and power coefficient actualize from different wind speed and tip speed ratio.

Methodology

Design Parameters

A five bladed 2D model of vertical axis wind turbine was designed in SolidWorks and reviewed in ANSYS Fluent, 2D model was preferred to cut down the computation time and to clearly understand the unsteady aerodynamics around the turbine blades. Numerical analysis is carried out with the turbine blades and neglecting hub and the supporting arm. Straight-bladed Darrieus rotor is characterized byNACA0018symmetric blades with a chord length(c) of 0.2 m, diameter of the turbine (D) 2.5m with an aspect ratio of 12.5 has been taken into account [11]. The geometry of the turbine is branched by a rectangular outer domain and circular inner domain. The rectangular domain host the inlet, far field wall and outlet, rotor is place in the center of the domain. Both inlet and outlet are placed at a distance of 25m from

the center of the rotor in order to capture the wake vortices formulation and the width of the domain is 20m. The inner circular domain with an interface boundary condition has a diameter of 3m mounts the turbine blades. To mimic the rotation of the turbine sliding mesh technique in FLUENT is adopted, which is a time-periodic technique suitable of analyzing rotating parts. The model is developed ideal for moving frame of reference problems or rotating object in a domain. The rotor mesh slides with respect to the fixed mesh surface in the interface region. The mesh from both the domains should match each other to transfer data through interpolation [12].

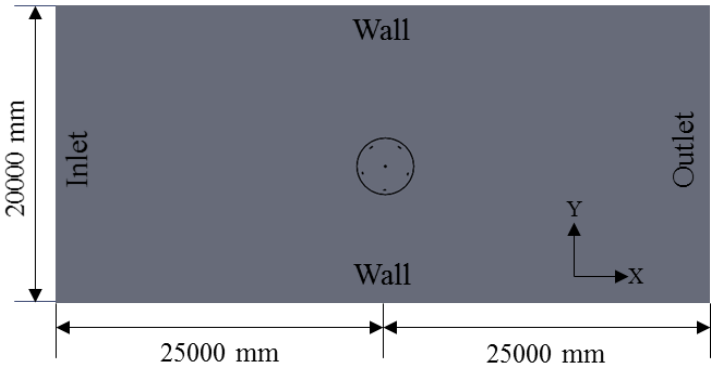


Figure 1.Turbine geometry and boundary condition

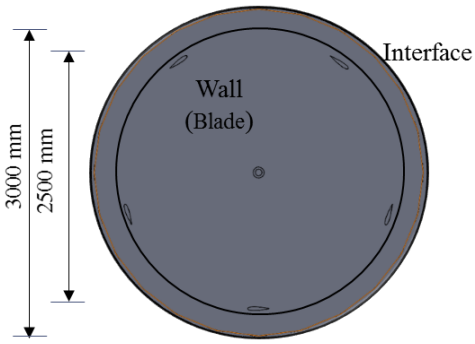


Figure 2. Rotating domain geometry

Table 1. Specification of the geometry

Design Aspects	
Number of blades	5
Domain	50 X 20 m
Blade position	72 degree apart
Chord Length (c)	0.2 m
diameter of the turbine (D)	2.5m
Interface	3m

Simulation and Discretization Parameters

A uniform velocity distribution is set along the x-direction at the velocity inlet. No-slip (U=0) condition obliged on the blade surface, top and bottom of the domain and in the outlet pressure outlet (Δp=0) boundary condition is imposed. Standard atmospheric air density and viscosity measures were chosen for the flow properties. A 2D, coupled, pressure based, transient, incompressible solver setting has been urged for the study. Pressure-Based solver has been applied with Gauss node based Interpolation Methods is used to discretize pressure gradients with second order accuracy which is more suitable for triangular mesh, same discretization is adopted for the turbulence equation. The pressure-velocity is set as coupled algorithm implicit discretization, which is highly effective scheme towards single phase flow and performs superior than the other segregated solvers [13]. The scheme is potential strong while dealing with large time step.

Turbulence modeling and solver setting

The Navier-strokes equation is solved by finite volume method, to measure the unsteady flow properties around the turbine blades. Selecting an equitable turbulence model to figure an

accurate result depends on the domain, nature of the flow and shape of the model. A fully developed turbulent flow encounters with high velocity fluctuation, infinite degree of freedom, highly nonlinear, three dimensional and riotous [14].So selecting an elegant model reduces computation time and benefits accurate result. Classical turbulence model based on Reynolds Averaged Navier-Stokes (RANS) advisable for 2D flow simulation tabbed for the study are: Standard, RNG, Realizable k-ε turbulence model Standard k-ω turbulence model and SST model. The above discussed model are two equation models except SST model with two PDEs to represent the turbulent kinetic energy and the turbulent dissipation rate with a turbulence intensity of 0.5% [15]. So five set of simulation are formulated and analyzed to resolve the exact model.

Time step sensitive review

Subsequently three case were setup to optimize the sensitivity of time step in which the flow property change with time in a simulation. Time step is more sensitive while solving the governing equation for rotating motion and for turbo machinery [16].Rotating motion hatch unsteady flow and is time periodic, so acceptable time step has to been chosen to get preferred convergence, lesser computational time and good accuracy of results. Taking this in account three different time steps has been inspected $\Delta t = 0.001$, $\Delta t = 0.003$, $\Delta t = 0.005$. And the Max iteration for each time step is set to 20, finally appropriate time step is chosen for getting accurate result and lesser computational resource.

Grid Refinement Survey

Complex aerodynamic flows around the turbine can be solved only with reliable mesh generation which adversely won't affect the factualness of the investigation. Improper grid resolutions leads to numerical error and incorrect results in the simulation [17], so mesh quality definitely has a huge influence on simulation accuracy. Grid dependency test was carried for three different mesh qualities: course with one lakh elements, medium with five lakhs element and fine with one million elements. Y-plus value less than 30 has been adopted for all the study with wall functions. To accurately visualize the wake formulation and flow separation from the turbine blades dense fine unstructured mesh triangular mesh refined on the surface of blades. The region of detect are both the blades and interface were meshed with edge sizing while the outer domain with structured rectangular mesh. Maximum skewness around 80 and aspect ratio around 23 has been maintained in the study.

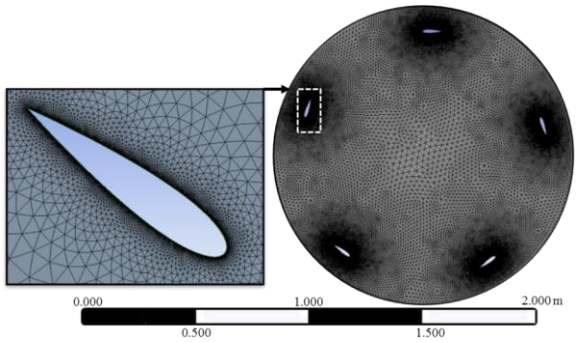


Figure 3. Detailed mesh discretization around the blade and rotor interface

Table 2. Detail of grid elements and simulation time

Refinement	Number of elements	Simulation time
Coarse	169430	6 h and 43min
Medium	510482	13h and 38 min
Fine	1036249	25 h and 54 min

Result and discussion

All the above observations are examined with a speed of U_∞ = 5m/s and tip speed ratio λ = 3, the reason to pick the particular tip speed is that from several published data's the performance of the turbine is hit its peak at between λ = 3. To interpret the correct selection of grid is a huge obstacle, to point out the ideal mesh quality three different mesh refinements were made namely coarse, medium and fine. Finally the primary goal was to select a pertinent mesh resolution to reduce cost, resource of computation and mainly veracity result. Figure shows torque generated along the unit length of single blade for one complete revolution.

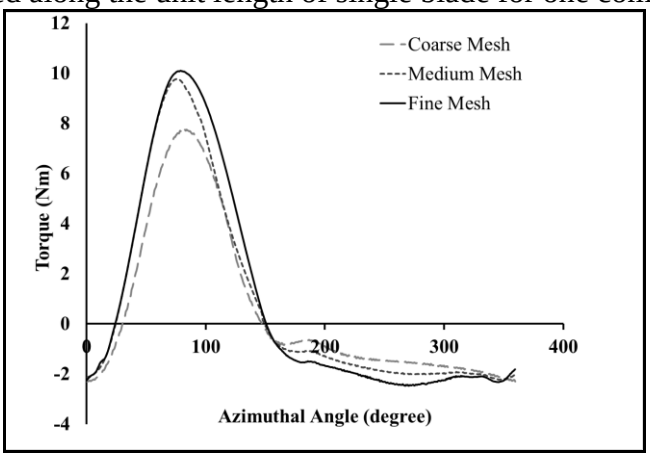


Figure 4. Grid refinement survey for 2D turbine

Varying results obtained from the three different mesh models, comparing all the three meshes both fine and medium mesh show very close results with little difference. The coarse grid provides much vague results which is unsuitable for further simulation. But better arrangement has been seen in both medium and fine mesh form 0 degree to 60 degree later on little dissimilarity subsequently. In general from the both the medium and fine mesh figures the more accurate positive peak torque between 0^o to 150^o, further negative torque below 150^o till 360^o. Positive torque falls on the upstream side while negative torque is generated in the downstream. Eventually from all the three mesh study to get an accurate result from a computational analysis reckon on the number of mesh elements. From the overall view the mesh study with one million elements is consider for all the forthcoming simulation as it is accurately predicting the torque despite it takes computational time.

Time step is another variable which may perturb the numerical result for a little extend, so three different time steps were carefully chosen for the time periodic study notable Δt = 0.001s, Δt = 0.003s and Δt = 0.005s. Torque generated from unit length of the blade is observed from each time step as shown in fig 6 along unit length of the blade which points out Δt = 0.001s is suitable for the study as the generated data in that particular time step is more precise. But time step, Δt calculation for turbo machinery and rotating simulations are proposed as [16]

$$\text{Time step, } \Delta t = \frac{1}{10}\frac{\text{Number of blades}}{\text{Rotational Velocity}} \tag{1}$$

The above equation yields a time step off, Δt = 0.033s for a ω = 12rad/s and radius, r = 1.25m. Validations is simulated for the calculated time step which predicts the minimal value of peak torque and has a huge dissimilarity when compared with the other probed time steps. So for better accuracy the time step is set to Δt = 0.001s throughout the entire 2D simulation.

Five different RANS based turbulence modelling were considered in this numerical analysis namely standard k–ε model, Realizable k-epsilon model, RNG k-epsilon model, standard k–

ω modeland SST model. For this case the fine mesh and time step Δt = 0.001s were selected from the above study. Fig 6 shows the torque force generated from a single blade for one rotation for each turbulence models analyzed. From the plot its shows that both standard k–ω model and SST model under predicts the torque in the upstream flow while both Realizable k-epsilon model and RNG k-epsilon model shows a good argument with higher torque generation in the upstream flow, also lesser dissimilarities in the downstream.

From the previous study it shows both realizable k-epsilon model and RNG k-epsilon model predicts more accurate and similar results and also simulates the detailed vortices during dynamic stall[18]. So realizable k-epsilon model has been chosen as the baseline model for all the postliminary simulations. The main prospect in choosing realizable k-epsilon model is that it performs well for large boundary layer separated flows and also for swirl or rotating flow withadverse pressure gradients which can be visualized during turbine operation.

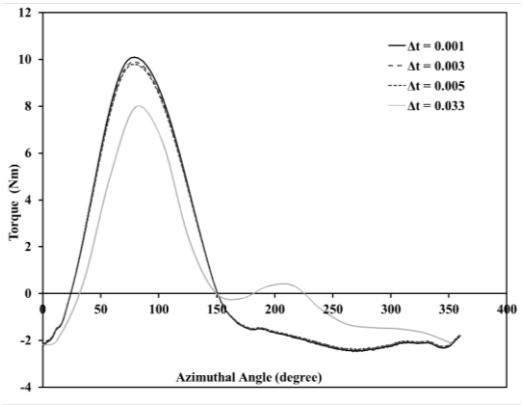


Figure 5. Time step sensitive study for the turbine

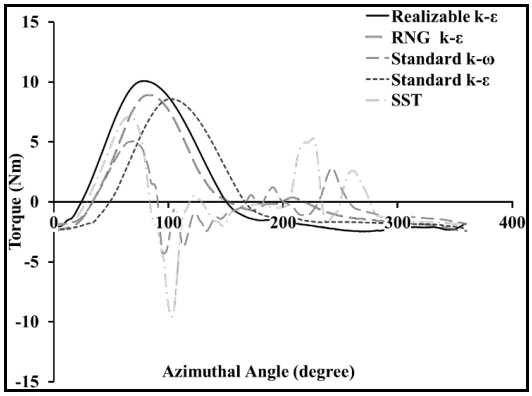


Figure 6. Turbulence modeling study for the turbine

Flow driven Characterization

Acceleration of the turbine about its own axis influenced by the induced air flow is simulated for free load condition considering the moment of inertia of the turbine, I and Aerodynamic moment, C_m generated by the incident air.

From Newton's second law of motion, the revolutionary motion of the flow drive turbine is measured as:

$$I\alpha = C_m \tag{2}$$

Where α is the angular acceleration of the turbine.

From equation (2) the revolutionary motion of the flow drive turbine is calculated as:

$$\omega = \int_0^t \frac{C_m}{I}\, dt \tag{3}$$

Where, ω angular velocity of the turbine

Similarly as the turbine starts to rotate the experience varying azimuthal angle, hence the equation is transformed to

$$\theta = \int_0^t \omega\, dt \tag{4}$$

Hence the above loop the passive rotation of the turbine for free load condition is achieved, Fluid driver structure coupling defines the torque generated by the turbine, calculating the shear stress and the pressure developed by the aerodynamic moment and also constraining the turbine rotational direction.

The torque generated by the turbine under free load condition for all blades and single blade is show in fig 7 and fig 8. The term free load defines the capability of turbine to accelerated about its own axis by the impact of the induced air without any counter torque been supplied [19]. The moment of inertia and mass of aluminum is considered for the study with an

induced flow velocity of U_∞= 5m/s with passive motion. Peak fluctuating torque around 11Nm has been obtained from the turbine during the accelerating period but after t=2s starts the ideal equilibrium period with a peak value of torque around 5Nm. Turbine attains its equilibrium state much prior if the induced flow velocity is higher but at lower wind speed the turbine poses "start and stop loop" because the turbine won't be able to overcome the inertia leads to low self-starting performance of straight bladed turbine [20]. This shows the passive model study accurately predicts the self-start efficiency of the turbine under no-condition.

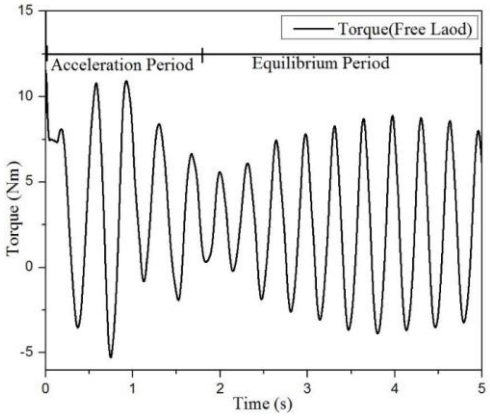


Figure 7. Flow driven torque generated by the turbine

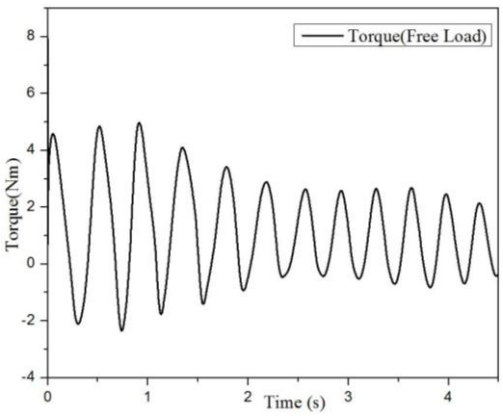


Figure 8. Flow driven torque generated on single blade

CFD validation

Realistic study of straight bladed turbines was compared against the numerical setup, the main purview was justify whether simulation falls in the right direction with that of experimental data. So more admissible empirical study has been considered, experimental setup figured by Qing 'an LI [21] was more promising and most significant for validation. The main benchmarks considered are: Straight five bladed Darrieus turbine, airfoil chord, symmetrical airfoil, diameter and solidity. The following considerations are listed below in the table 3.

Table 3. Comparison between the two turbines

Framework	2D simulation	Experimental
Number of blade	5	5
Airfoil profile	NACA 0018	NACA 0021
Airfoil chord (m)	0.2	0.265
Turbine diameter (m)	2.5	2
Blade pitch angle (β)	0 degree	10 degree
Solidity(σ)	0.8	1.325

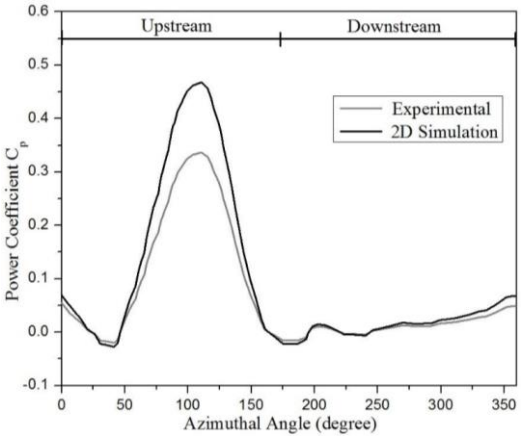


Figure 9. Power coefficient comparison between two turbines

Numerical simulation is performed with the same aerodynamic conditions applied in the experimental set up with a constant wind velocity of 8m/s and tip speed ratio 1.39. The figure 9 shows the validation for coefficient of power (C_p) versus Azimuthal angle for unit length of single blade, Comparison results of both the plot shows a promising trend with a discrepancy factor of 2, with a more similar curve. Both the experiment and numerical results show a maximum coefficient of power (C_p) at $\theta \approx 100^o$. Peak C_p of around 0.45 for the computational simulation and around 0.32 for the experimental study. The slight dissimilarity between the results are due to the effect of blade profile, blade pitch angle and also the effect of wing tip vortices visualized in the experimental study which is not observed in the 2D simulation. All the above numerical simulation is able to predict the physical behavior of the vertical axis wind turbine.

Dynamic stall

Another parameter to look into while investigating unsteady aerodynamic behavior of the turbine is to compute the lift and drag coefficients, as the turbine rotates the flow parameters such as the Reynolds number and angle of attack vary which tend to occur dynamic stall on blade surface with drag and lift hysteresis[16]. The lift and drag coefficients are probed over one rotation of the turbine for five different tip speed ratios from $\lambda = 2$ to $\lambda = 4$ as shown in figure 10 and 11. It figures that the tip speed ratio λ is directly proportional to that of lift and drag coefficients on the airfoil. As the tip speed ratio increases both lift and drag increases durably. Stall effect on the blades are characterized by vortex formation in the leading edge which leads to separation of flow and finally get detached to the trailing edge [22]. This progression generates huge hysteresis of flow with unsteady lift and drag coefficients.

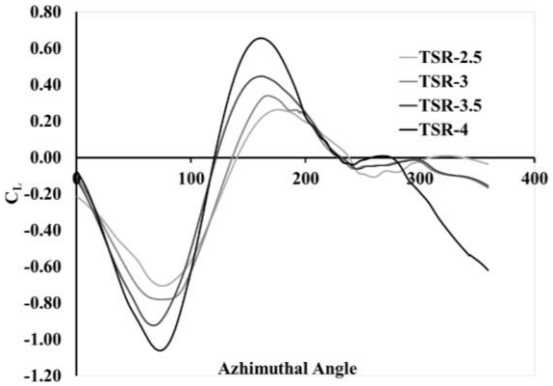


Figure 10. Lift coefficient vs. azimuthal angle for one rotation

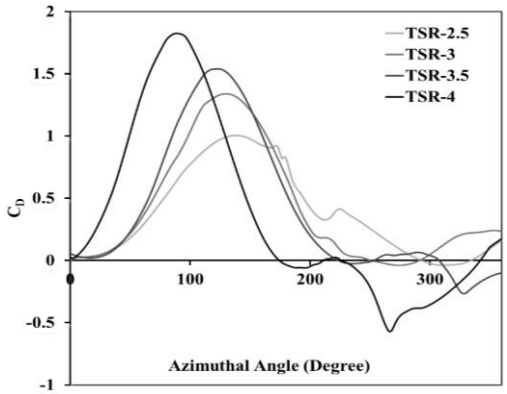


Figure 11. Drag coefficient vs. azimuthal angle for one rotation

Table 4.Stall at different TSR

Λ	$C_{L\ stall}$	$\theta_{\ stall}$
2.5	-0.69	59.956
3	-0.77	60.073
3.5	-0.96	63.438
4	-1.45	74.902

Performance Analysis

The power coefficients computed in this numerical simulation for different tip speed ratio were based on acknowledging the results obtained from the mesh, time step and turbulence

modeling research. According to simulation results fine mesh with time step $\Delta t = 0.001s$ and realizable k-epsilon model were considered for the Performance analysis. The power coefficient analysis is formulated for four different tip speed ratio with three different wind velocity U_∞ 4, 6 and 8 (m/s). A total of 12 simulations were carried out with a computational of 3 days for each simulation with 10 interactions loop per time step.

Fig 12 shows the average power coefficient obtained from three different wind flows, all the three curves follows the same trend for the considered wind speeds. The maximum power coefficient is achieved at $\lambda = 3$ for the three different incident wind velocity, with a maximum C_p around 0.3 at 8 m/s. This figures that higher the wind velocity the value of C_p increases for the particular tip speed ratio, further increase in tip speed ratio the graph shows a sudden drift in the C_p and it tend to decrease at $\lambda = 3.5$ and generates negative torque at $\lambda = 4$.

Fig 13 shows average torque coefficient at three different wind velocity as a function of Tip speed ratio. The plot shows a good argument with similar trend for all the three wind velocity, it is seen that as the tip speed increases the turbine achieves its maximum steady peak at $\lambda =$ 3and later on decreases linearly as the tip speed ratio increases. As the wind velocity increases there will be a steady increase in the turbine torque. Even here all the three curves follows the similar trend , coefficient of torque increases with TSR and later decreases after reaching a peak at $\lambda = 3$.

Table 5. Flow conditions for the performance analysis

Velocity (m/s)	U_∞	$\lambda = 2.5$	$\lambda = 3$	$\lambda = 3.5$	$\lambda = 4$
		Angular Velocity (ω) rad/s			
4		8	9.6	11.2	12.8
6		12	14.4	16.8	19.2
8		16	19.2	22.4	25.6

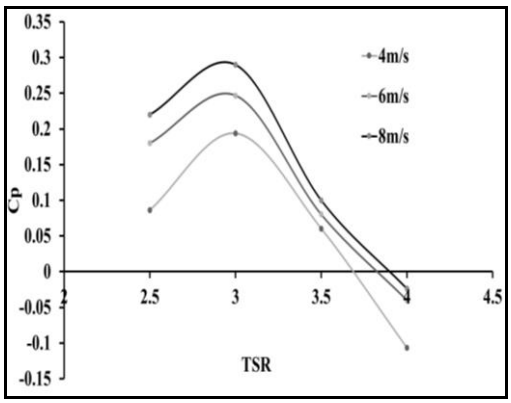


Figure 12.Power coefficient versus TSR at different wind velocity

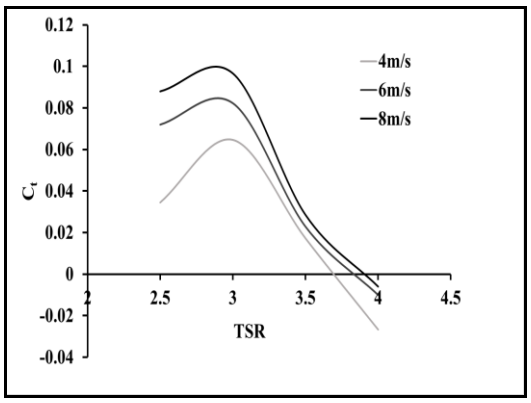


Figure 13. Torque coefficient versus TSR at different wind velocity

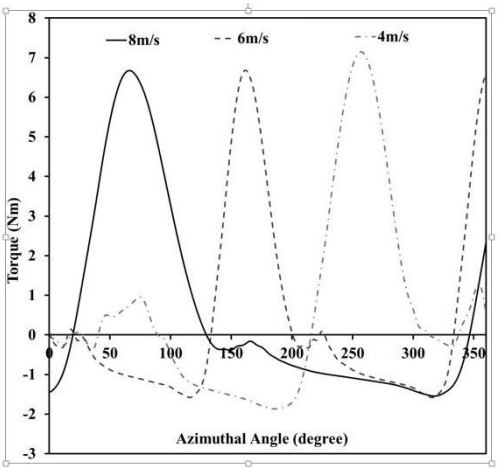


Figure 14. Torque generated at λ = 3 for three different wind speed on single revolution

Fig. 14 shows the torque obtained at λ = 3 for one complete revolution, as the velocity increase positive peak torque is achieved earlier. At 4 m/s the positive peak is obtained around 250^o but as the velocity increases the peak torque is shifted more forward so at 8m/s peak torque is obtained around 70^o. On an average the torque falls in the positive note also it states that the turbine won't be able to self-start at minimum velocity of airflow and also at certain azimuthal angle. As we have seen from the plot negative torques falls at lower induced wind flow and also minimum torque at certain azimuthal angle. Overall all the three curves follows similar trend with slight variation.

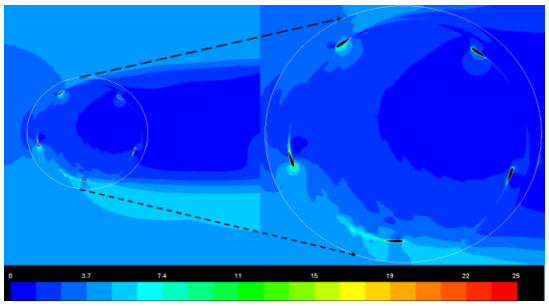


Figure 15. Velocity plot for TSR = 3 at 4m/s

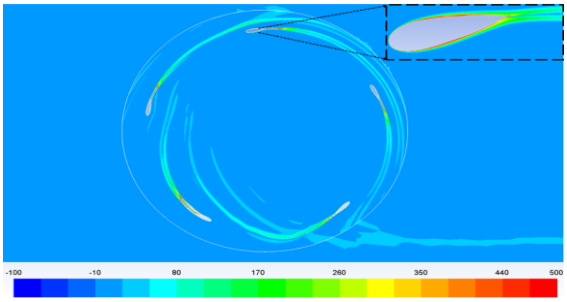


Figure 16. Vorticity magnitude for TSR = 3 at 4m/s

Conclusion

In this research a detailed examination is carried out to simulate the unsteady flow around a rotatory turbine. Characterization is performed in such a way is to select a reasonable grid resolution, convenient time step and favorable turbulence modeling for a sequential numerical simulation in order to conclude with accurate result. Aerodynamic coefficients generated in the simulation of the rotating turbine relay on the validated turbulence modeling and fine mesh resolution. Irregularities in probing the aerodynamic coefficients cause the simulation with weaken peak torque and invalid performance result. All these combined parameters predict that the wind energy harvested by the turbine lay on the upstream with net negative torque in the downstream. This also leads to the low self-staring behavior of the turbine as it poses gradual fluctuation in net torque and the turbine has to overcome the mass and inertia force acts on it. CFD validations is performed against a five bladed turbine with similar configuration shows a similar trend in result. Form the performance point of view maximum

Cp and torque is generated at higher wind speed and also at higher tip speed ratio and minimum power at low TSR which again proves that the turbine is not a self-starting.

References

[1] Wilson, R.E., (1980) Wind-turbine aerodynamics, *Journal of Wind Engineering and Industrial Aerodynamics*, *5*(3-4), pp.357-372.

[2] Dumitrescu, h. And cardos, v., (1981) A free wake method for vertical-axis wind turbine performance prediction, *Romanian journal of technical sciences-applied mechanics*, 1(54), pp.87-100.

[3] Lauffer, J.P., Carne, T.G. and Nord, A.R., (1984) Mini-modal testing of wind turbines using novel excitation No. SAND-84-0984C, CONF-840178-1, Sandia *National Labs*, Albuquerque, NM (USA).

[4] Amet, E., Maaztre, T., Pellone, C. and Achard, J.L., (2009) 2D numerical simulations of blade-vortex interaction in a Darrieus turbine, *Journal of fluids engineering*, 131(11), p.111103.

[5] Österberg, David (2010) Multi-Body Unsteady Aerodynamics in 2D Applied to a Vertical-Axis Wind Turbine Using a Vortex Method.

[6] Lain, S., and C. Osorio. "Simulation and evaluation of a straight-bladed Darrieus-type cross flow marine turbine." (2010).

[7] Edwards, J.M., Danao, L.A. and Howell, R.J., (2012) Novel experimental power curve determination and computational methods for the performance analysis of vertical axis wind turbines, *Journal of Solar Energy Engineering*, 134(3), p.031008.

[8] Castelli, M.R., De Betta, S. and Benini, E., (2012) Effect of blade number on a straight-bladed vertical-axis Darreius wind turbine, *World Academy of Science*, Engineering and Technology, 61, pp.305-3011.

[9] Zhu, Jianyang, Hailin Huang, and Hao Shen, (2015) Self-starting aerodynamics analysis of vertical axis wind turbine, *Advances in Mechanical Engineering*, 7.12: 1687814015620968.

[10] Gang, D. and Kau, W.C., (2015) Unsteady Flow Numerical Simulation of Vertical Axis Wind Turbine, *Procedia Engineering*, 99, pp.734-740.

[11] M.S.U Khalid, T.Rabbani, I Akhtar, N.Durrani, and M.S Siddiqui, (2015) Reduced-Order Modeling of Torque on a Vertical-Axis Wind Turbine at Varying Tip Speed Ratios, *J. Comput. Nonlinear. Dy*., 10, no.4, p.041012.

[12] Sánchez-Caja, A., Rautaheimo, P. and Siikonen, T., (1999) Computation of the incompressible viscous flow around a tractor thruster using a sliding-mesh technique, *In Proceedings of the 7th International Conference on Numerical Ship Hydrodynamics*, France.

[13] Fluent, A., (2009) *12.0 User's guide*, User Inputs for Porous Media, 6.

[14] Tony Saad, T., (2011) *Turbulence modeling for beginners*, University of Tennessee space institute.

[15] Wilcox, D.C., (1998) *Turbulence Modelling for CFD DCW Industries Inc*, La Canada, California.

[16] Alaimo, A., Esposito, A., Messineo, A., Orlando, C. and Tumino, D., (2015) 3D CFD analysis of a vertical axis wind turbine. Energies, 8(4), pp.3013-3033.

[17] Venditti, D.A. and Darmofal, D.L., (2003) Anisotropic grid adaptation for functional outputs: application to two-dimensional viscous flows. *Journal of Computational Physics*, 187(1), pp.22-46.

[18] Durrani, N., Hameed, H., Rahman, H. and Chaudhry, S., (2011) A Detailed Aerodynamic Design and Analysis of a 2-D Vertical Axis Wind Turbine Using Sliding Mesh in CFD, *In 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition (p. 541)*.

[19] Le, T.Q., Lee, K.S., Park, J.S. and Ko, J.H., (2014) Flow-driven rotor simulation of vertical axis tidal turbines: A comparison of helical and straight blades, *International Journal of Naval Architecture and Ocean Engineering*, 6(2), pp.257-268.

[20] Liu, Z., Qu, H. and Shi, H., (2016) Numerical Study on Self-Starting Performance of Darrieus Vertical Axis Turbine for Tidal Stream Energy Conversion. *Energies*, 9(10), p.789.

[21] Qing'an LI, Furukawa, K. and Yamamoto, M., (2014) Aerodynamic Models and Wind Tunnel for Straight-bladed Vertical Axis Wind Turbines, *IOSR J Eng*, 4(06), pp.35-44.

[22] Nobile, R., Vahdati, M., Barlow, J. and Mewburn-Crook, A., (2011) November. Dynamic stall for a vertical axis wind turbine in a two-dimensional study, *World Renewable Energy Congress-Sweden*, Linköping, Sweden No. 57, pp. 4225-4232.

Studying Mechanical Properties of Wing Blade with Hole Based on Reverse Engineering

***Heng Cai[1], Chenchen Chu[1], Xiaojun Qiao[2], †Junjie Ye[1]**

[1] Key Laboratory of Ministry of Education for Electronic Equipment Structure Design, Research Center for Applied Mechanics, Xidian University, Xi'an 710071, China

[2] Taiyuan Heavy Machinery Group Co., LTD Technology Center, Taiyuan 030024, China

*Present author: hcai_1@stu.xidian.edu.cn

†Corresponding author: ronkey6000@sina.com

ABSTRACT

The paper is establishing a model of wind blade by Reverse Engineering(RE), which can obtain the fitting surfaces of complex structure. In order to investigate the mechanical behaviors of blade with circular hole, the digital model is applied to the finite element calculation. Then the static loading tests are carried out to obtain strains at different positions of surfaces by Fiber Bragg Grating sensors. By comparing the finite element simulation value with the experimental value, it is verified that the finite element modeling based on Reverse Engineering is suitable.

Keywords: Reverse Engineering; wind blade; Finite Element Method

1.Instruction

With the development of the green energy, wind power generation is one of the most effective approach to get power[1].However, under the influence of external factors, the structural health of the wind blades will gradually degrade. In order to avoid the loss of economy and personnel, it is necessary to research the mechanical behaviors of wind blades.

Due to the complex characteristics of wind blade with hole, the free-form surfaces can not be rebuilt by Computer- Aided Design(CAD).Reverse engineering can be used to recover the actual geometric model of products with special structure. Therefore, the Reverse Engineering is more suitable for modeling in this case[2]. Besides, using the reverse modeling, the error of structural parameters is reduced.

In order to decrease the counting amount about the high order element[4], Mindlin proposed the shell theory which applies for laminate structure. In this paper, the finite element method is used to simulate the static loading test. For improving the accuracy of measurement, the Fiber Bragg Grating sensors are applied[3].It researches mechanical behaviors of composite laminate structure under the different loading through combining the measuring results with the simulation value. This study aims to research the influence of stress concentration, which provides the basis for the failure study of blade.

2.Reverse Engineering

Reverse Engineering uses three measuring instruments to acquire data from existing products and collect the data to reconstruct the model[2], the process of which can be divided into the following stages(Fig.1):

1.Data acquisition: the use of three-dimensional measuring instruments to measure the physical model of the model surface 3D data。

2.Data preprocessing: Pretreatment of measurement data, simplification, filtering, triangulation and so on;

3.Data segmentation: Because the measurement model is usually composed of a number of different geometric features of the surface, so it is necessary to block the measurement data;

4.Surface reconstruction: According to the geometric characteristics of the sub-surface, the surface is fitted, and the complete surface model is obtained;

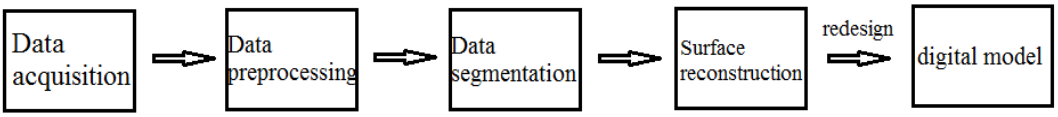


Fig.1.Reverse engineering process

If the existing complex model can be rebuilt, the requirement for the equipment precision is high. In this experiment, the scanning equipment MetraSCAN 3D is manufactured by Creaform with the seven laser cross-line, which adopts the portable design to get rid of the scanning method of the traditional mechanical arm. The highest scanning accuracy can reach 0.030mm, the scanning area of the measurement speed can reach 480000 times/sec.

3.Reserve Modeling

Because of Geomagic Studio containing the comprehensive functions, the processing applied the software to rebuild the wind blade. In the paper, the main challenge is the topology of the original surface is preserved, while sharp features and surface boundaries are hard to reproduce accurately in the reconstructed surface. Thus, the following ways are adopted to solve problems:(1)The key structure are scanned several times, and the points cloud data are spliced by the curvature variation of local features, which can reduce the influence of variable density, noise and outliners in one acquisition process.(2)If several triangular patches with sharp feature scan not match the actual model, the patches must be filled from the surface boundary that has been constructed, in order to conform to curvature variation of geometric boundary structure.

The pre-processing can get regular surface patches. Then the project finds smooth surface with repairing steps to remove the spindle, relax the boundary and simplify processing. Finally, the surface model can be used for finite element modeling. The software interface is shown in Fig.2.

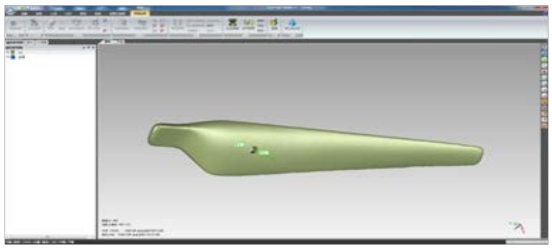


Fig.2.Geomagic Studio software interface

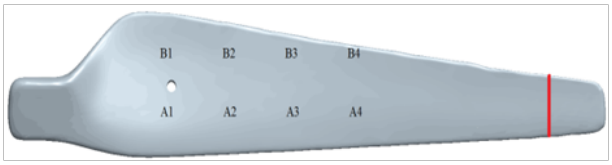


Fig.3.The position of the sensors and the load

4.Mechanicas Behaviors of The Wind Blade

Figure 3 is a three-dimensional digital model of the blade used in the experiment, the red is the loading position and the static loads are 1.8kg, 2.8kg, 3.8kg, 4.8kg, 5.1kg, 7.1kg, respectively. The clamping length of root is 50mm,the distance between the hole (D=13mm) and the blade root is 270mm, which is located in the center of the chord length. The direction of the A-group sensor and the B-group sensor is parallel to the blade orientation. A1 is located below the chord long line of the hole and B1 is located down it. The vertical distances from hole to A1 and B1are 24mm and 32mm respectively. The distances between A1, A2, A3, and A4 are 240mm, 125mm and 130mm.

The relationship between the wavelength variation $\Delta\lambda_B$ and the axial strain ε of the fiber grating can be expressed by the equivalent conversion coefficient p_e into the following formula:

$$\varepsilon = \frac{\Delta\lambda_B}{p_e} \times 1000 = \frac{\lambda_B - \lambda_0}{p_e} \times 1000 \tag{4-1}$$

The range of center wavelength of the Fiber Bragg Grating demodulator is 1510-1590nm, the strain of the small wind blade does not exceed 3000 micro strain. In order to ensure that the wavelength of each measuring point is not confused in the demodulation, the wavelength interval between different measuring points is set 6nm. The center wavelengths corresponding to the above-mentioned Fiber Bragg Grating (FBG)[4] sensors are shown in Table.1.

Table.1.The central wavelength of the measurement points in different groups

Sensor Number	A1	A2	A3	A4	B1	B2	B3	B4
Central wavelength/nm	1541	1547	1532	1538	1556	1535	1538	1552

The displacement vector of the shell element's deformation in Mindlin Theory is defined as[4]:

$$\begin{cases} u_x(x,y,z) = u + z\theta_y \\ u_y(x,y,z) = v + z\theta_x \\ u_z(x,y,z) = w \end{cases} \tag{4-2}$$

In this equations, $u = u(x,y)$ and $v = v(x,y)$ are the displacements in the x and y directions of the mid-plane; $\theta_x = \theta_x(x,y)$ and $\theta_y = \theta_y(x,y)$ are respectively x- axis negative direction of rotation angle andy- axis positive direction of rotation angle ; $w = w(x,y)$ is deflection variation.

Strain-displacement relation and mid-plane tensile strain are shown as:

$$\begin{Bmatrix} \varepsilon_{xx} \\ \varepsilon_{yy} \\ \gamma_{xy} \end{Bmatrix} = \begin{Bmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{Bmatrix} + z \begin{Bmatrix} \kappa_{x0} \\ \kappa_{y0} \\ \kappa_{xy0} \end{Bmatrix} = e(u) + zk(u) \tag{4-3}$$

$$e(u) \equiv \begin{Bmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{Bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & 0 & 0 \\ 0 & \frac{\partial}{\partial y} & 0 & 0 & 0 \\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 & 0 & 0 \end{bmatrix} \begin{Bmatrix} u \\ v \\ w \\ \theta_x \\ \theta_y \end{Bmatrix} \equiv L^e u \tag{4-4}$$

The fiber volume ratio of blade composites is 0.45, the material properties used in this study for the glass fiber include a Young's elastic modulus of 72GPa, a shear modulus of 40Gpa, a longitudinal Poisson's ratio of 0.2,and the epoxy matrix include a Young's elastic modulus of 3.45GPa, a shear modulus of 3.89Gpa, a Poisson's ratio of 0.35. By the finite element analysis, strains in loading direction and equivalent stresses are shown in Fig.4. The experimental data of the measured positions and the calculated strains in loading direction are provided in Fig.5.

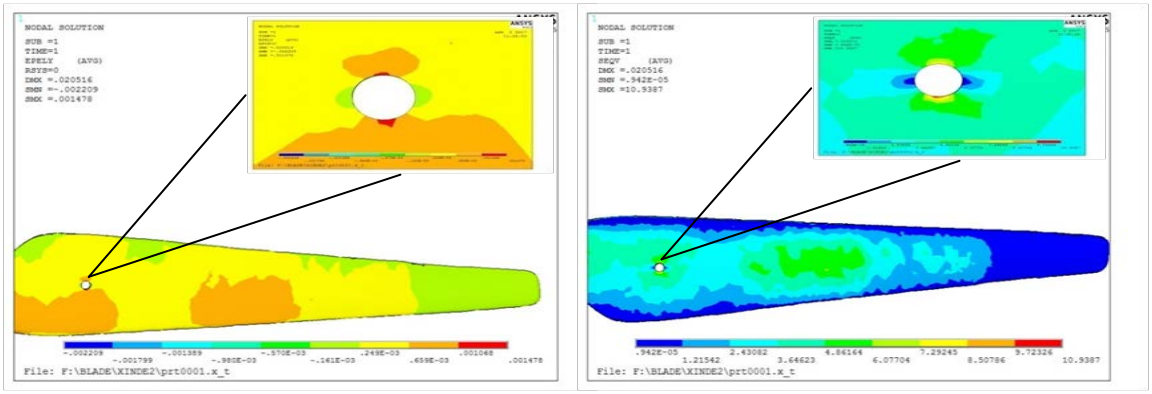


Fig.4. strain and stress contour plot

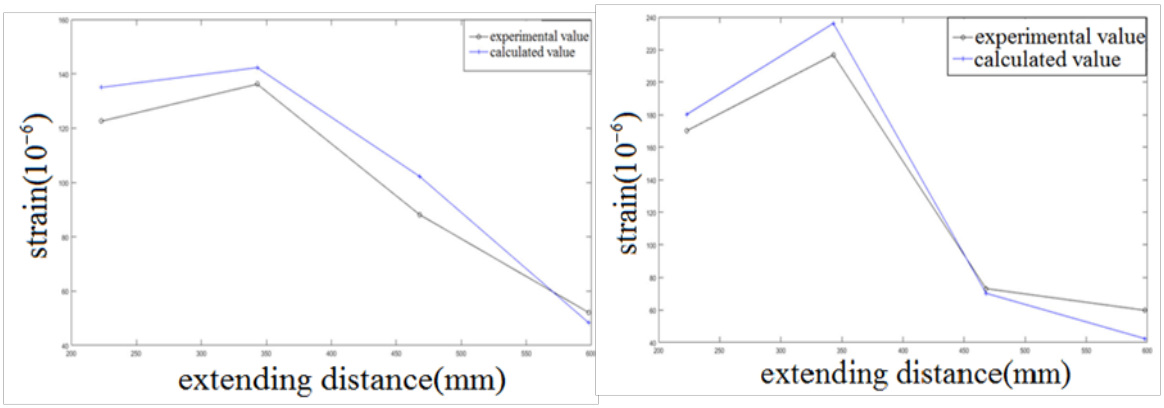


Fig.5.Comparing experimental strains with calculated strains in A group and B group under loading 5.1kg

5.Conclusion

In this paper, the Reverse Engineering is applied to study the mechanical properties of wind blades. The results of finite element analysis are compared with the experimental tests, which lays a good foundation for the study of complex structures. To improve calculation precision, the simulation should apply the higher-order elements with Finite Element Method. While it takes more time to get accurate results, choosing the proper method is also useful for mechanical behaviors of composite laminate.

Acknowledgements

This work was supported by the National Natural Science Foundation of China(Nos.51675397, 51305320)andNaturalScienceBasicResearchPlaninShaanxiProvinceofChina(No. 2015JQ5146), the 111 Project B14042.

Reference

[1] EWEA. World wind energy report 2009[R].Bonn: World Wind Energy Association,2010.

[2] Josef Hoschek and WErner. Reverse Engineering[J].Stuttgart:Teubner,1996

[3] Schoeder K, Ecke W, Apitz J, et al. A Fiber Bragg Grating sensor system monitors operational load in a wind turbine[J]

[4] Ziekiewicz OC, Taylor RL. The finite element method[M], 5th ed, volume 2: solid mechanics,2000

GPU parallel computation of topology optimization based on EFG method

***†S. G. Gong, Q. L. Liu, G. L. Xie, H. L.Lu, and J. P. Zhang**

School of Mechanical Engineering, Xiangtan University, Xiangtan 411105, China

*Presenting author: gongsg@xtu.edu.cn

†Corresponding author: gongsg@xtu.edu.cn

Abstract

Aiming at the problem of low efficiency and time-consuming in topology optimization based on element-free Galerkin (EFG) method, a new implementation of topology optimization based on EFG method and Graphic Processing Unit (GPU) parallel computing is presented by using node-by-node method and interacting nodes pairs to carry out the sensitivity analysis, and the corresponding computational formulas are derived. The GPU parallel execution model is designed to be used in the sensitivity analysis of objective function and implementing of optimization criterion (OC) method, and the flowchart is also given. The two examples of topology optimization are achieved, and the results obtained show that the proposed method is verified to be efficient and feasible. On the premise that the calculating accuracy is met, the 24 times speedup is obtained.

Keywords: Topology Optimization; GPU; Parallel Computing; EFG method; Sensitivity Analysis

1. Introduction

Structural topology optimization has played an important role in lightweight design of industrial product parts [Zargham et al. (2016); Chen et al. (2016)]. Its numerical methods are mainly based on element based method, such as the finite element. The numerical instabilities, such as checkerboards, mesh-dependencies and local minima, can occur in application of topology optimization because of using element based method [Sigmund O (1998)]. To get clear topological outline, the different topology optimization methods or filtering technologies have been put forward, such as level set method, evolutionary structural optimization method(ESO) , and so on [Zargham et al. (2016)].

The Element-Free Galerkin (EFG) method requires only nodal data, and the element connectivity between nodes is eliminated [Lu et al. (1994)]. As the EFG method has high rate of convergence, high computing accuracy and good computing stability, the scholars have tried to us it in topology optimization. Gong et al. [2012] has presented the topology optimization method based on EFG method by selecting the nodal density as the design variables. His research results show that the numerical instabilities in topology optimization have significantly improved. The same conclusion has also achieved by authors [Yang et al. (2016)]. Despite all this, the EFG method also has some shortcomings that are the EFG method has poor computational efficiency and time-consuming [Belytschko et al. (1996)]. For this reason, many scholars try to make use of parallel algorithm to improve computing efficiency of EFG method. Singh [2004] has researched parallel assembling of stiffness matrix and parallel solving of linear equations, and achieved the parallel computing of EFG method. Zeng et al. [2008] discussed the parallel computing the shape functions and their derivatives of Moving Least Squares method, and his research results show the EFG method

has large parallel potential and high parallel efficiency.

In particular, as the NVIDIA Company released Compute Unified Device Architecture (CUDA) and the Open Compute Language (OpenCL) in recent years, parallel acceleration of Graphic Processing Unit (GPU) has received considerable attention. Karatarakis et al. [2013] presented the GPU parallel algorithm to assemble stiffness matrix of EFG method. And the GPU parallel algorithm of solving linear equations by using conjugate gradient method has been presented by authors [Bolz et al. (2003)]. Gong et al. [2015] presented a GPU acceleration parallel algorithm of EFG method by using node pair-wise approach to assemble the stiffness matrix, the maximum speedup ration is up to 17 times. Stephan et al. [2011] investigates the GPU parallel computing of topology optimization with the solid isotropic material with penalization approach based on FEM, and they found that the computing efficiency of GPU is faster than a 48 core shared memory CPU system. In addition, Challis et al. [2014] has carried out the parallel computing of topology optimization with level set method based on FEM. They have also found that the GPU is utilized more effectively at the higher scale problem. To solve large FE model in topology optimization, Martínez-Frutos et al. [2016] has presented a multi-GPU system in terms of memory consumption and processing time. Cai et al. [2016] proposed the parallel computing method of Bi-directional Evolutionary Structural Optimization (BESO) based on MatLab and GPU. In a word, the GPU parallel computing has become a very popular accelerating computational method.

Hence, in this paper, our purpose is to present an entire topology optimization method based on the GPU parallel computing and EFG method. The parallel algorithm of sensitivity analysis in topology optimization and optimization criterion (OC) method will be explored in detail, and the flow chart of GPU parallel computing is given then. The two numerical examples are achieved based on the proposed method, and the influence of nodes to speedup will be discussed.

2. Topology optimization based on EFG method

In this work, we utilize the penalty function method to impose the essential boundary condition because the shape functions in the EFG method do not satisfy the Kronecker delta property. The minimum structural compliance c is chosen as the objection function, and the volume ratio f as the constraint, the relative density of nodes ρ_I as the design variables, the topology optimization model by using Solid Isotropic Material with penalization (SIMP) interpolation can is now given as follows.

$$\begin{cases} \text{find} & \rho(x) \\ \min & \mathrm{c}(\boldsymbol{\rho}) = \left(\boldsymbol{F} + \boldsymbol{F}^a\right)\boldsymbol{U} \\ \text{s.t} & V = fV_0 \\ & \left(\boldsymbol{K} + \boldsymbol{K}^a\right)\boldsymbol{U} = \boldsymbol{F} + \boldsymbol{F}^a \\ & 0 < \rho_{\min} \le \rho_I \le 1 \end{cases} \tag{1}$$

where $\boldsymbol{K}$ is the global stiffness matrix, $\boldsymbol{K}^{a}$ is the global penalty stiffness matrix, $\boldsymbol{F}$ is the global force vector. $\boldsymbol{F}^{a}$ denotes the global penalty force vector, and V represents the material volume of design domain, These parameters can be defined as follows, respectively.

$$\boldsymbol{K}=\int_{\Omega}\rho_{I}^{(p)}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{B}\mathrm{d}\Omega,\qquad \boldsymbol{K}^{\alpha}=\alpha\int_{\Gamma_{u}}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{\Phi}\cdot\mathrm{d}\Gamma,$$

$$\boldsymbol{F}=\int_{\Gamma_{t}}\boldsymbol{\Phi}^{\mathrm{T}}\bar{\boldsymbol{t}}\mathrm{d}\Gamma+\int_{\Omega}\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{b}\mathrm{d}\Omega,\qquad \boldsymbol{F}^{\alpha}=\alpha\int_{\Gamma_{u}}\boldsymbol{\Phi}^{\mathrm{T}}\bar{\boldsymbol{u}}\mathrm{d}\Gamma,\qquad V=\int_{\Omega}\rho\mathrm{d}\Omega$$

in which $\boldsymbol{B}$ is the geometric matrix, $\boldsymbol{D}$ is the elasticity matrix. α is the penalty factor, in general, $\alpha=\left(10^{3}\sim10^{7}\right)E$, and E is the elastic modulus. $\boldsymbol{\Phi}$ denotes the shape function; $\bar{\boldsymbol{t}}$ are the traction forces applied on the natural boundary, $\boldsymbol{b}$ is the body force vector, $\bar{\boldsymbol{u}}$ is the known displacements on the essential boundary, $\boldsymbol{U}$ is the approximation nodal displacement, and V_{0} is the initial volume of the design domain. In the meantime, in order to avoid the singularity in the optimal process, the lower limiting value of relative density $\rho_{\min}$ is set to 0.001. Meanwhile, the relative density of any node in the design domain is given as follow

$$\rho=\sum_{I=1}^{N_{e}}\phi_{I}\rho_{I}$$

On the other hand, we select the Optimization Criterion (OC) method [Sigmund (2001)] to update the design variables. The iteration scheme is given by

$$\rho_{I}^{(k+1)}=\begin{cases}\max\left(\rho_{\min},\rho_{I}^{(k)}-m\right) & \text{if } \max\left(\rho_{\min},\rho_{I}^{(k)}-m\right)\geq\left(\mathrm{B}_{I}^{(k)}\right)^{\eta}\rho_{I}^{(k)}\\ \left(\mathrm{B}_{I}^{(k)}\right)^{\eta}\rho_{I}^{(k)} & \text{if } \max\left(\rho_{\min},\rho_{I}^{(k)}-m\right)<\left(\mathrm{B}_{I}^{(k)}\right)^{\eta}\rho_{I}^{(k)}<\min\left(1,\rho_{I}^{(k)}+m\right)\\ \min\left(1,\rho_{I}^{(k)}+m\right) & \text{if } \left(\mathrm{B}_{I}^{(\mathrm{k})}\right)^{\eta}\rho_{I}^{(\mathrm{k})}\geq\min\left(1,\rho_{I}^{(\mathrm{k})}+m\right)\end{cases}\qquad(2)$$

where superscript k is the number of iteration. In order to guarantee the stable of iteration, the damping coefficient η and a positive move-limit m are introduced, respectively, and B_{I} can be found from the optimality condition as

$$B_{I}=(-\frac{\partial c}{\partial\rho_{I}})\Big/(\lambda\frac{\partial V}{\partial\rho_{I}})$$

where λ is a Lagrange multiplier that can be obtained by a bi-section algorithm.

The sensitivity analysis of the objective function and volume can be expressed as follows, respectively.

$$\frac{\partial V}{\partial \rho_I} = \int_\Omega \phi_I \mathrm{d}\Omega \tag{3}$$

$$\frac{\partial c}{\partial \rho_I} = -\boldsymbol{U}^{\mathrm{T}} \frac{\partial \boldsymbol{K}}{\partial \rho_I} \boldsymbol{U} \tag{4}$$

where $\frac{\partial \boldsymbol{K}}{\partial \rho_I}$ is given by

$$\frac{\partial \boldsymbol{K}}{\partial \rho_I} = \int_\Omega p\rho^{p-1} \phi_I \boldsymbol{B}^{\mathrm{T}} \boldsymbol{D}\boldsymbol{B} \mathrm{d}\Omega \tag{5}$$

where superscript p is used as a penalty factor enforcing a 0/1 distribution as intermediate values are greatly reduced.

More details on topology optimization based on EFG method in general can reference to the literature [Gong et al. (2012); Gong et al. (2009)].

3. Implementation of GPU parallel algorithm

3.1 Date storage and access

NVIDIA's GPU employing the isolated storage system of multi-level memory is an independent computing system, but there exist giant difference in feature of different memory, and their location, access permission and life cycle are different. So the rational allocation and use of GPU memory in the CUDA architecture is important for improving the performance of GPU program, especially topology optimization based on EFG method requires processing huge amount of data, and using frequencies and ways between different data are also different. Aiming at these problems, we take into account following ways to carry out optimization of data storage and access.

1) Subject to the limit of transmission efficiency of the PCI-E port, the data transmission bandwidth between CPU and GPU is much less than that of video memory. In order to avoid time-consuming in the data transmission, according to the needs of the program design, the prepared basic data will be transferred once from CPU to GPU instead of being transferred repeatedly, and only the residual error will be returned to CPU for looping control in the each iteration process. Finally, the optimization results are also transmitted to the host memory after the optimization.

2) The memory capacity of stiffness matrix in the EFG method is much bigger than the transitional FEM. Meanwhile, the data including the model data, updated data in the iterative

can be only stored in the GPU global memory. But the global memory has high access delay, and it is easy to become the access bottleneck. In order to achieve the highest access rate, the address of the first element in each row of data array is aligned by programming. This approach will be beneficial to satisfy the combined access requirements and ensure the maximum effective bandwidth.

3) Each adding a processor in the GPU has a cache to speed up reading operation from constant memory, and compared with the global memory, the constant memory has smaller delay and faster speed for access operation. Hence, for the sake of effectively reducing the access to global memory and improving process performance, the constants accessed frequently, for example, penalty factor, numerical damping coefficient, material properties, etc, will be stored in the constant memory.

4) To take full advantage of the register in the GPU chip, which it has the characteristic of high bandwidth, low delay and small capacity, the number of intermediate variables should be minimized. If the intermediate variables are overmuch, they should be stored in shared memory instead of register, and the registers should be allocated rationally for each thread according to the difference task.

3.2 Parallel algorithm of sensitivity analysis

In the conventional computing of topology optimization, the sensitivity analysis of objective function can be achieved by looping for integral point. That is to say, first we calculate the sensitivity component of all nodes in the influent domain of integral point, and then accumulate them by node number to get the sensitivity array. Due to the node may also appear in the influent fields of multiple Gauss integral points simultaneously. It means that the address of each node in the sensitivity array will be accessed repeatedly, as shown in Fig. 1. This way is not suitable for parallel computing because it is easy to lead to the conflicts of data storage. Therefore, we present a parallel algorithm to carry out the sensitivity analysis by using the interacting nodes pairs that is two nodes in which influence domain between them has overlap region [Gong et al. (2015)], as shown in Fig. 2. This parallel scheme adopts node-by-node method to circularly develop the sensitivity computing of objective function, and the address of node in sensitivity array will be accessed only once. This approach can eliminate the conflicts of data storage and be suitable for parallel computing.

Moreover, there exist a lot of the intermediate variables in the traditional calculation. These variables will appear in an array form to participate in operation. It will seriously degrade the global performance of GPU programs because these variables can only store in global memory. Aiming at this problem, by combining the parallel algorithm as shown in Fig. 2, we propose a following processing method for the sensitivity computing of objection function.

For two-dimension problem, the global stiffness matrix $\boldsymbol{K}$ is a $n \times n$ matrix, and each element $\boldsymbol{K}_{LM}$ in matrix is a 2×2 sub-blocks. Therefore, $\dfrac{\partial \boldsymbol{K}}{\partial \rho_I}$ is also comprised by a series of

sub-blocks $\frac{\partial \boldsymbol{K}_{LM}}{\partial \rho_I}$. According to Eq. (5), $\frac{\partial \boldsymbol{K}_{LM}}{\partial \rho_I}$ can be expressed as

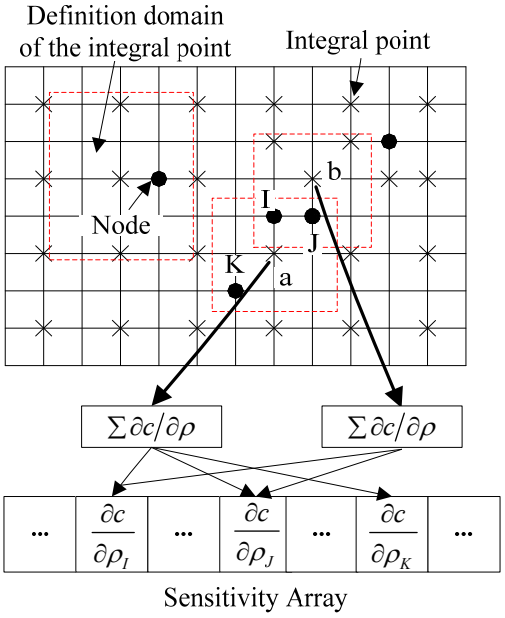


Fig. 1 Assembly the sensitivity matrix by looping for integral point

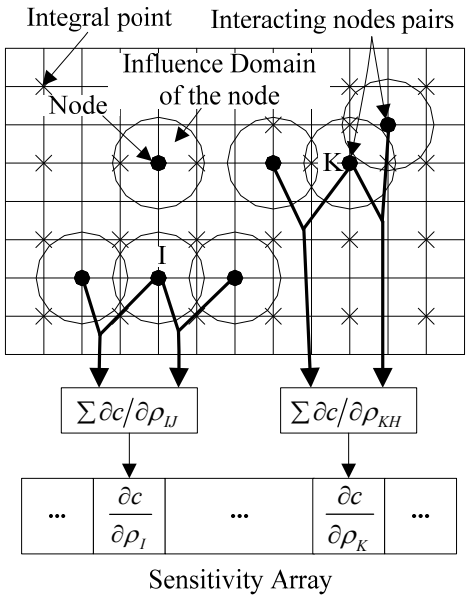


Fig. 2 Assembly the sensitivity matrix by using node-by-node method

$$\frac{\partial \boldsymbol{K}_{LM}}{\partial \rho_I} = \int_\Omega p\rho^{p-1}\phi_I \boldsymbol{B}_L^{\mathrm{T}} \boldsymbol{D}\boldsymbol{B}_M \mathrm{d}\Omega \tag{6}$$

In the EFG method, the global stiffness matrix is also banded sparse matrix, and it means that the value of the most of 2×2 blocks in $\frac{\partial \boldsymbol{K}}{\partial \rho_I}$ is zero. Only the block element value corresponding to the interacting nodes pairs of node I is not zero. So we can obtain the following equation according to Eq. (4).

$$\frac{\partial c}{\partial \rho_I} = -\sum_{i=1}^{NP_I}\left\{\left(\boldsymbol{u}_L^{\mathrm{T}} \frac{\partial \boldsymbol{K}_{LM}}{\partial \rho_I}\boldsymbol{u}_M\right)_i\right\} \tag{7}$$

where $\boldsymbol{u}_L = \left[u_{Lx}, u_{Ly}\right]^{\mathrm{T}}$ is the displacement vector of node L, and NP_I is the number of interacting nodes pairs of node I.

Substituting Eq. (6) into the Eq. (7), we have

$$\frac{\partial c}{\partial \rho_I} = -\sum_{i=1}^{NP_I}\left\{\left(\boldsymbol{u}_L^{\mathrm{T}}\left(\int_\Omega p\rho^{p-1}\phi_I \boldsymbol{B}_L^{\mathrm{T}} \boldsymbol{D}\boldsymbol{B}_M \mathrm{d}\Omega\right)\boldsymbol{u}_M\right)_i\right\} \tag{8}$$

The Eq. (8) can also be written as

$$\frac{\partial c}{\partial \rho_I}=\sum_{i=1}^{NP_I}\left\{\left(\begin{array}{l}\phi_I p\rho_I^{(p-1)}\dfrac{E}{1-\nu^2}HJ(u_{Lx}u_{Mx}\phi_{L,x}\phi_{M,x}+u_{Lx}u_{Mx}\phi_{L,y}\phi_{M,y}\dfrac{1-\nu}{2}+\\ \nu u_{Ly}u_{Mx}\phi_{L,y}\phi_{M,x}+u_{Ly}u_{Mx}\phi_{L,x}\phi_{M,y}\dfrac{1-\nu}{2}+\nu u_{Lx}u_{My}\phi_{L,x}\phi_{M,y}+\\ u_{Lx}u_{My}\phi_{L,y}\phi_{M,x}\dfrac{1-\nu}{2}+u_{Ly}u_{My}\phi_{L,y}\phi_{M,y}+u_{Ly}u_{My}\phi_{L,x}\phi_{M,x}\dfrac{1-\nu}{2})\end{array}\right)_i\right\} \tag{9}$$

where E is Young's modulus and ν is Poisson's ratio; H , J are the weight coefficient and the Jacobi coefficient of integral point, respectively. $\phi_{L,x}$, $\phi_{L,y}$ denote the derivative for shape function ϕ_L with respect to x and y, respectively.

Although this approach may increase the complexity of programming, it can ensure that most of the intermediate variables could be stored in the registers of GPU. And it could be beneficial to improve the performance of program.

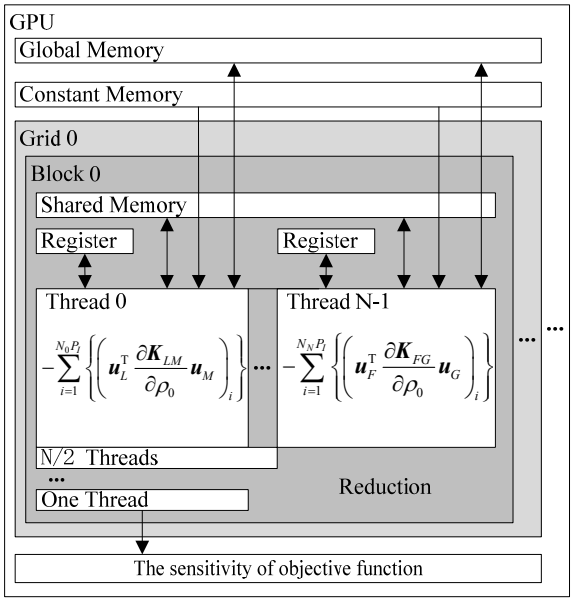


Fig.3 Parallel execution model based on CUDA architecture for the sensitivity of objective function

Synthesizing the above ideas, we present the parallel execution model based on CUDA architecture to achieve parallel computing for the sensitivity analysis of objective function, as shown in Fig. 3. That is, each thread reads data from the global memory and constant memory, and the registers and shared memory is used to store the intermediate variables. And then computing the sensitivity value of objective function of each node by using parallel Reduction Algorithm [Gong et al. (2015)], and its final results will store in global memory. Meanwhile, there are two parallel hierarchies in this algorithm. One or outer loop layer is the computing of nodes, that is, each thread blocks will calculate the objective function sensitivity of one node. The other or inner loop layer is the integral points in the influence domain of node, and each thread in the thread blocks will deal with the computing of an integral point. Finally, we

can obtain the sensitivity component of this node objective function.

3.3 Parallel algorithm of OC method

It shares less time in the whole topology optimization that the OC method is used to update the design variables, but to obtain the final optimal result need repeatedly iterate in the structural topology optimization. Only if the OC method is implemented in the CPU serial computing，a large amounts of data need to be transferred repeatedly between CPU and GPU, and it will seriously reduce the effectiveness of other parallel algorithm. On the other hand, the computing between each step in the iteration process is a serial operation, but some operation including the replacement of design variable, the calculation of the relative density, etc, may be a parallel operation in data level. So these operations can also make use of GPU to carry out parallel computation. Therefore, according to heterogeneous of CPU and GPU as well as the CUBLAS library, we present a parallel algorithm of implementing OC method on GPU, and it is controlled by using a sub-function in this paper. Detailed algorithm flow is expressed as follows.

1) Initialize the CUBLAS library.
2) Define the upper limit *l*1 and lower limit *l*2 of finite interval in host computer, which this interval is used to calculate the Lagrange multiplier by using bisection algorithm, and assigning their initial value at the same time.
3) Copy $\rho^{(k+1)}$ to $\rho^{(k)}$ by calling library function *cublasDcopy*() in CUBLAS library.
4) while(*l*2-*l*1>1.0e-4)
   {
   4.1) Calculate the Lagrange multiplier $\lambda_{(i+1)}$ by using bisection algorithm.
   4.2) After updating the Lagrange multiplier, Calculate the design variables $\rho_{(i+1)}^{(k+1)}$ by calling sub-function: *rou_Kernel*<<<*noumnod*, 1>>>() where *numnod* is the total number of nodes in the design domain. This sub-function is written according to Eq.(2), and there is a one-to-one mapping between thread block and node, that is, each thread block will calculate the density $\left(\rho_{(i+1)}^{(k+1)}\right)_I$ of one node.
   4.3) While Lagrange multiplier is $\lambda_{(i+1)}$, calculate the relative density of integral point, that is, $\rho_{in_{(i+1)}} = \sum_{I=1}^{N_c}\left[\phi_I\left(\rho_{(i+1)}^{(k+1)}\right)_I\right]$, by calling sub-function: *rou_integral_Kernel*<<<*numq*2, 32>>>() that is written according to calculation formula of $\rho_{in_{(i+1)}}$, where *numq*2 is the total number of integral points in design domain, and N_c is the number of nodes in influence field of integral point. There

are two parallel hierarchies in this sub-function. One is integral point layer, that is, each thread block computes the relative density of an integral point. The other is node in influence field of this integral point, that is, there is one-to-one mapping between each thread in thread block and node in influence field of this integral point. When the relative density component of each node is gained by using calculation formula $\left[\phi_I\left(\rho_{(i+1)}^{(k+1)}\right)_I\right]$, the relative density of integral point can be obtained by utilizing Reduction Algorithm to accumulate each component.

4.4) While Lagrange multiplier is $\lambda_{(i+1)}$, calculate the total volume V by calling sub-function *Vtol_Kernel*<<<1, 1024>>>() that is written by using expression $V_{(i+1)}=\int_\Omega \rho_{in}\mathrm{d}\Omega$. In this sub- function, each thread in the thread blocks only computes the volume component of a integral point, and the total volume $V_{(i+1)}$ of design domain can be obtained by accumulating volume component.

4.5) Judge the size of design domain volume $V_{(i+1)}$. if less than, then $l2=\lambda_{(i+1)}$, else $l1=\lambda_{(i+1)}$.

}

5) Clear the environment of CUBLAS library

6) Return the design variable value of node $\rho_{(i+1)}^{(k+1)}$.

3.4 Flowchart of parallel algorithm based on GPU

Combining the above ideas with the traditional topology optimization algorithm based on EFG method, in this paper, we propose the heterogeneous parallel program structure of CPU and GPU on CUDA framework, and the flowchart of parallel algorithm is shown in Fig.4. The CPU is responsible for the main control operation, calculation with shorter time-consuming or inconvenient parallel operation, and the final result output, etc. The GPU is responsible for iterative calculation, calculation of the shape function and its derivatives, etc, which they are the most time-consuming in structural topology optimization.

Meanwhile, the pretreatment in Fig.4 mainly include the following contents and procedure.

1) Set up integral points.

2) Calculate influence field radius of node and definition field radius of integral point.

3) Establish the list of integral points in influence field of nodes, and nodes in definition field of integral points.

4) Determine interacting nodes pairs.

5) Calculate global force vector and global penalty force vector.

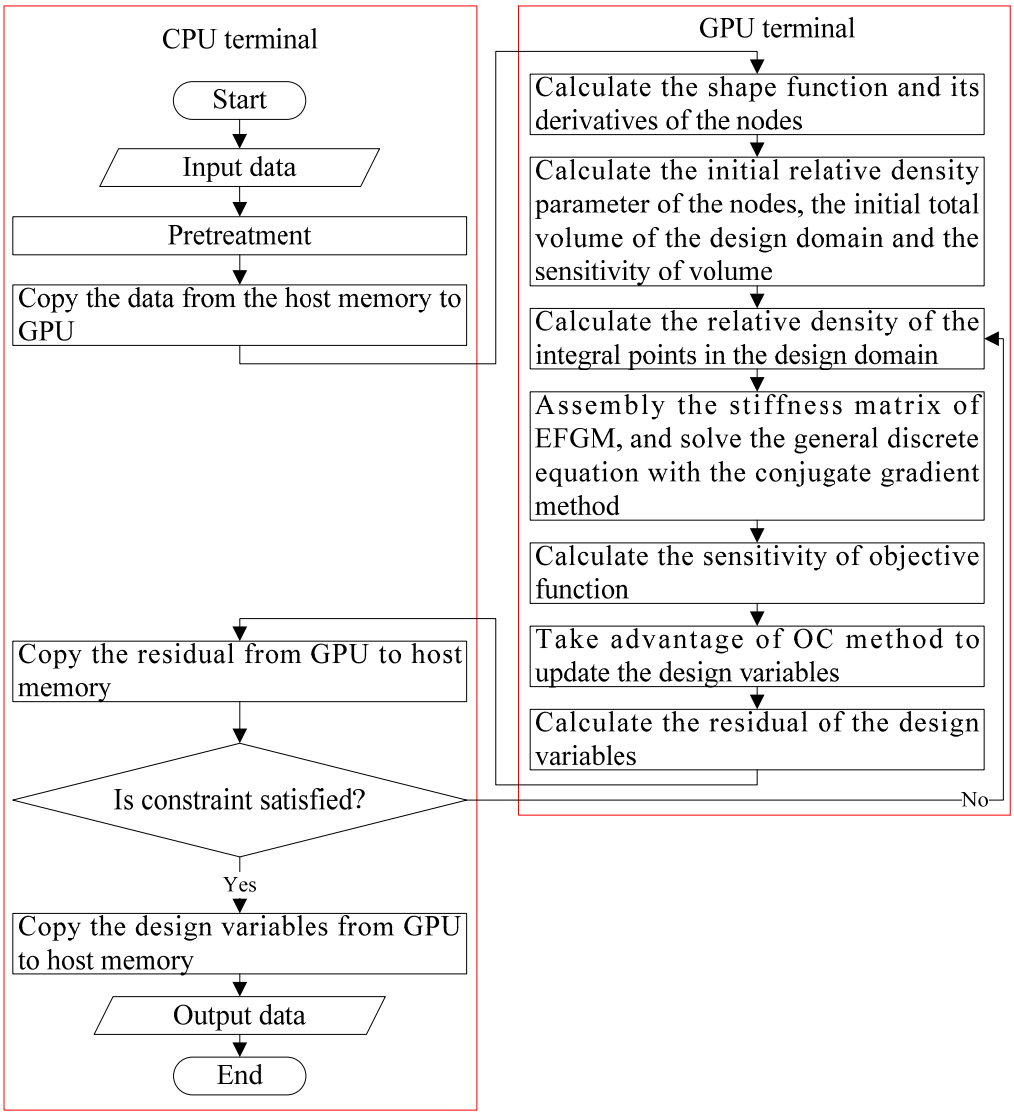


Fig. 4 Flowchart of GPU parallel algorithm for structural topology optimization

4 Numerical examples

The numerical examples are used to verify the feasibility of presented method in this paper, and are run on the following hardware and software, which CPU is Intel Core i5-3330 that has four physical cores at 3.0GHz, and GPU is the GeForce GTX 660 with 2GB GDDR5 memory, the version of CUDA is 3.0, and the operating system is 64 bit WIN 7 system. On the other hand, the same parameters are used in the following examples, that is, Penalty factor is $p = 3.0$, volume constraint is 30%, numerical damping coefficient is $\eta = 0.5$ and move-limit is $m = 0.2$. The termination condition of optimal iterative is $\left|\rho_I^{(k+1)} - \rho_I^{(k)}\right| \le 0.1$. The influence domain radius of node $\boldsymbol{x}_I$ is 2.5 times the size of the minimum distance between node $\boldsymbol{x}_I$ and other nodes, and the 2×2 Gauss points is assigned in each integration cell.

4.1 Example I- cantilever beam

The cantilever beam is a classical topology optimization problem, and its model is shown in Fig. 5(a). The cantilever assumed to be in a state of plane stress has characteristic height:

$H = 20\text{m}$ and width: $L = 40\text{m}$ and is considered to be of unit depth. It is fixed the left side and loaded with a point force $t = 1\text{N}$ at the lower of the right side. The design domain is discretized by 17×32 nodes uniformly distributed, as shown in Fig. 5(b), and is assigned by 496 integration cells. The elastic property of material is: Young's modulus $E = 1\text{Pa}$, Poisson' ratio $\nu = 0.3$.

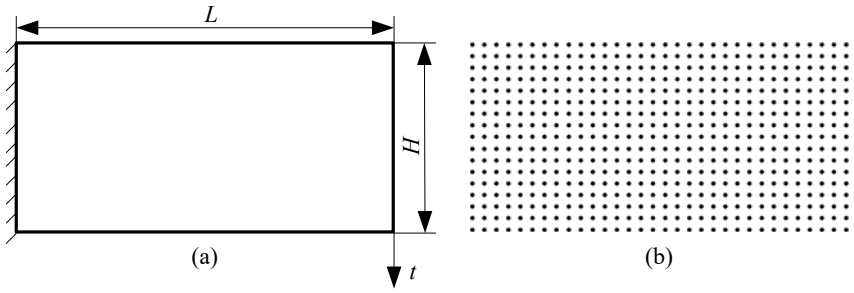


Fig. 5 Cantilever beam (a) geometry model, (b) discrete nodes

After 22 times of iteration, the program can reach the convergence precision. The optimization results obtained by using the GPU parallel algorithm and traditional algorithm are shown in Fig. 6, respectively. The figure 6 shows that both of results are good agreement. It means that the presented GPU parallel algorithm is feasibility and effective.

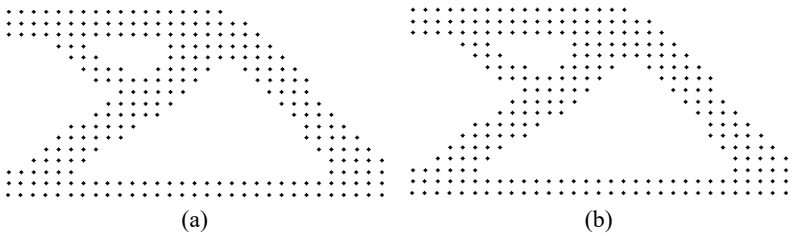


Fig. 6 Optimization results (a) traditional algorithm, (b) GPU parallel algorithm

When other parameters remain unchanged, the optimal results under different number of nodes can also be obtained by using GPU parallel algorithm, as shown in Fig.7.

In this paper, the speedup is defined as following

$$s_p = \frac{t_{\text{CPU}}}{t_{\text{GPU}}} \tag{10}$$

where t_{CPU}, t_{GPU} denotes for the run time of the traditional computation and parallel computation, respectively. The relation between speedup and nodes is shown in Fig.8.

In the case of different nodal spacing, the different topology results can be found in Fig.7, but their topological layout is similar. Meanwhile, it also implies that the GPU parallel algorithm presented is feasible to find the tiny structure in topological optimization.

The figure 8 shows that the speedup will increase with increasing of the number of nodes, and the maximum speedup is reached about 24 times. It is because GPU has a powerful floating-point computing power and is very suitable for computing of large scale problems. Meanwhile, it also means that the GPU parallel algorithm presented in this paper can effectively improve the computing efficiency of topology optimization.

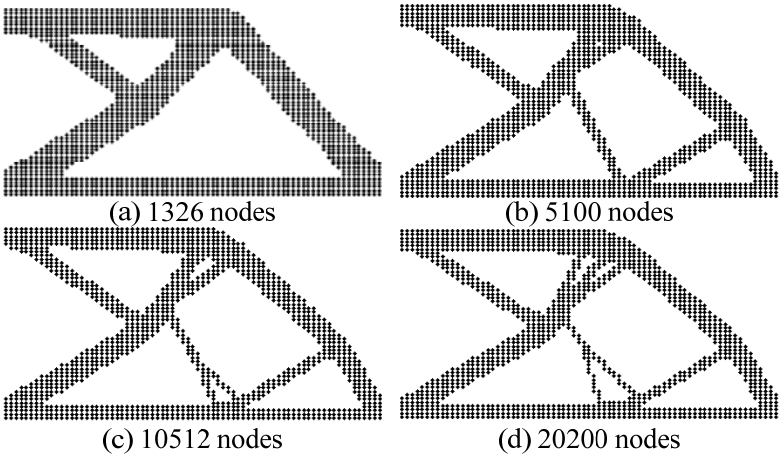


Fig. 7 Optimization results of cantilever beam under different number of nodes

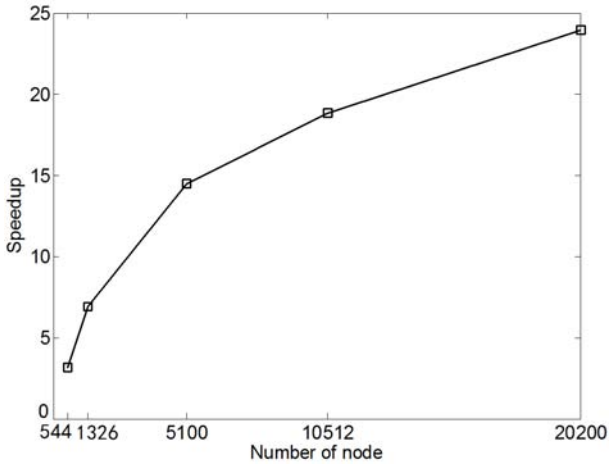


Fig.8 Relation between speedup and node

4.2 Example II-Deep beam with opening hole

A deep beam model with opening hole is showed in Fig. 9(a). It is subjected to a point load $P = 3000\text{kN}$, and Young's modulus is $E = 20820$ MPa; Poisson's ratio is $\nu = 0.15$. On the other hand, the initial width of this concrete beam is $b = 400\text{mm}$. The design domain is discretized by 545 nodes as shown in Fig. 9(b).

After iterating 20 times, the program reach the convergence precision. The final optimization results are shown in Fig. 10. The relation between speedup and number of nodes can be also obtained, as shown in Fig.11, and Fig.12 shows the optimal results under different number of nodes.

Figure 10 shows that the result obtained by using GPU parallel algorithm is good agreement with one gained by traditional algorithm. When the number of nodes is different, the basic topology layout is similar, but tiny structure in topological results can be found in Fig.12. And with the increasing of number of nodes, the speedup will also increase, as shown in Fig.11. The maximum speedup will reach 22.3 times.

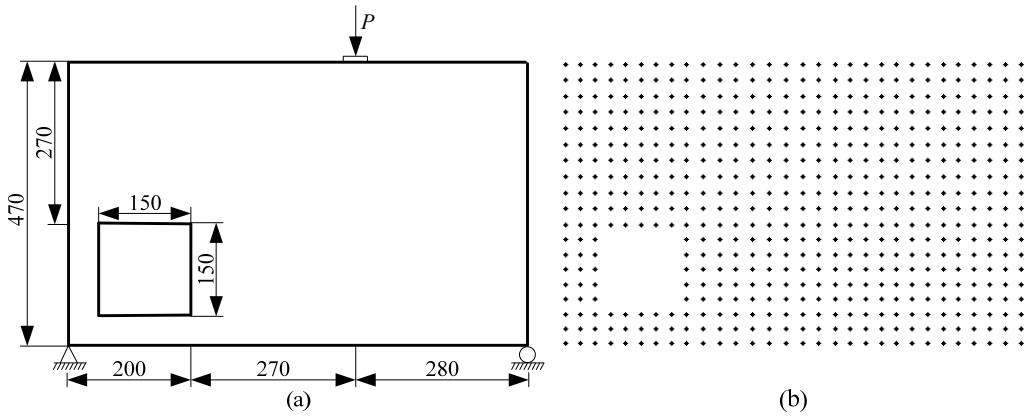


Fig. 9 Deep beam with opening hole (a) geometry model, (b) distribution of nodes

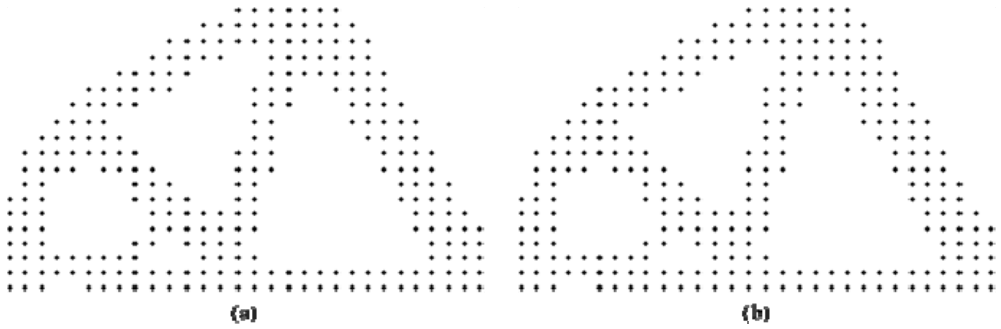


Fig. 10 Optimization result (a) traditional algorithm, (b) GPU parallel algorithm

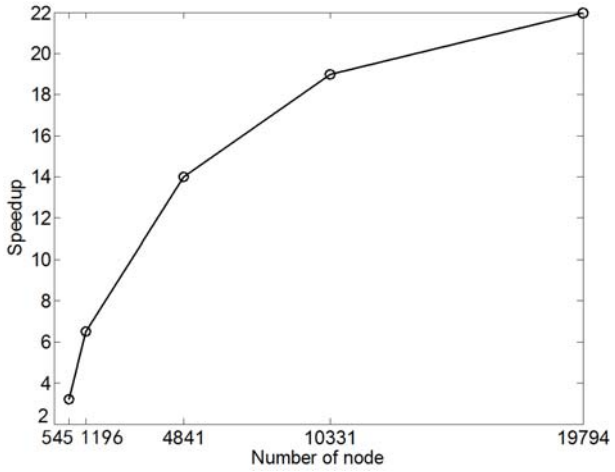


Fig.11 Speedup for beep beam model

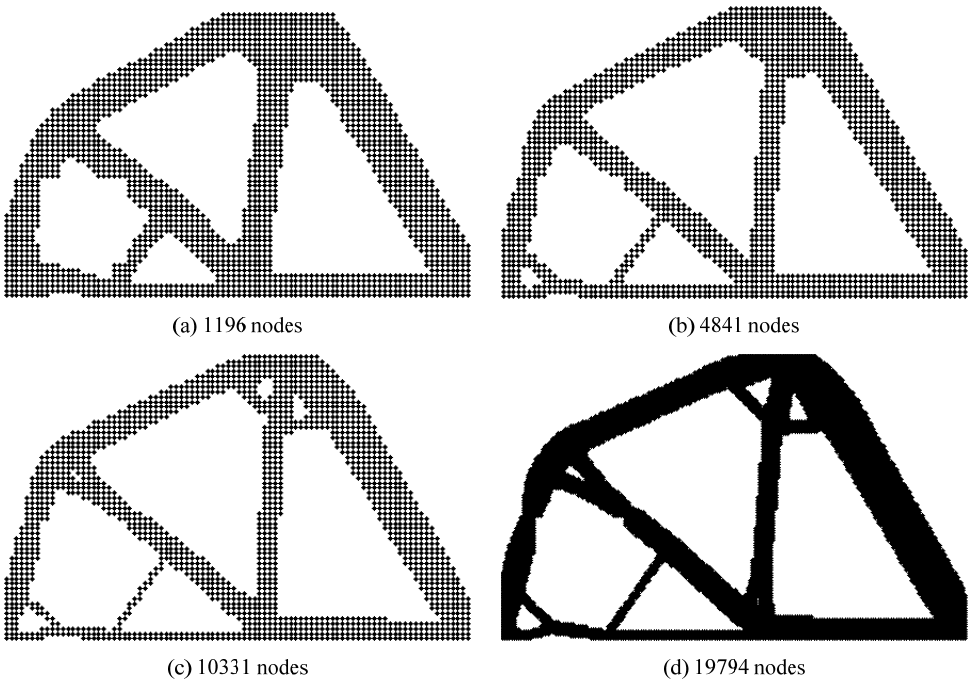


Fig. 12 Optimization results of beep beam under different number of nodes

Comparing Fig.12 and Fig.8, we can find that both of them are basically the same, and the size of speedup is related to the number of node.

5 Conclusions

In this paper, we present a new implementation of topology optimization based on GPU parallel acceleration by using the EFG method to carry out the numerical analysis. Its purpose is to reduce the computational cost of topology optimization. The parallel algorithm of the sensitivity analysis and OC method is discussed by using node-by-node method and interacting nodes pairs, and the corresponding flowchart and thread handling is given, respectively. Meanwhile, to improve the whole performance of GPU program and reduce the number of intermediate variables, the computational equation of sensitivity analysis for objection function is derived, so that the intermediate variables can be stored in the register of GPU. The presented method is verified by two numerical examples, and topological layouts obtained by using GPU parallel acceleration algorithm are good agreement with that using traditional algorithm. By studying on the speedup, it can be seen that the speedup will increase with the increase of number of node, and the maximum speedup may reach about 24 times in our given example. It means that the presented method in this paper is feasible and valid, and it can observably save the computation time of topology optimization.

Acknowledgment

This research is funded by National Natural Science Foundation of China [51375417, 51405415, 51475403]. The financial support to the first author is gratefully acknowledged.

References

[1] Zargham, S., Ward, T. A., Ramli, R. and Badruddin, I. A. (2016) Topology optimization: a review for structural designs under vibration problems, *Structural & Multidisciplinary Optimization*, **53**(6), 1157-1177.

[2] Chen, X. (2016) Topology optimization of microfluidics - A review, *Microchemical Journal*, **127**, 52-61.

[3] Sigmund, O. (1998) Numerical instabilities in topology optimization A survey on procedures dealing with checkboards, mesh-dependencies and local minima, *Structural Optimization*, **16**, 68-75.

[4] Lu, Y. Y., Belytschko, T. and Gu, L. (1994) A new implementation of the element free Galerkin method, *Computer Methods in Applied Mechanics & Engineering*, **113**(3-4), 397-414.

[5] Gong S. G., Chen M., Zhang J. P. and He R. (2012) Study on modal topology optimization method of continuum structure based on EFG method, *International Journal of Computational Methods*, **9**(1):1240005(12pags).

[6] Yang, X., Zheng, J. and Long, S. (2016) Topology optimization of continuum structures with displacement constraints based on meshless method, *International Journal of Mechanics & Materials in Design*, 1-10.

[7] Belytschko, T., Krongauz, Y., Organ, D., Fleming M. and Krysl P. (1996) Meshless methods: an overview and recent developments, *Computer methods in applied mechanics and engineering*, **139**(1), 3-47.

[8] Singh, I. V. (2004) Parallel implementation of the EFG method for heat transfer and fluid flow problems, *Computational Mechanics*, **34**(6), 453-463.

[9] Zeng, Y. S., Zeng, Q. H. and Lu, D. T. (2008) Parallel computing of element-free Galerkin method for elasto-dynamics, *Chinese Journal of Computational Mechanic*s, **25**(3), 385-391.

[10] Karatarakis, A., Metsis, P. and Papadrakakis, M. (2013) GPU-Acceleration of stiffness matrix calculation and efficient initialization of EFG meshless methods, *Computer Methods in Applied Mechanics and Engineering*, **258**, 63-80.

[11] Bolz, J., Farmer, I., Grinspun, E. and Der P. (2003) Sparse matrix solvers on the GPU: conjugate gradients and multigrid. *ACM Transactions on Graphics*, **22**(3), 917-924.

[12] Gong, S. G., Liu, Q. L., Lu, H. S., Zhou Z. Y. and Zhang J. (2015) Parallel computing and application of Element-free Galerkin method for GPU acceleration, *Chinese Journal of Computational mechanics*, **32**(6), 745-751.

[13] Stephan, S. and Volker, S. (2011) A 2589 line topology optimization code written for the graphics card, *Computing and Visualization in Science*, **14**, 249-256.

[14] Challis, V. J., Roberts, A. P. and Grotowski, J. F. (2014) High resolution topology optimization using graphics processing units(GPUs), *Structural and Multidiscipline Optimization*, **49**(2), 315-325.

[15] Martínez-Frutos, J. and Herrero-Pérez, D. (2016) Large-scale robust topology optimization using multi-GPU systems, *Computer Methods in Applied Mechanics & Engineering*, **311**, 393-414.

[16] CAI, Y. and Li S. (2016) Graphics processor unit parallel computing in Matlab and its application in topology optimization, *Journal of Computer Applications*, **36**(3), 628-632.

[17] Sigmund, O. (2001) A 99 line topology optimization code written in Matlab, *Structural and Multidiscipline Optimization*, **21**(2), 120-127.

[18] Gong, S. G., Liu, X., Xie, G. L. and Zhang J. P. (2009) Topology optimization under multi-load cases based on element-free Galerkin method, *Journal of Mechanical Engineering*, **45**(12), 137-142.

Design of pedestrian friendly vehicle frontal protection system using computer modelling and simulation

J.Q. Chen*[1], G. Haidar[2], S. Emilie[1], Y.T. Gu[1]

[1] Chemistry, Physics and Mechanical Engineering School, Queensland University of Technology, 2 George St, Brisbane, Queensland, 4000, Australia
[2] Product Development, Conversion & Accessories, Toyota Motor Corporation
*Corresponding Author: email: jiaqi.chen@student.qut.edu.au

Abstract

Globally, Vehicle Frontal Protection Systems (VFPS) need to satisfy pedestrian safety crash test requirements. In Australia, the existing designs of VFPS do not consider pedestrian safety due to an absence of pertinent regulatory requirements. This paper develops a design and validation framework for a new pedestrian-friendly concept of VFPS, which is in demand to meet global lower legform pedestrian safety requirements. The design and crash test simulations are carried out in Finite Element Analysis (FEA) program LS-DYNA. A physical crash test under the Euro NCAP Transport Research Laboratory (TRL) lower legform test condition is conducted to confirm the new VFPS design. It has been found the computer simulation results based on the new VFPS design agree with the results obtained by the experiment, satisfying Euro NCAP performance requirements. A Flexible Pedestrian Legform Impactor (Flex-PLI) model is further evaluated on the validated design, which also produces satisfactory predictions for future pedestrian safety testings.

Keyword: Finite Element Analysis, Vehicle Frontal Protection System, Pedestrian safety, Pedestrian protection, Lower legform, TRL legform, Flex-PLI legform

Introduction and background

Vehicle Frontal Protection Systems (VFPS), also known as bull bars and nudge bars in Australia, are the frontal devices fitted to the vehicle for vehicle and occupant protection in the event of a frontal collision, such as a kangaroo strike. While prevalent in the market, mainstream metal VFPSs have been questioned for their perceived aggressiveness towards pedestrians during collisions [1]. Government attempted to propose pertinent regulations to address pedestrian friendly design in VFPS [2][3], however the implementation was not successful because of the backlash from the manufacturers [4][5]. Therefore, the pedestrian safety issue of VFPSs remains unresolved to future society.

This paper is an industry response towards this future trend by researching and developing a new generation of VFPS product. The author commenced the preliminary design work utilising numerical modelling. Finite Element Analysis (FEA), for its characteristic of time-efficiency and cost-saving, was chosen to be the ideal tool to evaluate the new design. In literatures, vehicle pedestrian safety design has been studied using FEA in the past decades, but there is limited FEA R&D work undertaken on VFPS products after Europe and United Nations introduced regulations to mandate a high level of pedestrian friendliness, which caused the European VFPS market to plummet and never recover. Among these works, Ptak et.al investigated several geometry parameters of VFPS on their influences on the TRL

legform injury using LS-DYNA [7][8][10][15]. Brooks undertook his research about the material and structural requirements for a VFPS to meet 2005/55/EC criteria [9]**.** Pohlak [6] conducted the parameter study on VFPS bracket designs to investigate the impacts on pedestrian safety performance. There are separate design variable studies in these papers however neither new structure was implemented for VFPS nor real-world product developments were researched or published.

This paper employs the advanced capability of FEA in reconstructing explicit crash simulation using LS-DYNA, abiding by the latest protocol of Euro NCAP using lower legform practice to evaluate the design. As the current version of TRL legform will be replaced by Flex-PLI in Australia post 2018, TRL is used as a validation tool together with designed experiment tests in the lab to confirm the modelling, and generation 2.0 Flex-PLI legform performance is subsequently predicted on the validated model. The research demonstrates the validity of the FEA modelling in helping engineers design the product and the excellent performance of the VFPS design herein included.

Method and tools

In engineering design, computer simulations have been widely adopted across all Original Equipment Manufacturers (OEMs). In pedestrian-vehicle crash simulation where subsystems impactor or pedestrian dummy biomedical injury performance is assessed, FEA has the advantage of accurately predicting the local deformations which can accurately reflect the injury patterns of pedestrian victims. Figure 1 shows the lower legform test, where the lower legform is fired in the direction of the stagnant VFPS-mounted vehicle at an impact speed of 11.1 m/s. The injury parameters on legforms are extracted and used for the product rating.

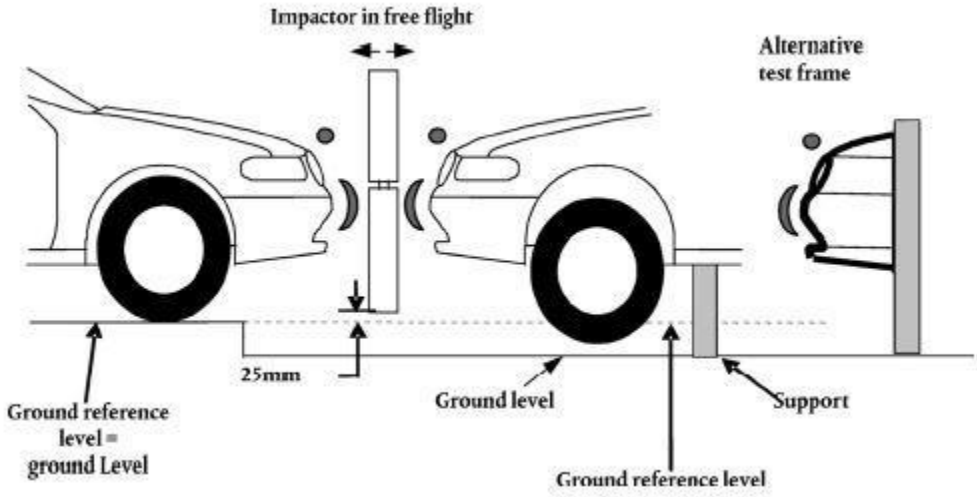


Figure 1 Lower legform to VFPS test (1)

This research involves the application of test legform tool models in LS-DYNA: TRL legform and Flex-PLI legform. TRL legform accords with the impactor in EC regulations and the Euro NCAP pre-2018 protocol [11]. As shown in Figure 2 the measuring parameters on TRL legform are: accelerator node unidirectional acceleration, knee bending angel of femur and tibia, and the shear displacement of the knee.

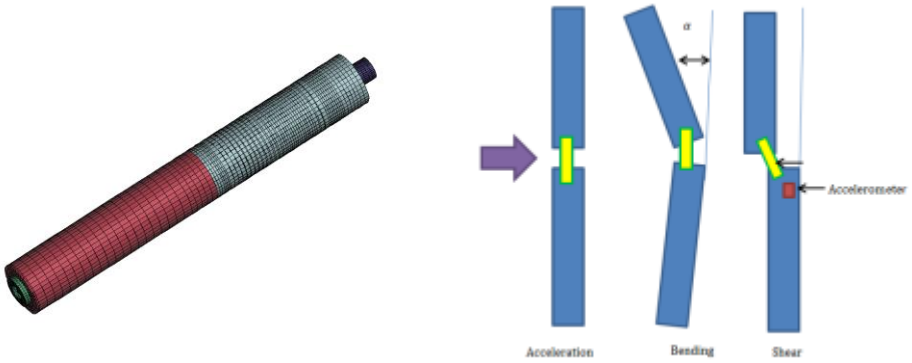


Figure 2 Left: TRL legform structure; Right: injury

Flex-PLI legform is a complex bio-fidelic legform, containing seven location bending moments and complex knee ligament injuries, used for Euro NCAP post 2018 in Australia (Figure 3) [12]. As for the injury, four tibia bending moments and four major knee ligaments: Medial collateral ligament (MCL), Anterior cruciate ligament (ACL), Posterior cruciate ligament (PCL) and Lateral collateral ligament (LCL) are designed for vehicle ratings. The FEA model is acquired through Humanetics, which is fully validated and verified for the use [14].

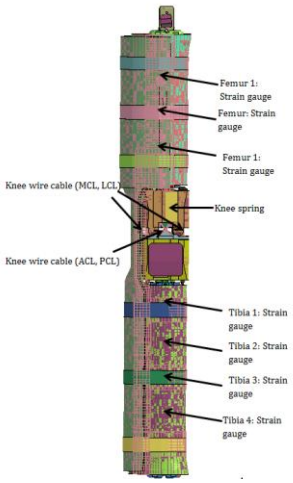


Figure 3 Left: Flex-PLI FEA model

VFPS Design Model

The VFPS model designed in this paper is a parametrically optimised nudge bar, based on the existing industry product retrofitted with a sandwich structure, consisting of a plastic cover, foam filling and metal back plate bonded with glues (Figure 4). Geometry, material, pan positioning, thicknesses are carefully chosen through an extensive parametric study conducted by the authors. The product model is designed to be tested without mounting to the vehicle, being held together by two rigid side constraints angle steels, creating worse scenario constraints for crash development.

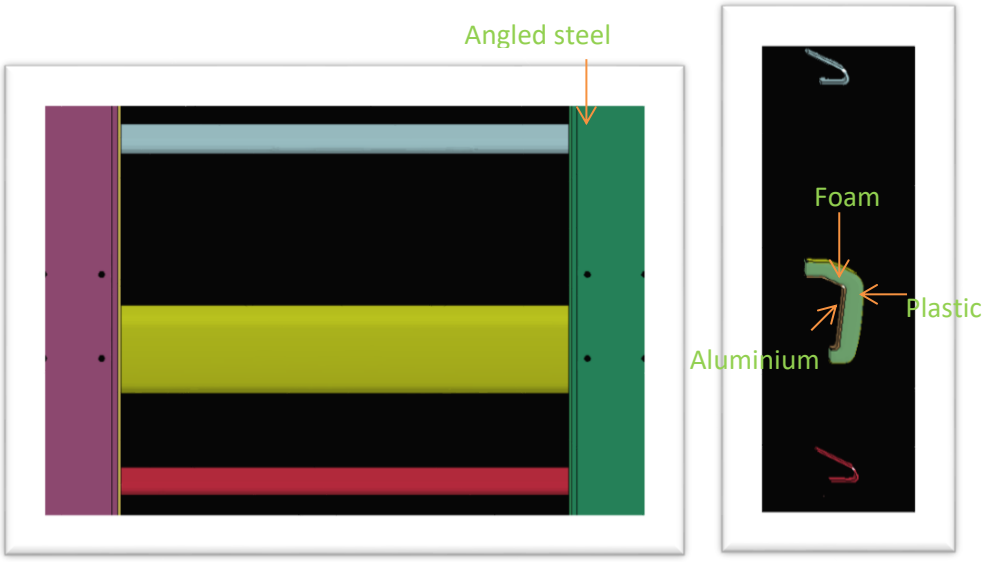


Figure 4 Left: VFPS frontal view; Right: VFPS cross section view

Part elements

The CAD CATIA model is processed and discretised in ANSYS for preparation of FEA simulation. Element-wise, Belyschoko Lin-Tsay ELFORM=2 one-integration shell elements are used across most of the parts except foam layers. Foam layer is modelled with ELFORM=2 fully integrated S/R solid elements to eliminate hourglass zero mode. To combat negative volume stability issue, besides using hex meshes instead of tet meshes, a method of coating the foam elements with a stiffer null shell closely around the surfaces was applied in the model. One tenth of the foam density was assigned to the null material *MAT_NULL so that the weight of the shell could be neglected in computation. *SECTION_SHELL was activated for a thin thickness and larger Young's modulus input to maintain desirable stability while the over-penetrations into the adjacent parts were avoided. Contact definitions are then therefore placed on the null shell.

Mesh-wise, uniformly-sized meshes are generated to produce a smooth transition between elements which leads to a more stable solution. Quadrilateral meshes (Quad) are preferred over triangular element (Tria) for shell elements. Hexahedral meshes for foam are chosen over tetrahedral. The mesh sizes in Table 1.

Table 1 Mesh sizes

Part name	Mesh size
Upper rail	5mm
Central rail	5mm
Lower rail	5mm
Foam	5mm
Plastic cover	5mm
Upright	7mm
Angled steel	7mm

Contact

As the sandwich structure is bonded together using glue, TIED_CONTACT_SURFACES_TO_SURFACES_OFFSET with SOFT=2 pinball segment-based contact is activated to tie the structure together with a modelled distance which also combats geometry irregularities. TIED_CONTACT_NODES_TO_SURFACE_OFFSET is used for foam shell bonding to the upright bracket. A CONTACT_INTERIOR is defined for foam part to resist overcompression by generating internal forces inside the foam. AUTOMATIC_SURFACE_TO_SURFACE two-sides search is used for legform to outer surfaces of the bar with SOFT=2 definition.

Material

For an accurate FEA design model, real-world materials are researched through laboratory testings to extract the most accurate information for the design. MAT_PIECEWISE_LINEAR_PLASTICITY (24) is used for steel and plastic of the sandwich, which is characterised by multi-linear strain-stress behaviour, and isotropic hardening. The necessary inputs of this material are density, Young's modulus, Poisson ratio, yield strength, failure strain and stress-strain curve in plastic region. The material tensile testings are designed to record the stress-strain behaviour of aluminium alloy and plastic cover (Figure 5a and b). Eight tests for alloy samples and ten tests for Acrylonitrile butadiene styrene (ABS) plastic samples are repeated to ensure average stress and strain behaviours are recorded. Among various foam models that can be used for FEA, MAT_FU_CHANG_FOAM (83) is chosen to simulate the polyurethane foam supplied for the sandwich because of this model's simplicity, efficiency, accuracy, and the inclusion of strain-rate effect for foam modelling. Four compression tests with four different velocities (50mm/min, 500mm/min, 1500mm/min and 3000mm/min) are designed to monitor its compression behaviour and its unloading characteristics (Figure 5 c).

Figure 5 Left: Aluminium alloy tensile test; Middle: ABS plastic tensile test; Right: foam compression test

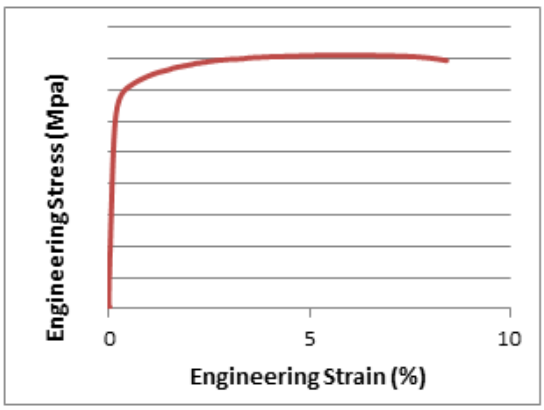
Stress-strain curves required as input for aluminium alloy and ABS plastic are expressed as true uniaxial stress and true plastic strain which are equivalent to Von Mises stress and effective plastic strain in the uniaxial case. The experimental data from the tensile test are engineering stress and strain, which are converted to true stress and strain and effective plastic strain through the formula:

$$\text{True strain} = \ln(1 + \text{engineering strain}) \quad (1)$$

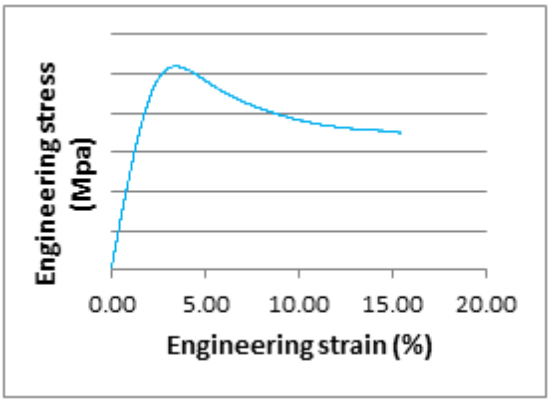
$$\text{True stress} = (\text{engineering stress}) * \exp(\text{true strain}) = (\text{engineering stress}) * (1 + \text{engineering strain}) \quad (2)$$

$$\text{Effective plastic strain} = \text{total true strain} - \frac{\text{True stress}}{\text{Young's Modulus}} \quad (3)$$

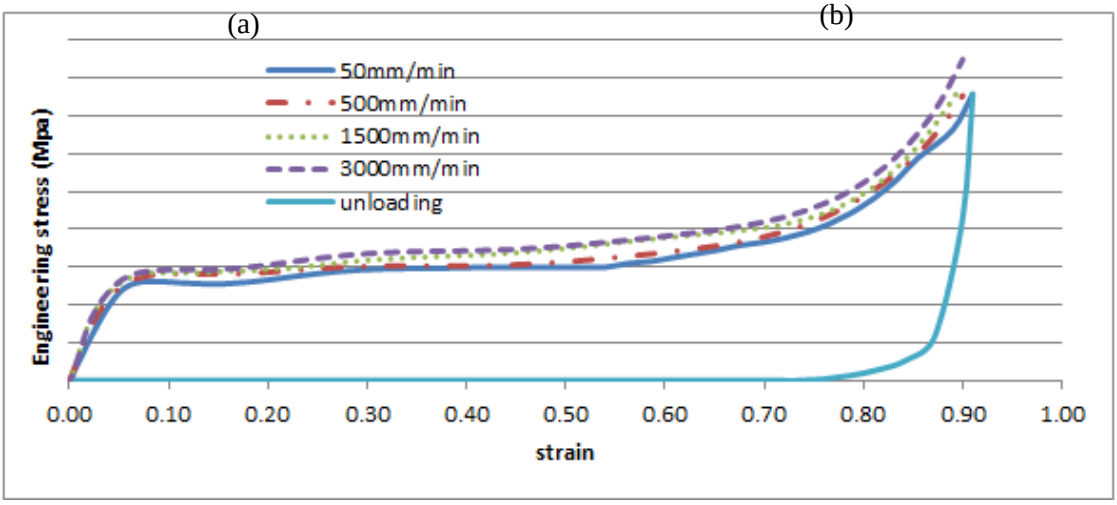
It is recommended that curves utilise minimal number of points constituting a smooth curve. A number of approximately 50 points are selected for the input curve. ABS plastic effective S-S curve is straightened up horizontally from the yield point in plastic region as MAT24 cannot handle the plastic softening [16]. MAT_FU_CHANG_FOAM allows the use of engineering stress-strain curve to define the model. The averaged engineering stress-strain curves of aluminium alloy, ABS plastics, and foams are as Figure 6.



(a)



(b)



(c)

Figure 6 Engineering Stress-strain curves. (a) Al alloy; (b) ABS plastic. (c) Polyurethane foam

As the unloading behaviour of the foam cannot simply be represented by inputting unloading curve, the alternative method of specifying HU (the hysteretic unloading factor) and SHAPE factors is used for this model as the following formula:

$$\sigma_{\varepsilon,unloading} = (1 - d)\sigma_{\varepsilon,loading} \quad (4)$$

$$d = (1 - HU)(1 - (\frac{W_{current}}{W_{max}})^{SHAPE}) \quad (5)$$

where W is the value of the absorbed hyperelastic energy per unit deformed volume [11].

Result

The virtual FEA model of the new VFPS is prepared in LS-PREPOST firstly with the TRL legform. The legform moves towards the VFPS centreline with initial velocity of 11.1m/s. Bolt holes on angle steel that are used to fix the bar to the chassis rail are locked up with Single Point Constraint nodes. Contact between legform neoprene outer skin and the PART_SET of upper rail, lower rail and plastic cover are defined by AUTOMATIC_SURFACE_TO_SURFACE. The relative height is designed to be Euro NCAP condition when the VFPS is mounted on the vehicle. In post-processing, the animations frames are recorded for the first 25ms where impact happens.

In the meantime, the VFPS is prototyped to undertake the physical crash test in the laboratory. As Figure 7, VFPS prototype is mounted on the mock chassis rail simulating the mounting on the vehicle, and a TRL legform is held by a hydraulic pusher designed with ballistic compensation to make sure the legform hits the designed first contact spot after free flight distance. The Euro NCAP standard high-speed camera captures all the happenings during the test. The data acquisition system is wired to the legform to record the critical injury parameters for the test.

Figure 7 Experimental crash test TRL legform

Comparing motion results obtained by FEA (Figure 8 (a)) and the experiment (Figure 8 (b), a similar motion trajectory can be captured in motion pictures within the quick crash window of 25ms, as shown in Figure 8.

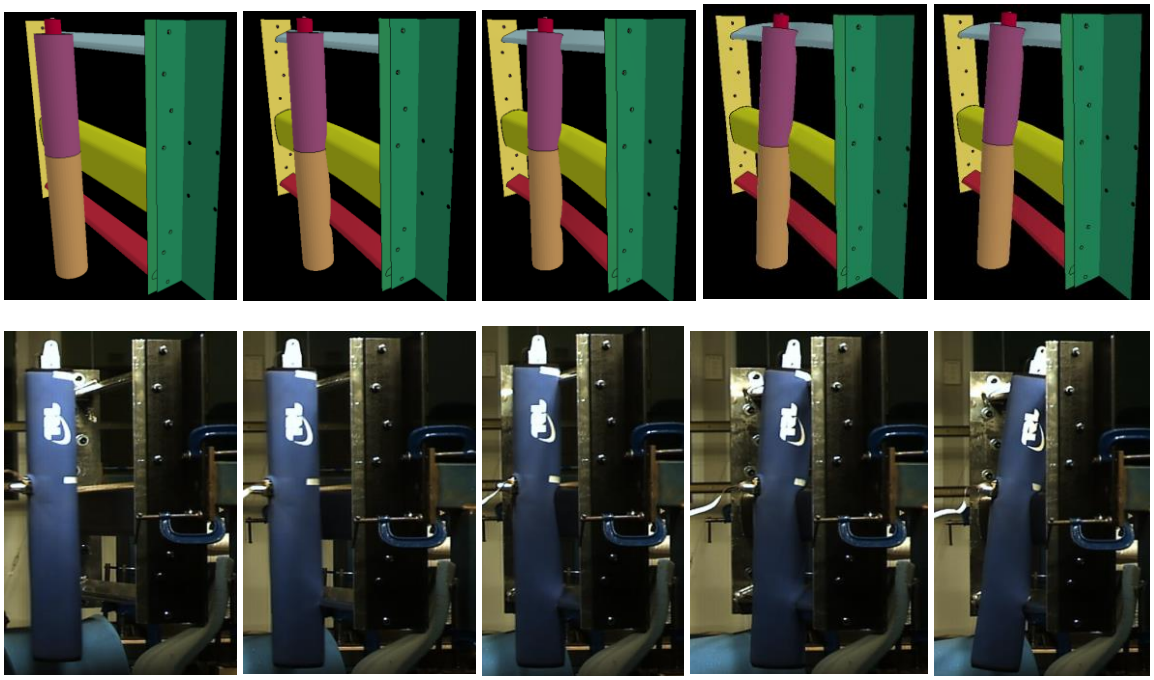


Figure 8 (a) FEA and (b) test crash frames

Figure 9 presents the comparisons of the three subtle legform injury parameters (tibia acceleration, knee shear displacement and knee bending angle) changed with time after sifting with a SAE filter of 180Hz as specified by Euro NCAP. The FEA and test curves illustrate very good agreements. This correlation has demonstrated an excellent capability of the FEA model developed in predicting the lower legform injury, in a crash scenario.

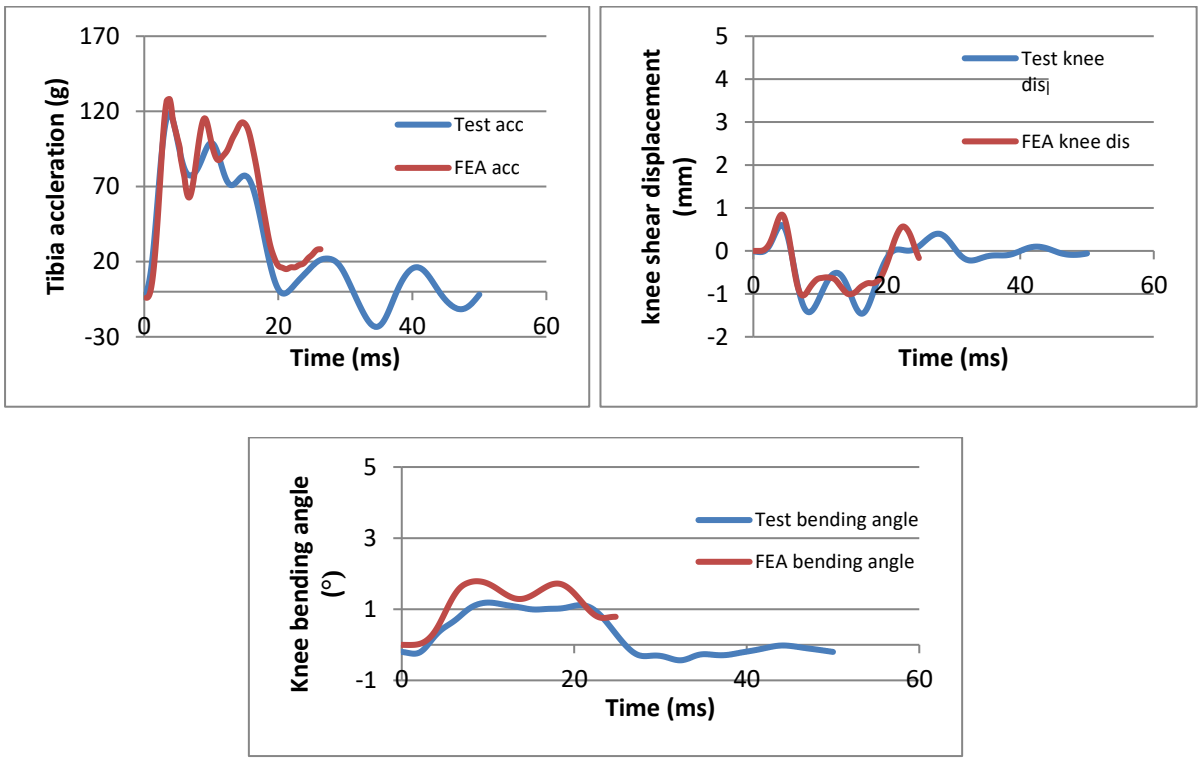


Figure 9 Test and FEA measurement results

Additionally, the leg crash injury severity caused is significantly low. All three injury parameters well exceed the Euro NCAP 5 star full performance criteria (Table 2). It is also worth noting that with this VFPS design the maximum knee displacement and maximum knee bending angle are very low (1.5mm and 1.195°) when the five star margin are much still higher (6mm and 15°). This means the safety design of this model is extremely effective in protecting and cushioning human knees and legs. This performance signifies an exceedingly high industry standard for vehicle product in terms of safety.

Table 2 Test results and Euro NCAP protocol 5.3.1 [12]

Injury parameters	Euro NCAP 5 star full mark	Test	FEA
Maximum tibia acceleration	<150g	118g	128g
Maximum knee shear displacement	<6mm	1.5mm	1.03mm
Maximum knee bending angle	<15°	1.195°	1.78°

Having built the validated VFPS model, it is of many manufacturers' interests to witness how the generation 2.0 Flex-PLI legform performs when crashes the VFPS design, with a design vision of the future. Flex-PLI FEA model is set up with a height of its bottom 50mm lower than the TRL legform in accordance with the Euro NCAP protocol 8.1 [13] (Figure 10). The contact of AUTOMATIC_SURFACE_TO_SURFACE is defined between the Flex-PLI skin null shell and the VFPS contact surfaces. Four lower tibia bending moments are measured and knee ligament injuries (MCL, PCL, ACL and LCL) are reflected from the discrete elements change length.

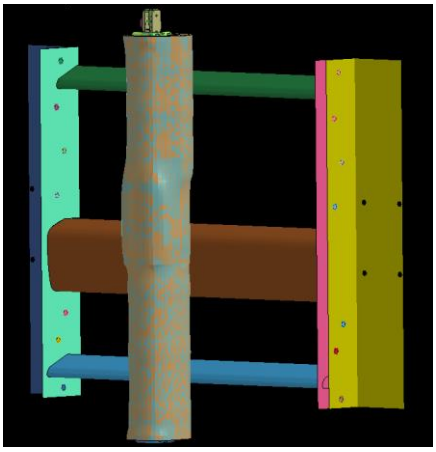


Figure 10 Flex-PLI FEA crash test

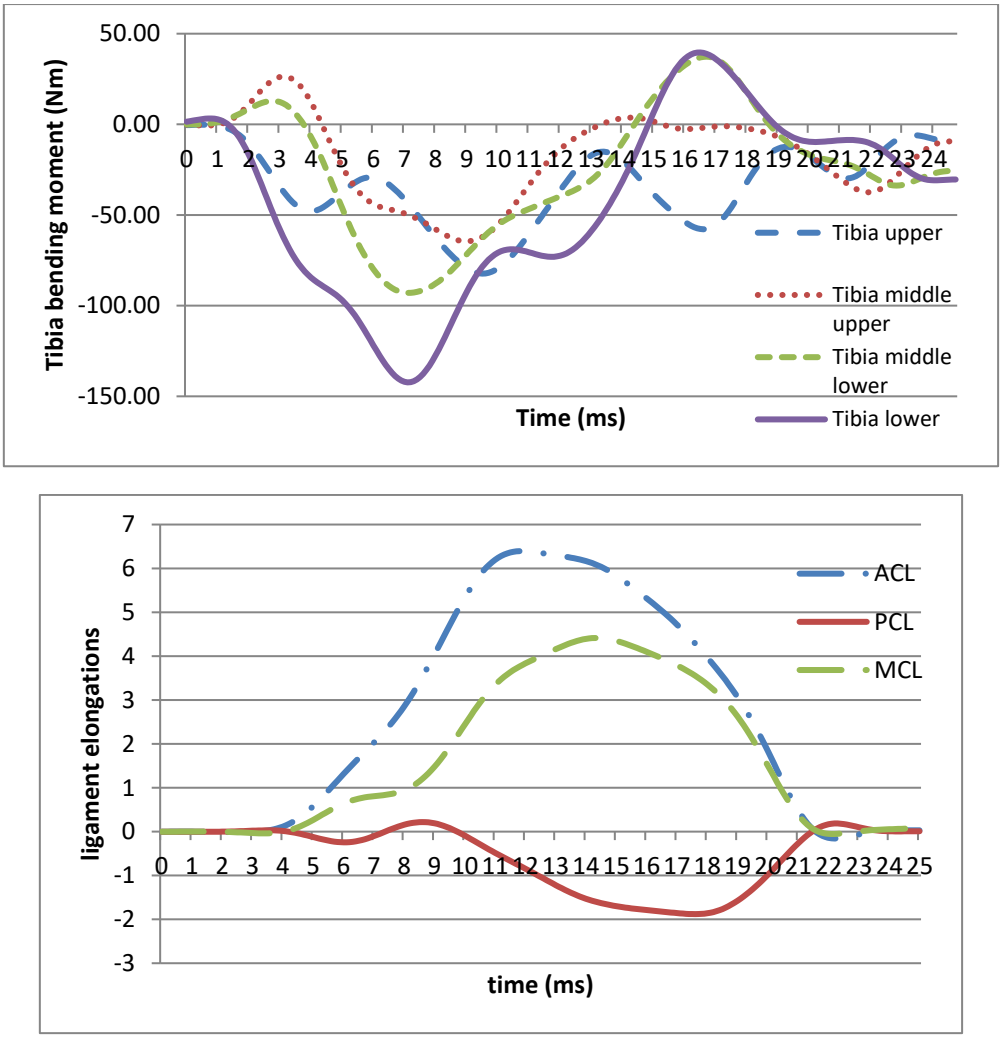


Figure 11 Flex-PLI Test results

As shown in Figure 11, the Flex-PLI crash test also produced a five-star rating for the VFPS as the maximum tibia bending moment 141.97Nm is much lower than the 2018 five star criteria of 200Nm, maximum ACL,PCL elongations 6.39mm are within 10mm, and maximum MCL elongation is 4.41mm as shown in Table 3. This reflects the VFPS also satisfies the top performance criteria of the future Flex-PLI legform.

Table 3 Flex-PLI test results and Euro NCAP protocol 8.1 [13]

Injury parameters	Euro NCAP 5 star full mark	FEA
Maximum Tibia acceleration	<282Nm	141.97Nm
MCL Elongation	<19mm	4.41mm
ACL/PCL Elongation	<10mm	6.39mm

Discussion and conclusions

In this paper, a sandwich structure three-pan VFPS is designed and evaluated using FEA model and experimental testing to meet current and future Euro NCAP pedestrian safety requirements. The correlated results in TRL test indicated that the VFPS FEA model developed is highly reliable in predicting pedestrian safety crash results for design of the new frontal protection system. The following main conclusions can be drawn from this study.

The proposed VFPS design well exceeds the highest performance criteria in terms of the requirements of Euro NCAP protocol 5.3.1 for TRL legform test. The injury readings from both FEA and experimental tests report satisfactory peak legform tibia acceleration, knee shear displacement and knee bending angle.

Using the validated model to evaluate a Flex-PLI legform crash test, the design can also well meet all optimum safety requirements. The maximum tibia bending moment of 141.97Nm, maximum MCL of 4.41mm and maximum ACL/PCL of 6.39 all well satisfy the five star performance level. The results will satisfy the criteria of Euro NCAP protocol 8.1.

The successful design and FEA modelling of the VFPS are well demonstrated through this process. The choice of the design parameters can well address the pedestrian injury during a frontal crash. Further work will be carried out with the assistance of this virtual and physical platform in the optimal product design study using Flex-PLI.

Reference

[1] Attewell, R. and K. Glase, Bull Bars and Road Trauma. 2000.

[2] Department of Infrastructure and Transport. Regulation Impact for Statement for Pedestrian Safety, Branch Department of Infrastructure and Transport Standards and International Vehicle Safety Standards, Editor. 2011.

[3] Australia Standards, 4876.1-2002 Motor vehicle frontal protection systems Part 1: Road user protection. 2002.

[4] Charity, S., Vehicle Frontal Protection Systems (Bull bars). 2010, Australian Automotive Aftermarket Association Ltd.
[5] King, C., Government will not ban bull bars. 2011.

[6] Pohlak, M., J. Majak, and M. Eerme, Engineering optimization of a car frontal protection system component. Estonian Journal of Engineering, 2009. 15(1): p. 61-72.

[7] Ptak, M., Rusinski, E., Kopczynski, A., Harnatkiewicz, P., Kaczynski, P., Virtual testing in terms of pedestrian safety improvements. 2009.

[8] Kopczyński, A., M. Ptak, and P. Harnatkiewicz, The influence of frontal protection system design on pedestrian passive safety. Archives of Civil and Mechanical Engineering, 2011. 11(2): p. 345-364.

[9] Brooks, R., A materials and structure perspective on the feasibility of automotive frontal protection systems meeting the proposed pedestrian safety test criteria. Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials Design and Applications, 2006. 220(2): p. 67-78.

[10] Mariusz Ptak, J.K., Numerical Investigation of the Frontal Protection System for Pedestrian Safety Enhancement, in IRCOBI Conference 2013. 2013.

[11]Hallquist, J.O., LS-DYNA theory manual. Livermore software Technology corporation, 2006. 3.

[12] Euro NCAP, EUROPEAN NEW CAR ASSESSMENT PROGRAMME (Euro NCAP) PEDESTRIAN TESTING PROTOCOL Version 5.3.1, E. NCAP, Editor. 2011. p. 63.

[13] Euro NCAP, EUROPEAN NEW CAR ASSESSMENT PROGRAMME (Euro NCAP) PEDESTRIAN TESTING PROTOCOL Version 8.1 , E. NCAP, Editor. 2015. p. 57.

[14] Humanetics (2016). "Flex-PLI-GTR." from http://www.humaneticsatd.com/crash-test-dummies/pedestrian/flex-pli-gtr.

[15] Kopczyński, A., M. Ptak, and P. Harnatkiewicz, The influence of frontal protection system design on pedestrian passive safety. Archives of Civil and Mechanical Engineering, 2011. 11(2): p. 345-364.

[16] Anton, Schmailzl, Amann Thomas, Glockner Markus, and Fadanelli Martin. "Finite element analysis of thermoplastic probes under tensile load using LS-DYNA compared to ANSYS WB 14 in correlation to experimental investigations." In Proceedings of the ANSYS conference & 30th CADFEM users' meeting. 2012.

Figures

(1) European Commission, COMMISSION REGULATION (EC) No 631/2009 of 22 July 2009 laying down detailed rules for the implementation of Annex I to Regulation (EC) No 78/2009 of the European Parliament and of the Council on the type-approval of motor vehicles with regard to the protection of pedestrians and other vulnerable road users, amending Directive 2007/46/EC and repealing Directives 2003/102/EC and 2005/66/EC, 2009. Figure 0-32 Lower legform to VFPS test; p.54; Figure 0-33 Left: Upper legform to bonnet leading test. Right: Upper legform to VFPS leading edge test; p.56

Study on the effects of the psoas major muscle and facet joint orientation on the intradiscal stress of the lumbar spine

Shuo Chen[1], Qiang Chen*[,1], Zhi-Yong Li*[,1,2]

[1]Biomechanics Laboratory, School of Biological Science & Medical Engineering, Southeast University, 210096, Nanjing, P.R. China

[2]School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology(QUT), Brisbane, QLD 4001, Australia

Abstract:

Background: The psoas major muscle plays an important role in sharing the external loads applied on the lumbar spine, and the facet joint orientation is considered as a congenital anatomical factor of lumbar degeneration. However, since the effects of the two factors on the stress distribution of the lumbar spine have rarely been studied, the influences of the psoas major muscle and fact joint orientation on the intradiscal stress are here investigated.

Method: To consider the influence of the psoas major muscle, a lumbar spine of a health volunteer was first scanned, and the obtained CT images were used to establish the finite element (FE) lumbar model A by the software Simpleware. Then, based on the model A, a model B with the muscle was developed, and further the model B was employed to construct three facet joint orientations (F45, F50, F55) of the L3-L4 lumbar segment. All the FE model were import into the FE software ABAQUS to perform simulations of six motions including flexion, extension, left and right lateral bending, left and right rotation.

Result: Under flexion, the intradiscal stress of the model B was 18.99% lower than the model A, and 23.42% lower in extension. In the left and right lateral bending, the stress was reduced by 36.63% and 27.20%, respectively. But in the left and right rotation, the changes of intradiscal stress were only 1.49% and 2.97%. In the lumbar flexion and extension, the orientation F55 showed the maximum intradiscal stress, and the orientationF45 showed the maximum intradiscal stress in the rotation, but all the orientations (F45, F50 and F55)share the similar intradiscal stress in lateral bending.

Conclusion: The psoas major muscle alleviates the intradiscal stress, and different facet joint orientations dominate the stress in different motions. This work could be useful for the researchers and clinicians to evaluate the stress distribution of the lumbar spine.

Keywords: Lumbar spine, Psoas major muscle, Facet joint orientation, Intradiscal stress, Finite element analysis.

1. Introduction

Degenerative lumbar scoliosis (DLS) often develops in the L4-L5 lumbar spine, and its occurrence is increasing with age. To prevent the lumbar degeneration, it is necessary to identify the factors that promote the degeneration.

Facet joint is formed by the superior facet and inferior facet, and it is regarded as a congenital anatomical factor in lumbar degeneration, and as one of the primary structures of the spine, it prevents excessive translation and rotation, and thus stabilizes the spinal motion. The facet joint orientation is defined as the direction of the joint surface and the coronal plane, and it influences intervertebral joint translation and rotational displacement, in other words, with different facet joint angles, the displacement of the intervertebral joint can be very different. SamartzisD et al. [1] reported that the sagittal facet angle was associated with L4-L5 degeneration in lumbar spine through the study of 349 patients, and the critical value of the facet joint angle was about 32 degrees. Therefore, the facet joint orientation is a potential anatomical predisposing factor of the lumbar degeneration that may lead to early degeneration of the intervertebral disc or degenerative spondylolisthesis [2]. Moreover, literature reported that experiments revealed the stabilizing effect of the psoas major muscle on the lumbar spine under the static state, and driving effect of the muscle on the dynamic extension of spine [3], but the biomechanical effects have not yet been validated.

Finite element method is often used to simulate the complex spinal system and using this method, Kim et al. [2] showed that a facet tropism structure could make the L3-L4 segment more vulnerable to external loads, and Christophy et al. [4] modelled the psoas major muscle in a finite element lumbar spine as springs which connected the L1-L4 vertebral body and the small rotor of femur. Hence, the finite element method is expected to be an effective tool to quantify the biomechanical effect of facet joint orientation and psoas major muscle.
In this study, the biomechanical effect of the psoas major muscle and the facet joint orientation on the intradiscal stress under six specific motions (i.e., flexion, extension, left and right lateral bending, left and right torsion) are studied.

2. Finite Element Method

To investigate the effects of the psoas major muscle and the facet joint orientation on the intradiscal stress, a normal lumbar model A without the muscle and model B with the muscle were established, and based on the model B, three facet joint orientations were considered, thus there are four cases in total. For the four cases, six motions (i.e., flexion, extension, left and right lateral bending, left and right torsion) were considered.

Geometrical models

Model A without the psoas major muscle: L1-L5 CT images of the lumbar spine from a healthy adult male (37-years-old, 75 kg, 178 cm)were used to establish a normal finite

element lumbar model A by the ScanIP module of the commercial software Simpleware (*Synopsys Inc*, UK), and the model mesh was generated by its mesh module (Figure 1B).

Model B with the psoas major muscle: According to Christophy et al [4], we first determined the connecting position of the psoas major muscle, namely, one end of the muscle is located at the rear side of the low edges of L1-L4 vertebrae, and the other is at the end of the small rotor of femur(Figure 1C). Here, under the six motions, the pelvis was considered to be motionless, and this was reasonable when the motions occur from up-right standing state. Therefore, two fixed reference points replaced the two connecting positions at the small rotor of the femur, and the coordinates of the reference points were determined by measuring the distance between the small rotor and the lumbar spine on the CT images of the volunteer. After determining the connection of the muscle, it was incorporated into the meshed model A in the commercial finite element software Abaqus (*SimuliaInc*, USA), see Figure 1D.

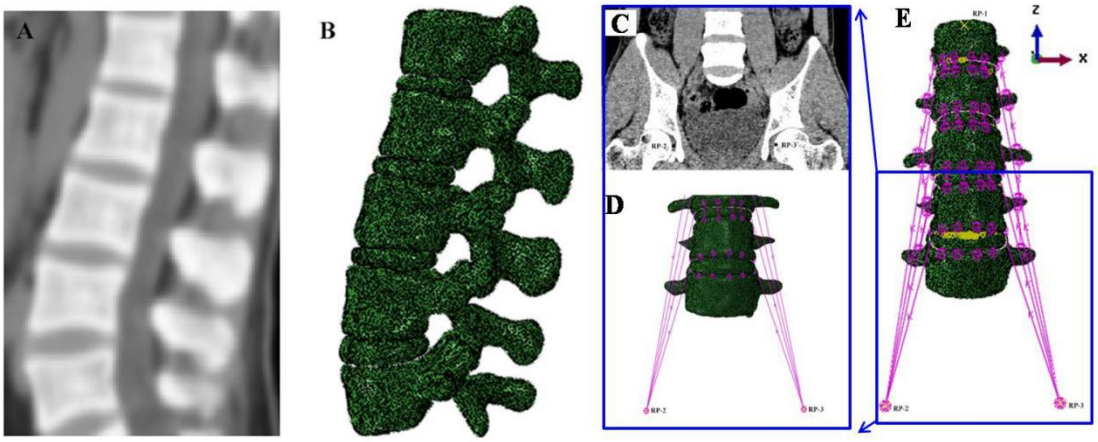


Figure 1.Three-dimensional finite element model of a lumbar spine of the volunteer. (A) A CT image with L1-L5 lumbar segments on the saggital plane; (B) The normal finite element lumbar model; (C) The CT image with L4/L5 lumbar segment and pelvis; (D) L4/L5 lumbar segment with the psoas major muscle. (E) The whole model with psoas major muscle.

Facet joint orientations of L3-L4 lumbar segment: Here, on the basis of clinical experience, different facet joint orientations were only considered to occur at the L3-L4 lumbar segment of the model B, which have a facet joint angle of 50^{o} (marked by F50). To construct the other two lumbar models with L3-L4 facet joint orientations of 45^{o} and 55^{o} (marked by F45 and F55, respectively), the CT images of L3-L4 lumbar segment were modified in the Simplware by the fusion and segmentation operations.

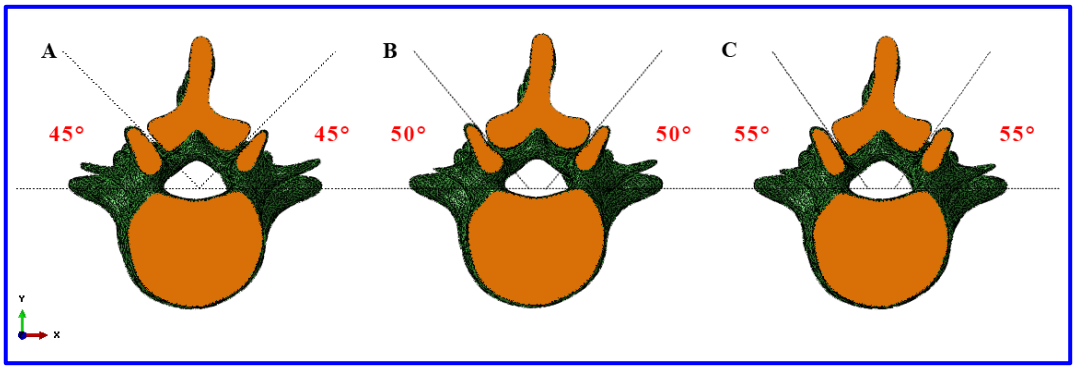


Figure 2.Facet joint orientations with different facet joints angles. (A) F45; (B)F50; (C)F55.

Material properties of components and boundary conditions

The lumbar model was divided into two parts, the vertebral body (L1-L5) and the intervertebral disc (D1-D4).For the vertebral body, it was constituted by four components, i.e., articular cartilage, cortical bone, cancellous bone, posterior structure where the facet joint belongs, andthe contact between the four components were all coupled. The thickness of the articular cartilage was defined as 0.5 mm, and that of the cortical bone as 3 mm, and they were assigned by shell element; and the cancellous bone and posterior structure were assigned by the C3D4solidelement.Moreover, the surface-to-surface contact was used between the superior and inferior articular process. For the intervertebral disc, it consisted of fibrous ring and nucleus pulposus, and the contact between vertebral body and fibrous ring was tied, plus the ligament and muscle were tied to the vertebral body, and they were assigned with the spring element [5].

All the materials in the models were simplified to be isotropic and linear elastic, and nucleus pulposus was approximately modelled as an incompressible material. Their Young's modulus and Poisson's ratios were listed in Table 1.

For all the motions, a 400N concentration force was applied on the reference point located at the centre of the top surface of the first lumbar segment (L1) to simulate the body weight above the L1, and the loading direction is along the z axis. Plus, due to bending moment induced by the six motions, a 10N·m moment was also applied on the reference point [9].Then, all the casesweresimulated by the software Abaqus (*SimuliaInc*, USA) to study the intradiscal stress undersix motions.

Table 1.Material properties employed in the FE models

Components	Young's modulus(MPa)	Poisson's ratio
Cortical bone	12000	0.3[6]
Cancellous bone	100	0.2[6]
Posterior element	3500	0.25[6]
Articular cartilage	24	0.4[7]
Fibrous ring	4.2	0.45[6]
nucleus pulposus	1.0	0.4999[8]
Anterior longitudinal ligament	20	0.3[9]
Posterior longitudinal ligament	70	
Ligamentum flavum	50	
Interspinous ligament	28	
Supraspinous ligament	28	
Intertransverse ligament	28	
Capsular ligament	20	
Psoas major	45[4]	

3. Model Validation

The rotational displacement of each intervertebral disc with respect to its bottom surface or the rotational stiffness defined as the moment to the rotational displacement, were calculated in flexion and extension states to validate the model. For the rotational stiffness, a concentrate force of 400N and a moment of 5 N·m were applied to the model A to compare the present result with Kim et al. [10], see Figure 3A. It shows a weak difference of the results by Kim et al., and this study presents similar stiffness of D1 and D2 as the literature, but D3 and D4 showed a large discrepancy between the present study and literature. Moreover, to further validate the model, a concentrate force of 150N and a moment of 10 N·m were applied to the model A to compare with the results of the rotational displacement by Yamamoto et al.[12] and Shirazi-Adl et al.[13]. Similarly as the comparison of the rotational stiffness, the results of D1 and D2 are comparable, and D3 and D4 are apparently different, in particular, the rotation angle of D4 was almost zero, see Figure 3B,C. This is because the model in the literature contains the sacrum and an extra disc between the L5 and the sacrum, whose bottom was fixed, and the model is more flexible, thus the stiffness is less than the present work, but the rotational displacement is greater than the present study, which does not have the sacrum and fixed at the bottom of the L5.

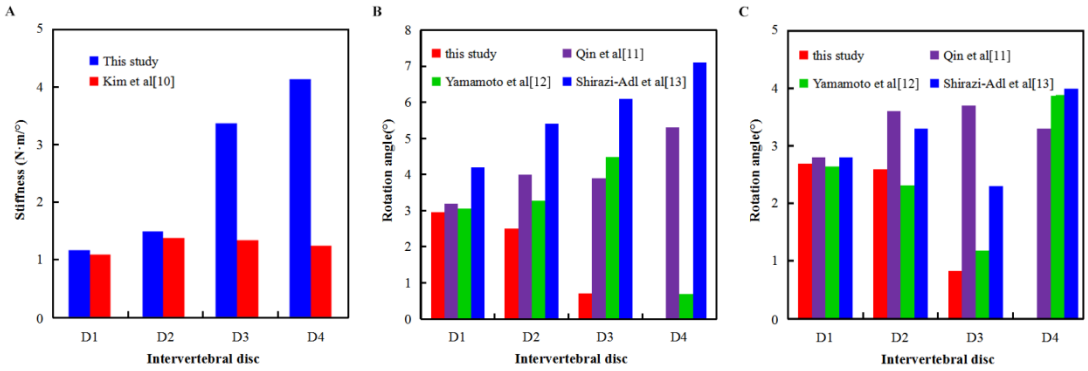


Figure 3.Validation of the original model. (A) Stiffness comparison with Kim et al.[10]; (B/C) Rotation angle comparison with Qin, Yamamoto and Shirazi-Adl et al. under flexion/extension.

4. Results

Effect of the psoas major muscle

Figure 4 shows the effect of the psoas major muscle on the greatest intradiscal stress of all discs for all the motions. Generally, the stress of the model A without the muscle is greater than that of the model B, and the maximum stress is on the D1. Moreover, the stress is in the order of D1 > D2 > D3 > D4 except in the extension, and under the extension, the variation of stress on last D2-D4 discs was opposite to the others. In detail, without the muscle, the stress of the model A increased by 18.99% under the flexion; under extension, the increment is 23.42%; the stress increase was 36.63% for left bending and 27.20% for right bending, respectively; but the stress increase was only 1.49% (left rotation) and 2.97% (right rotation). This indicates that the psoas major muscle bears a part of the load induced by the motions.

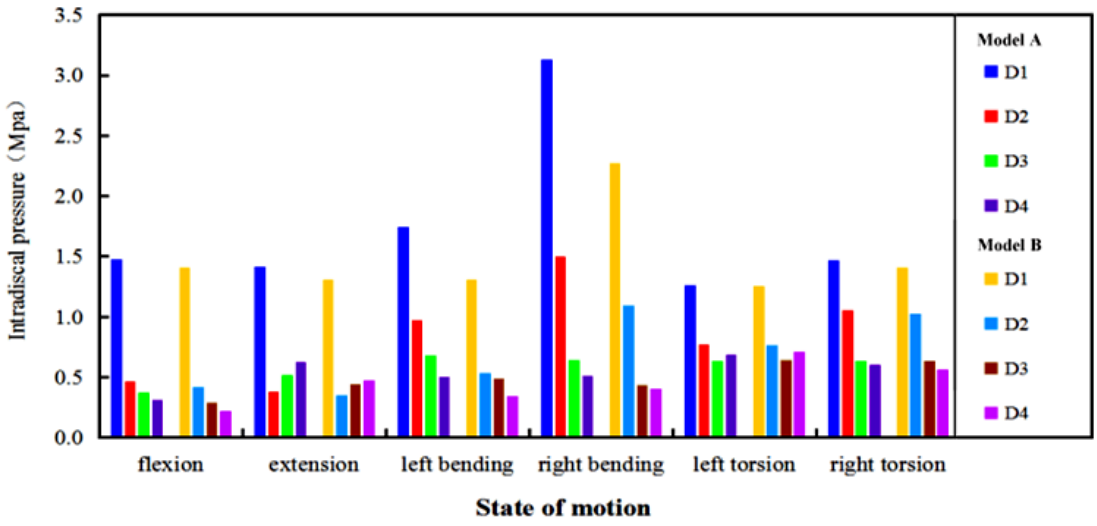


Figure 4.The intradiscal stress of the models A and B under six motions

Effect of facet joint orientation

Figure 6 shows the effect of three facet joint orientations on the greatest intradiscal stress of all discs of the model B for all the motions. Generally, they eakly differ, and the greater stress and stress discrepancy between the three orientations is at the upper disc, i.e., D1, and there is not common variation. In detail, as mentioned in Section 2, only the facet joint orientation of the L3-L4 was changed. In this regard, under the flexion and extension, the orientation produce a common influence on the D3, i.e., F45<F50<F55; under the left and right bending, the influence on the D3 are very weak, especially for the right bending; for the left and right rotation, the common order isF50<F55<F45 for the stress on the D3.This means that people with larger facet joint angle would more easily get injured for the flexion and extension, and people with less facet joint angle get injured for the rotation.

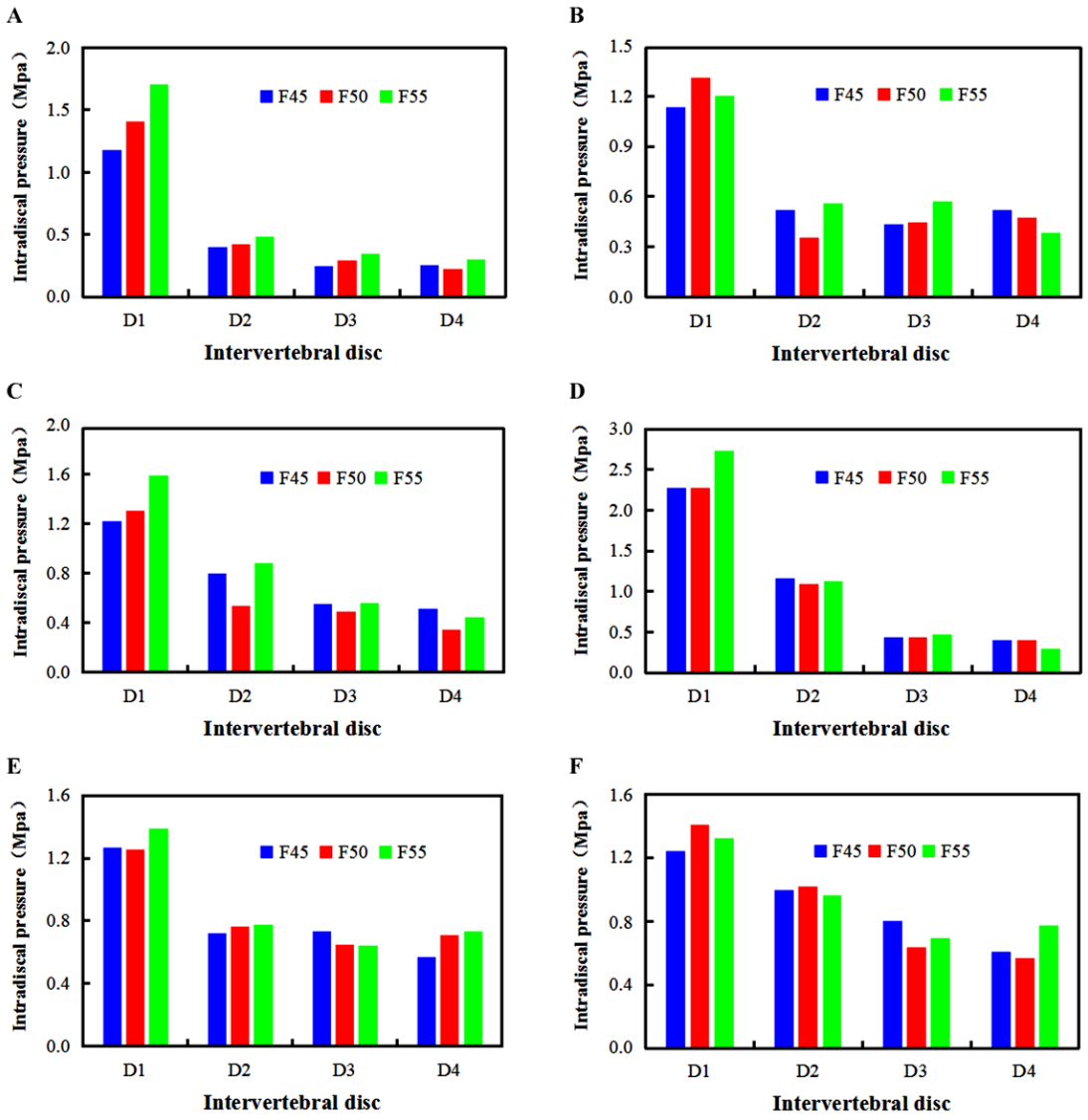


Figure 5.Intradiscalstress of D1-D4intervertebral discs in the three orientations (F45, F50 and F55). (A) Flexion (B) Extension (C) Left bending (D) Right bending (E) Left rotation (F) Right rotation.

5. Discussions and Conclusions

Since the psoas major muscle and facet joint orientation play an important role maintaining the stability of lumbar spine, this study aims to use finite element method to study their effects on the intradiscal stress in several motions.

Firstly, the model validation by comparing the present result with the literature, to some extent, they are in an agreement, and the difference mainly caused by the different geometrical models, but it still illustrate the availability of the present model.

For the effect of the muscle, it proved that the inclusion of the muscle in the model indeed shared the load of the body weight and applied moment, and reduces the stress magnitude of the intradiscal stress. This also illustrates that the finite element models in literature without the muscle over-estimated the stress distribution. As for the facet joint orientation, compared to other literature, our study provides an effective method to construct the model with different orientations, and a comprehensive research on the biomechanical mechanism of the facet joint orientation in occurrence and development of degenerative lumbar scoliosis based on the intradiscal stress. It is worth mentioning that only the facet joint orientation D3 is considered. Plus, the greatest intradiscal stress always locates at the fibrous ring, which may indicate the position of back pain or other clinical symptoms.

In summary, we have studied the effects of the psoas major muscle and facet joint orientation on the intradiscal stress of the lumbar spine. The maximum intradiscal stress always located at the first disc D1, and the psoas major muscle indeed reduces the stress of the intradiscal stress and improves the prediction. Different facet joint orientation induces maximum intradiscal stress under different motions. The present study could be helpful to the researchers or clinician to treat relevant lumbar diseases.

Acknowledgements

This work is partially supported by the National Natural Science Foundation of China (NSFC) (Nos. 31300780, 11272091, 11422222, 31470043), the Fundamental Research Funds for the Central Universities (No. 2242016R30014), and ARC (FT140101152).

References

[1] Samartzis D, Cheung JPY, Rajasekaran S, et al. Critical values of facet joint angulation and tropism in the development of lumbar degenerative spondylolisthesis: An international, large-scale multicenter study by the AO spine asiapacificresearch collaboration consortium. Global Spine Journal, 2016, 6:414–21.

[2] Kim HJ, Chun HJ, Lee HM, et al. The biomechanical influence of the facet joint orientation and the facet tropism in the lumbar spine. The Spine Journal, 2013, 13(10):1301–8.

[3] Wei YZ, Xie B, Tan SS, et al. The tension effects of psoas muscle on the extended stress of the spine: Biomechanical research. Chinese Journal of Clinical Anatomy, 2008.

[4] Christophy M, Senan NAF, Lotz JC, et al. A musculoskeletal model for the lumbar spine. Biomechanics and Modeling in Mechanobiology, 2012, 11(1):19–34.

[5] Lorenz M, Patwardhan A, Vanderby R Jr. Load-bearing characteristics of lumbar facets in normal and surgically altered spinal segments. Spine (Phila Pa 1976) 1983, 8(2):122–30.

[6] Du CF, Yang N, Guo JC, et al. Biomechanical response of lumbar facet joints under follower preload: a finite element study. BMC Musculoskeletal Disorders, 2016, 17(1):1.

[7] Moramarco V, del Palomar AP, Pappalettere C, et al. An accurate validation of a computational model of a human lumbosacral segment. Journal of Biomechanics, 2010, 43(2):334–42.

[8] Huang J, Li H, Wu H. Simulation calculation on biomechanical properties of lumbar disc herniation. Journal of Medical Biomechanics, 2012, 27(1):96–101. (In Chinese)

[9] Wang L, Zhang BK, Chen S, et al. A validated finite element analysis of facet joint stress in degenerative lumbar scoliosis. World Neurosurgery, 2016, 95:126–33.

[10] Kim KT, Lee SH, Suk KS, et al. Biomechanical changes of the lumbar segment after total disc replacement: Charite, Prodisc and Maverick using finite element model study. Journal of Korean Neurosurgical Society, 2010, 47(6):446–53.

[11] Qin JS, Wang Y, Peng XQ, et al. Three-dimensional finite element modeling of whole lumbar spine and its biomechanical analysis. Journal of Medical Biomechanics, 2013, 28(3):03–05. (Chinese)

[12] Yamamoto I, Panjabi MM, Crisco T, et al. Three-dimensional movements of the whole lumbar spine and lumbosacral joint. Spine, 1989, 14(11):1256–60.

[13] Shirazi-adl SA, Shrivastava SC, Ahmed AM. Stress analysis of the lumbar disc-body unit in compression-A three dimensional nonlinear finite element study. Spine, 1984, 9(2):120–34.

Investigating the Effect of Rock Pore Size Distribution on Reservoir Production Performance

†S. Rezaei-Gomari , and F. Amrouche*

School of science and Engineering, Teesside University, Middlesbrough TS1 3BX, United Kingdom

*Presenting author: faridaamrouche@yahoo.fr

†Corresponding author: S.Rezaei-Gomari@tees.ac.uk,

Abstract

Rock pore size distribution is one of the most important parameter that can affect the reservoir depletion during the water flooding process. In this paper the effect of the pore size distribution from micro, to macro pore sizes on the capillary pressure curves, relative permeability curves and the remaining oil saturation is investigated.
Gaussian distribution was developed for micro, meso and macro pore sizes assuming that the pores are bundle of tubes with cylindrical shape for simplicity. The results from capillary pressure curves illustrate the invasion paths in pores while the relative permeability curves present the rock grain sorting. Purcell's correlation is used to calculate the data related to the relative permeability. The obtained fluid flow curves for different pore size distributions are then implemented in a simple reservoir model to evaluate the reservoir respond for water flooding. The results from simulation show that the oil recovery increases with increasing of the number of the macro pores follows by the meso and micro pore sizes. Moreover it is observed that the time of water breakthrough is highly dependent on the type of pores where the late water breakthrough time is obtained for the micro pores. This study concludes that the reservoir performances during water flooding and hence reservoir fluid production are highly dependent on the rock typing and the pore size distribution.

Keywords: Pore size distribution, Capillary pressure, Relative permeability, Residual oil saturation, Gaussian distribution.

Introduction

Residual Oil Saturation or Remaining Oil Saturation (ROS) refers to the oil saturation left after the injection of a specific volume of displacement fluid (gas or water) into an oil bearing reservoir [1]. The assessment of ROS is necessary to determine whether an oil bearing reservoir has reached a limit beyond which production will not be economically possible. The proportion of ROS that can be moved with a required number of pore volumes of flooding fluid (gas or water) helps to determine the economic viability of production compared with the cost of fluid injection. The ROS is a function of the volume of fluid injected, wettability, heterogeneity of the reservoir and pore size distribution.

According to Shedid [2], pore-size distribution is considered critical to the displacement efficiency of oil bearing reservoirs. The pore-size distribution of rocks influences porosity and permeability of the reservoir, thereby affecting the ROS; hence it is important to study the pore-

size distribution in rocks to understand the flow processes within a porous matrix and the performance of the reservoir in general. Beiranvand [3] explained that pore-size distribution is also related to a variation in the sorting and packing of grain size of rocks. Gaussian distributions were carried out to study variety of rock pore size distributions of macro, micro, and meso. The reservoir performance in respond to the aforementioned pore size distributions are then investigated using standard oil and gas reservoir software know as Eclipse 100.

Methods and techniques

In the oil and gas reservoirs, the link between permeability and hydrocarbon saturation is the distribution of pore channel sizes, represented by 1/r. More over the pore sizes govern the minimum threshold pressure which in turn results in wetting phase being displaced by non-wetting phase. To study the reservoir performance in term of hydrocarbon production and its dependency to pore size distribution following steps are deployed.

Step 1: Three well defined pore size distributions are considered namely micro, meso and macro.

$$P_C = \frac{2 \times \sigma \times \cos(\theta)}{r} \quad (1)$$

Step 2: Compute capillary pressure incrementally for the selected pore sizes from higher values to lower with an incremental value of 1μm assuming pores as bundle of tubes by applying the following equation:

Step 3: The pore size distributions are modelled with Gaussian Distribution Function (GDF).

Step 4: The total volume of invaded pores are computed by calculating of the volume of each single pore size multiplied by the number of pores invaded for three different chosen models.

Step 5: The developed porous models are used to compute the relative permeability using Purcell Equations

Step 6: The computed relative permeabilities for three sets of porous media are included into a simple Eclipse model to investigate the reservoir performance in response to the changes in pore size distribution.

The source of data to calculate capillary pressure, relative permeabilities as well as the pore volume are given in the following table.

Table 1. Sources of the data used in this study

Parameters	Values	Source
Interfacial tension σ (dynes/cm)	27	Ling [4]
Contact angle θ in degrees	0	Ling [4]
Core Length (inch)	2	Ling [4]
Inter-particle pore size	Micro: [10-50μm]	Lønøy [5]
	Mesopres : [50-100μm]	Lønøy [5]
	Macro : >100μm	Lønøy [5]

Results and discussion

Effect of pore size distribution on capillary pressure curves and relative permeability curves

Fig.1a shows the capillary pressure curves for different ranges of inter-particles pore-size. It can be observed that the calculated Pc for macro size pores is the lowest at equal saturation with meso and micro. This indicates that a small pressure is required to invade the macro size pores compared to meso and micro. In line with the Pc, the required threshold pressure (entry pressure) for macro size pore is smaller than meso and micro which reveals that when there is an increase in permeability of the pore (macro), there would be subsequent decreases in capillary pressure.

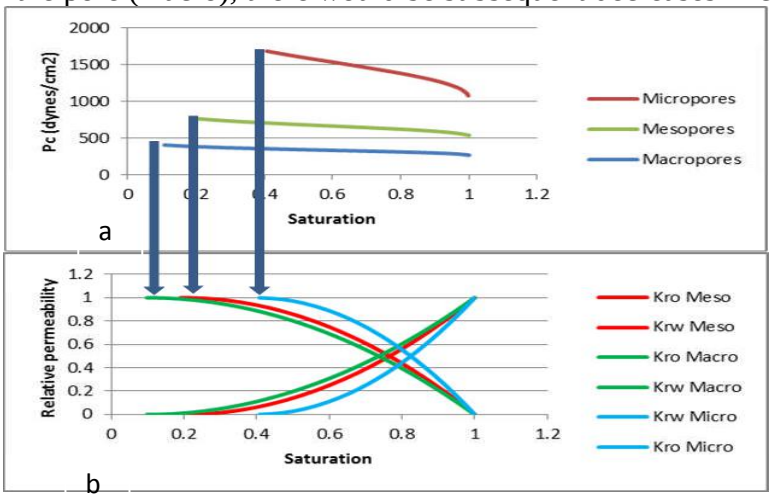


Figure 1. Calculated capillary pressure (a) and relative permeability curves (b) for the three pre-defined pore size distribution models.

For the defined rock models, the relative permeability curves are also calculated using Purcell approach and results are presented in Figure 1b. The different values of irreducible water saturation Swir (Swir micro =0.4, Swir meso=0.2 and swir macro 0.1) are shown in Fig.1b. It is obvious that micro represent the relative permeability with lower performance (shifted to the

right). The change in position of Kro–Krw curve depends on the pore size distribution. Therefore, every pore type has a unique relative permeability signature.

Reservoir performance to the pore size distribution (Simulation results)

To study the influence of pore size distribution on hydrocarbon reservoir performance, a simplifiedreservoir models (Black Oil model built by using Eclipse 100) are built and the calculated capillary pressure and relative permeability as the main contributors are introduced to the model. The three models (micro, meso, and macro) were run successfully for and production were reported for 30 years. The main field performance parameters were extracted and results are presented in Fig.2 to Fig.5.

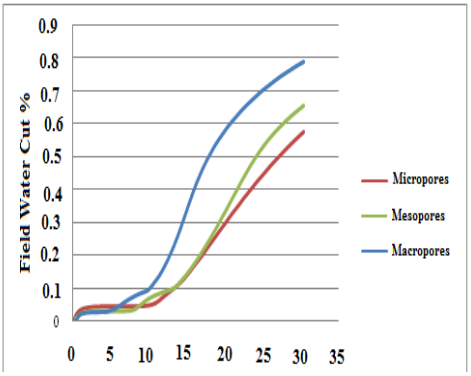


Figure 2.Trend of water cut for three pore size distribution cases.

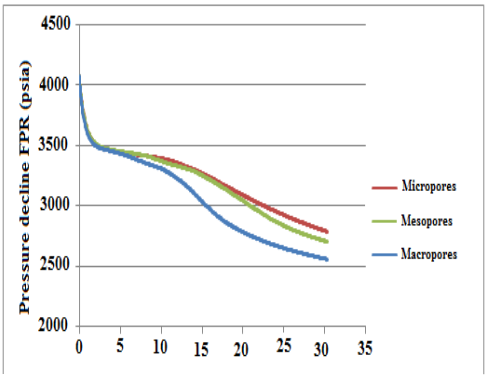


Figure 3.Pressure decline curves for three pore size distribution cases.

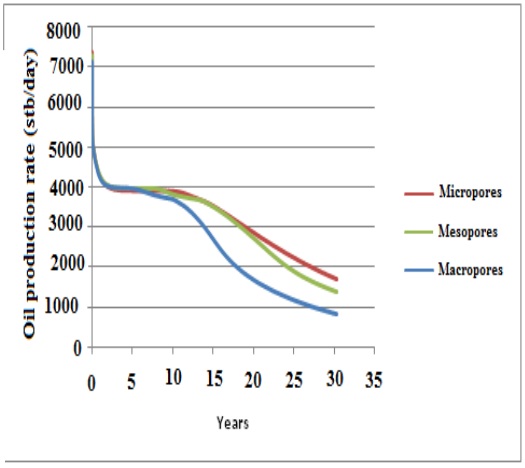


Figure 4. Oil production rate for three pore size distribution cases

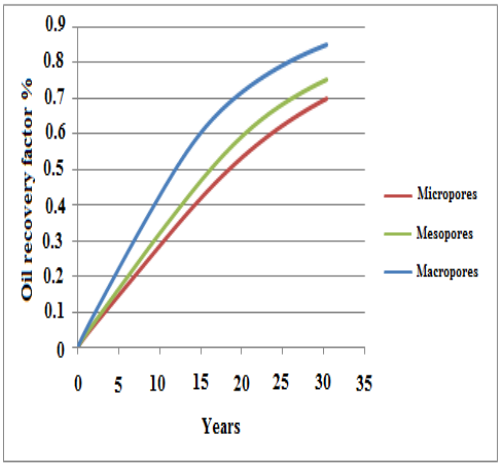


Figure 5. Oil recovery factor for three pore size distribution cases

From reservoir management consideration, water production has server impact on hydrocarbon production, hence on field development planning. Fig.2 shows the plot of field water cut for the cases owning three different rock pore size distribution. The Figure demonstrates a small increase in water cut for all three cases after which the trends remained steady till after about 7 years. Thereafter, there is a sharp increase in water cut for all cases with the macro case as the highest. The early water breakthrough event is observed for the macro pore size case, which means reservoir engineer should prepared for the intervention job at earlier stage compared to the other cases.

Fig.3 depicts the pressure decline during the production period. After the drastic drop in pressure caused by natural depletion, then the drop mainly governed by the rock pore size distribution where the pressure decline curve for microporous rock decreases further apart from the other curves followed by the average reservoir pressure is mesoporous and microporous. This result provides a guide to reservoir engineer to plan to maintain the reservoir pressure for macro case at early stage of reservoir production life.
It is also worth mentioning that any plan for water injection would be affected by the pore size distribution where the rate and injectivity of the injected water strongly affected by pore size distribution. This can be simply anticipated from the results in the Fig.3 where for macro case higher rate on injection is expected to depress the pressure decline.

The rate of oil production is also affected by the pore size distribution similar as pressure and water production and injection. Fig.4 shows the trend for the oil production rate for all cases. As it is demonstrated in this Figure, a sharp decrease in production rate in the early years is observed irrespective to the rock pore size. The production rate is maintained until the tenth year, after which it decreases sharply till 30 years. The initial decrease in oil production rate is due to the reduction in oil in the immediate production zone. The continuous flow of oil towards the producer causes the rates to sharply decrease further. There is a further decrease in the macro curve compared to othersbecause of the presence of large pores which can be attributed to the early water breakthrough presented in Fig.2.

Finally, the oil recoveries from the reservoirs having different pore size are evaluated using the following equation:

$$R_f = \frac{\overline{S_w} - S_{wir}}{1 - S_{wir}} \tag{2}$$

Where,

$\bar{S}_W$ is The average saturation behind the front calculated from the well know Buckley- Leverett equation [6], S_{wir} is Irreducible water saturation, and R_f is represents the recovery factor.

Fig.5 shows the plot of field oil efficiency for the three curves. It can be clearly seen that high oil recovery is associated with the macro pores despite the early time of breakthrough; this is because there is higher oil in place within macro (storage capacity) and hence sweep efficiency. The lower required capillary pressure and insignificant viscus and gravity forces are among other reason to ease the mobilization of oil in the macro-porous rocks than the other two studied cases.

Conclusions

This study showed that the pore size distribution has a great influence on the number of invaded pores, the capillary pressure, the relative permeability, and the field performance.
The following conclusions were drawn from this work:

- The water invasion into the reservoir rock starts from zone with the highest population of macro pores followed by mesoporous, and microporous.
- The favourable relative permeability is associated with the macroporous rocks.
- The early time of water breakthrough is witnessed within macroporous.
- The highest oil recovery is obtained for macroporous sample which is associated to the highest storage capacity of the oil, even there is early time of water breakthrough.

References

[1] Dernaika, M., Hannon, L., Serag El Din, S. and Kalam, M.Z. (2013) The Effect of Rock Properties on Remaining and Residual Oil Saturation in Heterogeneous Carbonate Rocks, *SPE 164141 presented at the Middle East Oil and Gas Show and Conference.10-13 March, Manama, Bahrain.*

[2] Elgaghah, S. (2007) A Novel Technique for the Determination of Microscopic Pore Size Distribution of Heterogeneous Reservoir Rocks, *SPE 107750 presented at the Asia Pacific Oil and Gas Conference and Exhibition,30 October-1 November, Jakarta, Indonesia.*

[3] Beiranvand, B. (2003) Quantitative characterization of carbonate pore systems by mercury-injection method and image analysis in a homogeneous reservoir, *SPE-81479 presented at the Middle East Oil show, 9-12 June, Bahrain.*

[4] Ling, K., Han, G., Shen, Z., Ghalambor, A. He, J., Pei, P. (2014) Calculating pore size distribution by using capillary pressure, *SPE-168183 presented at the International Symposium and Exhibition on Formation Damage , 26-28 February, Lafayette, Louisiana, USA.*

[5] Lønøy, A. (2006) Making sense of carbonate pore systems, *AAPG Bulletin **90**, 1381-1405.*

[6] Ahmed, T. (2010) *Reservoir Engineering Handbook*, 5th edn, Elsevier, USA.

Analytiacl Solution for Sandwiches Cantilever Beam

†Z. Wang[1], Y.J.Niu[2]

[1] College of Mathematics,Dalian University of Technology, China.
[2] College of Science, Liaoning University of Technology, China.

*Presenting author:
†Corresponding author: wangzhen@dlut.edu.cn

Abstract:

The sandwiches cantilever beam which often appeared in composite materials science. Firstly, we give the governing equations and interface condition for solid mechanics problems from the elastic theory. Then a stress function is supposed, which can be used to describe the strain and displacement function.Finally according to the boundary conditions, we can achieve the stress and displacement solution.

Keywords: Sandwiches cantilever beam; analytical solution; strain; displacement

1. Introduction

Most problems in engineering mechanics can be described in the form of differential equation, integrals and all kinds of algebraic form. While we hope to obtain the analytical solutions for the most of the practical problems, we can only achieve the numerical solution insteads. Therefore most numerical methods have been developed, such as Finite Element Methods[1-3], Finite Difference Methods, Finite Volume Methods, and recently Meshfree Methods[4]. And the analytical solution[5-7] are always used to check the accuracy and reliability of the numerical methods. In this paper, we obatain the analytical solution of the sandwiches cantilever beam which is fixed in one end. The analytical solution can be used to as a standard problem to test the the accuracy and reliability of the numerical methods.

2. Preliminary knowledge

Based on elastic mechanics[8] the discontinuous material problem defined in domain Ω bounded by Γ_e .(Fig.1) can be described by equilibrium equation. $\sigma_{ji,j}+b_i=0$, $\Omega=\Omega^+\bigcup\Omega^-$, where σ_{ij} is the component of stress tensor and b_i is the component of body force, and Ω^+ and Ω^- are two different materials. Boundary condintions are given as follows $\sigma_{ij}n_j=\overline{t}_i,(x,y)\in\Gamma_t$; $u_i=\overline{u}_i,(x,y)\in\Gamma_u$ .where t_i is the vector of tractions, u_i is the vector of displacement , n_j is the unit outward normal of Ω .By the continuity of displacements and tractions on Γ_e ,we can obtain

$$[|t_i|]=t_i^+-t_i^-=0. \tag{1}$$

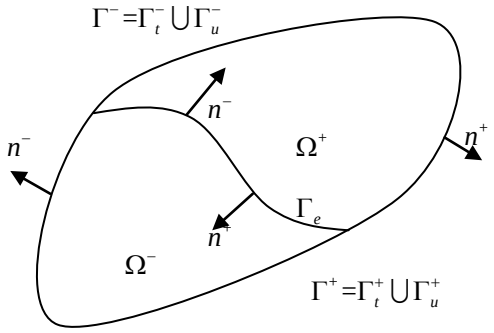
$$[|u_i|]=u_i^+-u_i^-=0. \tag{2}$$

Where $[|\cdot|]$ denotes the jump on the surface of the material,i.e. discontinuity of the Ω . By Cauchy's equation, the equation (1) can be written as

$$\sigma_{ij}^+n_j^+-\sigma_{ij}^-n_j^-=0. \tag{3}$$

Where n^+ and n^- are the unit outward normal of Ω^+ and Ω^- respectively.

Constitutive equation $\sigma_{ij}=C_{ijkl}\varepsilon_{kl}$,where C_{ijkl} is the elasticity modulus and $\varepsilon_{kl}=\frac{1}{2}(u_{k,l}+u_{l,k})$ is the strain tensor.



3.The Expression of strain, stress and displacement functions.

Consider a two-dimensional multilayer material problem. Suppose that $\varepsilon_{ij}, \sigma_{ij}$, ν_i, E_i are the strain, stress, Poisson's ration,and Young's modulus respectively, where $i, j = 1,2,3$. For each layer of the multilayer material, all have the same elasticity equation form. Firstly, the static equilibrium equation

$$\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0, \frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{xy}}{\partial x} = 0. \tag{4}$$

By constitutive equations of plane problem $\sigma = D\varepsilon$,we can obtain

$$\varepsilon_x = \frac{1}{E}(\sigma_x - \nu\sigma_y), \varepsilon_y = \frac{1}{E}(\sigma_y - \nu\sigma_x), \gamma_{xy} = 2\frac{1+\nu}{E}\tau_{xy}. \tag{5}$$

By $\varepsilon = Lu$

$$\varepsilon_x = \frac{\partial u}{\partial x}, \varepsilon_y = \frac{\partial v}{\partial y}, \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}. \tag{6}$$

From the formula (6) we can get the equation of strain compatibility

$$\frac{\partial^2 \varepsilon_x}{\partial y^2} + \frac{\partial^2 \varepsilon_y}{\partial x^2} = \frac{\partial^2 \gamma_{xy}}{\partial x \partial y} \tag{7}$$

Let U_1, U_2, U_3 be the stresses from top to bottom in Sandwiches Cantilever Beam problem, which satisfy the following equations respctively

$$\sigma^i{}_x = \frac{\partial^2 U_i}{\partial y^2}, \sigma^i{}_y = \frac{\partial^2 U_i}{\partial x^2}, \tau^i{}_{xy} = -\frac{\partial^2 U_i}{\partial x \partial y}, i = 1,2,3. \tag{8}$$

By the elastic theory[9], we can obtain

$$\sigma^i{}_y = \frac{\partial^2 U_i}{\partial x^2} = f_i(y). \tag{9}$$

Integrating both sides of (9) the stress can be expressed as

$$U_i = \frac{x^2}{2} f_i(y) + x f_{i,1}(y) + f_{i,2}(y). \tag{10}$$

Substituting (10) into (8), we have

$$\sigma_x^i = \frac{\partial^2 U_i}{\partial y^2} = \frac{x^2}{2}\frac{\partial^2 f_i(y)}{\partial y^2} + \frac{\partial^2 f_{i,1}(y)}{\partial y^2} + \frac{\partial^2 f_{i,2}(y)}{\partial y^2} \tag{11}$$

$$\tau_{xy}^i = -\frac{\partial^2 U_i}{\partial x \partial y} = -x\frac{\partial f_i(y)}{\partial y} - \frac{\partial f_{i,1}(y)}{\partial y} \tag{12}$$

Substituting (9),(11),(12) into (5) we have

$$\begin{cases} \varepsilon_x^i = \dfrac{1}{E_i}[\dfrac{x^2}{2}\dfrac{\partial^2 f_i(y)}{\partial y^2} + x\dfrac{\partial^2 f_{i,1}(y)}{\partial y^2} + \dfrac{\partial^2 f_{i,2}(y)}{\partial y^2} - \nu_i f_i(y)], \\ \varepsilon_y^i = \dfrac{1}{E_i}[f_i(y) - \nu_i(\dfrac{x^2}{2}\dfrac{\partial^2 f_i(y)}{\partial y^2} + \dfrac{\partial^2 f_{i,1}(y)}{\partial y^2} + \dfrac{\partial^2 f_{i,2}(y)}{\partial y^2})], \\ \gamma_{xy}^i = 2\dfrac{1+\nu_i}{E_i}(-x\dfrac{\partial f_i(y)}{\partial y} - \dfrac{\partial f_{i,1}(y)}{\partial y}). \end{cases} \tag{13}$$

We need to confirm the expressions of $f_i(y)$, $f_{i,1}(y)$, $f_{i,2}(y)$ to obtain U_i, by (5) and (13) we have

$$\begin{cases} \dfrac{\partial^2 \varepsilon_x{}^i}{\partial y^2} = \dfrac{1}{E_i}[\dfrac{x^2}{2}\dfrac{\partial^4 f_i(y)}{\partial y^4} + x\dfrac{\partial^4 f_{i,1}(y)}{\partial y^4} + \dfrac{\partial^4 f_{i,2}(y)}{\partial y^4} - \nu_i \dfrac{\partial^2 f_i(y)}{\partial y^2}], \\ \dfrac{\partial^2 \varepsilon_y{}^i}{\partial x^2} = \dfrac{1}{E_i}(-\nu_i \dfrac{\partial^2 f_i(y)}{\partial y^2}), \\ \dfrac{\partial^2 \gamma_{xy}{}^i}{\partial x \partial y} = -2\dfrac{1+\nu_i}{E_i}\dfrac{\partial^2 f_i(y)}{\partial y^2}. \end{cases} \tag{14}$$

Substituting (14) into (7) and simplifying ,we have

$$\frac{x^2}{2}\frac{\partial^4 f_i(y)}{\partial y^4} + x\frac{\partial^4 f_{i,1}(y)}{\partial y^4} + \frac{\partial^4 f_{i,2}(y)}{\partial y^4} + 2\frac{\partial^2 f_i(y)}{\partial y^2} = 0.$$

Since $f_i(y)$, $f_{i,1}(y)$, $f_{i,2}(y)$ can be taken arbitrarily, we have

$$\frac{\partial^4 f_i(y)}{\partial y^4} = 0, \frac{\partial^4 f_{i,1}(y)}{\partial y^4} = 0, \frac{\partial^4 f_{i,2}(y)}{\partial y^4} + 2\frac{\partial^2 f_i(y)}{\partial y^2} = 0. \tag{15}$$

Integrating the first two equations of (15) we can get the expressions of $f_i(y)$, $f_{i,1}(y)$

$$f_i(y) = A_i y^3 + B_i y^2 + C_i y + D_i,\ f_{i,1}(y) = F_i y^3 + G_i y^2 + H_i y.$$

By the third equation of (15) and (16) we have

$$f_{i,2}(y) = -\frac{A_i}{10} y^5 - \frac{B_i}{6} y^4 + M_i y^3 + N_i y^2.$$

Substituting (16),(17) into (10) the stress can be expressed as

$$\begin{aligned} U_i = {} & \frac{x^2}{2}(A_i y^3 + B_i y^2 + C_i y + D_i) + x(F_i y^3 + G_i y^2 + H_i y) - \\ & \frac{A_i}{10} y^5 - \frac{B_i}{6} y^4 + M_i y^3 + N_i y^2, \end{aligned} \tag{18}$$

Where A_i, B_i, C_i, D_i, F_i, G_i, H_i, M_i, $N_i, i = 1,2,3$ are undetermined coefficients.

Thus $\sigma_x^i, \sigma_y^i, \tau_{xy}^i$ can be denoted as

$$\begin{cases} \sigma_x{}^i = \dfrac{\partial^2 U_i}{\partial y^2} = \dfrac{x^2}{2}(6A_i y + 2B_i) + x(6F_i + 2G_i) - 2A_i y^3 - 2B_i y^2 + 6M_i y + 2N_i, \\ \sigma_y{}^i = \dfrac{\partial^2 U_i}{\partial x^2} = A_i y^3 + B_i y^2 + C_i y + D_i, \\ \tau_{xy}{}^i = -\dfrac{\partial^2 U_i}{\partial x \partial y} = -x(3A_i y^2 + 2B_i y + C_i) - 3F_i y^2 - 2G_i y - H_i. \end{cases} \tag{19}$$

By (6) and (19), the strain can be expressed as

$$\varepsilon_x^{\,i} = \frac{x^2(3A_iy + B_i) + x(6F_i + 2G_i) - 2A_iy^3 - 2B_iy^2 + 6M_iy + 2N_i}{E_i} - \nu_i \frac{A_iy^3 + B_iy^2 + C_iy + D}{E_i}, \tag{20}$$

$$\varepsilon_y^{\,i} = \frac{A_iy^3 + B_iy^2 + C_iy + D}{E_i} - \nu_i \frac{x^2(3A_iy + B_i) + x(6F_i + 2G_i) - 2A_iy^3 - 2B_iy^2 + 6M_iy + 2N_i}{E_i}, \tag{21}$$

$$\gamma_{xy}^{\,i} = -\frac{2(1+\nu_i)[-x(3A_iy^2 + 2B_iy + C_i) - 3F_iy^2 - 2G_iy - H_i]}{E_i}. \tag{22}$$

Integrating both sides of the following formulae $\varepsilon_x = \frac{\partial u}{\partial x}, \varepsilon_y = \frac{\partial v}{\partial y}$, we have

$$u_i = \frac{-(D_i\nu_i - 2N_i - 6M_iy + 2A_iy^3 + 2B_iy^2 + C_i\nu_iy + A_i\nu_iy^3 + B_i\nu_iy^2)}{E_i}x + \frac{G_i + 3F_iy}{E_i}x^2 + \frac{B_i + 3A_iy}{3E_i}x^3 + h_i(y), \tag{23}$$

$$v_i = (\frac{y(D_i - 2N_i\nu_i)}{E_i} + \frac{y^4(A_i + 2A_i\nu_i)}{4E_i} + \frac{y^3(B_i + 2B_i\nu_i)}{3E_i} + \frac{y^2(C_i - 6M_i\nu_i)}{2E_i})x - \frac{\nu_iy(2G_i + 3F_iy)}{E_i}x^2 - \frac{\nu_iy(2B_i + 3A_iy)}{2E_i}x^3 + g_i(x). \tag{24}$$

By the formulae $\gamma_{xy} = 2\frac{1+\nu}{E}\tau_{xy}$ and $\gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}$, we have

$$6F_iy + 4G_iy + 3\nu_iF_iy^2 + 2\nu_iG_iy + \frac{dh_i(y)}{dy}E_i + 2xG_i + x^3A_i + 3x^2F_i + 6M_ix + 2H_i + \nu_iC_ix + 2\nu_iH_i + \frac{dg_i(x)}{dx}E_i = 0 \tag{25}$$

where $i = 1, 2, 3.$

Since $h_i(y)$ and $g_i(x)$ can be taken arbitrarily, then suppose that

$$6F_iy + 4G_iy + 3\nu_iF_iy^2 + 2\nu_iG_iy + \frac{dh_i(y)}{dy}E_i = 0$$
$$2xG_i + x^3A_i + 3x^2F_i + 6M_ix + 2H_i + \nu_iC_ix + 2\nu_iH_i + \frac{dg_i(x)}{dx}E_i = 0 \tag{26}$$

Then we obtain

$$h_i(y) = -\frac{(2+\nu_i)(F_iy^3 + G_iy^2)}{E_i} + I_iy + J_i,$$
$$g_i(x) = -\frac{A_ix^4 + 4F_ix^3 + 4G_ix^2 + 12M_ix^2 + 2\nu_iC_ix^2 + 8(1+\nu_i)H_ix}{4E_i} - I_ix + L_i. \tag{27}$$

Substituting $h_i(y)$, $g_i(x)$ into (23),(24), the expression of displacement

$$u_i = \frac{-(D_i\nu_i - 2N_i - 6M_i y + 2A_i y^3 + 2B_i y^2 + C_i\nu_i y + A_i\nu_i y^3 + B_i\nu_i y^2)}{E_i} x + \frac{G_i + 3F_i y}{E_i} x^2 + \frac{B_i + 3A_i y}{3E_i} x^3 - \frac{(2+\nu_i)(F_i y^3 + G_i y^2)}{E_i} + I_i y + J, \tag{28}$$

$$\nu_i = [\frac{y(D_i - 2N_i\nu_i)}{E_i} + \frac{y^4(A_i + 2A_i\nu_i)}{4E_i} + \frac{y^3(B_i + 2B_i\nu_i)}{3E_i} + \frac{y^2(C_i - 6M_i\nu_i)}{2E_i} - \frac{2(1+\nu_i)H_i}{E_i} - I_i]x - [\frac{2\nu_i y(2G_i + 3F_i y) + 2G_i + 6M_i + \nu_i C_i}{2E_i}]x^2 - \frac{\nu_i y(2B_i + 3A_i y) + 2F_i}{2E_i} x^3 - \frac{A_i}{4E_i} x^4 + L_i. \tag{29}$$

4.Determine the coefficients by the boundary conditions

The following is to determine the coefficients $A_i, B_i, C_i, D_i, F_i, G_i, H_i, I_i, J_i, L_i, M_i, N_i$, where $i = 1, 2, 3$. The natural boundary conditions:

On the top layer, i.e. $y = h_1$,where $\sigma_y^1 = 0, \tau_{xy}^1 = 0$. By the second equation of (19) we have

$$A_1 h_1^3 + B_1 h_1^2 + C_1 h_1 + D_1 = 0. \tag{30}$$

Combine the third equation of (19) we obtain

$$\tau_{xy}^1 = -x(3A_1 h_1^2 + 2B_1 h_1 + C_1) - 3F_1 h_1 - 2G_1 h_1 - H_1 = 0,$$

since the value of x is not unique,then

$$3A_1 h_1^2 + 2B_1 h_1 + C_1 = 0,\ 3F_1 h_1 + 2G_1 h_1 + H_1 = 0 \tag{31}$$

On the bottom layer, i.e. $y = -h_1$, where $\sigma_y^3 = 0, \tau_{xy}^3 = 0$. Similarly

$$-A_3 h_1^3 + B_3 h_1^2 - C_3 h_1 + D_3 = 0, \tag{32}$$

$$3A_3(-h_1)^2 + 2B_3(-h_1) + C_3 = 0, \tag{33}$$

$$3F_3(-h_1) + 2G_3(-h_1) + H_3 = 0. \tag{34}$$

When $x = l$, the integral form of the natural boundary condition should satisfy the following

$$\int_{-h_1}^{-h_2} \sigma_x^3 dy + \int_{-h_2}^{h_2} \sigma_x^2 dy + \int_{h_2}^{h_1} \sigma_x^1 dy = 0, \tag{35}$$

$$\int_{-h_1}^{-h_2} \sigma_x^3 y dy + \int_{-h_2}^{h_2} \sigma_x^2 y dy + \int_{h_2}^{h_1} \sigma_x^1 y dy = 0, \tag{36}$$

$$\int_{-h_1}^{-h_2} \tau_{xy}^3 dy + \int_{-h_2}^{h_2} \tau_{xy}^2 dy + \int_{h_2}^{h_1} \tau_{xy}^1 dy = -q. \tag{37}$$

Substituting $\sigma_x^i, \tau_{xy}^i (i = 1, 2, 3)$ in (19) into (35)~(37), deforming obtain the equations of A_i, $B_i, F_i, G_i, M_i, N_i (i = 1, 2, 3)$.

$$\begin{aligned}
&\frac{l}{2}(h_2^2 - h_1^2)(3l^2 - (h_2^2 + h_1^2))A_3 + (-h_2 + h_1)(l^2 - \frac{2}{3}(h_1^2 + h_1 h_2 + h_2^2))B_3 + 3l(h_2^2 - h_1^2)F_3 + \\
&2l(-h_2 + h_1)G_3 + 3M_3(h_2^2 - h_1^2) + 2N_3(-h_2 + h_1) + h_2(2l^2 - \frac{4}{3}h_2^2)B_2 + 4h_2 l G_2 + 4h_2 N_2 + \\
&\frac{l}{2}(h_1^2 - h_2^2)(3l^2 - (h_1^2 + h_2^2))A_1 + (h_1 - h_2)(l^2 - \frac{2}{3}(h_1^2 + h_1 h_2 + h_2^2))B_1 + \\
&3l(h_1^2 - h_2^2)F_1 + 2l(h_1 - h_2)G_1 + 3M_1(h_1^2 - h_2^2) + 2N_1(h_1 - h_2) = 0
\end{aligned} \tag{38}$$

$$[l^2(h_1^3-h_2^3)-\frac{2}{5}(h_1^5-h_2^5)]A_3+\frac{1}{2}(h_2^2-h_1^2)(l^2-(h_1^2+h_2^2))B_3+2l(h_1^3-h_2^3)F_3+$$
$$l(h_2^2-h_1^2)G_3+2(-h_2^3+h_1^3)M_3+(h_2^2-h_1^2)N_3+2h_2^3(l^2-\frac{2}{5}h_2^2)A_2+4lh_2^3F_2+$$
$$4h_2^3M_2+[l^2(h_1^3-h_2^3)-\frac{2}{5}(h_1^5-h_2^5)]A_1+\frac{1}{2}(h_1^2-h_2^2)(l^2-(h_1^2+h_2^2))B_1+ \tag{39}$$
$$2l(h_1^3-h_2^3)F_1+l(h_1^2-h_2^2)G_1+2(h_1^3-h_2^3)M_1+(h_1^2-h_2^2)N_1=0.$$

$$l(h_1^3-h_2^3)A_3+l(h_2^2-h_1^2)B_3+l(h_1-h_2)C_3+(h_1^3-h_2^3)F_3+(h_2^2-h_1^2)G_3+$$
$$(h_1-h_2)H_3+2lh_2^3A_2+2lh_2C_2+2h_2^3F_2+2h_2H_2+l(h_1^3-h_2^3)A_1+ \tag{40}$$
$$l(h_1^2-h_2^2)B_1+l(h_1-h_2)C_1+(h_1^3-h_2^3)F_1+(h_1^2-h_2^2)G_1+(h_1-h_2)H_1=-q.$$

The essential boundary conditions :

When $(x,y)=(0,0)$,we have $u_2(0,0)=0, v_2(0,0)=0, \frac{\partial v_2}{\partial x}(0,0)=0.$ By the displacement expressions (28),(29) obatin

$$J_2=0,\ L_2=0,\ -\frac{2(1+\nu_2)H_2}{E_2}+I_2=0. \tag{41}$$

By the continuity of displacement and stress and (2),(3) obtain

$$u_1=u_2,\ v_1=v_2,\ \sigma_y^1=\sigma_y^2,\ \tau_{xy}^1=\tau_{xy}^2,\ \text{当 } y=-h_2 \text{ 时},$$
$$u_3=u_2,\ v_3=v_2,\ \sigma_y^3=\sigma_y^2,\ \tau_{xy}^3=\tau_{xy}^2,\ \text{当 } y=h_2 \text{ 时},$$

Since $u_1=u_2$,where $y=-h_2$ then

$$\{\frac{-\left(D_1\nu_1-2N_1+6M_1h_2-2A_1h_2^3+2B_1h_2^2-C_1\nu_1h_2-A_1\nu_1h_2^3+B_1\nu_1h_2^3\right)}{E_1}+$$
$$\frac{\left(D_2\nu_2-2N_2+6M_2h_2-2A_2h_2^3+2B_2h_2^2-C_2\nu_2h_2-A_2\nu_2h_2^3+B_2\nu_2h_2^3\right)}{E_2}\}x+$$
$$(\frac{G_1-3F_1h_2}{E_1}-\frac{G_2-3F_2h_2}{E_2})x^2+(\frac{B_1-3A_1h_2}{3E_1}-\frac{B_2-3A_2h_2}{3E_2})x^3-$$
$$\frac{(2+\nu_1)(-F_1h_2^3+G_1h_2^2)}{E_1}-I_1h_2+J_1+\frac{(2+\nu_2)(-F_2h_2^3+G_2h_2^2)}{E_1}+I_2h_2-J_2=0.$$

And the value of x is not unique, we have

$$\frac{-\left(D_1\nu_1-2N_1+6M_1h_2-2A_1h_2^3+2B_1h_2^2-C_1\nu_1h_2-A_1\nu_1h_2^3+B_1\nu_1h_2^2\right)}{E_1}+$$
$$\frac{\left(D_2\nu_2-2N_2+6M_2h_2-2A_2h_2^3+2B_2h_2^2-C_2\nu_2h_2-A_2\nu_2h_2^3+B_2\nu_2h_2^2\right)}{E_2}=0, \tag{42}$$

$$\frac{G_1-3F_1h_2}{E_1}-\frac{G_2-3F_2h_2}{E_2}=0,\ \frac{B_1-3A_1h_2}{3E_1}-\frac{B_2-3A_2h_2}{3E_2}=0, \tag{43}$$

$$-\frac{(2+\nu_1)(-F_1h_2^3+G_1h_2^2)}{E_1}-I_1h_2+J_1+\frac{(2+\nu_2)(-F_2h_2^3+G_2h_2^2)}{E_1}+I_2h_2-J_2=0. \tag{44}$$

Similarly since $\nu_1=\nu_2$,when $y=-h_2$ then we have

$$-\frac{h_2(D_1-2N_1\nu_1)}{E_1}+\frac{h_2^4(A_1+A_1\nu_1)}{4E_1}-\frac{h_2^3(B_1+2B_1\nu_1)}{3E_1}+\frac{h_2^2(C_1-6M_1\nu_1)}{2E_1}-\frac{2(1+\nu_1)H_1}{E_1}-I_1+$$
$$\frac{h_2(D_2-2N_2\nu_2)}{E_2}-\frac{h_2^4(A_2+A_2\nu_2)}{4E_2}+\frac{h_2^3(B_2+2B_2\nu_2)}{3E_2}-\frac{h_2^2(C_2-6M_2\nu_2)}{2E_2}+\frac{2(1+\nu_2)H_2}{E_2}+I_2=0, \quad (45)$$

$$-\frac{-2\nu_1h_2(2G_1-3F_1h_2)+2G_1+6M_1+\nu_1C_1}{2E_1}+\frac{-2\nu_2h_2(2G_2-3F_2h_2)+2G_2+6M_2+\nu_2C_2}{2E_2}=0, \quad (46)$$

$$-\frac{-\nu_1h_2(2B_1+3A_1y)+2F_1}{2E_1}+\frac{-\nu_2h_2(2B_2+3A_2y)+2F_2}{2E_2}=0, \quad (47)$$

$$-\frac{A_1}{4E_1}+\frac{A_2}{4E_2}=0, \quad (48)$$

$$L_1-L_2=0. \quad (49)$$

From $\sigma_y^1=\sigma_y^2$ obtain

$$-(A_1-A_2)h_2^3+(B_1-B_2)h_2^2-(C_1-C_2)h_2+D_1-D_2=0. \quad (50)$$

From $\tau_{xy}^1=\tau_{xy}^2$ obtain

$$-x(3A_1h_2^2-2B_1h_2+C_1-3A_2h_2^2+2B_2h_2-C_2)-3F_1h_2^2-2G_1h_2-$$
$$H_1+3F_2h_2^2+2G_2h_2+H_2=0.$$

And the value of x is not unique, we have

$$3A_1h_2^2-2B_1h_2+C_1-3A_2h_2^2+2B_2h_2-C_2=0, \quad (51)$$

$$-3F_1h_2^2-2G_1h_2-H_1+3F_2h_2^2+2G_2h_2+H_2=0. \quad (52)$$

Similarly when $y=h_2$,since $u_3=u_2$

$$-\frac{D_3\nu_3-2N_3-6M_3h_2+2A_3h_2^3+2B_3h_2^2+C_3\nu_3h_2+A_3\nu_3h_2^3+B_3\nu_3h_2^2}{E_3}+$$
$$\frac{D_2\nu_2-2N_2+6M_2h_2+2A_2h_2^3+2B_2h_2^2+C_2\nu_2h_2+A_2\nu_2h_2^3+B_2\nu_2h_2^2}{E_2}=0, \quad (53)$$

$$\frac{G_3+3F_3h_2}{E_3}-\frac{G_2+3F_2h_2}{E_2}=0, \quad (54)$$

$$\frac{B_3+3A_3h_2}{3E_3}-\frac{B_2+3A_2h_2}{3E_2}=0. \quad (55)$$

$$-\frac{(2+\nu_2)(F_2h_2^3+G_2h_2^2)}{E_2}+I_2h_2+J_2+\frac{(2+\nu_3)(F_3h_2^3+G_3h_2^2)}{E_3}-I_3h_2-J_3=0. \quad (56)$$

since $\nu_3=\nu_2$ then

$$\frac{h_2(D_3-2N_3\nu_3)}{E_3}+\frac{h_2^4(A_3+2A_3\nu_3)}{4E_3}+\frac{h_2^3(B_3+2B_3\nu_1)}{3E_3}+\frac{h_2^2(C_3-6M_3\nu_3)}{2E_3}-\frac{2(1+\nu_3)H_3}{E_3}-I_3-$$
$$\frac{h_2(D_2-2N_2\nu_2)}{E_2}-\frac{h_2^4(A_2+A_2\nu_2)}{4E_2}-\frac{h_2^3(B_2-2B_2\nu_2)}{3E_2}-\frac{h_2^2(C_2-6M_2\nu_2)}{2E_2}+\frac{2(1+\nu_2)H_2}{E_2}+I_2=0, \quad (57)$$

$$-\frac{2\nu_3h_2(2G_3+3F_3h_2)+2G_3+6M_3+\nu_3C_3}{2E_3}+\frac{2\nu_2h_2(2G_2+3F_2h_2)+2G_2+6M_2+\nu_2C_2}{2E_2}=0, \quad (58)$$

$$-\frac{\nu_3 h_2(2B_3+3A_3 y)+2F_3}{2E_3}+\frac{\nu_2 h_2(2B_2+3A_2 h_2)+2F_2}{2E_2}=0, \tag{59}$$

$$-\frac{A_3}{4E_3}h_2^4+\frac{A_2}{4E_2}h_2^4=0, \tag{60}$$

$$L_3-L_2=0. \tag{61}$$

By $\sigma_y^1=\sigma_y^3$ and $\tau_{xy}^1=\tau_{xy}^3$ obtain

$$(A_3-A_2)h_2^3+(B_3-B_2)h_2^2+(C_3-C_2)h_2+D_3-D_2=0, \tag{62}$$

$$3A_3h_2^2+2B_3h_2+C_3-3A_2h_2^2-2B_2h_2-C_2=0, \tag{63}$$

$$3F_3h_2^2-2G_3h_2+H_3-F_2h_2^2-2G_2h_2-H_2=0. \tag{64}$$

From the above we can solve equations (30)~(34),(38)~(64) for the coefficients $A_i, B_i, C_i, D_i, F_i, G_i, H_i, I_i, J_i, L_i, M_i, N_i,\ i=1,2,3.$,then substitute them into the analytical solutions of displacement and stress as following

$$u_1(x,y)=\frac{1}{4E_2(E_2h_2^3-E_3h_2^3+h_1^3E_3)}\cdot q[6h_2E_3(h_2^2-h_1^2)(1+\nu_2)+$$
$$E_2(6(1+\nu_3)(h_2-y)h_1^2+3h_2^2y(\nu_3-\nu_2)-(4\nu_3+2\nu_2+6)h_2^3+$$
$$y^3(2+\nu_3)+3xy(2l-x))],$$

$$v_1(x,y)=\frac{-q[3(l-x)(\nu_3 y+h_2^2(\nu_2-\nu_3))+x^2(3l-x)]}{4(E_2h_2^3-E_3h_2^3+h_1^3E_3)},$$

$$u_2(x,y)=\frac{yq[(1+\nu_2)(6E_3(h_2^2-h_1^2)+E_2(2y^2-6h_1^2))+3xE_2(2l-x)]}{4E_2(E_2h_2^3-E_3h_2^3+h_1^3E_3)},$$

$$v_2(x,y)=-\frac{q(3y^2\nu_2 l+3x^2 l-x^3-3y^2\nu_2 x)}{4(E_2h_2^3-E_3h_2^3+h_1^3E_3)},$$

$$u_3(x,y)=\frac{1}{4E_2(E_2h_2^3-E_3h_2^3+h_1^3E_3)}\cdot q[E_2(h_2+y)(\nu_3(y(y-h_2)+4h_2^2-6h_1^2)-6h_1^2)+$$
$$E_2(6xyl-3x^2y+2y^3)-6E_3h_2(1+\nu_2)(h_2^2-h_1^2)-E_2h_2^2(\nu_2(3y-2h_2)-6h_2)],$$

$$v_3(x,y)=-\frac{q[3(l-x)(\nu_3 y+h_2^2(\nu_2-\nu_3))+x^2(3l-x)]}{4(E_2h_2^3-E_3h_2^3+h_1^3E_3)},$$

In the top layer $[h_1,h_2]\times[0,L]$,the analytical solution of stress

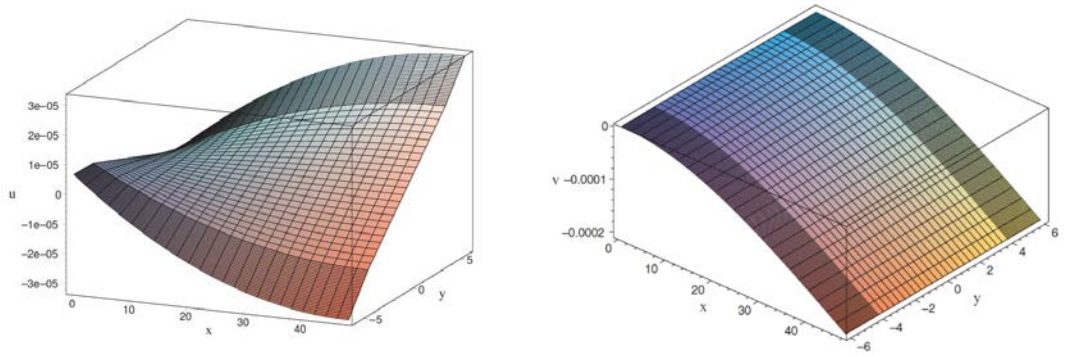
$$\sigma_x^1=\frac{3}{2}\frac{(l-x)yE_3q}{E_2h_2^3-E_3h_2^3+h_1^3E_3},\ \sigma_y^1=0,\ \tau_{xy}^1=\frac{3}{4}\frac{(y^2-h_1^2)E_3q}{E_2h_2^3-E_3h_2^3+h_1^3E_3}.$$

In the top layer $[h_2,-h_2]\times[0,L]$,the analytical solution of stress

$$\sigma_x^2=\frac{3}{2}\frac{(l-x)yE_2q}{E_2h_2^3-E_3h_2^3+h_1^3E_3},\ \sigma_y^2=0,\ \tau_{xy}^2=\frac{3}{4}\frac{(-h_2^2E_2+y^2E_2+h_2^2E_3-E_3h_1^2)q}{E_2h_2^3-E_3h_2^3+h_1^3E_3}.$$

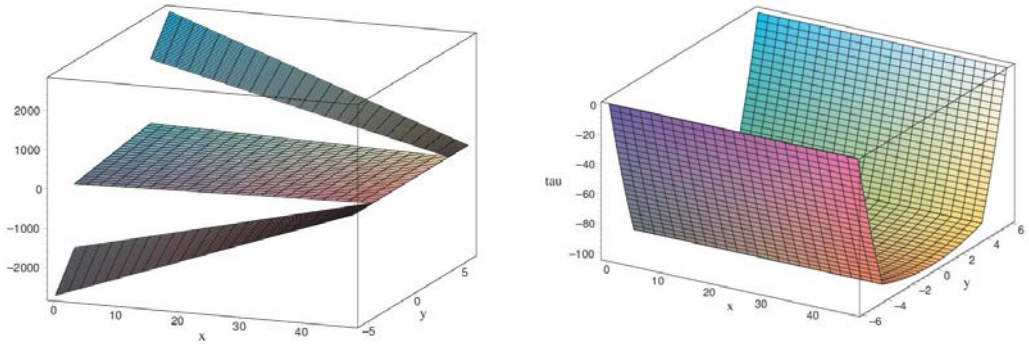
In the bottom layer $[-h_2,-h_1]\times[0,L]$,the analytical solution of stress

$$\sigma_x^3=\frac{3}{2}\frac{(l-x)yE_3q}{E_2h_2^3-E_3h_2^3+h_1^3E_3},\ \sigma_y^3=0,\ \tau_{xy}^3=\frac{3}{4}\frac{(y^2-h_1^2)E_3q}{E_2h_2^3-E_3h_2^3+h_1^3E_3}.$$



(a) (b)

Fig.2 Displacement distribution for sandwich beam (a) in x-direction (b) y-irection.



(a) (b)

Fig.3 Distribution of stress for sandwich beam (a) normal stress σ_x ans (b) shear stress τ_{xy} .

5.Conclusions

This paper focus on the analytical solution of the sandwiches cantilever beam. Firstly, we obtain the governing equations and the parametric expressions of stress and displacement by the elastic theory. Then by the interface condition and the boundary conditions we can achieve the analytical expression of stress and displacement.

References

[1] D.B.P.Huynh, T.Belytschko, The extended finite element method for fracture un composite materials, Int.J.Numer.Meth.Eng., 1999, 77(2):214-239.
[2] Zienkiewicz,O.C. Taylor, The finite element method, Butterworth Heinemann, Oxford, UK, 2000.
[3] BSusanne C. Brenner, L. Ridgway Scott, The mathematical theory of finite element methods,Springer, New York, 1994.
[4] G.R.Liu, Meshfree method:Moving Beyond the Finite Element Method, CRC Press,Boca Raton, USA, 2000.
[5] G.R. Liu, Z.Wang, G.Y.Zhang, Z.Zong, S.Wang, An edge-based smoothed point interpolation method for material discontinuity, Mech Adv Mater Struc, 2012, 19:3-17.
[6] R.J.Mackinnon, G.F.Garey, Treatment of material discontinuity in finite element computations, Int. J. Numer. Meth, 1987, 24:393-417.
[7] R.C.Batra, M. Porfiri , D. Spinello, Treatment of material discontinuity in two meshless local Petrov-Galerkin (MLPG) formulations of axisymmetric transient heat conduction, Int. J. Numer. Meth. Engng, 2004, 61:2461–2479.
[8] S.P.Timodhrnko, J.N.Goodier, Theory of Elasticity, New York:McGraw-Hill, 1951.
[9] J.L.Wu, Elastic Mechanics, Higher Education Press, Beijing, 2001.

An efficient class of fourth-order Jarratt-type methods for nonlinear equations

Ioannis K. Argyros[1], †Munish Kansal[2], V. Kanwar[2], and Raj Bala[3]

[1]Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA.
[2]University Institute of Engineering and Technology, Panjab University, Chandigarh-160 014, India.
[3]Department of Mathematics, Government College, Barwala, Panchkula-134-118, India.

†Corresponding author: mkmaths@gmail.com

Abstract

In this paper, we present a new two-point fourth-order Jarratt-type scheme based on Hansen-Patrick's family for solving nonlinear equations numerically. In terms of computational cost, each method of the proposed class requires only three functional evaluations (viz. one evaluation of function and two first-order derivatives) per full step to achieve optimal fourth-order convergence. Moreover, the local convergence analysis of the proposed methods is also given using hypotheses only on the first derivative and Lipschitz constants. Furthermore, the proposed scheme can also determine the complex zeros without having to start from a complex initial guess as would be necessary with other methods. Numerical examples and comparisons with some existing methods are included to confirm the theoretical results and high computational efficiency.

Keywords: Nonlinear equations, Jarratt-type methods, Kung-Traub conjecture, Local Convergence

Introduction

The construction of fixed point iterative methods for approximating simple zeros of a real valued function is an important task in theory and practice, not only in applied mathematics, but also for many applied scientific branches. In this paper, we consider iterative methods for solving a nonlinear equation of the form

$$f(x) = 0, \tag{1}$$

where $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is a scalar function defined on an open interval D. Analytical methods for solving such equations are almost non-existent and therefore, it is only possible to obtain approximate solutions by relying on numerical methods based on iterative procedure. One of the most famous and basic tool for solving such equations is the Newton's method [18] given by $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},\ n \geq 0$. It converges quadratically for simple roots and linearly for multiple roots. In order to improve its local order of convergence, many higher-order methods have been proposed and analyzed in [1, 4, 19]. One such well-known scheme is the classical cubically convergent Hansen-Patrick's family [6] defined by

$$x_{n+1} = x_n - \left[\frac{\beta + 1}{\beta \pm \{1 - (\beta + 1)L_f(x_n)\}^{1/2}}\right] \frac{f(x_n)}{f'(x_n)}, \tag{2}$$

where $L_f(x_n) = \frac{f''(x_n)f(x_n)}{f'^2(x_n)}$ and $\beta \in \mathbb{R}\backslash\{-1\}$. This family includes Ostrowski's square-root method for $(\beta = 0)$, Euler's method for $(\beta = 1)$, Laguerre's method for $\left(\beta = \frac{1}{\nu-1}, \nu \neq 1\right)$ and as a limiting case, Newton's method. Despite the cubic convergence, this scheme is considered less practical from a computational point of view because of the expensive second-order derivative evaluation. This fact motivated many researchers to investigate the idea of developing multipoint iterative methods for solving nonlinear equations.

Multipoint iterative methods [13, 14] for solving nonlinear equations are of great practical importance since they circumvent the limitations of one-point methods regarding the convergence order and computational efficiency. The main objective in the construction of the new iterative methods is to obtain the maximal computational efficiency. In other words, the aim is to attain convergence order as high as possible with fixed number of functional evaluation per iteration. According to the Kung-Traub conjecture [18], the order of convergence of any multipoint method without memory requiring d functional evaluations per iteration, cannot exceed the bound 2^{d-1}, called the optimal order. Consequently, convergence order of an optimal iterative method without memory consuming three functional evaluations cannot exceed four. Also, efficiency of an iterative method is measured by the efficiency index [18] defined as $E = p^{\frac{1}{d}}$, where p is the order of convergence. King's family [10], Ostrowski's method [18] and Jarratt's method [7, 12] are the well-known fourth-order multipoint methods without memory. The Jarratt method is widely considered and applied for its computational efficiency.

The fourth-order Jarratt's method which uses one evaluation of the function and two evaluations of the first derivatives is defined by

$$x_{n+1} = x_n - J_f(x_n)\frac{f(x_n)}{f'(x_n)}, \tag{3}$$

where $J_f(x_n) = \frac{3f'(y_n)+f'(x_n)}{6f'(y_n)-2f'(x_n)}$ and $y_n = x_n - \frac{2}{3}\frac{f(x_n)}{f'(x_n)}$. It satisfies the following error equation

$$e_{n+1} = \left(c_2^3 - c_2c_3 + \frac{c_4}{9}\right)e_n^4 + O(e_n^5).$$

Recently, Soleymani et al. [16] proposed two-point fourth-order Jarratt-type methods for obtaining simple roots of nonlinear equations, which are defined as follow:

$$\begin{cases} y_n = x_n - \dfrac{2}{3}\dfrac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \dfrac{2f(x_n)}{f'(x_n)+f'(y_n)}\left[\left(1+\left(\dfrac{f(x_n)}{f'(x_n)}\right)^3\right)\left(2 - \dfrac{7}{4}\dfrac{f'(y_n)}{f'(x_n)} + \dfrac{3}{4}\left(\dfrac{f'(y_n)}{f'(x_n)}\right)^2\right)\right], \end{cases} \tag{4}$$

and its error equation is given by

$$e_{n+1} = \frac{1}{9}\left(-9 + 33c_2^3 - 9c_2c_3 + c_4\right)e_n^4 + O(e_n^5),$$

and

$$\begin{cases} y_n = x_n - \dfrac{2}{3}\dfrac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \dfrac{f(x_n)}{2}\left[\dfrac{1}{f'(x_n)} + \dfrac{1}{f'(y_n)}\right]\left[\left(1+\left(\dfrac{f(x_n)}{f'(x_n)}\right)^4\right)\left(1 - \dfrac{1}{4}\left(\dfrac{f'(y_n)}{f'(x_n)} - 1\right)\right.\right. \\ \qquad \left.\left. + \dfrac{1}{2}\left(\dfrac{f'(y_n)}{f'(x_n)} - 1\right)^2\right)\right], \end{cases} \tag{5}$$

where it satifies the following error equation

$$e_{n+1} = \left(\frac{79}{27}c_2^3 - c_2c_3 + \frac{c_4}{9}\right)e_n^4 + O(e_n^5).$$

In [9], Khattri and Abbasbandy proposed an optimal fourth-order variant of Jarratt's method using one function evaluation and two first-order derivatives, which is defined as follows:

$$\begin{cases} y_n = x_n - \dfrac{2}{3}\dfrac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \left[1 + \dfrac{21}{8}\dfrac{f'(y_n)}{f'(x_n)} - \dfrac{9}{2}\left(\dfrac{f'(y_n)}{f'(x_n)}\right)^2 + \dfrac{15}{8}\left(\dfrac{f'(y_n)}{f'(x_n)}\right)^3\right]\dfrac{f(x_n)}{f'(x_n)}. \end{cases} \tag{6}$$

It satisfies the following error equation

$$e_{n+1} = \frac{1}{9}\left(85c_2^3 - 9c_2c_3 + c_4\right)e_n^4 + O(e_n^5).$$

In this work, we are interested in designing a new two-point fourth-order class of iterative methods from a view point of Hansen-Patrick type methods, which does not require any second-order derivative evaluation for obtaining simple roots of nonlinear equations. Each method requires only one evaluation of the given function and two evaluations of the first-order derivative per iteration. It is also observed that the body structures of our proposed methods are simpler than the existing two-point fourth-order methods mentioned above. We also present the local convergence analysis of the proposed methods using hypotheses only on the first-order derivative and Lipschitz constants. Moreover, it is shown by way of illustration that the proposed schemes can determine the complex zeros without having to start from a complex number as would be necessary with other methods. It can be easily seen that the proposed schemes are highly efficient in multi-precision computing environment.

Uni-parametric family of Jarratt-type methods

In this section, we intend to develop a new optimal class of fourth-order Hansen-Patrick type methods, not requiring the computation of second-order derivative. For this purpose, let

$$y_n = x_n - \alpha u(x_n), \tag{7}$$

where $u(x_n) = \frac{f(x_n)}{f'(x_n)}$ and α is non-zero real parameter. Now, expanding $f'(y_n) = f'(x_n - \alpha u(x_n))$ about a point $x = x_n$ by Taylor series expansion, we have $f'(y_n) \approx f'(x_n) - \alpha u(x_n)f''(x_n)$, which further yields

$$f''(x_n) \approx \frac{f'(x_n) - f'(y_n)}{\alpha u(x_n)}. \tag{8}$$

Using this approximate value of $f''(x_n)$ in formula (2), and using weight function technique in the second step, we obtain a modified family of methods free from second-order derivative as follows:

$$\begin{cases} y_n = x_n - \alpha\dfrac{f(x_n)}{f'(x_n)}, \quad \alpha \in \mathbb{R}\backslash\{0\}, \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}\left[\dfrac{\beta+1}{\beta+\left\{1-(\beta+1)L_f^*(x_n)\right\}^{\frac{1}{2}}}\right]H(\tau), \end{cases} \tag{9}$$

where $\beta \in \mathbb{R}$, $L_f^*(x_n) = \frac{f'(x_n)-f'(y_n)}{\alpha f'(x_n)}$ and $H : \mathbb{R} \to \mathbb{R}$ is a real variable weight function with $\tau = \frac{f'(y_n)}{f'(x_n)} = 1 + O(e_n)$. Theorem 1 illustrates that under what conditions on weight function, convergence order of the family (9) will arrive at the optimal level four.

Convergence analysis

Theorem 1 *Assume that function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is sufficiently differentiable and has a simple zero $x^* \in D$. If an initial guess x_0 is sufficiently close to $x^* \in D$, then the iterative scheme defined by* (9) *has optimal fourth-order convergence when*

$$\alpha = \frac{2}{3},\ H(1) = 1,\ H'(1) = 0,\ H''(1) = -\frac{9}{16}(\beta-1),\ |H'''(1)| < \infty, \tag{10}$$

where $\beta \in \mathbb{R}$. It satisfies the following error equation

$$e_{n+1} = \left[\Big(2 + \frac{32}{81}H'''(1) - \frac{3\beta}{2} - \frac{\beta^2}{2}\Big)c_2^3 - c_2c_3 + \frac{c_4}{9}\right]e_n^4 + O(e_n^5). \tag{11}$$

Proof Let $e_n = x_n - x^*$ be the error at n^{th} iteration and $c_k = \dfrac{1}{k!}\dfrac{f^{(k)}(x^*)}{f'(x^*)},\ k = 2, 3, \ldots$. Taking into account that $f(x^*) = 0$, we can expand $f(x_n)$ and $f'(x_n)$ about $x_n = x^*$ with the help of Taylor's series expansion. Therefore, we get

$$f(x_n) = f'(x^*)\left(e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + O(e_n^5)\right), \tag{12}$$

and

$$f'(x_n) = f'(x^*)\left(1 + 2c_2e_n + 3c_3e_n^2 + 4c_4e_n^3 + 5c_5e_n^4 + O(e_n^5)\right). \tag{13}$$

From (12) and (13), we obtain

$$\frac{f(x_n)}{f'(x_n)} = e_n - c_2e_n^2 + \left(2c_2^2 - 2c_3\right)e_n^3 + \left(-4c_2^3 + 7c_2c_3 - 3c_4\right)e_n^4 + O(e_n^5). \tag{14}$$

Using (14) in the first step of (9), we get

$$\begin{aligned} y_n = x_n - \alpha\frac{f(x_n)}{f'(x_n)} = (1-\alpha)e_n + \alpha c_2e_n^2 - 2\left(\alpha\left(c_2^2 - c_3\right)\right)e_n^3 + \alpha\left(4c_2^3 - 7c_2c_3 + 3c_4\right)e_n^4 \\ + O(e_n^5). \end{aligned} \tag{15}$$

Again, by using the Taylor series, we can easily get the following expansion of $f'(y_n)$ around simple zero x^*:

$$\begin{aligned} f'(y_n) = f'(x^*)\Big(1 + 2(1-\alpha)c_2e_n + \left(2\alpha c_2^2 + 3(1-\alpha)^2c_3\right)e_n^2 + \Big(-4\alpha c_2\left(c_2^2 - c_3\right) \\ +6(1-\alpha)\alpha c_2c_3 + 4(1-\alpha)^3c_4\Big)e_n^3 + \Big(3\left(\alpha^2c_2^2 - 4(1-\alpha)\alpha\left(c_2^2 - c_3\right)\right)c_3 \\ +12(1-\alpha)^2\alpha c_2c_4 + 2\alpha c_2\left(4c_2^3 - 7c_2c_3 + 3c_4\right) + 5(1-\alpha)^4c_5\Big)e_n^4 + O(e_n^5)\Big). \end{aligned} \tag{16}$$

Using equations (13) and (16), we obtain

$$\left[\frac{\beta+1}{\beta+\left(1-\frac{(\beta+1)(f'(x_n)-f'(y_n))}{\alpha f'(x_n)}\right)^{1/2}}\right]\frac{f(x_n)}{f'(x_n)} = e_n + \frac{1}{2}\left(-c_2^2+\beta c_2^2+2c_3-3\alpha c_3\right)e_n^3 + \frac{1}{2}\Big(2c_2^3-3\beta c_2^3+\beta^2 c_2^3-6c_2c_3 +6\alpha c_2c_3+6\beta c_2c_3-3\alpha\beta c_2c_3 +6c_4-12\alpha c_4+4\alpha^2 c_4\Big)e_n^4+O(e_n^5), \tag{17}$$

and

$$\tau=\frac{f'(y_n)}{f'(x_n)} = 1-2\left(\alpha c_2\right)e_n+3\left(2\alpha c_2^2-2\alpha c_3+\alpha^2 c_3\right)e_n^2-4\Big(4\alpha c_2^3-7\alpha c_2c_3 +3\alpha^2 c_2c_3+3\alpha c_4-3\alpha^2 c_4+\alpha^3 c_4\Big)e_n^3+O(e_n^4). \tag{18}$$

Since, it is clear from (18) that $\tau-1$ is of order $O(e_n)$. Therefore, we can expand the weight function $H(\tau)$ in the neighborhood of one using Taylor series expansion up to third-order terms as follows:

$$H(\tau)=H(1)+H'(1)\tau+\frac{1}{2!}H''(1)\tau^2+\frac{1}{3!}H'''(1)\tau^3+O(\tau^4). \tag{19}$$

Using equations (12)-(19) in scheme (9), we obtain the following error equation

$$e_{n+1}=(1-H(1))e_n+2\alpha H'(1)c_2e_n^2+\sum_{s=3}^{4}R_se_n^s, \tag{20}$$

where $R_s=R_s\left(c_2,c_3,c_4,\alpha,\beta,H(1),H^i(1)\right)$ for $i=1,2,3$.
From (20), it is clear that by inserting the following values:

$$H(1)=1,\ H'(1)=0, \tag{21}$$

we obtain atleast third-order convergence. Further, using (21) into $R_3=0$, we obtain two independent relations as follows:

$$4\alpha^2H''(1)+\beta-1=0,\quad 3\alpha-2=0, \tag{22}$$

which implies

$$\alpha=\frac{2}{3},\ H''(1)=-\frac{9}{16}\left(\beta-1\right). \tag{23}$$

Finally, using the above equations (21), (23) in (20), we obtain the following error equation

$$e_{n+1}=\left(\Big(2+\frac{32}{81}H'''(1)-\frac{3\beta}{2}-\frac{\beta^2}{2}\Big)c_2^3-c_2c_3+\frac{c_4}{9}\right)e_n^4+O(e_n^5).$$

This reveals that the modified family of Hansen-Patrick type methods (9) attains fourth-order convergence requiring only three functional evaluations, viz., $f(x_n)$, $f'(x_n)$ and $f'(y_n)$, per step.

Finally, by using (10) in (19), we get

$$H(\tau) = 1 - \frac{1}{2!}\left\{\frac{9}{16}(\beta - 1)\right\}\tau^2 + \frac{1}{3!}H'''(1)\tau^3, \quad \tau = \frac{f'(y_n)}{f'(x_n)}. \tag{24}$$

For the sake of simplicity, we take $H'''(1) = 0$ and get a wide general class of Hansen-Patrick type methods defined by

$$\begin{cases} y_n = x_n - \dfrac{2}{3}\dfrac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}\left[\dfrac{\beta + 1}{\beta + \left\{1 - \frac{(\beta+1)(f'(x_n)-f'(y_n))}{\alpha f'(x_n)}\right\}^{\frac{1}{2}}}\right]\left[1 - \dfrac{9(\beta-1)}{32}\left(\dfrac{f'(y_n)}{f'(x_n)}\right)^2\right]. \end{cases} \tag{25}$$

It satisfies the following error equation

$$e_{n+1} = \left(\left(2 - \frac{3\beta}{2} - \frac{\beta^2}{2}\right)c_2^3 - c_2c_3 + \frac{c_4}{9}\right)e_n^4 + O(e_n^5). \tag{26}$$

It is interesting to note that for $\beta = 1$ in (25), we get optimal fourth-order method proposed by Kou [11].

Special cases

In this section, we discuss some interesting special cases of our proposed scheme (9) based on different forms of weight function $H(\tau)$. In the forementioned cases, it can be easily checked that weight function $H(\tau)$ satisfies all the conditions of Theorem 1.

Case 1. Let us consider the following weight function

$$H(\tau) = \frac{1}{1 + \delta_1(\tau - 1)^2}, \tag{27}$$

where $\delta_1 = \frac{9}{32}(\beta - 1)$.

It is straight forward to see from above that the weight function has one free disposable parameter, namely β. Therefore, for different particular values of β, we get various optimal fourth-order Hansen-Patrick type methods but some of the important cases are described in Table 1.

Case 2. Now, we consider the following weight function

$$H(\tau) = \frac{1 + \delta_3(\tau - 1) + \delta_1(\tau - 1)^2}{1 + \delta_4(\tau - 1) + 2\delta_1(\tau - 1)^2}, \tag{28}$$

where δ_1 is defined by (27) and δ_3, δ_4 are free disposable parameters.

Particular sub-case of (28)*:*

For $\delta_3 = \frac{3}{2}$ and $\delta_4 = \frac{1}{2}$, weight function reads as:

$$H(\tau) = \frac{32 + 48(\tau - 1) + (25 - 9\beta)(\tau - 1)^2}{16(2 + 3(\tau - 1) + (\tau - 1)^2)}. \tag{29}$$

Similarly, by varying free parameter β, we obtain various cases but some of the important cases are displayed in Table 2.

Case 3. Now, we consider the following weight function

$$H(\tau) = \frac{1 + \delta_2(\tau - 1)}{1 + \delta_2(\tau - 1) + \delta_1(\tau - 1)^2}, \tag{30}$$

where δ_1 is defined by (27) and δ_2 is any free disposable parameter.
Particular sub-case of (30)*:*
For $\delta_2 = 1$, weight function reads as:

$$H(\tau) = \frac{\tau}{\tau + \delta_1(\tau - 1)^2}. \tag{31}$$

Hence, by varying free parameter δ_1, one can get several different cases but some of the important cases are displayed in Table 3.

Table 1: Sub-cases of weight function (27) **and their error equations**

S.No	Particular values of β	Sub-cases and their error equations
1.	$\beta = 0$	$H(\tau) = \frac{1}{1-\frac{9}{32}(\tau-1)^2}$,
	(Ostrowski's square-root type)	$e_{n+1} = (2c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
2.	$\beta = \frac{1}{2}$	$H(\tau) = \frac{1}{1-\frac{9}{64}(\tau-1)^2}$,
	(Laguerre's type)	$e_{n+1} = (\frac{9}{8}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
3.	$\beta = \frac{3}{4}$	$H(\tau) = \frac{1}{1-\frac{9}{128}(\tau-1)^2}$,
	(Laguerre's type)	$e_{n+1} = (\frac{19}{32}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.

Table 2: Sub-cases of weight function (29) **and their error equations**

S.No	Particular values of β	Sub-cases and their error equations
1.	$\beta = 0$	$H(\tau) = H(\tau) = \frac{1+\frac{3(\tau-1)}{2}+\frac{25(\tau-1)^2}{32}}{1+(\frac{3}{2}+\frac{(\tau-1)}{2})(\tau-1)}$,
	(Ostrowski's square-root type)	$e_{n+1} = (c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
2.	$\beta = \frac{1}{2}$	$H(\tau) = H(\tau) = \frac{1+\frac{3(\tau-1)}{2}+\frac{41(\tau-1)^2}{64}}{1+(\frac{3}{2}+\frac{(\tau-1)}{2})(\tau-1)}$,
	(Laguerre's type)	$e_{n+1} = (\frac{5}{8}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
3.	$\beta = \frac{3}{4}$	$H(\tau) = H(\tau) = \frac{1+\frac{3(\tau-1)}{2}+\frac{73(\tau-1)^2}{128}}{1+(\frac{3}{2}+\frac{(\tau-1)}{2})(\tau-1)}$,
	(Laguerre's type)	$e_{n+1} = (\frac{11}{32}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.

Table 3: Sub-cases of weight function (31) and their error equations

S.no	Particular values of β	Sub-cases and their error equations
1.	$\beta = 0$	$H(\tau) = \frac{\tau}{\tau - \frac{9}{32}(\tau-1)^2}$,
	(Ostrowski's square-root type)	$e_{n+1} = (\frac{4}{3}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
2.	$\beta = \frac{1}{2}$	$H(\tau) = \frac{\tau}{\tau - \frac{9}{64}(\tau-1)^2}$,
	(Laguerre's type)	$e_{n+1} = (\frac{19}{24}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.
3.	$\beta = \frac{3}{4}$	$H(\tau) = \frac{\tau}{\tau - \frac{9}{128}(\tau-1)^2}$,
	(Laguerre's type)	$e_{n+1} = (\frac{41}{96}c_2^3 - c_2c_3 + \frac{c_4}{9})e_n^4 + O(e_n^5)$.

Local convergence

The local convergence analysis of method (9) was based in the previous sections on Taylor expansions and hypotheses reaching atleast the fifth derivative of function f. These hypotheses restrict the applicability of method (9).

As a motivational example, define function f on $D = [-\frac{1}{2}, \frac{5}{2}]$ by

$$f(x) = \begin{cases} x^3 \ln x^2 + x^5 - x^4, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Choose $x^* = 1$. We have that

$$\begin{aligned} f'(x) &= 3x^2 \ln x^2 + 5x^4 - 4x^3 + 2x^2, \ f'(1) = 3, \\ f''(x) &= 6x \ln x^2 + 20x^3 - 12x^2 + 10x, \\ f'''(x) &= 6 \ln x^2 + 60x^2 - 24x + 22. \end{aligned}$$

Then, obviously, function f''' is unbounded on D. Notice that, in particular there is a plethora of iterative methods for approximating solutions of nonlinear equations. These results show that if initial point x_0 is sufficiently close to the solution x^*, then the sequence $\{x_n\}$ converges to x^*. But how close to the solution x^*, the initial guess x_0 should be? These local results give no information on the radius of convergence ball for the corresponding method. We address this question for method (9). The same technique can be used to other methods.

In particular, we use only hypotheses on the first derivative to show the local convergence of method (9) and Lipschitz constants.

Let $L_0 > 0$, $L > 0$, $M \geq 1$, $\alpha \in \mathbb{R}\backslash\{0\}$ and $\beta > 0$ be given parameters. Define function g_1 on the interval $[0, \frac{1}{L_0})$ by

$$g_1(t) = \frac{Lt + 2|1-\alpha|M}{2(1 - L_0 t)},$$

and parameters r_1 and r_A by

$$r_1 = \frac{2(1 - |1-\alpha|M)}{2L_0 + L}$$

and

$$r_A = \frac{2}{2L_0 + L}.$$

Suppose that

$$M|1-\alpha| < 1. \tag{32}$$

By (32) and the preceding definitions

$$0 < r_1 < r_A < \frac{1}{L_0},$$

$g_1(r_1) = 1$ and $0 \leq g_1(t) < 1$ for each $t \in [0, r_1)$.

Moreover, define functions p and h_p on the interval $[0, \frac{1}{L_0})$ by

$$p(t) = |\frac{\beta+1}{\alpha}|\frac{L_0(1+g_1(t))t}{1-L_0t}$$

and

$$h_p(t) = p(t) - 1.$$

We have that $h_p(0) = -1 < 0$ and $h_p(t) \to +\infty$ as $t \to \frac{1^-}{L_0}$. It follows from intermediate value theorem that function h_p has zeros in the interval $(0, \frac{1}{L_0})$. Denote by r_p the smallest such zero.

Let $\varphi : [0, \frac{1}{L_0}) \to [0, +\infty)$ be a continuous function. Furthermore, define functions g_2 and h_2 on the interval $[0, \frac{1}{L_0})$ by

$$g_2(t) = \frac{1}{2(1-L_0t)}\Big(Lt + \frac{2L_0M(1+\beta)(1+g_1(t))t}{\beta|\alpha|(1-L_0t)} + \frac{2M(1+\beta)}{\beta}\varphi(t)\Big)t,$$

and $h_2(t) = g_2(t) - 1$.

Then, again we have that $h_2(0) = -1 < 0$ and $h_2(t) \to +\infty$ as $t \to \frac{1^-}{L_0}$. Denote by r_2 the smallest zero of function h_2 in the interval $(0, \frac{1}{L_0})$. Set

$$r = min\{r_1, r_p, r_2\}. \tag{33}$$

Then, we have that

$$0 < r < r_A, \tag{34}$$

$$0 \leq g_1(t) < 1, \tag{35}$$

$$0 \leq p(t) < 1, \tag{36}$$

and

$$0 \leq g_2(t) < 1 \tag{37}$$

for each $t \in [0, r)$.

Let $U(w, \rho)$ and $\bar{U}(w, \rho)$ denote, respectively the open and closed balls in $\mathbb{R}$ with center $w \in \mathbb{R}$ and of radius $\rho > 0$. Next, we present the local convergence analysis of method (9) using the preceding notation.

Theorem 2 *Let $f : D \subseteq \mathbb{R} \to \mathbb{R}$ be a differentiable function. Suppose that there exist $x^* \in D$ and $L_0 > 0$ such that for each $x \in D$*

$$f(x^*) = 0, \ f'(x^*) \neq 0, \tag{38}$$

and

$$|f'(x^*)^{-1}(f'(x)-f'(x^*))| \leq L_0|x-x^*|. \tag{39}$$

Moreover, suppose that for each $x, y \in D_1 := D \cap U\left(x^*, \frac{1}{L_0}\right)$ *there exist* $\alpha \in \mathbb{R}\backslash\{0\}, L > 0, \beta > 0, M \geq 1$ *and a continuous nondecreasing function* $\varphi : [0, \frac{1}{L_0}) \to [0, +\infty)$ *such that*

$$M|1-\alpha| < 1,$$

$$|f'(x^*)^{-1}(f'(x)-f'(y))| \leq L|x-y|, \tag{40}$$

$$|f'(x^*)^{-1}f'(x)| \leq M, \tag{41}$$

$$|H(\tau)-1| \leq \varphi(|x-x^*|), \tag{42}$$

and

$$\bar{U}(x^*, r) \subseteq D, \tag{43}$$

where the radius of convergence r *is defined by* (33) *and* $\tau = f'(x)^{-1}f'(x-\alpha f'(x)^{-1}f(x))$. *Then, the sequence* $\{x_n\}$ *generated for* $x_0 \in U(x^*, r)\backslash\{x^*\}$ *by method* (9) *is well defined, remains in* $U(x^*, r)$ *for each* $n = 0, 1, 2, \ldots$ *and converges to the solution* x^*. *Moreover, the following estimates hold*

$$|y_n - x^*| \leq g_1(|x_n - x^*|)|x_n - x^*| \leq |x_n - x^*| < r \tag{44}$$

and

$$|x_{n+1} - x^*| \leq g_2(|x_n - x^*|)|x_n - x^*| \leq |x_n - x^*|, \tag{45}$$

where the "g" functions are defined previously. Furthermore, for $q \in [r, \frac{2}{L_0})$, *the limit point* x^* *is the only solution of equation* $f(x) = 0$ *in* $D_2 := D \cap U(x^*, q)$.

Proof The estimates (44) and (45) shall be shown using mathematical induction. By hypothesis $x_0 \in U(x^*, r)$-$\{x^*\}$ and (39), we have that

$$|f'(x^*)^{-1}(f'(x_0)-f'(x^*))| \leq L_0|x_0-x^*| < L_0 r < 1. \tag{46}$$

Using (46) and the Banach lemma on invertible functions [2, 3, 15, 17, 18], we get that $f'(x_0) \neq 0$,

$$|f'(x_0)^{-1}f'(x^*)| \leq \frac{1}{1-L_0|x_0-x^*|} \tag{47}$$

and y_0 is well defined by the first substep of method (9) for $n = 0$. We can write by (38) that

$$f(x_0) = f(x_0) - f(x^*) = \int_0^1 f'(x^* + \theta(x_0 - x^*))(x_0 - x^*)d\theta. \tag{48}$$

Notice that $|x^* + \theta(x_0 - x^*) - x^*| = \theta|x_0 - x^*| < r$, so $x^* + \theta(x_0 - x^*) \in U(x^*, r)$. Then, by (41) and (48), we get that

$$|f'(x^*)^{-1}f(x_0)| \leq M|x_0 - x^*|. \tag{49}$$

Then, using (33), (35), (38), (40), (47) and (49), we obtain in turn that

$$\begin{aligned}
|y_0 - x^*| &= |x_0 - x^* - f'(x_0)^{-1}f(x_0) + (1-\alpha)f'(x_0)^{-1}f(x_0)| \\
&\leq |f'(x_0)^{-1}f'(x^*)|| \int_0^1 f'(x^*)^{-1}\Big(f'(x^* + \theta(x_0 - x^*)) - f'(x_0)\Big)(x_0 - x^*)d\theta| \\
&+ |1-\alpha||f'(x_0)^{-1}f'(x^*)||f'(x^*)^{-1}f(x_0)| \\
&\leq \frac{L|x_0 - x^*|^2}{2(1 - L_0\|x_0 - x^*\|)} + \frac{|1-\alpha|M|x_0 - x^*|}{1 - L_0|x_0 - x^*|} \\
&= g_1(|x_0 - x^*|)|x_0 - x^*| \leq |x_0 - x^*| < r,
\end{aligned} \tag{50}$$

which shows (44) for $n = 0$ and $y_0 \in U(x^*, r)$. By (33), (36), (47) and (50), we have that

$$\begin{aligned}|(\beta+1)L_f^*(x_0)| &\leq |\frac{\beta+1}{\alpha}|\frac{1}{1-L_0|x_0-x^*|}\\ &\quad\Big[|f'(x^*)^{-1}(f'(x_0)-f'(x^*))|+|f'(x^*)^{-1}(f'(y_0)-f'(x^*))|\Big]\\ &\leq |\frac{\beta+1}{\alpha}|\frac{L_0\Big(|x_0-x^*|+|y_0-x^*|\Big)}{1-L_0|x_0-x^*|}\\ &\leq |\frac{\beta+1}{\alpha}|\frac{L_0(1+g_1(|x_0-x^*|))|x_0-x^*|}{1-L_0|x_0-x^*|}\\ &= p(|x_0-x^*|)<p(r)<1.\end{aligned} \tag{51}$$

In view of (51), $1-(\beta+1)L_f^*(x_0)\geq 0$. Hence, x_1 is well defined by the second substep of method (9) for $n=0$. Then, we can write

$$\begin{aligned}x_1-x^* &= x_0-x^*-f'(x_0)^{-1}f(x_0)+f'(x_0)^{-1}f(x_0)\Big[1-\frac{\beta+1}{\beta+\sqrt{1-(\beta+1)L_f^*(x_0)}}\Big]\\ &\quad+f'(x_0)^{-1}f(x_0)\Big[H(\tau)-1\Big]\frac{\beta+1}{\beta+\sqrt{1-(\beta+1)L_f^*(x_0)}}.\end{aligned} \tag{52}$$

Then, using (33), (37), (42), (47), (49), (50), (51), (52) and the triangle inequality, we get in turn that

$$\begin{aligned}|x_1-x^*| &\leq |x_0-x^*-f'(x_0)^{-1}f(x_0)|+|f'(x_0)^{-1}f'(x^*)||f'(x^*)^{-1}f(x_0)|\\ &\quad|\frac{(\beta+1)L_f^*(x_0)}{(\beta+\sqrt{1-(\beta+1)L_f^*(x_0)})(1+\sqrt{1-(\beta+1)L_f^*(x_0)})}|\\ &\quad+|f'(x_0)^{-1}f'(x^*)||f'(x^*)^{-1}f(x_0)||\frac{(\beta+1)}{\beta+\sqrt{1-(\beta+1)L_f^*(x_0)}}|\varphi(|x_0-x^*|)\\ &\leq \frac{L|x_0-x^*|^2}{2(1-L_0|x_0-x^*|)}+\frac{ML_0(1+\beta)(1+g_1(|x_0-x^*|))|x_0-x^*|^2}{(1-L_0|x_0-x^*|)^2\beta|\alpha|}\\ &\quad+\frac{M(\beta+1)\varphi(|x_0-x^*|)|x_0-x^*|}{|\beta|(1-L_0|x_0-x^*|)}\\ &= g_2(|x_0-x^*|)|x_0-x^*|\leq |x_0-x^*|<r,\end{aligned} \tag{53}$$

which shows (45) for $n=0$ and $x_1\in U(x^*,r)$.

The rest of the proof for estimates (44) and (45) follows using induction by simply replacing x_0, y_0, x_1 by x_n, y_n, x_{n+1} in the preceding estimates. Then, from the estimate

$$|x_{n+1}-x^*|\leq c|x_n-x^*|\leq |x_n-x^*|<r,\qquad c=g_2(|x_0-x^*|)\in[0,1),$$

we deduce that $\lim_{n\to\infty}x_n=x^*$ and $x_{n+1}\in U(x^*,r)$. Finally, to show the uniqueness part, let $Q=\int_0^1 f'(x^*+\theta(y^*-x^*))d\theta$ with $f(y^*)=0$ and $y^*\in D_2$. In view of (39), we get that

$$\begin{aligned}|f'(x^*)^{-1}(Q-f'(x^*))| &\leq L_0\int_0^1|y^*+\theta(x^*-y^*)-x^*|\\ &\leq L_0\int_0^1(1-\theta)|(x^*-y^*)|d\theta=\frac{L_0}{2}q<1.\end{aligned} \tag{54}$$

Hence, $Q\neq 0$. Then, in view of the identity $0=F(x^*)-F(y^*)=Q(x^*-y^*)$, we deduce that $x^*=y^*$.

Remark 0.1 *1. It follows from* (39) *that condition* (41) *can be dropped, if we set*

$$M(t) = 1 + L_0 t$$

or

$$M = 2,$$

since $t \in [0, \frac{1}{L_0})$.

2. The point r_A *is the convergence radius of Newton's method*

$$x_{n+1} = x_n - f'(x_n)^{-1} f(x_n), \text{ for each } n = 0, 1, 2, \ldots \tag{55}$$

given by us in [2]. It follows from (33) *that the convergence radius* r *of method* (9) *is smaller than* r_A.

3. Let us show how to choose function φ, *when* H *is defined by*

$$H(x) = 1 - \frac{9(\beta - 1)}{32}\Big(\frac{f'(x - \alpha f'(x)^{-1} f(x))}{f'(x)}\Big). \tag{56}$$

In view of the proof of Theorem 2 and (56), *we have that*

$$\begin{aligned} |H(\tau) - 1| = \frac{9}{32}|\beta - 1||\frac{f'(x^*)^{-1} f'(y_n)}{f'(x^*)^{-1} f'(x_n)}|^2 &\leq \frac{9}{32}|\beta - 1|\frac{M^2}{(1 - L_0|x_n - x^*|)^2} \\ &= \varphi(|x_n - x^*|). \end{aligned}$$

If we choose

$$\varphi(t) = \frac{9}{32}|\beta - 1|\Big(\frac{M}{1 - L_0 t}\Big)^2. \tag{57}$$

Next, we complete this section with some examples by choosing function φ as in (57).

Example 0.1 *Let* f *be a function defined on* $D = \bar{U}(0, 1)$, *which is given as follows*

$$f(x) = e^x - 1.$$

Then, $f'(x) = e^x$ *and* $x^* = 0$. *We get that* $L_0 = e - 1 < L = e^{\frac{1}{L_0}}$ *and* $M = e^{\frac{1}{L_0}}$. *The parameters using method* (9) *are:*

$$r_1 = 0.154407, \ r_p = 0.100312, \ r_2 = 0.138045, \ r_A = 0.382692$$

and as a consequence

$$r = 0.100312.$$

Example 0.2 *Returning back to the motivational example, we have that* $L = L_0 = 146.6629073$, $M = 2$ *and* $L_1 = L$. *The parameters using method* (9) *are:*

$$r_1 = 0.001515, \ r_p = 0.001138, \ r_2 = 0.004589, \ r_A = 0.00454557$$

and as a consequence

$$r = 0.001138.$$

Numerical experiments

In this section, we shall check the convergence behavior of newly proposed scheme (9) using weight functions (27), (29) and (31) (for $\beta = \frac{3}{4}$) to solve some nonlinear equations given in Table 4, which serve to check the validity and efficiency of theoretical results. These methods are denoted by $OM1$, $OM2$ and $OM3$, respectively. We compare them with existing robust methods, namely, Jarratt's method JM (3), Soleymani's methods (4) ($S1$), (5) ($S2$), Khattri and Abbasbandy method (6) (ABK), respectively. All computations have been performed using the programming package $Mathematica$ 7 [5] in multiple precision arithmetic environment. We have considered 2000 digits floating point arithmetic so as to minimize the round-off errors as much as possible.

To check the theoretical order of convergence, we calculate the computational order of convergence *(COC)* [8] denoted by ρ_c using the following formula

$$\rho_c = \frac{\log\left(|f(x_n)/f(x_{n-1})|\right)}{\log\left(|f(x_{n-1})/f(x_{n-2})|\right)}, \ n = 2, 3, ...,$$

by taking into consideration the last three approximations in the iteration process. We have considered variety of test functions of different nature to compute the errors $|x_n - x^*|$ of approximations.

For instance, we consider the following Planck's radiation law problem which calculates the energy density within an isothermal blackbody and is given by [20]:

$$\Psi(\lambda) = \frac{8\pi cP\lambda^{-5}}{e^{\frac{cP}{\lambda BT}} - 1}, \tag{58}$$

where λ is the wavelength of the radiation, T is the absolute temperature of the blackbody, B is the Boltzmann constant, P is the Planck constant and c is the speed of light. We are interested in determining wavelength λ which corresponds to maximum energy density $\Psi(\lambda)$.

Further, $\Psi'(\lambda) = 0$ implies that the maximum value of Ψ occurs when

$$\frac{\frac{cP}{\lambda BT} e^{\frac{cP}{\lambda BT}}}{e^{\frac{cP}{\lambda BT}} - 1} = 5. \tag{59}$$

If $x = \frac{cP}{\lambda BT}$, then (59) is satisfied when

$$f_1(x) = e^{-x} + \frac{x}{5} - 1 = 0. \tag{60}$$

Therefore, the solutions of $f_1(x) = 0$ give the maximum wavelength of radiation λ by means of the following formula:

$$\lambda \approx \frac{cP}{x^* BT}, \tag{61}$$

where x^* is a solution of (60).

Now, let us consider the test function f_3 is a polynomial of *Wilkinson's type* with real zeros $1, 2, 3, 4, 5$. It is well-known that this class of polynomials is ill-conditioned and small perturbations in polynomial coefficients cause drastic variations of zeros. Therefore, most of the iterative methods encounter serious difficulties in finding the zeros of *Wilkinson-like polynomials*. The errors $|x_n - x^*|$ of approximations to the corresponding zeros of test functions and computational order of convergence ρ_c are displayed in Table 5, where

Table 4: Test functions and their zeros

$f(x)$	*Root* (x^*)	*Initial Guess* (x_0)
$f_1(x) = e^{-x} + \frac{x}{5} - 1$	4.965114	5.2
$f_2(x) = (x-1)^3 - 1$	2	2.8
$f_3(x) = \prod_{i=1}^{5}(x-i)$	4	4.5
$f_4(x) = e^{-x^2+x+2} - \cos(x+1) + x^3 + 1$	-1	0.2
$f_5(x) = xe^{x^2} - \sin x^2 + 3\cos x + 5$	$-1.201576\ldots$	0.9
$f_6(x) = e^{-x^2+x+2} - 1$	-1	-0.6

Table 5: Comparison of different optimal fourth-order methods

$f(x)$		*JM*	*S*1	*S*2	*ABK*	*OM*1	*OM*2	*OM*3
f_1	$\lvert x_1 - x^*\rvert$	0.708e−6	0.301e−2	0.704e−3	0.800e−6	0.703e−6	0.701e−6	0.701−6
	$\lvert x_2 - x^*\rvert$	0.709e−28	0.821e−100	0.245e−150	0.136e−270	0.684e−28	0.675e−28	0.671e−28
	$\lvert x_3 - x^*\rvert$	0.107e−48	0.454e−400	0.107e−480	0.107e−480	0.104e−43	0.107e−48	0.107e−48
COC (ρ_c)		4.0000	4.0000	4.0427	4.0000	4.0000	4.0000	4.0000
f_2	$\lvert x_1 - x^*\rvert$	0.528e−1	0.176e−1	0.480e−1	0.550e−1	0.272e−2	0.222e−2	0.595e−3
	$\lvert x_2 - x^*\rvert$	0.459e−5	0.201e−6	0.109e−4	0.586e−4	0.141e−10	0.254e−12	0.117e−13
	$\lvert x_3 - x^*\rvert$	0.296e−21	0.384e−26	0.373e−19	0.107e−15	0.104e−43	0.433e−52	**0.178**e−56
COC (ρ_c)		3.9647	3.9855	3.9502	3.9176	3.9987	4.0002	3.9998
f_3	$\lvert x_1 - x^*\rvert$	0.108e+0	*CUR*	*CUR*	0.217e+0	0.107e+0	0.105e+0	0.106e+0
	$\lvert x_2 - x^*\rvert$	0.292e−3	*CUR*	*CUR*	0.393e+0	0.224e−3	0.139e−3	0.164e−3
	$\lvert x_3 - x^*\rvert$	0.851e−14	*CUR*	*CUR*	0.111e−1	0.238e−14	**0.297**e−15	**0.612**e−15
COC (ρ_c)		4.1747	—	—	3.8388	4.1652	4.1170	4.13333
f_4	$\lvert x_1 - x^*\rvert$	0.217e+0	*Div*	*Div*	0.219e+0	0.217e+0	0.217e+0	0.217e+0
	$\lvert x_2 - x^*\rvert$	0.176e−4	*Div*	*Div*	0.198e−4	0.176e−4	0.175e−4	0.175e−4
	$\lvert x_3 - x^*\rvert$	0.746e−20	*Div*	*Div*	0.586e−20	**0.752**e−20	**0.753**e−20	**0.753**e−20
COC (ρ_c)		3.7643	—	—	3.8464	3.7615	3.7598	3.7604
f_5	$\lvert x_1 - x^*\rvert$	0.266e−2	0.167e+0	0.119e+0	*Div*	0.661e−2	0.819e−2	0.450e−2
	$\lvert x_2 - x^*\rvert$	0.205e−10	0.468e−2	0.935e−3	*Div*	0.150e−8	0.651e−8	0.499e−9
	$\lvert x_3 - x^*\rvert$	0.475e−34	0.428e−8	0.451e−11	*Div*	**0.513**e−34	0.269e−32	**0.476**e−34
COC (ρ_c)		3.9994	3.6276	3.8074	—	4.0001	4.0022	4.0009
f_6	$\lvert x_1 - x^*\rvert$	0.102e−1	0.121e−1	0.160e−1	0.352e−1	0.423e−2	0.316e−2	0.354e−2
	$\lvert x_2 - x^*\rvert$	0.105e−7	0.846e−7	0.242e−6	0.171e−4	0.106e−9	0.585e−11	0.112e−10
	$\lvert x_3 - x^*\rvert$	0.120e−31	0.216e−27	0.138e−25	0.123e−17	0.417e−40	**0.710**e−46	0.112e−44
COC (ρ_c)		3.9950	3.9895	3.9859	3.9449	3.9976	3.9976	3.9989

CUR: Convergence to undesired root and *Div*: stands for Divergence
Note: Bold-face numbers denote the least error among the displayed methods.

$A(-h)$ denotes $A \times 10^{-h}$. On the accounts of results obtained in the Table 5, it can be concluded that the proposed methods are highly efficient as compared to the existing robust methods, when the accuracy is tested in the multi-precision digits. Additionally, the computational order of convergence *(COC)* of these methods also confirmed the above conclusions to a great extent.

Furthermore, we have also included two pathological examples to show that our proposed methods (9) will converge to the complex root without having to start with a complex number.

Example 1: $g_1(x) = x^3 - 3x^2 + 2x + \frac{2}{5}$.

In this pathological example, starting from the real initial guess $x_0 = 1.5$, our methods namely, $OM1$, $OM2$, $OM3$ for $(\beta = 3/4)$ takes only 5 iterations to converge to the complex root $1.57985 - 0.0932014I$ with error in the approximation as $2.4264e\text{-}6 - 4.39e\text{-}8I$. On the other hand, other existing methods fail to give complex roots starting from any real guess.

Example 2: $g_2(x) = x^3 + 2x^2 + 5$.

The zeros here are -2.69065, $0.345324 - 1.31873I$ and $0.345324 + 1.31873I$. Starting with real initial guess x_0 in our methods, we shall get a complex root. For instance, starting from the real initial guess $x_0 = 0.4$, our methods namely, $OM1$, $OM2$, $OM3$ for $(\beta = \frac{3}{4})$ takes only 6 iterations to converge to the complex root $0.345324 - 1.31873I$ with error in the approximation as $0.000e\text{-}17 + 2.6375I$. The other existing methods get no solution, no matter how many iterations are performed. This also demonstrates the advantage of our methods in finding complex roots without having to start with a complex initial guess.

Similar numerical experiments have been carried out on variety of problems which confirm the above conclusions to a great extent. Finally, we observe that our proposed methods have better stability and robustness as compared to the other existing methods.

Conclusions

In this study, we contribute further to the development of the theory of iteration processes and propose new fourth-order variants of Jarratt type methods for solving nonlinear equations numerically. The presented scheme is optimal in the sense of Kung-Traub conjecture and includes optimal modifications of Ostrowski's square root method, Euler's method and Laguerre's method for different values of free disposable parameter. Moreover, the local convergence of these methods is also given using hypotheses only on the first derivative and Lipschitz constants. The another most striking feature of this contribution is that the proposed methods can locate the complex roots without having to start from a complex number as would be necessary with other methods. Finally, the asserted superiority of the proposed methods is also corroborated in the numerical section.

References

[1] Amat, S., Busquier, S. and Gutiérrez, J.M (2003) Geometric constructions of iterative functions to solve nonlinear equations, *J. Comput. Appl. Math.* **157**, 197–205.

[2] Argyros, I.K. (2008) Convergence and Applications of Newton-type Iterations, Springer.

[3] Argyros, I.K. and Hilout S. (2013) Computational Methods in Nonlinear Analysis, *World Scientific Publ. Comp.* New Jersey.

[4] Gutiérrez, J.M. and Hernández, M.A. (1997) A family of Chebyshev-Halley type methods in Banach spaces, *Bull. Austral. Math. Soc*, **55**, 113–130.

[5] Hazrat, R. (2010) Mathematica: A Problem-Centered Approach. Springer, New York.

[6] Hansen, E. and Patrick, M. (1977) A family of root finding methods, *Numer. Math.* **27**, 257–269.

[7] Jarratt, P. (1969) Some efficient fourth-order multipoint methods for solving equations, *BIT* **9**, 119-124.

[8] Jay, I.O. (2011) A note on Q-order of convergence, *BIT Numerical Mathematics*, **41**, 422–429.

[9] Khattri, S.K. and Abbasbandy, S. (2011) Optimal fourth order family of iterative methods, *Mat. Vesnik*, **63**, 67–72.

[10] King, R.F. (1973) A family of fourth order methods for nonlinear equations, *SIAM J. Numer. Anal.* **10**, 876–879.

[11] Kou, J. (2007) Fourth-order variants of Cauchy's method for solving non-linear equations, *Appl. Math. Comput.* **192**, 113–119.

[12] Magreñán, Á.A. (2014) Different anomalies in a Jarratt family of iterative root finding methods, *Appl. Math. Comput.* **233**, 29–38.

[13] Argyros, I.K. and Magreñán, Á.A. (2015) On the convergence of an optimal fourth-order family of methods and its dynamics, *Appl. Math. Comput.* **252**, 336–346.

[14] Petković, M.S., Neta, B., Petković, L.D. and Džunić, J. (2013) Multipoint Methods for Solving Nonlinear Equations, Academic Press, Elsevier.

[15] Potra, F.A., Pták, V. (1984) Nondiscrete introduction and iterative processes. *Research Notes in Mathematics*, **103**, Pitman, Boston, MA.

[16] Soleymani, F., Khattri, S.K. and Karimi Vanani, S. (2012) Two new classes of optimal Jarratt-type fourth-order methods, *Appl. Math. Lett.* **25**, 847–853.

[17] Rheinboldt, W.C. (1978) An adaptive continuation process for solving systems of nonlinear equations. Mathematical models and numerical methods (A.N.Tikhonov et al. eds.) pub.3, (19), 129–142 Banach center, Warsaw, Poland.

[18] Traub, J.F. (1964) Iterative Methods for the Solution of Equations, Prentice-Hall, New Jersey.

[19] Weerakoon, S. and Fernando, G.I. (2000) A variant of Newton's method with accelerated third-order convergence, Appl. Math. Lett. **17**, 87–93.

[20] Jain, D. (2013) Families of Newton-like method with fourth-order convergence, *Int. J. Comput. Math.* **90 (5)**, 1072–1082.

The virtual node polygonal element method for fatigue crack growth simulation

***Z.H. Teng[1], † D.M. Liao[1], S.C. Wu[2], Z. B. Zhang[1], T. Chen[1], F. Sun[1]**

[1] State Key Lab. of Material Processing and Die & Mould Technology, Huazhong University of Science & Technology, China

[2] State Key Lab. of Traction Power, Southwest Jiaotong University, China

*Presenting author: tengzihao@hust.edu.cn

†Corresponding author: liaodunming@hust.edu.cn

Abstract

Simulation of fatigue crack propagation and prediction of structural life is crucial to ensure the safety of engineering structure. The standard Extended Finite Element Method (XFEM) requires sufficiently fine mesh for crack propagation problems, which is computationally expensive and inconvenient to operate. In the present work, a simple dynamic adaptive mesh refinement method is proposed using the Virtual-node Polygonal Element method within the framework of the XFEM. Through this method, a multi-level refinement of the custom region near the crack tip can be realized. The refinement area changes dynamically with the crack tip position during the crack propagation process, so that the demands of the computational cost and accuracy can be reconciled. The domain based interaction integral approach is used to obtain the stress intensity factors(SIFs). The propagation direction of the crack is determined by the maximum circumferential stress criterion and the fatigue life of the cracked structure is evaluated by Paris Law. Then two simple examples are presented and the accuracy and convergence of the algorithm are verified. Compared with the standard XFEM, the proposed method requires far fewer degrees of freedom for the same accuracy. Finally, two fatigue crack growth problems are solved, and the SIFs and fatigue life cycle show good agreement with the results available in literature.

Keywords: Polygonal element method, Dynamic adaptive mesh refinement, Fatigue crack growth

Introduction

Many engineering structures will bear the effect of cyclic load during the service process. It is very important to study the propagation behavior of the crack under cyclic loading to ensure the safety of the structure. The study of fatigue crack propagation path and the prediction of structure life cycle has being a research focus in recent years. Over the last few decades, the Finite Element Method has been widely used in the analysis of fracture problems, and a lot of commercial software and open source software packages have emerged. Due to the fact that the mesh must be aligned with the crack surfaces in FEM, re-meshing is needed once the crack is propagated, which limits the application of FEM in dynamic crack propagation simulation. In order to solve this problem, researchers have proposed many new methods, such as meshless

method[1]-[4], boundary element method (BEM)[5][6], superposition FEM[7][8], and extended finite element method (XFEM)[9]-[13]. Among these algorithms, XFEM has gained the most attentions for dynamic fracture problems with the discontinuities[14]-[16]. The key point is that augmented XFEM space of alive cracks is achieved by adopting the enrichment functions in terms of partition of unity (PU) and mesh is not required to be aligned with the crack surfaces.

Although XFEM can achieve better results in simulations of fatigue crack propagation, it is difficult for XFEM to balance the computational accuracy and computational efficiency for large and complex structures[17]. To ensure the accuracy of the stress intensity factors(SIFs) near the crack tip, it is necessary to set a fine mesh around the crack tip. On the other hand, to accurately characterize the actual crack propagation path, the crack growth increment should not be too large at each step, which also requires that the mesh near the crack tip should not be too coarse[16]. However, the location of crack tips are changing since the crack is propagated, a fine mesh for the whole domain or a local mesh refine method is needed to tackle the above problem. A fine global mesh will greatly decrease the computational efficiency while local mesh refine always needs to bring in transition elements and it will face the problem of mesh coarsening when the position of crack tips changes.

To solve the problems mentioned above, a multi-level adaptively refined mesh in XFEM at the crack tip is formulated by introducing virtual node polygonal element method (VPM) for solving elastic fracture mechanics. The foremost merit of proposed dynamic mesh refinement strategy is that the refine domain can be automatically moved with the crack tip without transition elements. Moreover, this method only needs to modify the shape function of the element and can be easily added to the existing XFEM program framework. This method has been applied to transient temperature field soldering in laser welding[19] and thermal fatigue crack propagation prediction[17].

Brief on the VP-XFEM

Small-deformation based homogeneous isotropic and linear-elastic cracked domain is considered here to derive the partial differential equations with VP-XFEM.

Governing equations

A 2D homogenous problem domain Ω without considering any discontinuities can be discretized into *N* polygonal elements with arbitrary nodes in VP-XFEM. The boundary Γ can be partitioned into the displacement Γ_u, the traction Γ_t and the traction free Γ_c. Thus the equilibrium conditions and boundary conditions are given as

$$\nabla \cdot \boldsymbol{\sigma} + \mathbf{b} = 0 \quad \text{in } \Omega \tag{1}$$

$$\mathbf{u} = \mathbf{u}_\Gamma \quad \text{at } \Gamma_u \tag{2}$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}_\Gamma \text{ at } \Gamma_t \tag{3}$$

$$\boldsymbol{\sigma} \cdot \mathbf{n} = 0 \quad \text{at } \Gamma_c \tag{4}$$

where ∇ is the divergence operator, $\boldsymbol{\sigma}$ is the Cauchy stress tensor and $\mathbf{b}$ is the body force term, $\mathbf{u}_\Gamma$ and $\mathbf{t}_\Gamma$ are the vectors of the prescribed displacements and tractions, respectively. The unit vector $\mathbf{n}$ is defined as the outward normal to the boundary Γ.
The constitution relationship is given by

$$\boldsymbol{\sigma} = \mathbf{D}\boldsymbol{\varepsilon} \tag{5}$$

$$\boldsymbol{\varepsilon} = \left[\nabla\mathbf{u} + \left(\nabla\mathbf{u}\right)^{\mathrm{T}}\right]/2 \tag{6}$$

where $\mathbf{D}$ is the matrix of material constants and $\boldsymbol{\sigma}^{\mathrm{T}}=\{\sigma_{xx}, \sigma_{yy}, \sigma_{xy}\}$ and $\boldsymbol{\varepsilon}^{\mathrm{T}}=\{\varepsilon_{xx}, \varepsilon_{yy}, \varepsilon_{xy}\}$ are the vectors of the stress and strain tensor, respectively. $\mathbf{u}=\{u, v\}^{\mathrm{T}}$ is the vector of the assumed displacement with newly-developed VP-XFEM.

VPM shape function

Assuming that a problem domain Ω bounded by Γ is discretized by N polygonal elements. For a given n-nodes polygon that can be divided into n virtual sub-triangles, the shape function of VPM consists of two components: the least-squares method (LSM) and the constant strain triangular element (CST). Assuming a point of interest p_l with the Cartesian coordinates $\mathbf{x}^{\mathrm{T}}=$(x, y, z), the shape function in VPM can be defined as

$$\Phi_l\left(\mathbf{x}\right) = W_{\mathrm{I},i}\left(\mathbf{x}\right)\left[\left(\delta_{il}+\delta_{jl}\right)\varphi_l\left(\mathbf{x}\right) + \varphi_k\left(\mathbf{x}\right)\psi_l\left(\mathbf{x}\right)\right] + W_{\mathrm{II},i}\left(\mathbf{x}\right)\psi_l\left(\mathbf{x}\right) \tag{7}$$

where the subscript i of weight function W_{I} from CST and weight function W_{II} from LSM represents that the point p_l is located inside the given sub-triangle T_i, the subscript k stands for the k-th polygonal element in the discretized domain. Note that weight functions W_{I} and W_{II} are assigned as $W_{\mathrm{I}}=\varphi_i+\varphi_j$ and $W_{\mathrm{I}}=-\varphi_k$ in which the subscript i, j and k represent the local nodal number of a sub-triangle. φ and ψ are shape functions based on the area or volume coordinates of CST and LSM, respectively.

VP-XFEM approximations

The displacement vector $\mathbf{u}_{\mathrm{XFEM}}$ in a cracked domain is

$$\mathbf{u}_{\mathrm{XFEM}}\left(\mathbf{x}\right) = \sum_{i\in I}\mathbf{u}_i\phi_i\left(\mathbf{x}\right) + \sum_{j\in J}\phi_j\left(\mathbf{x}\right)H\left(\mathbf{x}\right)\mathbf{a}_j + \sum_{k\in K}\phi_k\left(\mathbf{x}\right)\sum_{\alpha=1}^{4}\chi_\alpha\left(\mathbf{x}\right)\mathbf{b}_{k\alpha} \tag{8}$$

where I, J and K are the number of all nodes, that of interested nodes of bisected support and that of enriched nodes of bounded crack tip support, respectively; $\mathbf{u}_i$ and ϕ_i or ϕ_j are the displacement vector and shape function associated with node i or j of standard FEM, respectively; $\mathbf{a}_j$ are the additional degree of freedoms (DOFs) of bisected support by a crack path, $\mathbf{b}_k$ are the enriched DOFs associated with node k of crack tip region. $H(\mathbf{x})$ is the Heaviside function and χ_α is the component of crack tip enrichment function $\psi_\alpha(\mathbf{x})$ in which α=1, 2, 3, 4 as below

$$H\left(\mathbf{x}\right) = 1 \text{ if } \left(\mathbf{x}-\mathbf{x}'\right)\cdot\mathbf{n} \geq 0; \text{ otherwise } H\left(\mathbf{x}\right) = -1 \tag{9}$$

$$\psi_\alpha\left(r,\theta\right) = \left\{\sqrt{r}\sin\frac{\theta}{2},\quad \sqrt{r}\cos\frac{\theta}{2},\quad \sqrt{r}\sin\frac{\theta}{2}\sin\theta,\quad \sqrt{r}\cos\frac{\theta}{2}\sin\theta\right\} \tag{10}$$

where r and θ are the polar coordinates of interested point $\mathbf{x}$ with its origin crack tip.

Since the equation (8) is inconvenient for the imposing of essential displacement boundary conditions, the displacement requires to be further modified by replacing the enrichment functions with VP-XFEM as[14]

$$\mathbf{u}_{\text{VP-XFEM}}(\mathbf{x})=\sum_{i\in I}\mathbf{u}_i\phi_i(\mathbf{x})+\sum_{j\in J}\phi_j(\mathbf{x})\left[H(\mathbf{x})-H(\mathbf{x}_j)\right]\mathbf{a}_j+\sum_{k\in K}\phi_k(\mathbf{x})\sum_{\alpha=1}^{4}\left[\chi_\alpha(\mathbf{x})-\chi_\alpha(\mathbf{x}_k)\right]\mathbf{b}_{k\alpha} \tag{11}$$

Dynamic adaptive mesh refinement

VP Shape function continuity

Taking the quadrilateral for an exemplified data structure (see Fig. 1), the sub-quadrilateral element ① is further divided into four elements of ①, ⑤, ⑥ and ⑦. Thus the elements of ② and ③ have the hanging nodes of d and e, respectively. However in the VP-XFEM, there are no hanging node any more. All elements are recalled as polygonal element with variable *n* nodes, which can be treated by a VP shape function.

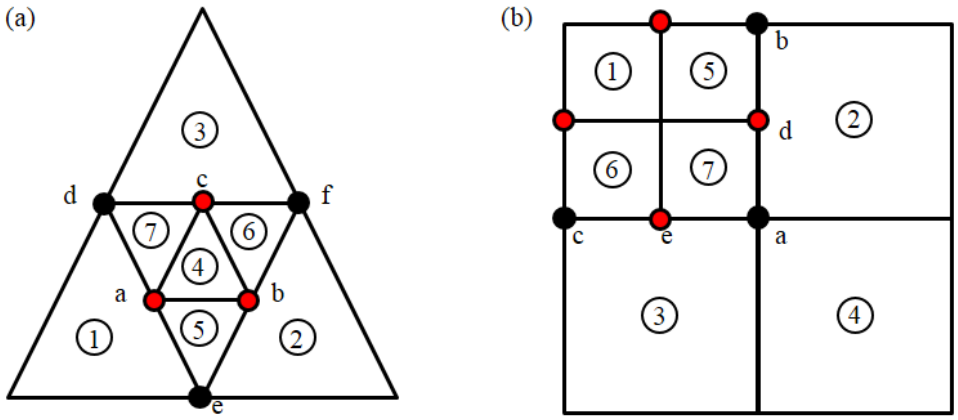


Fig. 1. The locally refined scheme with hanging nodes assigned as polygonal elements with variable *n* nodes: (a) the triangle based refinement; (b) the quadrilateral based refinement.

Based on the VP shape function in Eq. (7), the shape functions inside mixed or hybrid elements are depicted in Fig. 2. It is clearly observed that the VP shape functions possess a good continuity even near the hanging nodes. Such excellent property provides a feasible method to refine the mesh at the crack tip.

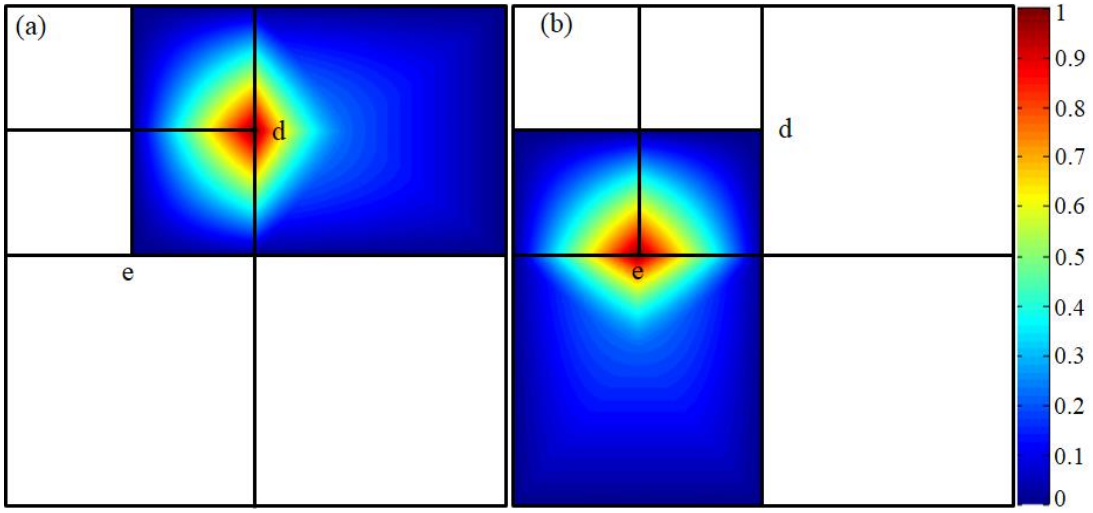


Fig. 2. The VP shape functions near hanging nodes of quadrilateral elements, in which the shape functions in (a) for d and (b) for e are zero over parts of the surrounding elements.

Single-step dynamic refinement

Based on linear elastic fracture mechanics, fatigue crack growth can be considered as a successive incident happened in the cyclic plastic zone ahead of a growing crack tip. An increment of crack extension is usually assumed to simulate the cracking process. To highlight the accuracy, a sing-step dynamic refinement algorithm is proposed along the crack path for an example of the four-point bending beam[17].Firstly, the refinement level L_r (≤ 4) and initial refinement dimension R_d are given in terms of $R_d=\alpha_r L_{max}$, in which α_r and L_{max} are the modeling parameter (α_r = 0.1~0.2 in view of the computational efficiency) and the maximum length of meshing box, respectively. Then the refinement radius of the i-th refinement level for a cracked problem can be taken as

$$r_i = R_d \beta^{i-1} \left(0 < \beta < 1\right) \tag{12}$$

where β is the refinement coefficient with an optimum value of 0.5.

For the nodes located in a circle with the radius r_i, all elements connected with those nodes require to be refined. Note that the identified circle is centered by the crack tip as shown in Fig. 3. However for multiple cracks in red herein, a domain union is created when multi cracks are available, as illustrated in Fig. 3.

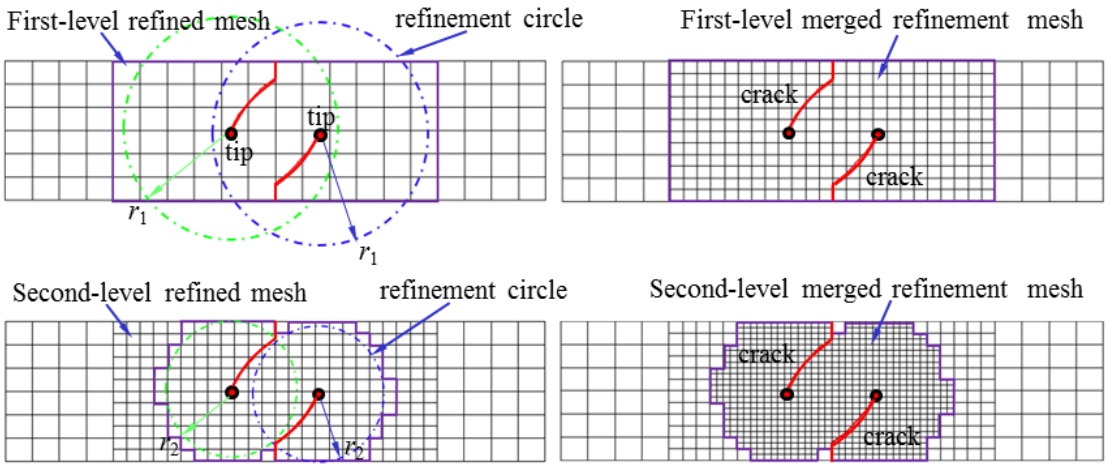


Fig. 3. The dynamic refinement of typical quadrilateral elements and the merged meshing regions around the crack after two meshing refinement operations.

Multi-step dynamic refinement

To achieve a good balance between accuracy and efficiency, two sets of meshing are introduced to deal with the refinement process. The initial background mesh with coarse elements (called the base mesh here) is stored by the first meshing set and keeps the same throughout the dynamic crack growth. The second meshing set is created dynamically with a growing crack tip, which is refined only at the crack tip with a single-step method. Fig. 4 presents the flowchart of the multi-step dynamic refinement around the crack tip with VP-XFEM.

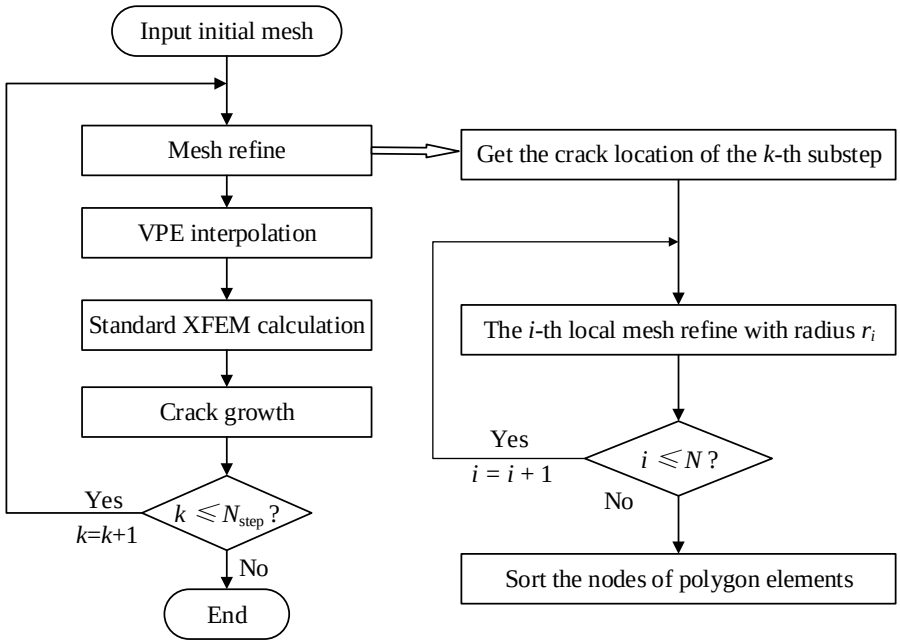


Fig. 4. A flowchart for the multi-step dynamic mesh refinement concept only around the crack tip during a typical fatigue crack growth simulation.

It is evidently seen from the flowchart that the initial background mesh is kept constant during the whole simulation and the local refinement from Eq. (12) is always conducted according to the position of a growing crack tip. The refined meshes on the crack path are retrieved to the

initial background mesh when the crack tip goes through the region. Here the four-point bending is used to present the concept as illustrated in Fig. 5. With the VP-XFEM in above Fig. 5, a four-level refined elemental meshing has been ideally realized with the movement of the crack tip for different incremental steps. The most significant advantage of current adaptively refined XFEM with the VP shape functions is actually independent of standard XFEM codes.

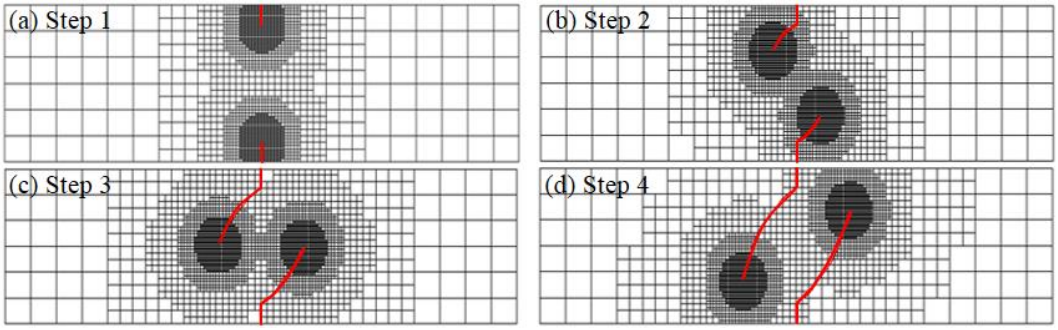


Fig. 5. The implementation of the dynamically local refinement meshing always at the crack tips with VP-XFEM, where the refined elements are bounded in terms of Eq. (12).

Crack growth modeling

For a 2D engineering crack under complex stress state, the domain based interaction integral approach is used to obtain the SIFs[10], and the maximum circumferential stress criterion is used to determine the dynamic crack growth direction

$$\theta_c = 2\arctan\left(\frac{K_I \pm \sqrt{K_I^2 + 8K_{II}^2}}{4K_{II}}\right), \quad -\pi < \theta_c < \pi \tag{13}$$

Under constant fatigue loading, the SIF range ΔK can be calculated by

$$\Delta K = K_{max} - K_{min} \tag{14}$$

where K_{max} and K_{min} are the maximum and minimum SIFs, respectively. The effective SIF range of ΔK_{eff} can thus be determined in terms of mode I cracking SIF range or ΔK_I and mode-II cracking SIF range or ΔK_{II} when approaching the critical angle θ_c by

$$\Delta K_{eff} = \Delta K_I \cos^3\left(\frac{\theta_c}{2}\right) - 3\Delta K_{II}\cos^2\left(\frac{\theta_c}{2}\right)\sin\left(\frac{\theta_c}{2}\right) \tag{15}$$

To predict the residual life of a cracked component based on elastic fracture mechanics, the proper selection of fatigue crack growth rate law is a fundamental issue for long cracks, and classical Paris model is probably the most useful in practice as

$$\frac{da}{dN} = C\left(\Delta K_{eff}\right)^m \tag{16}$$

where a and N are the crack extension and reversed cycles, respectively. C and m are the material parameters determined also dependent of testing conditions. Once crack growth resistance and crack driving force are known for a single standard crack, fatigue growth life can be integrated in terms of Eq. (16) under constant crack extension Δa. On the other hand, the

maximum crack extension $\Delta a_{\max}$ is assumed to be constant in case of an arbitrary-orientated crack or multi cracks. Therefore, the incremental length for any crack front point from VP-XFEM can be calculated by

$$\Delta a = \Delta a_{\max} \left(\frac{\Delta K_{\text{eff}}}{\Delta K_{\text{eff,max}}} \right)^{m} \tag{17}$$

where $\Delta K_{\text{eff,max}}$ is the maximum nodal ΔK_{eff} among all front points of a crack. It should be noted that the cracking simulation is stopped automatically when nodal $\Delta K_{\text{eff,max}}$ at a point of crack front is larger than measured fracture toughness K_{mat} of a material.

Numerical results

This sections verify the novel VP-XFEM models in terms of accuracy, convergence and efficiency by using different pre-cracked problems with or without analytical solutions. For quantitative examinations of the performance of proposed VP-XFEM in contrast with standard XFEM, two types of normalized error norms are introduced including the displacement error norm E_{u} and the energy error norm E_{e} as

$$E_{\text{u}} = \left\| \mathbf{u}^{\text{exa}} - \mathbf{u}^{\text{app}} \right\|_{L_2(\Omega)} = \sqrt{\frac{\int_{\Omega} \left(\mathbf{u}^{\text{exa}} - \mathbf{u}^{\text{app}} \right)^{\text{T}} \left(\mathbf{u}^{\text{exa}} - \mathbf{u}^{\text{app}} \right) \text{d}\Omega}{\int_{\Omega} \left(\mathbf{u}^{\text{exa}} \right)^{\text{T}} \left(\mathbf{u}^{\text{exa}} \right) \text{d}\Omega}} \tag{18}$$

$$E_{\text{e}} = \left\| \frac{\boldsymbol{\varepsilon}^{\text{exa}} - \boldsymbol{\varepsilon}^{\text{app}}}{\boldsymbol{\varepsilon}^{\text{exa}}} \right\| = \sqrt{\frac{\int_{\Omega} \left(\boldsymbol{\varepsilon}^{\text{exa}} - \boldsymbol{\varepsilon}^{\text{app}} \right)^{\text{T}} \text{D} \left(\boldsymbol{\varepsilon}^{\text{exa}} - \boldsymbol{\varepsilon}^{\text{app}} \right) \text{d}\Omega}{\int_{\Omega} \left(\boldsymbol{\varepsilon}^{\text{exa}} \right)^{\text{T}} \text{D} \left(\boldsymbol{\varepsilon}^{\text{exa}} \right) \text{d}\Omega}} \tag{19}$$

where the subscripts of u and e represent the displacement and energy norms, respectively. While the superscripts of exa and app stand for the approximate and exact or reference solutions to the displacement, strain and stress, respectively.

Linear patch test

The first example is to perform a patch test that is widely used in standard FEM. Three patches as shown in Fig. 6 are examined in terms of computational accuracy. Fig. 7 (a) gives a patch with 20 normal quadrilaterals, Fig. 8 (b) gives one with 26 polygonal elements where two rectangles are further divided into four quadrilaterals, thus producing 3 hanging nodes of each element. Finally, Fig. 9 (c) shows one with arbitrary-distributed and different-sized polygonal or mixed elements.

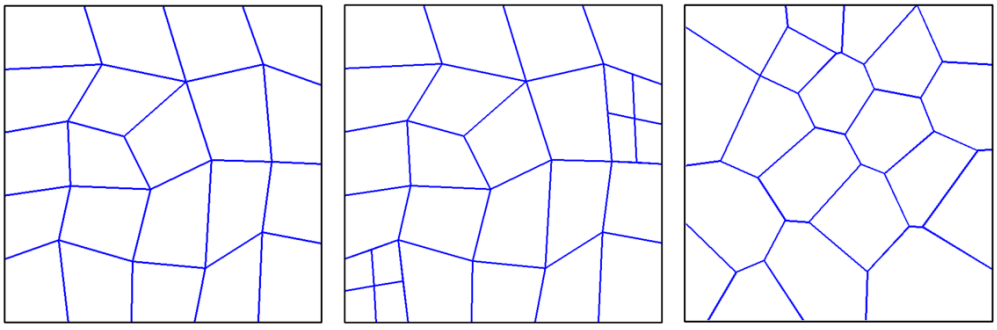


Fig. 10. Three typical meshes used for the patch tests: (a) perturbed quadrilateral mesh, (b) locally refined quadrilateral mesh and (c) full-field polygonal mesh.

The material parameters used in the testing are assumed as: Young's modulus of E=1.0 and Poisson's ratio of v=0.3. Note that in the work the international standard unit system is based unless specially denoted. For all patches, the displacements are prescribed along all boundaries by a linear basis of x and y as[20]

$$u = 1 + x + y \tag{20}$$

$$v = 1 - x + 2y \tag{21}$$

For a fully discontinuous domain or a problem without a crack, the XFEM is actually equivalent to standard FEM as built in Eq. (8). As a relatively fair comparison, 2×2 Gaussian integration scheme is employed for standard XFEM. While only one integration point into each virtual triangles adopted for VP-XFEM proposed. Table 1 lists the errors of different meshing schemes for both XFEM and novel VP-XFEM. It is found that newly-proposed VP-XFEM using local polygonal shape functions can pass the patch tests within machine precision, showing that current VP-XFEM can exactly reconstruct a linear function of imposed displacements.

Table 1. Relative errors for three mesh cases of standard finite elements, hybrid finite elements with hanging nodes and full-filed polygonal elements.

Error norms	Models	Mesh A	Mesh B	Mesh C
E_u	XFEM	3.50057e-14	—	—
	VP-XFEM	2.90270e-15	8.55441e-15	2.52088e-14
E_e	XFEM	1.11148e-14	—	—
	VP-XFEM	1.77698e-15	4.40240e-15	1.25195e-15

Besides, current VP-XFEM presents higher accuracy compared with standard XFEM. Furthermore, for locally refined elements with hanging nodes and irregularly polygonal elements, current VP-XFEM model with at least linear consistency can still satisfy the patch test at better accuracy. In contrast, original XFEM cannot cope with the problems with hanging nodes and polygonal elements.

Plate with an edge crack

The next problem is an edge-crack finite width plate of size 90mm×108mm under cyclic tensile load of F_{min}=8kN and F_{max}=16kN[16], the initial crack length a=45mm and the thickness of the plate is 6mm as shown in Fig. 7, some material parameters are taken as:Young modulus E=200GPa, Poisson's v=0.30, Paris exponent m=2.1, Paris constant C=7×10^{-8}.

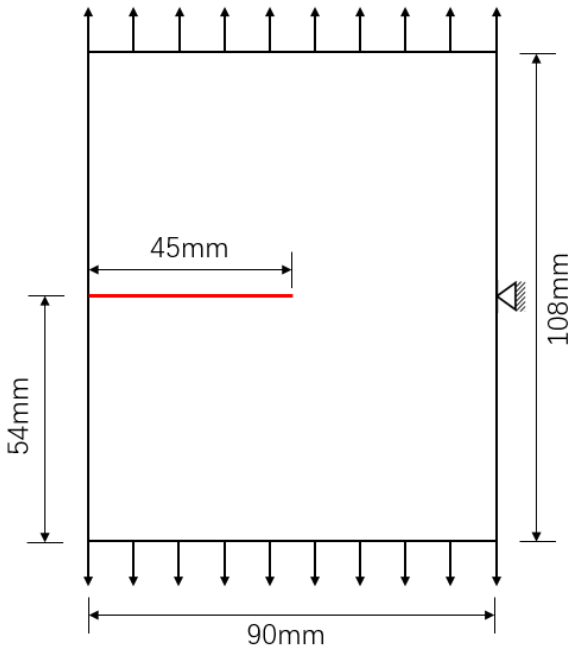


Fig. 11. Edge crack problem with a pre-crack subjected to cyclic tensile load.

The problem is solved by standard XFEM and VP-XFEM, the results are compared with the VNXFEM[16] and the experiment[21]. A uniform mesh of size 30×40 nodes is used with crack growth increment of 4.0mm for standard XFEM while the same mesh but with crack growth increments of 2.0mm and 4.0mm is used for VNXFEM. As for VP-XFEM, a coarse mesh of size 20×30 nodes is used and the refine radius α_r=0.15, refine level N=3 with crack growth increment of 2.0mm. Fig. 8 shows the variation of ΔK_{eff} as the crack grows. It can be seen that the three methods are all in good agreement with the experiment result. Besides, the VP-XFEM result is closer to the experiment through the partial amplification drawings. Moreover, the total DOFs of VP-XFEM is about 2500 to 2600 while the XFEM DOFs is around 2500, which demonstrates the computational efficiency of VP-XFEM. It is worth mentioning that the crack growth increment must be larger than the mesh size in standard XFEM, the VNXFEM breaks this limit by the partition of crack tip element, while the novel VP-XFEM achieves the same goal through the refinement of elements around the crack tip.

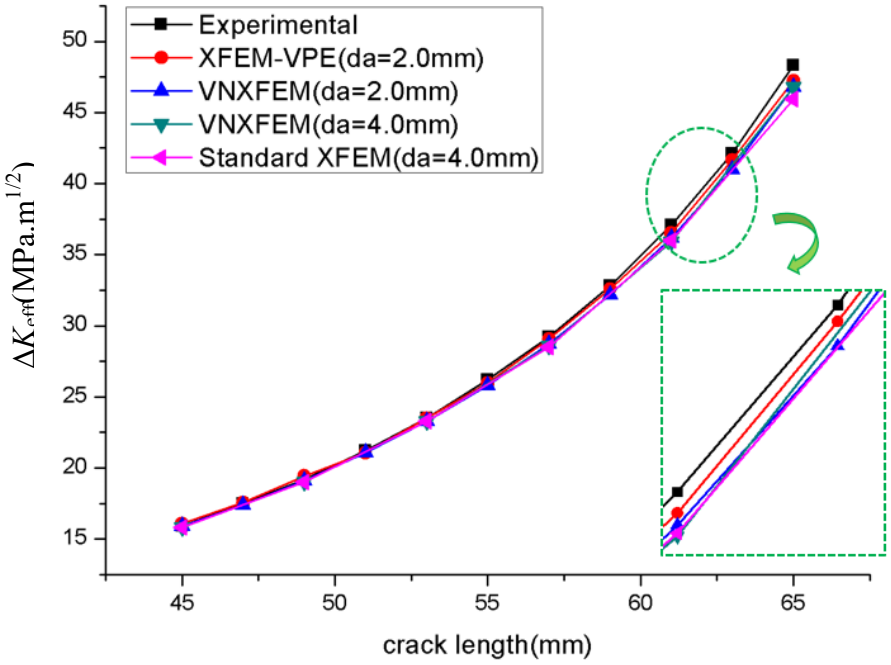


Fig. 12.The ΔK_{eff} variation with crack length of different cases.

Two internal non-colinear cracks

Based on newly-proposed VP-XFEM, this section performs a nonlinear analysis of fatigue crack growth of a rectangular plate ($90\times180\text{mm}^2$) subjected to reversed cyclic loading. In the middle of the plate under a cyclic tension ($\sigma_{\max}$=160MPa, $\sigma_{\min}$=0, then the stress ratio R=0, see Fig. 13) at both remote edges, there are two non-collinear and parallel straight cracks with the initial length a_0=10mm[4]. Moreover, the horizontal and vertical distances between two crack tip points (A) are 15mm and 5mm, respectively.

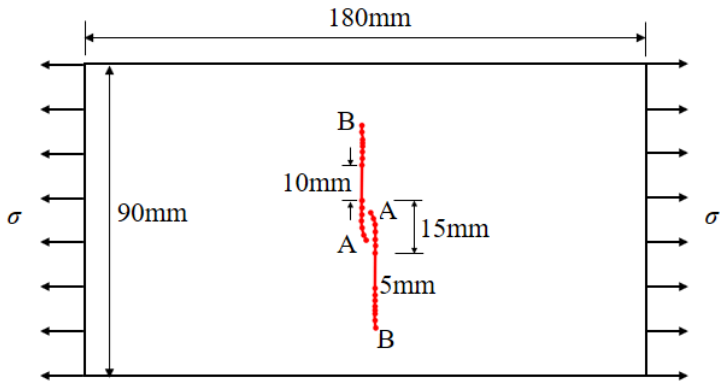


Fig. 13. Calculated cracking paths in a finite sized rectangular plate with two non-collinear cracks under cyclic tension loading at the stress ratio of R=0.

The material properties of fatigue resistance are taken as: E=74GPa, v=0.3, the fracture toughness of K_{mat}=1897.35N/mm$^{3/2}$, the Paris law based fatigue cracking parameters of C=2.087136×10^{-13} and m=3.32. The problem domain is discretized with nodes of 30×60 as an

initial background mesh for the VP-XFEM. In the fatigue cracking simulation, some parameters are taken as: the initial refinement radius is $0.1L_{max}$, the refinement level is L_r=4 and the crack increment is Δa_{max}=2mm. Note that for dynamic fracture mechanics, the large refinement level and high refinement radius should be adopted for better accurate SIFs solutions to determine the cracking direction and rate.

The evolutions of the SIF ranges ($\Delta K_{I,A}$, $\Delta K_{I,B}$, $\Delta K_{II,A}$ and $\Delta K_{II,B}$) and the effective SIF range ($\Delta K_{eff,A}$ and $\Delta K_{eff,B}$ in terms of Eq. (15)) at the most interior crack tips (A) and the crack tips near the edge (B) with the crack extension a_i is plotted in Fig. 14 and Fig. 15, respectively. As a comparison in this figure, numerical solutions from meshless method are available with totally 1416 nodes to discretize the same rectangular plate.

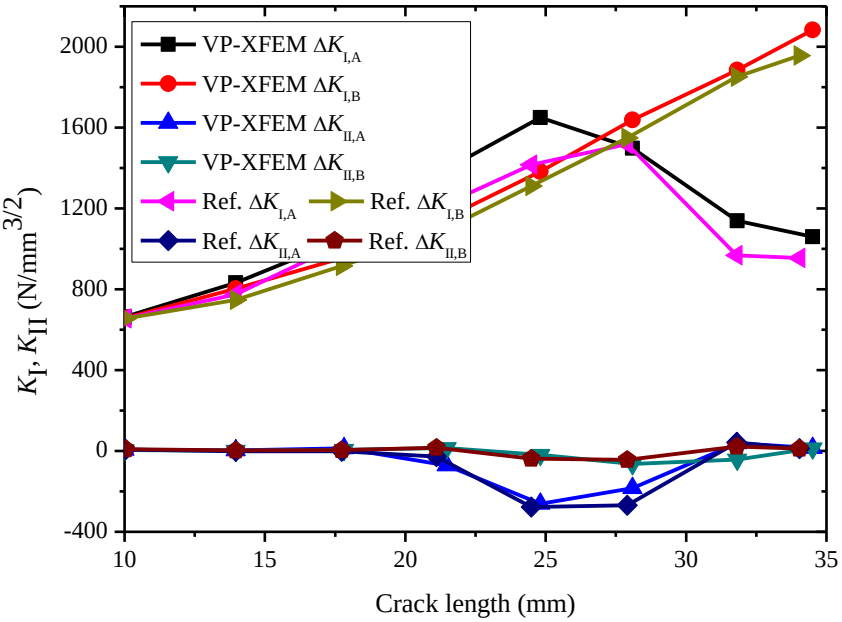


Fig. 14. The curves of SIF ranges of both points A and B with the crack length under a cyclic tension, in which meshless solutions are also provided for a comparison.

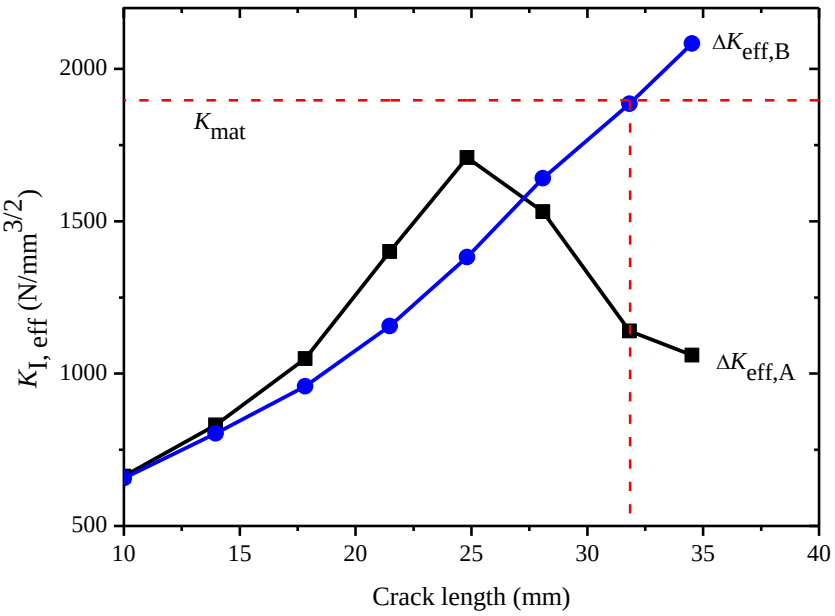


Fig. 15. The curves of effective SIF ranges of both points A and B with the crack length under a cyclic tension, in which the fracture toughness is provided to bound the cracking process.

Compared with meshless solutions with almost the DOFs, the present VP-XFEM always provides the larger SIF ranges at both points A and B, which shows an over-estimation results for a damaged body. At the starting of the cyclic loading, both $\Delta K_{II,A}$ and $\Delta K_{II,B}$ are basically zero to be a purely mode-I cracking problem. In contrast, the mode I fracture parameter $\Delta K_{I,A}$ of point A increases more rapidly than $\Delta K_{II,B}$ of point B until the mode II fracture parameter $\Delta K_{II,A}$ becomes negative. It is reasonably believed that the interaction due to crack tip stress fields happens with the crack extension of 25mm. However this phenomenon vanishes rapidly once the overlapping of two crack tips A takes place due to stress relaxation, showing a decreasing of the SIF ranges.

The effective SIF ranges in Fig. 15 are therefore feasible to evaluate the fatigue cracking process in terms of near-zero mode II fracture parameters. When the effective SIF range at point B exceeds the fracture toughness, an unstable fracture occurs. It is clearly seen that the predicted crack paths with lines in red are good agreement with both results of meshless method and experiments[4]. Based on the calculated SIF ranges of $\Delta K_{eff,A}$ and $\Delta K_{eff,B}$, the final fatigue crack growth life can be acquired in terms of Eq. , as illustrated in Fig. 16 for both points A and

(16)

B together with total values. It is found that for the same crack extension, the VP-XFEM provides a short fatigue life. The residual life of the similar cracked component is predicted as 6630 cycles, which well coincides with the experiment.

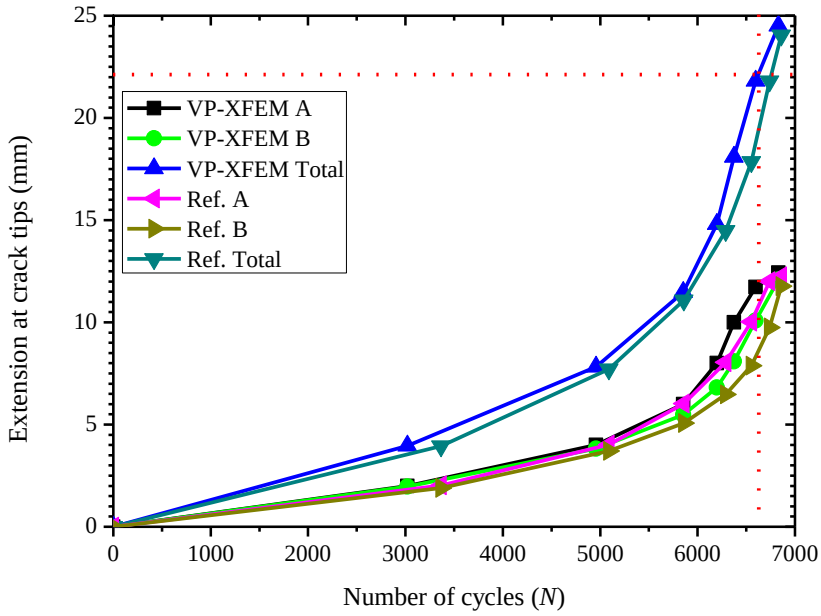


Fig. 16. Crack extension-Number of cyclic loading diagram.

Conclusions

In this paper, a theoretical study and an intensive numerical investigations on the newly-developed virtual polygonal elements based XFEM (VP-XFEM) have been carried out in terms of the local adaptive refinement strategy at the crack tip. The numerical results of VP-XFEM fully illustrate the advantages of the proposed method. The balance of computational accuracy and efficiency can be achieved by adjusting the refine parameters. Moreover, this method can be easily expanded to three dimensional problems.

References

[1] Belytschko T, Lu YY, Gu L. Element-free galerkin methods. Int J Numer Methods Eng 1994;37:229–256.

[2] Belytschko T, Lu YY, Gu L. Crack propagation by element-free Galerkin methods. Engrg.Fracture Mech 1995;51:295–315.

[3] Belytschko T, Krysl P. The Element Free Galerkin method for dynamic propagation of arbitrary 3-D cracks. Int J Numer Methods Eng 1999;44:767–800.

[4] Marc Duflot, Hung Nguyen-Dang. Fatigue crack growth analysis by an enriched meshless method. J Comput Appl Math 2004;168:155–164.

[5] Portela A, Aliabadi M, Rooke D. The dual boundary element method: effective implementation for crack problem. Int J Numer Methods Eng 1991;33:1269–1287.

[6] Yan AM, Nguyen-Dang H. Multiple-cracked fatigue crack growth by BEM. Comput Mech 1995;16:273–280.

[7] Fish J, Markolefas S, Guttal R, Nayak P. On adaptive multilevel superposition of finite element meshes for linear elastostatics. Appl Numer Math 1994;14:135–164.

[8] Kikuchi M, Wada Y, Shintaku Y. Fatigue crack growth simulation in heterogeneous material using s-version FEM. Int J Fatigue 2014;58:47–55.

[9] Melenk J, Babuska I. The partition of unity finite element method: basic theory and applications. Comput Methods Appl Mech Eng 1996;139:289–314.

[10] Moes N, Dolbow J, Belytschko T. A finite element method for crack growth without remeshing. Int J Numer Methods Eng 1999;46:131–150.

[11] Sukumar N, Moes N, Moran B, Belytschko T. Extended finite element method for three-dimensional crack modelling. Int J Numer Methods Eng 2000;48:1549–1570.

[12] Moes N, Gravouil A, Belytschko T. Non-planar 3D crack growth by the extended finite element and level sets-Part I:Mechanical model. Int J Numer Methods Eng 2002;53:2549–2568.

[13] Sukumar N, Prevost J-H. Modeling quasi-static crack growth with the extended finite element method Part I: Computer implementation. Int J Solids Struct 2003;40:7513-7537.

[14] Singh IV, Mishra BK, Bhattacharya, Patil RU. The numerical simulation of fatigue crack growth using extended finite element method. Int J Fract 2012;36:109-119.

[15] Pathak H, Singh A, Indra VS. Fatigue crack growth simulations of 3-D problems using XFEM. Int J Mech Sci 2013;76:112-131.

[16] Kumar S, Singh IV, Mishra BK, Rabczuk T. Modeling and simulation of kinked cracks by virtual node XFEM. Comput Methods Appl Mech Engrg 2015;283:1425–1466.

[17] Wu SC, Zhang SQ, Xu ZW. Thermal crack growth-based fatigue life prediction due to braking for a high-speed railway brake disc. Int J Fract 2016;87:359-369.

[18] Tang XH, Wu SC, Zheng C, Zhang JH. A novel virtual node method for polygonal elements. Appl Math Mech 2009;30(10):1233–1246.

[19] Wu SC, Peng X, Zhang WH, Stephane SPA. The virtual node polygonal element method for nonlinear thermal analysis with application to hybrid laser welding. Int J Heat Mass Transf 2013;67(3):1247–1254.

[20] Tian R, Wen LF. Improved XFEM—An extra-dof free, well-conditioning, and interpolating XFEM. Comput Methods Appl Mech Engrg 2015;285:639–658.

[21] Ma S, Zhang XB, Recho N, Li J. The mixed-mode investigation of the fatigue crack in CTS metallic specimen. Int J Fract 2006;28:1780-1790.

A reduced-order modeling technique for nonlinear buckling analysis

***Ke Liang, Qin Sun**

School of Aeronautics, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, China

Abstract:

A reduced-order modeling technique, termed the Koiter-Newton method, is presented for the elastic nonlinear structural problems. It is a combination of Koiter analysis and Newton arc-length method so that it is accurate over the whole equilibrium path but is also efficient in the presence of buckling. Various numerical examples are presented to evaluate the performance of the method.

Keywords: Koiter-Newton approach, Nonlinear structural problems, Koiter analysis, Buckling

Introduction

Nonlinear static analysis of structures is an essential step of the design of flight vehicles. It is also important in many situations of practical interest. For example, it is crucial when the displacements and/or rotations are large. Even more importantly for flight vehicles is the case where the structure (or some of its components) are prone to buckling. In many cases, it is crucial to assess the loads at which buckling occurs as well as the behavior of the structure beyond the buckling point (usually termed post-buckling analysis) [1].

Traditionally, there have been two major approaches to this problem. The first approach is the reduction method which is based on the physics-based reduced order models[2][3]. The basic idea is to significantly reduce the number of degrees of freedom in a nonlinear finite element model. Several orders of magnitude reduction in model size is possible using this approach[4] [5]. This method can be implemented in a finite element environment[2][4] and applied to moderately complex structures. Basically, there are two kinds of reduction methods. One is the Koiter's reduction method which is based on the Koiter's celebrated initial post-buckling theory [6]-[8]. It is very good for dealing with buckling sensitive structures and closely spaced modes. But it is based on just one perturbation expansion and is valid only in a small range around the buckling point. The other one is the general reduction method based on the power series expansion[9][10]. The expansions are carried out on some points along the equilibrium path in a step by step manner so that it can trace the whole nonlinear path. However, it is not good for dealing with the buckling sensitive structures. In addition, for both of the two reduction methods, there is no further link between the original finite element model and the reduced order model. Thus, the range of validity of the approximate model needs to be assessed by comparison to a full finite element analysis. This situation greatly limits the applicability of the new approach. The second approach is the finite element analysis. In this approach, the nonlinear response is traced along the equilibrium path starting from the non-

deformed position by a traditional Newton method. Now, this approach finds difficulty in tracing the response of buckling sensitive structures especially if the structure has closely spaced buckling modes[11]. In addition, this method is usually very expensive for computing the large nonlinear equations.

To achieve greater applicability, a combination of Koiter analysis and Newton arc-length method is proposed in this paper. In this Koiter-Newton approach, a reduced order model (ROM) is constructed based on the Koiter's initial post-buckling theory. This ROM is used to make an initial prediction of the response of the structure. At the new predicted point, the exact unbalanced force residual is calculated using the full finite element model. Then in a corrector step, this residual is driven to zero similar to traditional Newton arc-length methods. As the solution proceeds to higher and higher load levels, the quality of the ROM are assessed (based on the norm of force residuals) and if needed the ROM is updated to reflect changes in structural stiffness and load distribution. The proposed approach will significantly improve the efficiency of nonlinear static finite element analysis by incorporating information from Koiter's analysis while retaining the complete generality usually associated with finite element modeling.

Koiter Newton Approach

The nonlinear equilibrium equations can be written as the following simple form. It is ended with the third order about the displacement u,

$$L(u)+Q(u,u)+C(u,u,u)=f\cdot\lambda=F \tag{1}$$

where L is a linear operator, Q is a quadratic one and C is a cubic one. f is a matrix whose columns are formed by the sub-loads f_p. λ is the load parameter vector. The multiple load F is a summation of the sub-loads multiplied by the corresponding load parameters,

$$F=\sum_{p=1}^{m+1}\lambda_p f_p \tag{2}$$

where, m+1 is the total number of degrees of freedom in the reduction method. m is the number of degrees of freedom which is used in the analysis for describing the buckling branches and it is associated with the number of the closed buckling modes of the structure. 1 is the general degree of freedom for the primary path.

The displacement is also expanded to the third order with respect to the perturbation parameter a,

$$u=a_i u_i+a_i a_j u_{ij}+a_i a_j a_k u_{ijk} \tag{3}$$

where, the subscripts i,j,k=1,2,...,m+1. In the first order displacements u_i, u_1 is the displacement in the primary path; u_i(i=others) is the buckling branches. The second order displacements u_{ij} and third order displacements u_{ijk} describe the interaction effect of different first order displacement fields.

The finial reduced order model is assumed to be,

$$\lambda = \bar{L}(a) + \bar{Q}(a,a) + \bar{C}(a,a,a) \tag{4}$$

where, the $\bar{L}$, $\bar{Q}$ and $\bar{C}$ are separately the linear, quadratic and cubic operator. Introducing the equation (3) and (4) to the both sides of the equilibrium equation (1) and equating the coefficients of the various powers of a to zero, it will yield three linear equations.

$$\begin{bmatrix} K_t & -f \\ -f^T & 0 \end{bmatrix} \begin{Bmatrix} u_i \\ \bar{L}_i \end{Bmatrix} = \begin{Bmatrix} 0 \\ -E_i \end{Bmatrix} \tag{5}$$

$$\begin{bmatrix} K_t & -f \\ -f^T & 0 \end{bmatrix} \begin{Bmatrix} u_{ij} \\ \bar{Q}_{ij} \end{Bmatrix} = \begin{Bmatrix} -Q(u_i,u_j) \\ 0 \end{Bmatrix} \tag{6}$$

$$\bar{C}_{pijk} = C(u_i,u_j,u_k,u_p) - \frac{2}{3}\left(u^t_{ij} \cdot L(u_{pk}) + u^t_{jk} \cdot L(u_{pi}) + u^t_{ki} \cdot L(u_{pj})\right) \tag{7}$$

where, the subscripts i,j,k,p=1,2,...,m+1. $K_t = \bar{L}$ is the tangent stiffness matrix. In the vector E_i, only the i_{th} component is equal to one and the other components are all equal to zero. It is easy to see that only the first two linear equations need the matrix triangulation and they two have the same coefficient matrix. After solving them, the $\bar{L}$, $\bar{Q}$, $\bar{C}$, u_i, u_{ij} can be obtained. Then, the specific expression of the ROM is generated. Using the arc-length method to solve the ROM, the relationship for the load parameter λ and perturbation parameter a will be known. Introducing this relationship into the expansion of the displacement (3), the nonlinear response of the structure(λ - u) will be obtained.

In order to have an efficient algorithm, the analysis of the range of validity and the definition of a new starting point should be automatic, i.e. we have to automatically determine the values of the displacement u, over which the reduced solution will not satisfy a given accuracy.

In each step (or expansion) of the Koiter-Newton approach, this ROM is used to make an initial prediction of the response of the structure. During solving the ROM, the exact unbalanced force is calculated using the full finite element model at the end of each solution step. A criterion about the unbalanced force is given to judge when the initial prediction should be ended. If the criterion is not satisfied, the initial prediction will be stopped. Then in the following corrector step, this residual RF will be driven to zero similar to traditional Newton arc-length methods. The convergent point on the equilibrium path will be a new starting point for the next expansion. Until now, one whole step for the Koiter-Newton approach is ended. The path-tracing strategy of the proposed method is illustrated in Fig. 1. The proposed method has a larger step size to trace the nonlinear equilibrium path of the structure, compared to the conventional Newton method. This makes the method be a computationally efficient technique.

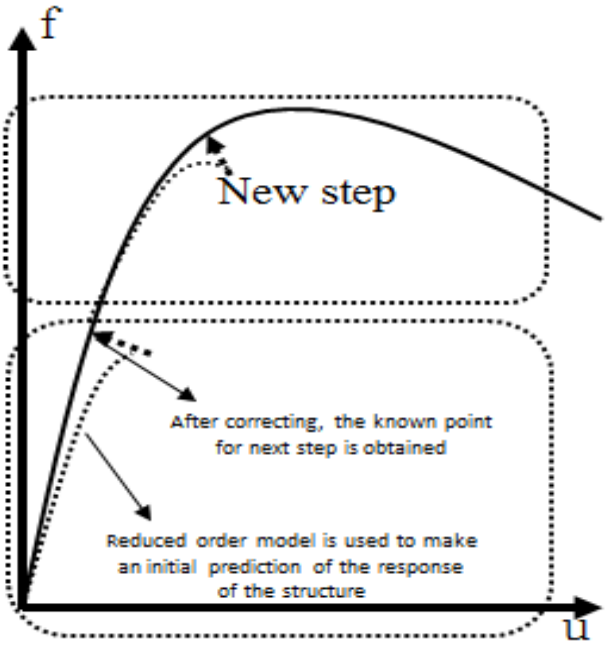
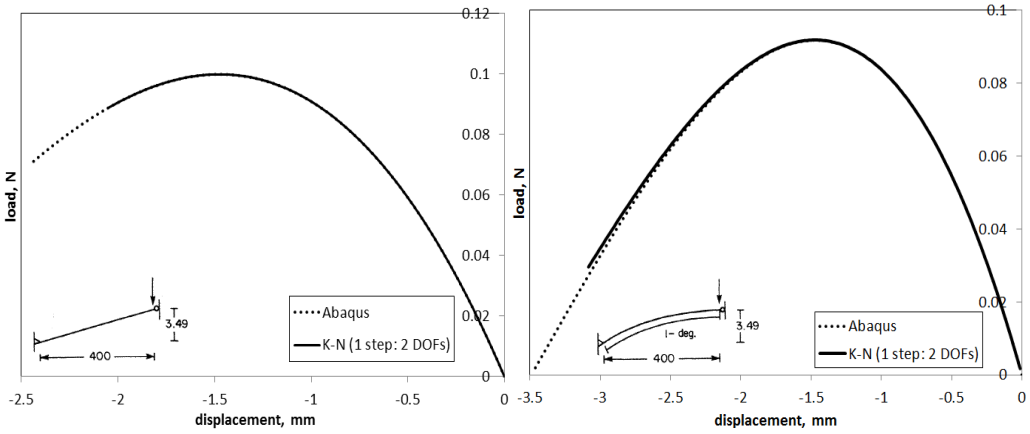
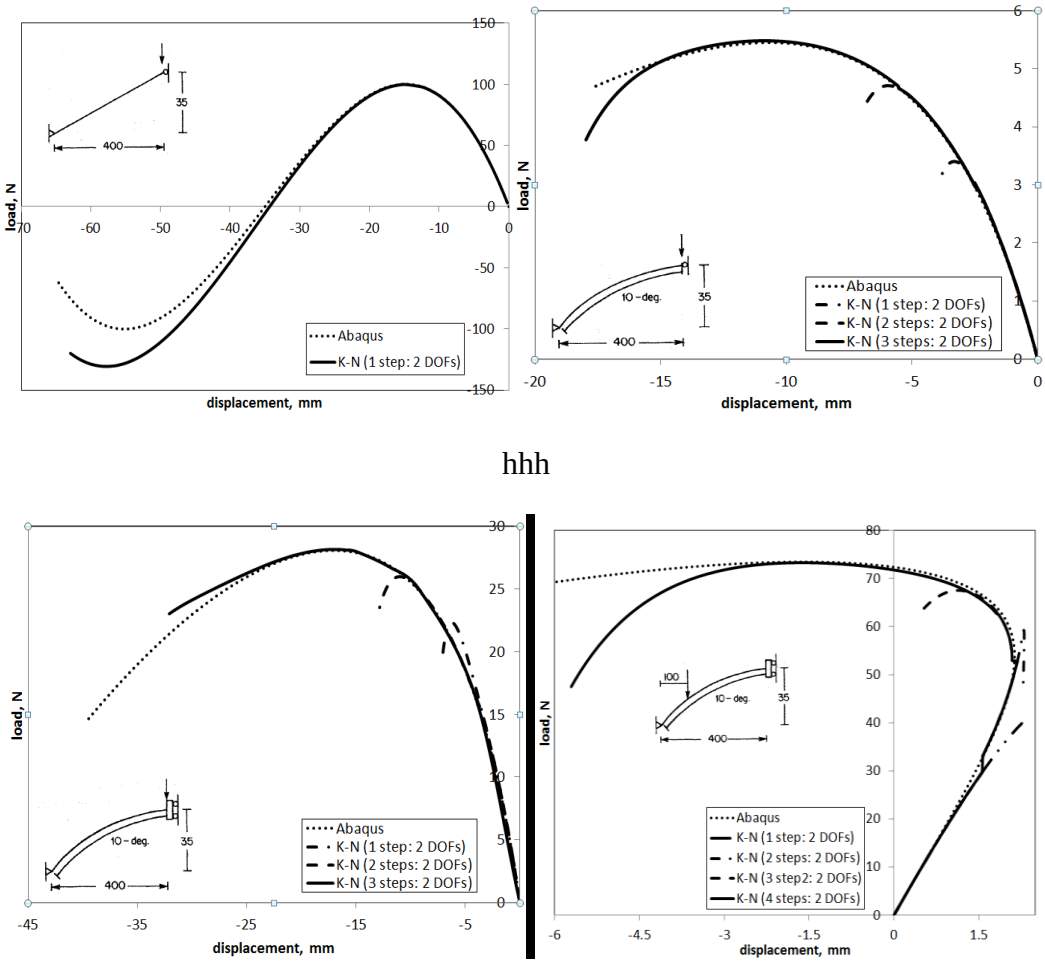


Figure 1. Path-following strategy of the proposed method

Numerical Examples

Six beams[12]which all have a nonlinear prebuckling state are analyzed in this example. They have the different shape, depth, constrain condition and loading position, as showed in figure 2. Young's modulus are all 2000MPa. The area and moment of inertia of the cross section are 391mm^2 and 2000mm^4, respectively.





hhh

Figure 2 Buckling response curves for the six single beams

Koiter-Newton approach is used to analyze these six beams. Because the first buckling load are not closed with the others, only the first buckling mode will be chosen and the number of degrees of freedom in the reduced order model is two. The nonlinear response curves (vertical displacement on the loading point vs. loading) compared with the Abaqus are in figure 2. For the beams(a)~(c), the figures show that only one perturbation step is enough to obtain an accurate buckling response(including prebuckling state, limit load and initial postbuckling state). However, because of the extremely nonlinearity of beams(e)~(g), 3, 3 and 4 steps will be needed to follow the nonlinear buckling paths, separately.

The computing consumption for reaching the same point on the postbuckling path is compared with the Abaqus. Here, the numbers of the linear equations needed to be solved are listed on table 1 for comparison. It is obviously that the Koiter-Newton approach is much more efficient than Abaqus.

Table 1. Comparison of the computing time

Beam examples	(a)	(b)	(c)	(d)	(e)	(f)
Abaqus	6	9	10	28	39	56
Koiter-Newton method	1	1	1	9	12	12

Conclusions

Based on the Koiter's initial post-buckling theory and the incremental iterative technique of the Newton method, a new reduction method, that is the Koiter-Newton method, which can trace the whole nonlinear equilibrium path automatically, is proposed. Co-rotational elements are successfully implemented into this new method. Some classical numerical examples are used to evaluate the Koiter-Newton method. If prebuckling nonlinearity is not very serious, only one perturbation step is enough to obtain the buckling characteristic. Otherwise, more steps will be needed due to the serious nonlinearity of the prebuckling. By comparing the results with Abaqus which adopts the full nonlinear finite element method, it proves that the Koiter-Newton method is automatically, accuracy and more efficient.

Reference

[1] Koiter, W. T. (1945) On the stability of the elastic equilibrium. *PhD thesis*, Delft University of Technology (in Dutch).

[2] Gallagher, R. H. (1975) Perturbation procedures in nonlinear finite element structural analysis in computational mechanics. *Lecture Notes in Mathematics*. Vol. 461, Springer Verlag, Berlin.

[3] Spottswood, S. M., Hollkamp, J. J. and Eason, T. G. (2008) On the use of reduced order models for a shallow curved beam under combined loading. *In 49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Schaumburg, Illinois, AIAA 2008-2235.

[4] Tiso, P. (2006) Finite element based reduction methods for static and dynamic analysis of thin-walled structures. *PhD thesis*, Delft University of Technology.

[5] Gordon, R. W. and Hollkamp, J. J. (2006) Reduced-order modeling of the random response of curved beams using implicit condensation. *In 47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, Newport, Rhode Island, AIAA 2006-1926.

[6] Przekop, A. and Rizzi, S. A. (2006) Nonlinear reduced order random response analysis of structures with shallow curvature. *AIAA Journal*, 44(8):1767-1778.

[7] Salerno, G. and Lanzo, A. D. (1997) A nonlinear beam finite element for the post-buckling analysis of plane frames by Koiter's perturbation approach. *Computer Methods in Applied Mechanics and Engineering*, 146:325-249.

[8] Carlo, A. D. and Pignataro, M. (1982) On nonlinear beam models from the point of view of computational post-buckling analysis. *International Journal of Solids Structures*, 18:327-347.

[9] Cochelin, B. (1994) A path-following technique via an asymptotic-numerical method. *Computers structures*, 53(5):1181-1192.

[10] Lopez. A. (1998) Scheme for tracing the equilibrium path in perturbation methods. *Computer methods in applied mechanics and engineering*, 154:193-202.

[11] Kasagi, A. and Sridharan. S. (1995) Modal interaction in composite cylinders under hydrostatic pressure. *International journal of solids structures*, 32:1349-1369.

[12] Chang, C. and Chen, J. J. (1986) Effectiveness of linear bifurcation analysis for predicting the nonlinear stability limits of structures. *International Journal for Numerical Methods in Engineering*, 23, 831-846.

Tensile Mechanical Properties and Its Failure Modes of the Basalt Fiber/Epoxy Resin Composite Material

†J.J. He, *J.P. Shi, X.S. Cao, T.L. Han

School of civil engineering and architecture, Xi'an University of Technology China
†Presenting author: hejing_86@126.com
*Corresponding author: shijp@xaut.edu.cn

Abstract

The uniaxial tensile tests of 16 types of basalt fiber / epoxy (BF/EP) composite material, formed by 4 different fiber orientation and 4 different fiber volume fraction, are carried out, with the tensile mechanical properties and its failure modes of the BF/EP composite material analyzed. The results show that the tensile strength, elastic modulus and limit strain of epoxy resin composite material increased significantly after being mixed with basalt fiber. With increasing fiber orientation angle, the tensile strength, elastic modulus and limit strain of BF/EP composite material decreased with the addition of a certain amount of fiber. However, with increasing fiber volume fraction, the tensile strength, elastic modulus and limit strain of BF/EP composite material increased for certain fiber orientation angle. There is a certain degree of agglomeration phenomenon in epoxy resin when mixed with basalt fiber with more than 1.2% of volume fraction.

Keywords: Basalt fiber; Epoxy resin; Composite material; Failure mode

Introduction

Basalt fiber is a new type of mineral fiber made from the melting of natural basalt in high temperature（1400~1500℃）[1]. Because of its characteristics such as high elastic modulus and tensile strength, it is widely applied in areas including machine building industry, aviation industry and building materials industry[2, 3，4, 5]. Basalt Fiber reinforced epoxy resin is made from a certain addition (volume mixing rate) of basalt fiber into epoxy resin. With the addition of basalt fiber, internal stress of small epoxy resin (matrix) and mechanical properties of epoxy resin can be effectively reduced [6].

At present, researches on Basalt Fiber reinforced epoxy resin are focused on the influence of fiber surface modification on mechanical properties [7, 8], yet there are fewer researches of the influence of fiber distribution on tensile properties of BF/EP composite material. However, fiber distribution can exert a significant influence on mechanical properties of BF/EP composite material [9,10,11].

As the main parameter reflecting fiber in matrix, fiber orientation is an important factor influencing mechanical properties of BF/EP composite material [12、13]. Therefore, some scholars made attempts to use flow field orientation （internal fluid viscous force）before matrix curing to control fiber orientation. For example, Yang Binxin and other scholars [14] conducted researches on flow field formed in the interval of two concentric rotating cylinders and the fiber movement and orientation in the flow field. The result shows that fiber has its movement and orientation along flow field. With the use of numerical algorithm of finite volume method and finite difference method, Zhang [15] analyzed the flowing behavior of fiber reinforced polymer melt in contraction flow chamber. The result shows that when shearing motion dominates, fiber has the orientation of cyclonic rotation; when stretching exercise dominates, fiber has the stretching orientation along monopodium.

In the thesis, with the use of handmade chute device, composite materials of oriented basalt fiber/epoxy resin composite material of different mixing rate are produced. After the completion of matrix solidification, along the fiber orientation, experiments are conducted concerning its axial tensile properties and destroying morphological analysis of its samples, with the cutting included angles of 0°,15°,30° and 45° of tensile samples. The thesis is to make people know the influence of fiber orientation and mixing rate on tensile properties of BF/EP composite material and to provide some experimental bases for the design and engineering application of BF/EP composite material.

Experiments and Researches of Tensile Properties of BF/EP composite material

Material Parameter & Raw Materials

Curing agent uses polyamide, and diluent uses acetone. Matrix is compounded according to mass ratio of M_1(epoxy resin): M_2(curing agent): M_3(diluent)=38:25:1. It is tested that after solidification, the tensile intensity of matrix is 16.67 MPa, limit strain 0.002, shearing strength 21.73MPa and curing residual stress 0.31 MPa. The fiber is Chopped discontinuous basalt fiber, and physical and mechanical indexes of basalt fiber are shown in sheet1.

Table. 1 Physical and mechanical indexes of basalt fiber

Density (g/cm^3)	Tensile strength (MPa)	Limit elongation (%)	Tensile modulus (GPa)	Diameter (mm)	Length (mm)
2.65	3300	3.2	100	0.01	12

The Design and Technology of BF/EP composite material

The volume addition i of basalt fiber range between 0.6%, 0.9%, 1.2% and 1.5%. The orientation angle θ of fiber in matrix ranges between 0°, 15°, 30° and 45°. The size of samples is 10mm×20mm×300mm and 3samples is one group.

To mix epoxy resin, curing agent and diluent according to a certain proportion, matrix of epoxy resin can be formed, which is added by basalt fiber of a certain addition and then pour it into a chute. Because of gravity, basalt fiber will flow into test mode through chute and elongation flow field of matrix will be formed which will lead to the fiber orientation approaching tensile direction[16、17]. Through the control of include angles between chute and surface, different gravity flow fields are formed. When mixing rate of fiber is 0.6%, 0.9%, 1.2% and 1.5% and included angle of chute is 15°, 23°, 30° and 45°, the fiber orientation in matrixes are relatively consistent. After the solidification of its matrix (about 48hours, 20±2℃), along the direction of fiber orientation, design sizes are cut in the angles of 0°, 15°, 30°and 45°. Therefore, 4types of fiber orientation and 4 different fiber volume fractions under 16 kinds of BF/EP composite material are gained.

Analysis of the Experiment Result

Uniaxial tensile test is conducted by adopting WAW-1000 universal testing machine and stress-strain curves are collected in the loading rate of 5N/s. The tensile stress-strain curves of different fiber orientation are given in fig. 1a、b、c、d when the volume additions of fiber are 0.6%, 0.9%, 1.2% and 1.5%.According to fig.1a, when i is 0.6% and θ is 0°,15° and 30°, the tensile strength of fiber to matrix has increased 16%, 8% and 1%, and limit strain of fiber to matrix has increased 22%, 5% and 1%. According to fig. 1b, when i is 0.9% and θ is 0°,15°, 30° and 45°, the tensile strength of fiber to matrix has increased 45%, 36%, 22% and 12%, and limiting strain of fiber to matrix has increased 29%, 22%, 8% and 4%. According to

fig. 1c, when i is 1.2% and θ is 0°,15°, 30° and 45° , the tensile strength of fiber to matrix has increased 90%, 81%, 56% and 34%, and limiting strain of fiber to matrix has increased 88%, 54%, 35% and 20%. According to fig. 1d, when i is 1.5% and θ is 0°,15°, 30° and 45° , the tensile strength of fiber to matrix has increased 106%, 98%, 80% and 58%, and limit strain of fiber to matrix has increased 114%, 75%, 38% and 21%. Therefore, the addition of basalt fiber has different improving impacts on the tensile strength and limiting strain of fiber and will increase as the mixing rate becomes higher; with increasing fiber orientation angle, the tensile strength and limit strain of the complex decreased with the addition of a certain amount of fiber.

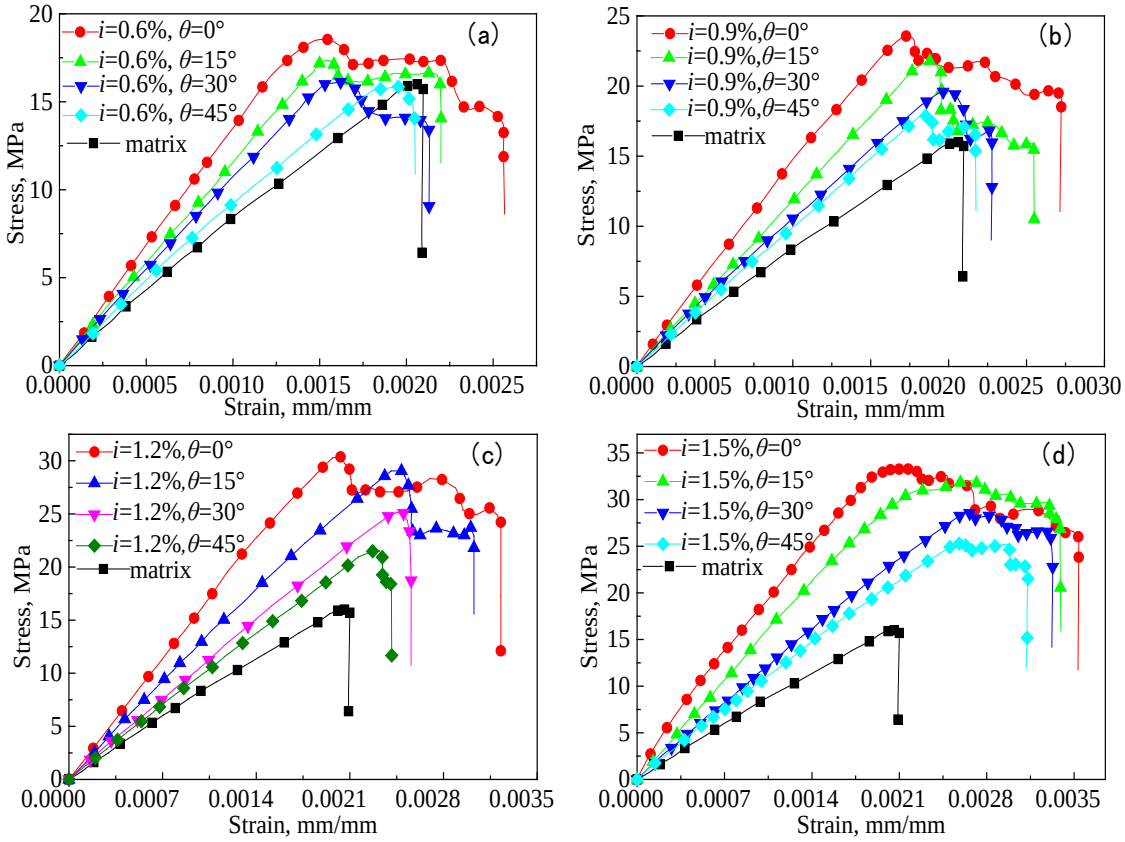


Fig.2 Tensile stress-strain curve of BF/EP composite material with different orientation and different volume mixing rate of fibers

In conclusion, the tensile strength and limit strain of BF/EP composite material have some relation with decreased with the orientation and volume mixing rate of fiber. In order to have further analysis of the influence of the orientation and volume mixing rate of fiber on tensile properties of BF/EP composite material, enhancement coefficients β is introduced, which is

$$\beta = \frac{I_f}{I_m} \tag{1}$$

In it, If is the tensile property indicator of BF/EP composite material, such as its tensile strength σ_f、 elastic modulus E_f, and limit strain ε_f; Im is the tensile property indicator of matrix, such as its tensile strength σ_m, elastic modulus E_m and limit strain ε_m; therefore, it can be got from fig.1 the enhancement coefficient of tensile strength β_t , the enhancement coefficient of elastic modulus β_e, the enhancement coefficient of limit strain β_s in different orientation and volume mixing rate of fiber . Fig. 3a, b, c are showing quadric surface fitting of β_s, β_e, β_t changing with fiber volume ratio i and orientation angle θ, whose fitting

surface function are

$$\beta_s = 1.274 - 0.3661i - 0.008179\theta + 0.4552i^2 - 0.00007442i\theta + 0.00004652\theta^2 \quad (2)$$

$$\beta_e = 1.469 + 0.05436i - 0.008847\theta + 0.2491i^2 - 0.01271i\theta + 0.000182\theta^2 \quad (3)$$

$$\beta_t = 0.3937 + 1.301i + 0.003205\theta - 0.08973i^2 - 0.008662i\theta - 0.00006437\theta^2 \quad (4)$$

Correlation coefficients of formula (2), (3) and (4) are 0.9677, 0.9783 and 0.9791. It can be known that surfaces of fig. 3a, b and c have objective laws representing β_s、β_e、β_t changing with i、θ. It can be known from fig. 3a, b, and c that a certain addition of basalt fiber can have a significant improvement on the tensile strength, elastic modulus and limit strain of epoxy resin composite material ; With increasing fiber orientation angle, the tensile strength, elastic modulus and limit strain of BF/EP composite material decreased with the addition of a certain amount of fiber. However, with increasing fiber volume fraction, the tensile strength, elastic modulus and limit strain of BF/EP composite material increased for certain fiber orientation angle.

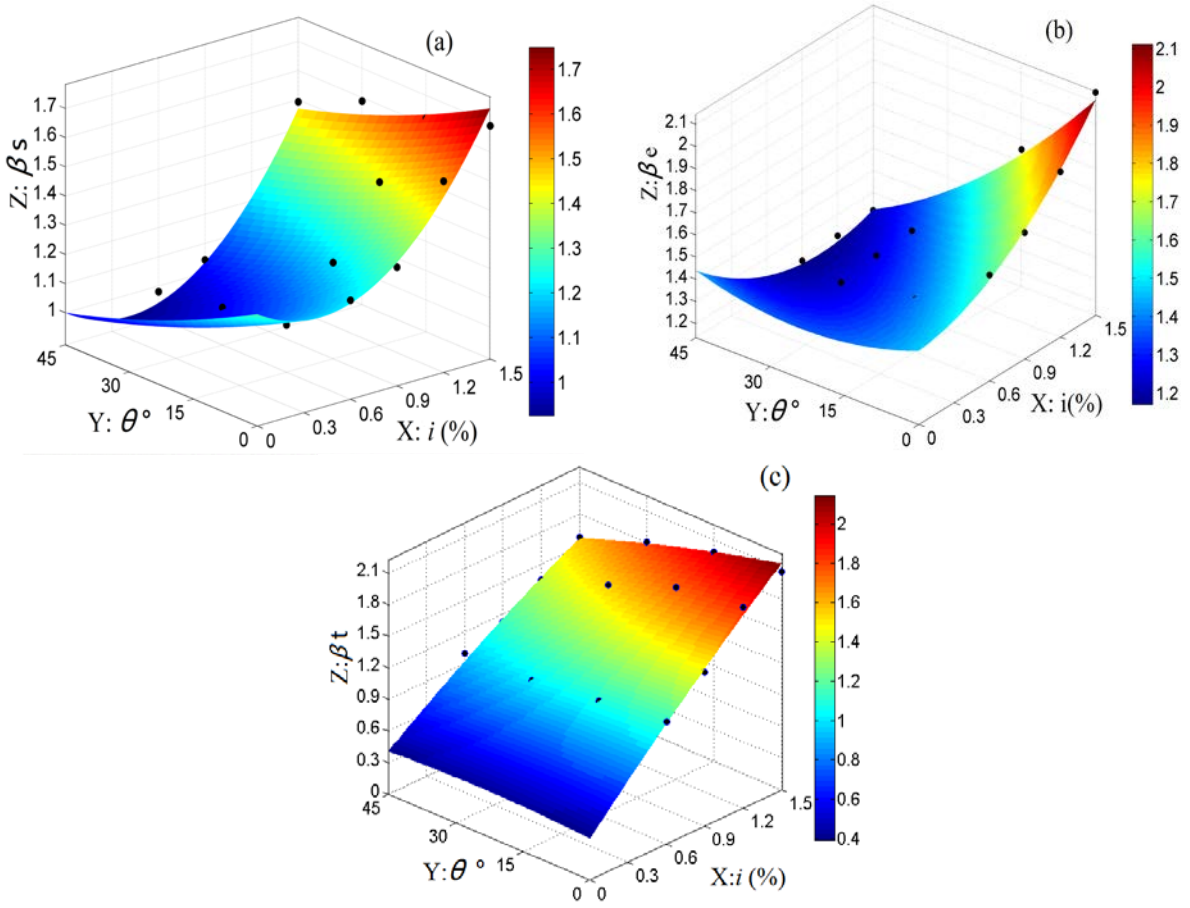


Fig.3 The fitting surface of fiber reinforcement coefficient with fiber orientation angle and fiber addition rate

Analysis of BF/EP composite material Tensile Failure Mode

To analyze the tensile failure mode of samples and to observe the facture surface of samples through SEM technology, it can be found that there is a certain degree of agglomeration phenomenon in epoxy resin when mixed with basalt fiber with more than 1.2% of volume fraction. Fig. 4 gives e-sports scan results of BF/EP composite material sample fracture

surface. Fig. 4a and 4b show agglomeration phenomenon of fiber in fracture surface when fiber volume mixing rates are 1.2% and 1.5%. It is because fiber and matrix enjoy different mixing technology, and it makes fiber hard to distribute evenly in epoxy resin and only to presenting in group or bundle [18].

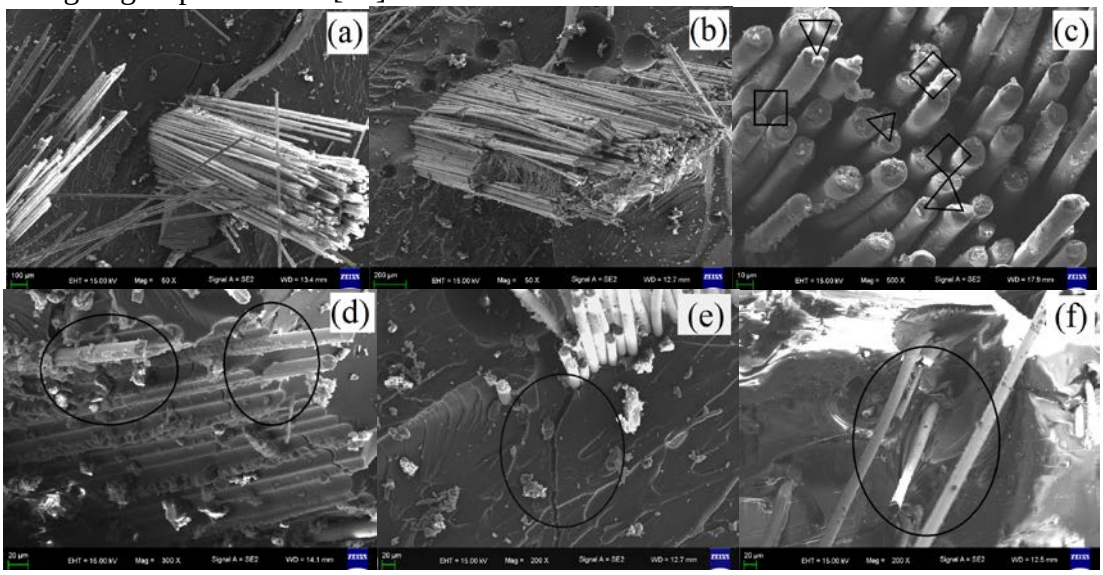


Fig.4 e-sports scan results of BF/EP composite material sample fracture surface

This kind of fiber agglomeration can be called “fiber-group effect”. It can be known that because of fiber-group effect, firstly, in and between fibers, there is only small gap, which makes it hard for matrix to enter fiber agglomerates and there is crevice in and between fiber agglomerates, seeing fig. 4c. Fibers in fiber agglomerates losing bound with matrix, most of its failure mode presents as the fiber pulling-out, while peripheral fiber of fiber agglomerates bounding well with matrix, most of its failure mode presents as the fiber pulling-off, seeing fig.4. According to Tsai Theory [19,20], matrix cracks for its tensile strength under tense force and all its load is transferred to fiber to its pull-off. Fiber approaches its yield strength and destroys with matrix, and the compounding strength of fiber and matrix is greater than its tensile strength among which fiber plays an enhanced role in matrix. However, matrix is destroyed before the pull-out fiber getting to its yield limitation, and the bonding strength of fiber and matrix is smaller than its tensile strength among which fiber doesn’t play an enhanced role in matrix. Similarly, crevice in fiber group leads to matrix cracking in advance to some degree and the tensile strength of composite decreasing, seeing fig.4e. Compared with sample fracture surface of fiber group, the fracture surface of samples is rough and has big bumps when fiber has even distribution, see fig.4f. It is because when fiber distributes evenly, it bonds well with matrix, distracting stresses concentration phenomenon within matrix and changing the stress path of matrix cracking, thus leading to the greater roughness of fracture surface. Researches show that tensile strength of solid material has some relation with its 3-dimension roughness of its fracture surface[21、22]. In conclusion, fiber-group effect has reduced the utilization of fiber to some degree and the improvement effect of fiber to matrix tensile strength. So in the design and engineering application of BF/EP composite material, attempts should be made to decrease and avoid fiber-group phenomenon.

Conclusion

The thesis analyzes different fiber orientation and BF/EP composite material tensile strength experiments and researches as well as failure mode of different fiber volume mixing rate, it can be concluded:

(1) With a certain amount of basalt fiber, the tensile strength, elastic modulus and limit strain of epoxy resin composite material have significant improvement effect. With increasing fiber orientation angle, the tensile strength, elastic modulus and limit strain of BF/EP composite material decreased with the addition of a certain amount of fiber. However, with increasing fiber volume fraction, the tensile strength, elastic modulus and limit strain of BF/EP composite material increased for certain fiber orientation angle.
(2) The internal fiber of the group cannot bond completely with the matrix for "fiber group-effect", and the tensile failure modes of BF/EP composite material are fiber pull-out and fiber pull-off. When the pulled-off fiber researched its yield limit, the BF/EP composite material failed. These situations are more associated with peripheral fiber of the group and non-group fiber. However, failure of the BF/EP composite material happened even when the pulled-out fiber did not research its yield limit. These situations are associated with internal fiber of the group.

References:

[1] Colombo, C., Vergani, L., & Burman, M. (2012) Static and fatigue characterisation of new basalt fibre reinforced composites, *Composite Structures* **94**, 1165−1174.
[2] Fiore, V., Bella, G. D., & Valenza, A. (2010) Glass–basalt/epoxy hybrid composites for marine applications. *Materials & Design* **32**, 2091−2099.
[3] Wei B, Cao H, Song S. (2010) Tensile behavior contrast of basalt and glass fibers after chemical treatment[J]. *Materials & Design* **31**,4244−4250.
[4] Lopresto, V., Leone, C., & Iorio, I. D. (2011) Mechanical characterisation of basalt fibre reinforced plastic. *Composites Part B Engineering* **42**, 717-723.
[5] Todic, A., Nedeljkovic, B., Cikara, D., & Ristovic, I. (2011) Particulate basalt–polymer composites characteristics investigation, *Materials & Design* **32**, 1677−1683.
[6] Kim, M. T., Kim, M. H., Rhee, K. Y., & Park, S. J. (2011) Study on an oxygen plasma treatment of a basalt fiber and its effect on the interlaminar fracture property of basalt/epoxy woven composites, *Composites Part B Engineering* **42**, 499−504.
[7] Lee, J. H., Rhee, K. Y., & Park, S. J. (2010) The tensile and thermal properties of modified cnt-reinforced basalt/epoxy composites, *Materials Science & Engineering A* **527**, 6838−6843.
[8] Huonnic, N., Abdelghani, M., Mertiny, P., & Mcdonald, A. (2010) Deposition and characterization of flame-sprayed aluminum on cured glass and basalt fiber-reinforced epoxy tubes, *Surface & Coatings Technology* **205**, 867−873.
[9] Fletcher, A. J., & Oakeshott, J. L. (1994) Thermal residual micro stress generation during the processing of unidirectional carbon fibre/epoxy resin composites: random fibre arrays, *Composites* **25**, 806−813.
[10] Hobbiebrunken, T., Hojo, M., Jin, K. K., & Ha, S. K. (2008) Influence of non-uniform fiber arrangement on microscopic stress and failure initiation in thermally and transversely loaded CF/epoxy laminated composites, *Composites Science & Technology* **68**, 3107−3113.
[11] Yang, C., Huang, H. X.,& Li, K. (2010) Investigation of fiber orientation states in injection-compression molded short-fiber-reinforced thermoplastics, *Polymer Composites* **31**, 1899–1908.
[12] Pochiraju, K., & Jovanović, V. (2005). Modeling material property heterogeneity in fiber-reinforced injection molded plastic parts. *Polymer Composites*, **26**(1), 98-113.
[13] Gui, L. L., Zhou, H. W., & Wang, H. W. (2011). Theoretical analysis of young's modulus for unidirectional fiber reinforced composites with different fiber orientations. *Advanced Materials Research*, **216**, 773-776.
[14] Yang Bin-Xin, Ouyang Jie, Zhou Wen, Wang Fang & Li Xue-Juan. (2015). Dynamic simulation of fiber orientation in the gap flow field between two rotating cylinders. *Acta Phys. Sin.* **64**(11), 404-413.
[15] Zhang, H., & Jie, O. (2007). Fiber orientation and stress analysis for fiber reinforced polymeric melt. *Acta Materiae Compositae Sinica*, **24**(6), 153-159.
[16] Ye, Z., Lee, D., Campbell, S. A., & Cui, T. (2011). Thermally enhanced single-walled carbon nanotube microfluidic alignment. Microelectronic Engineering, **88**(9), 2919-2923.
[17] Chandrasekaran, V. C. S., Advani, S. G., & Santare, M. H. (2010). Role of processing on interlaminar shear strength enhancement of epoxy/glass fiber/multi-walled carbon nanotube hybrid composites. Carbon, **48**(13), 3692-3699.
[18] Yang, C., Huang, H. X.,& Li, K. (2010) Investigation of fiber orientation states in injection-compression molded short-fiber-reinforced thermoplastics, *Polymer Composites* **31**, 1899–1908.
[19] Tsai, S. W., & Wu, E. M. (1971) A general theory of strength for anisotropic materials, *Journal of Composite Materials* **5**, 58−80.
[20] Liu, K. S., & Tsai, S. W. (1998) A progressive quadratic failure criterion for a laminate 1, *Composites Science & Technology* **58**, 1023-1032.

[21] Trefilov, V. I., Kartuzov, V. V., & Minakov, N. V. (2001). Fractal dimension of fracture surfaces. *Metal Science and Heat Treatment*, **43**(3), 95-98.
[22] Li, Y., Yan, Q., & Du, X. (2012). Relationship between autogenous shrinkage and tensile strength of cement paste with scm. *Journal of Materials in Civil Engineering*, **24**(10), 431.

2-D Numerical Simulation of Grounded Electrical-source Airborne Transient Electromagnetic Exploration based on Meshfree Method

***T.Z. Huang ¦ Y.Jiang ¦†Y.J. Ji [1, 2], S.S. Guan[1]**

[1] College of Instrumentation and Electrical Engineering, Jilin University, China.
[2] Key Laboratory of Geo-Exploration Instrumentation, Ministry of Education, Jilin University, China

*Presenting author: tingzhe_huang@163.com
†Corresponding author: jiyj@jlu.edu.cn

Abstract

Grounded Electrical-source Airborne Transient Electromagnetic（GREATEM）is a new geophysical exploration method. When using the traditional numerical simulation methods to deal with the forward modeling, there exits some problems such as mesh generation and low accuracy. In view of the above problems, the meshfree collocation method based on the radial basis function is applied to calculate the response of GREATEM. We derive the control iterative equations based on the Maxwell's equations, the boundary condition and stability condition are also discussed. We compare the meshfree solution with the analytical and the finite difference method in a homogeneous half space model, and calculate a circular low resistivity abnormal body model. The calculation results show that the meshfree collocation method to discrete control equation is simple and has high calculation precision, it is easy to simulate the complex models. The meshfree method is expected to be widely used in the numerical simulation of grounded airborne transient electromagnetic exploration, it will provide a new idea for the geophysical exploration modeling.

Keywords: GREATEM; meshfree; collocation; iterative equations.

Introduction

Grounded Electrical-source Airborne Transient Electromagnetic（GREATEM）is a new and hot geophysical electromagnetic exploration style in recent years. In this method. the current source is transmitted on the ground, while the receiver in the flying platform receives the electromagnetic signal. GREATEM combines the advantages of ground time-domain electromagnetic system and air time-domain electromagnetic system, it has the advantages of large depth exploration, high vertical resolution, simple and easy operation. It has been widely used in the investigation of mineral resources and engineering environment (Ito et al. 2011; Allah et al. 2013; Ji et al. 2016). Numerical simulation is an effective way to study the electromagnetic response variation law of GREATEM. The commonly used electromagnetic numerical simulation methods include finite difference method (FDM), finite element method(FEM), finite volume method(FVM) etc. These methods are based on grids, the solution domain is often divided into several certain shape meshes. In the process of dealing with the irregular abnormal body or the undulating terrain, it is necessary to divide the solution domain into small part meshes, which wastes lots of time, and the simulation result is not better.

Meshfree method is a new numerical calculation method, it does not require the predefined meshes to construct the shape functions, which can completely or partially eliminate the dependency of meshes. Meshfree method has the characteristics of simple pre-processing and easy to simulate the complex models, which has been widely concerned and used in the field of engineering calculation. In the field of geophysical electromagnetic exploration, Dai et al. (2014) carried out the 2D Ground Penetrating Radar (GPR) forward simulation with the

improved Sarma boundary condition by using Eelement-Free Galerkin method. Wittke and Tezkan (2014) used MLPG method to simulate the 2-D magnetotelluric response. Ji et al. (2016) studied the magnetotelluric response under the undulating terrain and the anisotropic media. At present, the application of meshfree method in the field of geophysics is mainly concentrated in the frequency domain and the weak form, the relevant results based on strong form (collocation method) has not been published yet.

Meshfree collocation method is a pure meshfree method, which has the advantages of direct, simple and high efficiency. In view of the characteristics of GREATEM, we apply the meshfree method to the numerical simulation of the 2-D GREATEM forward modeling in this paper. Based on the Maxwell's equations, the diffusion equation of the electric field along the strike direction is derived. We use the radial basis functions to discretize the spatial domain, and the C-N difference scheme to discretize the time, the numerical simulation is realized, and some electromagnetic response characteristics are also discussed.

Theory

GREATEM governing equation

The measurement principle of GREATEM is shown in Fig. 1. It adopts a grounded long conductor laid on the ground as the transmitting source to establish the primary field , when the current turns off, the aircraft equipped with the receiving sensor receives the second induced electromagnetic field which excites from the underground anomaly. GREATEM not only has the advantages of large depth and high resolution, but also has the advantages of wide range and high speed, it is especially suitable for mountains, forest coverages, swamps and other special landscape areas to detect the resources.

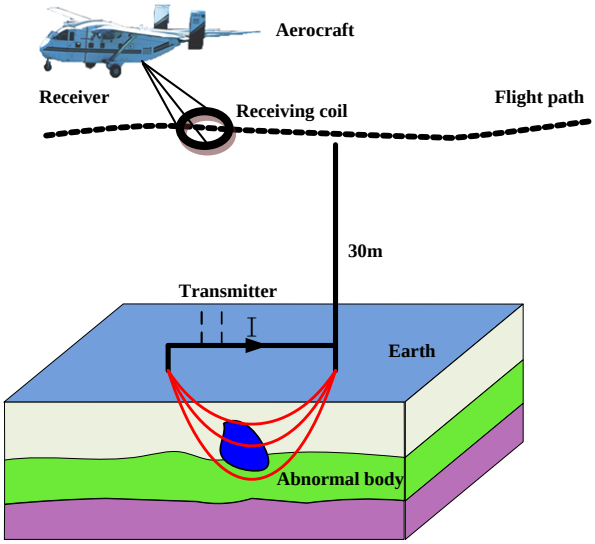


Figure 1. The measurement principle of GREATEM

Maxwell's equations are the basis of the electromagnetic exploration theory, when the electromagnetic waves propagate in a homogeneous, lossy, non-magnetic medium, the Maxwell's equations can be written as:

$$\begin{cases} \nabla \times E = -\dfrac{\partial B}{\partial t} \\ \nabla \times H = \varepsilon \dfrac{\partial E}{\partial t} + \sigma E \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \nabla \cdot E = 0 \\ \nabla \cdot H = 0 \end{cases} \tag{1}$$

Where E is the electric field intensity, H is the magnetic field intensity, B is the magnetic flux density, σ is the conductivity, ε is the dielectric constant, t is the time.

When the low frequency electromagnetic wave propagates in the lossy earth, the displacement current is relatively small, the conduction current is dominant, so the Maxwell's equation can be approximated under the quasi-static condition, we use the vector identity after ignoring the displacement current:

$$\nabla \times \nabla \times \mathrm{A} = \nabla(\nabla \cdot \mathrm{A}) - \nabla^2 \mathrm{A} \tag{2}$$

Then we can derive the diffusion equation of the electric field:

$$\nabla^2 \boldsymbol{E}(r,t) - \mu\sigma(r)\frac{\partial \boldsymbol{E}(r,t)}{\partial t} = 0 \tag{3}$$

When the length of the grounded conductor line is much greater than the distance from the observation point, the line source can be considered as an infinite 2-D source. We set the strike direction of the line source is along y axis, then there are only three components of the electromagnetic field:

$$E(x,z,t) = E_y y \qquad H(x,z,t) = H_x x + H_z \hat{z} \tag{4}$$

Therefore, the 2-D GREATEM control equation is:

$$\nabla^2 \boldsymbol{E}_y(r,t) - \mu\sigma(r)\frac{\partial \boldsymbol{E}_y(r,t)}{\partial t} = 0 \tag{5}$$

Meshfree collocation iterative equation

The Radial Basis Function (RBF) is based on the spatial distance between the center point and the collocation point, which has the advantages of simple form and isotropy (Lai et al., 2008). It is very suitable for the engineering calculation. Generally, the field function value $u(\boldsymbol{r})$ in the domain $\Re$ can be approximated by the RBF of each central node $\boldsymbol{r}_i$:

$$u(\boldsymbol{r}) \approx u^h(\boldsymbol{r}) = \sum_{i=1}^{N} \lambda_i \phi_i(\boldsymbol{r}) \qquad r \in \Re \tag{6}$$

Where $u^h(\boldsymbol{r})$ is the approximate field function, λ_i is the undetermined coefficient, $\phi_i(\boldsymbol{r}) = \phi(|\boldsymbol{r} - \boldsymbol{r}_i|)$ is the RBF between collocation point $\boldsymbol{r}$ and the central node $\boldsymbol{r}_i$. N is the number of central nodes in the domain.

Considering the MQ function has good interpolation properties and has been widely used (Kansa 1990; Cheng et al. 2003), so we use the MQ function to solve the 2-D GREATEM modeling. The basic form of difference approximation is shown in table 1.

The forward and backward difference schemes are conditionally stable, while the C-N scheme is the best of the four difference schemes, it has the highest precision and is used in this paper.

For any time t_0 in equation, the field $\boldsymbol{E}_y$ can be separated into

$$\boldsymbol{E}_y(r,t_0)=T(t_0)\boldsymbol{E}_y(r) \tag{7}$$

The C-N difference scheme is used to iterate each time step. The iterative relation in the space domain is:

$$u^{n+1}=\left(\nabla^2-\frac{\Delta t}{2K}\right)^{-1}\left[\nabla^2+\frac{\Delta t}{2\mu\sigma}\right]u^n \tag{8}$$

Table 1. Time domain difference scheme for the control equation

Difference scheme	Discrete form
Forward scheme	$\nabla^2(\boldsymbol{E}_y{}^n)-\mu\sigma\frac{\boldsymbol{E}_y{}^{n+1}-\boldsymbol{E}_y{}^n}{\Delta t}=0\Rightarrow\boldsymbol{E}_y{}^{n+1}=\boldsymbol{E}_y{}^n+\frac{\Delta t}{\mu\sigma}\nabla^2(\boldsymbol{E}_y{}^n)$
Backward scheme	$\nabla^2(\boldsymbol{E}_y{}^n)-\mu\sigma\frac{\boldsymbol{E}_y{}^n-\boldsymbol{E}_y{}^{n-1}}{\Delta t}=0\Rightarrow\boldsymbol{E}_y{}^n=\left(1-\frac{\Delta t}{\mu\sigma}\nabla^2\right)^{-1}\boldsymbol{E}_y{}^{n-1}$
Richardson scheme	$\nabla^2(\boldsymbol{E}_y{}^n)-\mu\sigma\frac{\boldsymbol{E}_y{}^{n+1}-\boldsymbol{E}_y{}^{n-1}}{2\Delta t}=0\Rightarrow\boldsymbol{E}_y{}^{n+1}=\boldsymbol{E}_y{}^{n-1}+\frac{2\Delta t}{\mu\sigma}\nabla^2(\boldsymbol{E}_y{}^n)$
Crank-Nilcoson scheme	$\frac{1}{2}(\nabla^2(\boldsymbol{E}_y{}^{n+1})+\nabla^2(\boldsymbol{E}_y{}^n))-\mu\sigma\frac{\boldsymbol{E}_y{}^{n+1}-\boldsymbol{E}_y{}^n}{\Delta t}=0\Rightarrow\boldsymbol{E}_y{}^{n+1}=\left(\nabla^2-\frac{\Delta t}{2\mu\sigma}\right)^{-1}\left[\nabla^2+\frac{\Delta t}{2\mu\sigma}\right]\boldsymbol{E}_y{}^n$

Bring the RBF approximation into Eq.(8):

$$\sum_{i=1}^{N}\lambda_i^{n+1}\phi(r_i)=\left(\nabla^2-\frac{\Delta t}{2K}\right)^{-1}\left(\nabla^2+\frac{\Delta t}{2K}\right)\sum_{i=1}^{N}\lambda_i^{n}\phi(r_i) \tag{9}$$

We rewrite it as matrix form:

$$\Phi_{M_1\times N}\lambda_{N\times 1}^{n+1}=\alpha(x)_{M_1\times 1} \tag{10}$$

The boundary matrix can be obtained by substituting the boundary conditions:

$$\Phi_{M_{2\times N}}\lambda_{N\times 1}^{n+1}=\psi(t)\Big|_{M_2\times 1}^{t=(n+1)\Delta t} \tag{11}$$

In the next time step, the weight coefficients should satisfy both the regional governing equation and the boundary condition, we combine the above matrices and then obtain λ^{n+1}:

$$\begin{bmatrix}\Phi_{M_1\times N}\\ \Phi_{M_2\times N}\end{bmatrix}\begin{bmatrix}\lambda_1^{n+1}\\ \lambda_2^{n+1}\\ \cdots\\ \lambda_N^{n+1}\end{bmatrix}=\begin{bmatrix}\alpha(x)_{M_2\times 1}\\ \psi(t)\Big|_{M_2\times 1}^{t=(n+1)\Delta t}\end{bmatrix} \tag{12}$$

Initial condition

Assuming that the earth is homogeneous, the response of the 2D current source in the homogeneous half space is taken as the initial field, its expression is (Oristaglio M L,1982):

$$E_y(x,z,t)=\frac{I}{\pi\sigma}\left\{\left[\frac{z^2-x^2}{R^2}+\frac{2z^2}{T}\right]\frac{e^{-R^2/T}}{R^2}-\frac{2ze^{-z^2/T}}{\sqrt{\pi}R^2}\left[\frac{1}{T^{1/2}}-2xF\left(xT^{-1/2}\right)\left(\frac{1}{T}+\frac{1}{R^2}\right)\right]\right\}+\frac{1}{\pi\sigma}\frac{x^2-z^2}{R^4}\left[1-erf\left(zT^{-1/2}\right)\right] \tag{13}$$

On the earth surface, it can be simplified as:

$$E(x,z=0,t)=\frac{I}{\pi\sigma}\frac{1}{x^2}(1-e^{-x^2/T}) \tag{14}$$

Boundary condition

The electric field in the TE mode is always continuous in the solution region. When the calculation domain is large enough and the boundary is far away from the abnormal body, we can set it be zero. On the ground-air boundary, the initial value can be calculated by the analytical formula. The electric field in the air can be realized by upward continuation theory (Wang and Hohmann,1993).

In order to save computation time, the step length is gradually increased with time, and the maximum time step can be:

$$\Delta t_{\max}=\frac{\left(\mu\sigma_{\min}t\right)^{1/2}\Delta_{\min}}{2} \tag{15}$$

Where Δ is the minimum node spacing.

Model calculation

Algorithm verification

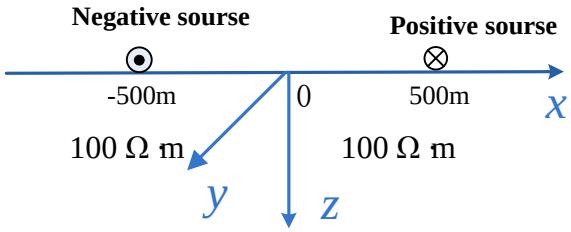


Figure 2 Homogeneous half space model

In order to verify the correctness of the meshfree method, a homogeneous half space model is built as shown in figure 2, the resistivity is ρ=100 Ω·m , there are two emission line source ,the positive source is located at x=500m, while the negative one is in x=-500m. The nodes are distributed evenly, the nodes near the source are dense while the nodes far from the source is gradually increasing. The number of the nodes and the node spacing are shown in Table 2.

Table 2. Node number and node spacing

x direction		z direction	
Node No.（i）	spacing（m）	Node No.（j）	spacing（m）

1-10	240	1-53	10
11-15	120	54-58	15
16-20	60	59-63	30
21-25	30	64-68	60
26-35	15	69-73	120
36-165	10	74-78	240
166-175	15		
176-180	30		
181-185	60		
186-190	120		
191-200	240		

We respectively calculate the homogeneous half space model by using analytical method , the Dufort-Frankel finite difference method (Oristaglio et al, 1984) and the meshfree collocation method under the same conditions. The calculation results of the meshfree collocation method and the analytical method at different time are shown in Figure 3. The error comparison curve of the meshfree method and the finite difference method is shown in figure 4.

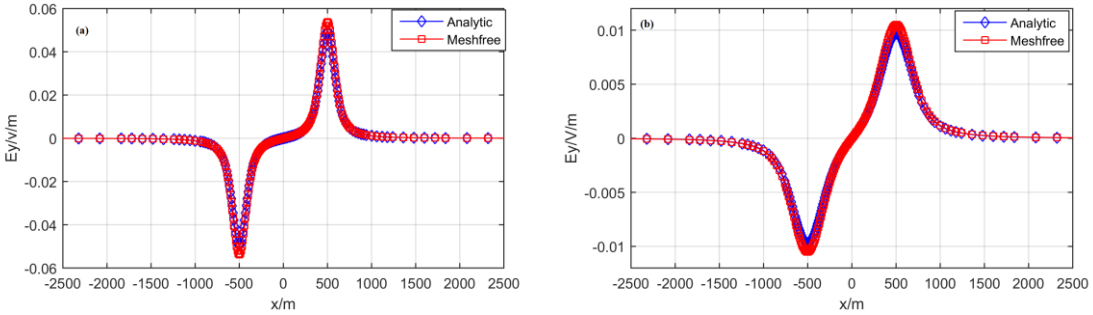


Figure 3 The comparison between the meshfree method and the analytical method （a）0.2ms （b）1ms

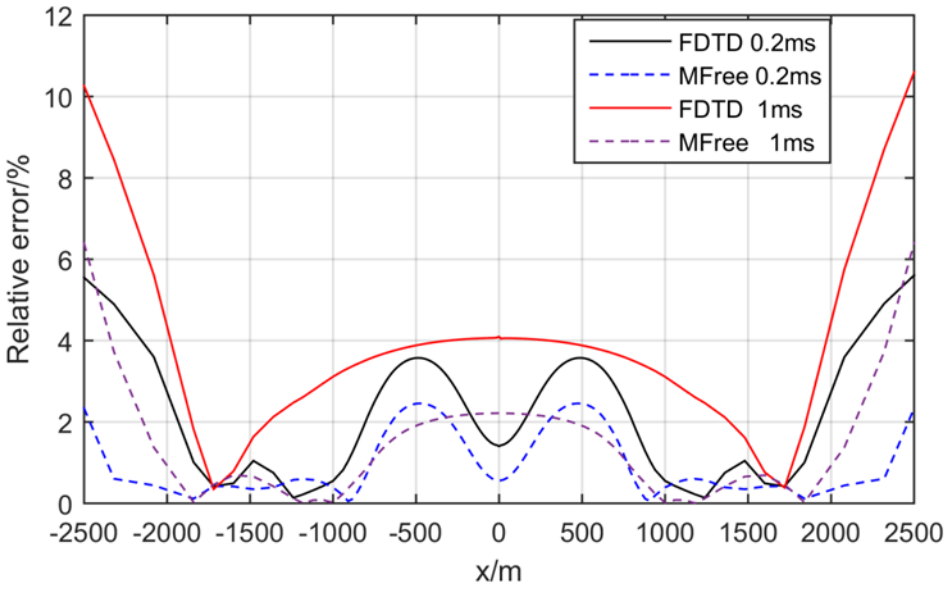


Figure 4 The error comparison curve of the meshfree method and the finite difference method

As ban be seen from Figure 3, the meshfree method and the analytical method fit well at different time, the maximum relative error is less than 3% in 1ms. The results show that meshfree method can effectively calculate 2D GREATEM response. Under the same time step and node distribution, the Dufort-Frankel finite difference solution is also calculated. As can be seen from Figure 4, after 1000 iterations to 0.2ms, the average relative error of FDTD is about 2.54%, while the meshfree method is about 1.21%, After 9000 iterations to 1ms, the average relative error of FDTD is about 3.84%, while the meshfree method is about 1.75%, its calculation precision is higher than the finite difference method. At the same time, the meshfree method pretreatment is simple and direct, which solves the problem of mesh generation and local approximation.

Anomaly model

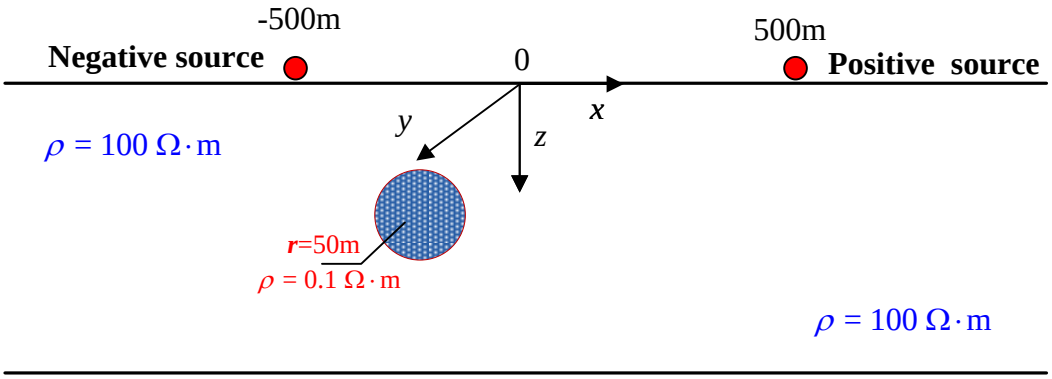


Figure 5 Circular low resistivity anomaly model

We establish a low resistivity anomaly model as shown in Figure 5, the resistivity of the homogeneous half space is 100 Ω·m, the resistivity of the circular anomaly bodyis 10 Ω·m. The center of the anomalous body is located in x=-200m, z=100m, its radius is 50m. The emission current is 100A, which is located at x=±500m. The node distribution are same as Table 2. The electric field section and profile at different time are respectively shown in Figure 6 and Figure 7.

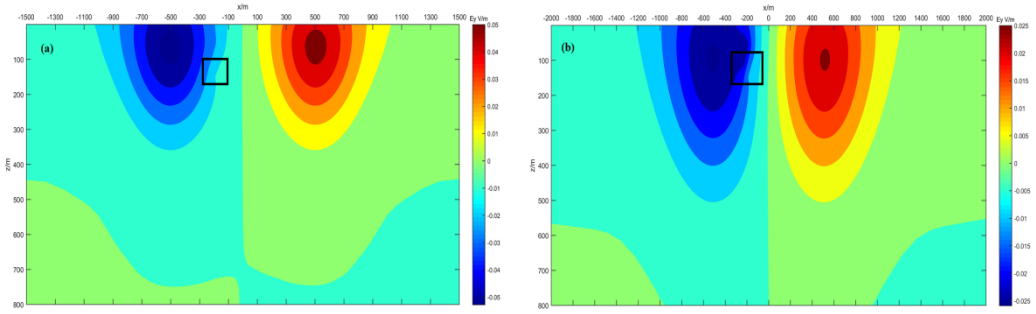


Figure 6 The electric field section of circular low resistivity anomaly model （a）0.2ms （b）0.5ms

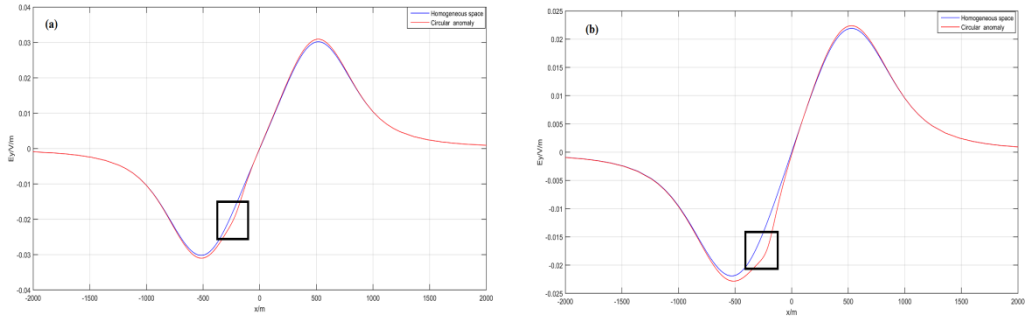


Figure 7 The electric field profile of circular low resistivity anomaly model （a）0.2ms （b）0.5ms

The anomaly body in Figure 5 is circular, and it is difficult to simulate with the traditional mesh-based method. Although it can be simulated by using a sufficiently small mesh or adaptive finite element method, it greatly increases the cost of computation. The meshfree method is out the restriction of the grid, and is more flexible than the grid-based method in dealing with the terrain undulation interface, it is easy to simulate this model.

When the homogeneous earth contains low resistivity anomalous body, the diffusion of electric field is distorted due to the attraction of the low resistivity anomalous body. The electric field contour near the low resistivity abnormal body becomes denser and the gradient becomes larger. The anomalous body has an aggregation effect on the electric field, and the diffusion velocity of the induced eddy current becomes slow. Figure 6 is the electric field section under the circular low resistivity body. It can be seen that the electric field near the low resistivity anomaly body distorts, we can judge the position of the circular anomaly body by the position of the electric field distortion. Figure 7 is the comparison of the electric field profile and the homogeneous half space , as can be seen from the figure, the electric field profile is no longer symmetrical, a large separation in the anomalous body position are engendered. So the GREATEM has a good ability to distinguish the perfect conductor, the modeling also provides a theoretical basis for the GREATEM to detect the metal ore and water geological structures.

Conclusions

In this paper, the 2-D numerical simulation of GREATEM is studied based on the theory of electromagnetic exploration and meshfree method, the principle and key techniques of the meshfree collocation method for the numerical simulation are presented. The main conclusions are as follows:

1. Radial basis functions have good fitting characteristics. The meshfree collocation method based on RBF to discrete control equations are simple and direct, it is easy to simulate the complex models and make up for the deficiency of grid-based methods to some extent. It will become a new geophysical numerical simulation method;
2. The homogeneous half space and the low resistivity anomaly model are respectively calculated , the results show that the accuracy of the meshfree collocation method is higher than that of the finite difference method under the same conditions. GREATEM is more sensitive to the low resistivity anomalies, it has a good ability to distinguish the perfect conductor and provides a basis for the data processing and inversion.
3. Radial basis function is a function of distance, which has the characteristics of simple form and isotropy. It is easy to extend to the analysis of high dimensional problems. The application of the meshfree method to the numerical simulation of 3D electromagnetic exploration will be the next step in our research.

References

[1] Ito H, Mogi T, Jomori A, Yuuki Y, Kiho K, Kaieda H, Suzuki K, Tsukuda K, Allah SA. 2011. Further investigations of underground resistivity structures in coastal areas using grounded-source airborne electromagnetics.*Earth Planets Space*, **63:** e9-e12.

[2] Allah SA, Mogi T, Ito H, Jymori A, Yukki Y, Fomenko E, Kiho K, Kaieda H, Suzuki K, Tsukuda K. 2013 Three-dimensional resistivity characterization of a coastal area: Application of Grounded Electrical-Source Airborne Transient Electromagnetic (GREATEM) survey data from Kujukuri Beach, Japan. *Journal of Applied Geophysics*, **99**:1-11.

[3] Ji YJ. Li DS. Yu MM. Wang Y. Wu Q. Lin J. 2016 .A de-noising algorithm based on wavelet threshold-exponential adaptive window width-fitting for ground electrical source airborne transient electromagnetic signal. *Journal of Applied Geophysics* , **128:**1-7.

[4] Dai QW, Feng DS, Wang HH, Chen DP.2014. Element free method forward modeling for GPR based on improved sarma absorbing boundary. *Journal of Environmental & Engineering Geophysics*. **19(4):**196-199

[5] J Wittke， B Tezkan.2014. Meshfree magnetotelluric modelling. *Geophysical Journal International*. **198(2):**1255-1268.

[6] JI Yan-Ju, HUANG Ting-Zhe, HUANG Wan-Yu et al .2016. 2D anisotropic magnetotelluric numerical simulation using meshfree method under undulating terrain.*Chinese Journal Geophysics*,**59(12):** 4483-4493.

[7] Lai S J, Wang B Z, Duan Y.2008. Meshless Radial Basis Function Method for Transient Electromagnetic Computations. *IEEE Transactions on Magnetics.* **44(10)**:2288-2295.

[8] Kansa EJ. Multiquadrics—A scattered data approximation scheme with applications to computational fluid-dynamics—I surface approximations and partial derivative estimates.1990. *Computers & Mathematics with Applications.* **19(8–9)**:127-145.

[9] Cheng HD, Golberg MA， Kansa EJ， Zammito G. Exponential convergence and H-c multiquadric collocation method for partial differential equations.2003. *Numerical Methods for Partial Differential Equations.* **19(19):**571-594.

[10] Oristaglio M L. 1982.Diffusion of electromagnetic fields into the eath from a line source of current. Geophysics, **47(11)：** 1585-1592.

[11] Wang T, Hohmann GW.A finite-difference, time-domain solution for 3-dimensional electromagnetic modeling.1993. **58(6):**797-809.

Large scale traffic evacuation simulation based on multi-agent modeling

Liu Yi, Yu Shuiping, Qian Jing, Yuan Shengcheng*

Institute of Public Safety Research (IPSR), Department of Engineering Physics,

Tsinghua University, Beijing, P. R. China

* Corresponding author: yuancc06@qq.com

Abstract

Traffic evacuation is one of the most important issues in the area of emergency management, and the selection of evacuation exit is the key to improve evacuation efficiency. Due to great difficulties in field experiments of emergency traffic evacuation, this paper presents a new type of traffic evacuation simulation system which could analyze the impact of the number and the location of evacuation exits and population density on evacuation time, and decision-making support can be obtained on the basis of it.

Keywords: Traffic evacuation simulation system, evacuation exit, population density, decision-making support

Introduction

Traffic evacuation is the main way of large-scale emergency evacuation, and is an important part of emergency response. Experts and scholars in the field have constructed the simulation methods of different scenarios by the abstraction, integration and dynamic coupling of various traffic simulation models.

In earlier research, the evacuation is considered a special traffic event which involves completely different driving behaviors and traffic management comparing with the normal way. On this account, a class for a certain type of disaster and evacuation simulation software emerged, such as NETVAC[1] was developed for evacuation of nuclear leakage accident, and MASSVAC[2] was for emergency evacuation of hurricane in city.

In recent years, with the mature and widely used ITS (Intelligent Traffic System) technology, the evacuation research tends to use simulation software based on ITS. There existing some popular such software, such as Paramics[3], CORSIM[4], MATSim[5], Integration[6], etc. ITS based traffic simulation software provides a lot of convenience for the evacuation simulation research, otherwise recent studies have shown that those software can't simulate emergency evacuation scenarios very well, and there are two reasons for this phenomenon.

There exists too many conditions and default information to evacuation simulation, for that most of the evacuees are likely to temporarily adjust the evacuation route.

The existing normal simulation software concerns only with traffic distribution and traffic control measures, while ignoring important contents such as background information of evacuation.

The traffic evacuation simulation system proposed in this paper is developed by Tsinghua University, and the accuracy and effectiveness of the system has been tested [7]. The system is on the analysis of microscopic evacuation model, characterizes heterogeneity based on dynamic parameters and models, establishes the microcosmic simulation models with heterogeneity and the numerical calculation models, and deduces the theoretical calculation errors, finally realizes the dynamic control of the calculation errors.

Agent and driving behavior models

Transportation is a complex system. Due to the driving variability, much attention has been paid on the simulation in normal situations and focus on various driving behaviors. The microscopic evacuation models used in the system are shown in Table 1. Car-following model consists of Gipps' model [8], Optimal Velocity Model (OVM) [9], Tampere model [10] and Intelligent Driver Model (IDM) [11]. Lane-changing models like MOBIL [12], describing when and how drivers change their lanes, were usually used with car-following models together. Intersection model is made up of signal light model [13], and models like Doniec's model (Doniec) [14] expanded the scope of car-following models from roads to crowded intersections. The path selection model uses the shortest path model [15] based on the A* algorithm, and also the model based on the potential energy network (simulated navigation equipment) [16]. The analysis model uses logit discrete choice model to make decision of departure time and destination [17][18]. The system also developed two simplified psychological cognitive models on the basis of the psychological model proposed by Spielberger [19] and Helbing [20].

The study uses a fuzzy value from 0 to 1 to measure the nervousness of an agent and use it to change other agent's parameters. It needs an external function to change the nervousness value. This paper assumes that it follows the logistic differential function. The value increases if the agent moves slowly and decreases if it drives fast.

$$\frac{dn_i(t)}{dt} = (2\eta_i(t) - 1)\frac{1}{t_r}\ n_i(t)\ \left(1 - n_i(t)\right),$$

$$n_i(0) = n_i^{(0)} \tag{1}$$

$$\eta_i(t) = \begin{cases} 0, & v_i(t) \geq V_{low}, \\ 1, & otherwise. \end{cases}$$

where t is the simulation time, n is the nervousness value of agent i, $n_i(0)$ is the initial value of n_i, t_r is the nervousness reaction time which refers to the time of an agent gaining its nervousness from 0.5 to $1/(1+e^{-1}) \approx 0.731$, Vlow is the threshold of speed in which an agent thinks it drives slowly, and $\eta_i(t)$ is the event that agent i drives in low speed. The system records the maximum nervousness value of each agent during simulation. An agent whose nervousness value is greater than 0.8 is regarded as a "panic" agent which will be focused in the experiment.

$$n_M(i) = \max_t n_i(t) \tag{2}$$

$$\xi(i) = \begin{cases} 1, & n_M(i) \geq 0.8 \\ 0, & otherwise. \end{cases} \tag{3}$$

where t is the simulation time, i is the agent's id number, $n_M(i)$ is the maximum nervousness value of agent i, and $\xi(i)$ is the panic event that agent i has been panic during the simulation.

An agent follows different behavior under anxiety or not, for example, people have a higher probability to run red lights and other herd behavior. The corresponding relationship between anxiety state and behavior pattern is shown in Table 2, and the corresponding relationship with the agent attribute is shown in Table 3.

Table1 Model library used in this system

Category	Model Name	Description	Literature
Car Following	Gipps Model	Driving model for safety	[8]
	OVM Model	Driving model for speed	[9]
	Tampere Model	Driving model for stability	[10]
	IDM Model	Driving model for smart	[11]

Lane Change	MOBIL Model	Deterministic lane changing model	[12]
	Lv Model	Probabilistic lane changing model	[21]
Intersection	Doniec Model	Intersection model with no signal	[14]
	Feng Model	Intersection model with signal	[13]
Path Selection	Shortest Path Model	Route selection model based on shortest distance	[15]
	Minimum Potential Energy Model	Route selection model with minimum driving time	[16]
Requirement analysis	Departure Logit Model	Discrete choice model for determining departure time	[17]
	Destination Logit Model	Discrete choice model for determining destination	[18]
Psychological model	Nervousness Model	A model for simulating the anxiety degree of evacuees	[19]
	Conformity Model	Herd behavior model of evacuees	

Table 2 The relationship between anxiety state and behavior pattern

Category	Non Anxiety	Anxiety
Car Following	Conservative Parameters	Radical parameters
Lane Change	MOBIL Model	Lv Model
Path Selection	Minimum Potential Energy Model	Shortest Path Model
Conformity Behavior	Low Probability	High Probability

Table 3 Relationship between anxiety state and agent attribute

Parameter	Non Anxiety $n_i = 0$	Anxiety $n_i = 1$
Maximum Acceleration (m/s^2)	2.0	4.0
Maximum Deceleration (m/s^2)	-2.8	-6.0
Maximum Deceleration of Front Vehicle (m/s^2)	-2.8	-6.0
Static Following Distance（m）	7	4
Lane Change（°）	10	20

The attributes of agent in anxiety state and non-anxiety state change in the linear relationship.

$$p = (1 - n_i)p_0 + n_i p_1$$

Where p is the agent's behavioral parameters, p_0 means the value in normal state and p_1 means the value in panic state.

Case Study

The evacuation time is of great significance in traffic evacuation. In this paper, a specific experiment is designed to study the relationship between the evacuation exits and evacuation time.

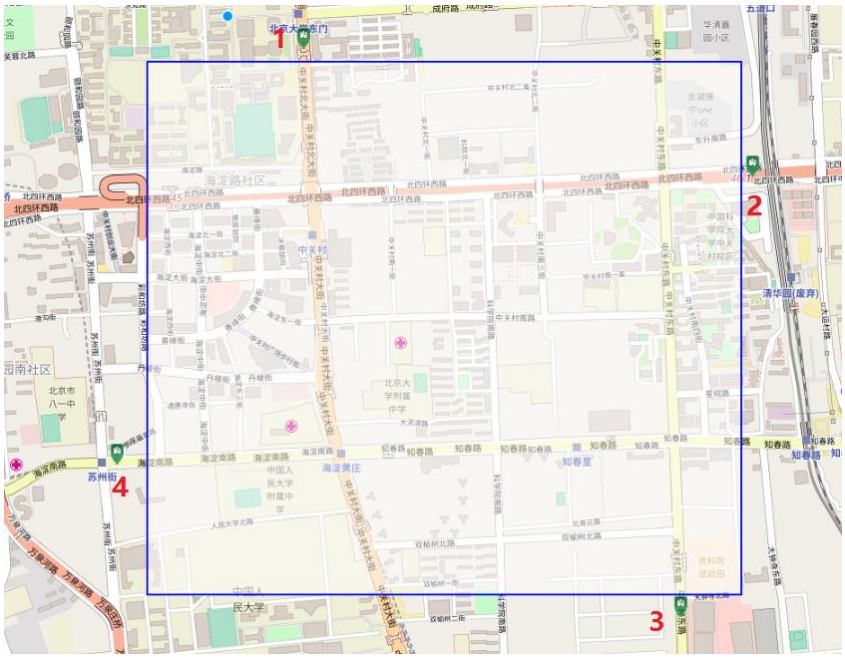


Figure 1 Emergency evacuation area

As shown in Figure 1, the experimental area is a busy area in Beijing, and simulation area is of 5.2694 square kilometers, including 4 exits. The study carried out five groups of experiments, designed as shown in Table 4, each group involves 9 tests through the establishment of different initial evacuation of the population.

Table 4 Experimental design

Number	Exit amount	Exit number	evacuation population
Group 1	1	1	Initial 5000 persons，each experiment increasing 2500，up to 25000.
Group 2	2	1&2	
Group 3	2	1&3	
Group 4	3	1，2&3	
Group 5	4	1，2，3&4	

Evacuation simulation results

Table 5 lists the linear fitting parameters of the evacuation time data after 5 groups of experiments which each of them consist 9 initial settings.

Table 5 The linear fitting parameters of the evacuation time data

Number	p_1*x+p_0 [p_1;p_0]	95% confidence interval	correlation coefficient	standard deviation
Group 1	0.269 10.01	(0.2615,0.2766) (-13.51,33.53)	0.9990	11.7787
Group 2	0.08785 52.47	(0.08314,0.09256) (37.8,67.14)	0.9964	7.3443

Group 3	0.1078 34.9	(0.1054,0.1102) (27.34,42.45)	0.9994	3.7883
Group 4	0.05429 52.8	(0.04871,0.05988) (35.39,70.2)	0.9869	8.7429
Group 5	0.04889 27.81	(0.0399,0.05787) (-0.213,55.84)	0.9594	14.0001

According to the data in table 5, the correlation coefficient of each experiment is relatively high, the correlation degree is 0.9990, 0.9964, 0.9994, 0.9869, and 0.9594, respectively. It can be found there existing a high linear correlation between evacuation time and population density. The more exits, the more random selection of the crowd, which leads to the decrease of the evacuation time correlation and improvement of data dispersion.

Figure 2 shows the relationship between population density and evacuation time, the horizontal axis represents the population density (per square kilometer), the vertical axis represents the evacuation time (minutes). Each increase of 500 people per square kilometer, the average change of evacuation time corresponding to each group is increasing 269 minutes, 88 minutes, 108 minutes, 54 minutes and 49 minutes (Fig. 3). The second group and third group of experiments show that under the same condition of evacuation, the diagonal evacuation exits (such as the export of 1 and 3) performs much more better than side exits (such as the export of 1 and 2) , it can be proved by experimental evacuation time (108 minutes >88 minutes). Figure 4 shows the ratio of evacuation time between multiple exit and single exit at the same population density, the ratio is 0.4044, 0.4505, 0.2821 and 0.2239, respectively. Multi group data show that the result of the third experiment group is worse than second groups, and it can be known by Figure 1, the density of the roads and buildings around exit 1 and exit 2 is high, there needs to evacuate more population.

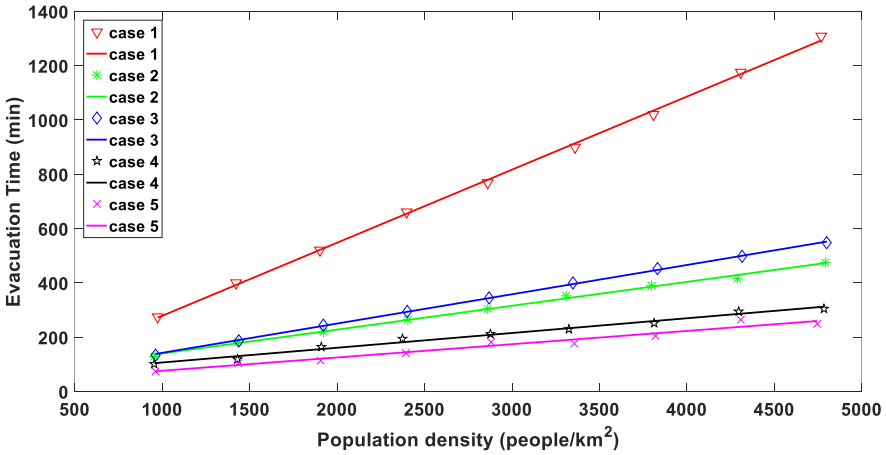


Figure 2. The relationship between population density and evacuation simulation time

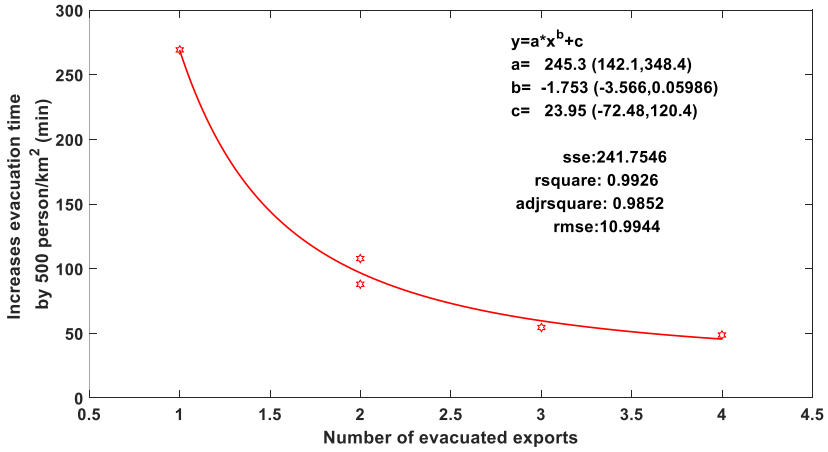


Figure 3 Population density increased by 500 people per square kilometer, the changes of evacuation time with different amount of exits

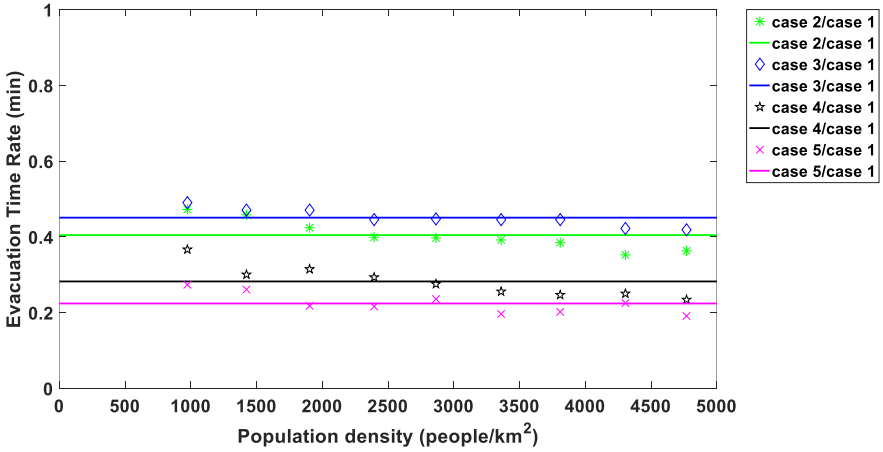


Figure 4. The ratio of evacuation time between multiple exit and single exit

Conclusions

This paper introduces a new type of traffic evacuation system, emphatically discusses the influence of the exit amount, location and population density on evacuation time, and the results can be used to traffic evacuation decision support.

In this paper, five groups of evacuation simulation experiments under different initial conditions are designed, and the experimental data are analyzed in detail. The main results are as follows:

(1) There is a highly positive linear correlation between evacuation time and evacuation population density. The evacuation time is related to the evacuation exit and nonlinear.

(2) Evacuation exits should be selected in areas with high population density.

In the future, we will consider additional factors (e.g., different types of vehicles, pedestrians, or bicycles) in the simulations for better understanding of evacuation-related decisions.

Acknowledgement

Funded by National Natural Science Foundation of China (No.71673158, No.91646101, No.91324022, and No.91646201) and Project of National Key Technologies R&D Program of China (No. 2015BAK12B01).

References

[1] Sheffi Y, Mahmassani H, Powell W B. (1982) A transportation network evacuation mode[J], *Transportation Research Part A Policy & Practice* **16**, 209-218.

[2] Hobeika A G, Jamei B. (1985) MASSVAC: A Model for Calculating Evacuation Times under Natural Disasters[J]. *Emergency Planning*: 23-28.

[3] Cova T J, Johnson J P. (2002) Microsimulation of Neighborhood Evacuations in the Urban-Wildland Interface[J]. *Environment & Planning A* **34**, 2211-2229.

[4] Williams B M, Tagliaferri A P, Meinhold S S, et al. (2007) Simulation and Analysis of Freeway Lane Reversal for Coastal Hurricane Evacuation[J]. *Journal of Urban Planning & Development* **133**, 61-72.

[5] Lammel G, Grether D, Kai N. (2003) The representation and implementation of time-dependent inundation in large-scale microscopic evacuation simulations[J]. *Transportation Research Part C Emerging Technologies* **18**, 84-98.

[6] Mitchell S W, Radwan E. (2006) Heuristic prioritization of emergency evacuation staging to reduce clearance time[C]// *Proceedings of the 85th Annual Meeting Transportation Research Board*, Washington, DC, USA.

[7] Shengcheng Y. (2016) Study on Simulation Method of Traffic Evacuation for Emergency Decision Support. [D]. *Tsinghua University*.

[8] Gipps P G. (1981) A behavioural car-following model for computer simulation[J]. *Transportation Research Part B Methodological* **15**, 105-111.
[9] Bando M, Hasebe K, Nakayama A, et al. (1995) Dynamical model of traffic congestion and numerical simulation [J]. *Physical Review E Statistical Physics Plasmas Fluids & Related Interdisciplinary Topics* **51**, 1035-1042.
[10] Tampere C M J. (2004) Human-kinetic multiclass traffic flow theory and modelling: with application to advanced driver assistance systems in congestion[D]. *TRAIL Thesis Series*.
[11] Treiber M, Hennecke A, Helbing D. (2000) Congested traffic states in empirical observations and microscopic simulations[J]. *Physical Review E Statistical Physics Plasmas Fluids & Related Interdisciplinary Topics* **62**, 1805-1824
[12] Kesting A, Treiber M, Helbing D. (2007) General Lane-Changing Model MOBIL for Car-Following Models[J]. *Transportation Research Record* **1999**, 86-94.
[13] Weidong F，Lihua Y. (1999) Simulation of Signalized Intersection Based on Automatic Cell Model [J]. *Application Research of Computers* **7**, 47-48
[14] Doniec A, Mandiau R, Piechowiak S, et al. (2008) A behavioral multi-agent model for road traffic simulation[J]. *Engineering Applications of Artificial Intelligence* **21**, 1443-1454.
[15] Hart P E, Nilsson N J, Raphael B. (1968) A Formal Basis for the Heuristic Determination of Minimum Cost Paths[J]. *IEEE Transactions on Systems Science & Cybernetics* **4**, 100-107.
[16] Dommety G, Jain R. (1998) Potential networking applications of global positioning systems (GPS). arXiv preprint cs/9809079.
[17] Hasan S, Ukkusuri S, Gladwin H, et al. (2011) Behavioral Model to Understand Household-Level Hurricane Evacuation Decision Making[J]. *Journal of Transportation Engineering* **137**, 341-348.
[18] Mesa-Arango R, Hasan S, Ukkusuri S V, et al. (2013) A Household-Level Model for Hurricane Evacuation Destination Type Choice Using Hurricane Ivan Data[J]. *Natural Hazards Review* **14**,11-20
[19] Spielberger C D, Jacobs G, Russell S. (1983) Assessment of Anger: The State-Trait Anger Scale[J]. *Advances in Personality Assessment* **2**, 159-187.
[20] Helbing D, Farkas I J, Molnar P, et al. (2002) Simulation of pedestrian crowds in normal and evacuation situations[M]// *Pedestrian and Evacuation Dynamics*, 21-58.
[21] Lv W, Song W G, Liu X D, et al. (2013) A microscopic lane changing process model for multilane traffic[J]. *Physica A Statistical Mechanics & Its Applications* **392**, 1142-1152.

Interval field model and interval finite element analysis

†*C. Jiang[1], B.Y. Ni[1]

[1] State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,
College of Mechanical and Vehicle Engineering, Hunan University, Changsha City, P. R. China.

*Presenting author: jiangc@hnu.edu.cn
†Corresponding author: jiangc@hnu.edu.cn

Abstract

Uncertain parameters with inherent spatial variability are commonly encountered in engineering. Modeling of this kind of spatial uncertainty plays a fundamental role in structural uncertainty analysis, which provides a necessary basis for subsequent uncertainty propagation through the system. An interval field model for quantification of spatial uncertain parameters is proposed, by which only the upper and lower bounds of the spatial uncertain parameters rather than their precise probability distributions are required. The dependency can be fully considered by the proposed interval field model. With the information of dependency, an interval K-L expansion is presented as a combination of deterministic functions with uncorrelated standard interval variables, through which the continuous spatial interval field with dependency can be expressed only by very limited intervals. Necessary mathematical illustrations are provided for the proposed interval field model and the interval K-L expansion. The sampling method for the interval field model is given, providing a robust numerical analysis basis for subsequent structural uncertainty analysis. When the interval field model is applied in finite element analysis of structures with spatial uncertain parameters, the non-deterministic equilibrium equations with interval factors is then formulated. The MCS method and the perturbation method are developed for solution of the derived interval equilibrium equations. The feasibility and validity of the proposed interval field model and corresponding interval finite element methods are verified by numerical examples, where the upper and lower bounds of the responses such as the displacements and the stresses of structures with spatial uncertain parameters are computed and compared.

Keywords: Interval field model; Spatial uncertainty; Interval K-L expansion; Interval finite element method

1. Introduction

The modeling of uncertain input parameters with inherent spatial variability are commonly encountered in engineering. These include material properties of the heterogeneous media such as concrete or porous rock, geographical parameters such as soil permeability over the scale of meters. This kind of uncertainties generally present with spatially varying properties, which traditionally can be well quantified by random field models [1]. And solutions of the stochastic problems where the properties of the structures are modeled as random fields can be found by stochastic finite element method (SFEM) [2]. As a primary non-deterministic framework, the probabilistic methods have been tremendously developed over the last decades. However, a large amount of information is required to determine the credible probability density function (PDF) for construction of the probabilistic model, which is quite impractical or very costly to obtain in many engineering problems.

In this work, an interval field model is proposed for quantification of a spatially uncertain parameter, which requires only the upper and lower bounds of the parameter rather than its precise probability distributions. An interval field can be denoted as $\left\{H(\mathbf{x}) \in H^I(\mathbf{x}), \mathbf{x} \in D\right\}$, where D refers to spatial domain. In Fig. 1, an interval field with constant upper and lower

bounds is given. For arbitrary location $\mathbf{x}_k$ in the two-dimensional domain, the variation range of the variable $H(\mathbf{x}_k)$ is strictly limited within the interval $H^I(\mathbf{x}_k)=[H^L(\mathbf{x}_k),H^U(\mathbf{x}_k)]$. In many practical circumstances, although the values of a spatial uncertain parameter differ with location, dependency exists between these spatial uncertainties especially for those adjacent ones. For example, the material property such as the elasticity modulus of a concrete structure may be spatially uncertain because of its inhomogeneity, however the value of the elasticity modulus at arbitrary location is likely to be close to that at another location nearby. For this reason, the covariance function $C(\mathbf{x},\mathbf{x}')$ and the correlation coefficient function $R(\mathbf{x},\mathbf{x}'),\mathbf{x},\mathbf{x}'\in D$ are also defined to reflect the dependency degree of the interval field at different locations.

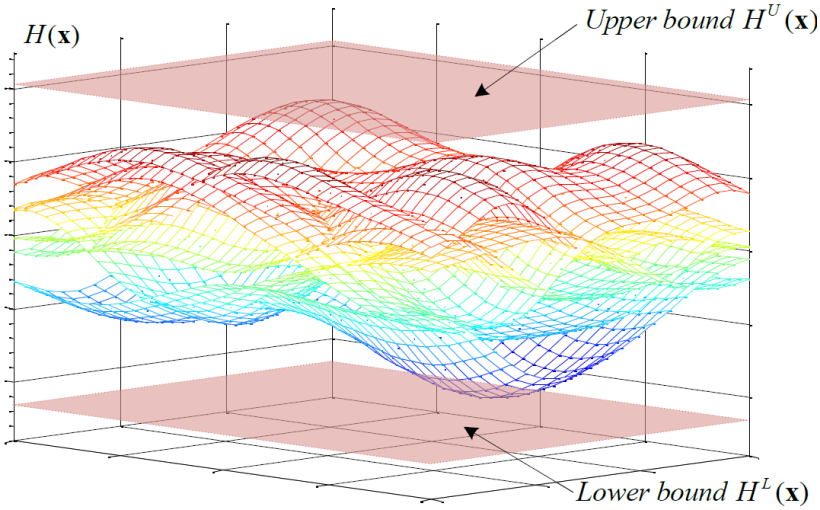


Fig.1. Interval field model

Similar to the Karhunen-Loève (K-L) expansion of a random field [2], the interval K-L expansion is created to represent an interval field as an infinite linear combination of orthogonal functions multiplied with uncorrelated standard interval variables, which can be expressed as:

$$H(\mathbf{x})=H^c(\mathbf{x})+\sum_{j=1}^{\infty}H^r(\mathbf{x})\sqrt{\lambda_j}\varphi_j(\mathbf{x})\zeta_j \tag{1}$$

where $H^c(\mathbf{x})$ and $H^r(\mathbf{x})$ are the midpoint function and the radius function of the interval field, respectively; $\zeta_j\in\zeta^I=[-1,1],\ j=1,2,\ldots$ are standard uncorrelated interval variables that satisfy $\sum_{j=1}^{\infty}\zeta_j^2\le 1$; and $\lambda_j\in[0,\infty)$, $\varphi_j(\mathbf{x}):D\to\mathbb{R}$ are respectively the eigenvalues and eigenfunctions of the correlation coefficient function $R(\mathbf{x},\mathbf{x}')$. Corresponding mathematical foundations for this interval field model and the interval K-L expansion are also established. In practical engineering, it is not only impossible but also unnecessary to use infinite terms with interval variables for quantification of the spatial uncertainty. Generally, most of the characteristics of a spatially uncertain parameter can be reflected considerably by those principle terms. Therefore for practical implementation, the series is generally approximated by sorting the eigenvalues λ_i and the corresponding eigenfunctions $\varphi_j(\mathbf{x})$ in a descending order and truncating the expansion after *M* terms. The error analysis of the truncated form in the representation of a spatially uncertain parameter is also given, from which an index that evaluates the degree of approximation is suggested.

When the interval field model is applied to the finite element analysis of structures with

spatial uncertain parameters such as material properties and distributed loads, the derivation of the interval finite element method (IFEM) [3, 4] is then formulated. According to the sources of spatial uncertainty $\alpha(\mathbf{x})$ in the finite element system, the non-deterministic equilibrium equation can be classified into the following three types:

$$\begin{aligned} &\text{Type-I} \quad \mathbf{K}(\alpha)\mathbf{u}=\mathbf{p} \\ &\text{Type-II} \quad \mathbf{K}\mathbf{u}=\mathbf{p}(\alpha) \\ &\text{Type-III} \quad \mathbf{K}(\alpha)\mathbf{u}=\mathbf{p}(\alpha) \end{aligned} \tag{2}$$

For the existence of the spatially uncertain parameters with variation bounds, the responses such as the displacements and stresses of a structure also presents with bounded uncertainties. Both the Monte Carlo simulation (MCS) method and the perturbation method are developed to solve the interval equilibrium equations derived by the IFEM. The procedure of the MCS method is given in Fig. 2, which provides a robust numerical analysis framework and can be used as a standard reference for other numerical solutions.

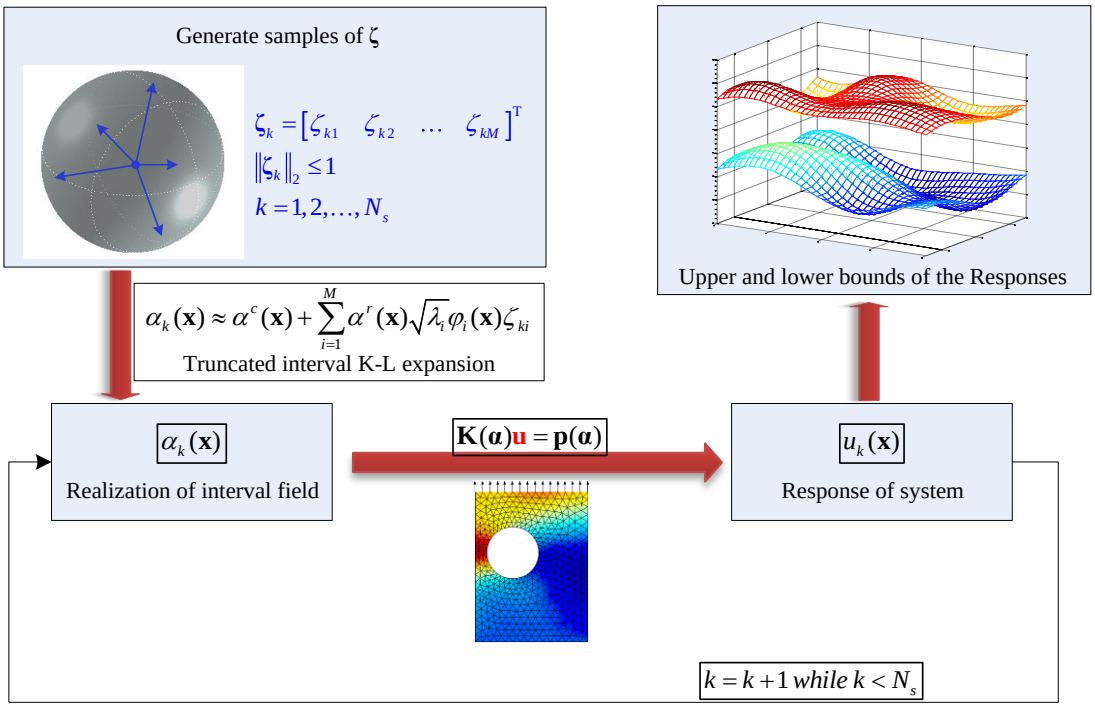


Fig. 2. MCS method for evaluation of response bounds

2. Results and discussions

Displacement analysis of a concrete quadrate plate subjected to distributed forces at the two sides is implemented. Due to the inhomogeneity and the spatial uncertainty of the material properties of the concrete plate, the Young's modulus is described as an interval field $E(\mathbf{x})\in E^I(\mathbf{x})$ with constant midpoint function $E^c(\mathbf{x})=32.5\text{GPa}$ and radius function $E^r(\mathbf{x})=10\%E^c(\mathbf{x})$. The correlation coefficient function $R(\mathbf{x},\mathbf{x}')$ is given as exponential form. With the truncated interval K-L expansion, the continuous spatial uncertain Young's modulus $E(\mathbf{x})$ is approximated only by 24 standard uncorrelated interval variables with approximation degree $\kappa=90.42\%$. The response bounds of horizontal displacements by the perturbation method are depicted in Fig. 3. The region enveloped by the upper bound and the lower bound indicates the variation domain of all possible responses under all realizations of the spatially uncertain Young's modulus $E(\mathbf{x})$. In general, the perturbation method can be regarded as an effective approach for problems with uncertainty of degree not higher than 10%. For problems with large uncertainty, the MCS method can be applied, and other more effective

methods are required to be developed in future.

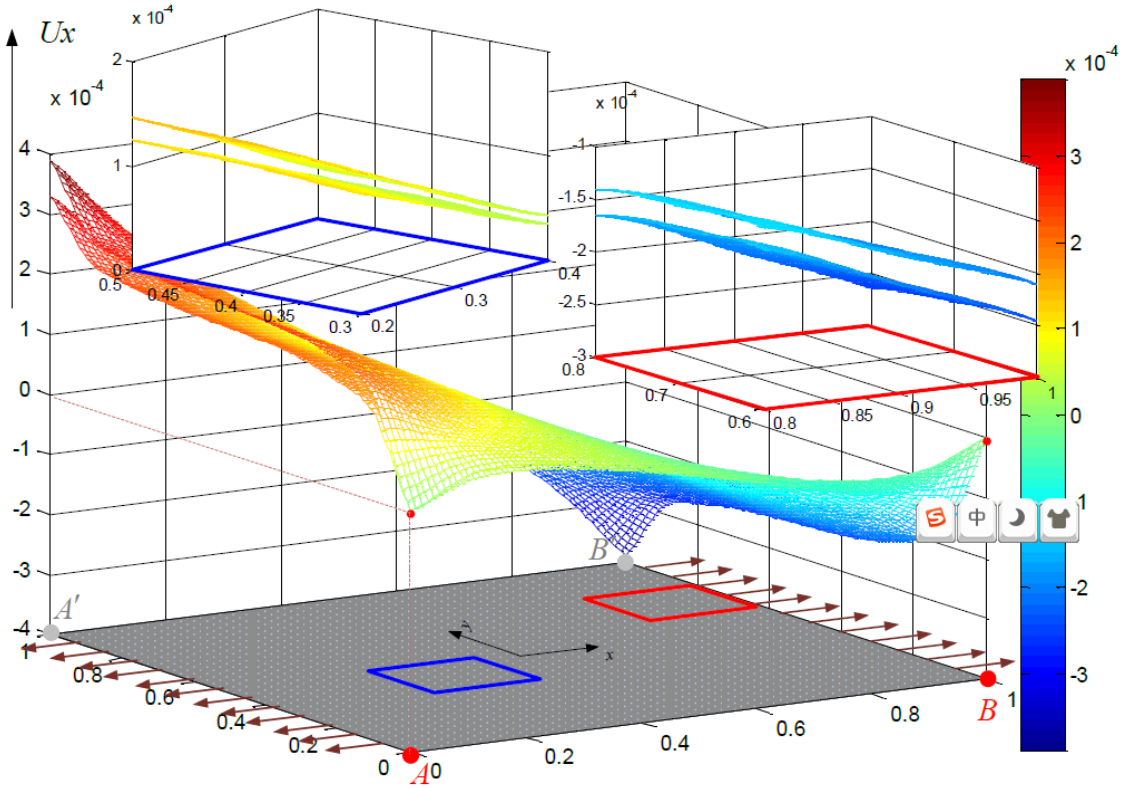


Fig. 3. Response bounds of horizontal displacement by the perturbation method

3. Conclusions

In this work, an interval field model for quantification of spatially uncertain parameters is proposed. The interval field model requires only the upper and lower bounds of the spatial uncertain parameters rather than their precise probability distributions. The dependency can be fully considered by the proposed interval field model. With the information of dependency, an interval K-L expansion is presented as a combination of deterministic functions with uncorrelated standard interval variables, through which the continuous spatial interval field with dependency can be expressed only by very limited intervals. When the interval field model is applied in finite element analysis of structures with spatial uncertain parameters, the non-deterministic equilibrium equations with interval factors is then formulated. Solutions by MCS method and the perturbation method are developed and compared in numerical examples. The MCS method is applicable to cases of large uncertainties and strong nonlinearities, but it generally costs much computational time. The perturbation method is developed based on the assumption of small uncertainty of spatial parameters; it is an efficient approximation method for structural response analysis. In numerical examples, the upper and lower bounds of structural responses such as displacements and stresses are computed, by which the feasibility and validity of the proposed interval field model and corresponding interval finite element analysis method are illustrated.

References

[1] Vanmarcke, E. (1983) *Random Fields: Analysis and Synthesis*, The MIT Press, Cambridge, Massachussets.

[2] Ghanem, R. and Spanos, P.D. (1991) *Stochastic Finite Elements: A Spectral Approach*, Springer-Verlag, Berlin.

[3] Muhanna, R.L., Mullen, R.L. and Zhang, H. (2007) Interval finite elements as a basis for generalized models of uncertainty in engineering mechanics, *Reliable Computing* **13(2)**, 173–194.

[4] Moens, D. and Vandepitte, D. (2005) A survey of non-probabilistic uncertainty treatment in finite element analysis. *Computer Methods in Applied Mechanics and Engineering* **194**, 1527–1555.

On Improving Evolutionary Algorithms and Acceleration Techniques Based on Estimation of Convergence Point Population for Chosen Optimization Problems of Mechanics

Janusz Orkisz, *†Maciej Glowacki

Institute for Computational Civil Engineering, Cracow University of Technology, Poland

*Presenting author: mglowac@gmail.com
†Corresponding author: mglowac@gmail.com

Abstract

Several issues regarding development of highly accelerated and efficient Evolutionary Algorithms (EA) for solving large, non-linear, constrained optimization problems are considered in this work. In particular, we briefly present here advances in development of already proposed acceleration techniques, including smoothing and balancing, adaptive step-by-step mesh refinement, as well as a'posteriori error analysis and related techniques. Our most recent research has been focused mainly on searching of efficient combination of the proposed techniques and their parameters, as well as on development of some new concepts based on estimation of the convergence point of population. The improved EA-based approach provides significant speed-up of solution process and/or possibility of solving such large problems, when the standard EA methods fail.

Keywords: Evolutionary Algorithms, acceleration techniques, large non-linear constrained optimization problems, convergence point of population

Introduction

Many important problems of computational mechanics may be formulated in terms of constrained optimization. Complexity of these problems may result mostly from their non-linearity, as well as from a large number of decision variables and constraints involved. Thus, we consider here a wide class of large, non-linear, constrained optimization problems. Due to the size and complexity of such problems, this research is focused, first of all, on the significant efficiency increase of the solution algorithms applied. Our solution approach is based on the EA, which on the contrary to most deterministic methods may be successfully applied to the both convex and non-convex problems [1]. However, general efficiency of the standard EA is rather low. Therefore, significant acceleration of the solution process is often needed. The forthcoming engineering objective of this long-term research includes residual stresses analysis [2][6] in railroad rails, and vehicle wheels, as well as a wide class of problems resulting from the Physically Based Approximation (PBA) of experimental and/or numerical data [4].

General problem formulation and solution algorithms

In the analyzed wide class of optimization problems, a function given in the discrete form, e.g. expressed in terms of its nodal values, is sought. These nodal values are defined on a mesh formed by arbitrarily distributed nodes. The optimal solution usually has to satisfy numerous equality, and inequality constraints. To obtain discrete formulation of optimization problem, any discretization method can be applied, including Finite Element, as well as Meshless Finite Difference Methods used here.

The EA are understood here as real-value coded genetic algorithms consisting of selection, crossover, and mutation operators [1]. We have proposed and tested so far several new, simple but effective EA acceleration techniques with various variants, including solution smoothing and balancing, an adaptive step-by-step mesh refinement, as well as a'posteriori solution error analysis and related techniques [3]. Appropriate constraint handling techniques were investigated as well. Our most recent research has been focused on further development of techniques based on various variants of estimation of the convergence point of a population considered. Reference [5] introduces a general idea and a few methods for estimation of the convergence point for the moving vectors of individuals between two subsequent generations. Such convergence point indicates the neighborhood of the optimum (see Fig. 1). It presents a powerful individual for the optimization process.

Considered is a population of M individuals

$$\mathbf{u}^j = [u_1^j, u_2^j, u_3^j, ..., u_n^j] \ , \quad j = 1, 2, 3, ..., M \tag{1}$$

in a n-dimensional space. Moving vectors are calculated between individuals $\mathbf{u}^{j,k}$ from k-th generation and their offspring $\mathbf{u}^{j,k+1}$ from (k+1)-th generation. $\tilde{\mathbf{u}}^k$ is a convergence point, and $\bar{\mathbf{u}}$ is the optimum point.

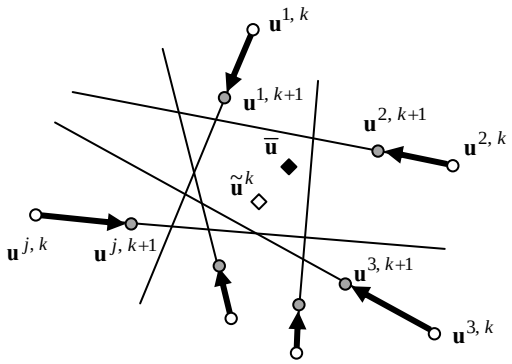


Figure 1. General idea of estimation of the convergence point of population

Various approaches, namely the exact, approximated and iterative ones are discussed in [5]. Approximated approach, which was presented as the best one, is based on the truncated Neumann series expansion. In this case, estimated convergence point is calculated in a following way:

$$\tilde{\mathbf{u}}^k \approx \tfrac{1}{M} \sum_{j=1}^{M} \mathbf{u}^{j,k} - \tfrac{1}{M} \sum_{j=1}^{M} ((\mathbf{u}^{j,k})^{\mathrm{T}} \mathbf{b}^{0\,j,k})\, \mathbf{b}^{0\,j,k} \tag{2}$$

where

$$\mathbf{b}^{j,k} = \mathbf{u}^{j,k+1} - \mathbf{u}^{j,k} \ , \text{ and } \ \mathbf{b}^{0\,j,k} = \mathbf{b}^{j,k} / \| \mathbf{b}^{j,k} \| \tag{3}$$

All general approaches presented in [5] can be applied to almost any population-based computations. We have proposed and preliminarily evaluated a specific formulation and implementation of these general approaches used for the EA acceleration. We have also

proposed several new, original concepts for estimation of the convergence point of population, which can be used alternatively for these proposed in [5]. However, they still need further evaluation and development.

On benchmark problems

The efficiency of the new algorithms was examined using various demanding benchmarks involving large number of decision variables and constraints, including residual stress analysis in chosen elastic-perfectly plastic bodies, such as thick-walled cylinder, under various cyclic loadings. These benchmarks allow to choose almost any number of decision variables involved. The largest executed numerical tests involved more than 3000 decision variables. Several inverse problems were analyzed as well, including reconstruction of residual stresses. Such analysis used experimentally measured data, and the PBA approach. For example, the following sample optimization problem given in the polar coordinates for residual stresses in the thick-walled cylinder under cyclic internal pressure was investigated [6].

Find the minimum of the total complementary energy:

$$\min_{\sigma_r^r,\sigma_t^r,\sigma_z^r} \tfrac{1}{2E} 2\pi L\int_a^b [(\sigma_r^r-\sigma_t^r)^2+(\sigma_t^r-\sigma_z^r)^2+(\sigma_z^r-\sigma_r^r)^2]\,r\,dr \tag{4}$$

subject to the equilibrium equation

$$\tfrac{\partial \sigma_r^r}{\partial r}+\tfrac{\sigma_r^r-\sigma_t^r}{r}=0 \tag{5}$$

boundary conditions

$$\sigma_{r|a}^r=0, \qquad \sigma_{r|b}^r=0 \tag{6}$$

the incompressibility equation

$$\sigma_z^r=v(\sigma_r^r+\sigma_t^r) \tag{7}$$

and the yield condition

$$\phi(\sigma_r^r,\sigma_t^r,\sigma_z^r,\sigma^E)\le\sigma_Y \tag{8}$$

where $\sigma_r^r,\sigma_t^r,\sigma_z^r$ are respectively the radial, circumferential and longitudinal stresses, $\sigma^E=\{\sigma_r^E,\sigma_t^E,\sigma_z^E\}$ is the purely elastic solution of the problem, σ_Y is the yield stress, a,b are respectively the internal and external radii, L is the cylinder length, and E is the Young modulus.

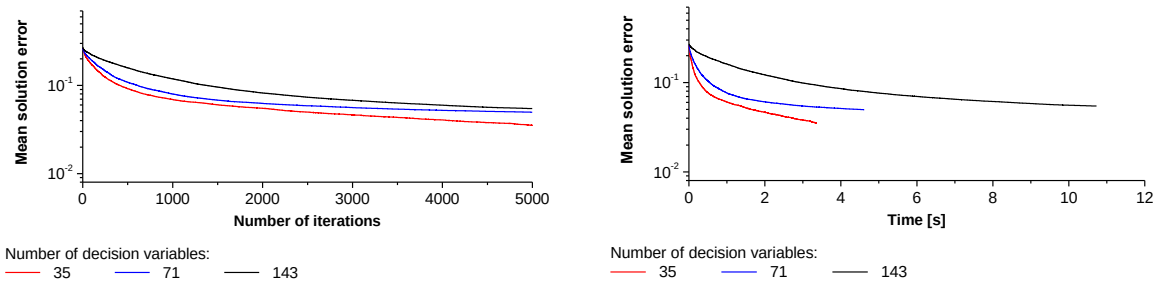
Sample of numerical results

In Fig. 2-5 one may find typical numerical results obtained for our efficiency analysis. They present convergence of mean solution error for residual stress analysis in cyclically pressurized thick-walled cylinder used as a benchmark problem. Due to stochastic nature of evolutionary computation, all results shown here were averaged over 20 independent solution processes.

Fig. 2 shows results obtained for the standard EA approach, consisting of selection, crossover, and mutation operators only, without any additional acceleration techniques. Results for three different number of decision variables are presented. Each decision variable corresponds to one nodal value of residual stresses searched.

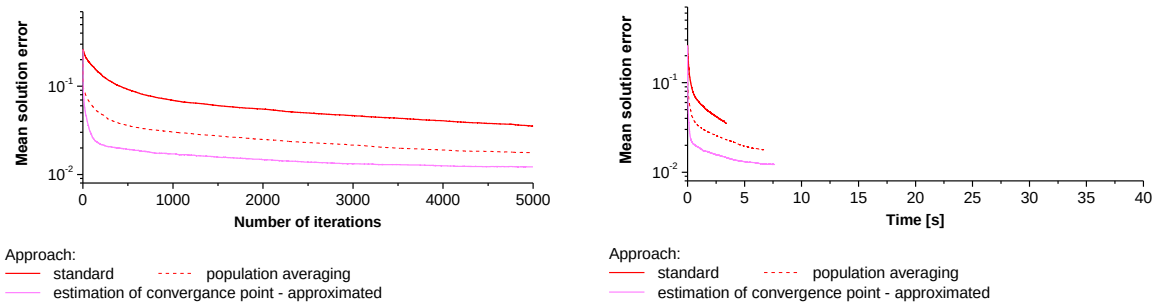
Comparison of accelerated EA is shown in Fig. 3-5. Approach based on approximated estimation of convergence point of population is compared to the standard EA, as well as to EA using simple averaging of population.

All optimization processes were calculated for the same number of generations (iterations) of EA – see figures (a). On the other hand, in figures (b) one may find time of computation needed to process these iterations.



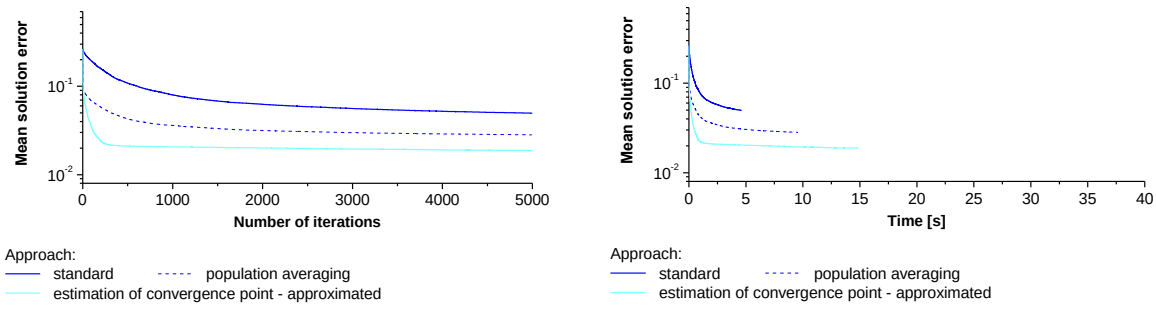
(a) (b)

Figure 2. Efficiency analysis of the standard EA for different number of decision variables



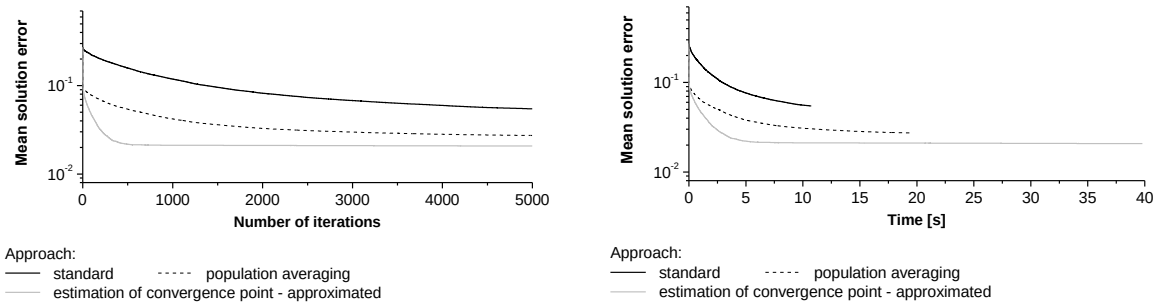
(a) (b)

Figure 3. Comparison of accelerated EA for 35 decision variables



(a) (b)

Figure 4. Comparison of accelerated EA for 71 decision variables



(a) (b)

Figure 5. Comparison of accelerated EA for 143 decision variables

Techniques based on estimation of the convergence point of population allowed to obtain acceleration up to about 40 times. It is still less than in the case of our earlier approach based on step-by-step mesh refinement combined with smoothing and a'posteriori error analysis (about 140 times), but this methods may be still improved.

In numerical results presented, approach based on approximated estimation of convergence point is better than EA using simple averaging of population. However, this method was not so efficient in all benchmark problems considered. Thus, averaging of population should also be taken into account as one of possible acceleration techniques worth applying.

Final remarks

Numerical results obtained indicate possibility of practical application of the improved EA to real complex optimization problems involving large number of decision variables and constraints. Numerical analysis also shows possibilities of further development of speed-up techniques considered, e.g. by means of combining various variants.

Future research will be mostly focused on application of the improved EA to engineering problems of mechanics.

References

[1] Engelbrecht, A.P. (2007) *Computational intelligence: an introduction*, Wiley, Chichester.
[2] Hill, R. (2004) *The Mathematical Theory of Plasticity*, Oxford University Press, New York.
[3] Glowacki, M., Orkisz, J. (2015) On increasing computational efficiency of evolutionary algorithms applied to large optimization problems, *2015 IEEE Congress on Evolutionary Computation*, 2639-2646.
[4] Karmowski, W., Orkisz, J. (1993) Physically based method of enhancement of experimental data – concepts, formulation and application to identification of residual stresses, *IUATAM Tokyo 1992 Symposium on Inverse Problems in Engineering Mechanics*, Springer-Verlag, 61-70.
[5] Murata, T., et al. (2015) Analytical estimation of the convergence point of populations, *2015 IEEE Congress on Evolutionary Computation*, 2619-2624.
[6] Orkisz, J. (1992) Prediction of actual residual stresses by constrained minimization of energy, *Residual Stress in Rails*, Vol. 2, Kluwer Acad. Publisher, 101-124.

Thermoelastic analysis of FGM beam using meshless weighted least-square method

Zhou H.M.

School of Mechanical and Power Engineering, Henan Polytechnic University, JiaoZuo, China

Corresponding author: zhm_1979@163.com

Abstract

The paper analyzes the thermoelastic problem of the FGM beam using meshless weighted least-square method (MWLS). The MWLS as a meshless method is fully independent of mesh, and a discrete function is used to construct a series of linear equations, which avoided the troublesome task of numerical integration. The effectiveness and accuracy of the approach are illustrated by a clamped-clamped FGM beam which is subjected with interior heat source. The volume fraction of FGM beam is assumed to be given by a simple power law distribution. Material properties of the FGM beam are assumed to be temperature independent and calculated by Mori-Tanaka method. The results shows that a good agreement is achieved between the proposed meshless method and commercial COMSOL Multiphysics.

Keywords: thermoelastic analysis; FGM beam; Meshless weighted least-square method.

1 Introduction

FGM object can resist high temperatures and are profient in reducing the thermal stress, and have received a considerable attention from the researchers [1]. Various numerical techniques, such as the finite difference method [2], finite element method [3-4], boundary element method [5] or more recently developed meshless methods[6-11], have been developed for analyzing these thermoelastic and heat conduction problems. Because of the complexity of the relevant governing equation, analytical solutions are usually difficult to obtain for those arbitrary geometry and complex boundary conditions. Compared with FEM, FDM and BEM, the meshless methods is associated with a class of numerical techniques that approximate a given differential equation or a set of differential equations using global interpolations on the discrete nodes or background mesh, exhibiting the advantages of avoiding mesh generation, simple data preparation, easy post-processing and so on.

Liu and Gu [12] introduced meshless methods and their programming, such as the element-free Galerkin (EFG) method, the hp-clouds method, the meshless local Petrov-Galerkin (MLPG) method, meshless Galerkin method using radial basis functions, the least-square method and meshless point collocation method. The main advantage of the MLPG method [6-7] compared with regular Galerkin-based methods is that no background mesh is used to evaluate various integrals appearing in the local weak formulation of problem, but it requires a high-order quadrature rule to obtain converged results and thus needs much more computational effort in terms of CPU time than that for the FEM. Katsikadelis [10]

employed the meshless analog equation method to solve the 2D elastostatic problem for inhomogeneous anisotropic. A meshless algorithm of fundamental solution coupling with radial basis functions based on analog equation theory was proposed to simulate the static thermal stress distribution in 2D FGMs [11]. Bhavani etc.[13] solved thermoelastic equilibrium equations for a functionally graded beam to obtain the axial stress distribution. Sohn and Kim[14] analysized static and dynamic stabilities of FG panels which are subjected to thermal and aerodynamic loads.

Zhou etc. presented steady-state[15] and transient-state[16] heat conduction analysis of heterogeneous material using the meshless weighted least-square method. In this paper the pure meshless method (MWLS) was then extended to solve problems of thermoelastic analysis for the FGM beam with interior heat source. The volume fractions of constituent materials composing the FGM beam are assumed to be given by a simple power law distribution. Material properties of the FGM are obtained by Mori-Tanaka method. The paper is divided as follows: in section 2, we give problem description and MWLS analysis about the thermoelastic problem. In order to demonstrate the efficiency and accuracy of the proposed method, numerical implementation is given in section 3. The last section includes some conclusions.

2. MWLS analysis of the thermoelastic problem

The basis of MWLS analysis is described in this section. The shape functions in MWLS analysis is a moving least-squares approximation scheme which is originally developed for the smooth interpolation of irregularly distributed data.

2.1 The Moving Least-square (MLS) approximation scheme

The local approximate function of $f(\boldsymbol{x})$ is expressed as

$$f(x) \approx f^h(\boldsymbol{x}) = N_I(\boldsymbol{x}) f_I = \boldsymbol{p}^T(\boldsymbol{x})\boldsymbol{a}(\boldsymbol{x}) \tag{1}$$

Where $\boldsymbol{p}^T(\boldsymbol{x})$ is the basis function and the quadratic basis $\boldsymbol{p}^T(\boldsymbol{x}) = \{1, x, y, x^2, xy, y^2\}$ is used in this paper; $\boldsymbol{a}(\boldsymbol{x})$ is the coefficient, which is determined by minimizing a functional of weighted residual

$$J = \sum_{I=1}^{N} \omega_I(\boldsymbol{x})[f^h(\boldsymbol{x}, \boldsymbol{x}_I) - f(\boldsymbol{x}_I)]^2 = \sum_{I=1}^{N} \omega_I(\boldsymbol{x})[\boldsymbol{p}^T(\boldsymbol{x})\boldsymbol{a}(\boldsymbol{x}) - f(\boldsymbol{x}_I)]^2 \tag{2}$$

The minimum value of J may be obtained through differentiating with respect to $\boldsymbol{a}(\boldsymbol{x})$

$$\frac{\partial J}{\partial a_j(\boldsymbol{x})} = 2\sum_{I=1}^{N} \omega_I(\boldsymbol{x})[\sum_{i=1}^{m} p_i(\boldsymbol{x}_I) a_i(\boldsymbol{x}) - f_I] p_j(\boldsymbol{x}_I) = 0 \qquad j=1,2,\ldots,m \tag{3}$$

Where $\boldsymbol{x}_I$ are the positions of the N nodes, f_I is the nodal parameter of the field variable at node I. $\omega_I(\boldsymbol{x})$ is the weighting function and usually a compactly supported function that is

only nonzero in a small neighborhood called the "support domain" of node $\boldsymbol{x}_I$. The gauss function is used in this paper.

$$\omega(r)=\begin{cases}(\exp(-r^2\beta^2)-\exp(-\beta^2))/(1-\exp(-\beta^2)) & r\le 1\\ 0 & r>1\end{cases} \qquad (4)$$
$$r=\left\|\boldsymbol{x}-\boldsymbol{x}_I\right\|/d_{mI}$$

Solving $\boldsymbol{N}(\boldsymbol{x})$ from Eq.(1) and Eq.(3), the shape function is given by:

$$\boldsymbol{N}(\boldsymbol{x})=\boldsymbol{p}^T(\boldsymbol{x})\boldsymbol{A}^{-1}(\boldsymbol{x})\boldsymbol{B}(\boldsymbol{x}) \qquad (5)$$

Where the matrices $\boldsymbol{A}(\boldsymbol{x})$ and $\boldsymbol{B}(\boldsymbol{x})$ are defined as

$$\begin{cases}\boldsymbol{A}(\boldsymbol{x})\boldsymbol{a}(\boldsymbol{x})=\boldsymbol{B}(\boldsymbol{x})f\\ \boldsymbol{A}(\boldsymbol{x})=\sum_{I=1}^{N}\omega_I(\boldsymbol{x})\boldsymbol{p}(\boldsymbol{x}_I)\boldsymbol{p}^T(\boldsymbol{x}_I)\\ \boldsymbol{B}(\boldsymbol{x})=[\omega_1(\boldsymbol{x})p(\boldsymbol{x}_1)\quad \omega_2(\boldsymbol{x})p(\boldsymbol{x}_2)\quad \ldots\quad \omega_N(\boldsymbol{x})p(\boldsymbol{x}_N)]\end{cases} \qquad (6)$$

In Eq.(4) the radius of the circular support domain, d_{mI}, is chosen to make the matrix $\boldsymbol{A}(\boldsymbol{x})$ nonsingular everywhere in the domain, i.e., the support domain must have enough neighborhood nodes. Through finding the kk^{th} nearest points of the evaluation point $\boldsymbol{x}$, the smallest support domain including these points can be obtained. The selection of kk is to compare some numerical examples with their analytical solution in Zhou et al.[16].

2.2 Thermoelastic analysis of FGM object

Consider the 2D FGM anisotropic linear elastic body occupying the domain Ω with boundary Γ of the xy plane. The governing equation and boundary condition are as follows,

$$\sigma_{ij,j}=0 \quad in \quad \Omega \qquad (7)$$

$$\begin{aligned}&\text{stress boundary condition: } \sigma_{ij}n_j-\overline{t}_i=0 \quad \text{on} \quad \Gamma_t\\ &\text{displacement boundary condition: } u_i=\overline{u}_i \quad \text{on} \quad \Gamma_u\end{aligned} \qquad (8)$$

where σ_{ij} is the components of the Cauchy stress tensor. A comma followed by index j denotes the partial differentiation with respect to coordinate x_j of a material point. u_i are the displacement components, and $\overline{u}_i$ are the prescribed displacements on Γ_u and $\overline{t}_i$ are the given tractions on Γ_t where Γ_u and Γ_t are the complementary parts of the boundary Γ.

$$\sigma_{ij,j}(\boldsymbol{x}_k)=0 \qquad \boldsymbol{x}_k\in\Omega, i,j=1,2; k=1,2,...,N_\Omega \qquad (9)$$
$$\sigma_{ij}(\boldsymbol{x}_k)n_j=\overline{t}_i(\boldsymbol{x}_k) \qquad \boldsymbol{x}_k\in\Gamma_t, i,j=1,2; k=1,2,...,N_t \qquad (10)$$
$$u_i(\boldsymbol{x}_k)=\overline{u}_i(\boldsymbol{x}_k) \qquad \boldsymbol{x}_k\in\Gamma_u, i=1,2; k=1,2,...,N_u \qquad (11)$$

Substituting the approximate shape function $\boldsymbol{f}$ of Eq.(9~11) into Eq.(1),

$$\sum_{I=1}^{N}\boldsymbol{H}_I(x_k)\boldsymbol{f}_I=0 \qquad x_k\in\Omega, k=1,2,...,N_\Omega \qquad (12)$$

$$\sum_{I=1}^{N} \boldsymbol{Q}_I(x_k)\boldsymbol{f}_I = \overline{\boldsymbol{t}}_k \qquad x_k \in \Gamma_t, k = 1,2,...,N_t \tag{13}$$

$$\sum_{I=1}^{N} \boldsymbol{N}_I(x_k)\boldsymbol{f}_I = \overline{\boldsymbol{u}}_k \qquad x_k \in \Gamma_u, k = 1,2,...,N_u \tag{14}$$

Where

$$\boldsymbol{H}_I(x_k) = \frac{E}{1-\upsilon^2}\begin{bmatrix} \dfrac{\partial^2 N_I(x_k)}{\partial x^2} + \dfrac{1-\upsilon}{2}\dfrac{\partial^2 N_I(x_k)}{\partial y^2} & \dfrac{1+\upsilon}{2}\dfrac{\partial^2 N_I(x_k)}{\partial x \partial y} \\ \dfrac{1+\upsilon}{2}\dfrac{\partial^2 N_I(x_k)}{\partial x \partial y} & \dfrac{\partial^2 N_I(x_k)}{\partial y^2} + \dfrac{1-\upsilon}{2}\dfrac{\partial^2 N_I(x_k)}{\partial x^2} \end{bmatrix} \tag{15}$$

$$\boldsymbol{Q}_I(x_k) = \frac{E}{1-\upsilon^2}\begin{bmatrix} l\dfrac{\partial N_I(x_k)}{\partial x} + m\dfrac{1-\upsilon}{2}\dfrac{\partial N_I(x_k)}{\partial y} & l\upsilon\dfrac{\partial N_I(x_k)}{\partial y} + m\dfrac{1-\upsilon}{2}\dfrac{\partial N_I(x_k)}{\partial x} \\ m\upsilon\dfrac{\partial N_I(x_k)}{\partial x} + l\dfrac{1-\upsilon}{2}\dfrac{\partial N_I(x_k)}{\partial y} & m\dfrac{\partial N_I(x_k)}{\partial y} + l\dfrac{1-\upsilon}{2}\dfrac{\partial N_I(x_k)}{\partial x} \end{bmatrix} \tag{16}$$

$$\boldsymbol{N}_I = \begin{bmatrix} \boldsymbol{N}_I(x_k) & 0 \\ 0 & \boldsymbol{N}_I(x_k) \end{bmatrix}, \quad \boldsymbol{f}_I = \begin{bmatrix} f_{1I} \\ f_{2I} \end{bmatrix}, \overline{\boldsymbol{t}}_k = \begin{bmatrix} \overline{t}_1(x_k) \\ \overline{t}_2(x_k) \end{bmatrix}, \quad \overline{\boldsymbol{u}}_k = \begin{bmatrix} \overline{u}_1(x_k) \\ \overline{u}_2(x_k) \end{bmatrix} \tag{17}$$

Where, $l = \cos(No, x)$, $m = \cos(No, y)$, No is the nomal vector of any point. $\boldsymbol{H}$,$\boldsymbol{Q}$ and $\boldsymbol{N}$ are the shape function.

Substituting the approximate function $\boldsymbol{f}$ of Eq.(7~8) into Eq.(1), the residuals are minimized in a least-squares manner,

$$\Pi = \int_{\Omega} \sigma_{ij,j}\sigma_{ik,k}\,d\Omega + \int_{\Gamma u} (u_i - \overline{u}_i)(u_i - \overline{u}_i)\,d\Gamma u + \int_{\Gamma t} (\sigma_{ij}n_j - \overline{t}_i)(\sigma_{ij}n_j - \overline{t}_i)\,d\Gamma t \tag{18}$$

The system equations of the MWLS method for solving thermoelastic problem are obtained as

$$\boldsymbol{K}\boldsymbol{d} = \boldsymbol{P} \tag{19}$$

Where

$$\boldsymbol{K} = \sum_{s=1}^{N} \boldsymbol{H}^T(x_s)\boldsymbol{H}(x_s) + \sum_{s=1}^{N_u} \boldsymbol{N}^T(x_s)\boldsymbol{N}(x_s) + \sum_{s=1}^{N_t} \boldsymbol{Q}^T(x_s)\boldsymbol{Q}(x_s) \tag{20}$$

$$\boldsymbol{P} = -\sum_{s=1}^{N} \boldsymbol{H}^T(x_s)\overline{f}_s + \sum_{s=1}^{N_u} \boldsymbol{N}^T(x_s)\overline{u}_s + \sum_{s=1}^{N_t} \boldsymbol{Q}^T(x_s)\overline{t}_s \tag{21}$$

Where, $\boldsymbol{H}$,$\boldsymbol{Q}$ and $\boldsymbol{N}$ are obtained by Eq.(15~17), $\boldsymbol{d}$ denotes the displacement of x,y. $\boldsymbol{D}$ is the stiffness matrix for a linearly elastic, isotropic 2-D solid.

$$\boldsymbol{\sigma} = \boldsymbol{D}\boldsymbol{\varepsilon} - \boldsymbol{\beta}\theta \tag{22}$$

$$\boldsymbol{D} = \frac{\overline{E}(x)}{1-\overline{\upsilon}(x)^2}\begin{bmatrix} 1 & \overline{\upsilon}(x) & 0 \\ \overline{\upsilon}(x) & 1 & 0 \\ 0 & 0 & (1-\overline{\upsilon}(x))/2 \end{bmatrix} \tag{23}$$

in which $\bar{E} = E$; $\bar{\upsilon} = \upsilon$; $\boldsymbol{\beta} = \dfrac{\alpha E}{1-\upsilon}\begin{Bmatrix}1\\1\\0\end{Bmatrix}$, for plane stress with E、γ and α denoting the Young's modulus, Poisson's ratio, and coefficient of thermal expansion, respectively, and $\bar{E} = \dfrac{E}{1-\upsilon^2}$; $\bar{\upsilon} = \dfrac{\upsilon}{1-\upsilon}$; $\boldsymbol{\beta} = \dfrac{\alpha E}{1-2\upsilon}\begin{Bmatrix}1\\1\\0\end{Bmatrix}$ for plane stain.

2.3 material properties

There are two methods to describe the variance of the material properties. One is to use the specific functions for all kinds of material properties. The other is to employ the specific functions of volume fraction of FGM objects. In this paper, relevant material properties at the discrete points are determined based on Mori-Tanaka model [17]. It is the modified rule of mixtures and the effective material properties can be determined using the following relation,

$$k = k_1 + \frac{3k_1V_2(k_2 - k_1)}{3k_1 + V_1(k_2 - k_1)} \tag{24}$$

$$E = E_1 + \frac{V_2(3E_1 + 4\mu_1)(E_2 - E_1)}{3(1-V_2)(E_2 - E_1) + 3E_1 + 4\mu_1} \tag{25}$$

$$\mu = \mu_1 + \frac{V_2(\mu_1 + f_1)(\mu_2 - \mu_1)}{(1-V_2)(\mu_2 - \mu_1) + \mu_1 + f_1} \tag{26}$$

$$\alpha = \alpha_1 + \frac{E_2(E_1 - E)(\alpha_2 - \alpha_1)}{E(E_1 - E_2)} \tag{27}$$

Where, $E_1 = \dfrac{P_1}{3(1-2\upsilon_1)}$; $E_2 = \dfrac{P_2}{3(1-2\upsilon_2)}$; $\mu_1 = \dfrac{P_1}{2(1+\upsilon_1)}$; $\mu_2 = \dfrac{P_2}{2(1+\upsilon_2)}$; $f_1 = \dfrac{\mu_1(9E_1 + 8\mu_1)}{6(E_1 + 2\mu_1)}$, *P* may be modulus of elasticity, Poisson's ratio υ, bulk modulus *E*, shear modulus μ, thermal conductivity *k*, volume fraction *V* and coefficient of thermal expansion α.

3 Numerical results and discussions

A clamped-clamped FGM beam is shown in Figure 1, length L=1000mm, width D=500mm and thickness H=100mm, k1=233W/mK, k2=65W/mK, E1=7e10Pa, E2=4.27e11Pa, v1=0.3, v2=0.17, alfa1=2.34e-5/K, alfa2=4.3e-6/K, Q=5e5W/m3, the spatial variation of the volume fraction of material 1 is taken to be a power law distribution in the y-direction as V=(y/D)^a.

The beam is assumed to be in a state of plane strain normal to the xy plane, and the design region is discretized as 31×15. The effective material properties are determined by the Mori-Tanaka scheme. In order to verify our method, we do some compararions between MWLS and the commercial COMSOL Multiphysics for uniform material (a=0), relevant

results listed in Table 1. The results obtained with the two methods are in good agreement in Table 1. The maximum temperature 361.8K is in the center (0.5,0.25) of the beam.

For a=2, we analyzed the thermoelastic problem and heat conduction problem using MWLS, details of heat conduction analysis using MWLS can be found in Zhou et al.[15]. Temperature field distribution, x-displacement and y-dispalcement are shown in Figure2, Figure3 and Figure4, respectively. Figure 2 indicates that the maximum temperature 425.6K is higher that the uniform material, in the Cartesian Coordinates (0.5,0.179) of the beam. Figure 3 and Figure 4 show that, when subjected to temperature rise, the beam expands, the maximum y-displacement is located at the top middle of the beam. To make a comparison, we do thermoelastic analysis and obtain thermal stresses in the neutral axis of the beam among a=0,a=2 and a=3, as shown in Figure5. In Figure5 (a) and (c), the volume fraction of material 1 is gradually decreased from a=0 to a=3, the σx and σy stesses are be in an upward trend. The maximum thermal stress always occurred in the vicinity of neutral axis of the beam from Figure5 (a,c,d). The results also agreed well with the presented elasticity solutions of Ref [13].

Table 1. Compararion of MWLS method and COMSOL Multiphysics

method	Temperature/K		X displacement /mm		Y displacement/mm	
	maximum	minimum	maximum	minimum	maximum	minimum
MWLS	361.8	300	.596	-.873	2.7	-2.6
COMSOL Multiphysics	361.1	300	.639	-.639	3.81	-3.81

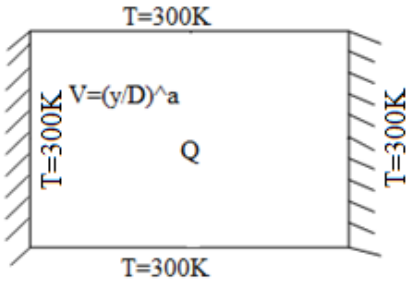


Figure1. Initial condition

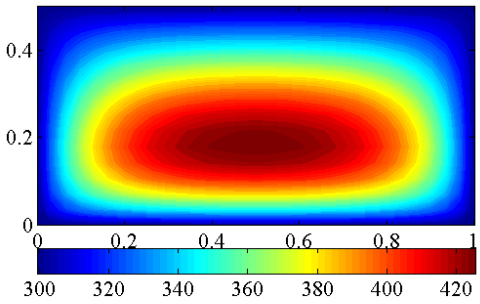


Figure 2. Distribution of temperature field

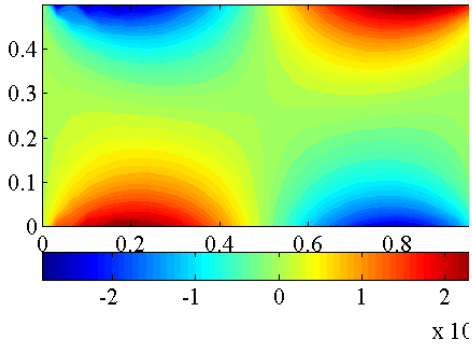


Figure 3. x-displacement/m

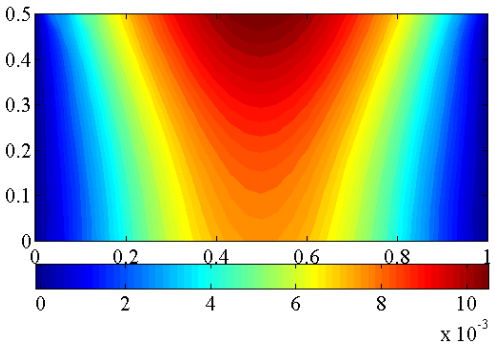
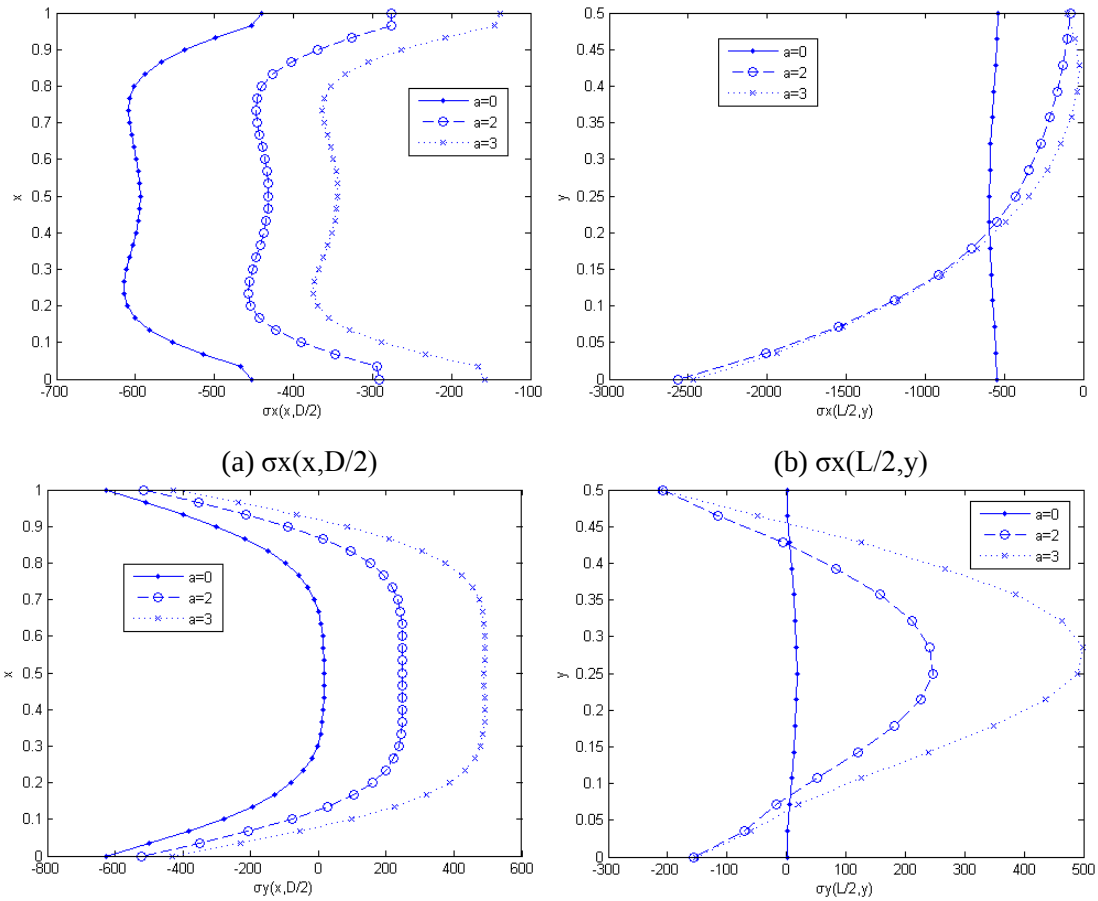


Figure 4. y-displacement /m



(a) σx(x,D/2) (b) σx(L/2,y)

(c) σy(x,D/2) (d) σy(L/2,y)

Figure 5. Stress results compararion of thermoelastic analysis among a=0,a=2 and a=3

4 Conlusion

In this paper, a noval thermoelastic analysis of FGM beam based on MWLS method was presented. We do thermoelastic and heat conduction analysis aimed at a clamped-clamped thick beam which is subjected with interior heat source. The FGM beam is assumed to be given by a simple power law distribution. Material properties of the FGM beam are obtained by Mori-Tanaka method. Through compared with the commercial software, it verified the effectiveness and accuracy. We also listed the compararion of thermal stresses with the varation of power law index. The present method of analysis will be also useful in the design and optimization of FGM objects.

Acknowledgements

The work described in this paper was supported by a grant from the National Natural Science Foundation of China (Projects No.51505131), the Fundamental Research Funds for the Universities of Henan Province and a doctor fund (No. B2013-032) and Program for Innovative Research Team (No.T2017-3) of Henan Polytechnic University. The correlative members of the projects are hereby acknowledged.

References

[1] K.Swaminathan, D.M.Sangeetha. (2017) Thermal analysis of FGM plates – A critical review of various modeling techniques and solution methods. Composite Structures 160, 43-60.

[2] Sergio Turteltaub. (2002) Optimal control and optimization of functionally graded materials for thermomechanical processes. International Journal of Solids and Structures 39, 3175-3197.

[3] Kou X.Y., Tan S.T. (2007) A systematic approach for Integrated Computer-Aided Design and Finite Element Analysis of Functionally-Graded-Material objects. Materials and Design 28, 2549-2565.

[4] M.Lezgy-Nazargah. (2015) Fully coupled thermo-mechanical analysis of bi-directional FGM beams using NURBS isogeometric finite element approach. Aerospace Science and Technology 45, 154-164.

[5] Alok Sutradhar, Glaucio H.Paulino. (2004) The simple boundary element method for transient heat conduction in functionally graded materials. Comput.Methods Appl. Mech. Engrg.193, 4511-4539.

[6] Ching H.K.,Chen J.K. (2007) Thermal stress analysis of Functionally Graded composites with temperature-dependent material properties. Journal of mechanics of materials and structures 2,633-653.

[7] H.K.Ching, S.C.Yen. (2005) Meshless local Petrov-Galerkin analysis for 2D functionally graded elastic solids under mechanical and thermal loads. Composites:Part B 36, 223-240.

[8] Hui wang, Qing-Hua Qin. (2008) Meshless approach for thermo-mechanical analysis of functionally graded materials. Engineering Analysis with Boundary Elements 32, 704-712.

[9] Andrew J.Goupee, Senthil S.Vel. (2006) Two-dimensional optimization of material composition of functionally graded materials using meshless analyses and a genetic algorithm. Comput. Mehtods Appl. Mech. Engrg.195, 5926-5948.

[10] John T.Katsikadelis. (2008) The 2D elastostatic problem in inhomogeneous anisotropic bodies by the meshless analog equation method (MAEM). Engineering Analysis with Boundary Elements.32, 997-1005.

[11] D.F.Gilhooley, J.R.Xiao, R.C.Batra. etc. (2008) Two-dimensional stress analysis of functionally graded solids using the MLPG method with radial basis functions. Computational Materials Science 41, 467-481.

[12] Liu G.R, Gu Y.T. (2005) An introduction to meshfree methods and their programming. Springer, Berlin.

[13] Bhavani V.Sankar, Jerome T.Tzeng. (2002) Thermal stresses in Functionally Graded Beams. AIAA Journal. Vol.40,1228-1232.

[14] K.J.Sohn, J.H.Kim. (2008) Structural stability of functionally graded panels subjected to aero-thermal loads. Composite Structures 82, 317-325.

[15] Zhou H.M., Liu Z.G. and Lu B.H. (2010) Heat conduction analysis of heterogeneous objects based on multi-color distance field. Mater Design 31, 3331-3338.

[16] Zhou H.M., Zhou W.H., Qin G. etc. (2017) Transient heat conduction analysis for distance-field-based irregular geometries using the meshless weighted least-square method. Numerical Heat Transfer, Part B:Fundamentals 71, 456-466.

[17] X.Peng, N.Hu, H.Zheng, H.Fukunaga. (2009) Evaluation of mechanical properties of particulate composites with a combined self-consistent and Mori–Tanaka approach. Mechanics of Materials 41, 1288-1297.

Numerical Study on Growth of Strata Disturbance Abscission Layer

Yu Guangming1，Li Gang1,3，Song Kun1，Yu Bingyao1，Lu Shibao1，Qin Yongjun2

[1]School of Civil Engineering, Qingdao Technological University, Qingdao 266033, Shandong, China

[2]School of Civil Engineering and Architecture, Xinjiang University, Urumqi 830046, Xinjiang, China

[3]Institute of Architectural Engineering, Qingdao Agricultural University, Qingdao 266109, Shandong, China

Abstract

In the sedimentary strata, rock strata exhibit layered feature and have an obvious level. Overlying rock often produces separation layer along the level after the underground excavation. However, due to the difference of each layer, but the rule of separated layer is extremely complex due to the difference of each layer, the change of the mining process and the height difference in the level and other factors. However, the disturbed overburden separation is an important basis for studying the subsidence of the surface and implementing the damage control. Therefore, in this paper, the phenomenon, feature and regularity of disturbed overburden separation are found by means of numerical simulation. The paper reveals the mechanical mechanism of the development of overburden separation, and it is also explores the way and method of disturbed overburden separation, and verified by experiment. It is of great importance to understand the law of surface subsidence through the overburden separation development process.

Keywords: perturbed; overburden separation; numerical simulation; mechanical mechanism; surface subsidence

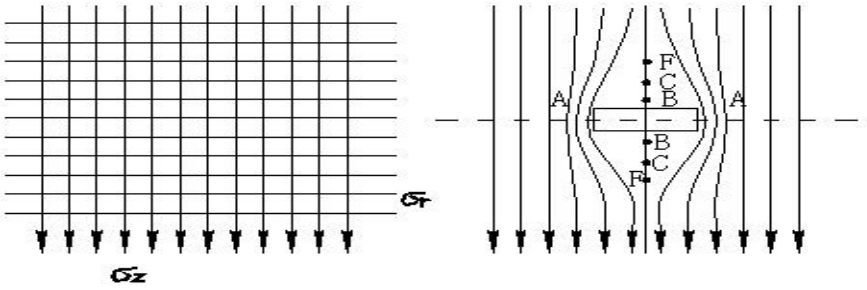
*The national natural science fund project (51374135, 51179080) and the national study abroad funded projects, relying on the blue economic zone construction and safety of collaborative innovation center.

（Contact person）Author’s brief introduction: Yu Guangming (1962-), male, graduated from China University of Mining and Technology in 1997 with a major in engineering mechanics. He is currently a professor and doctoral tutor. Mainly engaged in underground engineering, surface subsidence control and other direction of scientific research.

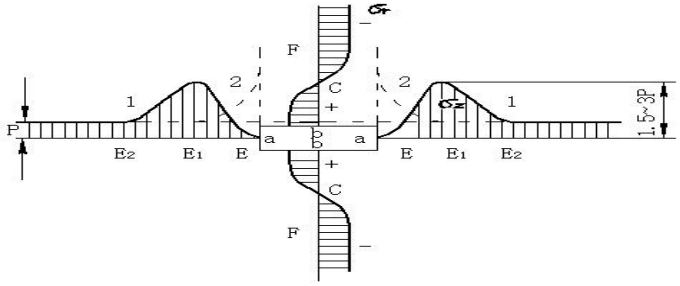
E-mail: yu-guangming@263.net. Tel: 13375557000

1 The Source of Surface Subsidence

In the sedimentary strata, underground excavation or useful minerals are excavated, the surrounding rock mass is damaged, resulting in stress redistribution (Fig.1), and seeking a new balance, so that the rock stratum and the surface produce movement and deformation (Fig. 2), and even produce non-continuous damage, this is collectively called "strata subsidence", which is defined as the slow or sudden deformation, destruction, prominence and movement in vertical and horizontal direction of the strata because of overheaded or forced, developing up to the surface to produce subsidence which is larger than he underground excavation[1]-[2].



(a) Balanceable state of initial stress (b) Redistribution of stress



(c) stress distribution of arounding rock in mining region

Figure 1. Redistribution of initial stress caused by mining

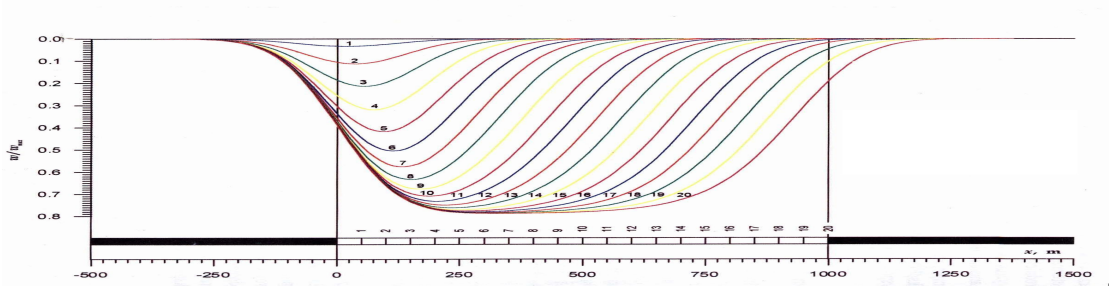


Figure 2. Subsidence of the earth's surface

2 Basic Characteristics of Perturbed Overburden Damage

Under the geological conditions of normal sedimentary strata, when the subterranean excavation is carried out, from its direct roof, and from bottom to up in order to occur creep, break, bend, separation, that is, in the overlying fracture, separation and synchronous bend field of the excavation area. In the creeping field, the rock mass is broken from the bottom, and produce a rearranged structure; In the fracture field, the rock mass is fractured or broken perpendicular to the plane, and its fracture depth and fracture frequency decrease with the vertical distance from the working face increase; In the separation field, the rock mass splits along the plane and develops into separation; In the synchronous bend field, the rock mass including the topsoil developed to synchronous bend under the action of geostatic stress(Figure3).

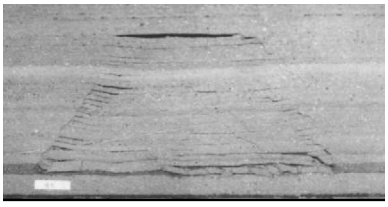


Figure 3. The picture of failured disturbed overburden separation

3 Mechanics Mechanism and Dynamic Process of The Development of Strata Disturbance Abscission Layer

3.1 Mechanics of The Formation and Development of Separation

After the underground excavation, in the overlying rock mass which is effected by disturbance, the direction of the maximum principal stress σ 1 changed from the vertical direction before excavation to the skew on both sides of the excavation area Due to the fact that some of the vertical stress protoliths on the top of the roof are transferred to both sides of the excavation area, the load on both sides of the rock mass is higher than that before exploitation, and form the bearing pressure belt on both sides of the excavation area. A large number of studies have shown that the bearing pressure belt takes the shape of "arch" in the overlying rock mass, called the bearing pressure arch. The two arch feet are located in the rear of the face and in front of the working area, and move forward with the advance of the working face; Because of the pressure of the upper strata is guided by the bearing pressure belt to the rock mass on both sides of the mined-out area, the vertical stress is lower than that before the mining, and it is called unloaded arch[3]-[4]. The rock itself in the unloaded arch moves to the surface of the mining area due to the elastic recovery and the self-weight action, resulting in bending deformation. Because the difference of lithology, thickness and height of the strata, the sinking of the upper and lower strata is not synchronized. When the deflection of the upper strata is smaller than that of the lower strata, the upper and lower strata undergo a abrupt instability of spallation and then produce separation.

In the first stage of force and destruction of the overburden in the excavation area, the overlying rock mass in the excavation area first forms a small bearing pressure arch and unloaded arch. In the continuous excavation stage, as the working face forward and pull back the top, the front arch of the bearing pressure arch and unloaded arch to move forward, and continue to expand. As the working face constantly advancing, the creeped rock at the back of the working face is compacted, and the stress state in the upper rock mass gradually returns to the original stress state. At this time, the bearing pressure arch and unloaded arch will continue to advance with the working face forward, as the road continuously to move forward. When entering the stop phase, the bearing pressure arch and unloaded arch will enter the static stage(Fig.4).

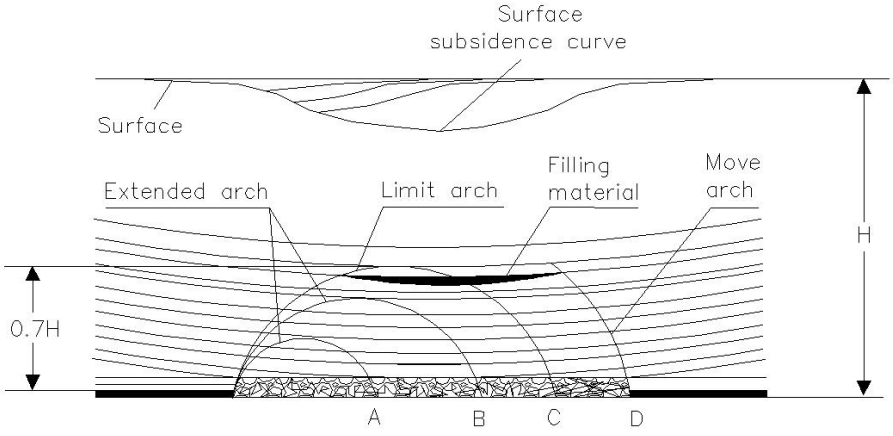


Figure 4. The distribution of the movement and separation of bearing pressure arch and unloaded arch

3.2 The Dynamic Process of The Formation and Development of Separation

With the three stages of the dynamic stress process of the overburden in the excavation area, the formation and development of the separation in the overlying strata have experienced three periods, and a series of dynamic changes.

After the first caving, the rock in the unloaded arch generates curve subsidence and split for the action of gravity, and then produce the separation, forming the initial four-domain distribution and the "arch-balanced structure" of overburden. With the continuous advance of the working face, the arch-beam equilibrium structure is expanding, the internal separation occur a series of changes, the original separation continuously expand, and then closed, the new separation produce and expand in the higher separation, but the vertical largest discontinuous deformation generally occurs at the top of the unloaded arch. However, this kind of arch-beam equilibrium structure is only a very quasi-static balance, but in reality with the advance of the working face it transits from a quasi-equilibrium state to another quasi-equilibrium state. It is a dynamic development process from small to large, low to high and back to forward. When the bearing pressure arch and unloaded arch reache the limit stage, the arch-beam equilibrium structure reaches the maximum period, called the ultimate equilibrium arch; the separation on the top of the arch is also developed to the highest position, called the separation limit height. And then with the advance of the working face, the space range of limited arch-beam equilibrium structure translates forward instead of expanding. The separation is no longer to develop upward, but forward expansion, and to a certain period of time, the rear part of the separation closed. Until the late mining, due to the continuous subsidence of the upper rock, the separation is closed continuously [5]-[9].

From the above analysis we can see that the separation start from the tension of the layers, first occurred separation from the gradient to the mutation, resulting in separation. And then with the advance of the working face, the separation expand and closed, and the distribution and development of separation is from small to large, from bottom to up, from back to forword, with the development of the internal space occupied by the separation is increased by less, then by the reduction of the number of changes.

4 Mining The Spallation of The Overburden Strata — Mechanics Mechanism of The Abscission Layer

The spallation of rock mass is refers to the cracking of the sedimentary rock mass along the rock interface under the action of mining stress, including shear (interlayer dislocation) and cracking (separation).Spallation is a common problem in mining subsidence. The sedimentary layered rock mass is relatively common. During the process of sedimentary rock, due to the change of external conditions, the sedimentary process is interrupted or temporarily interrupted. In the rock mass, the unconformity surface, the pseudo integration surface and the the lithology of different surfaces are formed. Due to the difference between the upper and lower layers of the rock and the existence of the structural plane, making it becomes the stress and displacement discontinuity. The amount of stress displacement transfer, depends on sedimentary surface properties and load conditions. With the action of additional stress in rock strata, the overlying strata of the mining area will be split, and the interlayer is also connected

in the vertical direction near the stope area, and the abscission layer is formed away from the stope.

The delamination (referred to as cracking) caused by transverse stretching occurs and develops under the self-weight of the pull-down strata. As shown in Fig. 5, the mechanical conditions can be simply expressed as:

$$\sigma_A \geq [\sigma_T] = C \tag{1}$$

In the formula: σ_A is the tensile stress of A point on the level； $[\sigma_T]$ is the level of unidirectional tensile strength， in fact, is the level of adhesion C.

The final result of the tensile fracture is the separation between the layer and the layer, and it is clear that the spallation will rapidly expand after the crack of the a point, and its extension will depend on the suspended span and the cohesive force between the layers.

Caused by the longitudinal shear layer crack (referred to as shear crack) is implemented in the process of strata subsidence curve, as shown in figure 6, the mechanical conditions as follows:

$$\tau = C + \sigma_n tg\varphi \tag{2}$$

In the formula: τ is the interlayer shear force; τ_n is the normal stress on the level ； φ is the friction angle between layers.

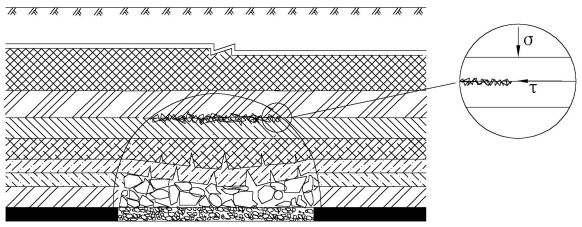


Figure 5. Bearing force state of strata lay

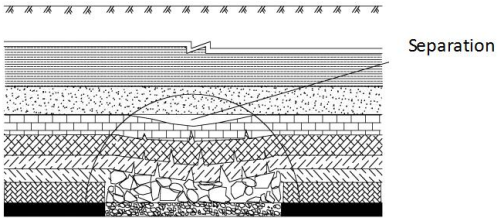


Figure 6. Leaving lay of covering rock

5 Numerical Calculation and Analysis of Disturbed Overburden Separation

According to the rupture process of disturbed rock mass, the actual state of engineering rock mass and the mechanism and regularity of overburden separation which reveals from above research, we use the RFPA simulation software[10] to analyze the failure process of rock mass, realize the visualization simulation of overlying strata, as shown in Figure 7 is the rupture process simulation of no jointed rock mining, Figure 18 is the numerical simulation of the stress in the overlying strata during the different excavation distance of underground excavation.

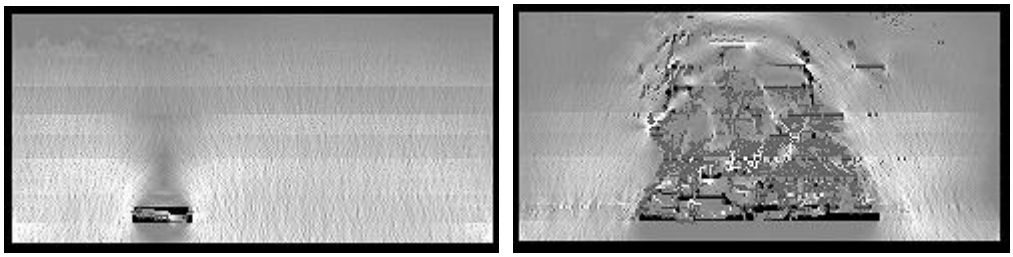


Figure 7. Mining damage process of rocks that have no crannies

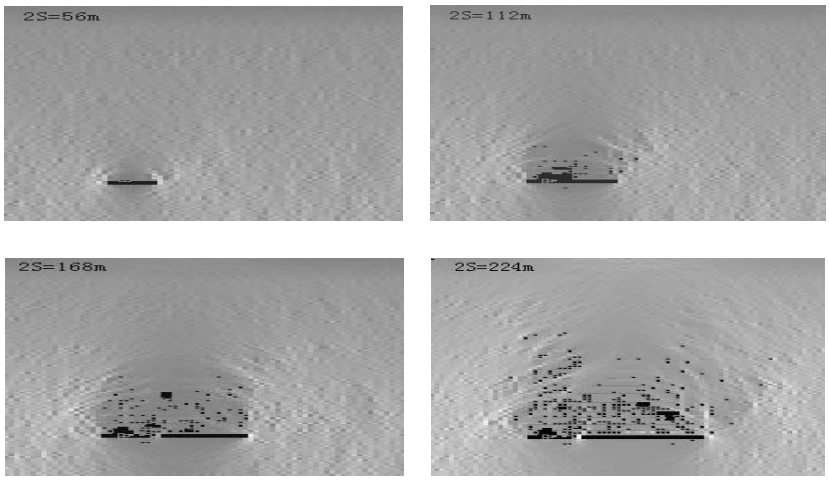


Figure 9. Stress and deformation simulation of covering rocks caused by mining

6 Concluding Remarks

Disturbed strata separation is one of the most important characteristics in the process of strata subsidence, and its evolution rules control the development law of strata subsidence to some extent. Therefore, it is a great theoretical significance and engineering guidance value to study the allometric evolution law of disturbed overburden strata. This paper revolves around the mechanism of disturbed strata separation, meanwhile, analyzing the forming conditions, mechanism and development characteristics of mining overburden rock, etc. The RFPA numerical simulation is used to study the growth evolution law of disturbed strata separation The knowledge of the stress state and the dynamic process of the separated strata are obtained. It makes sense to further study the formation of disturbed strata and its influence on the mechanical properties of overlying strata and surface subsidence.

Reference

[1] Yu Guangming,Yang Lun,Su zhongjie.Nonlinear Principle, Monitoring and Control of Mining Rock Mass Failure[National Natural
[2] Science Fund Monograph Publishing Fund.Jilin：Jilin University Press,2000
[3] Yu Guangming,Zhao Jianfeng,Xie Heping. Numerical Simulation of Fractal Interface Effect of Mining-Caused Activation of Fault.
[4] Discrete Dynamics in Nature and Society, 2002,7(3):151~155
[5] Борисов,А·А·,Wang Qingkang translation, Principle and Calculation of Mine Pressure,Coal Industry Press,1986
[6] S.S.Pen, Strata Control in Coal Mine
[7] Zhao Deshen、Su Zhongjie, et al. Experimental Study on The Surface Subsidence of Rock Stratum Filling,Journal of Coal science 1997.6.
[8] Fan Xueli、Liu Wensheng,et al. Theory and Practice of Mining Damage Protection in Northeast China Coal
[9] Mining Area,Coal Industry Press,1998.

[10] Su Zhongjie,Yu Guangming,Yang Lun. Mechanical Model of Overburden Deformation and Its Application.Journal of Geotechnical
[11] Engineering,2002,24(6): 778-781.
[12] Wang Jinshan,Wang Zhongchang. Three Dimensional Numerical Simulation of Mining
[13] Overburden Failure and Stratified Development. Coal Mining,2011,16(6):22-26.
[14] Dai Huayang,Deng Zhiyi,Yan Yueguan, et al. Study on Distribution Law of Overlying Strata and
[15] Normal Fracture in Deep Mining of Tangshan Coal Mine. Coal Mining,2011，16(2):8-11.
[16] Tang Chunan, Yu Guangming. Numerical Simulation of Rupture and Rock Mass Movement in Mining Rock[National Natural Science
[17] Fund Monograph Publishing Fund].Jilin：Jilin University Press,2003

A stabilization method of F-barES-FEM-T4 for dynamic explicit analysis of nearly incompressible solids

Ryoya Iida[1,a)], Yuki Onishi[1] and Kenji Amaya[1]

[1] Department of Systems and Control Engineering, Tokyo Institute of Technology, Japan

[a)]Corresponding and Presenting author: riida@a.sc.e.titech.ac.jp

Abstract

SymF-barES-FEM-T4 is proposed in order to ensure the stability of F-barES-FEM-T4 in dynamic analysis. This formulation aims to symmetrize stiffness matrix of F-barES-FEM-T4 so as to remove unstable deformation modes. The modification for stabilization is restricted to the internal force vector and does not cause increase of degrees of freedom (DOF). An example of analysis reveals that SymF-barES-FEM-T4 can show the comparable deformed shapes and acceptable pressure distributions without energy divergence, which cannot be realized by F-barES-FEM-T4. It also reveals that increase of the number of cyclic smoothings for SymF-barES-FEM-T4 does not always improve the pressure distributions no longer unlike F-barES-FEM-T4.

Introduction

The tetrahedral elements are the current only choice in finite element (FE) analysis for complex structures due to their ease of mesh generation. However, conventional tetrahedral FE formulations easily cause locking and pressure oscillation in nearly incompressible cases. The high-order formulations can resolve only shear locking but still suffer from volumetric locking and pressure oscillation. The u/p hybrid formulations[1], which is widely used to overcome incompressibility, are applicable only in implicit cases but inapplicable in explicit dynamics. Thus, FE formulations for dynamic explicit analysis of nearly incompressible materials with 4-node tetrahedral (T4) elements are still in research stage.

Recently, some T4 elements featured with the idea of smoothed finite element methods (S-FEMs)[2] have been proposed in order to realize explicit dynamics of nearly incompressible materials. Selective ES/NS-FEM-T4[3, 4, 5, 6, 7] decomposes Cauchy stress tensor into hydrostatic part and deviatoric part, which are derived from NS-FEM-T4 and ES-FEM-T4 respectively. Although it can analyze deformation without locking, it cannot completely overcome pressure oscillation. F-barES-FEM-T4[8, 9] decomposes deformation gradient into isovolumetric part and volumetric part in the same manner as F-bar method[10]. Isovolumetric part is derived in the same as ES-FEM-T4 and volumetric part is derived by smoothing between nodes and elements in a few times. This formulation can completely overcome locking and pressure oscillation in static analysis and be expected to show the same ability in explicit dynamics.

Our group extended F-barES-FEM-T4 to explicit dynamics in previous work[11]. The results can show comparable pressure distribution and deformed shapes without locking. However, they also revealed unstability of F-barES-FEM-T4 in dynamic problems due to their imaginary parts of eigenfrequencies raised by the asymmetric stiffness matrix; therefore, high-accurate results are restricted to short-term analysis.

This paper proposes a stabilized F-barES-FEM-T4, named SymF-barES-FEM-T4. The idea for stabilization is to symmetrize stiffness matrix of F-barES-FEM-T4 by replacing the formula to derive the internal force. In the following sections, the outline and an example of analysis for SymF-barES-FEM-T4 are explained. An example for explicit dynamics of nearly incompressible materials illustrates the stability and accuracy of SymF-barES-FEM-T4.

Methods

This section explains the outline of F-barES-FEM-T4 and that of proposed method named SymF-barES-FEM-T4.

Outline of F-barES-FEM-T4

F-barES-FEM-T4 is designed in order to realize highly accurate analysis for nearly incompressible materials. This formulation uses the ideas of F-bar method[10] and S-FEMs[2]. At first, deformation gradients at h-th edge ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}$ are decomposed into isovolumetric part ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{iso}}$ and volumetric part ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{vol}}$ in the same manner as F-bar method:

$$ {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}} = {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{iso}} \cdot {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{vol}}. \tag{1} $$

${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{iso}}$ is calculated in the same manner as ES-FEM-T4, namely,

$$ {}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{iso}} = {}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}^{\mathrm{iso}} = \frac{1}{{}^{\mathrm{Edge}}_{h}\widetilde{J}^{1/3}} {}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}, \tag{2} $$

where $\widetilde{\square}$ denotes the smoothed value at an edge and ${}^{\mathrm{Edge}}_{h}\widetilde{J}$ is relative volume change at h-th edge calculated as $\det({}^{\mathrm{Edge}}_{h}\widetilde{F})$. Meanwhile, volumetric part ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}^{\mathrm{vol}}$ is derived from the weighted mean of neighboring elements' relative volume change ${}^{\mathrm{Elem}}J$. Weight values are defined by cyclic smoothing among nodes and elements. More concrete procedure is described in [8, 9, 11]. The smoothed Cauchy stress tensor at h-th edge ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{T}}$ is derived from ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}$ and material constitutive model. The nodal force vector at h-th edge ${}^{\mathrm{Edge}}_{h}\boldsymbol{f}^{\mathrm{int}}$ is calculated as following:

$$ {}^{\mathrm{Edge}}_{h}f^{\mathrm{int}}_{P:p} = \frac{\partial {}^{\mathrm{Edge}}_{h}\widetilde{D}_{ij}}{\partial \dot{u}_{P:p}} {}^{\mathrm{Edge}}_{h}\overline{T}_{ij} {}^{\mathrm{Edge}}_{h}V, \tag{3} $$

where $\square_{P:p}$ indicates p-th direction of P-th nodal value, u is nodal displacement, $\dot{\square}$ indicates time derivative value, ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{D}}$ is stretching tensor derived from ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{F}}$ and ${}^{\mathrm{Edge}}_{h}V$ is assigned volume to h-th edge.

Outline of SymF-barES-FEM-T4

According to [11], F-barES-FEM-T4 holds unstability in dynamic analysis due to the imaginary parts of eigenfrequencies. Such imaginary parts are caused by asymmetry of the stiffness matrix in small deformation analysis. To ensure the stability in dynamic analysis, we modify ${}^{\mathrm{Edge}}_{h}\boldsymbol{f}^{\mathrm{int}}$ in the following manner[12]:

$$ {}^{\mathrm{Edge}}_{h}f^{\mathrm{int}}_{P:p} = \frac{\partial {}^{\mathrm{Edge}}_{h}\overline{D}_{ij}}{\partial \dot{u}_{P:p}} {}^{\mathrm{Edge}}_{h}\overline{T}_{ij} {}^{\mathrm{Edge}}_{h}V, \tag{4} $$

where stretching tensor ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{D}}$, derived from ${}^{\mathrm{Edge}}_{h}\overline{\boldsymbol{F}}$, is used instead of ${}^{\mathrm{Edge}}_{h}\widetilde{\boldsymbol{D}}$. This formulation, named SymF-barES-FEM-T4, ensures the symmetry of stiffness matrix in small deformation case; therefore, it can be expected that unstable deformation modes of F-barES-FEM-T4 are removed. Details of derivation for each term are summarized in [9].

Since this modification is restricted to the stretching tensor, SymF-barES-FEM-T4 can be formulated without increasing DOF and restrictions on material constitutive models as well as F-barES-FEM-T4. However, the calculation cost for internal force vector is more expensive than F-barES-FEM-T4.

Results

Figure 1 illustrates the outline of dynamic bending analysis for cantilever. The shape of cantilever is defined as $10 \times 1 \times 1$ m cuboid; its left side is perfectly constrained; a uniform initial velocity of 2.0 m/s in $-z$ direction is applied. The material constitutive model is Neo-Hookean hyperelastic model. The density, initial Young's modulus and initial Poisson's ratio are 920 kg/m^3, 6.0 MPa and 0.49, respectively. The analyses with ABAQUS/Explicit C3D4, Selective ES/NS-FEM-T4, SymF-barES-FEM-T4 and F-barES-FEM-T4 are performed with unstructured tetrahedral elements of 0.2 m global mesh seed size. The analysis with ABAQUS/Explicit C3D8 of 0.2 m global mesh seed size is also performed to obtain a reference solution. The number of cyclic smoothings c is 1 to 3, in the analyses with SymF-barES-FEM-T4 and F-barES-FEM-T4. All results of SymF-barES-FEM-T4 and F-barES-FEM-T4 are labeled with (c) such as "F-barES-FEM-T4(c)" In these analyses, the time integration scheme is Velocity Verlet, and the time increment is 1.0×10^{-4} s.

The comparison of the vertical displacements (u_z) at one of the corner node (◯ in Figure 1) is shown in Figure 2. Time histories of SymF-barES-FEM-T4s and F-barES-FEM-T4s agree with the reference in almost the same level; therefore, it

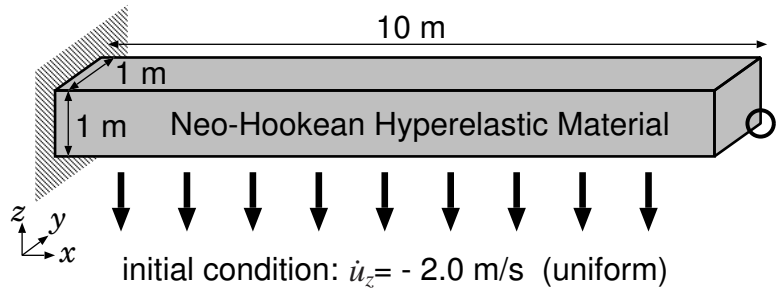


Figure 1. Outline of the dynamic bending analysis of a cantilever. The initial uniform velocity is -2.0 m/s in z direction.

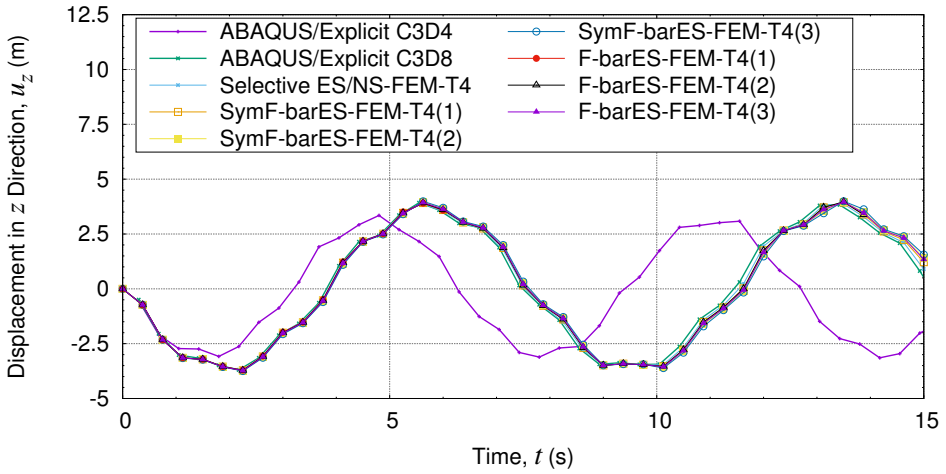


Figure 2. Comparison of the vertical displacement at the corner vs. time in the cantilever bending analysis.

can be concluded that symmetrization of SymF-barES-FEM-T4 doesn't spoil the locking-free property of F-barES-FEM-T4. Meanwhile, the result of ABAQUS/Explicit C3D4 shows far different time history due to the locking.

Figure 3 and 4 show the pressure distributions at 0.75 s and 4.50 s respectively. In these figures, the value above the range is colored in dark red, the one below the range is colored in dark blue and the contour ranges are [49.0, -76.8] (kPa) for 0.75 s and [0.10, -0.10] (MPa) for 4.50 s. The results of F-barES-FEM-T4 show comparable pressure distributions to the one of ABAQUS/Explicit C3D8 and increase of c improve the pressure oscillation. However, in Figure 4, F-barES-FEM-T4(1) shows the worst pressure distribution due to the energy divergence. F-barES-FEM-T4 cannot be applied to long-term analyses since they causes energy divergence in relative earlier stage than other formulation in this way.

SymF-barES-FEM-T4 can suppress pressure oscillation in the same level as not F-barES-FEM-T4 but slightly better than Selective ES/NS-FEM-T4 and increase of c does not improve pressure distribution no longer unlike F-barES-FEM-T4. Meanwhile, their accuracies are not spoiled even in Figure 4 since SymF-barES-FEM-T4 does not cause energy divergence.

Figure 5 illustrates the time histories of total energies for each formulation. SymF-barES-FEM-T4 does not cause energy divergence although F-barES-FEM-T4(1) and (2) cause energy divergence within this analysis time. This fact indicates symmetrization can suppress unstable deformation modes of F-barES-FEM-T4.

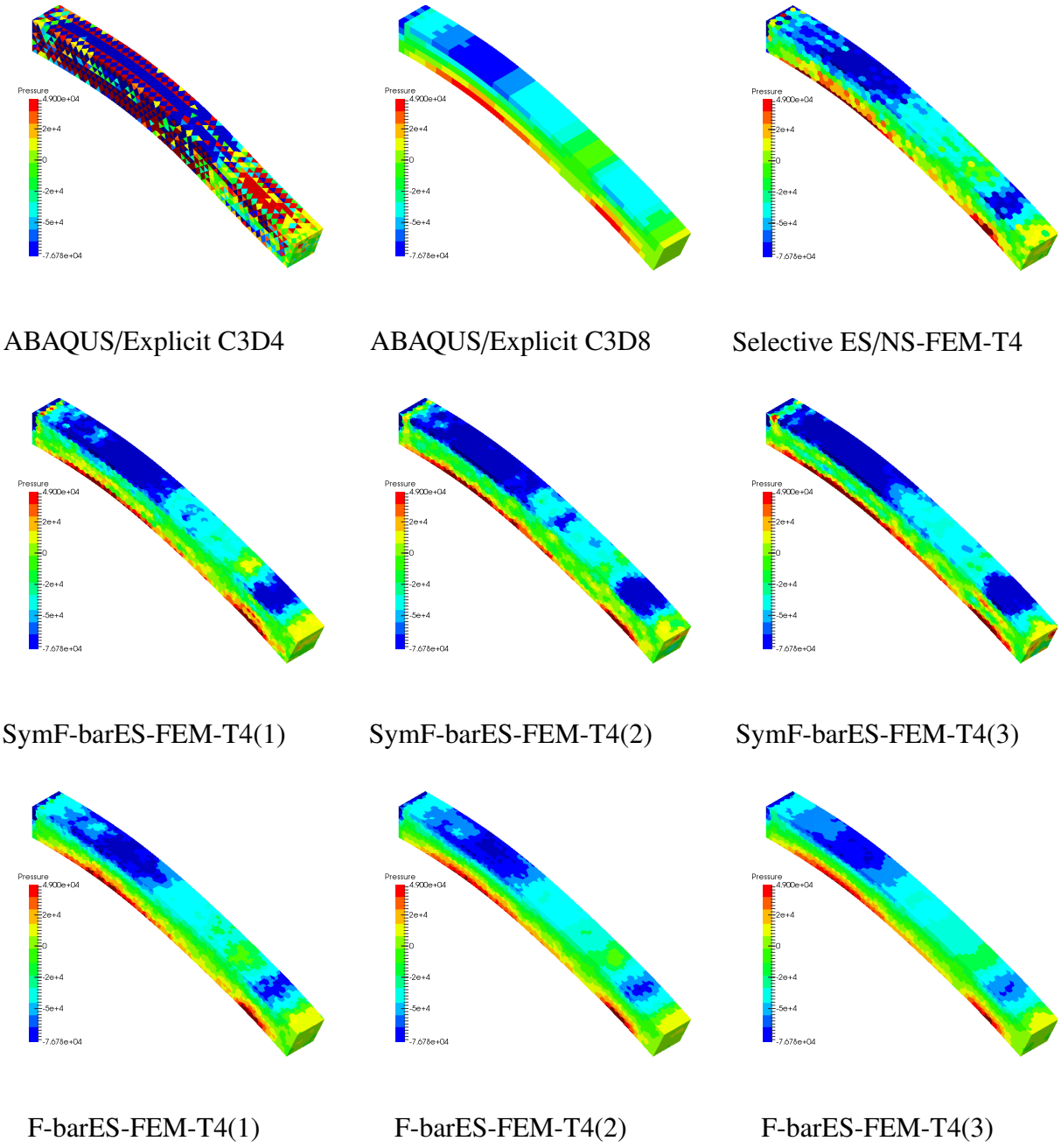


Figure 3. Deformed shapes and pressure distributions of the dynamic cantilever bending analysis at 0.75 s.

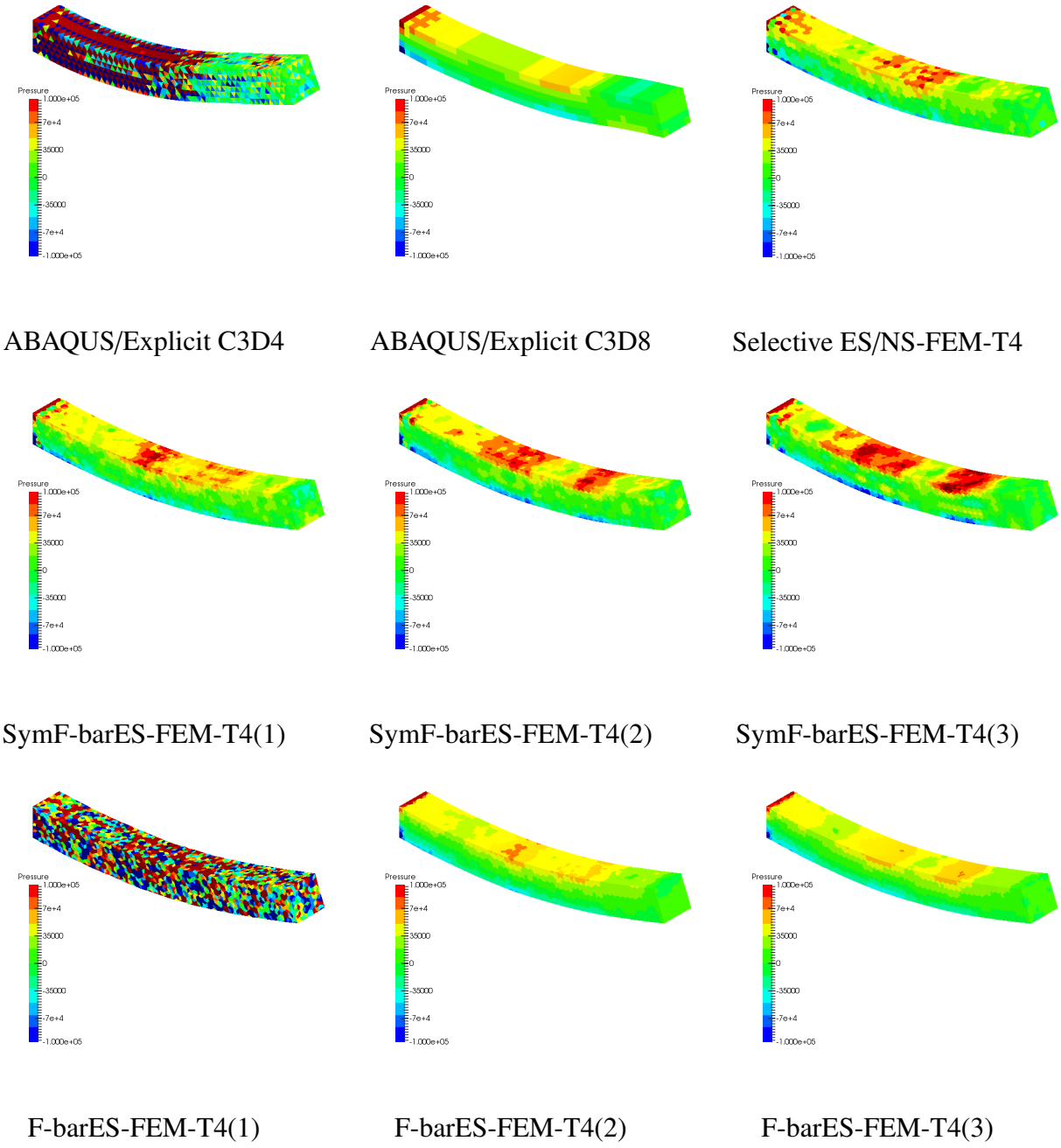


Figure 4. Deformed shapes and pressure distributions of the dynamic cantilever bending analysis at 4.50 s.

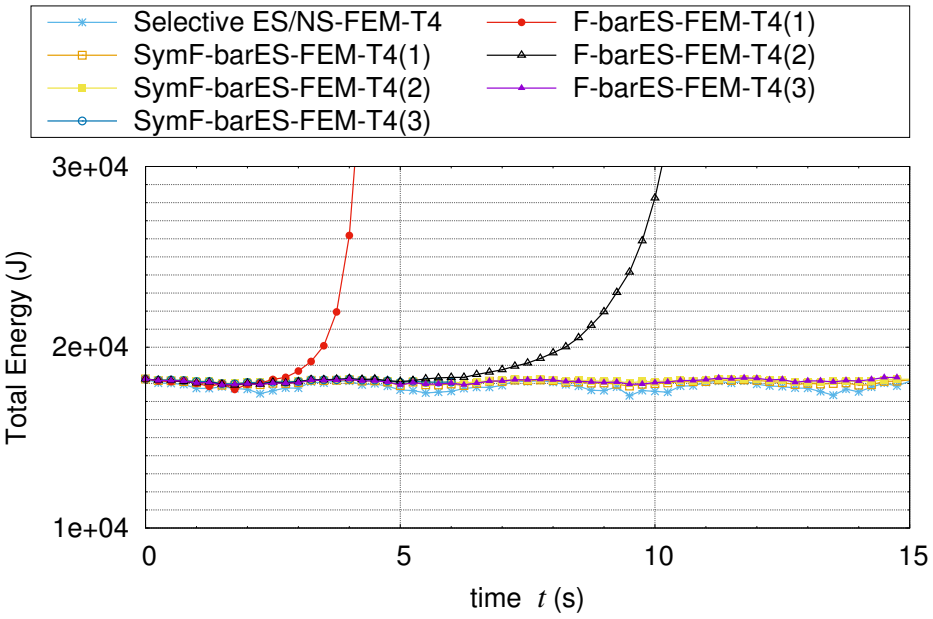


Figure 5. Time histories of the total energies for each formulation.

Conclusion

We propose stabilized F-barES-FEM-T4, named SymF-barES-FEM-T4, to realize long-term analysis for nearly incompressible materials. The advantages of SymF-barES-FEM-T4 are summarized in the followings:

- ✓ This formulation has following advantages which F-barES-FEM-T4 also has:
  - ✓ no increasing of DOF
  - ✓ no restrictions for material constitutive model
- ✓ Applicable for long-term analysis, which cannot be realized by F-barES-FEM-T4

The disadvantages of SymF-barES-FEM-T4 are summarized in the followings:

- ✗ Increasing of the number of cyclic smoothings does not improve pressure distributions unlike F-barES-FEM-T4
- ✗ suppression for pressure oscillation is worse than F-barES-FEM-T4

References

[1] Dassault Systèmes Simulia Corp. *ABAQUS 6.13 Theory Guide*. Dassault Systèmes Simulia Corp., Providence, RI, USA, 2013.

[2] G R Liu and T Nguyen-Thoi. *Smoothed Finite Element Methods*. CRC Press, Boca Raton, FL, USA, 2010.

[3] Eric Li, Zhongpu Zhang, C.C. Chang, G.R. Liu, and Q. Li. Numerical homogenization for incompressible materials using selective smoothed finite element method. *Composite Structures*, 123:216 – 232, 2015.

[4] Nguyen-Xuan Hung, Stéphane Pierre Alain Bordas, and Nguyen-Dang Hung. Addressing volumetric locking and instabilities by selective integration in smoothed finite elements. *Communications in Numerical Methods in Engineering*, 25(1):19–34, 2009.

[5] Chen Jiang, Zhi-Qian Zhang, Xu Han, and Gui-Rong Liu. Selective smoothed finite element methods for extremely large deformation of anisotropic incompressible bio-tissues. *International Journal for Numerical Methods in Engineering*, 99(8):587–610, 2014.

[6] Y. Onishi and K. Amaya. A locking-free selective smoothed finite element method using tetrahedral and triangular elements with adaptive mesh rezoning for large deformation problems. *International Journal for Numerical Methods in Engineering*, 99(5):354–371, 2014.

[7] Yuki Onishi and Kenji Amaya. Performance evaluation of the selective smoothed finite element methods using tetrahedral elements with deviatoric/hydrostatic split in large deformation analysis. *Theoretical and Applied Mechanics Japan*, 63:55–65, 2015.

[8] Y. Onishi and K. Amaya. F-bar aided edge-based smoothed finite element method with tetrahedral elements for large deformation analysis of nearly incompressible materials. *The 6th International Conference on Computational Methods*, 2015.

[9] Y. Onishi, R. Iida, and K. Amaya. F-bar aided edge-based smoothed finite element method using tetrahedral elements for finite deformation analysis of nearly incompressible solids. *Int. J. Numer. Meth. Engng.*, 2016.

[10] E.A. de Souza Neto, D. Peric, M. Dutko, and D.R.J. Owen. Design of simple low order finite elements for large strain analysis of nearly incompressible solids. *International Journal of Solids and Structures*, 33(20-22):3277–3296, 1996.

[11] R Iida and Y. Onishi K. Amaya. Performance evaluation of various smoothed finite element methods with tetrahedral elements in large deformation dynamic analysis. *The 7th International Conference on Computational Methods*, 2016.

[12] Thomas Elguedj, Yuri Bazilevs, Victor M Calo, and Thomas JR Hughes. B-bar and f-bar projection methods for nearly incompressible linear and non-linear elasticity and plasticity using higher-order nurbs elements. *Computer methods in applied mechanics and engineering*, 197(33):2732–2762, 2008.

Simulation of hydraulic fracturing process by using peridynamics

†Shuhui Li*, Fan Wu, Shasha Qiu, Zeyang Feng

Department of Engineering Mechanics, The State key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, Liaoning 116024, P.R.China.

*Presenting author: shuhuili@dlut.edu.cn
†Corresponding author: shuhuili@dlut.edu.cn

Abstract

A non-local bond-based peridynamic method is employed to simulate the hydraulic fracture process since it involves propagation of massive cracks in a brittle solid where crack branching and interaction present. Traditional numerical methods (e.g. XFEM) developed specially for cracks are awkward to simulate hydraulic fracture process because of its huge amount of cracks. Instead, peridynamic as a recently developed theory of solid mechanics replaces the partial differential equations of the classical continuum theory with integral equations. Hence, its basic equations are valid everywhere, regardless of continuities. This prominent advantage enables peridynamic to simulate hydraulic fracture process. This paper applies the peridynamic method to simulate the hydraulic fracturing of shale, containing horizontal well fracking with multiple perforations and initial natural cracks. A new scheme of tracing crack path and applying hydraulic fracture pressure is proposed. And some preliminary 2D results are presented to illustrate the validity of the proposed method. The numerical results show that hydraulic fracture cracks can restrain each other when they get close. With different angles of initial natural cracks, the crack pattern presents big difference.

Keywords: Peridynamic, Hydraulic fracture, Massive cracks, Fragmentation, Crack propagation

Introduction

Shale gas as a new energy source receives increasingly attentions these years. British Petroleum expects the shale gas revolution that has already transformed the U.S. natural gas market to continue apace. However, shale gas exploration faces many severe problems because of the depth of shale format, lean ore, environment pollution, etc [1]. To increase output, a mainstream technique called hydraulic fracture was developed. Hydraulic fracture, also known as fracking, injects high-pressure fluid into rocks deep underground, inducing the release of fossil fuels. Hence, using modern computer to simulate hydraulic fracture and consequently giving guidance on practical engineering is especially important.

Computational simulation is developed rapidly on account of its economic advantage. It's rather challenge to simulate problems involving massive cracks propagation, branching and interaction, such as hydraulic fracture process. Traditional numerical method such as finite element method, based on classical continuum mechanics which is the most popular adopted in commercial software, for example, has the assumption of continuity. This assumption, which contradicts the fact of physics, leads to the invalidation when it comes to discontinuity. Other numerical method such as extend finite element method specially developed for cracks also suffers severe problem dealing with massive cracks problem as fracking.

In light of the inadequacies of local classical continuum mechanics theories, the peridynamic theory, which is nonlocal, was introduced by Silling [2] in an attempt to deal with the

discontinuities. Basically, the peridynamics theory is a reformulation of the equation of motion in solid mechanics that is better suited for modeling bodies with discontinuities, such as cracks [3]. Consequently, peridynamic model is adopted in this paper to simulate hydraulic fracture, which many researchers have proved reliable [4]-[7].

This paper is organized as follows: section 2 introduces the basic theory of peridynamic model. Hydraulic fracture process is briefly introduced as well. The numerical implement of hydraulic fracture is discussed in section 3, containing the hydraulic fracture pressure applied by crack path tracing scheme developed in this paper. In section 4, several numerical results are presented. And conclusion is given in section 5.

Peridynamic model of hydraulic fracturing

Peridynamic basic theory

Peridynamic theory is a non-local theory which assumes that an arbitrary material point in a body interacts with other material points within its range, called horizon, as shown in Fig 1. It can be regarded as a continuum version of molecular dynamics. Each material point follows Newton's Second Law in the form of

$$\rho(\boldsymbol{x})\ddot{\boldsymbol{u}}(\boldsymbol{x},t)=\int_{H}\boldsymbol{f}\left(\boldsymbol{u}(\boldsymbol{x},t)-\boldsymbol{u}(\boldsymbol{x}',t),\boldsymbol{x}-\boldsymbol{x}'\right)dH+\boldsymbol{b}(\boldsymbol{x},t) \tag{1}$$

in which $\boldsymbol{\rho}$ is the material density. $\boldsymbol{u}(\boldsymbol{x},t)$ is the displacement of point $\boldsymbol{x}$. $\boldsymbol{b}$ is the applied force in the form of body force. H denotes the range of point $\boldsymbol{x}$ can act on, named horizon. All others points in a certain material point's horizon together is called this material point's family. The interaction between two family points is defined as a bond. $\boldsymbol{f}$ describes the internal force between each couple of material points, called pairwise force function. It has the dimension of force per volume squared.

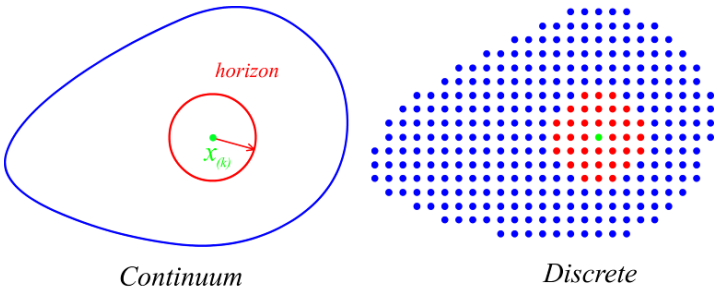


Figure 1. Peridynamic theory

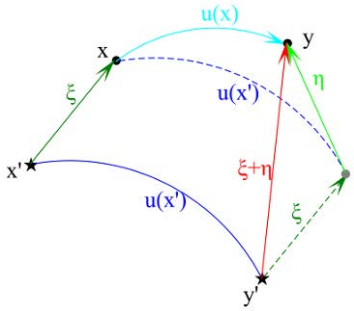


Figure 2. Relationship between relative position vector and relative displacement vector

Pairwise force function can be derived from bond's micro-potential ω. In micro-elastic material, for example:

$$\boldsymbol{f} = (\boldsymbol{\eta}, \boldsymbol{\xi}) = \frac{\partial \omega(\boldsymbol{\eta}, \boldsymbol{\xi})}{\partial \boldsymbol{\eta}} \tag{2}$$

$\boldsymbol{\xi} = \boldsymbol{x} - \boldsymbol{x}'$ is the initial bond vector called relative position vector. $\boldsymbol{\eta} = \boldsymbol{u} - \boldsymbol{u}'$ is the deformed bond vector called relative displacement vector as shown in Figure 2. In this paper, a PMB (prototype micro-elastic brittle) material [8] is considered, defined by

$$\omega(\eta, \xi) = \frac{c(\xi) s^2(\xi)}{2} \tag{3}$$

where $\xi = |\boldsymbol{\xi}|, \eta = |\boldsymbol{\eta}|$ are the magnitudes of the corresponding vectors. s is the bond stretch which represents the elongation of bond defined as

$$s = \frac{\zeta - \xi}{\xi} \tag{4}$$

$\zeta = |\boldsymbol{\zeta}| = |\boldsymbol{\xi} + \boldsymbol{\eta}|$ is the bond vector in the current reference configuration. $c(\xi)$ is a material parameter describes the stiffness of a single bond, called micro-modulus.

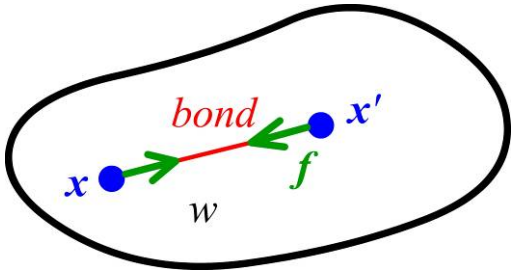


Figure 3. Peridynamic bond and micro-potential

Additionally, the pairwise force function for PMB material is obtained by differentiating the micro-potential

$$f = |\boldsymbol{f}| = cs \tag{5}$$

By Equaling strain energy density in the classical theory and peridynamic theory this peridynamic material parameter can be associated with bulk modulus k in classical theory. For linear elastic material, consider a large homogeneous body under isotropic extension

$$\eta = s\xi \qquad \forall \xi \tag{6}$$

Thus

$$f = cs = c\eta / \xi \tag{7}$$

It follows that

$$\omega = \frac{c\eta^2}{2\xi} = \frac{cs^2\xi}{2} \tag{8}$$

The strain energy density for peridynamic theory can be obtained by integrating the micro-potential associated within a certain point's horizon

$$W = \frac{1}{2}\int_{H_x} \omega(\boldsymbol{\eta}, \boldsymbol{\xi}) dV_\xi \tag{9}$$

1/2 means that each bond's strain energy density is shared by two connected points. Therefore, strain energy density for peridynamic theory can be obtained. For example in 3D

$$W_{PD} = \frac{1}{2}\int_{H_x} \omega(\boldsymbol{\eta}, \boldsymbol{\xi}) dV_\xi = \frac{1}{2}\int_0^\delta \left(\frac{cs^2\xi}{2} \right) dV_\xi \tag{10}$$

where δ is the radius of horizon. $W_{PD} = W_{CM}$ leads to

$$\begin{cases} c = \dfrac{6E}{\pi\delta^3(1-\nu)} & p-\sigma \\ c = \dfrac{6E}{\pi\delta^3(1-2\nu)(1+\nu)} & p-\varepsilon \\ c = \dfrac{12E}{\pi\delta^4} & 3D \end{cases} \tag{11}$$

where E is the elastic modulus, ν is the possion ratio [9]. In the bond-based peridynamic theory

$$\nu = \begin{cases} \dfrac{1}{3} & 2D \\ \dfrac{1}{4} & 3D \end{cases} \tag{12}$$

Failure is introduced at bond level. A bond breaks when its elongation s exceeds the critical relative elongation s_0

$$\begin{cases} \boldsymbol{f} = cs\dfrac{\boldsymbol{\xi}+\boldsymbol{\eta}}{|\boldsymbol{\xi}+\boldsymbol{\eta}|} & s \le s_0 \\ \boldsymbol{f} = \boldsymbol{0} & s > s_0 \end{cases} \tag{13}$$

Once a bond is broken, it stays broken. That means bonds can't heal. And it makes the model historical dependent.

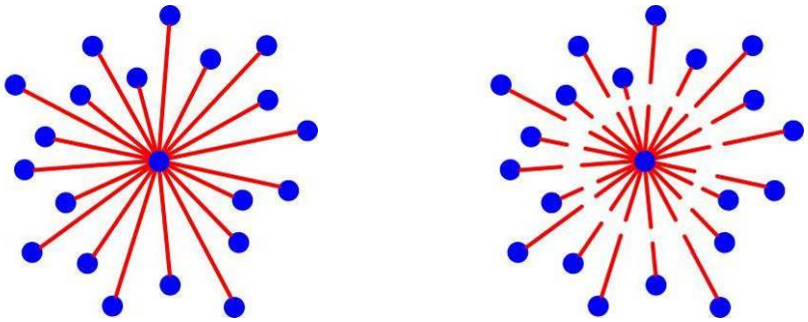


Figure 4. Damage status

Additionally, a quantity describing the damage status of a certain point is defined by

$$\varphi(x,t) = 1 - \frac{\int_H \mu(\mathbf{x},t,\xi)dV_h}{\int_H dV_h} \tag{14}$$

Simulation of hydraulic fracture via peridynamic

The hydraulic fracture process is shown in Figure 5. Fracking fluid is injected after perforating on the wall of horizontal well. Horizontal well is the one of the most widely used technique in shale gas production process because of its long length in horizontal direction. The difficulty in simulating horizontal well fracking is obvious: it involves massive cracks. Thus peridynamic theory mentioned above is applied for simulation in this paper.

Material parameters used are adopted from Nongan oil shale field in Jilin Province of China as listed in Table 1 below. Perforations are applied as initial cracks by cutting off all the bonds crossing the initial crack paths.

Table 1. Shale parameter

E(Gpa)	Density(kg/m^3)	Possion ratio	s_0
19.03	2400	1/3	0.00003

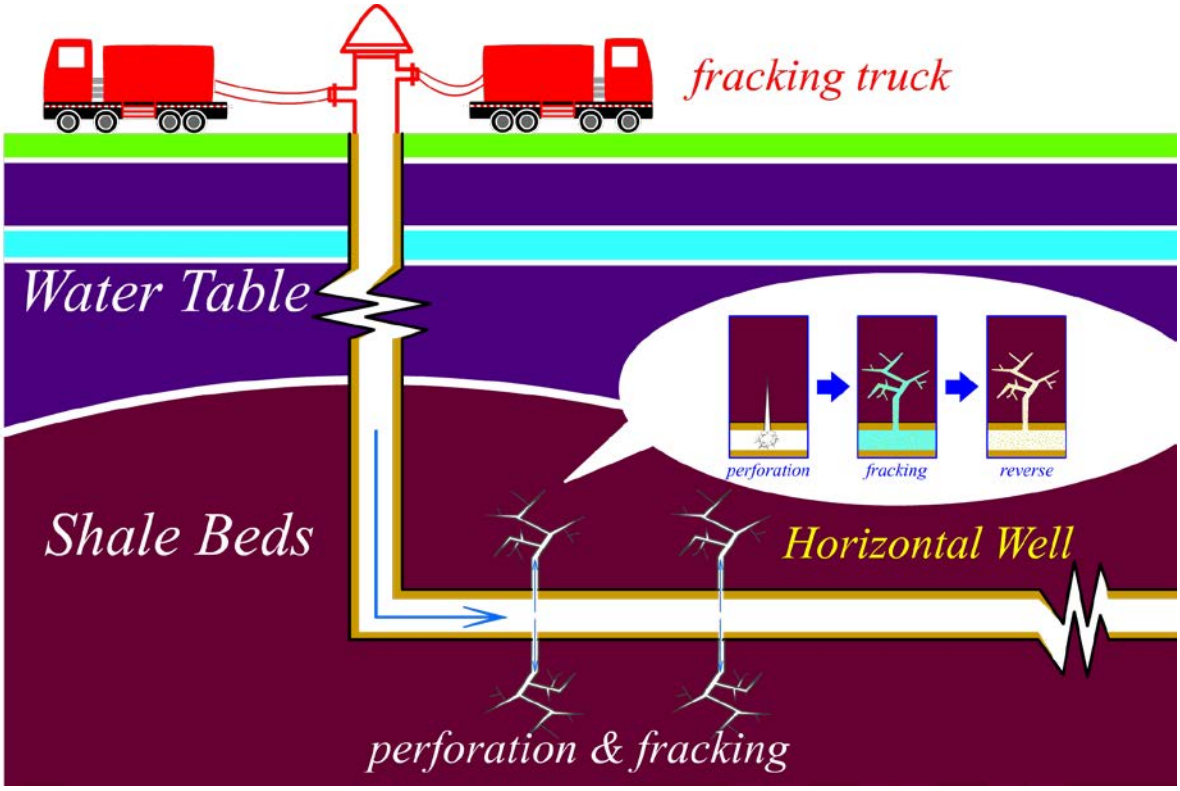


Figure 5. Schematic of hydraulic fracture process

Numerical implementation

Time integration

Velocity-Verlet algorithm [10] is adopted here as time integration, which is an explicit method. The Velocity-Verlet algorithm is

$$\begin{aligned} \dot{\boldsymbol{u}}_{n+\frac{1}{2}} &= \dot{\boldsymbol{u}}_n + \frac{\Delta t}{2}\ddot{\boldsymbol{u}}_n \\ \boldsymbol{u}_{n+1} &= \boldsymbol{u}_n + \Delta t\dot{\boldsymbol{u}}_{n+\frac{1}{2}} \\ \dot{\boldsymbol{u}}_{n+1} &= \dot{\boldsymbol{u}}_{n+\frac{1}{2}} + \frac{\Delta t}{2}\ddot{\boldsymbol{u}}_{n+1} \end{aligned} \tag{15}$$

where $\ddot{\boldsymbol{u}}$, $\dot{\boldsymbol{u}}$ and $\boldsymbol{u}$ are acceleration, velocity and displacement respectively.

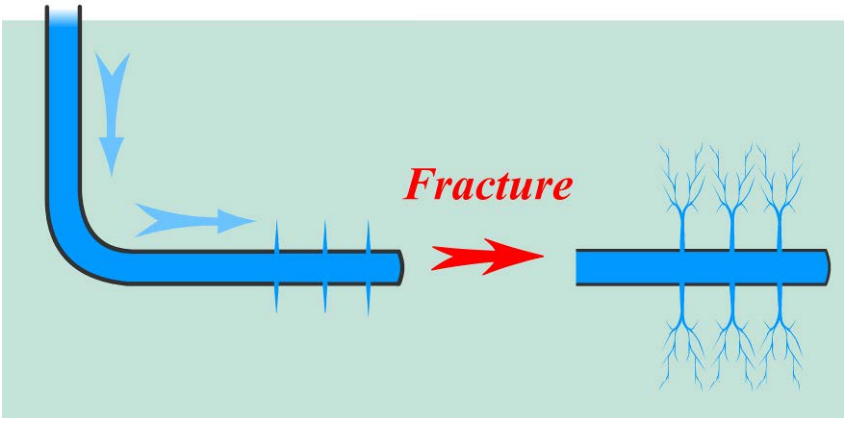


Figure 6. Schematic of crack network caused by hydraulic fracture

Hydraulic fracture pressure applied by crack path tracing scheme

One of the most important features of hydraulic fracture is that hydraulic fracture pressure needs to be applied on both initial cracks and new born cracks, as shown as in Fig 6. Thus a new scheme is proposed in this paper to trace the hydraulic fracture crack paths in peridynamic model, and the corresponding normal directions can also be obtained. Cracks might be very chaotic when cracks branching, especially fragmentation occurred, as Fig 7. To find the effective material points on crack paths, points with damage at certain level are adopted in this paper:

$$0.35 < \varphi < 0.50 \tag{16}$$

And the normal direction of a certain point on the crack path illustrated in Fig 8 is calculated by weighted average of its associated bonds' lengths

$$\alpha_b(\xi) = \sum_{i=1}^{\infty} p_i \left(\frac{\xi}{\Delta x} \right)^{i-1} \tag{17}$$

in which p_i is the coefficient of fitting polynomial. Consequently the corresponding direction of a certain point i can be obtained by

$$\begin{aligned} \cos\theta_i &= \sum_{b}^{\infty} \alpha_b \cdot \cos\theta_b \\ \sin\theta_i &= \sum_{b}^{\infty} \alpha_b \cdot \sin\theta_b \end{aligned} \tag{18}$$

where the subscript b donate the bond's number associated with point i.

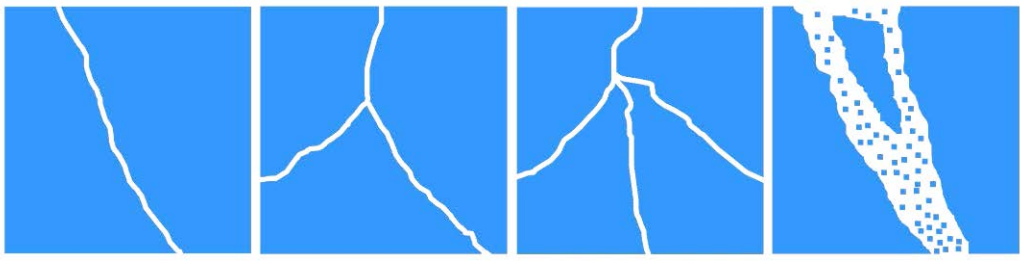


Figure 7. Schematic of several kinds of crack path

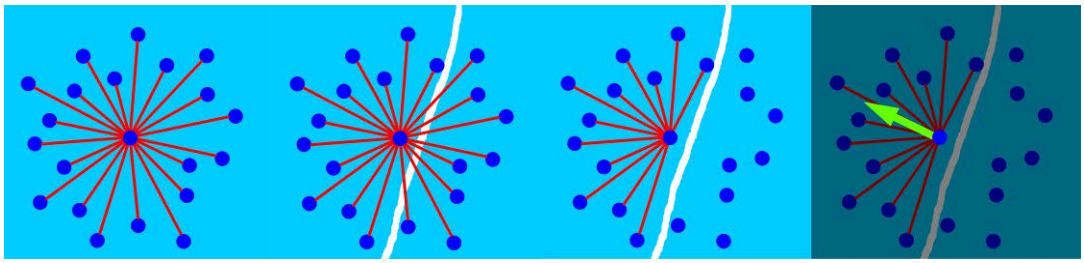


Figure 8. Determination of hydraulic fracture pressure direction

Results

In this section, longitudinal section of horizontal well is studied as shown in Figure 9. Half of the longitudinal section is omitted because of its symmetry. The domain is $40m \times 20m$, with perforations on the bottom wall of horizontal well. The hydraulic fracture pressure applied is shown in Figure 10 which is increasing over time. Time step is $\Delta t = 10^{-5} s$, perforation depth is $a = 1.0m$, and the mesh size is $\Delta x = 0.1m$.

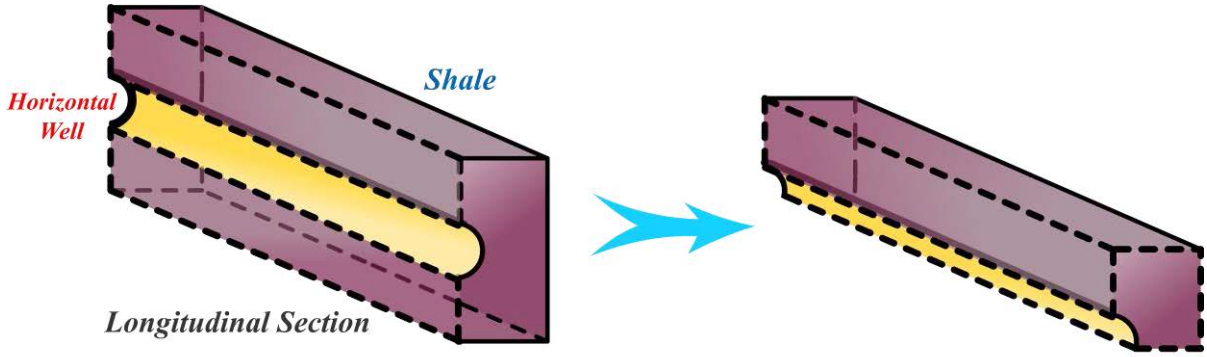


Figure 9. Schematic of longitudinal section of horizontal well

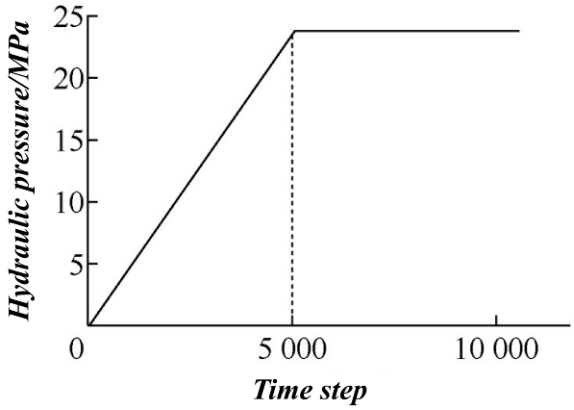


Figure 10. Fracking load

Numerical results of different numbers of perforations

The influence of perforations is studied here. One, two and three perforations on the wall of horizontal wells are set, as in Figure 11. The simulation results are shown in Figure 12. Cracks branching and secondary branching path occur. It proves that peridynamic model is convenient to deal with hydraulic fracture problem.

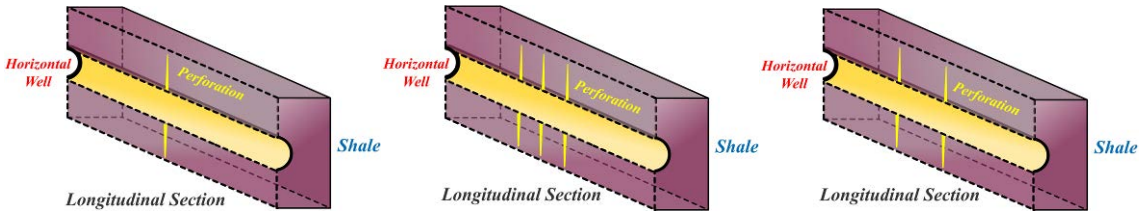


Figure 11. Schematic of horizontal well with different numbers of perforations

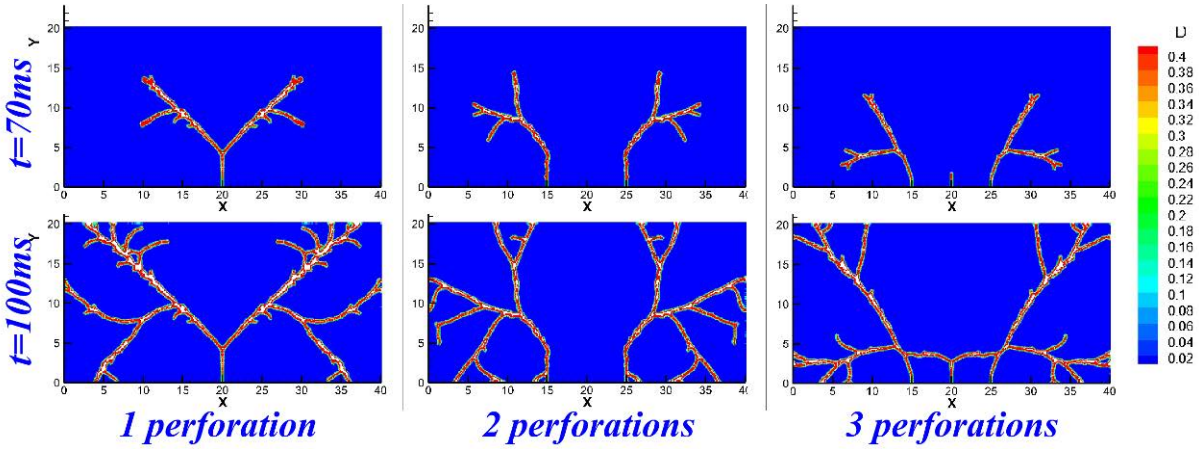


Figure 12. Different numbers of perforations fracking result

Additionally, it illustrates that in the hydraulic fracture process cracks can be restrained by each other when distance between cracks reduces during propagation. Figure 13 shows the relationship between time step and crack paths length which is most concerned. The crack paths lengths are represented by the numbers of material points applied in hydraulic fracture pressure. An interesting phenomenon is that the single perforation crack paths length reaches the maximum at last and nearly the same order of magnitude as the others all the time. Thus it proves that simply addition of perforation number in a certain area can't increase the crack path length proportionally. It's not wise to set perforations too dense to get better fracking results.

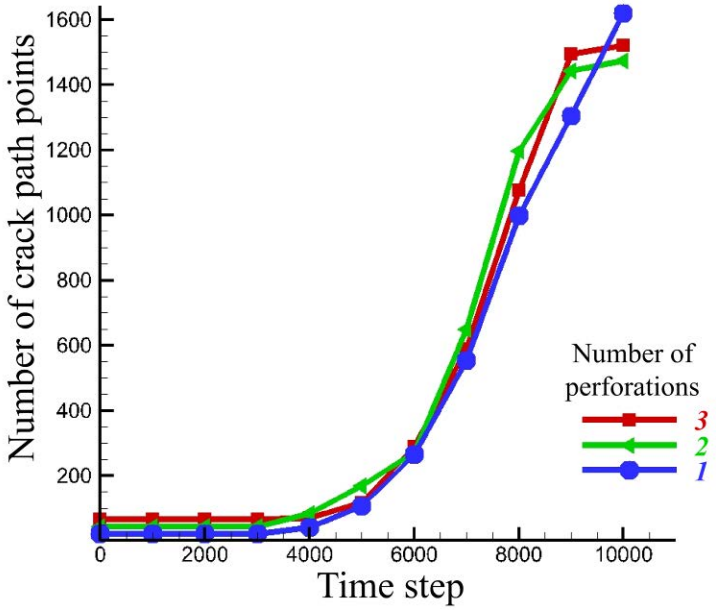


Figure 13. Relationship between time and crack path length

Numerical results of various angles of initial natural cracks

Furthermore, cases of two perforations with different angle of initial natural cracks fracking are investigated to discuss the influence of initial natural cracks. The computational domain is same as above. Fracking pressure applied, mesh size, perforation depth, time step are same as the previous case, too.

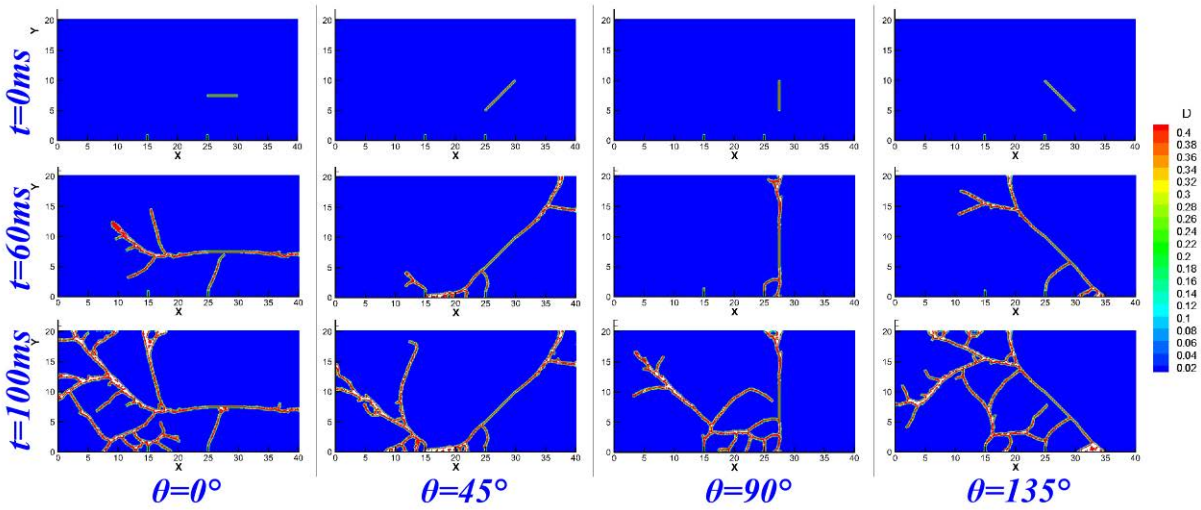


Figure 14. Various angles of initial natural cracks fracking result

The results are shown above in Figure 14. It reveals different angles of initial natural cracks influence the crack results remarkably. As shown in Figure 15 at $t = 100ms$ 0° natural crack case reaches the maximum by 1496 crack path points, while case with 45° natural crack gets the minimum crack path points by 1000. Since the shale is lamellar structure full of natural cracks with certain angles, investigation of the geological structure would be particularly important.

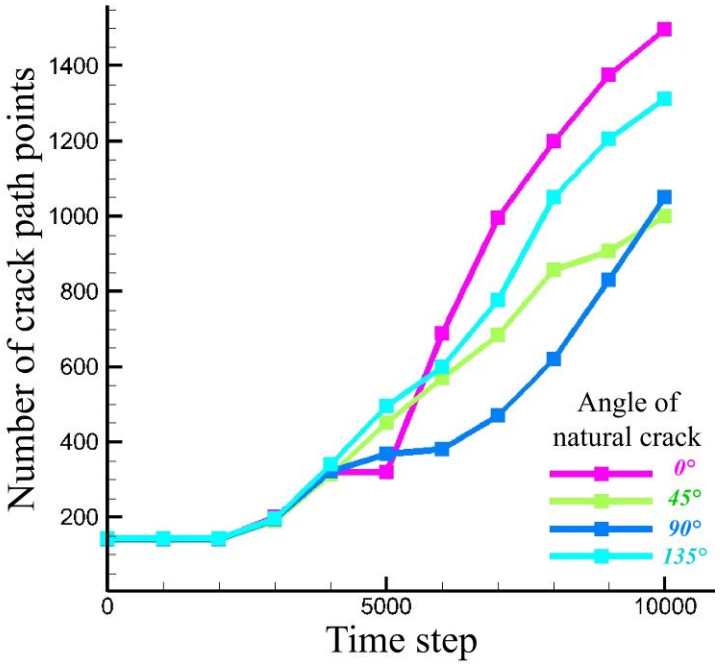


Figure 15. Relationship between time and crack path length

Conclusions

Peridynamic is a powerful method to deal with the problem involving massive cracks. In this paper, hydraulic fracture process is simulated by peridynamic model we developed. With the crack path tracing scheme we developed, crack branching and coalescence are observed in the numerical results which prove it is convenient and suitable for hydraulic fracture problems. The results also revealed that hydraulic fracture perforation numbers at a certain area are closely associated with crack patterns. And perforations restrain the cracks propagation when they set close enough. That means increasing the number of perforation simply in a certain area can't guarantees the increasing of crack proportionally. Besides, initial natural cracks distances remarkably influence the fracking results. Therefore, investigation of geologic structure is significant meaningful for hydraulic process before engineering construction.

Acknowledgement

The authors are pleased to acknowledge the support of this work by Science Challenge Project (No. JCKY2016212A502), the National Key Research and Development Program of China (2016YFC1402705, 2016YFC1402706), the National Natural Science Foundation of China (11372066, 11232003), the Fundamental Research Funds for the Central Universities (DUT15LK07), the open funds of the state key laboratory of water resources and hydropower engineering science (2015SGG03), the open funds of the state key laboratory of Geohazard Prevention and Geoenvironment Protection (SKLGP2016K007).

References

[1] Bowker, K. A. (2007) Barnett Shale gas production, Fort Worth Basin: Issues and discussion, Aapg Bulletin, **91**,523-533.

[2] Silling, S. A.. (2000) *Reformulation of elasticity theory for discontinuities and long-range forces*, Journal of the Mechanics & Physics of Solids **48,**175-209.

[3] Lee, J., Liu, W., and Hong, J. W. (2005) *Impact fracture analysis enhanced by contact of peridynamic and finite element formulations*, International Journal of Impact Engineering, **87**,108-119.

[4] Wu, F., Li, S. H., Duan, Q. L. and Li, X. K. (2017) *Application of the Method of Peridynamics to the Simulation of Hydraulic Fracturing Process*, Proceedings of the 7th International Conference on Discrete Element Methods. Springer Singapore.

[5] Wu F, Li S, Duan Q, Li X, and Zhang H. W. (2017) Numerical simulation of hydraulic fracturing process based on the method of peridynamics , Computer Aided Engineering, **26**(1),1-6

[6] Nadimi, S., Miscovic, I. and Mclennan, J. (2016) *A 3D peridynamic simulation of hydraulic fracture process in a heterogeneous medium,* Journal of Petroleum Science & Engineering, **145**,444-452.

[7] Ouchi, H., Katiyar, A., Foster, J. T. and Sharma, M. M. (2015) *A Peridynamics Model for the Propagation of Hydraulic Fractures in Heterogeneous, Naturally Fractured Reservoirs*, SPE Hydraulic Fracturing Technology Conference, Texas.

[8] Silling, S. A. and Askari, E. (2005) *A meshfree method based on the peridynamic model of solid mechanics*. Computers & Structures, **83**(17–18),1526-1535.

[9] Shojaei, A., Mudric, T., Zaccariotto, M. and Galvanetto, U. (2016) *A coupled meshless finite point/Peridynamic method for 2D dynamic fracture analysis*. International Journal of Mechanical Sciences, **119**,419-431.

[10] Hairer E, Lubich C and Wanner G. (2003) *Geometric numerical integration illustrated by the Störmer-Verlet method*. Acta Numerica, **12**(12), 399--450.

Computational method for geometric properties of arbitrary plane areas

†X. L. Chen[1], D. Y. Li[1], and L. Zhang[1]

[1]School of Mathematics and Physics, Chongqing University of Science and Technology, Chongqing, China.
†Corresponding author: xlchen80@163.com

Abstract

The centroid, the moment of inertia, and the product of inertia for an arbitrary polygon and an arbitrary plane area, are formulated in algebraic forms and programmed with Matlab software. Three numerical examples are shown and indicate the method is valid for geometric properties of simply or multiply connected plane areas with arbitrary boundaries.

Keywords: Centroid; moment of inertia; product of inertia; arbitrary boundaries; program

Introduction

Geometric properties of an area, including the centroid, the moment of inertia, the polar moment of inertia, and the product of inertia, are important quantities used in mechanics of materials, structural mechanics, fluid mechanics, and machine design, hence their computation by the computer is helpful for the engineer. Wen and Zhang [1] studied geometric properties of an arbitrary triangle and applied to convex polygons. Based on calculating the geometrical characteristic values of a triangle with a vertex on the coordinate origin, Cai [2] suggested a regular method for calculating geometrical and mechanical characteristic values of sections for structural bodies in a shape of prism.

Computational Method for Geometric Properties of an Arbitrary Polygon

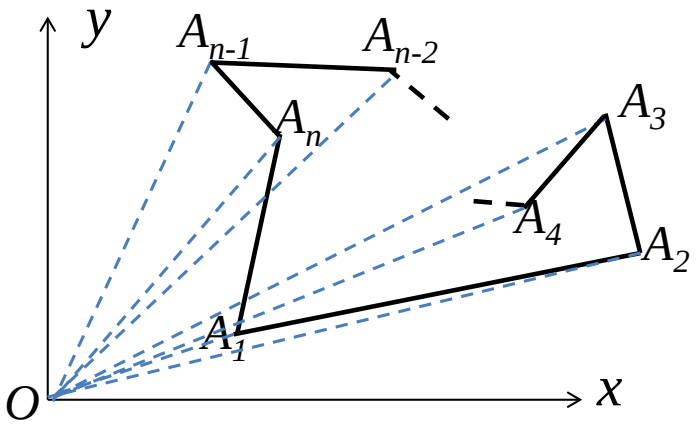


Figure 1. An arbitrary polygon

Consider an arbitrary polygon, shown in Fig. 1, which lies in the *x*-*y* plane. It should be noticed that the vertexes $A_i(i=1,2,...,n)$ must be continuously numbered with a counterclockwise direction, and A_{n+1} is assumed to coincide with A_1. x_i and y_i are the coordinates of the vertex A_i. S_i is the algebraic area of ΔOA_iA_{i+1} and can be expressed as Eq. (1). Since the area of a polygon is the sum of S_i, substituting geometric properties of an arbitrary triangle shown in the reference [2] into the general integral definition of geometric properties, we have geometric properties of an arbitrary polygon shown as Eq. (2)- Eq. (5), where x_C and y_C are the coordinates of the centroid C, I_x, I_y, and I_{xy} are the moment of inertia and the product of inertia respectively.

$$S_i = (x_i y_{i+1} - x_{i+1} y_i)/2 \quad (i = 1, 2, ..., n) \tag{1}$$

$$x_C = \frac{\sum_{i=1}^{n} S_i \cdot \frac{x_i + x_{i+1}}{3}}{\sum_{i=1}^{n} S_i}, y_C = \frac{\sum_{i=1}^{n} S_i \cdot \frac{y_i + y_{i+1}}{3}}{\sum_{i=1}^{n} S_i} \tag{2}$$

$$I_x = \sum_{i=1}^{n} S_i \cdot \frac{y_i^2 + y_i y_{i+1} + y_{i+1}^2}{6} \tag{3}$$

$$I_y = \sum_{i=1}^{n} S_i \cdot \frac{x_i^2 + x_i x_{i+1} + x_{i+1}^2}{6} \tag{4}$$

$$I_{xy} = \sum_{i=1}^{n} S_i \cdot \frac{2(x_i y_i + x_{i+1} y_{i+1}) + x_i y_{i+1} + x_{i+1} y_i}{12} \tag{5}$$

Matlab Program

The Matlab Program for the computation of area, centroid, moment of inertia, polar moment of inertia, and product of inertia for an arbitrary polygon is as follows:

```
clear;clc;
xy=load('coordinates.txt') ;          % vertex coordinate matrix
n=length(xy);
xy=[xy;xy(1,:)];                      % An+1=A1
for i=1:n
s(i)=xy(i,1)*xy(i+1,2)-xy(i+1,1)*xy(i,2);
xci(i)=(xy(i,1)+xy(i+1,1));
yci(i)=(xy(i,2)+xy(i+1,2));
ixi(i)=xy(i,1)^2+xy(i,1)*xy(i+1,1)+xy(i+1,1)^2;
iyi(i)=xy(i,2)^2+xy(i,2)*xy(i+1,2)+xy(i+1,2)^2;
ixyi(i)=2*(xy(i,1)*xy(i,2)+xy(i+1,1)*xy(i+1,2))+xy(i,1)*xy(i+1,2)+xy(i+1,1)*xy(i,2);
end
S=sum(s)/2                            % area of an arbitrary polygon
xc=sum(s.*xci)/area/2/3               % the centroid
yc=sum(s.*yci)/area/2/3               % the centroid
Ix=sum(s.*iyi)/12                     % the moment of inertia about x-axis
Iy=sum(s.*ixi)/12                     % the moment of inertia about y-axis
Ip=Ix+Iy                              % the polar moment of inertia
Ixy=sum(s.*ixyi)/24                   % the product of inertia
```

Numerical Examples

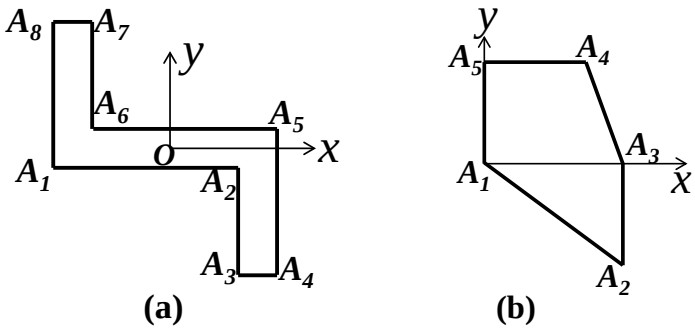


Figure 2. Polygonal cross-sectional areas

In Fig. 2a, the coordinates of the vertexes are $A_1(-300,-50)$, $A_2(200,-50)$, $A_3(200,-350)$, $A_4(300,-350)$, $A_5(300,50)$, $A_6(-200,50)$, $A_7(-200,350)$, and $A_8(-300,350)$, where all the length units are mm. Substituting these coordinates into Eq. (1)- Eq. (5), geometric properties of the polygon can be solved by the above Matlab program. Results are the area $S=120000\,mm^2$, the centroid $x_C=0$ and $y_C=0$, the moment of inertia $I_x=2.9\text{e}9\,\text{mm}^4$ and $I_y=5.6\text{e}9\,\text{mm}^4$, the polar moment of inertia $I_O=8.5\text{e}9\,\text{mm}^4$, and the product of inertia $I_{xy}=\text{-}3\text{e}9\,\text{mm}^4$. In Fig. 2b, the coordinates of the vertexes are $A_1(0,0)$, $A_2(9,-6)$, $A_3(9,0)$, $A_4(6,6)$, and $A_5(0,6)$, where all the length units are mm. Substituting these coordinates into Eq. (1)- Eq. (5), geometric properties of the polygon can be solved by the above Matlab program. Results are the area $S=72\,mm^2$, the centroid $x_C=4.625\,mm$ and $y_C=1\,mm$, the moment of inertia $I_x=648\,\text{mm}^4$ and $I_y=1971\,\text{mm}^4$, and the product of inertia $I_{xy}=81\,\text{mm}^4$. The above results agree with the reference [3].

Computational Method for Geometric Properties of Arbitrary Plane Areas

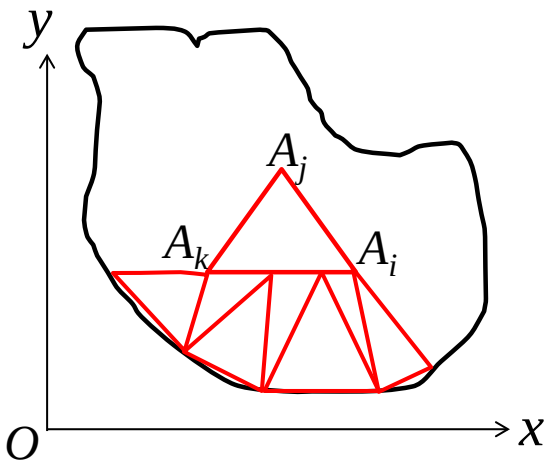


Figure 3. An arbitrary plane area

As shown in Fig. 3, an arbitrary plane area can be easily meshed into a finite number of triangles by the software such as ANSYS, and the coordinates of nodes can be automatically obtained. Substituting geometric properties of an arbitrary triangle shown in the reference [1] into the general integral definition of geometric properties, we have geometric properties of an arbitrary plane area shown as Eq. (6)- Eq. (10), where S_i is the area of meshed triangle

$\Delta A_i A_j A_k$, x_C and y_C are the coordinates of the centroid C, I_x, I_y, and I_{xy} are the moment of inertia and the product of inertia respectively.

$$S_i = \frac{1}{2}\left|x_i(y_j - y_k) + x_j(y_k - y_i) + x_k(y_i - y_j)\right| \tag{6}$$

$$x_C = \frac{\sum_{i=1}^{n} S_i \cdot \frac{x_i + x_j + x_k}{3}}{\sum_{i=1}^{n} S_i}, y_C = \frac{\sum_{i=1}^{n} S_i \cdot \frac{y_i + y_j + y_k}{3}}{\sum_{i=1}^{n} S_i} \tag{7}$$

$$I_x = \sum_{i=1}^{n} S_i \cdot \frac{y_i^2 + y_j^2 + y_k^2 + y_i y_j + y_k y_j + y_i y_k}{6} \tag{8}$$

$$I_y = \sum_{i=1}^{n} S_i \cdot \frac{x_i^2 + x_j^2 + x_k^2 + x_i x_j + x_k x_j + x_i x_k}{6} \tag{9}$$

$$I_{xy} = \sum_{i=1}^{n} S_i \cdot \frac{2(x_i y_i + x_j y_j + x_k y_k) + x_i(y_j + y_k) + x_j(y_k + y_i) + x_k(y_i + y_j)}{12} \tag{10}$$

Matlab Program

With the triangular elements and coordinates of nodes obtained by the ANSYS software, the Matlab Program for the computation of area, centroid, moment of inertia, polar moment of inertia, and product of inertia for an arbitrary plane area is as follows:

```
clear;clc;
node=load('node.txt');   % coordinates of nodes
b=load('ele.txt');       % triangular elements
n=length(b);
for m=1:n
  xi=node(b(m,7),2);yi=node(b(m,7),3);
  xj=node(b(m,8),2);yj=node(b(m,8),3);
  xk=node(b(m,9),2);yk=node(b(m,9),3);
s(m)=abs(xi*(yj-yk)+xj*(yk-yi)+xk*(yi-yj));
xci(m)=xi+xj+xk;
yci(m)=yi+yj+yk;
ixi(m)=yi^2+yj^2+yk^2+yi*yj+yj*yk+yk*yi;
iyi(m)=xi^2+xj^2+xk^2+xi*xj+xj*xk+xk*xi;
ixyi(m)=2*(xi*yi+xj*yj+xk*yk)+xi*(yj+yk)+xj*(yk+yi)+xk*(yi+yj);
end
S=sum(s)/2                  % area of an arbitrary plane area
xc=sum(s.*xci)/area/2/3     % the centroid
yc=sum(s.*yci)/area/2/3     % the centroid
Ix=sum(s.*iyi)/12           % the moment of inertia about x-axis
Iy=sum(s.*ixi)/12           % the moment of inertia about y-axis
Ip=Ix+Iy                    % the polar moment of inertia
Ixy=sum(s.*ixyi)/24         % the product of inertia
```

Numerical Examples

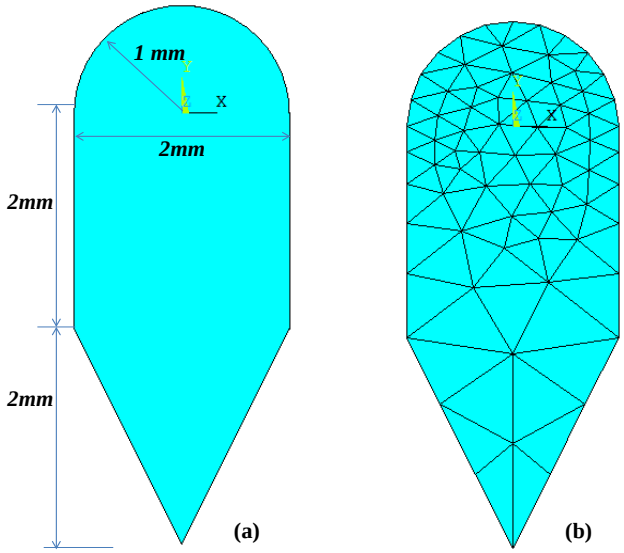


Figure 4. General cross-sectional areas

In Fig. 4, a general cross-sectional area is meshed into 112 triangular elements by the ANSYS software, and the coordinates of nodes can also be obtained by the software. Substituting these coordinates into Eq. (6)- Eq. (10), geometric properties can be solved by the above Matlab program. Results are the area $S = 7.5529\,mm^2$, the centroid $x_C = \text{-}1.1637\text{e-}17\,mm$ and $y_C = \text{-}1.1490\,mm$, the moment of inertia $I_x = 2.0505\,\text{mm}^4$ and $I_y = 20.3838\,\text{mm}^4$, the polar moment of inertia $I_O = 22.4343\,\text{mm}^4$, and the product of inertia $I_{xy} = 1.8619\text{e-}016\,\text{mm}^4$. While both the theoretical values of x_C and I_{xy} should be zero, it indicates numerical computation exits the rounding error.

Conclusions

Eq. (1)- Eq. (5) are exact for geometric properties of arbitrary simply connected polygons with straight edges. Eq. (6)- Eq. (10) can be applied to compute plane areas with arbitrary boundaries. They are easy to be executed by the computer program, however there may exits some little rounding errors. For the multiply connected polygon, it can be considered as a combination of some simply connected polygons and solved by Eq. (1)- Eq. (5), or solved directly by Eq. (6)- Eq. (10).

Acknowledgement

Project Supported by Scientific and Technological Research Program of Chongqing Municipal Education Commission (Grant No. KJ1713340).

References

[1] Wen, Z. M., Zhang, L. S. (2001) Using the unit of coordinate to find moment of area, the moment of inertia and the product of inertia, *Journal of Southern Institute of Metallurgy* **22**, 293–295.
[2] Cai, J. B. (1995) The regular method of calculating the geometrical and mechanical characteristic values of prismatic section. *Acta univ. agric. boreali-occidentalis* **23**, 69–73.
[3] Hibbeler, R. C. (2004) *Engineering Mechanics: Statics*, 10th edn, Higher Education Press, Beijing, China.

Coupling immersed boundary-lattice Boltzmann method with smoothed point interpolation method for large-displacement fluid-structure interaction problems

***S.Q. Wang[1], Y.N. Cai[1], †G.Y. Zhang[1,2,3] , S. Li[1,2,3], and J.H. Lu[1,3]**

[1]Liaoning Engineering Laboratory for Deep-Sea Floating Structures, School of Naval Architecture, Dalian University of Technology, Dalian, 116024, P. R. China

[2]State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, 116024, P. R. China

[3]Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai, 200240, P. R. China

*Presenting author: wangsqyzx@163.com

†Corresponding author: gyzhang@dlut.edu.cn

Abstract

The immersed boundary-lattice Boltzmann method (IB-LBM) has been verified to be an effective tool for fluid-structure interaction (FSI) simulation associated with thin and flexible body, and the newly developed smoothed point interpolation method (S-PIM) can handle the largely deformable solids owing to its soften model stiffness of model and insensitivity to mesh distortion. In this work, a novel method has been proposed in the present work by coupling IB-LBM with S-PIM for FSI problems with large-displacement solids. The proposed method preserves the efficiency of LBM for fluid solver, utilizes S-PIM to establish the realistic constitutive laws for nonlinear solids, and avoids the mesh regeneration based on the frame of immersed boundary method (IBM). Three benchmarking examples have been carried out to validate the accuracy, convergence and stability of the proposed method in consideration of comparative results with referenced solutions.

Keywords: immersed boundary-lattice Boltzmann method; smoothed point interpolation method; fluid-structure interaction (FSI); constitutive laws

Introduction

Fluid-structure interaction problems (FSI) are common in many fields in light of two-phase interaction when the movement or distortion of the body arises due to the external force exerted by the neighboring fluid and the corresponding response also has an impact on the fluid domain. The complicated mechanism requires an efficient and reliable tool especially for largely deformable nonlinear solids/structures. The immersed boundary method (IBM) has been successfully applied in the simulation of interaction between the fluid and the moving interface, and the mesh regeneration can be avoided using non-boundary-fitted grid [1]. It assumes the fluid around the interface is affected by a kind of body force which allows the solver of Navier-Stokes equations based on the fixed Euler grid, and the configuration together with shape of the boundary would not be taken into consideration. Hence, the solver process for FSI problems has been simplified in comparison with the boundary-fitted grid

method.

Generally, one can use the finite element method (FEM) or finite volume method (FVM) to handle the Navier-Stokes equations in FSI simulation. However, the solver for the nonlinear convection term may bring about the nonphysical numerical oscillation and the special format of FEM or FVM need be constructed such as pressure-stabilized Petrov-Galerkin (PSPG) formulation [2] and characteristic-based split (CBS) algorithm [3] which may consume additional calculation resource. And the pressure Poisson equation should be considered for the incompressible viscid flow based on these traditional methods which also affect the calculation efficiency. Then the lattice Boltzmann method (LBM) was introduced and has been widely used in the CFD field during the past two decades. Based on the explicit algorithm, LBM can avoid the shortcoming of FEM and FVM and guarantee the accuracy and efficiency with the simple form as well as easy operation. And the immersed boundary-lattice Boltzmann method (IB-LBM) was proposed for FSI simulation in consideration of the same discretization using Cartesian mesh in IBM and LBM [4]. The further development and improvement have been achieved for various FSI problems [5-7].

FEM is a popular solver for the transient analysis of nonlinear solids and structures. And the simple triangular element with three nodes (T3) for 2D or tetrahedron element with four nodes (T4) for 3D are very suitable for the preprocessing especially encountered with complex domain. However, some disadvantages have limited the extensive application like the overly-stiff performance and the poor accuracy. Then a class of gradient/stain smoothing methods have been proposed to improve the performance of T3/T4 cells including smoothed finite element methods (S-FEM) [8, 9] and smoothed point interpolation methods (S-PIM) [10, 11]. Compared to S-FEM, S-PIM allows the discontinuous displacement function in the smoothed domain by constructing a weakened-weak form in G space. Based on different smoothed domains, ES-PIM and NS-PIM can be constructed and used for linear and nonlinear analysis.

IB-LBM has been verified as an efficient tool for FSI simulation and S-PIM also can handle the complex nonlinear constitutive equation well. Coupling IB-LBM with S-PIM, this article puts forward a novel method for FSI problems involved with the large deformable nonlinear solids and structures. And Figure 1 has shown the general procedure. The fluid and solid can be separated from the FSI system, and solved based on fixed Euler grids by LBM and moving lagrange grids by S-PIM. The coupling force has been calculated by the frame of IBM and the information exchange can be implemented by the interpolation of the delta function.

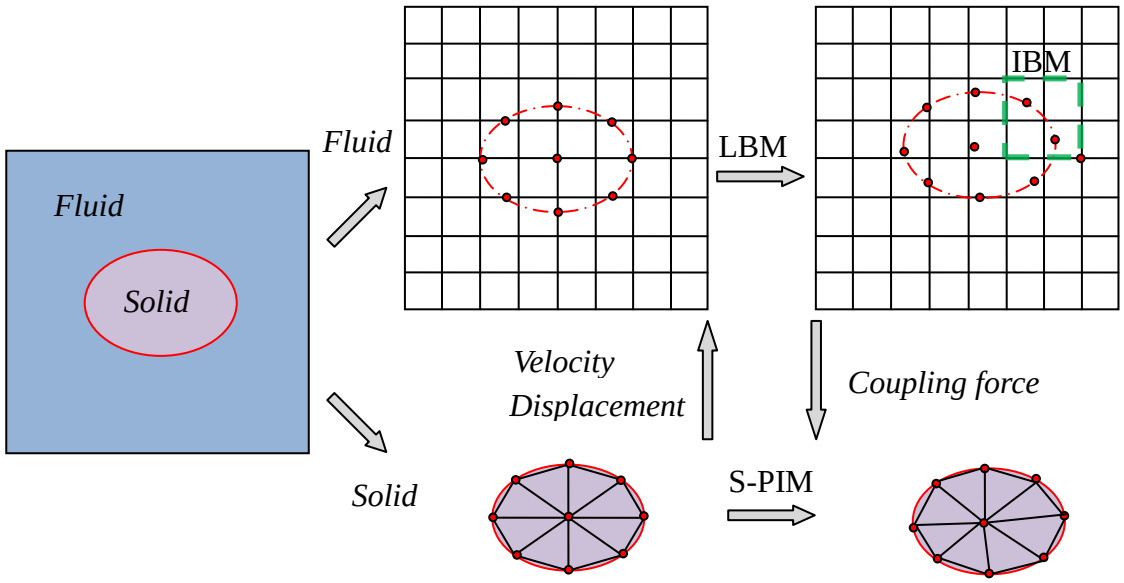


Figure 1 The general procedure of coupling S-PIM with IB-LBM

Immersed boundary-lattice Boltzmann method (IB-LBM)

For the incompressible fluid flows, the evolution equation of LBE [12] can be written as:

$$f_\alpha(\mathbf{x}+\mathbf{c}_\alpha\delta t, t+\delta t)-f_\alpha(\mathbf{x},t)=-\frac{f_\alpha(\mathbf{x},t)-f_\alpha^{eq}(\mathbf{x},t)}{\tau}+F_\alpha\delta t \tag{1}$$

where $\mathbf{x}$ is the lattice coordinate, $\mathbf{c}_\alpha$ are the velocities of particles, f_α is the distribution function of particles, f_α^{eq} is the corresponding equilibrium distribution, τ is the relaxation time, and F_α is the discrete force distribution function.

The particle velocity $\mathbf{c}_\alpha$ is defined as follows using D2Q9 model [13]:

$$\mathbf{c}_\alpha=\begin{cases}(0,0), & \alpha=0,\\ (\cos[(\alpha-1)\pi/2],\sin[(\alpha-1)\pi/2)c, & \alpha=1,2,3,4,\\ (\cos[(\alpha-1)\pi/2],\sin[(\alpha-1)\pi/2)c, & \alpha=5,6,7,8.\end{cases} \tag{2}$$

where the velocity $c=\delta x/\delta t$.

The equilibrium distribution function f_α^{eq} can be written in the following form:

$$f_\alpha^{eq}=\rho w_\alpha[1+\frac{\mathbf{c}_\alpha\cdot\mathbf{u}}{c_s^2}+\frac{(\mathbf{c}_\alpha\cdot\mathbf{u})^2}{2c_s^4}-\frac{\mathbf{u}^2}{2c_s^2}] \tag{3}$$

where the sound velocity $c_s=c/\sqrt{3}$, and the factor w_α is selected as $w_0=0$, $w_{1-4}=1/9$ and $w_{5-8}=1/36$.

The discrete force distribution function F_α can be expressed as follows:

$$F_\alpha = (1-\frac{1}{2\tau})\omega_\alpha[\frac{\mathbf{c}_\alpha - \mathbf{u}}{c_s^2} + \frac{\mathbf{c}_\alpha \cdot \mathbf{u}}{c_s^4}\mathbf{c}_\alpha]\cdot\mathbf{f} \tag{4}$$

where $\mathbf{f}$ is the external force.

By means of the mesoscopic model, the macroscopic variables like density and velocity can be obtained as:

$$\rho = \sum_{\alpha=0}^{8} f_\alpha \tag{5}$$

$$\rho\mathbf{u} = \sum_{\alpha=0}^{8} c_\alpha f_\alpha + \frac{1}{2}\mathbf{f}\delta t \tag{6}$$

The kinematic viscosity ν is determined by:

$$\nu = c_s^2(\tau - \frac{1}{2})\delta t \tag{7}$$

There are several measures to evaluate the boundary force, and the present work adopts the direct forcing technique which was illustrated in Kang's paper [14]. The boundary force density at the *m*th Lagrangian point $\mathbf{X}_m$ can be obtained as follows:

$$\mathbf{F}(\mathbf{X}_m,t) = 2\rho(\mathbf{U}^d(\mathbf{X}_m,t) - \mathbf{U}^*(\mathbf{X}_m,t))/\delta t \tag{8}$$

where $\mathbf{U}^d$ is the solid boundary velocity, and $\mathbf{U}^*$ is the evolution velocity of fluid in Lagrangian mesh without force modification which can be interpolated from the neighboring Eulerian points:

$$\mathbf{U}^*(\mathbf{X}_m,t) = \sum_{(i,j)} \mathbf{u}^*(\mathbf{x}_{ij},t)\delta_h(\mathbf{x}_{ij} - \mathbf{X}_m)h^2 \tag{9}$$

where $\mathbf{x}_{ij}$ is the Euler node coordinate, δ_h is a continuous kernel distribution to approximate the delta function, and h is the mesh size. $\mathbf{u}^*$ is the evolution velocity in Eulerian points and evaluated with the following formula:

$$\rho\mathbf{u}^* = \sum_{\alpha=0}^{8} c_\alpha f_\alpha \tag{10}$$

Once the boundary force density of the Lagrangian points is given, it can be distributed into the Eulerian points around it:

$$\mathbf{f}(\mathbf{x}_{ij},t) = \sum_{b=0}^{N} \mathbf{F}(\mathbf{X}_b,t)\delta_h(\mathbf{x}_{ij} - \mathbf{X}_b)h\Delta s \tag{11}$$

where N is the total number of boundary nodes and Δs is the distance of adjacent Lagrangian points. And the boundary force density $\mathbf{F}(\mathbf{X}_m,t)$ can be used to calculate the FSI force exerted on the solid by the boundary integrals.

Smoothed point interpolation method (S-PIM)

If a group of N_n^s nodes and N_{ele}^s triangular background cells are used to discretize the solid domain ${}^0\Omega^s$. The displacement u_i^s, velocity v_i^s and acceleration a_i^s can be interpolated by the proper shape function ${}^0\Phi_I^s$:

$$u_i^s = \sum_I {}^0\Phi_I^s u_{Ii}^s \ , \ v_i^s = \sum_I {}^0\Phi_I^s v_{Ii}^s \ , \ a_i^s = \sum_I {}^0\Phi_I^s a_{Ii}^s \tag{12}$$

where ${}^0\Phi_I^s$ is the FEM shape function calculated at the initial configuration.

The gradient/strain smoothing technique is introduced in S-PIM based on the smoothed domain to soften the model stiffness, which is the main difference from the FEM. Suppose the domain ${}^0\Omega^s$ is divided into N_{sd}^s smoothing domain ${}^0\Omega_{isd}^{sd}$ with boundaries ${}^0\Gamma_{isd}^{sd}$ (isd =1,2,…, N_{sd}^s). Each smoothing domain ${}^0\Omega_{isd}^{sd}$ is non-overlapped and covers the total calculation together. The smoothed displacement gradient in ${}^0\Omega_{isd}^{sd}$ can be achieved using the displacement gradient $u_{i,j}^s$ in the following form [15]:

$$\bar{u}_{i,j}^s(x_L) = \sum_I \left(\frac{1}{A_{isd}^{sd}} \int_{{}^0\Gamma_{isd}^{sd}} {}^0\Phi_I^s(x^s) n_j^{sd} d\Gamma\right) u_{Ii}^s = \sum_I \frac{1}{A_{isd}^{sd}} \sum_n W_n \left({}^0\Phi_I^s(x^s) n_j^{sd}\right) u_{Ii}^s \tag{13}$$

where, $\bar{u}_{i,j}^s(x_L)$ is the smoothed displacement gradient, A_{isd}^{sd} is the area of the smoothing domain, n_j^{sd} is the outward surface normal of the smoothing domain boundary ${}^0\Gamma_{isd}^{sd}$, I is the number of nodes in smoothing domain, n is the number of gauss points, and W_n is the weight coefficient.

In the nonlinear analysis of solids, the deformation gradient, $F_{ij} = u_{i,j} + \delta_{ij}$, is the primary strain measure. The smoothing operation on the deformation gradient F_{ij} yields the following smoothed deformation gradient $\bar{F}_{ij}^s$:

$$\bar{F}_{ij}^s(X_{isd}^s) = \frac{1}{A_{isd}^{sd}} \int_{{}^0\Gamma_{isd}^{sd}} {}^0\Phi_I^s(X^s) n_j^{sd} d\Gamma u_{Ii}^s + \delta_{ij} \tag{14}$$

Then other smoothed variables can be obtained such as the smoothed Green strain and the smoothed second Piola-Kirchhoff (PK2) stress using the smoothed deformation gradient.

The transient solutions of the nonlinear solids are achieved using the well-developed explicit time integration based on the central difference algorithm. The discretized equation of motion is given in the following form:

$$M_{IJ}^s a_{Ji}^s = f_{Ii}^{s,ext} - f_{Ii}^{s,int} \tag{15}$$

where M_{IJ}^s is the lumped mass matrix, $f_{Ii}^{s,int}$ is the internal force vector defined in the total Lagrangian formulation, and $f_{Ii}^{s,ext}$ is the external force vector in the standard FEM form:

$$f_{Ii}^{s,int} = \int_{{}^0\Omega^s} {}^0\Phi_{I,j}^s P_{ji}^s \mathrm{d}\Omega \tag{16}$$

$$f_{Ii}^{s,ext} = \int_{{}^0\Omega^s} {}^0\Phi_J^s \rho^s \mathrm{g}_i \mathrm{d}\Omega + \int_{{}^0\Gamma^s} {}^0\Phi_J^s \bar{T}_i \mathrm{d}\Gamma \tag{17}$$

Considering the solid immersed in the fluid domain, we can get:

$$\int_{{}^0\Gamma^s} {}^0\Phi_J^s \bar{T}_i \mathrm{d}\Gamma = f_i^{s,FSI} \tag{18}$$

where $f_i^{s,FSI}$ denotes the FSI force.

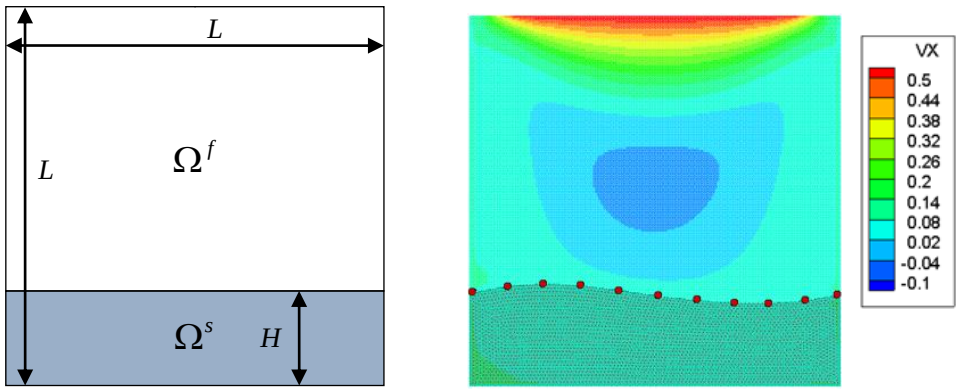
Numerical examples

Lid-driven cavity flow with a soft wall

In this example, a lid-driven cavity flow with a hyperelastic wall is analyzed which can be considered a benchmark FSI problem and has been verified by some previous algorithms including ALE [16], IFEM [17] and IS-FEM [15]. As is illustrated in Figure 2(a), the size of the square cavity is *L*=2 cm and the soft wall is located at the bottom with the length *L*=2 cm and the height *H*=0.5 cm. The fluid properties are given as the density $\rho^f = 1.0\ \mathrm{g/cm^3}$ and the viscosity $\mu^f = 0.2\ \mathrm{g/(cm \cdot s)}$. A simplified Mooney-Rivlin material is used to model the hyperelastic wall with the material constants $C_{10} = 0.1\ \mathrm{g/(cm \cdot s^2)}$, $C_{01} = 0$ and $\kappa = 0$ and the density is set as $\rho^s = 1.0\ \mathrm{g/cm^3}$. The top lid is driven by the following velocity distribution:

$$v_1^f = 0.5 \begin{cases} \sin^2(\pi x / 0.6) & x_1^f \in [0.0, 0.3] \\ 1.0 & x_1^f \in (0.3, 1.7) \\ \sin^2(\pi (x - 2.0) / 0.6) & x_1^f \in [1.7, 2.0] \end{cases} \tag{19}$$

The other boundaries of the fluid domain satisfy the non-slip boundary condition, and the pressure at the midpoint of the bottom edge is set as zero to be a reference value. The top edge of the solid is free while the others are fixed. The fluid domain is divided into 200×200 uniform grids and the solid wall is discretized by irregular triangle elements with 976 nodes. Figure 2(b) shows the result of fluid velocity contour and the configuration of elastic wall, which is consistent with the previous study.



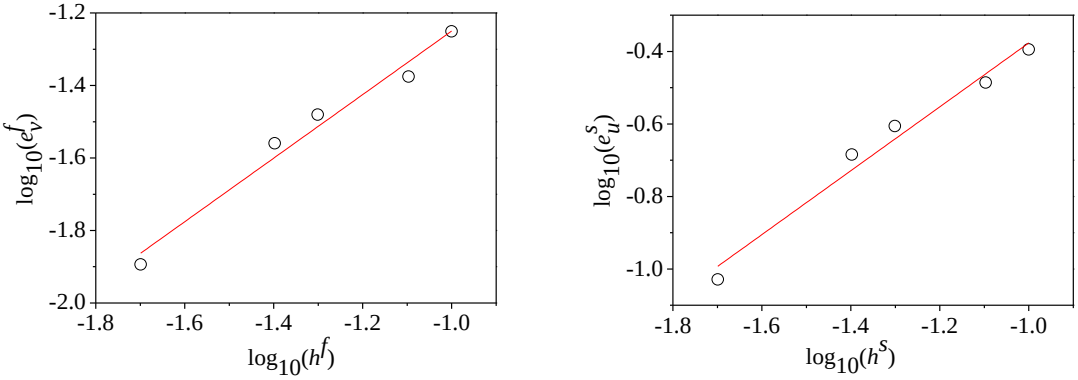
(a) (b)

Figure 2 A elastic wall in a lid-driven cavity flow (a) initial setting (b) Simulation result and the red circle denotes the result obtained by Ref [16]

Then the analysis of convergence and stability would be carried out for the present method by a group of meshes setting. For the test of the fluid, the grid sizes of fluid domain are set as h^f=0.1, 0.08, 0.05, 0.04, and 0.02, which are corresponding to the same solid element size of h^s=0.02. The reference solution is obtained using h^f =0.01 and h^s=0.02. And for the test of solid, the grid sizes of solid domain are set as h^s=0.04, 0.036, 0.032, 0.028, and 0.024, which are corresponding to the same fluid element size of h^f=0.04. The reference solution is obtained using h^f =0.04 and h^s=0.02. We calculate the L^2 norms of errors in the fluid velocity and solid displacement via the following formulas:

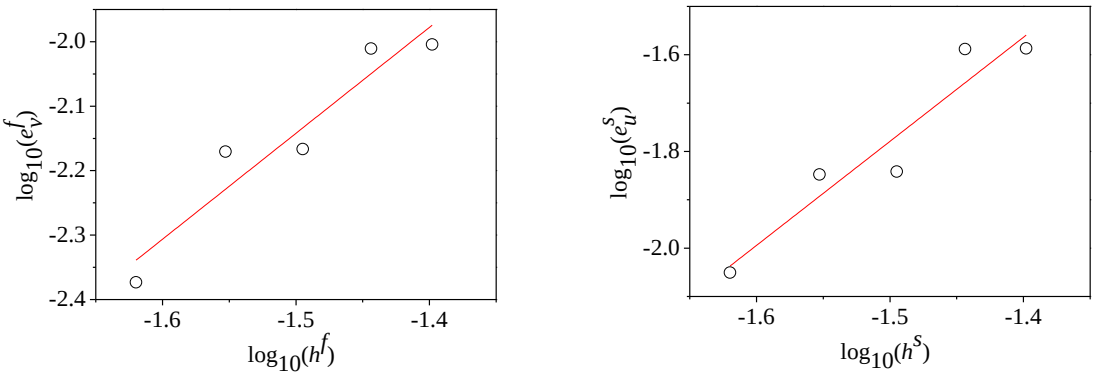
$$e_v^f = \left\| v_i^{f,num} - v_i^{f,ref} \right\|_{L_2} / \left\| v_i^{f,ref} \right\|_{L_2}, e_u^s = \left\| u_i^{s,num} - u_i^{s,ref} \right\|_{L_2} / \left\| u_i^{s,ref} \right\|_{L_2} \tag{20}$$

where $v_i^{f,num}$ and $u_i^{s,num}$ are the numerical solutions, and $v_i^{f,ref}$ and $u_i^{s,ref}$ are the reference solutions respectively. And the result can be seen in Figure 3 and 4 which verifies the good convergence and stability of the proposed method.



(a) (b)

Figure 3 The convergence analysis for the fluid domain (a) the fluid velocity (b) the solid displacement



(a) (b)

Figure 4 The convergence analysis for the solid (a) the fluid velocity (b) the solid displacement

A elastic beam in a fluid tunnel

Here, a steady problem is considered about a flexible beam fixed in the fluid tunnel which was simulated previously using IS-FEM [15]. The viscous fluid flows across the beam and leads to a large deformation because of the fluid force. The beam will come to a stead status after some time when the elastic force balances with the fluid force. As illustrated in Figure 5, The length and height of the fluid field is $L = 4$ cm and $H = 1$ cm. The distance between the fixed beam and the left edge of the fluid tunnel is $L/4$. The thickness of the beam is $a = 0.04$ cm and the height is $b = 0.8$ cm. Gravity is neglected for this problem. The bottom of the fluid domain satisfies the nonslip boundary condition. And the top satisfies a symmetric condition using $v_2^f = 0$.

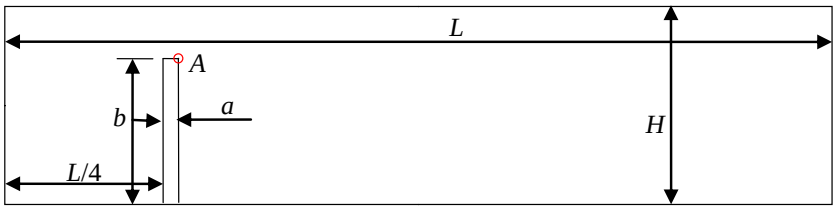


Figure 5 Calculation model of beam in a fluid tunnel

We use 800×200 uniform grids for fluid domain and irregular triangle elements with 409 nodes for the flexible beam. And Figure 6 have shown the result of velocity contour at the time t = 1s, 3s, 6s. A high velocity field arises on the upper of the beam because of the oscillation of the tip end and a low velocity field has been also formed behind the beam due to the barrier effect. And the fluid flow enforces the beam to bend along the fluid tunnel. The fluid force balances the elastic force gradually which brings out a stable state for the FSI system.

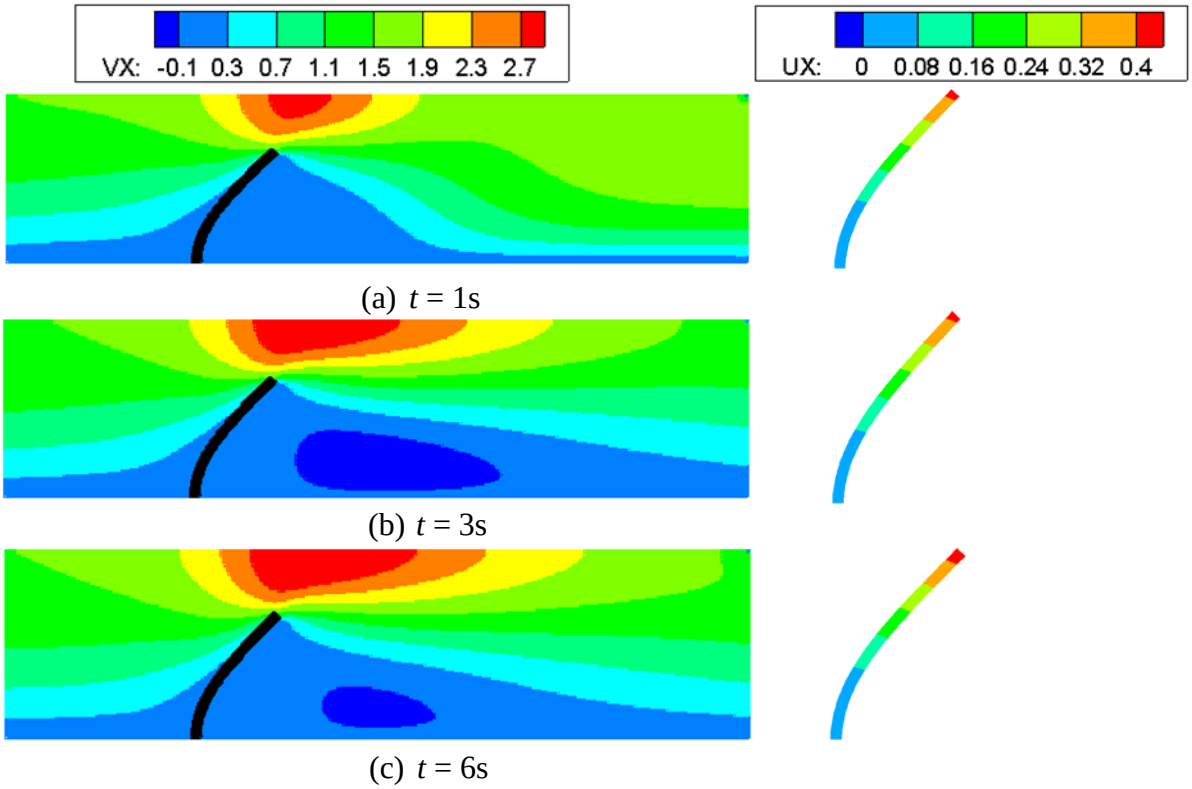
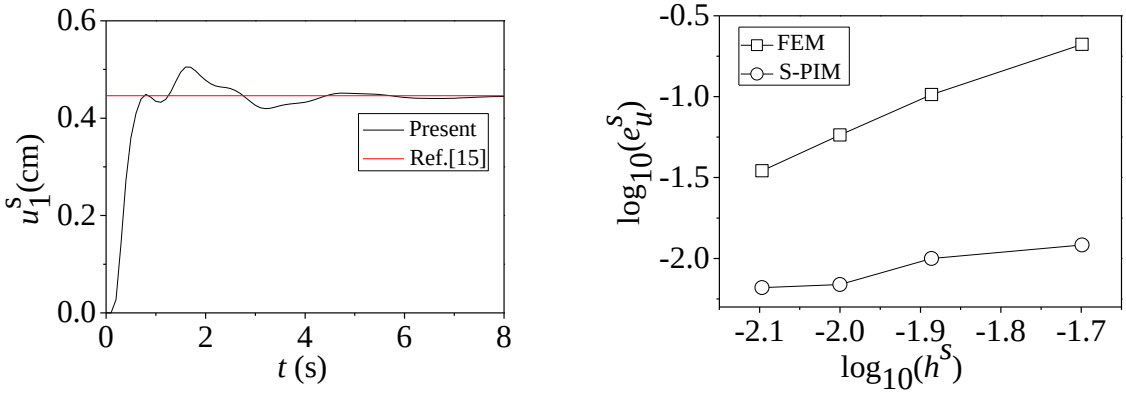


Figure 6 Velocity contour and configuration of beam at different time

(a) *t*=1s (b) *t*=3s (c) *t*= 6s

The horizontal displacement of the tip in the beam has also been investigated in comparison with the result of reference solution, which can be found in Figure 7(a). The displacement amplitude in the balance state keeps consistent with the reported solution. Furthermore, we set a group of meshes to calculate the displacement errors to verify the advantages of S-PIM with FEM. The mesh sizes of solid domain are set as h^s=1/50, 1/75, 1/100, and 1/125 with the same fluid grid of h^f=1/100, and the reference solution is set h^s=1/200 and h^f=1/100 using FEM for solid solver. Figure 7(b) has shown the comparison result which validates the advantage of S-PIM for the solution of elastic problem.



(a) (b)

Figure 7 (a) The curve of horizontal displacement in comparison with IS-FEM [15] (b) The comparison of displacement errors in S-PIM and FEM

Then the stress analysis of beam using S-PIM has been given in the Figure 8. And from the

figure, the beam encounters a primary tension-compression stress in y direction at the fixed bottom end and the fluid force leads to an obvious shear force around the middle of the beam.

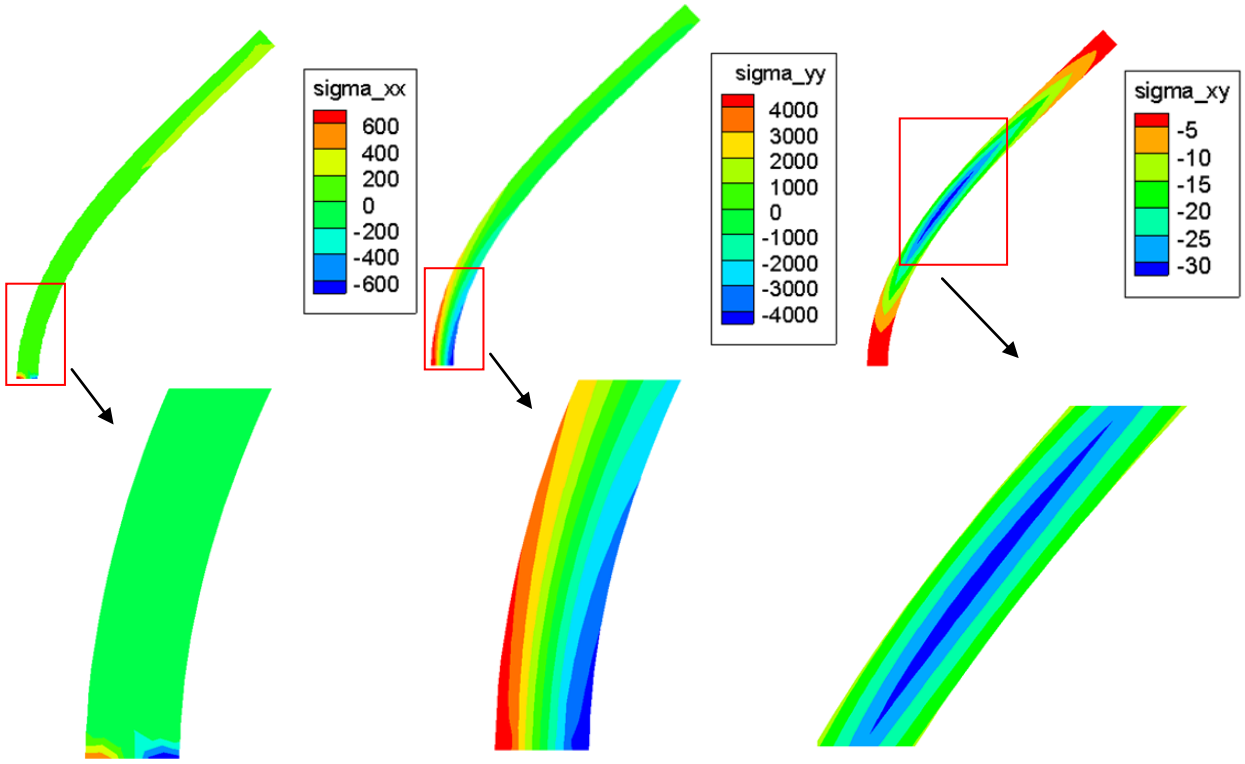


Figure 8 The stress contour at the steady state

Flow passing a cylinder with a flexible flag

A benchmark FSI problem of a cylinder with a flexible flag in the downstream side is analyzed to verify the reliability of IB-LBM with S-PIM. Here we consider the non-steady FSI case discussed by Turek and Hron [18]. As illustrated in Figure 9, the fluid domain is set as $L=2.5$ m and $H=0.41$ m, with a fixed circle of diameter $d=0.1$ m and centered at $C=(0.2,0.2)$ m. The elastic bar was attached at the right edge of the circle with the length $l=0.35$ m and height $h=0.02$ m. The fluid properties are given as $\rho^f=1.0\times10^3$ kg/m^3, $\mu^f=1$ kg/(m·s) which means flow with a Reynolds number of R_e=100. The solid materials are modeled by Saint Venant-Kirchhoff with the density $\rho^s=10\times10^3$ kg/m^3, Poisson's ratio $\nu^s=0.4$ and Young's modulus $E^s=0.5\times10^6$ kg/(m·s^2). The boundary conditions are such that there is no slip over the top and bottom of the fluid channel together with the surface of the solid. At the outlet the pressure is set to be zero, and the input velocity $v_x(t)$ has the following distribution:

$$v_x(t)=\begin{cases}\bar{v}\dfrac{1-\cos(\pi t/2)}{2} & t<2.0\\ \bar{v} & t\geq 2.0\end{cases} \tag{21}$$

where $\bar{v}=1.5\bar{U}y(H-y)/(H/2)^2$.

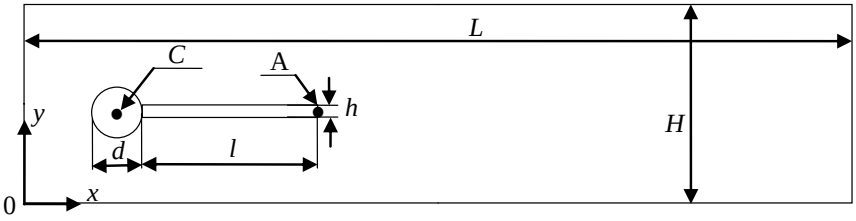
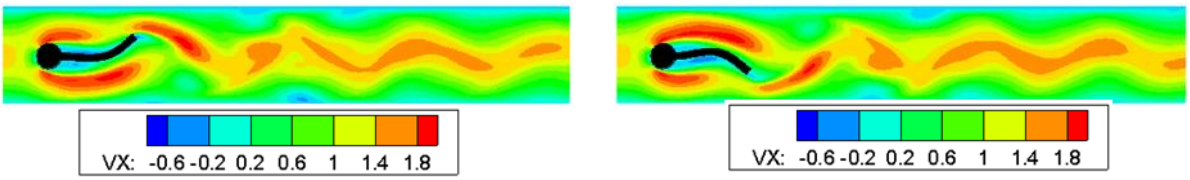


Figure 9 Problem setting of fluid flow past a cylinder with a flag

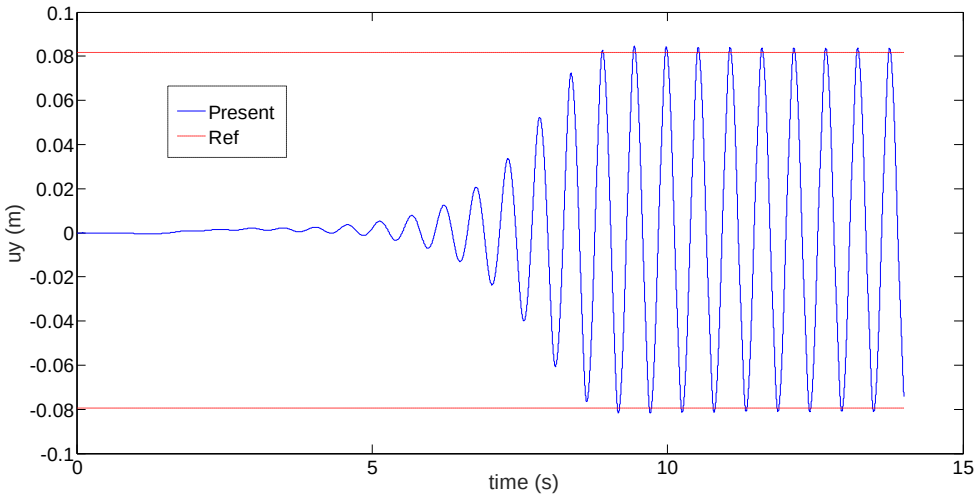
The fluid domain is discretized by 1000×164 uniform grids. And Figure 10 shows the velocity contour at the time t=10 s and t=13 s. The fluid force enforces the flexible to swing and the vortex sheds along with flag as time varies.



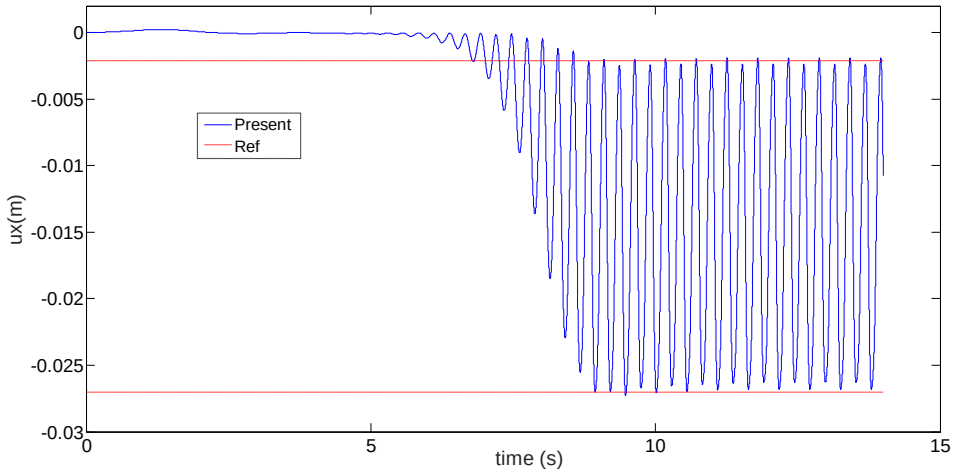
(a) (b)

Figure 10 Fluid velocity contours (a) t=10s (b) t=13s

Figure 11 shows the displacement curve of the tail end in comparison with the reference result by vertical displacement u_y and horizontal displacement u_x, respectively. And from it, we can see a periodical vibration can be obtained after some time.



(a)



(b)

Figure 11 The history of displacement at point A (a) vertical displacement u_y (b) horizontal displacement u_x

Conclusion

In this paper, we coupled IB-IBM with S-PIM to solve fluid-structure interaction problems with large deformable solids. Following conclusions can be obtained from the numerical examples:

1) The proposed method employs the framework of immersed boundary method which makes the method possess the advantages of avoiding re-meshing operation for moving interface.
2) The simple and efficient lattice Boltzmann method is used for incompressible viscid fluid flow with explicit evolution algorithm and avoids the solution of pressure Poisson equation.
3) The smoothed point interpolation method (S-PIM) is employed as solid solver which can soften the model stiffness and establish real constitutive equation for nonlinear analysis.
4) The better result can be obtained by S-PIM coupling with IB-LBM in comparison with coupled FEM for solid analysis.

Acknowledgements:

The authors wish to thank the support of the National Natural Science Foundation of China (No. 51579042), the Key Program of National Natural Science Foundation of China (No. 51639003), the Recruitment program of Global Young Experts (NO. D1007001) and the Fundamental Research Funds for the Central Universities (DUT16ZD218).

References

[1] Peskin, C. (1977) Numerical analysis of blood flow in the heart. *Journal of Computational Physics* 25, 220–252.

[2] Hughes, T.J.R., Franca, L.P. and Balestra, M. (1986) A new finite element formulation for computational fluid dynamics: V. Circumventing the babuška-brezzi condition: a stable Petrov-Galerkin formulation of the stokes problem accommodating equal-order interpolations. *Computer Methods in Applied Mechanics and Engineering* 59, 85-99.

[3] Zienkiewicz, O.C., Nithiarasu, P., Codina, R., Vázquez, M. and Ortiz, P. (1999) The characteristic - based - split procedure: an efficient and accurate algorithm for fluid problems. *International Journal for Numerical*

Methods in Fluids 31, 359-392.
[4] Feng, Z.G. and Michaelides, E.E. (2004) The immersed boundary-lattice Boltzmann method for solving fluid–particles interaction problems. *Journal of Computational Physics* 195, 602-628.
[5] Feng, Z.G. and Michaelides, E.E. (2005) Proteus: a direct forcing method in the simulations of particulate flows. *Journal of Computational Physics* 202, 20-51.
[6] Niu, X.D., Shu, C., Chew, Y.T. and Peng, Y. (2006) A momentum exchange-based immersed boundary-lattice Boltzmann method for simulating incompressible viscous flows. *Physics Letters A* 354, 173-182.
[7] Dupuis, A., Chatelain, P. and Koumoutsakos, P. (2008) An immersed boundary–lattice-Boltzmann method for the simulation of the flow past an impulsively started cylinder. *Journal of Computational Physics* 227, 4486-4498.
[8] Liu, G.R. and Nguyen-Thoi, T. *Smoothed Finite Element Methods*. 2010, Boca Raton: CRC Press
[9] Liu, G.R., Nguyen-Thoi, T. and Lam, K.Y. (2009) An edge-based smoothed finite element method (ES-FEM) for static,free and forced vibration analyses of solids. *J Sound Vib* 320, 1100-1130.
[10] Liu, G.R., Nguyen-Thoi, T. and Lam, K.Y. (2009) Edge-based smoothed point interpolation methods. *Int J Comp Meth-Sing* 5, 621-646.
[11] Liu, G.R. and Zhang, G.Y. *Smoothed Point Interpolation Methods: G Space and Weakened Weak Forms*. 2013, Singapore: World Scientific Press.
[12] Guo, Z., Zheng, C. and Shi, B. (2002) Discrete lattice effects on the forcing term in the lattice Boltzmann method. *Physical review. E, Statistical, nonlinear, and soft matter physics* 65, 046308.
[13] Qian, Y.H. (1992) Lattice bgk models for navier-stokes equation. *EPL (Europhysics Letters)* 17, 479-484.
[14] Kang, S.K. and Hassan, Y.A. (2011) A comparative study of direct-forcing immersed boundary-lattice Boltzmann methods for stationary complex boundaries. *International Journal for Numerical Methods in Fluids* 66, 1132-1158.
[15] Zhang, Z.Q., Liu, G.R. and Khoo, B.C. (2012) Immersed smoothed finite element method for two dimensional fluid-structure interaction problems. *Int J Numer Meth Eng* 90, 1292-1320.
[16] Dunne, T. (2006) An Eulerian approach to fluid–structure interaction and goal‐oriented mesh adaptation. *International Journal for Numerical Methods in Fluids* 51, 1017-1039.
[17] Wang, X. and Zhang, L.T. (2010) Interpolation functions in the immersed boundary and finite element methods. *Comput Mech* 45, 321-334.
[18] Turek S and Hron J, Proposal for Numerical Benchmarking of Fluid–Structure Interaction Between an Elastic Object and Laminar Incompressible Flow, in *Fluid–structure interaction(LectureNotes in Computational Science and Engineering)*, S. e. Bungartz H-J, Editor. 2006, Springer: Berlin, Heidelberg. p. 371–385.2006

Comparation with different interface capturing schemes based on gradient smoothing method using unstructured meshes

***D.Hui[1],†G.Y. Zhang[1,2,3] ,and Z Zong[1,2,3]**

[1]Liaoning Engineering Laboratory for Deep-Sea Floating Structures, School of Naval Architecture, Dalian University of Technology, Dalian, 116024, P. R. China

[2]State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, 116024, P. R. China

[3]Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai, 200240, P. R. China

*Presenting author: huida_answer@hotmail.com

†Corresponding author: gyzhang@dlut.edu.cn

Abstract

This paper focuses on comparing the present advection schemes to capture the interface without reconstruction. The VOF (volume of fluid) equation is solved based on gradient smoothing method. With the help of blending function, the interface capturing schemes are devised as a blend of high-resolution and compression schemes. There are three well-known schemes to be selected, including CICSAM (Compressive Interface Capturing Scheme for Arbitrary meshes CICSAM), FBICS (Flux-Blending Interface-Capturing Scheme) and CUIBS (Cubic Upwind Interpolation based Blending Scheme).Using gradient smoothing operation, the variables of upwind points can be calculated by interpolation on gradient smoothing domains. Two benchmark tests are adopted. Numerical results show that CICSAM scheme produces more numerical error with the increase of Courant number because of numerical diffusion, while FBICS and CUIBS schemes can obtain satisfactory predictions at different Courant numbers.

Keywords: gradient smoothing method (GSM), normalized variable diagram (NVD), unstructured meshes, volume of Fluid (VOF) method.

1 Introduction

In past two decades, computational fluid dynamics (CFD) plays an important role in the safety evaluation of ship and ocean structured (e.g., sloshing, ship slamming,

green water impact, etc). With the growing need in ship building industry, CFD as an engineering tool is facing new requirements. Multiple flows is a major challenge in solving naval hydrodynamics problems using CFD, especially the development of more accurate mathematical model using unstructured meshes. Thus, several interesting methods were introduced and developed to solve multiple flows problems, such as level set method [1], particles on interface [2] or smoothed particle hydrodynamics [3].

One convenient and powerful method based on Eulerian mesh is the Volume of Fluid (VOF) method, which was first developed by Nichols and Hirt [4]. In the VOF method, a volume fraction function is introduced, which represents the fraction of a local cell volume occupied by one of the fluids. And the volume fraction function is governed by a scalar convection equation through flow domain. For maintain the sharpness of the captured interface, one class of schemes is introduced with using interface reconstruction and high-resolution differencing schemes. Noh and Woodard [5] approximated the interface of each cell by vertical or horizontal lines, which is named simple line interface calculation (SLIC). For improving accuracy, the piecewise linear interface calculation (PLIC) was proposed by Youngs [6] using an oblique lines to reconstruct the interface. However, it is not difficult to see that the application is very complicated on unstructured meshes. To avoid reconstructing the interface, another class of approaches is to combine high-resolution schemes with compressive schemes. Over the past decades, many such improved schemes have been developed, among them: HRIC [7], CICSAM [8], STACS [9], FBICS [10] and CUIBS [11] schemes.

More recently, gradient smoothing methods (GSM) has been developed to solve compressible flows problems using unstructured meshes [12]. And GSM also was applied to solve the steady state and transient incompressible flow problems using the artificial compressibility method [13]. The method is effective for various types of fluid dynamics problems by combining with the major features of FVM and some meshfree techniques [14]. Because of different alternative smoothing functions and quadrature schemes for gradient approximation [15], the method has advantages on versatility and flexibility. Thus, the upwind variables can interpolated with the help of gradient smoothing operation, because the upwind points are need for constructing high-resolution schemes. For solving free surface problems using GSM, VOF is introduced in this paper. Thus, different advection schemes is performed and discussed.

Accordingly, several high-resolution, compressive advection schemes are compared in the context of GSM on unstructured meshes. In this article, a brief principle of the GSM is presented. Then the general methodology in interface-capturing schemes is clarified and concisely described, especially, three classical advection schemes are used. Finally, the results related to two advection cases obtained using several schemes at different Courant number values are presented and discussed.

2 Gradient smoothing method

Liu and Zhang developed generalized gradient smoothing technique [16]. Based G space and weakened weak formulation, smoothed point interpolation is presented and used for solving solid mechanics problems. Further, Liu and Xu introduced the method to solve strong–form governing equations for fluid dynamic problems [12]. Variable information is stored on the nodes and their derivatives at various locations are approximated with gradient smoothing operation over relevant gradient smoothing domains.

2.1 Gradient smoothing operation

The gradients of a field variable U at an arbitrary point at xi in domain Ω_i can be approximated in the form of

$$\nabla U_i \equiv \nabla U(X_i) \approx \int_{\Omega_i} \nabla U(X)\, \widehat{\omega}(X - X_i) dV \tag{1}$$

By integrating Eq. (1) by part and using divergence theorem, it becomes

$$\nabla U(X_i) \approx \int_{\Gamma_i} U(X)\, \widehat{\omega}(X - X_i) \boldsymbol{n} \mathrm{d}\Gamma_i - \int_{\Omega_i} U(X)\, \widehat{\omega}(X - X_i) \mathrm{d}\Omega_i \tag{2}$$

where □ is the gradient operator; $\widehat{\omega}$ is the smoothing function; Γ_i denotes the boundary of the gradient smoothing domain Ω_i; and $\boldsymbol{n}$ represents the outward-pointing unit normal vector on Γ_i, as shown in Fig. 1.

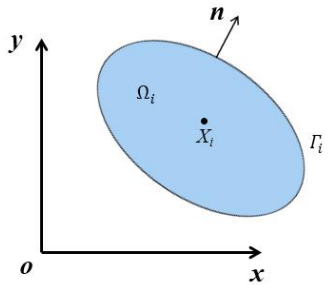


Figure 1. Smoothing domain on point x_i

Based on some essential conditions, e.g. the unity and compact conditions [17], the smoothing function is chosen properly to satisfy requirement of numerical solution. Accordingly, the smoothing functions in our study can be designed to be piecewise constant as follow

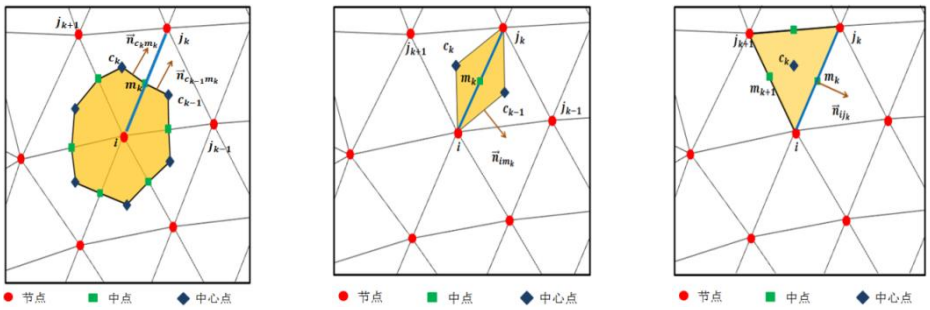
$$\widehat{\omega}(X - X_i) = \begin{cases} 1/A_i, X \in \Omega_i \\ 0 \quad\quad , X \notin \Omega_i \end{cases} \tag{3}$$

where A_i stands for the area of the gradient smoothing domain Ω_i; Thus, the second term on right-hand-side in Eq. (2) will vanishes, which reduces to

$$\nabla U(X_i) \approx \frac{1}{\mathrm{A_i}} \int_{\partial\Omega_i} U(X)\, \boldsymbol{n} \mathrm{d}S \tag{4}$$

2.2 Construction of smoothing domains

The smoothing domains are constructed based on a set of primitive cells which are connected by nodes in computational domain. And the values of field functions are stored at those nodes. There are three types of gradient smoothing domains in GSM, respectively, the node-based gradient smoothing domain (nGSD), midpoint-based gradient smoothing domain (mGSD) and centroid-based gradient smoothing domain (cGSD). The nGSD is formed by connecting the centroids of relevant triangles with midpoints of influenced cell-edges, as shown in Fig. 2 (a). The mGSD is the connection of two end-nodes of the edge with the centroids on the both sides of the cell-edge, as shown in Fig. 2 (b). And the cGSD is formed by a primitive cell, as shown in Fig. 2 (c).



(a) nGSD (b) mGSD (c) cGSD

Figure 2. Illustration of gradient smoothing domains

2.3 Approximations of spatial derivatives

2.3.1 First-order derivatives at nodes

One-point quadrature scheme (chosen as the midpoint) for each edge is used, and it is assumed that

$$U_{ck} = U_{ck-1} = U_{mk} \tag{5}$$

Using gradient smoothing operation of Equation (4), first-order derivatives at nodes can be approximated as

$$\begin{cases} \dfrac{\partial U_i}{\partial x} \approx \dfrac{1}{A_i^{nGSD}} \sum_{k=1}^{n_i} (\Delta \boldsymbol{S}_x)_{ij_k} U_{mk} \\ \dfrac{\partial U_i}{\partial y} \approx \dfrac{1}{A_i^{nGSD}} \sum_{k=1}^{n_i} (\Delta \boldsymbol{S}_y)_{ij_k} U_{mk} \end{cases} \tag{6}$$

where

$$(\Delta \boldsymbol{S}_x)_{ij_k} = (\Delta \boldsymbol{S}_x)_{ij_k}^{(L)} + (\Delta \boldsymbol{S}_x)_{ij_k}^{(R)}, \qquad (\Delta \boldsymbol{S}_y)_{ij_k} = (\Delta \boldsymbol{S}_y)_{ij_k}^{(L)} + (\Delta \boldsymbol{S}_y)_{ij_k}^{(R)} \tag{7}$$

In above equations, A_i^{nGSD} is the area of nGSD; n_i denotes the number of

supporting nodes around node i; $(\Delta \boldsymbol{S}_x)_{ij_k}$ and $(\Delta \boldsymbol{S}_y)_{ij_k}$ represent the sum of normal vectors of domain edges $m_k c_k$ and superscripts (L) and (R) are pointers to the two domain-edge associated with cell edge ij_k; $m_k c_{k-1}$ associated with cell edge ij_k over nGSD shown in Fig. 2 (a),

$$\begin{cases} (\Delta \boldsymbol{S}_x)_{ij_k}^{(L)} = r_{m_k c_k}(\boldsymbol{n}_x)_{m_k c_k}, & (\Delta \boldsymbol{S}_y)_{ij_k}^{(L)} = r_{m_k c_k}(\boldsymbol{n}_y)_{m_k c_k} \\ (\Delta \boldsymbol{S}_x)_{ij_k}^{(R)} = r_{m_k c_{k-1}}(\boldsymbol{n}_x)_{m_k c_{k-1}}, & (\Delta \boldsymbol{S}_x)_{ij_k}^{(R)} = r_{m_k c_{k-1}}(\boldsymbol{n}_x)_{m_k c_{k-1}} \end{cases} \tag{8}$$

where r is the length of domain face and n_x and n_y represent the two components of a domain edge vectors.

The values of field variables U at midpoint are evaluated by simple liner interpolation:

$$U_{mk} \approx \frac{U_i + U_{j_k}}{2} \tag{9}$$

2.3.2 First-order derivatives at midpoints

The gradient at midpoint can be approximated with Eq. (4) over mGSM shown in Fig. 2 (b). They are approximated as follows:

$$\begin{aligned} \frac{\partial U_{mk}}{\partial x} \approx [&\frac{1}{2}(\Delta \boldsymbol{S}_m^x)_{ic_{k-1}}(U_i + U_{c_{k-1}}) + \frac{1}{2}(\Delta \boldsymbol{S}_m^x)_{c_{k-1}j_k}(U_{c_{k-1}} + U_{j_k}) \\ &+ \frac{1}{2}(\Delta \boldsymbol{S}_m^x)_{j_k c_k}(U_{j_k} + U_{c_k}) + \frac{1}{2}(\Delta \boldsymbol{S}_m^x)_{c_k i}(U_{c_k} + U_i)] \frac{1}{A_i^{mG\ SM}} \end{aligned} \tag{10}$$

$$\begin{aligned} \frac{\partial U_{mk}}{\partial y} \approx [&\frac{1}{2}(\Delta \boldsymbol{S}_m^y)_{ic_{k-1}}(U_i + U_{c_{k-1}}) + \frac{1}{2}(\Delta \boldsymbol{S}_m^y)_{c_{k-1}j_k}(U_{c_{k-1}} + U_{j_k}) \\ &+ \frac{1}{2}(\Delta \boldsymbol{S}_m^y)_{j_k c_k}(U_{j_k} + U_{c_k}) + \frac{1}{2}(\Delta \boldsymbol{S}_m^y)_{c_k i}(U_{c_k} + U_i)] \frac{1}{A_i^{mGSM}} \end{aligned} \tag{11}$$

where A_i^{mGSD} is the area of mGSD; ΔS_m^x and ΔS_m^y represent the components of a respective face vector of mGSD; The face vectors is computed in the similar way as the face vectors for nGSD. And the values of field variables U at centroid are calculated by simple liner interpolation:

$$U_{ck} \approx \frac{U_i + U_{j_k} + U_{j_{k+1}}}{3} \tag{12}$$

2.3.3 First-order derivatives at centroids

Analogous to the discretization at nodes and midpoints described above, the gradient at centroids can be approximated over cGSM shown in Fig. 2 (c).

$$\begin{aligned} \frac{\partial U_{ck}}{\partial x} \approx [&\frac{1}{2}(\Delta \boldsymbol{S}_c^x)_{ij_k}(U_i + U_{j_k}) + \frac{1}{2}(\Delta \boldsymbol{S}_c^x)_{j_k j_{k+1}}(U_{j_k} + U_{j_{k+1}}) + \\ &\frac{1}{2}(\Delta \boldsymbol{S}_c^x)_{j_{k+1}i}(U_{j_{k+1}} + U_i)] \frac{1}{A_i^{cGSD}} \end{aligned} \tag{13}$$

$$\frac{\partial U_{ck}}{\partial y} \approx [\frac{1}{2}(\Delta \boldsymbol{S}_c^y)_{ij_k}(U_i + U_{j_k}) + \frac{1}{2}(\Delta \boldsymbol{S}_c^y)_{j_k j_{k+1}}(U_{j_k} + U_{j_{k+1}}) + \frac{1}{2}(\Delta \boldsymbol{S}_c^y)_{j_{k+1} i}(U_{j_{k+1}} + U_i)] \frac{1}{A_i^{cGSD}} \quad (14)$$

where $\Delta \boldsymbol{S}_c^x$ and $\Delta \boldsymbol{S}_c^y$ denote the two components of a respective face vector for cGSM and A_i^{cGSD} is the area of the cGSM.

In this paper, GSM is adopted to solve VOF equation. Because only first order derivative need to calculated, GSM can be treat as vertex-centered FVM. However, the gradient operation is applied for the reconstruction of upwind point. This will be introduced in Section 5.

3. The VOF model

The various fluids are assumed to be incompressible and solutions are obtained by solving the following the conservation of mass and momentum equations [10]:

$$\frac{\partial(\rho \vec{V}_j)}{\partial x_j} = 0 \quad (15)$$

$$\frac{\partial \rho \vec{V}_j}{\partial t} + \frac{\partial}{\partial x_j}(\rho \vec{V}_i \vec{V}_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \rho g_i \quad (16)$$

where $\vec{V}_j$is velocity vector, p is the pressure, $\tau_{ij}$is the viscous stress tensor and $g_i$is gravitational acceleration.

And the volume fractionφ is governed by a simple advection equation:

$$\frac{\partial \varphi}{\partial t} + \frac{\partial(\vec{V}_j \varphi)}{\partial x_j} = 0 \quad (17)$$

The density is calculated by $\rho = \varphi\rho_1 + (1-\varphi)\rho_2$ and viscosity by $\mu = \varphi\mu_1 + (1-\varphi)\mu_2$. The subscripts 1 and 2 respectively denote the two fluids.

In this work, VOF equation is solved without interface reconstruction explicitly. The key is the spatial discretization of the advection equation on unstructured meshes. Thus, the convection term in Eq. (17) over a cell can be approximate as

$$\int_\Omega \frac{\partial(\vec{V}_j \varphi)}{\partial x_j} d\,\Omega = \int_\Gamma \vec{n} \cdot (\vec{V}_j \varphi)\, d\Gamma \approx \sum_f (\vec{n}_f \cdot \vec{V}_f)\varphi_f S_f \quad (18)$$

where Γdenotes the boundary of the control volume Ω; S_f is the area of each face and f is the number of faces of the control volume Ω. For the temporal discretization, the Crank-Nicholson scheme is employed.

4. The present interface capturing schemes

From previous study, it is obvious that the key of the VOF method without interface

reconstruction focus on the interface capturing schemes used in advection equation. The schemes can ensure sharp resolution of the discontinuity, meanwhile avoid an over compressed interface. The design of interface scheme possess the two following basic fundamentals:(a) The interface scheme is a combination of Compressive (BD) and High-Resolution (HR) schemes;(b) Based on the angle between the interface direction and the grid orientation, a blending function should be obtained, preferably in a continuous fashion. Generally, the normalized value of φ at the control volume face can be obtained by blending the two schemes involved BD scheme and HR scheme:

$$\tilde{\varphi}_f = \tilde{\varphi}_f^{BD} f(\theta) + \tilde{\varphi}_f^{HR}[1 - f(\theta)] \tag{19}$$

where $f(\theta)$ is blending function which varies from 0 to 1 and θ is the angle between the normal unit vector of the interface and the vector pointing from centre point C to downstream point D. And $\tilde{\varphi}$is the normal value which is introduced by Gaskell and Lau [18] and Leonard [19]:

$$\tilde{\varphi} = \frac{\varphi - \varphi_U}{\varphi_D - \varphi_U} \tag{20}$$

where the index *U*, *C* and *D* denote upwind point, centre point and downwind point in GSM, respectively, as depicted in Fig. 3. It is clear that when fluid flows from the upwind cell to the interface, if the interface is parallel to the cell face, the compressive scheme should be employed; and if perpendicular to the cell face, only high-resolution is used.

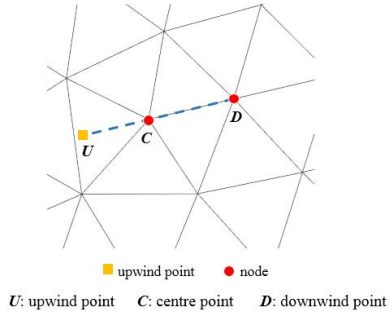


Figure 3. Illustration of the upwind, centre and downwind points on unstructured meshes

4.1 Compressive Interface Capturing Scheme for Arbitrary meshes, CICSAM

The CICSAM scheme was developed by Ubbink and Issa for interface capturing [20]. The scheme switches between the HYPER-C scheme and ULTIMATE QUICKEST (UQ) scheme. Both two schemes need to satisfy the Convection Boundedness Criterion (CBC). The HYPER-C scheme combined the CFL condition and CBC, and is expressed as:

$$\tilde{\varphi}_{fHYPER\ -C} = \begin{cases} \min\left(1, \frac{\tilde{\varphi}_C}{C_f}\right) & 0 \le \tilde{\varphi}_C \le 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \tag{21}$$

And the UQ scheme adopt a blend of upwind and QUICK schemes with a Courant number, the normal face value is defined as:

$$\tilde{\varphi}_{fUQ} = \begin{cases} \min\left(\frac{8C_f\tilde{\varphi}_C+(1-C_f)(6\tilde{\varphi}_C+3)}{8}, \tilde{\varphi}_{fHYPER-C}\right) & 0 \le \tilde{\varphi}_C \le 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \quad (22)$$

where C_f is the value of local Courant number and defined by $C_f = \frac{\sum_f \max(\boldsymbol{V_f S}_f \Delta t, 0)}{V}$.

Furthermore, using a blending function, CICSAM switches the both schemes and can be written as:

$$\tilde{\varphi}_f = \tilde{\varphi}_{fHYPER-C} f(\theta_f) + \tilde{\varphi}_{fUQ}[1 - f(\theta_f)] \quad (23)$$

where $f(\theta_f)$ is blending function of angle θ_f between the gradient of the volume fraction at the interface and the normal to the cell face. The blending function and the angle are calculated by

$$f(\theta_f) = \min[\frac{\cos(2\theta_f)+1}{2}, 1] \quad (24)$$

and

$$\theta_f = \arccos\left|\frac{\nabla\varphi_f \boldsymbol{d}_f}{\|\nabla\varphi_f\|\|\boldsymbol{d}_f\|}\right| \quad (25)$$

The NVD of CICSM is drawn in Fig. 4.

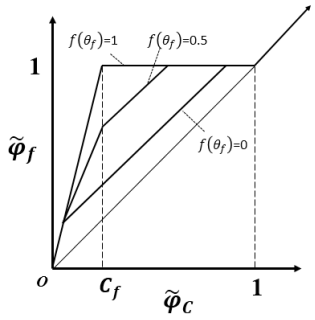


Figure 4. NVD of the CICSAM scheme

4.2Flux-Blending Interface-Capturing Scheme, FBICS

Tsui and co-workers have developed two interface-capturing scheme based on flux blending, FBICS-A and FBICS-B [10]. And in this paper, the FBICS-A is referred to simply as FBICS. Compared with other present schemes, FBICS scheme was the most accurate in capturing interface at different Courant number. FBICS uses Fromm's scheme as the basic scheme in HR and is built to satisfy CBS, is expressed:

$$\tilde{\varphi}_{fHR} = \begin{cases} 3\tilde{\varphi}_C & 0 < \tilde{\varphi}_C \leq \frac{1}{8} \\ \tilde{\varphi}_C + \frac{1}{4} & \frac{1}{8} < \tilde{\varphi}_C \leq \frac{3}{4} \\ 1 & \frac{3}{4} < \tilde{\varphi}_C \leq 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \tag{26}$$

$$\tilde{\varphi}_{fBD} = \begin{cases} 3\tilde{\varphi}_C & 0 < \tilde{\varphi}_C \leq \frac{1}{3} \\ 1 & \frac{1}{3} < \tilde{\varphi}_C \leq 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \tag{27}$$

The scheme is depicted on normalized variables diagram in Fig. 5. Different from CICSAM, FBICS is not dependent on the Courant number.

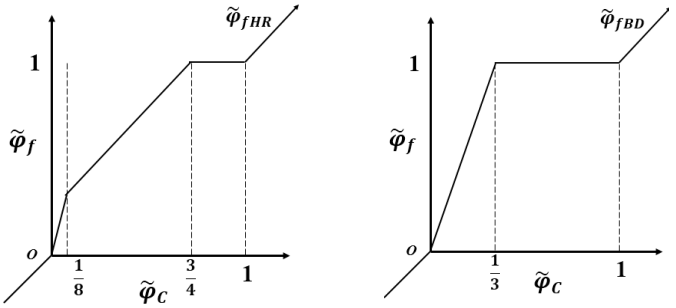


Figure 5. NVD of the FBICS scheme

4.3Cubic Upwind Interpolation based Blending Scheme, CUIBS

A new scheme is proposed for interface capturing, which is inspired by the study of Waterson and Deconinck based on the $GPL-\kappa$ class of schemes [21]. The CUIBS scheme is design to solve VOF model using unstructured meshes and shows a performance that is independent of Courant number [11]. In CUIBS scheme, limited CUI scheme is used as HR scheme and the BD scheme for the compressive is employed which is same as that used for FBICS. The normalized variable diagram of the CUIBS scheme is shown in Fig. 6. The HR and BD scheme is expressed as

$$\tilde{\varphi}_{fHR} = \begin{cases} 3\tilde{\varphi}_C & 0 < \tilde{\varphi}_C \leq \frac{2}{13} \\ \frac{5}{6}\tilde{\varphi}_C + \frac{1}{4} & \frac{2}{13} < \tilde{\varphi}_C \leq \frac{4}{5} \\ 1 & \frac{4}{5} < \tilde{\varphi}_C \leq 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \tag{28}$$

$$\tilde{\varphi}_{fBD} = \begin{cases} 3\tilde{\varphi}_C & 0 < \tilde{\varphi}_C \leq \frac{1}{3} \\ 1 & \frac{1}{3} < \tilde{\varphi}_C \leq 1 \\ \tilde{\varphi}_C & otherwise \end{cases} \tag{29}$$

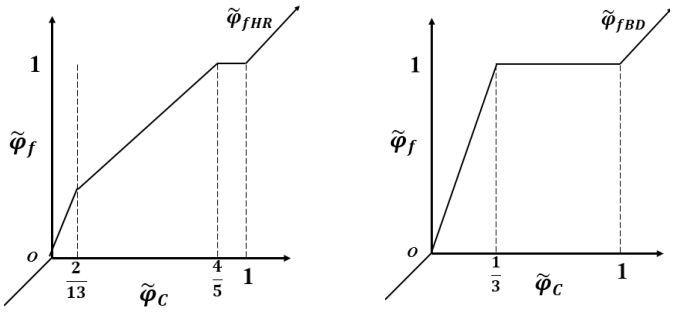


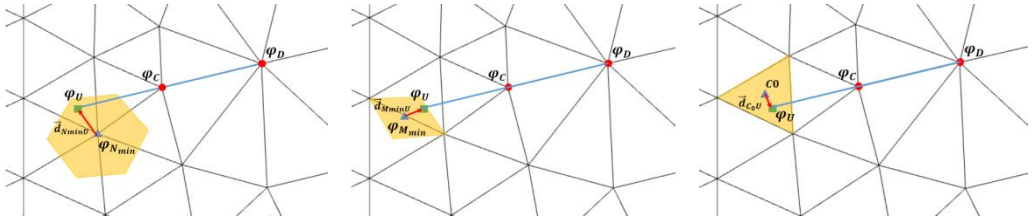
Figure 6. NVD of the CUIBS scheme

5. The upwind point reconstruction on unstructured meshes

Because of the more intricate geometrical computational field, it is difficult to implement TVD scheme on unstructured meshes. The value of point *C* and *D* can be easily obtained from known variables on unstructured meshes, but the location and variable value of upwind point are unknown. Three interpolation scheme (node gradient smoothing method, nGSM、midpoint gradient smoothing method, mGSM and centre gradient smoothing method, cGSM) based on gradient smoothing method are proposed in our previous study, as shown in Fig. 7, it is demonstrated that cGSM lead to a better performance in terms of accuracy and monotonicity. The information at upwind point can be calculated with interpolation on cGSD, is expressed as

$$\varphi_U = \varphi_U + \vec{\boldsymbol{d}}_{UC} \cdot (\nabla\phi)_C \tag{30}$$

where $\vec{\boldsymbol{d}}_{UC}$ is the vector from point U to centroid$C$ and $(\nabla\phi)_C$ is the gradient of centroid C. $(\nabla\phi)_C$ is calculated based on the gradient smoothing domain of centroid$C$.



(a) nGSM (b) mGSM (c) cGSM

Figure 7. Upwind point reconstruction using three schemes based on GSM

6. Numerical test

In this section, the three interface capturing schemes including CICSAM, FBICS and CUIBC schemes are compared to evaluate the relative performance. Two cases are selected for testing: (a) advection of hollow square in an oblique flow; (b) advection of a circle in shear flow. The tests are performed with low and high Courant number. Three Courant numbers are performed and denoted by low, medium and high in

present tests, they approximates 0.1, 0.5 and 0.7, respectively.

The relative error in numerical solutions is defined as

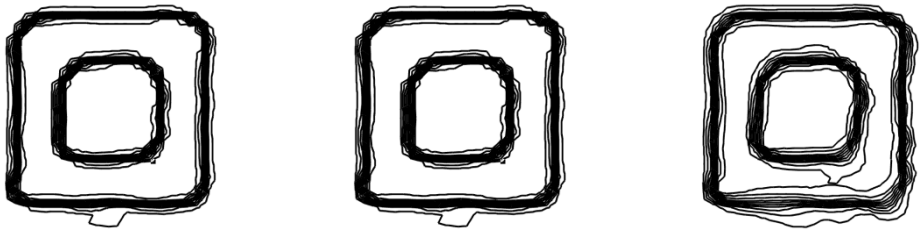
$$E = \frac{\sum_{i=1}^{N}|\phi_i^n - \phi_i^a|}{\sum_{i=1}^{N}|\phi_i^a|} \tag{31}$$

where N is the total number of nodes in the domain. ϕ_i^n and ϕ_i^a respectively denote the numerical solution and analytical solution.

6.1 Advection of hollow square in an oblique flow

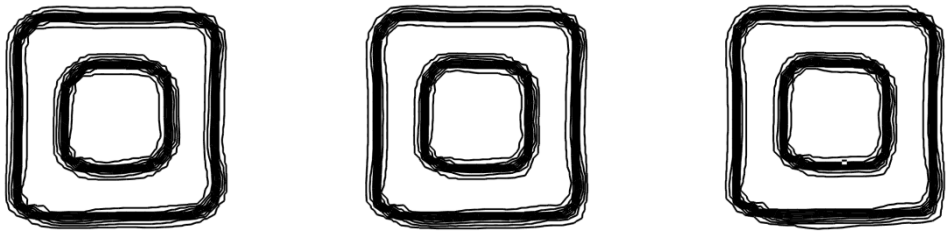
To confirm the performance of interface capturing schemes, a hollow square which the outer width is 0.8 and the inner width is 0.4 and initially centred at (0.8, 0.8) ,transports in an oblique velocity field $(\boldsymbol{V} = (u, v) = (2, 1))$. The domain is set to be 4×4 square. There are 11419 nodes and 22436 cells in unstructured triangular meshes.

After 1 unit of time, the contours of the volume fraction on unstructured meshes including CICSAM, FBICS and CUIBS schemes are depicted in Fig. 8- Fig. 10, and are over the range from 0.05 to 0.95 in interval of 0.1.



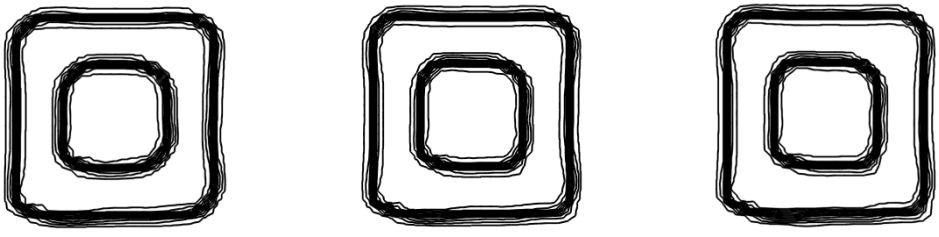
(a) low C_f (b) mediumC_f (c) high C_f

Figure 8. Contour plots for advection of hollow square in oblique flow using CICSAM scheme



(a) low C_f (b) mediumC_f (c) high C_f

Figure 9. Contour plots for advection of hollow square in oblique flow using FBICS scheme



(a) low C_f (b) mediumC_f (c) high C_f

Figure 10. Contour plots for advection of hollow square in oblique flow using CUIBS scheme

The result shows that CICSAM scheme deteriorate at high Courant number because of numerical diffusive, as shown in Fig. 8. And the other two schemes can capture a sharp interface, as depicted in Fig. 9 and Fig. 10. To verify the influence of Courant number, the relative error for the hollow square is calculated using Eq. (31) and summarized in Table 1. With the increasing of Courant number, the error of CICSAM increases, while FBICS and CUIBS scheme are just opposite. Thus, it is also demonstrated that the Courant number has effect on CICSAM scheme and the error of FBICS and CUIBS scheme change very little at different Courant number. This indicates FBICS and CUIBS perform satisfactorily, regard less of the Courant number.

Table 1. Relative error of hollow square in an oblique flow with different Courant numbers

Scheme	Low	Medium	High
CICSAM	0. 13793	0. 14222	0. 18928
FBICS	0. 14328	0. 13411	0. 13049
CUIBS	0. 14249	0. 13180	0. 12897

6.2Advection of a circle in a shear flow

To further compare the ability of the three schemes for capturing interface, a circle in shear flow as a benchmark was tested. The problem reflects the interface twisted by a shear flow field, which is subjected to flow straining and deforms continuously. The computational field was set as a square with the size of $\pi \times \pi$. There are a circle of radius 0.2πcentred at$\left(\frac{\pi}{2}, \frac{(1+\pi)}{5}\right)$. The velocity field is assumed

$$\begin{cases} u = \sin(x)\cos(y) \\ v = -\cos(x)\sin(y) \end{cases} \tag{32}$$

Simulations are performed using unstructured triangular meshes included 26142 nodes and 51682 cells. The circle is strained for *N* time units in forward step, then the velocity is reverses and the circle returned to its original configuration by the backward of *N* time units. Similar with advection of hollow square in oblique flow, the contours of the volume fraction for *N*=8 are depicted in Fig. 11-Fig. 13with different Courant numbers, which are over the range from 0.05 to 0.95 in interval of 0.1.

(a) low C_f (b) medium C_f (c) high C_f

Figure 11. Contour plots for advection of a circle in a shear flow of the forward and backward using CICSAM scheme

(a) low C_f (b) medium C_f (c) high C_f

Figure 12. Contour plots for advection of a circle in a shear flow of the forward and backward using FBICS scheme

(a) low C_f (b) medium C_f (c) high C_f

Figure 13. Contour plots for advection of a circle in a shear flow of the forward and backward using CUIBS scheme

By comparing the contours obtained over unstructured meshes, the results show that CICSAM scheme has evidently dependence on C_f and the predicted interface become evidently diffusive with increasing Courant number. Table 2 presents the relative error variation of the three schemes with Courant number. It should be noted that FBICS and CUIBS scheme lead to more accurate numerical predictions at medium and high C_f, while has slight more numerical diffusion at low C_f.

Table 2. Relative error of a circle in a shear flow with different Courant numbers

Scheme	Low	Medium	High
CICSAM	0. 02613	0. 04609	0. 43241
FBICS	0. 03009	0. 02533	0. 02727
CUIBS	0. 03067	0. 02549	0. 02889

6. Conclusions

Three present interface capturing schemes are implemented and compared in this study. The VOF model is solved by gradient smoothing method without explicitly interface reconstructing. On unstructured meshes, the variables on upwind points are calculated by cGSM model for improving the numerical accuracy. For comparing the ability of the three advection schemes for interface capturing, two benchmark tests are used at different Courant numbers. The results indicate that accuracy of CICSAM scheme is depended on Courant number and has serious numerical diffusion at high Courant number. While FBICS and CUIBS schemes can produce accurate numerical predictions even at high Courant number. Thus the two schemes will be alternative in application to free surface problems using GSM in future.

Acknowledgements:

The authors wish to thank the support of the National Natural Science Foundation of

China (No. 51579042), the Key Program of National Natural Science Foundation of China (No. 51639003), the Recruitment program of Global Young Experts (NO. D1007001) and the Fundamental Research Funds for the Central Universities (DUT16ZD218).

Reference

[1] Osher, S. and Sethian, J.A. (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics* 79, 12-49.

[2] Daly, B.J. (1969) A Technique for Including Surface Tension Effects in Hydrodynamic Calculations. *Journal of Computational Physics* 4, 97-117.

[3] Violeau, D. and Issa, R. (2010) Numerical modelling of complex turbulent free-surface flows with the SPH method: an overview. *International Journal for Numerical Methods in Fluids* 53, 277-304.

[4] Hirt, C.W. and Nichols, B.D. (1981) Volume of fluid (VOF) method for the dynamics of free boundaries ☆. *Journal of Computational Physics* 39, 201-225.

[5] Noh, W.F. and Woodward, P. *SLIC (Simple Line Interface Calculation)*. in *Some Methods of Resolution of Free Surface Problems*. 1976.

[6] Youngs, D.L. *Time-Dependent Multi-material Flow with Large Fluid Distortion*. 1982. 273-285.

[7] Musaferija, S. and Peri`C, M. *Computation of free-surface flows using interface-tracking and interface-capturing methods*. 1999. 59-100.

[8] Ubbink, O. and Issa, R.I. (1999) A Method for Capturing Sharp Fluid Interfaces on Arbitrary Meshes. *Journal of Computational Physics* 153, 26-50.

[9] Darwish, M. and Moukalled, F. (2006) Convective Schemes for Capturing Interfaces of Free-Surface Flows on Unstructured Grids. *Numerical Heat Transfer Fundamentals* 49, 19-42.

[10] Tsui, Y.Y., Lin, S.W., Cheng, T.T. and Wu, T.C. (2009) Flux-blending schemes for interface capture in two-fluid flows. *International Journal of Heat & Mass Transfer* 52, 5547-5556.

[11] Patel, J.K. and Natarajan, G. (2015) A generic framework for design of interface capturing schemes for multi-fluid flows. *Computers & Fluids* 106, 108-118.

[12] Liu, G. and Xu, G.X. (2008) A gradient smoothing method (GSM) for fluid dynamics problems. *International Journal for Numerical Methods in Fluids* 58, 1101-1133.

[13] Xu, G.X., Li, E., Tan, V. and Liu, G. (2012) Simulation of steady and unsteady incompressible flow using gradient smoothing method (GSM). *Computers & Structures* 90, 131-144.

[14] Yao, J., Lin, T., Liu, G.R. and Chen, C.L. (2015) An adaptive GSM-CFD solver and its application to shock-wave boundary layer interaction. *International Journal of Numerical Methods for Heat & Fluid Flow* 25.

[15] Xu, G. (2009) Development of gradient smoothing method (GSM) for fluid flow problems. *National University of Singapore, Singapore*.

[16] Liu, Y.L., Zhong, A.M., Wang, S.P. and Tian, Z.L. (2013) Study on bubble dynamics near plate with hole based on boundary element method. *Acta Physica Sinica* 62.

[17] LIU, G.R. (2011) A GENERALIZED GRADIENT SMOOTHING TECHNIQUE AND THE SMOOTHED BILINEAR FORM FOR GALERKIN FORMULATION OF A WIDE CLASS OF COMPUTATIONAL METHODS. *International Journal of Computational Methods* 5, 199-236.

[18] Gaskell, P.H. and Lau, A.K.C. (2010) Curvature Compensated Convective Transport: SMART, A New Boundedness Preserving Transport Algorithm. *International Journal for Numerical Methods*

in Fluids 8, 617-641.

[19] Leonard, B.P. (2010) Simple high-accuracy resolution program for convective modelling of discontinuities. *International Journal for Numerical Methods in Fluids* 8, 1291-1318.

[20] Ubbink, O. (1997) Numerical Prediction of Two Fluid Systems With Sharp Interfaces. *Imperial College London*.

[21] Watersona, N.P. and Deconinckb, H. (2007) Design principles for bounded higher-order convection schemes – a unified approach. *Journal of Computational Physics* 224, 182-207.

Numerical investigation of blast-induced fractures using smoothed particle hydrodynamics

***†Saba Gharehdash, Luming Shen, and Yixiang Gan**

School of Civil Engineering, The University of Sydney, NSW 2006, Australia.

*Presenting author: saba.gharehdash@sydney.edu.au
†Corresponding author: saba.gharehdash@sydney.edu.au

Abstract

In this study, smoothed particle hydrodynamics (SPH) is utilized to simulate the dynamic behavior of rock under blast. In the simulation, the Johnson-Holmquist (JH2) damage model is employed to model the rock damage and fracture under blasting loads. The effects of air and water as coupling material and copper inside of the borehole are considered. A penalty based node to node contact model is introduced on the interfaces of the different SPH parts to avoid interface effects. The simulated 3D blast-induced fractures are validated by comparing with available experiments. It is found that the crack propagation is asymmetrical around the borehole, despite the apparent isotropy and homogeneity of the simulated rock. The numerical results indicate that the SPH approach used in this work can be applied to effectively simulate densely cracked region, radial cracks and circumferential cracks of rock subjected to blast loading.

Keywords: smoothed particle hydrodynamics, blast loading, fracture, rock

Introduction

Blast-induced damage in rock is a significant yet poorly understood area in the hydrogeology [1], coal gasification [2], geothermal reservoirs [3], fracturing oil shale with explosives [4] and mining [5]. The prediction and control of blast damage has been traditionally done by approximate methods mostly based on experience rather than on understanding of the physical phenomenon. Perhaps the difficulties of experimentation and modeling in blasting added to the significant imperfections of natural rock masses at every scale, plus the limited knowledge on material behavior at very large stresses and loading rates, has significantly limited the research in this area and therefore its understanding. The study presented in this paper intends to contribute to this knowledge by providing a method to be applied to predict blast-induced damage in rock.

Among the different methods which are used for this complicated problem, computational methods are the most economical and efficient tool [6]. Zhu et al. [5] conducted a FEM method of blasting-induced damage in cylindrical rocks, Potyondy et al. [7] applied PFC3D, a 3-dimensional discrete-element program, to simulate rock fragmentation. Trivino and Mohanty [8] used a combined finite-discrete element method (FEM–DEM) to simulate blast-induced damage in a granitic outcrop. The numerical models presented in their works can be applied to treat the problem of blasting-induced crack initiation and propagation in rocks. Three basic fracture zones, i.e., crushed zone, severely fractured zone and spalling cracks have been successfully simulated. However, they could not extend their works to verify the predicted fracture dimension against controlled experiment. Their approach holds promise for this class of rock fracture problems. While the quantitative correlations of model predictions with experiment was not successful.

The mesh free methods such as smoothed particle hydrodynamics (SPH) [9], material point method [10] and discrete element method [11] have shown the potentials to simulate large deformation behaviour of rock medium by including elasto-plastic or damage models. In particular, SPH is widely applied to computational solid mechanics [9] due to its superiority in solving problems with large deformation. Recent developments in SPH make it promising for simulating rock fracture and fragmentation under blast load [12]. The numerical example

presented in their work showed the potentiality to estimate blast-induced crack initiation and propagation. The failure process of the rock medium can be separated into three failure zone namely, crush zone, radial cracks zone and spalling cracks zone. In many applications, it is convenient to combine SPH with another method, such as the finite element (FEM) or discrete element method (DEM) [13, 14]. Gharehdash et al. [13] used coupled SPH-FEM for investigating of rock fracturing. Presented SPH-FEM approach can be applied to treat the problem of blasting-induced crack propagation. It was found that the model is capable to capture some of the observed phenomena in rock blasting experiments. SPH-FEM approach may potentially lead to handling of large deformation problems with low computational cost, however, the interaction between SPH particles and FEM elements didn't represent the real interface between SPH and FEM. SPH-DEM for first time was used by Fakhimi and Lanari [14]. It was shown that the proposed hybrid model is capable of simulating the crack propagation in the rock material and the crushed zone and radial cracks, and surface spalling were all captured successfully. The interaction between DEM and SPH followed a perfectly plastic collision that caused partial penetration of smoothed particles into discrete elements. The contact model in hybrid methods can be problematic and cause numerical instability.

Although some realistic features such as the incorporation of various free surfaces and displacement of fragments have been achieved in 3D simulations, most previous investigations considered only 2D simulations [5]. In the present work, a 3D numerical simulation using SPH has been performed to predict the fracture patterns in rock under different blast conditions. The numerical investigation in this paper considers both quantitative and qualitative aspects of rock blasting.

Methodology

Smoothed particle hydrodynamics method

Smoothed Particle Hydrodynamics (SPH) is a Lagrangian-based numerical method used for simulating problems in fluid and solid mechanics. SPH was first developed to simulate nonaxisymmetric phenomena in astrophysical dynamics [15]. Due to its flexibility of meshless Lagrangian nature, ease of implementation is well employed within numerous branches of computational physics [9].
The SPH method considered as an interpolation method using the sifting property of the Dirac-delta function. Figure. 1 shows a particle neighbourhood domain centred at a particle of interest, labelled *i*. The circle of radius *2h* indicates the immediate neighbours.

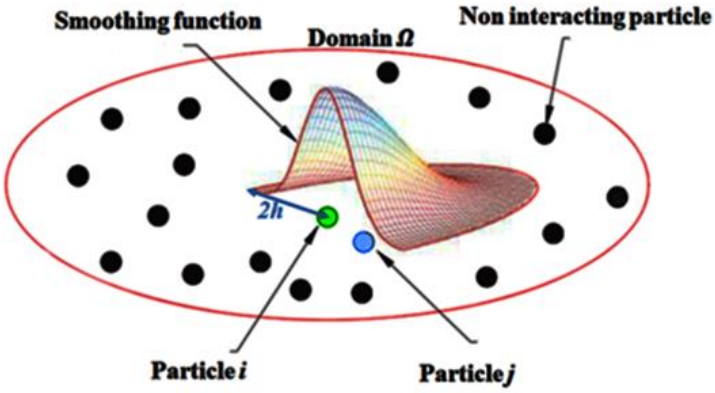


Figure 1. SPH particle neighborhood.

Consider a function *f*, a kernel *W* which has a width (support domain) determined by the parameter *h*, we define the kernel estimate as

$$f(x) = \int f(x_j) W(|x - x_j| / h) dx_j \quad (1)$$

where x_j is the value of the quantity x for particle j, h is smoothing length. If we go from an integral to a sum, then the discrete kernel estimate becomes the summation over neighbouring particles and the particle i itself as

$$f(x) \cong \sum_{j=1}^{N} m_j f(x_j) W(\left|x - x_j\right| / h) / \rho_j \tag{2}$$

where m_j is mass of particle j, ρ_j is density of particle j and N is the number of particles within support domain of kernel function. Particle equation for the gradient can be obtained as

$$\nabla f(x)_{x \approx x_j} \cong -\sum_{j=1}^{N} m_j f(x_j) \nabla W(\left|x - x_j\right| / h)_{\left|x \approx x_j\right.} / \rho_j \tag{3}$$

Approximations (2) and (3) to functions and their derivatives allows estimates of accelerations, strain rates, etc., in the continuum equations and forms the basis of SPH. The transformation of conservation equations of continuum mechanics into SPH framework described as

$$\frac{d\rho_i}{dt} = \rho_i \sum_j \frac{m_j}{\rho_j} \left(U_j - U_i\right) \cdot \nabla W_{ij} \tag{4}$$

$$\frac{dU_i}{dt} = \sum_j m_j \left(\frac{\sigma_i}{\rho_i^2} + \frac{\sigma_j}{\rho_j^2} - \Pi_{ij} I + \left(R_{\varepsilon i} + R_{\varepsilon j}\right) f_{ij}^{\,n} \right) \cdot \nabla W_{ij} \tag{5}$$

$$\frac{d\rho_i}{dt} = \rho_i \sum_j \frac{m_j}{\rho_j} \left(U_j - U_i\right) \cdot \nabla W_{ij} \tag{6}$$

where I is unit matrix, U is the velocity vector, and σ is the stress tensor. A given particle i has a density change determined by (4), an acceleration obtained from (5), and an internal energy change given by (6). The Π_{ij} term represents the artificial viscous pressure. The summations are over neighbouring j particles. For more comprehensive details on SPH method one can refer to Gingold and Monaghan [15].

SPH simulations

In the present study, SPH algorithm was implemented into the explicit finite element code LS-DYNA [16] in order to model rocks failure under blast loading. Some modifications in the SPH suit the needs of simulations in removing the tensile instability, interaction between different SPH parts and artificial viscosity. The proposed SPH is then applied to the simulation of well documented laboratory experiments in granitic rocks. The different aspects of SPH are described as follows. In order to prevent the particle from clumping with neighbour particles we used the most effective and successful treatment which was introduced by Monaghan [17]. The main idea is to introduce an artificial repulsive force in the neighbourhood of a SPH particle that is in tension. The repulsive force is introduced in the momentum equation in the form of an artificial stress term (Eq. 5). In Equation 5 $R_{\varepsilon i}$ and $R_{\varepsilon j}$ are artificial stress terms of particles i and j, respectively, with correction parameter ε [17]; n is exponent dependent on the smoothing kernel; and f_{ij} is defined as

$$f_{ij} = \frac{W_{ij}}{W(\Delta d, h)} \tag{7}$$

Where Δd is initial particle spacing. In this study, h is assumed to be $1.2\,\Delta d$ for the cubic spline kernel. For the problem described herein, exponent n and correction parameter ε are chosen to be 4 and 0.5, respectively, to remove the tensile instability of the SPH particle of the rock material [17]. This is of the greatest importance for applications in solid mechanics, and especially for problems involving damage, fracture and fragmentation resulting from transient tensile stress fields. On the other hand, Smoothed quantities of a particle show falsified values when densities and masses of neighbouring particles vary largely within the smoothing length. So in our blast simulations, a penalty based node to node contact model is introduced on the interfaces of the different SPH parts to avoid interface effects. The standard SPH method is based on an updated Lagrangian formulation. Rabczuk et al. [18] has shown that this updating of the coordinates was the main source of instability in solid mechanics applications. He hence proposed to use a total Lagrangian formulation of the equilibrium equations [18]. This formulation was used for updating the equilibrium equations in this work.

Description of simulated problem

For blast simulation a cylindrical rock containing a single centrally located line source of explosive coupling with different medium is considered (see Figure 2). The cylindrical rock measures 144 mm in diameter and 150 mm in length. The line source of explosive is represented by a single strand of detonating cord containing a core load of PETN explosive (1.65 mm in diameter) surrounded by a thin sheath of polyethylene, with the total diameter of 4.5 mm. The borehole diameter is 6.45 mm. The coupling materials of water and air are used in this simulation. A copper tube with a 1.2 mm thick was tightly installed in the borehole of each specimen. Copper can be deformed easily under the applied shock/stress loads and expand with the borehole without tearing, and thereby preventing any gas penetration into the resulting cracks to prevent fragmentation of rock. Figure 2 shows three dimensional view of rock specimen with a scaled close-up of the borehole region for different simulations.

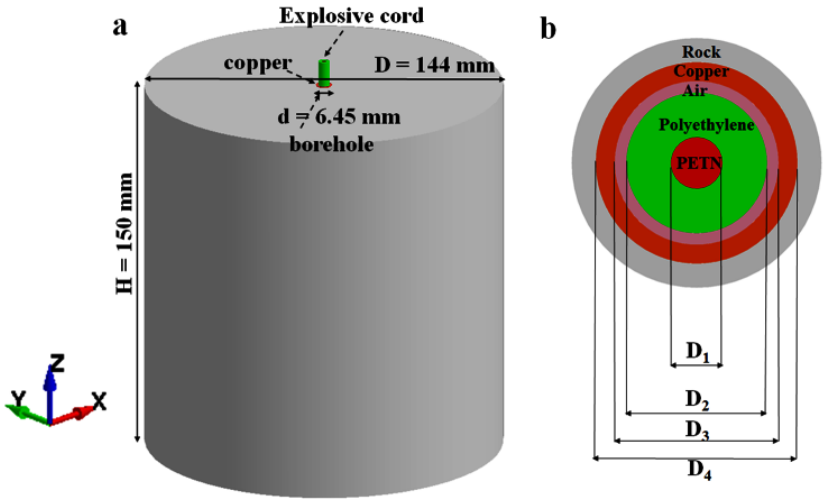


Figure 2. Rock cylinder under blast load a) 3D view, b) cross-section view of the borehole.

The *p*-wave velocity in the Barre granite is calculated with the following expression:

$$c_d = \sqrt{\frac{(\lambda + 2\gamma)}{\rho}} \tag{8}$$

where λ and γ are Lame constants and ρ is the material's mass density. This wave velocity is used for analysis of wave propagation in the numerical simulations. The transmitted pressure arrival from the top of the borehole is used to record arrival of the stress wave from blasting. The rock bottom and top surfaces are set as the fixed boundary in the Z direction, and its side surface is set as the free boundary. The location of the initiation points was set at the top of the blast hole. After mesh convergence study, particle distance of 0.15 mm was selected for dynamic simulation, giving a total 5457876 particles. Table 1 depicts all combinations of borehole diameters, detonating cords and coupling media used in the simulations.

Table 1. Different combinations of detonating cords, coupling media and borehole diameter (equal to outside diameter of copper tube)

Materials inside of borehole	Simulations		
	Air-coupled	Water-coupled	Without copper
	Outside diameter (mm)		
PETN (D_1)	1.65 mm	1.65 mm	1.65 mm
Polyethylene (D_2)	4.5 mm	4.5 mm	4.5 mm
Air (D_3)	5.25 mm	-	6.45 mm
Water (D_3)	-	5.25 mm	-
Copper (D_4)	6.45 mm	6.45 mm	-

Material model

In the SPH simulations, Johnson-Holmquist constitutive model [19] was used to model granitic rocks. This damage model is capable of considering pressure and strain rate dependencies, and allows softening of the material under investigation. A brief description of the model can be found in Johnson and Holmquist [19]. The material properties of the Barre granite rock for blast simulation is from Dehghan Banadaki and Mohanty [20]. Johnson and Cook model [21] was used to model the copper. This model requires an equation of state (EOS) to be incorporated to represent the behaviour of the copper under different phases, so Mie-Grunesien EOS was used for copper under shock load [21]. For air, water and polyethylene Material Type 9 of LS-DYNA (*MAT_NULL) [16] is used. As for the air, the polynomial EOS is usually employed, in which the pressure *P* is expressed as

$$p = C_0 + C_1\mu + C_2\mu^2 + C_3\mu^3 + (C_4 + C_5\mu + C_6\mu^2)e \tag{9}$$

where *e* is the internal energy per volume. The compression of the material is defined by the parameter $\mu = \frac{\rho}{\rho_0} - 1$, where ρ and ρ_0 are the current and initial density of the material, respectively. As a matter of fact, the air is often modelled as an ideal gas by setting $C_0 = C_1 = C_2 = C_3 = C_6 = 0$ and $C_4 = C_5 = 0.401$. Air mass density ρ_0 and initial internal energy e_0 are 1.255 kg/m^3 and 0.25 J/cm^3, respectively. For the water and polyethylene (sheath of the detonating cord) in the immediate vicinity of the explosive charge, shock EOS is used [20]

$$U_s = C_0 + su_p \tag{10}$$

where U_S is the shock wave velocity, C_0 is the bulk sound speed, u_p is the particle velocity and s is the material constant, listed in Table 2.

Table 2. Shock EOS parameters for the water and polyethylene used in the simulations [20].

Material	Bulk sound speed C_0 (m/s)	Material constant s
Polyethylene	2901	1.481
Water	1483	1.750

The explosive PETN was modelled using explosive burn constitutive model (*MAT_HIGH_EXPLOSIVE_BURN) [16]. For the detonation produced explosive gas, the standard Jones-Wilkins-Lee (JWL) equation of state is employed, which corresponds to a detonation velocity of 6690 m/s and a Chapman–Jouget (C-J) pressure of 16 GPa. The pressure of the explosive gas is

$$p = A\left(1-\frac{\omega\eta}{R_1}\right)e^{-\frac{R_1}{\eta}} + B\left(1-\frac{\omega\eta}{R_2}\right)e^{-\frac{R_2}{\eta}} + \omega\eta\rho_0 E \tag{11}$$

where η is the ratio of the density of detonation products to the initial density of the original explosive, E is the specific internal energy per unit mass, and A, B, R_1, R_2, ω are fitting coefficients. The values of the corresponding coefficients can be found in [20].

Simulations of blast testing

A mesh convergence study is conducted first. Comparisons of the fracture patterns for the different mesh sizes are presented in Figure 3. Six models, shown in Figure 3, with varying SPH mesh densities were created to compare the sensitivity of the predicted fracture pattern response to mesh size. Six meshes with particle sizes 0.6, 0.4, 0.25, 0.2, 0.15 and 0.1 mm for top cross section are modelled. It can be noticed that the crack patterns are very similar for particle sizes 0.25, 0.2, 0.15 and 0.1 mm although the crack paths are thinner for the finer meshes. The failure patterns for the four mesh densities are similar and there is failure pattern convergence as the mesh size decreases. Considering the balance between accuracy and efficiency, the mesh size 0.15 mm is used in all the simulations below. This mesh density study clearly illustrates the importance of performing a mesh study using the SPH method.

Four different simulations for rock under blast loading were performed as described in Table 1. In this section, we first show the pressure history results for air-coupled with copper simulation to validate the proposed SPH, and then analyse the crack patterns produced by the blast loading. Pressures at different distances from the borehole walls obtained from numerical simulations are shown in Figure 4. As seen, the simulated pressure distribution match the experimental data quite well.

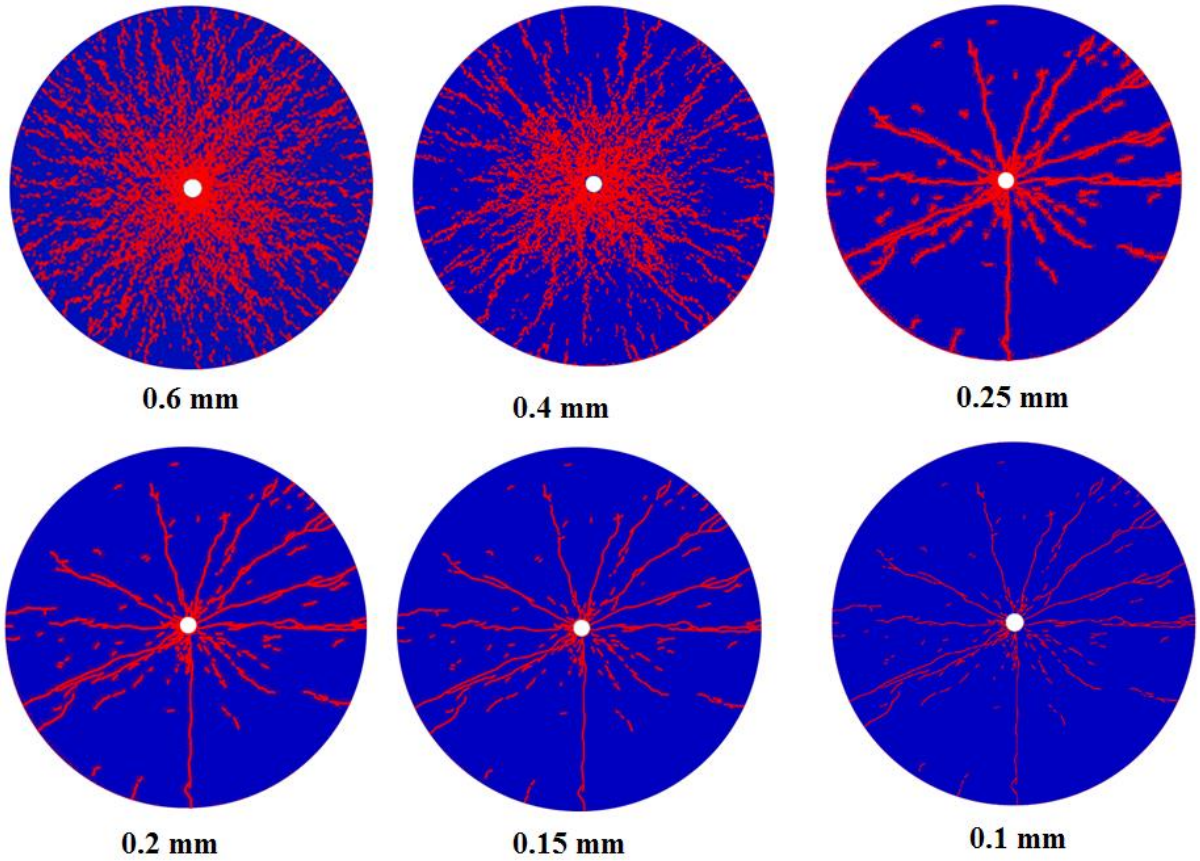


Figure 3. Failure patterns simulated using different mesh sizes.

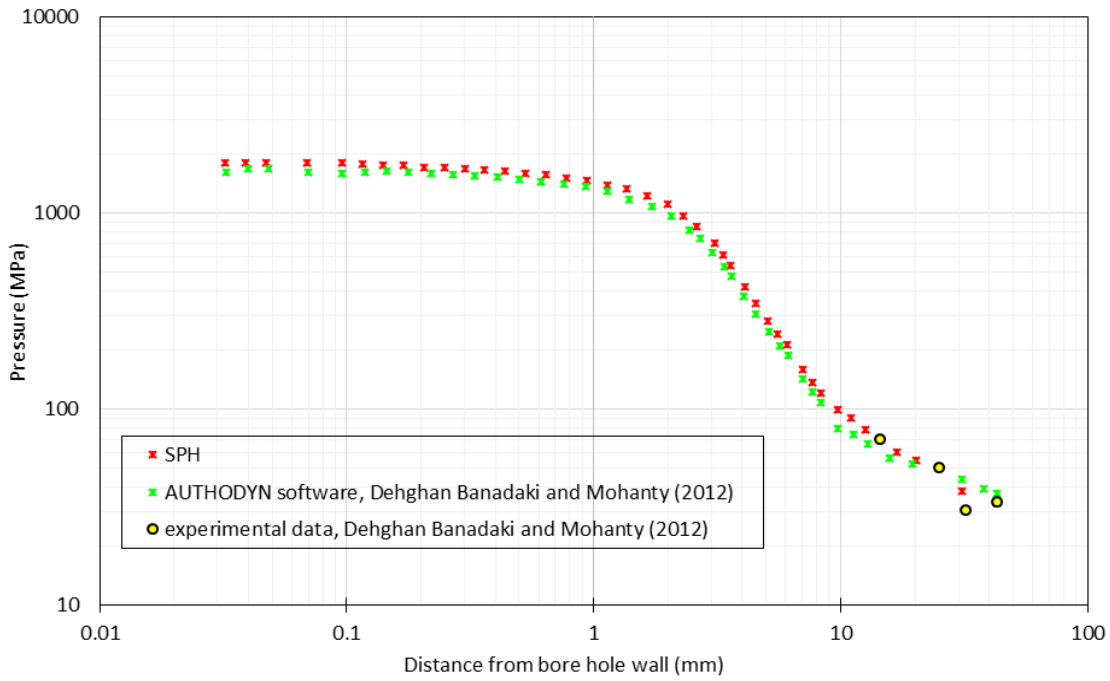


Figure 4. Comparison of pressures at different distances from the borehole walls obtained from experimental and numerical works.

In order to further validate our simulations with experimental data, damage distribution around the borehole was examined qualitatively. Exploded rock was cut at 25, 75 and 125 mm from the bottom surface of the rock specimen in experimental test. Figure 5 compares fracture patterns for the air-coupled with copper simulation. As seen, after explosion the resulted intense stress wave travels into the rock specimen and causes different types of failures. To investigate the formation mechanism of the different types of failures, the relationship among pressure, stresses and strength for the failed particles are presented. In the three failed particles as shown in Figure 5, the histories of stresses are recorded in order to analyse rock fracture mechanism. Throughout this work, tensile stress is positive and compressive stress is negative.

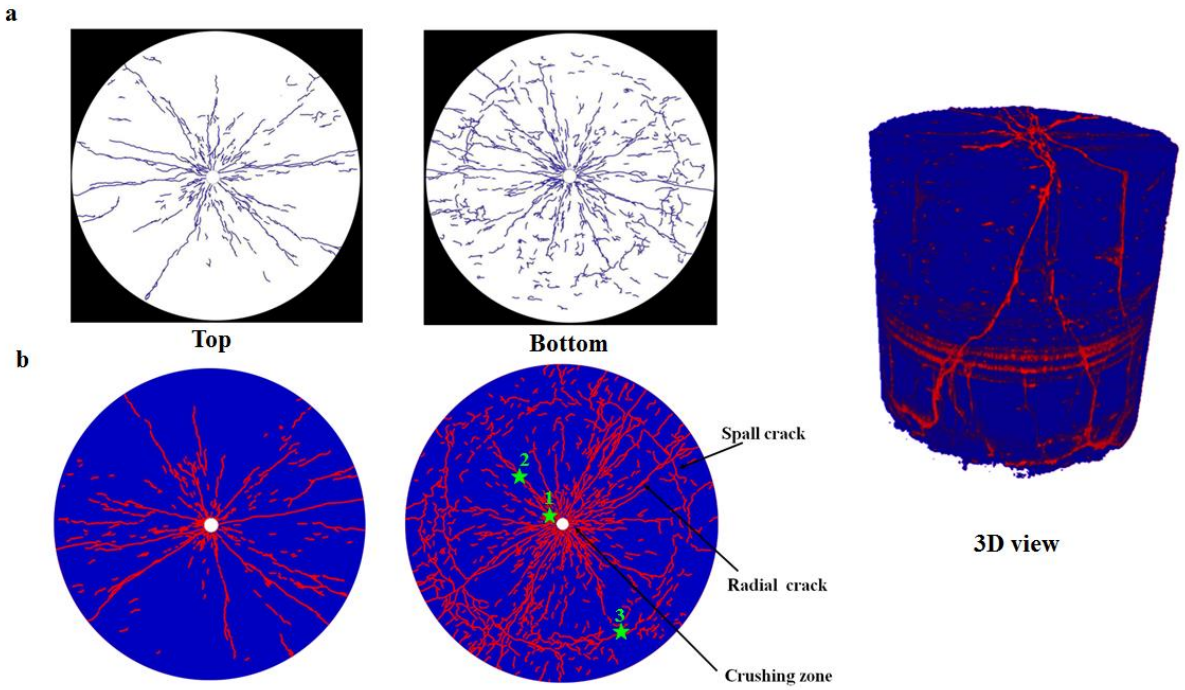


Figure 5. Comparison of damage created around the borehole for top and bottom sections, a) experimental [22], b) numerical simulation.

For the failed particle close to the borehole wall ($x = 8.54$ mm, $y = 4.43$mm, $z = 100$ mm) at times smaller than 72.5 μs, before the shock wave touches this particle, the particle pressure is equal to zero and yield stress is equal to 286 MPa (see Figure 6). Details of obtaining the yield stress can be found in [17, 18]. Then yield stress gradually increases with pressure and reaches its maximum value of 2499 MPa at 72.754 μs. As expected for the JH2 model, this clearly shows the pressure dependency of the strength. Rock behaves elastically from when it experiences the shock wave until the Mises stress contacts the yield surface. After the peak pressure has passed the particle, Mises stress contacts the yield surface at time 72.83 μs and the particle starts to deform plastically. During plastic deformation damage accumulates in the particle. Maximum and minimum principal stresses increase with time at the compressive status. At 72.8 μs, the maximum shear stress reaches up to 853.17 MPa, which is much larger than the rock dynamic shear strength (265 MPa). Then the particle fails in shear, and the all principal stresses, pressure, Mises stress and yield stress fall to zero.

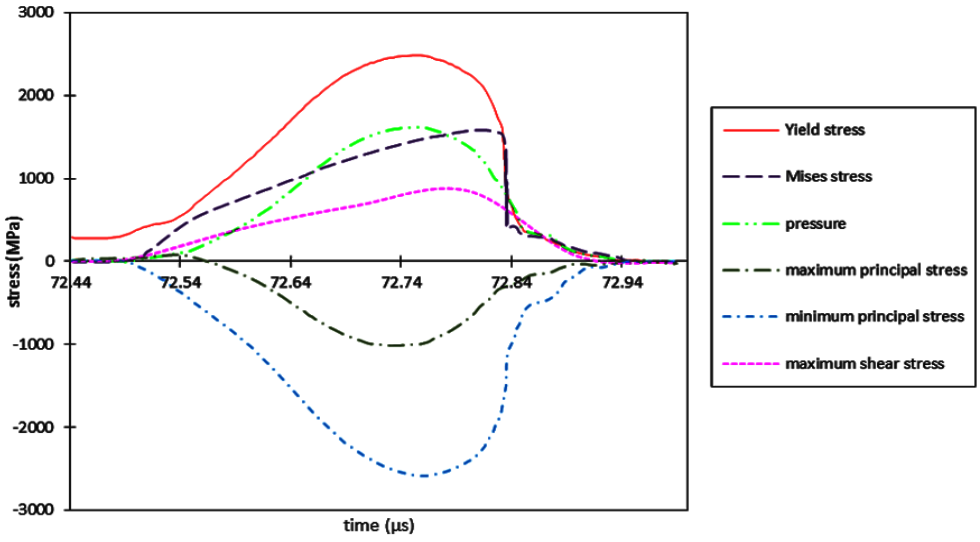


Figure 6. Relationship between pressure, stresses and strength versus time for failed point 1 close to the borehole (shown with a green star in Figure 5).

The failed particle in coordinate of ($x = 31.75$ mm, $y = 28.62$ mm, $z = 100$ mm) is selected to investigate the radial crack formation. The pressure, stresses and strength as a function of time are shown in Figure 7. Like previous case, Mises stress contact the yield surface after the passage of the compressive stress wave at time 73.61 μs. The tension tail of the compressive stress wave results in negative pressure in the particle followed by positive pressure. The minimum principal stress is always compressive stress. The maximum principal stress alternates between compression and tension status due to the coupling effect between the stresses and particle strain energy release; then it reaches 115 MPa at 73.43 μs, which is larger than the rock dynamic tensile strength (100 MPa) cause the tensile failure at this particle. The maximum shear stress doesn't exceed the dynamic shear strength of rock so the mechanism of failure cannot be in shear. Then when particle completely damaged all the stresses become zero.

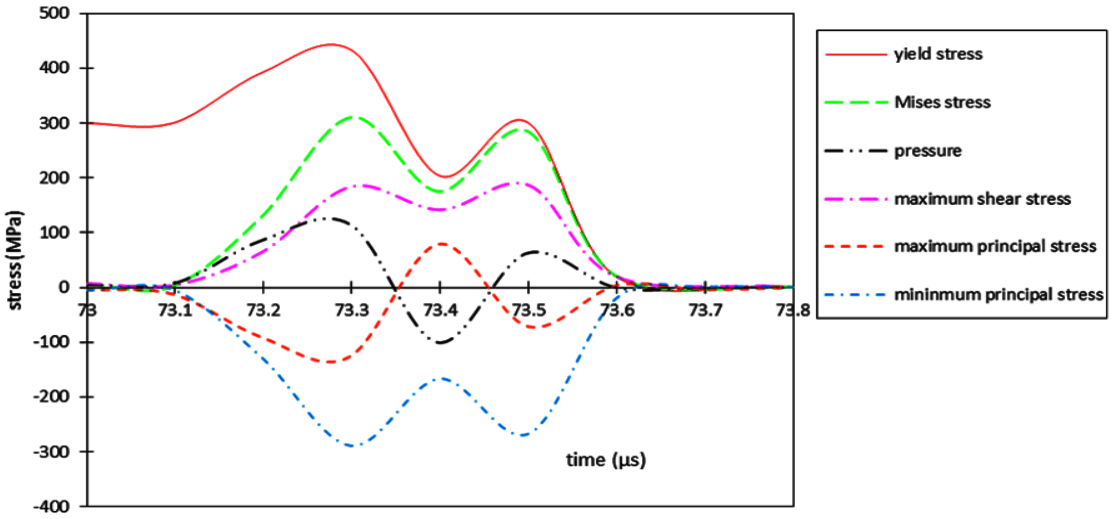


Figure 7. Relationship between pressure, stresses and strength versus time for failed point 2 (shown with a green star in Figure 5).

The relationships between pressure, stresses and strength for particle on the spall crack at position of ($x = 68.23$ mm, $y = -57.98$ mm, $z = 100$ mm) are shown in Figure 8. Unlike the previous cases, Mises stress does not contact the yield surface after the passage of the compressive stress wave. The stress wave propagates to the particle at 76.87 μs; then, the principal stresses increase with time at the compressive status. The maximum principal stress reaches 107 MPa at 87.65 μs, and the particle fails in tension. After a while, the reflected wave comes back from the free surface resulted in higher negative pressures in the selected particle. Although the particle fails because the maximum principal stress exceeds the rock dynamic tensile strength, its maximum shear principal stress is not close to the rock dynamic shear strength at this time. Thus the failure is not in shear. These results are in good agreement with work [22]. Experiments on rock also show similar signs of damage (Figure 5).

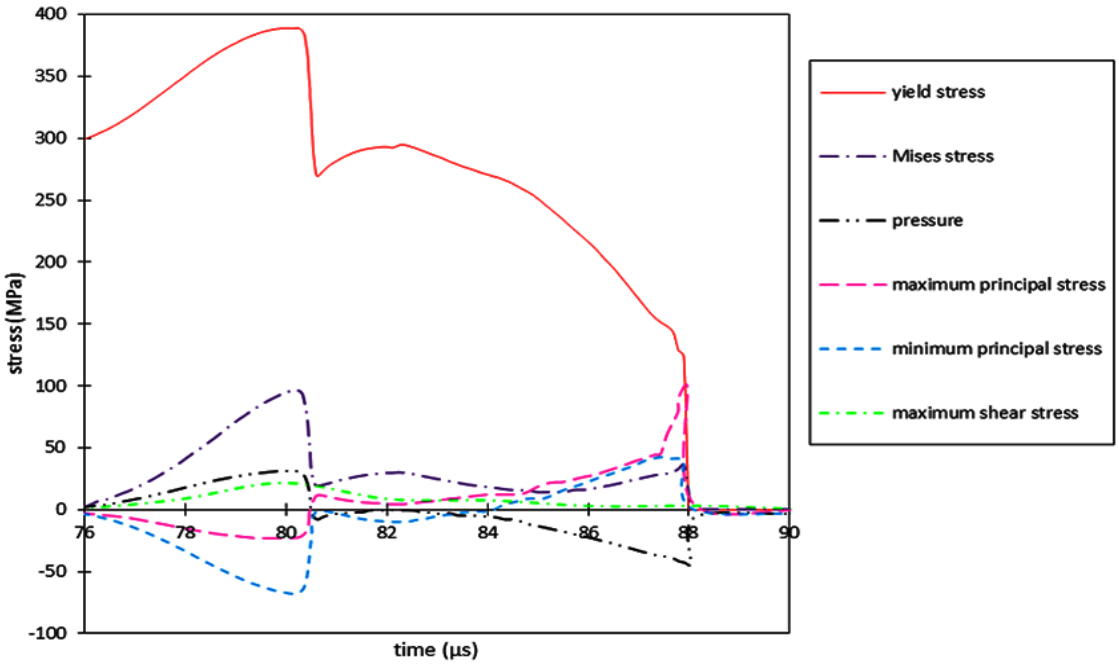


Figure 8. Relationship between pressure, stresses and strength versus time for failed point 3 (shown with a green star in Figure 5).

In Figures 6, 7 and 8, when the Mises stress becomes larger than the intact strength, it returns back to the yield surface and rock undergoes plastic deformation. With increasing permanent deformation, damage in the material accumulates and its strength gradually decreases from the intact strength to a lower strength [19].

In order to quantify the variations of crack patterns in their experimental tests, crack densities, defined as length of cracks per unit area, were calculated at different depths and radial distances from the specimen borehole [22]. Each image was then calibrated spatially and divided into three equal width zones from the centre to the edge with zone 1 being in the centre and zone 3 the most outside. Crack densities were then calculated by dividing the total length of all the cracks in each category by the zone area.

In calculating the crack density, each SPH particle is considered occupy space of a cube with size of 0.15 mm. The following equation was used to calculate the crack density

$$\Gamma = \frac{\sum_{i}^{n} D_i \times A_i^{1/2}}{A_{zone_j}} = \frac{L_i \times \sum_{i}^{n} D_i}{A_{zone_j}} \tag{12}$$

where Γ, D_i, L_i, A_i and $A_{(Zone\ j)}$ are crack density, the damage of particle i, length of particle i, area of particle i, area of zone j. In this work, any particle with D>0.8 is considered crack. Figure 9 shows the crack density in Barre granite for water-coupled and air-coupled with and without copper as well as the variation of measured crack densities for top, middle and bottom surfaces of rock, respectively. Crack density was measured in terms of mm/mm^2. Investigation of the crack patterns on cross sections of rock (Figure 5) reveals that by moving from top to the bottom of the specimen, the number of smaller cracks and the number of longer cracks are increased. This is related to the stronger stress field, due to confinement at greater depths, at distances away from the top surface of the specimen.

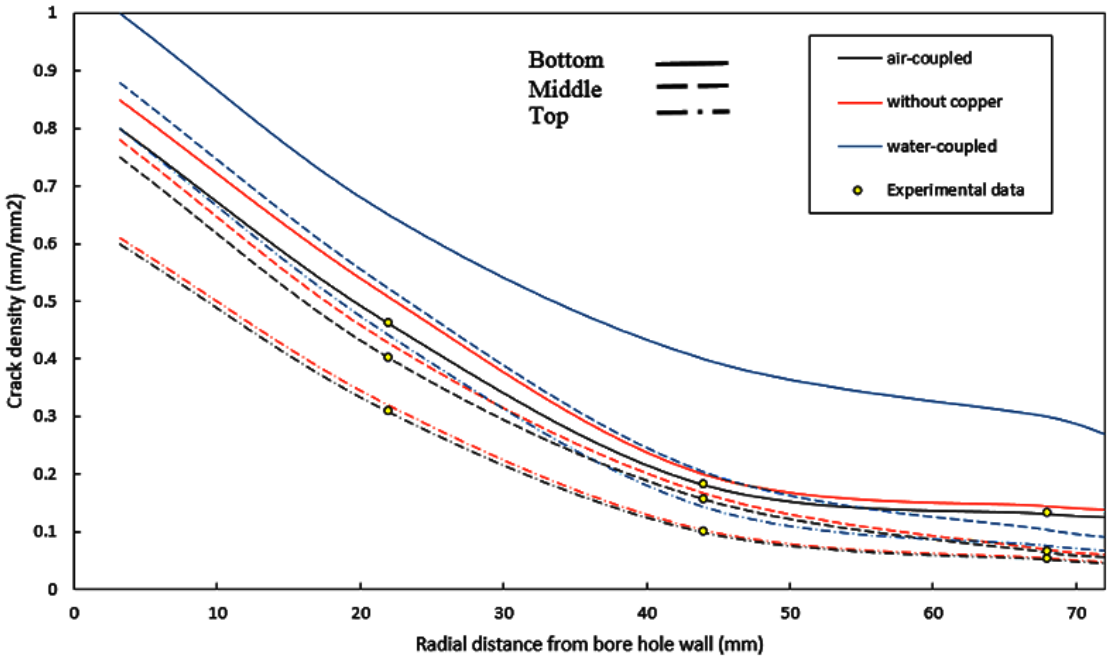


Figure 9. Variation of crack density in radial and axial directions of rock for experimental [22] and three simulations.

It is seen that for the three cross sections of rock, crack densities around the borehole are much higher than far distances from the borehole. This is not surprising since the rock adjacent to the borehole experiences higher stress intensities. On the other hand, maximum crack density is observed at the bottom region of rock. Close to the borehole, the difference between crack densities is minimum. In fact, near the borehole crack densities at top, middle and bottom regions are very similar. As shown in Figure 9, our simulations match well with the experimental results.

The difference in crack density from simulations with and without copper was observed to be marginal, especially at the top and mid-sections of the specimen, while a noticeable difference was noted at bottom section closer to the borehole. The number of damaged particles in bottom section for two simulations of air-coupled with and without copper are shown in Figure 10. The induced damage rate for simulation without copper is much higher than that

with copper. The rate of damaged particle generation is reduced as the redial cracks gradually lose their initial momentum to extend. This is suggested by the drastic difference in the slopes of the Figure 10. The difference between the number of damaged particles for simulations with and without copper is the kinetic and elastic strain energies that are absorbed by the copper as a liner.

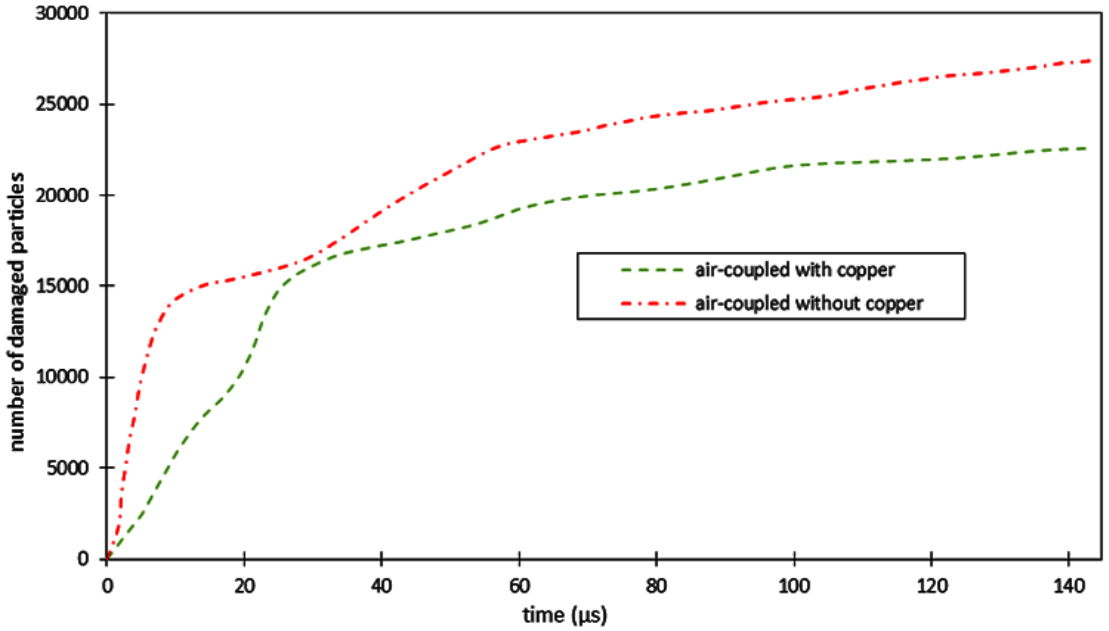


Figure 10. Number of damaged particles in bottom section for air-coupled simulation with and without copper.

Crack density is much higher for the water-coupled case in comparison with air-coupled case with and without copper. This is to be expected, as water is known to serve as an excellent medium for transmission of shock and stress wave, compared to air.

Crack distribution probability

Crack distribution probability is approximated by a function defined on damage level and number of damaged particles for comparison in radial direction. In this method, each cross section with normal direction along the Z axis, is divided into 72 segments with the bin size of 5 degree. Then the number of particles with damage level equal to or above 0.8 is counted and divided by the total number of damaged particles in that cross section.

$$PDF = \frac{\sum_{i}^{n} D_i}{\sum_{i}^{72}\sum_{i}^{n} D_i} \frac{1}{bin/360} \tag{13}$$

where *PDF* is probability distribution function and *n* is the number of damaged particles. For comparison of different simulations, PDF function is normalized.

Figure 11 shows the polar graphs which compare the normalized crack distribution probability for three simulations air-coupled, without copper and water-coupled from time 23 μs to 144.88 μs. It can be clearly seen that the normalized probability at time 23 μs in all direction (0^o-360^o) around the borehole is the same for these three simulations. The normalized probability for air-coupled, without copper and water-coupled are 0.000942,

0.001682 and 0.003558, respectively. At time 69 μs, the normalized probability is different from 0^o to 360^o direction around the borehole for three simulations, however, this distribution shows a symmetrical pattern around the borehole. The same results at time 69 μs are obtained for without copper and water-coupled simulations.

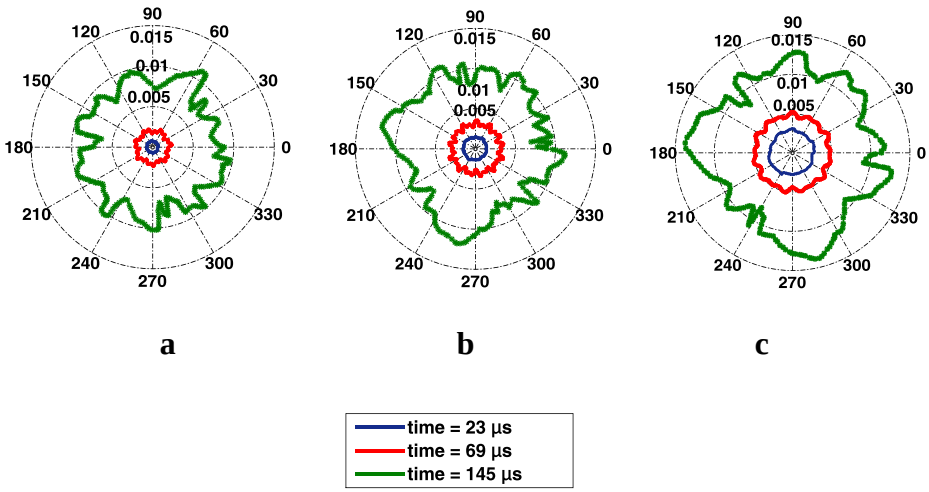


Figure 11. Normalized crack distribution probability in bottom section a) air-coupled b) without copper c) water-coupled.

At time between 69 μs and 144.88 μs, the normalized probability distribution shows an asymmetrical patterns around the borehole for three simulations. This is mainly due to the specific failure criterion applied. In the simulation, if a particle is failed, it will not be able to sustain any tensile and shear loadings, i.e., its tensile and shear stresses will release, which will significantly affect the stress field around this failed particle, therefore, the particle stresses could be different. Therefore, stress interaction between a fracture and a field of cracks in the three dimensional space generate irregular fracture pattern. It can be seen that the resulting probability distributions are quite different each other. Water-coupling produces the most extended crack zones probability, and air-coupling the least.

Conclusions

This study presented an approach that allows the prediction of 3D fracture structure under different blast loading by SPH method. The validity and predictability of models obtained by comparing fracture patterns to experimental works. The main conclusions are:

(1) It is found that the SPH simulations with the JH2 model are very promising in effectively elucidating the complicated and fundamental failure mechanisms of rocks. The more accurate type of simulation has been made possible by the development of the SPH method.

(2) Adopted SPH method permits the 3D quantification of the fracture patterns. The quantification results are in satisfying agreement with experimental measurements.

(3) The study showed that radial cracks propagate along different directions with different distributions and the propagation direction. The crack distribution probability was found to be asymmetrical around blast hole. Numerical results show that densely cracked region around the borehole, radial and spall cracks match well with experimental results. The agreement is both qualitative and quantitative.

Acknowledgments

The work was supported in part by the Australian Research Council through Discovery Projects (DP140100945 and DP170102886) and by the National Natural Science Foundation of China (Grant No. 11232003). This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government.

References

[1] Sarbhukan, M. M. (1990) Application of Bore - Blast technique for rock fracturing for drinking water: A case study of Ghotkarwadi, Taluka Akole District Ahmednagar. Proceedings of national seminar on Modern techniques of rain water harvesting, water conservation and artificial recharge for drinking water, afforestation, horticulture and agriculture, Government of Maharashtra (India), pp. 827-838.

[2] Stephens, D. R. (1974) Revised cost estimate for the LLL in-situ coal gasification concept. Lawrence Livermore Laboratory, Livermore, Calif., Rept. UCRL-51578.

[3] Austin, C. F. and Leonard, G. W. (1973) Chemical explosive stimulation of geothermal wells. Geothermal Energy Resources, Production and Stimulation, Stanford U. Press, Stanford, Calif. 269-292.

[4] Miller, J. S., Walker, C. J., and Eakin, J. L., (1974) Fracturing oil shale with explosives for in-situ oil recovery. USBM, Bartlesville, Okla., Rept. RI 7874.

[5] Zhu, Z., Xie, H., Mohanty, B. (2008) Numerical investigation of blasting-induced damage in cylindrical rocks. International Journal of Rock Mechanics and Mining Sciences **45**, 111-121.

[6] Shin, YS., Chisum, JE. (1997) Modelling and simulation of underwater shock problems using a coupled Lagrangian-Eulerian analysis approach. *Shock and Vibration* **4**, 1–10.

[7] Potyondy, D., Cundall, P., Sarracino, R. (1996) *Modelling of shock- and gas-driven fractures induced by a blast using bonded assemblies of spherical particles.* Rock fragmentation by blasting. Rotterdam, Netherlands: A.A. Balkema. p. 55–62.

[8] Trivino, L.F.n., Mohanty, B. (2015) Assessment of crack initiation and propagation in rock from explosion-induced stress waves and gas expansion by cross-hole seismometry and FEM–DEM method. *International Journal of Rock Mechanics & Mining Sciences* **77**, 287–299.

[9] Libersky, L.D., Randles, P.W., Carney, T.C., et al. (1997) Recent improvements in SPH modelling of hypervelocity impact. *International Journal of Impact Engineering* **20**, 525–532.

[10] Chen, Z., Hu, W., Shen, L., Xin, X., and Brannon, R. (2002) An Evaluation of the MPM for Simulating Dynamic Failure with Damage Diffusion, *Engineering Fracture Mechanics* **69**, 1873-1890.

[11] Flores-Johnson, E.A., Wang, S., Maggi, F., El-Zein, A., Gan, Y., Nguyen, G.D., and Shen, L. (2016) Coupled discrete element – finite element method for modelling dynamic impact of unsaturated sand. *International Journal of Mechanics and Materials in Design*, **12**(4), 495-507.

[12] Pramanik, R. Deb, D. (2014) Implementation of Smoothed Particle Hydrodynamics for Detonation of Explosive with Application to Rock Fragmentation. *Rock Mechanics Rock Engineering* **16**, 92–99.

[13] Gharehdash, S., Shen, L., Gan, Y., Flores-Johnson, E. A. (2016) Numerical Investigation on Fracturing of Rock under Blast Using Coupled Finite Element Method and Smoothed Particle Hydrodynamics. *Applied Mechanics and Materials* **846**, 102-107.

[14] Fakhimi, A., Lanari, M. (2014) DEM–SPH simulation of rock blasting. *Computers and Geotechnics* **55**, 158–164.

[15] Gingold, R., Monaghan, J. (1977) Smoothed particle hydrodynamics theory and application to non spherical stars. *Monthly Notices of the Royal Astronomical Society* **181**, 375–389.

[16] LS-DYNA LSTC. (2010) *keyword user's manual*, Version 970, Livermore Software Technology Corporation, Livermore, CA.

[17] Monaghan, J. (2000). SPH without a Tensile Instability, *Journal of Computational Physics* **159**, 290 311.

[18] Rabczuk, T., Belytschko, T., Xiao, SP. (2004) Stable particle methods based on Lagrangian kernels. *Computer Methods in Applied Mechanics and Engineering* **193**, 1035–63.

[19] Johnson, GR., Holmquist, TJ. (1999) Response of boron carbide subjected to large strains, high strain rates, and high pressures. *Journal of Applied Physics* **85**, 8060–73.

[20] Dehghan Banadaki, M.M., Mohanty, B. (2012) Numerical simulation of stress wave induced fractures in rock. *International Journal of Impact Engineering* **41**, 16-25.

[21] Johnson, G.R., Cook. W.H. (1983) *A constitutive model and data for metal subjected to large strains, high strain rates and high temperature*. In Proceedings of the 7th International Symposium on Ballistics. The Hague, Netherlands.

[22] Dehghan Banadaki, MM. (2010) *Stress-wave induced fracture in rock due to explosive action*. PhD thesis. Canada: University of Toronto.

Effect of turbulent Schmidt number on the scalar field simulation of a fluidic precessing jet flow

†Xiao Chen, *Zhao Feng Tian and Graham 'Gus' Nathan

[1]Center for Energy Technology,
The School of Mechanical Engineering
The University of Adelaide, Adelaide, SA 5005, Australia

*Presenting author, † Corresponding author: xiao.chen01@adelaide.edu.au

Abstract

A preliminary numerical study of the external scalar field of a fluidic precessing jet (FPJ) flow is reported. The unsteady Shear Stress Transport (SST) model, which showed generally good agreement with the measured velocity field of flows within a FPJ nozzle, despite some discrepancies, was adopted to assess the effect of the turbulent Schmidt number, Sc_t, on the simulated scalar field. The simulated jet axis concentrations with three turbulent Schmidt number values, Sc_t =0.5, 0.9 and 1.3 have been compared with the measured results in the literature. It is found that the SST model over-predicts the centreline concentration of the jet in the near field downstream from the exit of the FPJ nozzle, while under-predicts it in the far field. An increase in Sc_t number causes the simulated jet to be more distributed away from the axis of the confinement, which is in contrast to the measured data. The best agreement with the measured result was achieved by adopting a Sc_t number of 0.5. However, due to the complexity of the FPJ flow, it is not feasible for a two-equation URANS model to reliably reproduce the scalar field by simply adjusting the turbulent Schmidt number.

Keywords: scalar mixing, turbulent Schmidt number, CFD, URANS

Introduction

A fluidic precessing jet (FPJ) nozzle, which has been employed in industrial rotary kilns was proposed by Nathan [1] to generate the FPJ flow. Numerous previous investigations of the flows within the FPJ nozzle have conducted by experimental measurements [2, 3], analytical method [4] and recently computational fluid dynamics (CFD) [5, 6]. Compared with the many studies of flows within the FPJ nozzle, the study of scalar mixing downstream the FPJ nozzle is much less. Parham [7] and Parham et al. [8] measured the scalar field downstream the FPJ nozzle using a two colour planar laser-induced fluorescence technique. In these studies, the effects of co-flow, confinement and a shaping jet on the mixing characteristics are investigated. Nevertheless, a reliable CFD model of the scalar field downstream the FPJ nozzle is still lacking. Hence the main objective of this paper is to preliminarily assess the feasibility of the unsteady Reynolds-Averaged Navier-Stokes (URANS) turbulence model in predicting the scalar field of the FPJ flow.

In the previous CFD studies of the flows within the FPJ nozzle [5, 6] conducted in the authors' group, the authors have assessed the performance of the Shear Stress Transport (SST) model in modeling the turbulent flows within the FPJ nozzle. The SST model achieves reasonable agreement with the measured mean axial velocity profiles for pipe and contraction inlets [6], and the measured centerline velocity decay and equivalent diameters of the phase-averaged precessing jet for the contraction inlet case despite some discrepancies [5]. Therefore the current CFD model of the scalar field downstream the FPJ nozzle is based on the SST model.

When applying the URANS models for the scalar field simulation, the turbulent Schmidt number (Sc_t), which is defined as the ratio of the turbulent eddy viscosity (μ_t) and the turbulent diffusivity (Γ_t), was found to have great influence on the simulated turbulent mixing [9]. In a numerical simulation of a jet in a cross-flow simulated with the standard k-ε model [10], the range of the turbulent Schmidt number, from 0.5 to 0.9, was found to have slight influence on the simulated temperature field. In another study of temperature field in a cross-flow [11], it was also found that the simulated temperature distribution is not sensitive to the turbulent Schmidt number. However, it is found that the turbulent Schmidt number has a significant influence on the simulated scalar mixing and Sc_t =0.2 is recommended for the simulation of a jet in a cross-flow [11]. Nevertheless, how the turbulent Schmidt number will influence the simulated scalar mixing of a FPJ flow is unclear.

The main aim of the current paper is to assess the effect of turbulent Schmidt number on the simulated scalar field in a flow downstream a FPJ nozzle. The FPJ flow investigated in [7, 8] are simulated using the SST model. The simulated jet concentration based on the turbulent Schmidt number, Sc_t =0.9, is compared with measured values reported in [7]. The default value of the turbulent Schmidt number is normally set as 0.9 (e.g. in ANSYS/CFX and ANSYS/FLUENT). A lower Sc_t number of 0.5 is tested in the current study following the work [10], and a larger Sc_t number of 1.3 is also tested.

Numerical Model

ANSYS/Designmodeler 16.5 was used to generate the 3-dimensional CFD geometry. Figure 1 illustrates geometric configuration of the CFD domain that includes an FPJ nozzle with a contraction inlet and a confinement. This geometry is identical to the configuration in the experimental studies [7]. The confinement is a cylindrical domain which has a diameter of 390 mm and a length of 1100 mm. The FPJ nozzle has a length of 115 mm from the main flow inlet to the exit of the contraction and 110 mm from the contraction exit to the nozzle exit. The inner diameter of the FPJ nozzle is 38 mm and the diameter of the center body is 27 mm. More details of the FPD nozzle can be found in the literature [7].

Figure 2 presents the details of the CFD mesh that was generated with the software ANSYS/ICEM CFD 16.5. The O-grid method was adopted to generate the structured mesh to

ensure the mesh quality, especially in the near wall region, which makes the y+ values to be less than 1. A mesh independent test has been conducted and a final mesh of 8.6 million nodes is used for the study.

The CFD software ANSYS/CFX 16.5 was adopted for the simulations. The flow of the CFD model is a multiple component fluid that includes two fluids. Water at 25 °C was employed as the first fluid (fluid1). A second fluid (termed as fluid2 in the model) is used for the co-flow, while its properties are exactly the same as those of the fluid1 at the main inlet. This numerical approach matches the experiment [7,8] in which the jet fluid was water marked with a fluorescent dye. The dye in the experiment is with very low concentration and its effects on the water dynamic properties can be neglected. At the main flow inlet (see Figure 1b), mass fraction of fluid1 is 1, i.e. the mass fraction of fluid2 is 0. At the co-flow inlet, mass fraction of fluid1 is 0 and the mass fraction of fluid2 is 1.

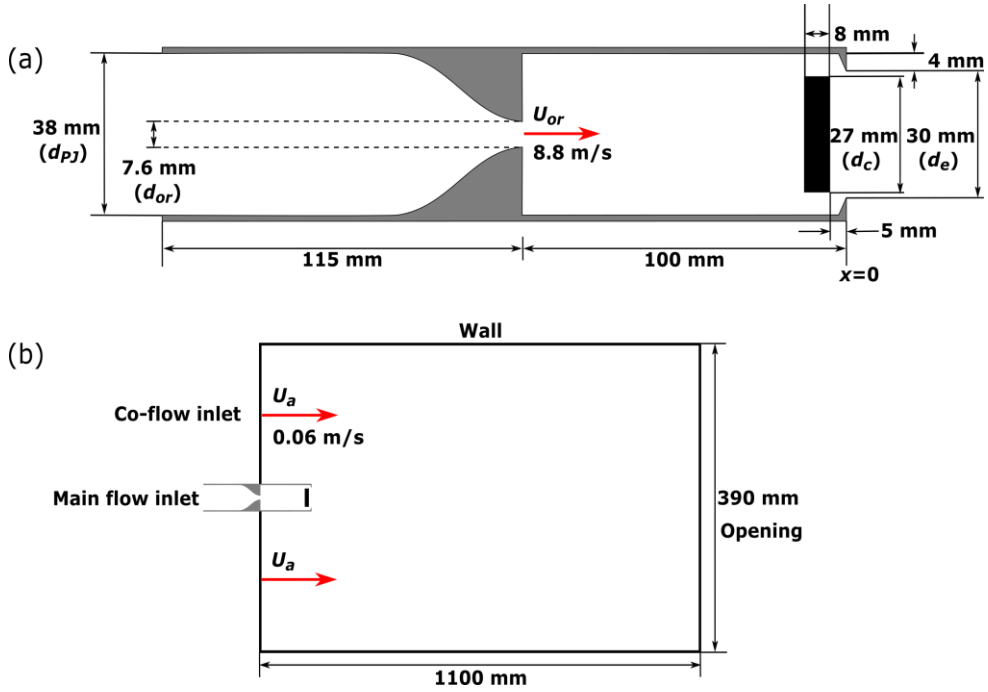


Figure 1: The geometric configurations of (a) the FPJ nozzle and (b) the whole fluid domain. Here d_{PJ}, d_{or}, U_{or} and U_a are the diameter of the nozzle, diameter of the inlet orifice, the inlet velocity at the orifice and the co-flow velocity, respectively.

For multi-component gases, the continuity equation and momentum equation after Reynolds averaging [12] are given as below:

$$\nabla \cdot U = 0 \tag{1}$$

$$\frac{\partial(\rho U)}{\partial t} + \nabla \cdot (\rho UU) = -\nabla p + \nabla \cdot \left\{ \mu \left[\nabla U + (\nabla U)^T \right] - \rho \overline{UU} \right\} + S_M \tag{2}$$

here U is the mean velocity vector, P the mean pressure, S_M the external momentum source, and ρ the mixture density that is calculated as:

$$\rho = \sum_{I=1}^{Nc} Y_I \rho_I \tag{3}$$

where ρ_I is the density of the component I. *Nc* is the number of modelled species in the mixture, and Y_I is the mass fraction of the species I that is solved by the following equation:

$$\frac{\partial(\rho Y_I)}{\partial t} + \nabla \cdot (\rho U Y_I) = \nabla \cdot (\Gamma_{I.eff} \nabla Y_I) + S_I \quad (4)$$

here S_I is the source term of the species. The effective diffusion coefficient of species I, $\Gamma_{I.eff}$, in Equation 4 is calculated as [12]:

$$\Gamma_{I.eff} = \Gamma_I + \frac{\mu_t}{Sc_t}, \quad (5)$$

Where Γ_I is the molecular diffusion coefficient of species, $\Gamma_I = \rho_I D_I$, D_I the kinematic diffusivity of the species I, μ_t the turbulent viscosity.

The root mean square (r.m.s) residuals are all under 5 x 10^{-5}. The high resolution scheme and the second order backward Euler scheme were adopted for the advective and transient terms, respectively.

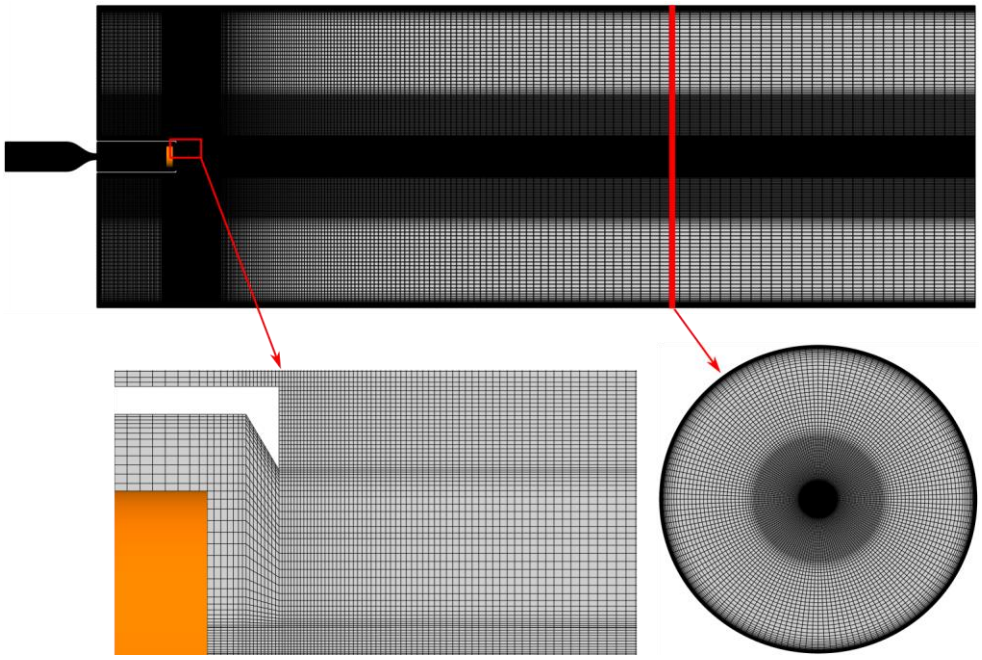


Figure 2: Mesh of the CFD model.

Preliminary Results and Discussion

Figure 3 compares the measured [7] and simulated instantaneous concentration of the jet on a cross-sectional plane downstream the nozzle exit with the Sc_t value of 0.5, 0.9 and 1.3, respectively. It can be seen that the small eddies and scalar concentration fluctuation in the measured instantaneous flow have not been reproduced with any of the three URANS approaches. The three simulated deflected angles between the instantaneous jets and the nozzle axis (indicated as the dashed lines) are similar, although they all appear to be larger than the measured data. It is observed that the value of Sc_t does not have a significant influence on the simulated flows within the FPJ nozzle. This observation can be explained by reviewing equations 1-4 and the inlet conditions in the flow. The turbulent Schmidt number is

only incorporated in Equations 4, which describes the transport of mass fraction of species. That is, it influences only the mixture density (Equation 3) and then the mixture flow field (Equations 1 and 2). If there is no mixing process, i.e. the mass fraction of a species is 1, the change of Sc_t will effect neither the mixture density nor the mixture species field. In the current case, the mass fraction of fluid1 is 1 at the main flow inlet and is nearly 1 anywhere inside the FPJ nozzle, except in a small region near to the nozzle exit where the mixture in the emerging field can be entrained into the FPJ nozzle. Nevertheless, as shown in Figure 4, the mass fraction of the mixture entrained into the FPJ nozzle exit region is very small and can be neglected.

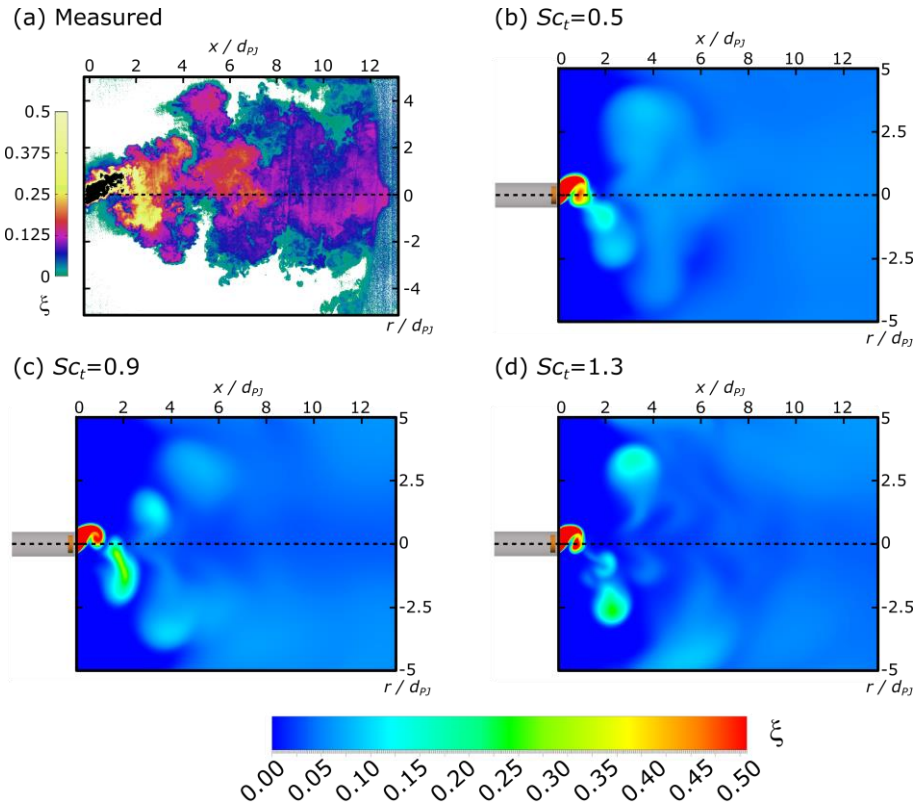


Figure 3: Instantaneous concentration cross-sectional contours of the FPJ flows that were (a) measured [7] and simulated using the SST model with the Sc_t value of (b) 0.5, (c) 0.9 and (d) 1.3.

The effect of the turbulent Sc_t number on the mixing process within the external flow is more pronounced than within the chamber, as is shown in Figure 3b-d. It can be seen that when the turbulent Sc_t number increases from 0.5 to 1.3, the diffusion of the fluid1 in the external flow decreases, leading to higher conentration of fluid1 in some flow vortexes. This can be explained by looking at Equation 5, when the turbulent Sc_t number increases, the effective diffusion coefficient of species I decreases, leading to lower diffusion of fluid1 as shown in Equation 4.

Figure 5 compares the measured [7] and simulated centreline concentration with three turbulent Schmidt numbers, namely, Sc_t = 0.5, 0.9 and 1.3. The measured centerline concentration of the jet exhibits a fast decay in the region from the FPJ exit (x/d_{PJ}=0) to an

"elbow point" (x/d_{PJ}≈1.4). At the "elbow point", the decay rate suddenly decreases and is nearly constant downstream from that point. This trend has been simulated with all the three models, however, the distance between the "elbow point" and the nozzle exit is over-predicted. For all CFD cases, the centreline jet concentration is over-predicted in the emerging field (say x/d_{PJ}<0.5) and is under-predicted in the region downstream from about x/d_{PJ}=1, especially in the region near to the "elbow point".

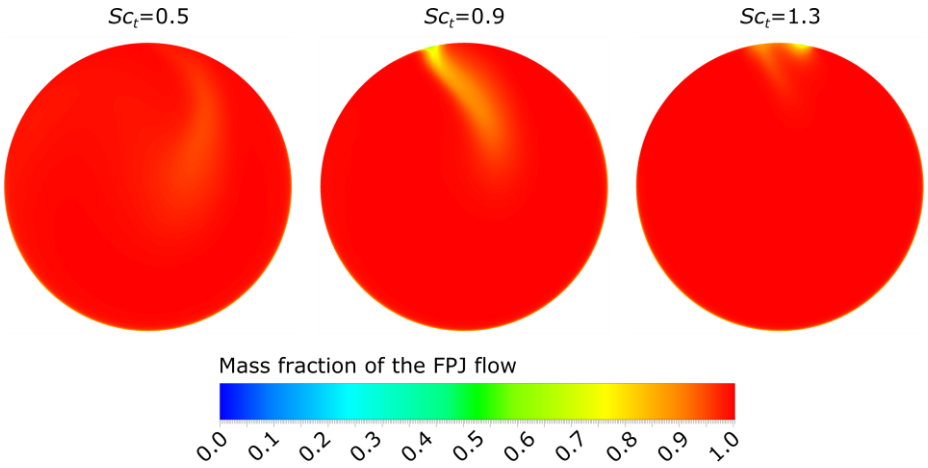


Figure 4: Instantaneous concentration contours of the Fluid1 on the nozzle exit (x=0) that were simulated using the SST model with the Sc_t value of 0.5, 0.9 and 1.3.

Figure 5 also shows that a decrease in Sc_t number leads to a decrease in centreline concentration, indicating a higher simulated mixing in the near field (x/d_{PJ}<0.5). This is consistent with Equations 4 and 2 that a decrease in Sc_t number causes the effective diffusivity to increase, hence improves the simulated scalar mixing. However in the far field (x/d_{PJ}>1), model with the lowest Sc_t number (Sc_t=0.5) simulated the highest centreline jet concentration. One possible reason is that, based on the observation of the simulated instantaneous jet concentration, the jet flow was simulated to be distributed preferentially in the near wall region as the increase of Sc_t number. Although the best agreement against the measured data was achieved by adopting a turbulent Schmidt number of 0.5, this model greatly under-predicts the centreline concentration of the jet in the region near to the "elbow point" (x/d_{PJ}≈1.4). This implies that the mechanism is not one of "turbulent" diffusion, but is rather controlled by the exit angle of the emerging jet, which is a function of large scale turbulent flow features.

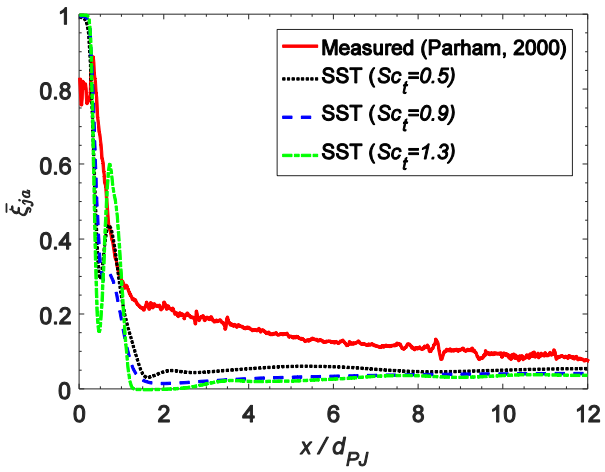


Figure 5: Measured [7] and simulated mean centerline concentration of the jet.

Summary

A decrease in Sc_t number causes the centreline concentration of the FPJ flow to decrease in the emerging field and increase in the far field. However, all the three approaches are found to over-predict the centreline concentration of the jet in the emerging field downstream the exit of the FPJ nozzle while under-predict it in the far field. This is due to large-scale features of the turbulent flow rather than to small scale "diffusion" processes, which are modelled by the turbulent Schmidt number. For this reason, the level of disagreement between the measured and predicted values of the scale field is not improved by changes to the turbulent Schmidt number. Indeed, the trends are opposite to what would be expected.

The flow in the far field of the confined, co-flowing turbulent FPJ flow is simulated to be distributed mostly in the region near to the wall of the confining cylinder, while the measured jet is preferentially distributed near to the axis. This discrepancy between the simulated and measured scalar field may be attributed to the over-predicted angle of the jet emerging from the FPJ nozzle.

Acknowledgements

The first author was supported by the Divisional Scholarship of the faculty of ECMS, the University of Adelaide. The authors gratefully acknowledge the support of the Australian Research Council through Grant DP150102230.

References

[1] Nathan, G., 1988, "The enhanced mixing burner." PhD Thesis, the University of Adelaide.
[2] Nathan, G., Hill, S., and Luxton, R., 1998, "An axisymmetric 'fluidic'nozzle to generate jet precession," Journal of Fluid Mechanics, 370, pp. 347-380.
[3] Wong, C., Lanspeary, P., Nathan, G., Kelso, R., and O'Doherty, T., 2003, "Phase-averaged velocity in a fluidic precessing jet nozzle and in its near external field," Experimental Thermal and Fluid Science, 27(5), pp. 515-524.
[4] Kelso, R., 2001, "A mechanism for jet precession in axisymmetric sudden expansions." In the Proceeding of 14th Australasian Fluid Mechanics Conference, pp. 829-832.

[5] Chen, X., Tian, Z., Kelso, R., and Nathan, G., 2017, "The topology of a precessing flow within a suddenly expanding axisymmetric chamber," ASME Journal of Fluids Engineering, 139(7), 071201.
[6] Chen, X., Tian, Z. F., Kelso, R. M., and Nathan, G. J., 2017, "New understanding of mode switching in the fluidic precessing jet flow," ASME Journal of Fluids Engineering, 139(7), 071102.
[7] Parham, J. J., 2002, "Control and optimisation of mixing and combustion from a precessing jet nozzle," PhD thesis, the University of Adelaide.
[8] Parham, K., Esmaeilzadeh, E., Atikol, U., and Aldabbagh, L., 2011, "A numerical study of turbulent opposed impinging jets issuing from triangular nozzles with different geometries," Heat and Mass Transfer, 47(4), pp. 427-437.
[9] Reynolds, A., 1975, "The prediction of turbulent Prandtl and Schmidt numbers," International Journal of Heat and Mass Transfer, 18(9), pp. 1055-1069.
[10] Chao, Y.-C., and Ho, W.-C., 1991, "Heterogeneous and nonisothermal mixing of a lateral jet with a swirling crossflow," Journal of Thermophysics and Heat Transfer, 5(3), pp. 394-400.
[11] He, G., Guo, Y., and Hsu, A. T., 1999, "The effect of Schmidt number on turbulent scalar mixing in a jet-in-crossflow," International Journal of Heat and Mass Transfer, 42(20), pp. 3727-3738.
[12] Tian, Z. F., Witt, P. J., Schwarz, M. P., and Yang, W., 2018, "Numerical Modelling of Pulverised Coal Combustion," Handbook of Multiphase Flow Science and Technology, G. Yeoh, ed., SPRINGER Singapore.

Riemann Function and Relativistic Structure

Wang Yiping [1,2]

[1] Senior Science & Technology Workers Association of Quzhou, Zhejiang 324000, China
[2] Research Institute of Mathematics and Power Engineering, Qianjiang, Zhejiang 324000, China
Author: wyp3025419@163.com

Abstract
It has pointed out that pointwise is a quantized unit concept of Riemann function, which features random "asymmetry, inhomogeneity and discontinuity". It establishes the abstract relativistic structure of dimensionless quantities(called circular logarithm and super symmetric matrix unit) by application of the principle of relativity so as to ensure the "normativity and invariability" of every pointwise numerical value, location, property, topology and zero error, and achieve, with Riemann function, "topological variational rules without any specific content and accurate resolution that the number of elements and critical points (1/2) are all at the $\{1/2\}^Z$straight line". This computing method is simple, stable, self-consistent and pragmatic and features extensive applicability for multiple disciplines.
Keywords: Riemann function, pointwise quantization, relativity structure (circular log), critical point of limit value

1. Foreword

In 1859 the German mathematician Bernhard Riemann proposed the Riemannξ function, a function composed of reciprocal sum. It has to meet the requirement by "On the Number of Prime Numbers Less Than A Given Value". "The real part of all non trivial zeros is (1/2)"is called the Riemann hypothesis. The difficulty is infinite proceduresand infinite sets (infinity) 【2】p35.

Via his book *Prime Obsession*【1】, John Derbyshirehas systematically introduced the researches into the Riemann function by mathematicians from various nations for over 150 years. In particular, he points out: the improved prime number theorem PNT 【1】p98, golden key 【1】p98-105, prime calculusform 【1】p111and final exact expressions of Riemann function J (x) 【1】p325have created conditions for theory research by upper generations.

In this paper, the Riemann function is integrated into the pointwise quantized function, and with the help of the principle of relativity, abstract dimensionless circular log (called relativistic structure or super symmetric unit matrix) is established, which results in gauge invariances such as the value, location, space, topology and properties of each element of the Riemann function with realization of "topological variation rule with no specific content, by which, primes are accurately resolved in the range of $[0\sim1/2\sim1]^{[0\sim1/2\sim1],}$ and it is proven that non normal zeroes are on the critical line of $\{1/2\}^Z$" It might become the latter part of *Prime Obsession.*

2. Pointwise Quantized Riemann Function

2.1.Basic Definition

Definition 1. Pointwise Space and Riemann Function:
Pointwise Riemann function in infinite elements at all levels (including random "symmetric and asymmetric, uniform and non-uniform, continuous and discontinuous, tangible and intangible space", making up the space and state of pointwise quantized Riemann function. It is briefly referred to as pointwise space.
With:

$$\xi(x)=\sum\{x_r\}^Z=\{x\}^{K(S\pm N)}=\{x\}^Z$$
$$=\{(1+2^{-s}+3^{-s}+4^{-s}+5^{-s}+6^{-s}+7^{-s}+\ldots)^{-1}\};$$

$$= Ax^{k(Z-0)}+Bx^{k(Z-1)}+\ldots+Px^{k(Z-p)}+\ldots+Qx^{k(Z-q)}; \quad (1)$$

Formula (1) re-inverses the sum of the reciprocal inverse of the Riemann function without losing generality. The convergence of the Riemann function is controlled via K= (+1,0, -1). The independent variable (-S) converts the infinite program power function K (Z-p).

Definition 2: Pointwise space and golden key [1] p104

Because:

$$1/\xi(x)=\prod\{x_p\}^Z=\prod_p(1-p^{-S})^{-1}$$
$$=\prod(1-1^{-S})(1-3^{-S})(1-5^{-S})(1-7^{-S})\ldots; \quad (2)$$

p stands for “taken over all the prime numbers”.

Namely:

$$\prod\{x_p\}^Z=\sum\{x_r\}]^Z; \quad (3)$$

Derbyshire called it the golden key, which was a way of expressing the Eratosthenes sieve method in 1737 by the Euler product.

Definition 3. Definition of the pointwise average space

The pointwise average space reflects the average combination of point state space, which has nothing to do with the internal components.$(C_{1+r}^{-1})\sum\{x_r\}^Z=(C_{1+p}^{-1})\prod\{x_p\}^Z$. Select (C_{1+h}^{-1}) to contain the number of the combination of the golden key (C_{1+r}^{-1}) natural number and (C_{1+p}^{-1}) prime number.

$$\xi(x_0)=\{x_0\}^Z=J(x)$$
$$=\sum[\ (C_{1+h}^{-1})\prod\{x\}^Z$$
$$=\sum[\ (C_{1+h}^{-1})\sum\{x\}^Z; \quad (4)$$

It is easy to prove that the round log helps reflect the combination of the pointwise quantized Riemann function that the products in the same level are equivalent of reciprocal combination.

Definition 4. Definition of Pointwise Combination Coefficient

In the infinite dimensional space, the pointwise states have a number of infinite regular combinations from "(S±0) -(0±S)" to "(S±N) -(N±S)”, which are called coefficients.

$$(C_{1+S}) = (C_{S+1})$$
$$=(S!)/(N!)= (S-0)\ (S-1)\ldots(S-p)!$$
$$/\ (N-0)\ (N-1)\ldots(1)!\ ; \quad (5)$$

In the formula: (C_{1+S}) represents the regularized coefficient of pointwise combination. (S!) represents that the factorial number of combination number N!) refers to the factorial of the order.

Definition 5. Definition of Power Function Equation

$$Z=Z/T,t= K(S\pm N\pm N\pm p)/T,t$$
$$= \{K(S\pm N\pm N\pm 0)+K(S\pm N\pm N\pm 1)+\ldots$$
$$+ K(S\pm N\pm N\pm p)+\ldots+ K(S\pm N\pm N\pm q)\}/T,t:$$
$$= \{\ (Z\pm 0)\ ,(Z\pm 1),\ldots,(Z\pm p),\ldots,(Z\pm q)\ \}/T,t; \quad (6)$$

Power function equation(S±N±N±p±p) contains time - thermal - properties - space as a whole, known as the “path integral exp”.

In the formula: (S±N) refers to infinite dimensional, and (±N)finite dimensional; (±N) refers to calculus order; (±P) polynomial sequence, T (temperature) and t (time). Based on the function of time, synchronization of the thermodynamic function and the pointwise space, there will be no separate description unless necessary.

Definition 6.Pointwise Calculus

The traditional calculus sign Mark $d\{x\}^N/dt^N$ (including partial differential method) and $\int^N f\{x\}dt^N$become the pointwise calculus:$F\{\ x^{K(S\pm N\pm p\pm N)}\}$;

Pointwise differential order:**N** =-**N**;

$$F\{x\}^{K(S\pm N\pm p\ -n)}:$$

Pointwise integral order: **N**=+**N**;

$$F\{x\}^{K(S\pm N\pm p+n)};$$

Pointwise calculus:

$$\{x\}^{Z}=\{x\}^{K(S\pm N\pm p\pm 0)}$$
$$+\{x\}^{K(S\pm N\pm p\pm 1)}+\ldots+\{x\}^{K(S\pm N\pm p\pm p)}+\ldots+\{x\}^{K(S\pm N\pm p\pm q)}$$
$$=\{x\}^{K(Z\pm 0)}+\{x\}^{K(Z\pm 1)}+\ldots+\{x\}^{K(Z\pm p)}+\ldots+\{x\}^{K(Z\pm q)};$$

In the formula, {} represents a collection of pointwise space, distinct from a set of logical algebra.

3. Pointwise Gauge Invariance Theorem

3.1.Theorem 1: Homology Logarithm

Definition of isomorphic circular logarithm: Each subitem of pointwise space is divided by (multiplied) by the total pointwise space to get “to be divided by itself, the maximum value is 1”. It is called the first type of gauge invariance.

To prove:

$$(1-\eta_h^{\,2})^{Z}\sim(\eta_h)^{Z}=\xi(x)_h\ /\xi(x)_H$$
$$=[\{x_h\}/\{x_H\}]^{Z}$$
$$=[x_1+x_2+\ldots+x_p+\ldots+x_q]^{Z}/\{x_H\}^{Z}$$
$$=[x_1^{\,2}+x_2^{\,2}+\ldots+x_p^{\,2}+\ldots+x_q^{\,2}]^{Z}/\{x_H^{\,2}\}^{Z}$$

$$=\begin{vmatrix} (1-\eta_{h1}^{\,2})^{K(Z\pm 0)} & 0 & 0\ldots 0\ldots 0 \\ 0 & (1-\eta_{h2}^{\,2})^{K(Z\pm 1)} & 0\ldots 0\ldots 0 \\ & \cdots\cdots & \\ 0 & 0\ldots & (1-\eta_{hp}^{\,2})^{K(Z\pm p)}\ldots 0 \\ 0 & 0\ldots\ 0\ \ldots & (1-\eta_{hq}^{\,2})^{K(Z\pm q)} \end{vmatrix}$$

$$=\begin{vmatrix} \{x_{h1}/x_H\}^{K(Z\pm 0)} \\ \{x_{h2}/x_H\}^{K(Z\pm 1)} \\ \cdots\cdots \\ \{x_{hp}/x_H\}^{K(Z\pm p)} \\ \{x_{hq}/x_H\}^{K(Z\pm q)} \end{vmatrix} \tag{7}$$

$$=(1-\eta_{h1}^{\,2})^{K(Z\pm 0)}+(1-\eta_{h2}^{\,2})^{K(Z\pm 1)}+\ldots$$
$$+(1-\eta_{hp}^{\,2})^{K(Z\pm p)}+\ldots+(1-\eta_{hq}^{\,2})^{K(Z\pm q)};$$
$$=\{0\sim 1\}^{Z} \tag{8}$$

Homology logarithmic expansion

$$(\eta_h^{\,2})^{Z}=(\eta_{h1}^{\,2})^{Z}+(\eta_{h2}^{\,2})^{Z}+\ldots+(_{hp}^{\,2})^{Z}+\ldots$$
$$+(\eta_{hq}^{\,2})^{Z}=\{1\}^{Z};\ \text{(Vector)} \tag{9.1}$$
$$(\eta_h)^{Z}=(\eta_{h1})^{Z}+(\eta_{h2})^{Z}+\ldots+(\eta_{hp})^{Z}+\ldots$$
$$+(\eta_{hq})^{Z}=\{1\}^{Z};\ \text{(Scalar)} \tag{9.2}$$

Coherent circular logarithmic secures quantized expansion of pointwise space (function) In particular, as for homology circle logarithmic in the pointwise space regularization combination, the value can be: zero, defect and non uniform, but the pointwise position cannot be vacant so as to ensure that the "non uniform pointwise status" automatically eliminates"error" and quantization stability.

3.2. Theorem Two: Isomorphic Logarithmic

In this paper, we define the isomorphic circular logarithm as the pointwise average space

divided by (multiplied by) the pointwise average space, and get the value of "0~1", that is, "a number divided by itself is not necessarily 1", which is called the second type of gauge invariance.)

To prove：

$$(1-\eta^{2})^{Z}\sim(\eta)^{Z}=\{\xi(x_0)_{H}/\xi(x_0)_{H}\}^{Z}$$

$$=(C_{i+0})\{x_{(X\pm X)}\}^{K(Z-0)}\bullet\{\xi(x_0)\}^{K(Z+0)}$$

$$+(C_{i+1})\{x_{(X\pm X)}\}^{K(Z-1)}\bullet\{\xi(x_0)\}^{K(Z+1)}+\cdots$$

$$+(C_{i+p})\{x_{(X\pm X)}\}^{K(Z-p)}\bullet\{\xi(x_0)\}^{K(Z+p)}+\cdots$$

$$+(C_{1+q})\{x_{(X\pm X)}\}^{K(Z-q)}\bullet\{\xi(x_0)\}^{K(Z+q)}$$

$$=\begin{vmatrix}(1-\eta^{2})^{K(Z\pm0)}0&\cdots&0&\cdots&0\\0&(1-\eta^{2})^{K(Z\pm1)}&\cdots&0&\cdots&0\\&\cdots\cdots&&&&\\0&0&(1-\eta^{2})^{K(Z\pm p)}&\cdots&0\\0&0&0&\cdots&(1-\eta^{2})^{K(Z\pm q)}\end{vmatrix}$$

$$=\begin{vmatrix}\{x_{01}/x_{0H}\}^{K(Z\pm0)}\\\{x_{02}/x_{0H}\}^{K(Z\pm1)}\\\cdots\cdots\\\{x_{0p}/x_{0H}\}^{K(Z\pm p)}\\\{x_{0q}/x_{0H}\}^{K(Z\pm q)}\end{vmatrix}$$

$$=\{(1-\eta^{2})^{K(Z\pm0)}+(1-\eta^{2})^{K(Z\pm1)}+\cdots$$

$$+(1-\eta^{2})^{K(Z\pm p)}+\cdots+(1-\eta^{2K(Z\pm q)}\}$$

$$=\{0\sim1/2\sim1\}^{Z};\qquad(10)$$

Derivation One: Conversion rules of the logarithmic order, order value sequence and sequence number of the circular logarithm contain isomorphic geometric grid change rules

$$(1-\eta^{2})^{Z}=(1-\eta^{2})^{K(S\pm N)}\cdot\{2\}^{K(\pm p\pm N)};\qquad(11)$$

Derivation Two:

The isomorphism of pointwise space has nothing to do with the position of coordinate origin, and it adapts to accurate grid computing.

There are two kinds of writing methods for the circular logarithm (super symmetric element matrix), horizontal and determinant. This paper focuses on the horizontal method (same below).

3.3. Theorem Three: ConvergentCircular Logarithm

By defining the pointwise space divided by the pointwise average space, we obtain the three kinds of properties of "convergence, flat and diffusion""positive, medium and negative" or "red, yellow, and blue" so as to ensure the consistency of the isomorphic circular logarithm, which is called the third type of gauge invariance.)

To prove：

$$(1-\eta^{2})^{Z}\sim(\eta)^{Z}=\xi(x)_{H}/\xi(x_0)$$

$$=(1-\eta^{2})^{+Z}+(1-\eta^{2})^{0Z}+(1-\eta^{2})^{-Z}$$

$$=\{0\sim1/2\sim1\}^{Z};\qquad(12)$$

Among it：

K=+1：$(1-\eta^{2})^{+Z}$ convergence：

$$\xi(x)_H / \xi(x_0) \leq \{1/2\}^Z;$$
K= 0：$(1-\eta^2)^{0Z}$ flat：
$$\xi(x)_H / \xi(x_0) = \{1/2\}^Z;$$
K=-1：$(1-\eta^2)^{-Z}$ diffusion：
$$\{1/2\}^Z \leq \xi(x)_H / \xi(x_0);$$

3.4. Theorem Four: Mapping of the Circular Logarithm

The quantized calculus of the Riemann function (±N) is converted to the circular logarithm, mapped (projected) as one dimensional, two dimensional, three dimensional and arbitrary dimensional circular logarithm:

(1).Vector field logarithm (three dimensional projection in spherical coordinates

$$(1-\eta_{[XYZ]}^2)^Z = (1-\eta_{[X]}^2)^Z \mathbf{i}+(1-\eta_{[Y]}^2)^Z \mathbf{j} +(1-\eta_{[Z]}^2)^Z \mathbf{K}; \quad (13.1)$$

(Or)：
$$(1-\eta_{[XYZ]}^2)^Z=(\eta_{h1[XYZ]}^2)^Z+(\eta_{h2[XYZ]}^2)^Z+\ldots +(\eta_{hp[XYZ]}^2)^Z+\ldots+(\eta_{hq[XYZ]}^2)^Z=1; \quad (13.2)$$

(2).Rotational field logarithm (four dimensional projection of spherical coordinates

$$\begin{aligned}(1-\eta_{[XX]}^2)^Z &= (1-\eta_{[Y]}^2)^Z-(1-\eta_{[Z]}^2)^Z=(1-\eta_{[YZ]}^2)^Z\\ (1-\eta_{[YY]}^2)^Z &= (1-\eta_{[Z]}^2)^Z-(1-\eta_{[X]}^2)^Z=(1-\eta_{[ZX]}^2)^Z\\ (1-\eta_{[ZZ]}^2)^Z &= (1-\eta_{[X]}^2)^Z-(1-\eta_{[Y]}^2)^Z=(1-\eta_{[XY]}^2)^Z \quad (14.1)\end{aligned}$$

Or：
$$(1-\eta_{[XYZ]}^2)^Z=(\eta_{[XX]}^2)^Z+(\eta_{[YY]}^2)^Z+\ldots+(\eta_{[ZZ]}^2)^Z = \{(1/3)^2+(2/3)^2+(2/3)^2+\} = 1; \quad (14.2)$$

Among them,

$\{x_h\}^Z$and $\{x_H\}^Z$represent the pointwise space; $\{x_0\}^Z$and $\{x_{0H}\}^Z$the average pointwise space. $\{x_{0H}\}$Z said the pointwise average space;

$(1-\eta_{[XYZ]}^2)^Z$ replaces the Hamiltonian operator.

3.5. Relativistic Structure -- Pointwise Gauge Field

Pointwise space has the adaptability to properties of symmetry and asymmetry, duality and non duality, getting:

$$\xi(x) = (1-\boldsymbol{\eta}^2)^Z \xi(x_0); \quad (15.1)$$
$$0 \leq (1-\boldsymbol{\eta}^2)^Z \leq 1; \quad (15.2)$$

3.6. Limit value (topological phase transition) critical point:

By definition 3 and definition 4

With：
$$\begin{aligned}(1-\eta^2)^Z &= \sum(1-\eta^2)^Z = \prod\{(1-\eta^2)^Z\\ (1-\eta^2)^{0Z} &= \sum\{(1-\eta^2)^{+Z}+(1-\eta^2)^{-Z}\}^Z = \{0,1\}^Z\\ (1-\eta^2)^{0Z} &= \prod\{(1-\eta^2)^{+Z}\cdot(1-\eta^2)^{-Z}\}^Z = \{0,1\}^Z \quad (16.1)\end{aligned}$$

Formula (16.1) simultaneous equations very easily gets stable (zero error):

$$(1-\eta^2)^Z = \{(0,1/2,1)^{(0,1/2,1)}\}^Z; \quad (16.2)$$

The formula (16.2) proves that with the Riemann's function, when the pointwise space of arbitrary asymmetry is transformed into that of relative symmetry, the non normal zero (1/2) is on a straight line $\{1/2\}^Z$, namely, "second RH problem".

3.7. On the interpretation of large O

The circular logarithm is a logarithm based on relatively variable circular function, whose four operations of the power function equation is contained in Napel logarithm $a^x=(1-\boldsymbol{\eta}^2)^Z$; Euler logarithm e^x， $(\ln x)=(1-\boldsymbol{\eta}^2)^Z$; Feynman integral $[\exp\{\pm(i/h)S_t(v)\}]=(1-\boldsymbol{\eta}^2)^Z$; modulus

square $|\psi(t,x)|^2=(1-\eta^2)^Z$; trigonometric function $(\cos x+\sin x)^Z=(1-\eta^2)^Z$, etc. They all have a fixed logarithm of the bottom function. The bottom function can not be synchronized with the power function and can not eliminate the error $O(x^{(1/2)+\text{infinitesimal}})$ [1] p234.
However, among converted circular logarithms based on relatively variable circular functions, error is automatically eliminated in the gauge invariance, with synchronization between bottom function and power function, and without the problem of $O(x^{(1/2)+\text{infinitesimal}})$. It is called pointwise path integral with reasonable treatment of the nonuniformity of prime number theorem, accurate acquisition of the number of primes, critical point or limit value, and the solution to the calculus equation of arbitrary high dimensional order.

4. Pointwise space and Riemann Function Equation

4.1 Pointwise Riemann function equation

According to the Brouwer center theorem (fixed point theorems)[2] p324, the pointwise Riemann function accordingly produces a balanced central pointwise space and becomes a pointwise quantized Riemann function equation:
Among them, the boundary conditions are:
(1). Discrete state:
(Or parallel continuous summation combinations)

$$\begin{aligned}&\xi(x)_{B1}=\{D\}\\&=\sum(1/C_{1+h})\{D_A+D_B+D_C+D_D+\ldots\};\\&=\{D_0\}^Z;\end{aligned}\tag{17.1}$$

(2). Entangled state:
(or serial combination and product combination)

$$\begin{aligned}&\xi(x)_{B2}=\{D\}\\&=\prod(1/C_{1+h})\{D_A\cdot D_B\cdot D_C\cdot D_D\ldots\}\\&=(\sqrt[KS]{D})^Z;\end{aligned}\tag{17.2}$$

Combination of Riemann's equations:

$$\begin{aligned}&\{\xi(x)_A\pm\xi(x)_B\}^Z\\&=F\{X\pm(\sqrt[KS]{D})\}^Z\\&=Ax^{K(Z\pm0)}+Bx^{K(Z\pm1)}+\ldots+Cx^{K(Z\pm p)}+\ldots+Dx^{K(Z\pm q)}\\&\pm(\sqrt[KS]{D})^Z\\&=Ax^{K(Z\pm0)}+Bx^{K(Z-1)}D^{K(Z+1)}+\ldots\\&+Px^{K(Z-p)}D^{K(Z+p)}+\ldots+Qx^{K(Z-q)}D^{K(Z+q)}\\&\pm(\sqrt[KS]{D})^Z;\\&=\{(1-\eta^2)^{K(Z\pm0)}+(1-\eta^2)^{K(Z\pm1)}+\ldots\\&+(1-\eta^2)^{K(Z\pm p)}+\ldots+(1-\eta^2)^{K(Z\pm q)}\}\\&=(1-\eta^2)^Z\{0,2\}^Z\{D_0\}^Z;\end{aligned}\tag{18.1}$$

$$\begin{aligned}&(1-\eta^2)^Z=(\sqrt[KS]{D})^Z/\{D_0\}^Z\\&=\{\xi(x)_A/\xi(x)_B\}^Z\\&=\{\xi(x)_A-\xi(x)_B\}^Z/\{\xi(x)_A+\xi(x)_B\}^Z\\&=\{\xi(x)_A-\xi(x_0)\}^Z/\xi(x_0)\}^Z\\&=\{\xi(x)_B-\xi(x_0)\}^Z/\xi(x_0)\}^Z\\&=\{0\sim1/2\sim1\}^Z;\end{aligned}\tag{18.2}$$

In the equation, $\{\xi(x)_A+\xi(x)_B\}^Z$ represents “grand balance or sum”
$\{\xi(x)_A-\xi(x)_B\}^Z$ represents “zero balance or spin”
In terms of logarithm:

$$\{\xi(x)_A+\xi(x)_B\}^Z=(1-\eta^2)^Z\{2\}^Z\xi(x_0);\tag{19}$$

Formula (19) proves that sum of any two sufficiently large prime numbers "$\{2\}^{Z}\xi(x_0)$"is an "even number".

Its practicalsignificance:The theory of relativity is structured to smoothly solve the combination of multi-element "asymmetry and uncertainty", and become the combination and decomposition of "relative symmetry and relative certainty".

4.2.Solution of High Dimensional Differential Equation

Formula (18.1) solution is to compute the corresponding prime (element) value of a known numerical value ($^{KS}\sqrt{D}$) (interactive entanglement state if encountered in engineering and discrete state if encountered in quantum computing.

4.3. Discriminant:

Judge the possibility of formula (18.1) solution

(1). Calculate the average state space $\{D_0\}$ (Select arbitrary P combination.

$$\{D_0\}^{Z}= [^{K(S\pm p)}\sqrt{(P/C_{1+p})}]^{[Z\pm p]} \quad ; \qquad (20)$$

Average state statistics are different from quantum state Einstein statistics

(2). Satisfying) $0\leq(1-\eta^2)=(^{KS}\sqrt{D} / D_0]^{[Z\pm p]}\leq 1$;

(3). Coefficients (A, B, C, D, etc.) are adjusted to be $\{D_0\}^{[Z\pm p]}$

(4). To meet the sum of the regularization coefficient: $\sum C_{1+p} = \{2\}^{[Z\pm p]}$; among them the element combination value can be incomplete, but the coefficient can not be vacant.

4.4. Solution

Choose the simplest second terms of the calculus equation (18.1) K[Z-1] to obtain the isomorphism.

$$\{D_0\}^{[Z-1]}=\{D_0\}^{[Z-p]}=\{D_0\}^{Z}:$$

With: $$\{D_0\}^{[Z-1]}=(B/C_{1+1}),$$

$$(1-\eta^2) = (^{KS}\sqrt{D})\}^{[Z-1]} / \{D_0\}^{[Z-1]};$$

$$(1-\eta_h^2)^{[Z-1]} =\sum(1-\eta_h^2)^{[Z-1]} =1; \qquad (21)$$

General formula: $\{D_0\}^{[Z-p]}=(P/C_{1+p})$,

$$(1-\eta^2)^{[Z-p]} = (^{KS}\sqrt{D})\}^{[Z-p]} / \{D_0\}^{[Z-p]}$$

$$(1-\eta_h^2)^{[Z-p]} =\sum(1-\eta_h^2)^{[Z-p]} =1; \qquad (22)$$

$$\{X_{h1}\}^{[Z-1]}=[(1-\eta^2)/ (1-\eta_{h1}^2)]^{[Z-1]} (B/C_{1+1}); \qquad (23)$$

$$\{X_{hp}\}^{[Z-p]}= [(1-\eta^2)/ (1-\eta_{hp}^2)]^{[Z-p]}(p/C_{1+1}); \qquad (24)$$

Formulas (23) and (24) in the analysis of the law of the composition of elements or prime number distribution theorem (PNT), it is easy to get: before a known value, determine several prime numbers, that is, "RH's first problem."

5. Engineering Application

5.1. "Explanation of the Yang-Mills Gauge Field"

When pointwise equations are comprised of non-uniform even multiplication or even addition functions, they become extended quantum state Yang-Mills gauge field in physics, namely; the unity of pointwise non-uniform Dirac dynamic equations (calculus equations), Maxwell electromagnetic equations (even power functions) and gravitational equations (odd power functions) and the gauge invariance, becomes the abstract computing method for pointwise state with no specific quality content.

$$\{L[\psi(x_A),A_\mu(x_A)]\pm L[\psi(x_B),A_\mu(x_B)]\}^{Z}$$

$$=(1-\eta^2)^{Z}\{0,2\}^{Z}\{L[\psi(x_0),A_\mu(x_0)]\}^{Z}; \qquad (25)$$

$$(1-\eta^2)^{Z} =L[\psi(x_A),A_\mu(x_A)]/L[\psi(x_B),A_\mu(x_B)] \qquad (26)$$

In the formula: $L[\psi(x_0),A_\mu(x_0)]^Z$ represents the set aggregation of combinations of various levels of pointwise mean Riemann functions.

5.2. Explanation of Energy Asymmetry

Under the condition of mass conservation, the vacuum excitation of the universe is caused by the asymmetry of the super high energy, which experiences the process of topological phase transition = vacuum excitation = Higgs particle mutation.

With: $\{(1-\eta^2)^{+Z}\rightarrow(1-\eta^2)^{0Z}\rightarrow(1-\eta^2)^{-Z}\}\cdot\{MC^2\}$；。

Get: $\{(1-\eta^2)^{+Z}\leq\leq(1-\eta^2)^{-Z}\}$；

Assume: 11 dimensional space of the universal energy particle(8 dimensional electromagnetic quantum +3 dimensional gravitational quantum),comprise two parallel equations of energy:

$$\{x\}^{[Z-11]}=\{x\}^{[Z-6]}+\{x\}^{[Z-5]}$$

$\{x\}^{[Z-6]}$equation：6prime numbers：1,3,5,7,11,13as of entanglement state；

$\{x\}^{[Z-5]}$equation：5 natural numbers：1,2,3,4,5 as of discrete state

Get: (Calculationprocedure omitted)

Mass energy ratio： $(1-\eta^2)^{+[Z-11]}$ ： $(1-\eta^2)^{-[Z-6]}$

（4.758845% ： 95.241155%）；

Energy ratio： $(1-\eta^2)^{+[Z-6]}$ ： $(1-\eta^2)^{-[Z-6]}$

(1： 40.027004)；

The above calculation results are in surprising agreement with astronomical observations and high energy particle collision test data.

6. Conclusion and Prospect

Pointwise space is the fusion of Riemann function and the theory of relativity, forming a relativistic structure (circular logarithm and super symmetric unit matrix) in that it is inclusive of traditional uniform elements and non-uniformity and is a macroscopic and microscopic unification. It has got rid of the interference of specific element contents, acquired the gauge invariance of linear (simple combination)and nonlinear combination (complex combination) and has become a new and broader calculation system. It enjoys the superiority of simple, self-consistent, accurate andzero-error interactions. It is widely applicable to mathematics, physics, astronomy, biology, chemistry, mechanics, statistics, big data and other areas.

Existing problem: In the Riemann function and engineering application, there is an interaction mechanism, which brings the problems of force’s coupling constants and topological phase transformation points. They are waiting for scientists to further determine and research. However, no matter whether or not the coupling constantsof force exist, it does not affect the circular logarithm topology or statistical calculation process. The results are adjusted only at the end of calculation.

Initial structure of the theory of relativity may inevitably be flawed. Sincerely welcome are criticism and suggestions for improvement.More experts and scholars are expected to communicate, promote its application for joint innovation and development.

Thanks to John Derbyshire, who, through the book “Love of Prime Numbers”, has put together the study of the Riemann function and prime numbers, and made important contributions to provision of guidance to coming generations to crack the Riemann conjecture.

Thanks to the 2016 China Conference on Computational Mechanics and International Synposium on Computational Mechanics for Chinesein Hangzhou (CCCM-ISCM 2016), which has recognized the *Study and Application of the Supersymmetric Unit Matrix (Circular Logarithm)* as a novel "abstract analysis".

Thanks to the Senior Science & Technology Workers Association of Quzhou, Zhejiangfor its long-term support and encouragement. It has set up a blog: Exploring the Free Sky (LKX0570), (since 2009 it has published more than 640 public blog articles) and reflects the difficult process of establishment of circular logarithm-theory of relativity structure.

Thanks to the US Matter Regularity academic journal, which published seven articles in 2014-2016, gradually establishing a system. (End)

References

[1] PRIME OBSESSION——BERNHAD RIEMANN AND THE GREATEST UNSOLVED PROLEM IN MATHEMATICS by John Derbyshire; ranslator: Chen Weifeng; Shanghai Science & Technology Education Publishing House, Aug.2009, 2nd printing

[2] Mathematical Thought From Ancient to Modern Times by Morris Kline, Copyright C 1972 Vol. 3, p353, Shanghai Science & Technology Education Publishing House, Aug. 2014, 2nd printing

[3] "Criticism of Berkeley's Paradox and Point Continuity and Related Issues" by Xu Lizhi, Journal of Higher Mathematical Research, 2013 No. 5 Issue, P33-35

[4] Relativity: The Special and General Theory by Albert Einstein, translated by Yang Runyin, Shanghai Science & Technology Education Publishing House, Apr.1979, 3rd printing

[5] Brief Introduction to Application of the Theory of Relativity in Mathematics (Chinese version) by Wang Yiping, [US] Matter Regularity, 2014/1, p38-60

[6] Big Data and Circular Logarithm Algorithm (English version) by Wang Yiping, [US] Matter Regularity, 2016/4, p1-10

An Improved Algorithm for Clustering

***Tsan-Jung He[1], Zhao-Yu Wang[1], †Shie-Jue Lee[1], and Shing-Tai Pan[2]**

[1]Department of Electrical Engineering National Sun Yat-Sen University, Kaohsiung, Taiwan
[2]Department of Electrical Engineering National University of Kaohsiung, Kaohsiung, Taiwan

*Presenting author: zlhe@water.ee.nsysu.edu.tw
†Corresponding author: leesj@mail.ee.nsysu.edu.tw

Abstract

In this paper, we propose a new clustering algorithm which is an improvement to a self-constructing clustering (SCC) method. The SCC processes all the data points incrementally. If the input data point is similar enough to an existing cluster, the point is added to the cluster. Otherwise, the data point forms a new cluster of its own. The method ends up with a set of clusters after it runs through the whole dataset once. However, once a data point is assigned to a cluster, there is no way to change the assignment afterwards. This may cause assignment errors and the efficacy of the clustering is reduced. In this paper, we adopt an iterative approach to overcome this shortcoming. A data point can be re-assigned to another cluster. Adding points into and removing points from a cluster are allowed to be done iteratively in the clustering process. The clustering work stops when all the assignments are stable, i.e., no assignment would be changed. The proposed approach can result in better clusters, and experimental results show that it performs better than SCC for real world datasets.

Keywords: data mining, clustering, self-constructing clustering, similarity, classification.

I. Introduction

In the field of artificial intelligence, clustering techniques play a very important role [3][9]. Clustering is an unsupervised classification technology, with a purpose of forming meaningful clusters for the objects under consideration. Usually, similar objects are grouped in the same cluster, and different objects are grouped in different clusters. The clustering concept is widely applied in a variety of different areas, such as bio-engineering [6][14], environmental monitoring [8], economic applications [12], and so on.

In the electronic text applications [4], the dimensionality of the data can be reduced to improve the efficiency of the operation through the clustering technology. In the recommendation applications of e-commerce [7], clustering is used to reduce the size of the information matrix to enhance the efficiency of the operation. In the application of regression, the reduction of information dimension is used. In power system, clustering is used to predict the electrical trend in the future [1]. In other areas, such as stock market and data regression [11][13], the clustering technology is an important and indispensable core key. Therefore, developing a better clustering technology is a very critical issue.

Ouyang et al. [5][10] proposed a clustering method, self-constructing clustering (SCC), which has been applied in various applications. It considers all the data points one by one. For an input point, its similarity to each existing cluster is calculated. If the point is similar enough to an existing cluster, the point is added in the cluster. On the other hand, if the point is not similar enough to any of the existing clusters, the point forms a new cluster. The algorithm proceeds until all the points have been processed once. SCC offers several advantages. First, since the algorithm runs through the data points once, it is fast. Second, it considers the

variation of the data under consideration. Third, the number of clusters is not to be specified in advance.

However, SCC has one disadvantage. Once a data point is assigned to a cluster, there is no way to change the assignment afterwards. This may cause assignment errors and the efficacy of the clustering is reduced. In this paper, we adopt an iterative approach to overcome this shortcoming of SCC. A data point can be re-assigned to another cluster. Adding points into and removing points from a cluster are allowed to be done iteratively in the clustering process. The clustering work stops with a desired number of clusters when all the assignments are stable, i.e., no assignment would be changed. The proposed approach can result in better clusters, and experimental results show that it performs better than SCC for real world datasets.

The rest of this paper is organized as follows. SCC is briefly reviewed in Section II. Our proposed improvement presented in Section III. Experimental results are shown in Section IV. Section V gives a conclusion.

II. Related Work

SCC [5][10] is a progressive clustering method using the Gaussian function as the membership function of the resulting clusters. For each cluster, its center and distribution are described by the mean and standard deviation, respectively, of the contained data points.

Data points are considered one by one sequentially. When the first data point comes in, the first cluster is created for it. Then, for each of the rest data points, SCC calculates the similarity between the input data point and each existing cluster. If the input data point is similar enough to an existing cluster, the data point is added into this cluster. Otherwise, a new cluster is created for the input data point. Given the input data point x, the similarity to cluster G for x is calculated as follows:

$$\mu_G = \prod_{i=1}^{p} exp[-(\frac{x_i - m_i}{\sigma_i})^2] \tag{1}$$

where p is the number of dimensions of the input data, and m_i and σ_i are the center and standard deviation of the i th dimension of cluster G, defined respectively by

$$m_i = \frac{\sum_{j=1}^{|G|} y_{ji}}{|G|}, \tag{2}$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{|G|} (y_{ji} - m_{ji})^2}{|G| - 1}}. \tag{3}$$

Note that $|G|$ is the total number of data points contained in cluster G, y_j, $j=1,\ldots,|G|$, are the data points contained in cluster G, and y_{ji} represents the ith dimension of y_j. When the similarity between the input data point x and the existing clusters is greater than a default threshold, the input is added into the cluster with the largest similarity. Let the cluster be G_t and its size be S_t. The mean and deviation of cluster G_t are then updated.

After all the data points are considered, SCC stops and a set of clusters are obtained. The algorithm can be described below.

Table 1. SCC

Algorithm 1 SCC	
1:	**Input:** dataset D
2:	**for** each data point x **do**
3:	compute similarity to each existing cluster
4:	**if** some similarity bigger enough **then**
5:	add x to the cluster with largest similarity
6:	**else**
7:	form a new cluster
8:	**end if**
9:	**end for**
10:	**Output:** all the clusters obtained

SCC offers several advantages. First, since the algorithm runs through the data points once, it is fast. Second, it considers the variation of the data under consideration. Third, the number of clusters is not to be specified in advance.

III. Proposed Method

However, SCC has one disadvantage. Once a data point is assigned to a cluster, there is no way to change the assignment afterwards. This may cause assignment errors and the efficacy of the clustering is reduced. We adopt an iterative approach to overcome this shortcoming of SCC. A data point can be re-assigned to another cluster. Adding points into and removing points from a cluster are allowed to be done iteratively in the clustering process. The clustering work stops with a desired number of clusters when all the assignments are stable, i.e., no assignment would be changed.

For convenience, our proposed approach is called New-SCC. New-SCC consists of several rounds of iteration. In the first round, the algorithm of the original SCC is performed.
Then we perform the second round and beyond, each round considering all the data points sequentially. In each succeeding round, for an input data point x, we first remove it from the cluster G_t, which x belongs to. Then we calculate the similarity between x and each existing cluster by Eq. (1). If the max similarity occurs with cluster G_a and it is higher than a specified threshold, x is added into G_a. However, if the max similarity is not higher than the specified threshold, a new cluster is created for x. A round of iteration ends when all the data points are gone through once. If one of the assignments has been changed for the data points in the current round, the next round of iteration begins. Otherwise, the assignments are stable and

New-SCC stops with a desired number of clusters. Let's have an example here to illustrate how New-SCC works. Suppose we have the following 12 data points:

$$x_1 = \langle 0.30,0.60 \rangle;$$
$$x_2 = \langle 0.70,0.35 \rangle;$$
$$x_3 = \langle 0.50,0.52 \rangle;$$
$$x_4 = \langle 0.78,0.20 \rangle;$$
$$x_5 = \langle 0.62,0.25 \rangle;$$
$$x_6 = \langle 0.40,0.65 \rangle;$$
$$x_7 = \langle 0.35,0.38 \rangle;$$
$$x_8 = \langle 0.28,0.48 \rangle;$$
$$x_9 = \langle 0.19,0.89 \rangle;$$
$$x_{10} = \langle 0.24,0.81 \rangle;$$
$$x_{11} = \langle 0.29,0.89 \rangle;$$
$$x_{12} = \langle 0.24,0.89 \rangle.$$

After performing SCC in the first round, we have 6 clusters: G_1, G_2, G_3, G_4, G_5, and G_6, with

$$m_1 = \langle 0.35,0.625 \rangle, \sigma_1 = \langle 0.2707,0.2354 \rangle;$$
$$m_2 = \langle 0.66,0.3 \rangle, \sigma_2 = \langle 0.2566,0.2707 \rangle;$$
$$m_3 = \langle 0.5,0.52 \rangle, \sigma_3 = \langle 0.2,0.2 \rangle;$$
$$m_4 = \langle 0.78,0.2 \rangle, \sigma_4 = \langle 0.2,0.2 \rangle;$$
$$m_5 = \langle 0.315,0.43 \rangle, \sigma_5 = \langle 0.2495,0.2407 \rangle;$$
$$m_6 = \langle 0.24,0.81 \rangle, \sigma_6 = \langle 0.2408,0.24 \rangle.$$

Note that G_1 contains data points 1 and 6, G_2 contains data points 2 and 5, G_3 contains data point 3, G_4 contains data point 4, G_5 contains data points 7 and 8, and G_6 contains data points 9, 10, 11, and 12. After the second round, we have 4 clusters: G_1, G_2, G_3, and G_4, with

$$m_1 = \langle 0.37,0.5625 \rangle, \sigma_1 = \langle 0.3013,0.2768 \rangle;$$
$$m_2 = \langle 0.7,0.2667 \rangle, \sigma_2 = \langle 0.28,0.2764 \rangle;$$
$$m_3 = \langle 0.35,0.38 \rangle, \sigma_3 = \langle 0.2,0.2 \rangle;$$
$$m_4 = \langle 0.24,0.87 \rangle, \sigma_4 = \langle 0.2408,0.24 \rangle.$$

Note that G_1 contains data points 1 3, 6, and 8, G_2 contains data points 2, 4, and 5, G_3 contains data point 7, and G_4 contains data points 9, 10, 11, and 12. After the third round, we have 3 clusters: G_1, G_2, and G_3, with

$$m_1 = \langle 0.366,0.526 \rangle, \sigma_1 = \langle 0.2882,0.3053 \rangle;$$
$$m_2 = \langle 0.7,0.2667 \rangle, \sigma_2 = \langle 0.28,0.2764 \rangle;$$
$$m_3 = \langle 0.24,0.87 \rangle, \sigma_3 = \langle 0.2408,0.24 \rangle.$$

Note that G_1 contains data points 1 3, 6, 7, and 8, G_2 contains data points 2, 4, and 5, and G_3 contains data points 9, 10, 11, and 12. After the fourth round, no assignment has been changed. New-SCC stops with three clusters G_1, G_2, and G_3, with

$$m_1 = \langle 0.366, 0.526 \rangle, \sigma_1 = \langle 0.2882, 0.3053 \rangle;$$
$$m_2 = \langle 0.7, 0.2667 \rangle, \sigma_2 = \langle 0.28, 0.2764 \rangle;$$
$$m_3 = \langle 0.24, 0.87 \rangle, \sigma_3 = \langle 0.2408, 0.24 \rangle.$$

IV. Experimental Results

In this section, some experimental results are presented. Comparisons between SCC and New-SCC have been done. Several real world datasets taken from the UCI Machine Learning Repository are used in experiments [2]. The characteristics of these datasets are listed in Table 2.

Table 2. Accuracy results with different window sizes

Dataset	# instances	# features	# classes
Breast	569	30	2
Ecoli	336	7	8
Glass	214	9	6
Heart	270	13	2
Iris	150	4	3
Libras	360	90	15
Wine	178	13	3
Yeast	1484	8	10

In this table, column 1 indicates the name of the dataset, and the remaining columns indicate the number of instances, the number of features, and the number of classes, respectively, associated with each dataset. For example, the Breast dataset contains 569 data instances, each instance has 30 features (or dimensions) and belongs to one of 2 classes. Note that these datasets are single-labeled, i.e., an instance belongs to only one class.

To evaluate the effectiveness of SCC and New-SCC, the following performance measures are adopted:

1. F-score. It is defined as

$$F-score = \sum_{j=1}^{k} \frac{n_j}{n} * max_{1 \le l \le L} \{ \frac{2 * \frac{n_{jl}}{n_j} * \frac{n_{jl}}{n_l}}{\frac{n_{jl}}{n_j} + \frac{n_{jl}}{n_l}} \} \quad (8)$$

where **k** is the number of classes, **L** is the number of clusters, **n** is the size of the entire data set, n_{jl} is the number of data instances belonging to class j in cluster l, n_l is the size of cluster l, and n_j is the size of class j.

2. RI. It is defined as

$$RI = \frac{a+b}{n(n-1)/2} \tag{9}$$

where **a** is the number of pairs of data objects having different class labels and belonging to different clusters, **b** is the number of pairs of data objects having the same cluster labels and belonging to the same clusters, and **n** is the size of the entire data set.

3. NMI. It is defined as

$$NMI = \frac{\sum_{j=1}^{k}\sum_{l=1}^{L} n_{jl} log(\frac{n * n_{jl}}{n_j * n_l})}{\sqrt{(\sum_{j=1}^{k} n_j log \frac{n_j}{n}) * (\sum_{l=1}^{L} n_l log \frac{n_l}{n})}} \tag{10}$$

where **k** is the number of classes, **L** is the number of clusters, **n** is the size of the entire data set, n_{jl} is the number of data instances belonging to class j in cluster l, n_l is the size of cluster l, and n_j is the size of class j.

All these measures have a common property: a higher measure indicates a better clustering performance.

Table 3 shows performance comparisons between SCC and New-SCC. In this table, the values for the three measures, F-score, RI, and NMI, are listed, and the CPU time elapsed in clustering is also listed in the last column. For the sake of fairness, we compare SCC and New-SCC under the condition of producing the same number of clusters for each dataset. Evidently, New-SCC performs better than SCC in F-score, RI, and NMI for most of the datasets. For example, for the Breast dataset, SCC has F-score = 0.7691, RI=0.6742, and NMI = 0.3349, while New-SCC has F-score = 0.9260, RI=0.8630, and NMI=0.6049. We can see that New-SCC provides a very significant improvement to SCC in this case. However, not all the datasets offer so much difference.

For some datasets, New-SCC is even inferior to SCC in some measure or another. For example, for the Ecoli dataset, SCC has F-score = 0.7333, RI = 0.8229, and NMI = 0.6261, while New-SCC has F-score = 0.7212, RI = 0.8339, and NMI=0.6447. NEW-SCC is better in RI and NMI, but is worse in Fscore. Note that, in general, New-SCC takes more CPU time in clustering than SCC. This is reasonable, since SCC performs one round of iteration while New-SCC performs two or more rounds of iteration. For example, for the Breast dataset, SCC takes 0.02 seconds while New-SCC takes 0.25 seconds for clustering.

Table 3. Accuracy results with different dimension sizes

Dataset		F-score	RI	NMI	# clusters	CPU time
Breast	SCC	0.7691	0.6742	0.3349	2	0.02
	New-SCC	0.9260	0.8630	0.6049	2	0.25
Ecoli	SCC	0.7333	0.8229	0.6261	8	0.015
	New-SCC	0.7212	0.8339	0.6447	8	0.29
Glass	SCC	0.5209	0.5506	0.3440	6	0.01
	New-SCC	0.5593	0.6472	0.4519	6	0.13
Heart	SCC	0.6506	0.5273	0.0938	2	0.01
	New-SCC	0.6470	0.5342	0.1075	2	0.056
Iris	SCC	0.8639	0.8589	0.7351	3	0.006
	New-SCC	0.8901	0.8781	0.7472	3	0.04
Libras	SCC	0.3871	0.7619	0.4434	15	0.023
	New-SCC	0.4767	0.8787	0.5603	15	0.37
Wine	SCC	0.7201	0.7073	0.5525	3	0.007
	New-SCC	0.7634	0.7546	0.6024	3	0.075
Yeast	SCC	0.4184	0.5398	0.1985	10	0.07
	New-SCC	0.4300	0.7288	0.2653	10	5.6

V. Conclusion

We have presented a new clustering algorithm, New-SCC, which is an improvement to the SCC clustering algorithm. SCC considers all the data points one by one sequentially. Clusters are created incrementally and automatically. If the input data point is similar enough to an existing cluster, the point is added to the cluster. Otherwise, the data point forms a new cluster of its own. SCC ends up with a set of clusters after it runs through the whole dataset once. However, once a data point is assigned to a cluster, there is no way to change the assignment afterwards. This may cause assignment errors and the efficacy of the clustering is reduced. New-SCC is aimed to overcome this shortcoming. A data point can be re-assigned to another cluster. Adding points into and removing points from a cluster are allowed to be done iteratively in the clustering process. New-SCC stops when all the assignments are stable, i.e., no assignment would be changed. As a result, New-SCC can produce better clusters. Experimental results have shown that NEW-SCC performs better than SCC for real world datasets.

References

[1] Alvarez, F. M., Troncoso, A., Riquelme, J. C. and Ruiz, J. S. A. (2011) Energy time series forecasting based on pattern sequence similarity. *IEEE Transactions on Knowledge and Data Engineering*, **23**, 1230–1243.

[2] Asuncion, A. and Newman, D. (2007) UCI machine learning repository.

[3] Haykin, S. (1999) *Neural Networks -- A Comprehensive Foundation*. Prentice-Hall Upper Saddle River, NJ, USA.

[4] Lee, S. J. and Jiang, J. Y. (2014) Multilabel text categorization based on fuzzy relevance clustering. *IEEE Transactions on Fuzzy Systems*, **22**, 1457–1471.

[5] Lee, S. J., Ouyang, C. S. and Du, S.H. (2003) A neuro-fuzzy approach for segmentation of human objects in image sequences. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **33**, 420–437.

[6] Li, W., Jaroszewski, L. and Godzik, A. (2001) Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics*, **17**, 282–283.

[7] Liao, C. L. and Lee, S. J. (2016) A clustering based approach to improving the efficiency of collaborative filtering recommendation. *Electronic Commerce Research and Applications*, **18**, 1–9, 2016.
[8] Wang, M., Yu, Y. and Lin, W. (2009) Adaptive neural-based fuzzy inference system approach applied to steering control. *In Proceedings of International Symposium on Neural Networks*, Springer, 1189–1196.
[9] Olson, D. L. and Shi, Y. (2007) *Introduction to business data mining*, **10**, McGraw-Hill/Irwin Englewood Cliffs.
[10] Ouyang, C. S., Lee, W. J. and Lee, S. J. (2005) A TSK-type neurofuzzy network approach to system modeling problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **35**, 751–767.
[11] Wang, Z. Y. and Lee, S. J. (2014) A neuro-fuzzy based method for TAIEX forecasting. *In Proceedings of International Conference on Machine Learning and Cybernetics (ICMLC)*, **2**, 579–584.
[12] Wei, C. C., Chen T. T. and Lee, S. J. (2013) K-NN based neuro-fuzzy system for time series prediction. *In Proceedings of 14th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, 569–574.
[13] Xu, R. F. and Lee, S. J. (2015) Dimensionality reduction by feature clustering for regression problems. *Information Sciences*, **299**, 42–57.
[14] Xu, Y., Olman, V. and Xu, D. (2002) Clustering gene expression data using a graph-theoretic approach: an application of minimum spanning trees. *Bioinformatics*, **18**, 536–545.

Numerical simulation for compression failure of bi-material interface by using cohesive zone model

Li-Ya Liu*, Qing-Sheng Yang†

Department of Engineering Mechanics, Beijing University of Technology

Beijing 100124, PR China

*Presenting author: liuliya@emails.bjut.edu.cn

†Corresponding author: qsyang@bjut.edu.cn

Abstract

The interface plays an important role in the performance of the composite. The cohesive zone model is an effective method to simulate the initiation and propagation of the interfacial crack by properly selecting the parameters. In this paper, a bilinear cohesive zone model is used to simulate the fracture behavior of the interface of bi-material under compression loads. The three-dimensional cohesive element with eight nodes is employed. The load-displacement curves and the interfacial damage modes are obtained for different combinations of materials. The effect of the interfacial parameters on the failure process of the bi-material interface is discussed. The results demonstrate that the compression failure of the bi-material interface severely depends on the properties and geometrics of the two materials. It is also found that the ultimate failure load will increase for the enhanced cohesion strength of the interface.

Keywords: Bi-material interface, Compression failure, Cohesive zone model

Introduction

The metal laminates have broad promising applications in automobile, aerospace, national defense and so on. Their excellent mechanical properties are closely dependent to the interface condition. The basic failure mode of laminates by the low inter-laminar bonding strength is the separation between layers, namely delamination. The failure characteristics is different from those of the general materials [1]. However, the interface of composite material, which is a system formed in the environment of thermology, chemistry and mechanics, has a very complicated structure. The study of bi-material interface plays a vital role in a fundamental understanding of the composite property [2]. In the early 1990s, the corresponding theory and numerical methods for the interfacial performance have been developed. To date, its research is still in the stage of continuous enrichment and improvement.

In the past few years, several cohesive zone models (CZMs) have been proposed to characterize interaction between bi-material components generated by the cohesive forces [3]. Nevertheless, CZM is a phenomenological model rather than an exact physical representation of the failure process in the fracture zone, which belongs to the energy viewpoint [4]. Actually, CZM is an effective method to simulate the failure of bi-material interface by varying the values of cohesive strength and fracture energy of the interface [5]. Atkinson [6] [7] believed that the properties on the both sides of interface of different materials are not abrupt, but gradual, and defined that the interface of different materials is an interface region with a certain thickness, on which cracks are likely to generate and propagate. Chandra et al. [8] used two different forms of CZMs (exponential and bilinear) to evaluate the response of

interfaces in titanium matrix composites reinforced by silicon carbide (SCS-6) fibers, and proposed that in addition to the primary parameters, the form of the traction–separation equations for CZMs plays a very critical role in determining the macroscopic mechanical response of the composite system. Hereafter, Kent [9] employed the trapezoidal CZM to analyze mixed mode failure behavior of a thin adhesive layer. Fiedler et al. [10] simulated the failure modes of unidirectional fiber reinforced composites under lateral loading by numerical simulation method. Zhang et al. [11] simulated the progressive debonding between circular inclusion in an infinite plate and matrix interface by using the surface-based cohesive method based on CZM in ABAQUS. However, there are limited studies on the compression failure behavior of bi-material interface now, especially the numerical simulation.

In this paper, the failure behavior of the interface of bi-material under compression was simulated by applying a cohesive finite element method. Then the influence of the interface parameters and geometry on the ultimate failure load of the bi-material laminates was investigated.

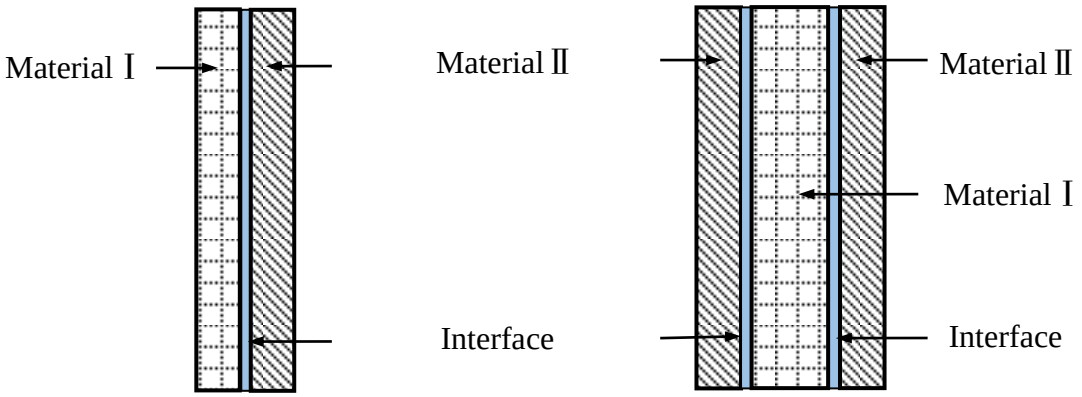
Cohesive finite element model of bi-metal laminates

CZM has been widely used in the simulation of interfacial failure behavior of laminates. In previous research, Hashagen et al. [12] [13] used an interface element based on plasticity theory to simulate delamination failure behavior of the laminates. Paul et al. [14] combined the cohesive interface element based on CZM with the fatigue damage rule to simulate the composite cantilever beam test.

There are different combinations of materials in this work. As shown in Fig. 1a, each of the two plates has a length of 30cm, a width of 10cm, and a thickness of 1cm. Material Ⅰ is hard aluminum alloy and material Ⅱ is rolled aluminum. Young's modulus of the hard aluminum alloy and the rolled aluminum were taken as 70GPa and 68GPa, respectively. The thickness of the hard aluminum alloy plate was gradually increased, and three models were established. The thickness ratios of the hard aluminum alloy plate and the rolled aluminum plate are 1:1, 1.1:1 and 1.2:1.

In another case, as illustrated in Fig.1b. The thickness of both sides of the laminate is 1cm respectively, and the thickness in the center is 2cm. The material of both sides of the laminate is material Ⅰ, the other is material Ⅱ. Other parameters are the same as those described in the previous paragraph. Similarly, the thickness of the hard aluminum alloy plate was gradually increased, and three models were established. The thickness ratios of the hard aluminum alloy plate and the rolled aluminum plate are 1:2, 1.1:1 and 1.3:1.

The static analysis of the interface of the metal laminates was carried out. In order to study the mechanical properties of bi-material interface in compression, the bottom of the plate was fixed, a normal strain was exerted on the top of the model, while the part of left and right sides were clamped along the thickness directions. In those cases, a three-dimensional cohesive element with eight nodes was employed for the interface and the details of the CZM will be given in Section 3.



(a) double-ply laminate **(b) three-ply laminate**

Figure 1. Model of the bi-metal laminates

CZM of bi-material laminates interface

The fracture behavior of bi-material laminates was characterized by using a bilinear cohesive law, in which the fracture energy G_c and the cohesive strength T_c are two pivotal parameters [15]. As shown in Fig.2, the cohesive constitutive model of ABAQUS is the bilinear model, where δ_0 denotes the interface damage and the maximum displacement δ_f denotes the interface failure. In order to observe the influence of geometry on the interfacial failure load of bi-material interface, we selected the proper parameters of CZM. In detail, the cohesive strength T_c = 45Pa, the fracture energy of the double-ply laminate G_c =1.5 J/m^2 and the fracture energy of the three-ply laminate G_c =1 J/m^2.

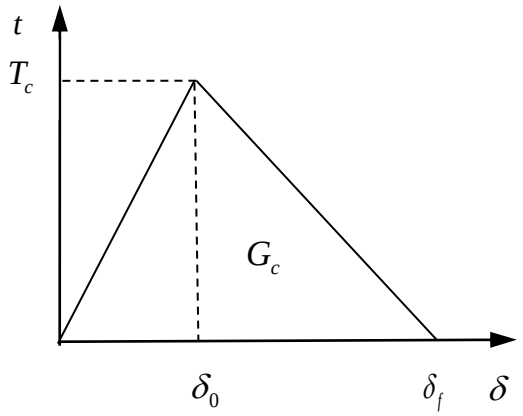


Figure 2. Constitutive relations of bilinear cohesive zone model

Results and discussion

The load-displacement curve of the bi-material interface can be divided into three stages according to the slope value, as shown in Fig.3. Obviously, the three stages of the curve correlating to the three deformation processes of the bi-metal laminates. The first stage is the linear compression stage, that is, the elastic deformation stage, which indicates that the interface has no damage. Afterward, the second stage is the nonlinear loading stage, which indicates the initiation and cumulation of the interface damage. Thereafter, there is a drop of

the load-displacement curve in the third stage, which indicates the complete failure of the cohesive interface and the delamination of the laminates.

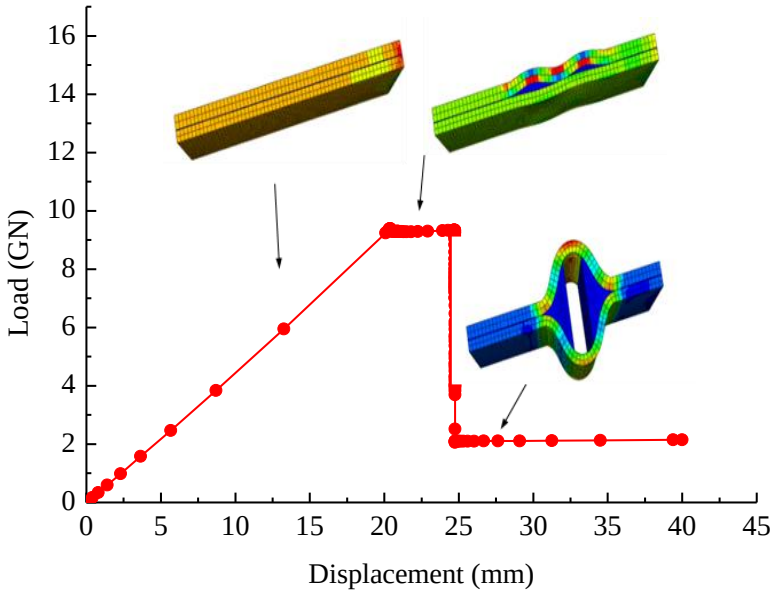


Figure 3. Interfacial debonding process of double-ply laminate

4.1 Effect of the thickness ratio

The prediction of the relationship between the interface bonding strength and the thickness of laminates has theoretical and practical significance for the effective utilization of composite laminates. As discussed in Section 2, the thickness of the aluminum alloy laminate was changed and the load-displacement curves at different thickness ratios were obtained, as shown in Fig.4and Fig.5.

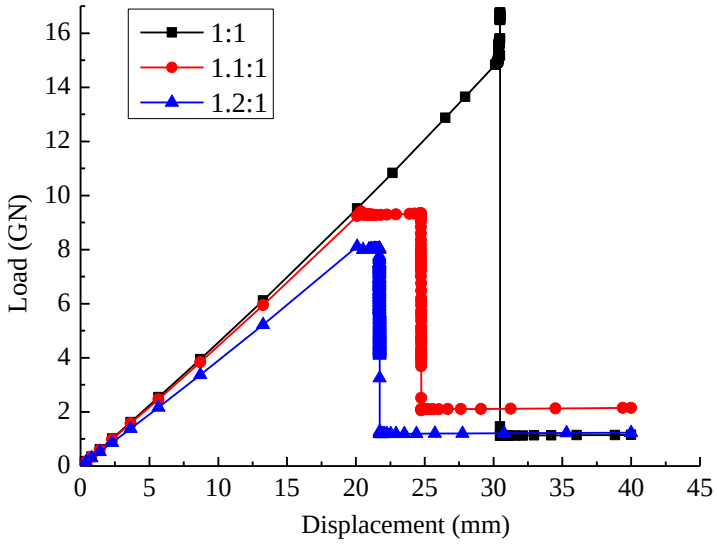


Figure 4. The load-displacement curves for various thickness ratio of double-ply laminate

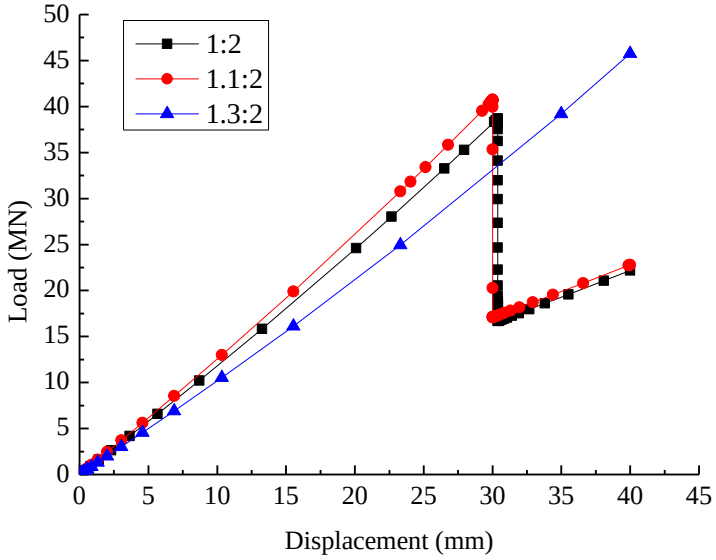


Figure 5. The load-displacement curves for various thickness ratio of three-ply laminate

The results (Fig.4) demonstrate that the ultimate load of the bi-material interface element of the double-ply laminate at critical failure decreases with the increase of thickness ratio. The bearing capacity of the bi-metal laminates is the strongest, when the thickness ratio is 1:1. By observing the linear stage of the curve, it is found that the bi-metal laminate at the thickness ratio of 1:1 is most difficult to bend.

On the contrary, as illustrated in Fig.5, the failure load of the bi-material interface element of the three-ply laminate increases by increasing the thickness ratio. Moreover, the laminate still has the bearing capacity after the failure of the interface. Besides, when the compression displacement is constant and the thickness ratio increases to a certain value, the bi-metal laminates is always in the linear compression stage. As shown in the blue line of Fig.5, the bi-material interface with a thickness ratio of 1.3:1 has no failure when the compression displacement is 4cm.

4.2 Effect of the interfacial parameters

In the practical engineering, the properties of the bi-material interface are closely related to the bonding of the laminates. Practically, the properties of the bi-material cohesive interface are controlled by two key parameters: the cohesive strength T_c and the fracture energy G_c.

In the above section, the typical load-displacement relation was obtained with the initial the double-ply laminate, whose thickness ratio is1:1. In order to study the influence of the two interface parameters on mechanical properties of the bi-material, the fracture energy stayed constant, and the cohesive strength is 55Pa, 45Pa, 35Pa respectively. By observing the load-displacement curves (Fig.6), it is noticed that the ultimate failure load increases with enhancing cohesion strength of the interface. Therefore, the mechanical properties of bi-material interface could be designed by properly increasing the cohesive strength.

In another case, keeping the cohesion strength of the interface unchanged, the value of fracture energy is 2 J/m^2, 1.5 J/m^2 and 1 J/m^2 respectively, and the corresponding load-displacement curves are obtained, as shown in Fig.7. It is found that the fracture energy has little influence on the ultimate failure load, but it also increases with the increase of the fracture energy.

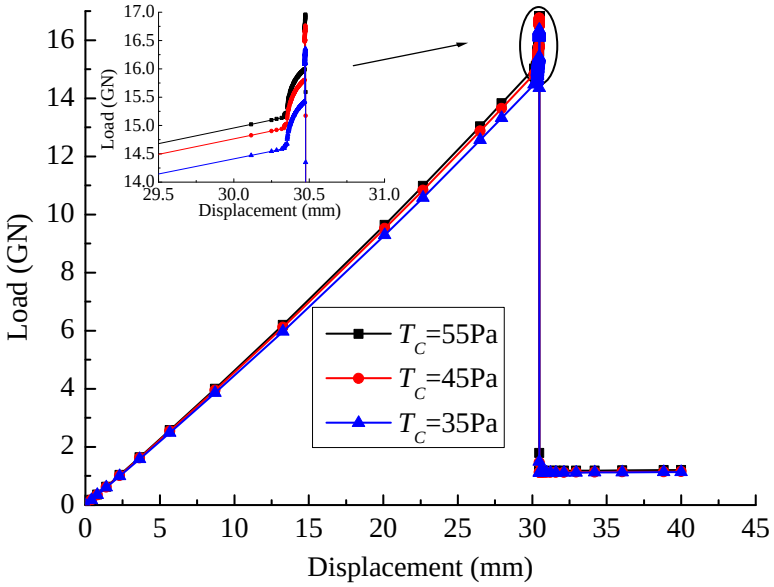


Figure 6. The load-displacement curves for various cohesive strengths

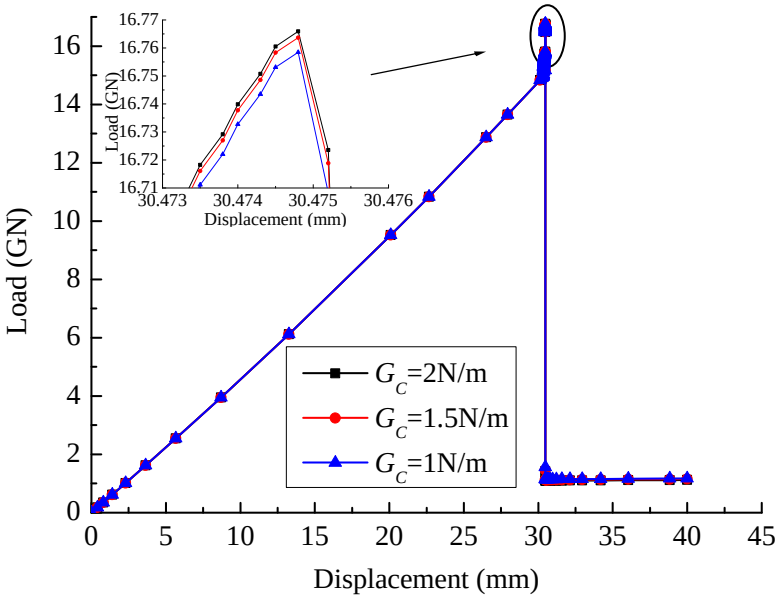


Figure 7. The load-displacement curves for various fracture energies

Conclusions

In this paper, a bilinear cohesive law was applied in cohesive finite element model to characterize the bi-material interface under compression loads. The load-displacement curves and the interfacial damage modes are obtained for different combinations of materials. The conclusions can be summarized as follows:

1. The ultimate load of the bi-material interface element of the double-ply laminate at critical failure decreases with the increase of thickness ratio. Moreover, it is found that the bi-metal laminate at the thickness ratio of 1:1 is most difficult to bend.
2. The failure load of the bi-material interface element of the three-ply laminate increases by increasing the thickness ratio. Besides, when the compression displacement is constant and the thickness ratio increases to a certain value, the bi-metal laminate is always in the linear compression stage.
3. The ultimate failure load increases with enhancing cohesion strength of the interface. Meanwhile, the fracture energy has little influence on the ultimate failure load.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under the Project Numbers 11502007, 11472020 and 11632005, which are gratefully acknowledged.

References

[1] Lesuer D R, Syn C K, Sherby O D, et al. Mechanical behaviour of laminated metal composites[J]. International Materials Reviews, 1996, 41(5):169-197.
[2] Wang F, Liu L, Liu Q, et al. Studies of Bimaterial Interface Fracture with Peridynamics[C]// International Power, Electronics and Materials Engineering Conference. 2015.
[3] Shet C, Chandra N. Analysis of Energy When Using Cohesive Zone Models to Simulate Fracture[J]. Transactions of the Asme Journal of Engineering Materials & Technology, 2002, 124:pág̃s. 440-450.
[4] Liu X, Yang Q, Su L. Interface analysis and design in carbon nanotube array composite based on cohesive finite element approach[J]. Materials Science & Engineering A, 2014, 592(2):83-87.
[5] Guo X, Leung A Y T, Chen A Y, et al. Investigation of non-local cracking in layered stainless steel with nanostructured interface[J]. Scripta Materialia, 2010, 63(4):403-406.
[6] Atkinson C. The interface crack with a contact zone[J]. International Journal of Fracture, 1982, 19(2):161-177.
[7] Atkinson C. The interface crack with a contact zone[J]. International Journal of Fracture, 1982, 19(2):131-138.
[8] Chandra N, Li H, Shet C, et al. Some issues in the application of cohesive zone models for metal–ceramic interfaces[J]. International Journal of Solids & Structures, 2002, 39(10):2827-2855.
[9] Salomonsson K. Mixed mode modeling of a thin adhesive layer using a meso-mechanical model[J]. Mechanics of Materials, 2008, 40(8):665-672.
[10] Fiedler B, Hojo M, Ochiai S. The influence of thermal residual stresses on the transverse strength of CFRP using FEM[J]. Composites Part A Applied Science & Manufacturing, 2002, 33(10):1323-1326.
[11] Zhang J, Huang Q Q, Qu Z. Elastic Fields of Interacting Elliptical Inhomogeneities for Two-Dimensional Problems Based on the Equivalent Inclusion Method[J]. Applied Mechanics & Materials, 2014, 501-504(5):2515-2519.
[12] Hashagen F, Borst R D, Vries T D. Delamination behavior of spliced Fiber Metal Laminates. Part 2. Numerical investigation[J]. Composite Structures, 1999, 46(2):147-162.
[13] Hashagen F, Borst R D. Numerical assessment of delamination in fibre metal laminates[J]. Computer Methods in Applied Mechanics & Engineering, 2000, 185(2–4):141-159.
[14] Harper P W, Hallett S R. A fatigue degradation law for cohesive interface elements – Development and application to composite materials[J]. International Journal of Fatigue, 2010, 32(11):1774-1787.
[15] Lu W B, Wu J, Song J, et al. A cohesive law for interfaces between multi-wall carbon nanotubes and polymers due to the van der Waals interactions[J]. Computer Methods in Applied Mechanics & Engineering, 2008, 197(41–42):3261-3267.

Nyström extrapolation algorithm for solving delay Volterra integral equations with weakly singular kernel☆

Li Zhang, Jin Huang, Yubin Pan, Hongyan Liu

School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu Sichuan 611731, China

Abstract

This paper concentrates on the numerical solutions of the weakly singular Volterra integral equations with vanishing delay by means of the Nyström method based on the modified rectangular quadrature formula. An extrapolation type arithmetic can be constructed according to the Euler-Maclaurin expansion of error. The algorithm is beneficial to increase the convergence rate of the approach and reaches higher accuracy than the former procedures. Moreover, a posterior error estimate is derived, which can be used to formulate self-adaptive algorithm. Numerical examples demonstrate that this method is effective and applicable.

Keywords: , weakly singular Volterra integral equations, vanishing delay, Nyström method, extrapolation, a posterior error estimate

1. Introduction

In recent years, Volterra integral equations (VIEs) with proportional delay have received a considerable amount of attentions.The VIEs play an important role in the fields of science,engineering and radiative heat transfer problems. There are many numerical techniques to solve these equations. In this paper, we mainly focus on the delay Volterra integral equations with weakly singular kernel as follows:

$$u(t) = g(t) + \int_0^{qt} s^{\alpha} k(t,s)u(s)ds, t \in [0,T], \tag{1}$$

where g and k are continuous when $t \in [0,T]$, $0 < q < 1, -1 < \alpha < 0$.

In the literature [1], Brunner H.introduced the collocation on graded mesh to solve the weakly singular VIEs. Then,in [2] Chen Y and Tang T assorted the spectral methods to solve weakly singular VIEs with smooth solutions. As to the delay equations, Ali I et al. employed the spectral method for the pantograph-type delay differential equations in [3] and the differential-integral equations with multiple delay in [4]. H.xie.et al.[5] have solved VIEs with vanishing delay by means of the collocation methods. In the recent

☆This work was supported by the National Natural Science Foundation of China (11371079)

Email addresses: lizhang_137363@163.com (Li Zhang), huangjin12345@163.com (Jin Huang), yubinpan2014@163.com (Yubin Pan), Hongyan_Liu1528@163.com (Hongyan Liu)

references, emerging a lot of algorithms, as like the new Bernouli wavelet method [6], multistep Legendre pseudo-spectral method [7], the least squares approximation method [8] and [9] Runge-Kutta method.

However, the researchers focus on solving the equations whose kernel are continuous, few of documents pay attention to the situation that the kernels are singular. Some of the methods mentioned above are lower order methods and have the characteristics of needing large amount of calculations and the storage space is large. Then we propose a kind of algorithm for delay VIEs with vanishing delays, which is called the Nyström extrapolation algorithm based on the mid-rectangular formula. By this means,the calculation scheme is simple and high accuracy. Through the form of the assignment the values only occupying small storage space contrasted with the collation methods.

The remainder of this paper as follows:In section 2, we introduce the Nyström methods based on the modify mid-rectangle formula.In section 3, we get some numerical results, which demonstrate this method is efficient and practical. The final section is for concluding remarks.

2. The numerical methods

In this section, we will construct the Nyström methods [10] based on the modify mid-rectangle formula.When the collocation points have choosed, then,the equation (1) is turned into fredholm type. We first consider the integral

$$Q(g) = \int_{\Omega} g(x)dx, \tag{2}$$

where $\Omega \in R^s$, suppose that there exists n-points approximate quadrature formula

$$Q_n(g) = \sum_{j=1}^{n} \omega_j^{(n)} g(x_j^{(n)}), \tag{3}$$

where $x_j^{(n)}$ are the basic quadrature points, $\{\omega_j^{(n)}\}_{j=1}^n$ are the corresponding integral weights, which satisfy the condition

$$\sum_{j}^{n} |\omega|_j^{(n)} \leq C, \tag{4}$$

and is a constant and independents of the n . When $n \to \infty, Q_n(g) \to Q(g)$. Consider

$$u(x) - \int_{\Omega} k(x,y)u(y)dy = f(x). \tag{5}$$

By means of the quadrature formula we can deduce the approximate equations

$$u_n(x) - \sum_{j}^{n} \omega_j^{(n)} k(x, x_j^{(n)}) u_n(x_j^{(n)}) = f(x), x \in \Omega, \tag{6}$$

Once $\{u_n(x_i), i = 1, \cdots, n\}$ are solved, by the Nyström interpolation methods, we have

$$u_n(x) = \sum_{j}^{n} \omega_j^{(n)} k(x, x_j^{(n)}) u_n(x_j^{(n)}) + f(x), x \in \Omega, \tag{7}$$

specially,when $\omega_i^{(n)} = h$, it's the mid-rectangle formula. In order to solve the weakly singular kernel, we must make some appropriate changes. For simplicity, we consider the type of linear,as follows:

$$u_i = g(x_i) + h \sum_{j=0}^{i-1} (x_i - x_{j+\frac{1}{2}})^{\alpha} k(x_i, x_{j+\frac{1}{2}}) \frac{(u_j + u_{j+1})}{2}, i = 1, \cdots, N \tag{8}$$

where $(x_i)_{i=1}^n$ are collocation points, and the original value $u_0 = g(0)$, Assume that $u \in C^2[0, T]$ is the solution of (1) , and the kernel satisfies Lipschitz continuity condition with constant s and t, Then, the expansion of errors is

$$u_i - u(x_i) = T_0^M(x_i) h^{2+\alpha} + O(h^2), i = 1, \cdots, N. \tag{9}$$

Therefore, in order to achieve high accuracy, we can utilize the technology of extrapolation with the aid of the error of gradual expansion. Furthermore, we consider a posteriori error estimate.First, we execute Richardson extrapolation with $h^{2+\alpha}$, get

$$u_i^h = \frac{2^{2+\alpha} u_i^{(\frac{h}{2})} - u_i^h}{2^{2+\alpha} - 1} = u(x_i) + O(h^2), i = 1, \cdots, N. \tag{10}$$

Now, we will construct the concrete scheme on delay integral equations with weakly singular kernel.First, we divide $I \in [0, T]$ into several intervals such that the stepsizes $h = \frac{T}{N}$,$t_i = ih, i = 0, 1, \cdots, N$. Let $t = t_i$, we get

$$\begin{aligned} u(t_i) &= g(t_i) + \int_0^{qt_i} k(t_i, s) u(s) ds \\ &= g(t_i) + \int_0^{t_{[qi]}} k(t_i, s) u(s) ds + \int_{t_{[qi]}}^{qt_i} k(t_i, s) u(s) ds \\ &= g(t_i) + I_1 + I_2, \end{aligned} \tag{11}$$

where $[qi]$ denote the integer part of qi. In order to structure the rectangular quadrature algorithm, only using modify mid-rectangle formula on I_1 and I_2,

$$I_1 = \int_0^{t_{[qi]}} k(t_i, s) u(s) ds \approx h \sum_{k=0}^{[qi]-1} k(t_i, t_{k+\frac{1}{2}}) u(t_{k+\frac{1}{2}}), \tag{12}$$

and

$$I_2 = \int_{t_{[qi]}}^{qt_i} k(t_i, s) u(s) ds \approx (qt_i - t_{[qi]}) k(t_i, \frac{qt_i + t_{[qi]}}{2}) u(\frac{qt_i + t_{[qi]}}{2}). \tag{13}$$

Taking into account $u(qt_i)$ are not the node values, so we can turn to the linear interpolation approximation with values of the adjacent points $u(t_{[qi]})$ and $u(t_{[qi]}+1)$, with $t_{[qi]} \leq qt_i \leq t_{[qi]+1}$. Then, there exists $\beta_i \in [0,1]$ such that $qt_i = \beta_i t_{[qi]} + (1-\beta_i)t_{[qi]+1}$ are established, where $\beta_i = 1 + [qi] - qi$. Clearly, we can get

$$u(\frac{qt_i + t_{[qi]}}{2}) = \frac{1+\beta}{2}u(t_{[qi]}) + \frac{1-\beta}{2}u(t_{[qi]+1}), \tag{14}$$

and $u(t_{k+\frac{1}{2}}) = \frac{1}{2}u(t_k) + \frac{1}{2}u(t_{k+1})$.

3. The numerical results

Two examples have been presented to show the efficiency of the Nyström extrapolation algorithms.For simplicity,we design a set of grids on the interval I ,we take $N = 40, 80$ and $h = \frac{1}{N}$ in the following two examples.

Example 1. Let us consider the second kind linear volterra integral equation of

$$u(t) = t^{\frac{5}{2}} - 0.243t^5 - 0.164025t^4 + \int_0^{qt} s^{-\frac{1}{2}}(t^2+s)u(s)ds, t \in [0,T], \tag{15}$$

with $q = 0.9, T = 1$,the initial value $u(0) = g(0) = 0$ and the analytical solution is $u(t) = t^{\frac{5}{2}}$. we denote the approximate values by u_h, The absolute errors $e_h(t) = \ u(t) - u_h(t)|, t \in [0,1]$ and the posteriori error have been presented in the Table.1.Futhermore, the $h^{2+\alpha}$ extrapolation values and the h^2 extrapolation values have also been shown in the Table.1. indicate that our algorithm provides high accuracy results for the weakly singular at endpoint.

Table 1: The results of example 1

t	e_h	$e_{\frac{h}{2}}$	$h^{2+\alpha}$ extrapolation	h^2 extrapolation	posteriori error
0.2	3.96e-06	1.01e-06	6.05e-07	2.79e-07	1.61e-06
0.4	2.42e-05	6.11e-06	3.82e-06	1.72e-07	9.93e-06
0.6	7.69e-05	1.91e-05	1.25e-05	2.12e-07	3.16e-05
0.8	1.87e-04	4.65e-05	3.01e-05	8.97e-07	7.67e-05
1	3.95e-04	9.87e-05	6.33e-05	1.26e-07	1.62e-04

Example 2. This example is to further demonstrate the effectiveness of the Nyström method based on the modify mid-rectangle formula.We consider the equation as follows:

$$u(t) = t - \frac{2}{3} \times 0.8^{\frac{3}{2}} t^{\frac{3}{2}} + \int_0^{qt} s^{-\frac{1}{2}}u(s)ds, t \in [0,T], \tag{16}$$

the analytical solution is $u(t) = t$,we take $q = 0.8, T = 1$ and the initial value $u(0) = g(0) = 0$ .We use the Nyström extrapolation scheme for solving (16).From the Table 2, we see that the approximate solution obtained by the $h^{2+\alpha}$ extrapolation and the h^2 extrapolation are effictive on the whole interval.

Table 2: The analysis of example 2

t	e_h	$e_{\frac{h}{2}}$	$h^{2+\alpha}$ extrapolation	h^2 extrapolation	posteriori error
0.2	3.2907e-04	1.3047e-04	2.1853e-05	5.0705e-06	1.0862e-04
0.4	4.5591e-04	1.7906e-04	2.7650e-05	6.3606e-06	1.5141e-04
0.6	5.7579e-05	2.2550e-04	3.3926e-05	7.9592e-06	1.9158e-04
0.8	6.9634e-04	2.7240e-04	4.0548e-05	9.8280e-06	2.3186e-04
1	8.2037e-04	3.2066e-04	4.7359e-05	1.1556e-05	2.7330e-04

4. Conclusion

From the above tables,we can conclude that the Nyström method based on the modify mid-rectangle formula is an efficient and accurate numerical technique for Volterra integral equations of second with weakly singular kernel. In particular, this approach is novel for various problem with cauchy singular kernel and the hyper-singular delay integral equations.

5. references

[1] Brunner H. The numerical solutions of weakly singular Volterra integral equations by collocation on graded mesh. Math Comp, 1985, 45: 417-437

[2] Chen Y, Tang T. Spectral methods for weakly singular Volterra integral equations with smooth solutions. J Comput Appl Math, 2009, 233: 938C950

[3] Ali I, Brunner H, Tang T. A spectral method for pantograph-type delay differential equations and its convergence analysis. J Comput Math, 2009, 27: 254C265

[4] Ali I, Brunner H, Tang T. Spectral method for pantograph-type differential and integral equations with multiple delays. Front Math China, 2009, 4: 49C61

[5] H. Xie , R. Zhang , H. Brunner , Collocation methods for general Volterra functional integral equations with vanishing delays, SIAM J. Sci. Comput. 33 (2011)3303C3332.

[6] P.K.Sahu,S.Saha Ray,A new Bernoulli wavelet method for accurate solutions of nonlinear fuzzy Hammerstein-Volterra delay integral equations,FSS,399(2017)131-144.

[7] Zhang Xiao-yong,Convergence analysis of the multistep Legendre pseudo-spectral method for Volterra integral equations with vanishing delays,J.Comput.Appl.Math,321(2017)284-301.

[8] Maryam Mosleh,Mehmood Otadi,Least squares approximation method for the solution of Hammerstein-Volterra delay integral equations,AMC,258(2015)105-110.

[9] Saberi-Nadiafi,M.Tamamagar,A generalized block-by-block method for solving linear Volterra integral equations,Appl.Math.Comput.188(2017)1969-1974.

[10] J.Huang,T.Lü,High accuracy algorithm of integral eqautions[M],Science press,2013.

A complex variable interpolating meshless method for two-dimensional transient heat conduction problem

***Yajie Deng[1], †Xiaoqiao He[1,2]**

[1] Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong
[2] Center for Advanced Structural Materials, City University of Hong Kong Shenzhen Research Institute, Shenzhen 518057, China

*Presenting author: yjdeng7-c@my.cityu.edu.hk
†Corresponding author: bcxqhe@cityu.edu.hk

Abstract

In this paper, on the basis of complex variable moving least squares approximations, a new complex variable moving least squares interpolating (CVMLSI) method is proposed. In this method, a complete basis function and singular weight function are introduced to form the new basis function through orthogonalization process. Then a new interpolating shape function is derived, which satisfy the property of Kronecker δ function.

Combining the CVMLSI method with the weak integral form of the two-dimensional transient heat conduction problem, a complex variable element free Galerkin interpolating (CVEFGI) method for transient heat conduction problem is obtained. Due to the fact that the essential boundary conditions can be applied directly, the final discrete matrix equation is more concise than that in the non-interpolating complex variable element free Galerkin method. Finally, a numerical example is presented to illustrate the advantages of the CVEFGI method.

Keywords: Complete basis function, interpolating meshless method, Kirchhoff plates

Introduction

As a numerical tool, the meshless method [1] has developed widely in engineering analysis. Different with the traditional mesh based numerical methods, such as the finite element method, and boundary element method [2][3], the meshless method is built on a series of discrete nodes. So when using the meshless method to solve some special and complicated problems, such as nonlinear large deformation of polymer gel, crack propagation problem and so on, the re-meshing techniques are not necessary during the computing process [4][5].

One of the most common used methods to build the trial function is the moving least squares (MLS) approximation [6]. On the base of the MLS approximation, Belytschko presented the element free Galerkin (EFG) method [7]. With the development of the meshless method, a variety of complex variable moving least squares method approximations were proposed on the foundation of the MLS approximation. In these complex variable moving least squares approximations, the basis functions $\boldsymbol{p}^{\mathrm{T}}(z)=(1,z)$ and $\overline{\boldsymbol{p}}^{\mathrm{T}}(z)=(1,\bar{z})$ are used to construct the trial functions [8][9]. Then the trial function of a two-dimensional problem can be formed with a one-dimensional basis function, which leads to the complex variable moving least squares approximations have higher efficiency.

However, the basis functions above mentioned are not complete basis functions, which can not express all functions in the problem domain and may reduce the computing accuracy. Besides, in most of the complex variable moving least squares approximations, the obtained

shape functions can not satisfy the property of Kronecker δ function. In the meshless method built on these approximations, special techniques are useful to apply the essential boundary conditions, such as Lagrange multiplier and penalty methods [1][10].

Trying to solve above two problems, in this paper, we introduce a complete basis function $\boldsymbol{p}^{\mathrm{T}}(z)=(1,z,\bar{z})$ and the singular weight function. Then reference the idea presented by Ren [11], improve the basis function with orthogonalization process to get the new interpolating shape function. Then a new complex variable moving least squares interpolating (CVMLSI) method is presented. Combining the CVMLSI method with the weak integral form of the two-dimensional transient heat conduction problem, a complex variable element free Galerkin interpolating (CVEFGI) method for heat conduction problem is obtained and the final matrix equation is derived. Finally, a numerical example is solved to validate the advantages of the CVEFGI method compared with non-interpolating complex variable element free Galerkin method.

Methodology

In this part, the CVMLSI method is introduced. According to the improved complex variable moving least squares (ICVMLS) approximation presented by Bai [9], the trial function can be expressed as

$$u^h(z)=u_1^h(z)+\mathrm{i}u_2^h(z)=\sum_{i=1}^{m}p_i(z)a_i(z)=\boldsymbol{p}^{\mathrm{T}}(z)\boldsymbol{a}(z),$$
$$(z=x_1+\mathrm{i}x_2\in\Omega), \tag{1}$$

where the $\boldsymbol{p}^T(z)=(p_1(z),p_2(z),\ldots,p_m(z))$ is the complete basis function vector. In the two-dimensional domain, the linear and quadratic basis function vectors are shown as

$$\boldsymbol{p}^{\mathrm{T}}(z)=(1,z,\bar{z}), \tag{2}$$
$$\boldsymbol{p}^{\mathrm{T}}(z)=(1,z,\bar{z},z^2,\bar{z}^2,z\bar{z}). \tag{3}$$

The local approximation at point z can be expressed as

$$u^h(z,\hat{z})=\sum_{i=1}^{m}p_i(\hat{z})a_i(z)=\boldsymbol{p}^{\mathrm{T}}(\hat{z})\boldsymbol{a}(z), \tag{4}$$

where $\hat{z}$ is the node whose influence domain covers the point z.

Then, combining the singular weight function to make the following improvement on the space $span(p_1,p_2,...,p_m)$, $p_1(z)\equiv 1$ is normalized as [12]

$$\beta_z^{(1)}(z)=\frac{p_1}{\|p_1\|_z}=\frac{1}{[\sum_{I=1}^{n}w(z-z_I)]^{\frac{1}{2}}}, \tag{5}$$

let $p_i(z)\ \ (i=2,3,...,m)$ be orthogonal to $\beta_z^{(1)}(z)$, then we can gain that

$$b_z^{(i)}(z)=p_i(z)-(p_i(z),\beta_z^{(1)}(z))_z\beta_z^{(1)}(z)$$
$$=p_i(z)-\frac{\sum_{I=1}^{n}p_i(z_I)w(z-z_I)}{\sum_{I=1}^{n}w(z-z_I)}$$

$$= p_i(z) - \sum_{I=1}^{n} p_i(z_I) v(z - z_I)\,,\ (i = 2,3,...,m)\,, \tag{6}$$

that is

$$b_z^{(i)}(z) = p_i(z) - p_i^{(s)}(z) = p_i(z) - \boldsymbol{v}^{\mathrm{T}}(z)\boldsymbol{p}_i\,. \tag{7}$$

Then, the $\beta_z^{(1)}(z), b_z^{(2)}(z), ..., b_z^{(m)}(z)$ are the new basis function, and the new interpolating function is

$$u^h(z) = (u, \beta_z^{(1)}(z))_z \beta_z^{(1)}(z) + \sum_{i=2}^{m} a_i(z) b_z^{(i)}(z)\,, \tag{8}$$

i.e.

$$u^h(z) = \boldsymbol{v}^{\mathrm{T}}(z)\boldsymbol{u} + \boldsymbol{b}^{\mathrm{T}}(z)\boldsymbol{a}(z)\,, \tag{9}$$

where

$$\boldsymbol{v}(z) = (v(z - z_1), v(z - z_2), ..., v(z - z_n))^{\mathrm{T}}\,, \tag{10}$$

$$\boldsymbol{u} = (u(z_1), u(z_2), ..., u(z_n))^{\mathrm{T}}\,, \tag{11}$$

$$\boldsymbol{a}(z) = (a_2(z), a_3(z), ..., a_m(z))^{\mathrm{T}}\,, \tag{12}$$

$$\boldsymbol{b}(z) = (b_z^{(2)}(z), b_z^{(3)}(z), ..., b_z^{(m)}(z))^{\mathrm{T}}\,. \tag{13}$$

In Eq. (10), $v(z - z_i)$ is the normalized weight function when consider the special condition $m = 1$, which is a weighted average of the function values at node z_I in the influence domain of point z. The expression form is

$$v(z - z_i) = \frac{w(z - z_i)}{\sum_{I=1}^{n} w(z - z_I)}\,, \quad (i = 1,2,...,n)\,. \tag{14}$$

Using the functional J in the ICVMLS approximation [9] and ensure the J is minimum, the corresponding unknown coefficient vector is obtained as,

$$\boldsymbol{a}(z) = \boldsymbol{A}_z^{-1}(z)\boldsymbol{B}_z(z)\boldsymbol{u}\,, \tag{15}$$

where

$$\boldsymbol{A}_z(z) = \overline{\boldsymbol{C}}_z^{\mathrm{T}}\boldsymbol{W}(z)\boldsymbol{C}_z\,, \tag{16}$$

$$\boldsymbol{B}_z(z) = \overline{\boldsymbol{C}}_z^{\mathrm{T}}\boldsymbol{W}(z)\,, \tag{17}$$

$$\boldsymbol{C}_z = \begin{bmatrix} b_z^{(2)}(z_1) & b_z^{(3)}(z_1) & \cdots & b_z^{(m)}(z_1) \\ b_z^{(2)}(z_2) & b_z^{(3)}(z_2) & \cdots & b_z^{(m)}(z_2) \\ \vdots & \vdots & \ddots & \vdots \\ b_z^{(2)}(z_n) & b_z^{(3)}(z_n) & \cdots & b_z^{(m)}(z_n) \end{bmatrix}, \tag{18}$$

$$\boldsymbol{W}(z) = \begin{bmatrix} w(z - z_1) & 0 & \cdots & 0 \\ 0 & w(z - z_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w(z - z_n) \end{bmatrix}. \tag{19}$$

Substituting Eq. (15) into Eq. (9), we can obtain

$$u^h(z) = \boldsymbol{v}^{\mathrm{T}}(z)\boldsymbol{u} + \boldsymbol{b}^{\mathrm{T}}(z)\boldsymbol{A}_z^{-1}(z)\boldsymbol{B}_z(z)\boldsymbol{u} = \boldsymbol{\Phi}(z)\boldsymbol{u} = \sum_{I=1}^{n} \Phi_I(z) u(z_I)\,, \tag{20}$$

where $\boldsymbol{\Phi}(z)$ is the new interpolating shape function

$$\boldsymbol{\Phi}(z)=(\Phi_1(z),\Phi_2(z),...,\Phi_n(z))=\boldsymbol{v}^{\mathrm{T}}(z)+\boldsymbol{b}^{\mathrm{T}}(z)\boldsymbol{A}_z^{-1}(z)\boldsymbol{B}_z(z)\,, \tag{21}$$

and

$$u_1^h(z)=\mathrm{Re}[\boldsymbol{\Phi}(z)\boldsymbol{u}]=\mathrm{Re}[\sum_{I=1}^{n}\Phi_I(z)u(z_I)]\,, \tag{22}$$

$$u_2^h(z)=\mathrm{Im}[\boldsymbol{\Phi}(z)\boldsymbol{u}]=\mathrm{Im}[\sum_{I=1}^{n}\Phi_I(z)u(z_I)]\,. \tag{23}$$

The following singular weight function is used as [13]

$$w(z-z_I)=\begin{cases}\dfrac{\rho^2}{|z-z_I|^2}\left(1-\dfrac{\rho}{|z-z_I|}\right)^2 & |z-z_I|\le\rho\\ 0 & |z-z_I|\ge\rho\end{cases}\,, \tag{24}$$

where ρ is the radius of the influence domain of the point z. This singular weight function can satisfy general properties of other weight functions, such as the cubic and quartic spline weight function [1].

This is the CVMLSI method. Because the new complete basis function and the singular weight function is used to construct the trial function, the CVMLSI method has higher accuracy and the shape function obtained from this method can satisfy the property of Kronecker δ function.

Numerical Example

The CVMLSI method is used to solve the two-dimensional transient heat conduction problem. The governing equation is [14]

$$\rho c\frac{\partial T(\boldsymbol{x},t)}{\partial t}-\frac{\partial}{\partial x_1}(k_1\frac{\partial T(\boldsymbol{x},t)}{\partial x_1})-\frac{\partial}{\partial x_2}(k_2\frac{\partial T(\boldsymbol{x},t)}{\partial x_2})-Q=0\,,$$
$$(\boldsymbol{x}=(x_1,x_2)\in\Omega), \tag{25}$$

with the following boundary conditions and initial condition

$$T(\boldsymbol{x},t)-\overline{T}=0\,,\qquad(\boldsymbol{x}\in\Gamma_1), \tag{26}$$

$$k_1\frac{\partial T(\boldsymbol{x},t)}{\partial x_1}n_1+k_2\frac{\partial T(\boldsymbol{x},t)}{\partial x_2}n_2-\overline{q}=0\,,\qquad(\boldsymbol{x}\in\Gamma_2), \tag{27}$$

$$k_1\frac{\partial T(\boldsymbol{x},t)}{\partial x_1}n_1+k_2\frac{\partial T(\boldsymbol{x},t)}{\partial x_2}n_2-h(T_a-T(\boldsymbol{x},t))=0\,,\quad(\boldsymbol{x}\in\Gamma_3), \tag{28}$$

$$T(\boldsymbol{x},0)=T_0\,, \tag{29}$$

where $T(\boldsymbol{x},t)$ is the temperature field function, Q is the internal heat generation per unit volume, $\overline{T}$ is the given temperature on the boundary Γ_1, $\overline{q}$ is the given density of the teat flux on the boundary Γ_2 and T_0 is the given initial temperature.

When build the CVEFGI method for the two-dimensional transient heat conduction problem, using the CVMLSI method to disperse the spatial domain and using the difference method to disperse the time, then we can obtain the final matrix equation

$$(\frac{\boldsymbol{C}}{\Delta t}+\theta\overline{\boldsymbol{K}})\boldsymbol{T}_{t+\Delta t}=(\frac{\boldsymbol{C}}{\Delta t}-(1-\theta)\overline{\boldsymbol{K}})\boldsymbol{T}_t+\theta\overline{\boldsymbol{F}}_{t+\Delta t}+(1-\theta)\overline{\boldsymbol{F}}_t, \tag{30}$$

where Δt is the time step,

$$\boldsymbol{C}=\int_\Omega \widetilde{\boldsymbol{\Phi}}^{\mathrm{T}}(z)\rho c\widetilde{\boldsymbol{\Phi}}(z)\mathrm{d}\Omega, \tag{31}$$

$$\overline{\boldsymbol{K}}=\boldsymbol{K}+\boldsymbol{H}, \tag{32}$$

$$\overline{\boldsymbol{F}}=\boldsymbol{F}^{(1)}+\boldsymbol{F}^{(2)}+\boldsymbol{F}^{(3)}, \tag{33}$$

$$\boldsymbol{K}=\int_\Omega \boldsymbol{B}^{\mathrm{T}}(z)\widetilde{\boldsymbol{k}}\boldsymbol{B}(z)\mathrm{d}\Omega, \tag{34}$$

$$\boldsymbol{H}=\int_{\Gamma_3} \widetilde{\boldsymbol{\Phi}}^{\mathrm{T}}(z)h\widetilde{\boldsymbol{\Phi}}(z)\mathrm{d}\Gamma, \tag{35}$$

$$\boldsymbol{F}^{(1)}=\int_\Omega \widetilde{\boldsymbol{\Phi}}^{\mathrm{T}}(z)Q\mathrm{d}\Omega, \tag{36}$$

$$\boldsymbol{F}^{(2)}=\int_{\Gamma_2} \widetilde{\boldsymbol{\Phi}}^{\mathrm{T}}\overline{q}\mathrm{d}\Gamma, \tag{37}$$

$$\boldsymbol{F}^{(3)}=\int_{\Gamma_3} \widetilde{\boldsymbol{\Phi}}^{\mathrm{T}}(z)hT_a\mathrm{d}\Gamma. \tag{38}$$

This is the CVEFGI method for the two-dimensional transient heat conduction problem. Compared with the improved complex variable element-free Galerkin (ICVEFG) method based on the ICVMLS approximation, the above expressions of matrices are more concise [14].

For a specific two-dimensional transient heat conduction problem with a square domain, the control equation is

$$\frac{\partial T(x_1,x_2,t)}{\partial t}-\frac{\partial^2 T(x_1,x_2,t)}{\partial x_1^2}-\frac{\partial^2 T(x_1,x_2,t)}{\partial x_2^2}+2=0,$$
$$x_1\in[0,\pi],\quad x_2\in[0,\pi], \tag{39}$$

and the essential boundary conditions are

$$T(0,x_2,t)=0, \tag{40}$$

$$T(\pi,x_2,t)=\pi^2, \tag{41}$$

$$T(x_1,0,t)=x_1^2, \tag{42}$$

$$T(x_1,\pi,t)=x_1^2. \tag{43}$$

The initial condition is

$$T(x_1,x_2,0)=x_1^2+\sin(x_1)\sin(x_2), \tag{44}$$

and the analytical solution is

$$T(x_1,x_2,t)=x_1^2+e^{-2t}\sin(x_1)\sin(x_2). \tag{45}$$

In the CVEFGI method, the linear basis function and 4×4 Gauss points are used. Try to make error analysis, the relative error is

$$e=\left|\frac{T_i^{num}-T_i^{exact}}{T_i^{exact}}\right|, \tag{46}$$

where T_i^{num} is the numerical solution and T_i^{exact} is the analytical solution.

In this example, 15×15 nodes are distributed uniformly in the square domain. The scaling parameter is $d_{\max}=2$, and the time step is $\Delta t=0.001s$. For the ICVEFG method, the penalty factor is $\alpha=1.0\times10^{4}$.

Fig. 1 and Fig. 2 compare the solutions of heat distributions at $x_2=\pi/2$ and $x_1=\pi/2$ respectively with different times. The numerical solutions obtained from the CVEFGI method and the ICVEFG method are in good agreement with the analytical solutions at different times. Fig. 3 and Fig. 4 show the relative errors of the CVEFGI method and the ICVEFG method at different times at $x_2=\pi/2$ and $x_1=\pi/2$ respectively. We can see compared with the ICVEFG method, the CVEFGI method has higher computing accuracy, especially on the borders.

Besides due to the essential boundary conditions can be applied directly in the CVEFGI method, the final discrete equation of this transient heat conduction problem is simpler than that in the ICVEFG method. And there is need to choose suitable penalty factor or Lagrange multiplier which will save more computing time than the ICVEFG method.

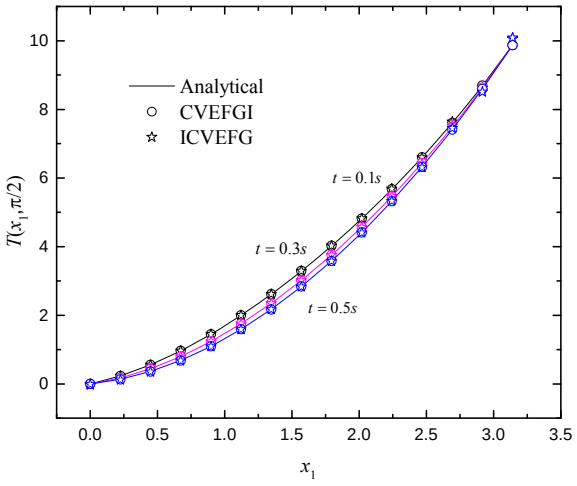


Figure 1. Heat distributions at $x_2=\pi/2$ with different times

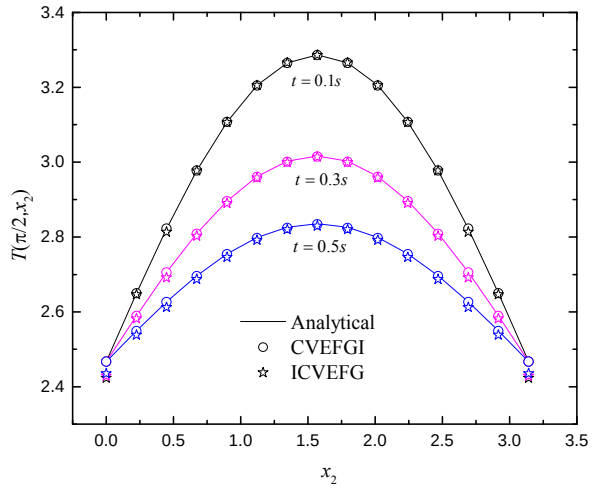


Figure 2. Heat distributions at $x_1 = \pi/2$ with different times

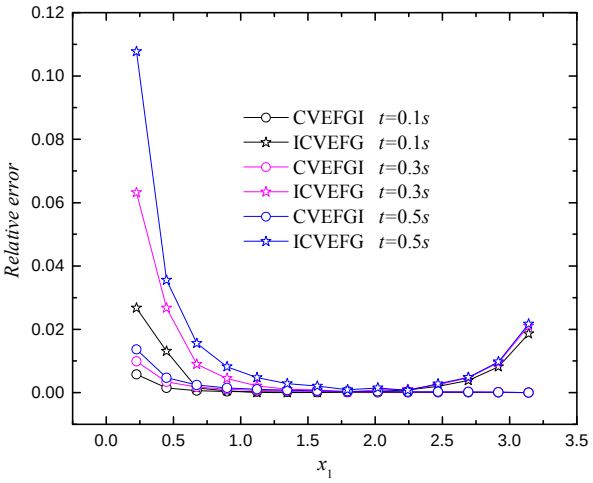


Figure 3. Relative errors at $x_2 = \pi/2$ with different times

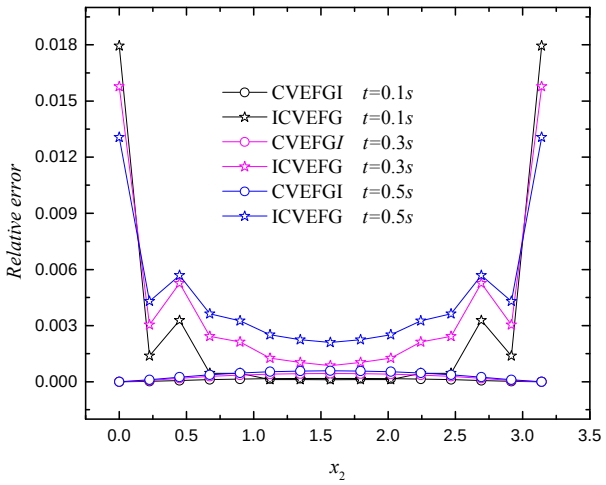


Figure 4. Relative errors at $x_1 = \pi/2$ with different times

Conclusions

In this paper, a complete basis function and singular weight function are introduced to derive the new shape function, and then the CVMLSI method is presented. In the CVMLSI method, the new shape function can satisfy the Kronecker δ function. So compared with the ICVMLS approximation with non-interpolating shape function, the CVMLSI method has higher accuracy. Combining the weak integral form of the two-dimensional transient heat conduction problem and the CVMLSI method, the CVEFGI method for the heat conduction problem is obtained. In the CVEFGI method, because the essential boundary conditions can be applied directly, the final matrix equation is more concise and it is unnecessary to choose appropriate penalty factor. The numerical example shows that the CVEFGI method is more accurate and efficient than the ICVEFG method.

Acknowledgements *Financial support from the Natural Science Foundation of China under Grant No. 11372264, the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No. CityU 114013] and the Science and Technology Innovation Commission of Shenzhen Municipality [Project No. JCYJ20160229165310679] are gratefully acknowledged.*

References

[1] Belytschko, T., Krongauz, Y., Organ, D., Fleming, M. and Krysl, P. (1996) Meshless method: An overview and recent developments, *Computer Methods in Applied Mechanics Engineering* **139**, 3-47.

[2] Zienkiewicz, O. C., Taylor, R. L., Zienkiewicz. O. C. and Taylor, R. L. (1977) *The finite element method*, McGraw-Hill, London.

[3] Banerjee, P. K. and Butterfield, R. (1981) *Boundary element methods in engineering science*, McGraw-Hill, London.

[4] Li, D. M., Zhang, Z. and Liew, K. M. (2014) A numerical framework for two-dimensional large deformation of inhomogeneous swelling of gels using the improved complex variable element-free Galerkin method, *Computer Methods in Applied Mechanics Engineering* **274**, 84-102.

[5] Zhang, Z., Liew, K. M., Cheng, Y. M. and Lee, Y. Y. (2008) Analyzing 2D fracture problems with the improved element-free Galerkin method, *Engineering Analysis with Boundary Elements* **32**, 241-250.

[6] Lancaster, P. and Salauskas, K. (1981) Surfaces generated by moving least squares methods, *Mathematics of Computation* **37**, 141-158.
[7] Belytschko, T., Lu, Y. Y. and Gu, L. (1994) Element free Galerkin methods, *International Journal for Numerical Methods in Engineering* **37**, 229-256.
[8] Peng, M. J., Liu, P. and Cheng, Y. M. (2009) The complex variable element-free Galerkin (CVEFG) method for two-dimensional elasticity problems, *International Journal of Applied Mechanics* **1**, 367-385.
[9] Bai, F. N., Li, D. M., Wang, J. F. and Cheng, Y. M. (2012) An improved complex variable element-free Galerkin method for two-dimensional elasticity problems, *Chinese Physics B* **21**, 020204.
[10] Zhu, T. and Alturi, S. N. (1998) Modified collocation method and a penalty function for enforcing the essential boundary conditions in the element free Galerkin method, *Computational Mechanics* **21**, 211-222.
[11] Ren, H. P. and Cheng, Y. M. (2012) The interpolating element-free Galerkin (IEFG) method for two-dimensional potential problems, *Engineering Analysis with Boundary Elements* **36**, 873-880.
[12] Ren, H. P., Cheng, J. and Huang, A. X. (2012) The complex variable interpolating moving least-squares method, *Applied Mathematics and Computation* **219**, 1724-1736.
[13] Deng, Y. J., Liu, C., Peng, M. J. and Cheng, Y. M. (2015) The interpolating complex variable element-free Galerkin method for temperature field problems, *International Journal of Applied Mechanics* **7**, 1550017.
[14] Wang, J. F. and Cheng, Y. M. (2012) A new complex variable meshless method for transient heat conduction problems, *Chinese Physics B* **21**, 120206.

Numerical solution of a bioheat transfer problem with transient blood temperature

†*Kuo-Chi Liu[1] and Fong-Jou Tu[2]

[1] Department of Mechanical Engineering, Far East University, Taiwan.
[2] Department of Electrical Engineering, Nan Jeon University, Taiwan

*Presenting author: kcliu@mail.feu.edu.tw
†Corresponding author: kcliu@mail.feu.edu.tw

Abstract

In the heat treatment process, blood perfusion starts up a negative feedback mechanism. The blood temperature undergoes a transient process before onset of equilibrium, and then changes the situation of temperature distribution. In substance, the blood temperature undergoes a transient process for heat exchange between blood and tissue. For more fully exploring the heat transfer behavior of biological tissue, this paper analyzes the bio-heat transfer problem with the non-constant blood temperature based on the Pennes bioheat equation. A numerical scheme based on the Laplace transform is proposed to solve the bio-heat transfer problem with simultaneous equations.

Keywords: Bio-heat transfer, blood temperature, Laplace transform, Pennes bioheat equation

Introduction

Based on experiment analysis, in1948, Pennes[1] proposed the first constitutive relationship between temperature and the blood flow rate. This relation is popularly known as Pennes' bioheat equation. The equation includes a special term that describes the heat exchange between blood flow and solid tissues. Many researchers used it to deal with various problems. The literature [2, 3] modeled small breast carcinomas surrounded by extended health tissue as a solid sphere and investigated the effect of dose on the temperature distribution. Kuznetsov[4] explored the temperature distribution with a transient thermal dose and investigated the effect of thermal dose accumulation during cooling. Michelea et al. [5] studied how the infusion behavior of magnetic nanofluids affects the thermal response in tissue. Lin et al. [6] numerically studied the bio-heat transfer problem in a bi-layered spherical tissue with blood perfusion and metabolism. Kudryashov and Shilnikov[7] used the Pennes bioheat model to describe he heat transfer in soft tissue during the thermal exposure to low temperature. Ma et al. [8] analyzed the effect of controlling the blood perfusion and temperature into the brain on brain hypothermia.

It is believed that even the applications with the estimated values do not affect explanation of the applicability of the bioheat transfer model [9]. For convenience of analysis, therefore, the above papers [2-8] regarded the blood temperature as a constant. In the heat treatment process, blood perfusion starts up a negative feedback mechanism. The blood temperature undergoes a transient process before onset of equilibrium, and then changes the situation of temperature distribution. In substance, the blood temperature undergoes a transient process for heat exchange between blood and tissue [10]. For more fully exploring the heat transfer behavior of biological tissue, this paper analyzes the bio-heat transfer problem with the non-constant blood temperature. A numerical scheme based on the Laplace transform is proposed to solve the present problem.

Problem Formulation

Energy conservation equation of bioheat transfer described in the Pennes model is

$$-\nabla \cdot \vec{q} + w_b c_b (T_b - T) + q_m + q_r = \rho c \frac{\partial T}{\partial t} \tag{1}$$

Here, ρ, c, and T denote density, specific heat, and temperature of tissue. c_b and w_b are, respectively, the specific heat and perfusion rate of blood. q_m is the metabolic heat generation and q_r is the heat source for spatial heating. T_b is the arterial temperature.

This work considers that the skin surface temperature could be kept constant as the skin contacts with a large steel plate at a high temperature. The assumption that heat flux approaches zero deep in tissue $x = L$ was made. The present work defines the heat transport in the skin with constant physiological parameters as the following equations.

$$\rho C\frac{\partial T}{\partial t} = k\frac{\partial^2 T}{\partial x^2} + W_b\rho_b C_b(T_b - T) + q_m + q_r \tag{2}$$

And then, the he boundary conditions can be written as

$$T(0,t) = T_o \quad \text{and} \quad \frac{\partial T(L,t)}{\partial x} = 0 \tag{3}$$

and the initial conditions

$$T(x,0) = T_{bi}\,,\ \frac{\partial T(x,0)}{\partial t} = 0\,,\ \text{and}\ q(x,0) = 0 \tag{4}$$

where T_{bi} is the initial blood temperature and is specified as 37 ^{o}C.

The blood temperature always was assumed to be constant arterial blood temperature for studying such problems. In substance, the blood temperature undergoes a transient process for heat exchange between blood and tissue. The transient process was defined by [6]

$$\varepsilon\rho_b c_b\frac{\partial T_b}{\partial t} = G(T - T_b) \tag{5}$$

where G is the coupling factor between blood and tissue. ε is a proportional rate.

Analytical Method

Two new variables H and T_B are defined as $H = T - T_{bi}$ and $T_B = T_b - T_{bi}$. Eq. (2) can be rewritten for $q_r = 0$ as

$$\rho c\frac{\partial H}{\partial t} + w_b\rho_b c_b H = k\frac{\partial^2 H}{\partial x^2} + w_b\rho_b c_b T_B + q_m \tag{6}$$

The boundary conditions become

$$H(0,t) = H_0 \tag{7}$$

$$\frac{\partial H(L,t)}{\partial x} = 0 \tag{8}$$

The initial conditions are rewritten as

$$H(x,0)=0\,,\ \frac{\partial H(x,0)}{\partial t}=0\,,\text{ and }\ q(x,0)=0 \tag{9}$$

Subsequently, the use of the Laplace transform technique maps the transient problem into the steady one. The differential equations (5) and (6) and the boundary conditions (7) and (8) are transformed in conjunction with the initial conditions (9) as

$$\frac{d^2\tilde{H}}{dr^2}-\lambda^2\tilde{H}=-f \tag{10}$$

and

$$\tilde{H}(0,s)=H_0/s \tag{11}$$

$$\frac{d\tilde{H}(L,s)}{dx}=0 \tag{12}$$

where

$$\lambda^2=\frac{1}{k}[\rho cs+w_b\rho_b c_b(1-\frac{1}{1+\tau_p s})] \tag{13}$$

$$f=\frac{q_m}{ks} \tag{14}$$

$$\tau_p=\frac{\varepsilon\rho_b c_b}{G} \tag{15}$$

, and s is the Laplace transform parameter for time t.

The present work divides the whole space domain into several sub-space domains. For continuities of heat flux and temperature within the whole space domain, the following conditions are required at the interface of the sub-space domain j-1, $[x_{i-1},x_i]$, and the sub-space domain j, $[x_i,x_{i+1}]$.

$$\tilde{H}_{j-1}(x_i)=\tilde{H}_j(x_i) \qquad i=1,2,\ldots,n\ ;\ j=i \tag{16}$$

$$\frac{d\tilde{H}_{j-1}(x_i)}{dx}=\frac{d\tilde{H}_j(x_i)}{dx} \qquad i=1,2,\ldots,n\ ;\ j=i \tag{17}$$

where the subscript i is the number of node. n is the total number of nodes.

In order to perform the derivation of the governing algebraic equations, $\tilde{H}$ is approximated by using the nodal temperatures and shape function within a small sub-space domain. The shape function in each sub-space domain is derived from the governing equation (10) with the following procedures.

For the sub-space domain j, $[x_i,x_{i+1}]$, the analytical solution of the governing equation (10) subjected to the boundary conditions

$$\tilde{H}_j(x_i) = \tilde{H}_i \quad \text{and} \quad \tilde{H}_j(x_{i+1}) = \tilde{H}_{i+1} \tag{18}$$

are easily obtained and can be written as

$$\tilde{H}_j(x_i) = \tilde{H}_i \quad \text{and} \quad \tilde{H}_j(x_{i+1}) = \tilde{H}_{i+1} \tag{18}$$

$$\tilde{H}_j = \frac{1}{\sinh \lambda\ell}\left\{(\tilde{H}_i - \frac{f}{\lambda^2}x_i)\sinh \lambda(x_{i+1} - x) + (\tilde{H}_{i+1} - \frac{f}{\lambda^2}x_{i+1})\sinh \lambda(x - x_i)\right\} + \frac{f}{\lambda^2} \tag{19}$$

where ℓ denotes the length of sub-space domain or the distance between two neighboring nodes.

Substituting Eq. (16) and the shape function (19) into Eq. (17) and then evaluating the resulting derivative can lead to the discretized form for the interior nodes as following

$$\tilde{H}_{i-1} - 2\cosh(\lambda\ell)\tilde{H}_i + \tilde{H}_{i+1} = \frac{f}{\lambda^2}[2 - 2\cosh(\lambda\ell)] \tag{20}$$

Eq. (20) in conjunction with the discretized forms of the boundary conditions can be rearranged as the following matrix equation

$$[B]\{\tilde{H}\} = \{F\} \tag{21}$$

where $[B]$ is a matrix with complex numbers, $\{\tilde{H}\}$ is a column vector in the Laplace transform domain, and $\{F\}$ is a column vector representing the forcing term. Thereafter, the value of H in the physical domain can be determined with the application of the Gaussian elimination algorithm and the numerical inversion of the Laplace transform [11].

Results and Discussion

Some thermal properties of the sample skin are regarded as w_b = 0.5 kg/m^3· s, k = 0.2 W/m· ^{o}C, ρ = 1000 kg/m^3 and $c = c_b$ = 4200 J/kg· ^{o}C [11]. The distance between the skin and body core is L = 0.01208 m and the value of H_o is specified with 12 ^{o}C [11]. The values of the other parameters are individually determined for each calculation.

The primary premise of the Pennes bioheat equation is that the blood temperature is to be constant arterial blood temperature, and immediately equilibrates(thermally) with the surrounding tissue. In substance, the blood temperature undergoes a transient process before onset of equilibrium, so the assumption $\varepsilon\rho_b c_b \partial T_b/\partial t = G(T - T_b)$ was made [10]. Figure 1 shows the temperature variations at x = 0.01 m for the assumptions, $\varepsilon\rho_b c_b \partial T_b/\partial t = G(T - T_b)$ and T_b = 37^{0}C. It is observed that the cooling function of blood is reduced for the assumption $\varepsilon\rho_b c_b \partial T_b/\partial t = G(T - T_b)$, because the blood temperature would be increased from T_{bi} to the tissue temperature T [12]. The variation rates of temperature shown in Fig. 1(a) and Fig. 1(b) are obviously different. It expresses that the effect of location significantly affects the temperature variation at the measurement point.

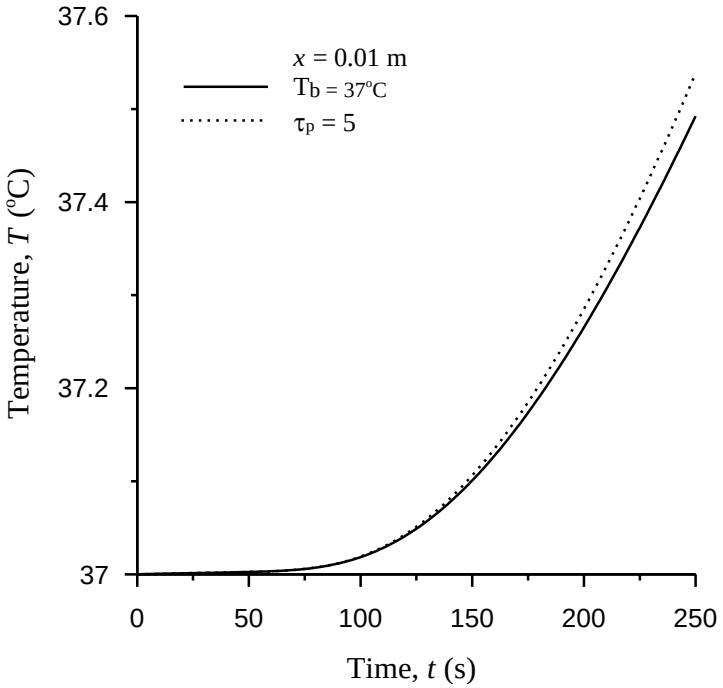


Figure 1. Temperature variations for the assumptions, $\varepsilon\rho_b C_b \partial T_b/\partial t = G(T - T_b)$ and $T_b = 37^0$C, at $x = 0.01$ m.

Figure 2 depicts the temperature variations at $x = 0.00208$ m for $\tau_p = 1$ and $\tau_p = 9$. It is observed from Fig. 1 that the temperature at $x = 0.00208$ m increases with the time. However, Ref. [11] indicated that as the blood temperature is specified as 37 ^{o}C, the temperature at $x = 0.00208$ almost has been in steady state after $t = 250$ *s*. This result implies that the cooling function of blood will be reduced as the blood temperature undergoes a transient process. In this case, the curves of temperature variation for $\tau_p = 1$ and $\tau_p = 9$ are coincident. It implies the effect of τ_p is not obvious in the present case.

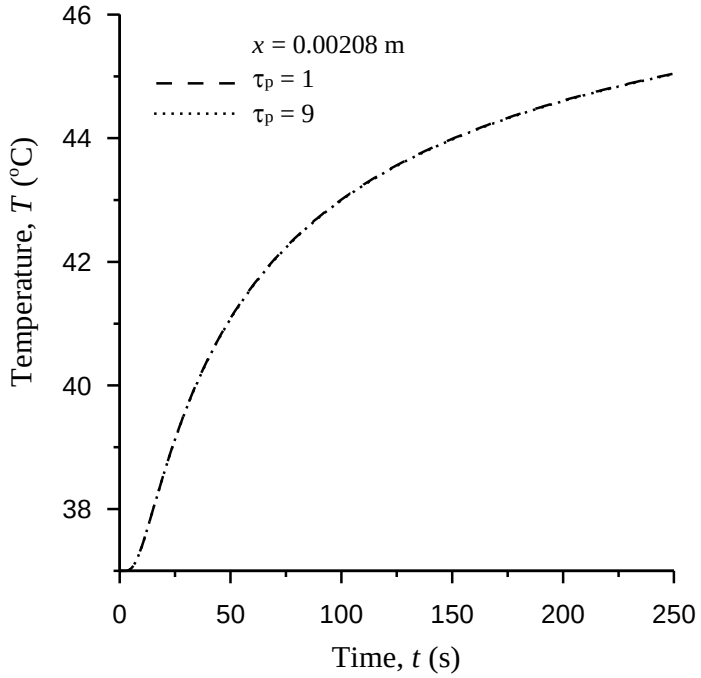


Figure 2. Temperature variations for $\tau_p = 1$ and $\tau_p = 9$ at $x = 0.00208$ m.

In order to further explore the effect of τ_p, the perfusion rate of blood is increased from w_b = 0.5 kg/m^3· s to w_b = 1.0 kg/m^3· s. Figure 3 shows the calculated results with w_b = 1.0 kg/m^3· s for τ_p = 1 and τ_p = 9. Two temperature variation curves are coincident. This phenomenon is same as that shown in Figure 2. It implies that the effect of τ_p is not strengthened with increasing the value of w_b in the present problem. The temperature distributions at t = 50 s and t = 150 s for τ_p = 1 and τ_p = 9 with w_b = 1.0 kg/m^3· s are also presented in Figure 4.

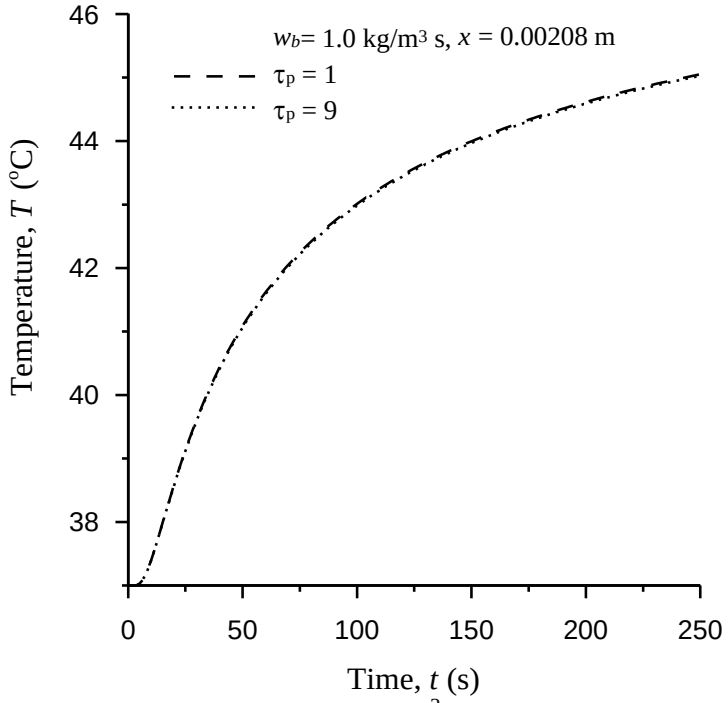


Figure 3. Temperature variations with w_b = 1.0 kg/m^3· s for τ_p = 1 and τ_p = 9 at x = 0.00208 m.

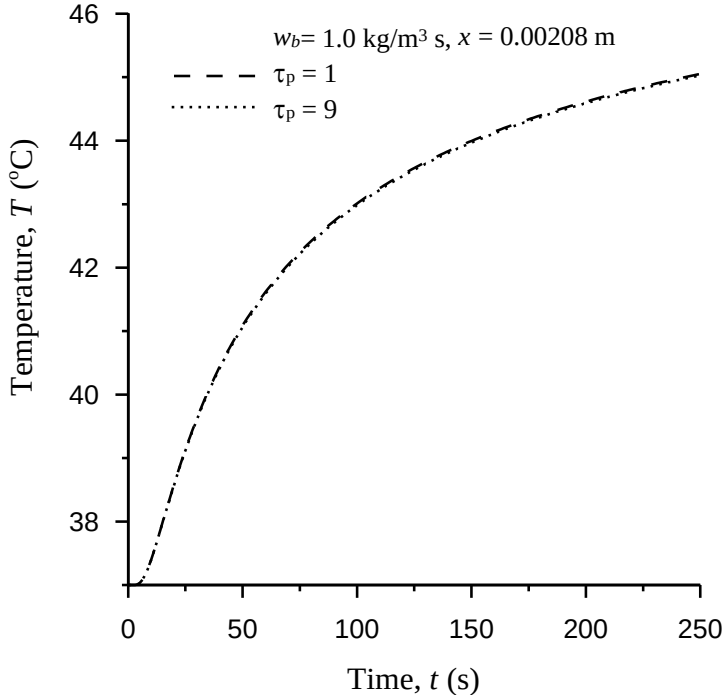


Figure 4. Temperature distributions at t = 50 s and t = 150 s for τ_p = 1 and τ_p = 9 with w_b = 1.0 kg/m^3· s.

Conclusions

A numerical scheme based on the Laplace Transform method is proposed for solving the Pennes bio-heat transfer equation with transient blood temperature. The results without constant blood temperature obviously differ from those with constant blood temperature. The effect of τ_p is not strengthened with increasing the value of w_b in the present problem. The present study depicts that the effects of τ_p and w_b are not obvious under that the blood temperature undergoes a transient process for heat exchange between blood and tissue.

References

[1] Pennes, H.H. (1948) Analysis on tissue arterial blood temperature in the resting human forearm, *J.Appl.Physiol.***1**, 93–122.

[2] Andrä, W., d'Ambly, C.G., Hergt, R., Hilger, I. & Kaiser, W. A.(1999). Temperature distribution as function of time around a small spherical heat source of local magnetic hyperthermia. *J. Magnetism and Magnetic Materials* **194**, 197-203.

[3] Shimizu, T. & Matsui, M.(2003). New magnetic implant material for interstitial hyperthermia. *Science and Technology of Advanced Materials*, **4**,469–473.

[4] Kuznetsov, A.V.(2006). Optimization problem for bioheat equation. *Int. Commun. Heat Mass Transfer* **33**, 537-543.

[5] Michelea, F. D., Pizzichelli, G.., Mazzolai, B. & Sinibaldi, E.(2015). On the preliminary design of hyperthermia treatments based on infusion and heating of magnetic nanofluids. *Mathematical Biosciences* **262**, 105–116.

[6] Lin, C.T. & Liu, K.C.(2009). Estimation for the heating effect of magnetic nanoparticles in perfused tissues. I*nternational Communications in Heat and Mass Transfer* **36**, 241-244.

[7] Kudryashov, N. A. and Shilnikov, K. E. (2015). Numerical modeling and optimization of the cryosurgery operations, *Journal of Computational and Applied Mathematics* **290**, 259 – 267.

[8] Ma, W., Liu, W., Li, M. (2016). Analytical heat transfer model for targeted brain hypothermia, *International Journal of Thermal Sciences* **100**, 66-74.

[9] Nakayama, A. and Kuwahara, F. (2008). A general bioheat transfer model based on the theory of porous media, *International Journal of Heat and Mass Transfer* **51**, 3190–3199.

[10] Zhang, Y. (2009). Generalized dual-phase lag bioheat transfer equations based on nonequilibrium heat transfer in living biological tissues. *International Journal of Heat and Mass Transfer* **52**, 4829-4834.

[11] Liu, K.C.(2008). Thermal propagation analysis for living tissue with surface heating. *International Journal of Thermal Sciences* **47**, 507-513.

[12] Liu, K.C. and Chen, Y.S. (2016). Analysis of heat transfer and burn damage in a laser irradiated living tissue with the generalized dual-phase-lag model, *International Journal of Thermal Sciences* **103**, 1-9.

3D meso-scale fracture modelling and validation of concrete based on in-situ X-ray CT images and cohesive crack model

***W.Y. Ren[1,2], †Z.J. Yang[3], R. Sharma[4], S.A. McDonald[5], P.M. Mummery[6]**

[1]College of Water Resources and Architectural Engineering, Northwest A&F University, China
[2] Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, China
[3]College of Civil Engineering and Architecture, Zhejiang University, China
[4]School of Engineering, Indian Institute of Technology Mandi, India
[5]Manchester X-ray Imaging Facility, School of Materials, the University of Manchester, UK
[6]School of Mechanical, Aerospace and Civil Engineering, the University of Manchester, UK

*Presenting author: wenyuange304@nwsuaf.edu.cn
†Corresponding author: zhjyang@zju.edu.cn

Abstract

A meso-scale finite element modelling technique is developed for simulation of complicated 3D fracture in multiphase concrete, based on realistic X-ray computed tomography (XCT) images and the cohesive crack model. It is validated against in-situ XCT tests of a concrete cube under Brazilian-like compression in terms of load-displacement curves and crack patterns. Meso-scale simulations under uniaxial compression and tension are also carried out.

Keywords: Concrete; Meso-scale fracture; X-ray computed tomography; Finite element simulation; Cohesive crack model

Introduction

As a quasi-brittle multiphasic composite, concrete has been widely used in many civil and industrial structures. Due to the random distribution of multiphases, i.e. cement/mortar, aggregates and pores, it exhibits heterogeneous mechanical properties in micro/meso-scales, which in turn determine the performances and reliability of structures at macro-scale. Therefore, understanding its micro and meso-scale mechanical behaviour, including damage and fracture, becomes an important and challenging engineering and scientific problem [1]-[3]. In particular, developing predictive numerical models capable of simulating realistic fracture propagation in three-dimensional (3D) multiphasic composites like concrete is still challenging due to difficulties in experimental characterisation as well as numerical models [4]. To date, most of the numerical models available in literature use either assumed micro/meso-scale morphologies [5] or assumed random field properties [3][6]. As such, the simulated fracture processes cannot be directly and accurately validated.

The X-ray Computed Tomography (XCT) technique is now increasingly used for characterisation of micro-structures of composite materials. Recently, we carried out in-situ micro XCT tests of concrete cubes under progressive compressive loading to elucidate the complicated 3D fracture process in concrete [7]. By directly converting high-resolution XCT images into finite element (FE) elements, realistic 3D and 2D models have been developed for complicated multi-cracking, using the continuum damage plasticity model [8] and the discrete cohesive crack model (CCM) [9], respectively, with very promising results. In this study, we further develop the XCT-image based CCM approach to 3D so that both the load-carrying capacity and discrete 3D fracture evolution can be directly validated by the in-situ XCT test.

Construction of 3D FE meshes

The in-situ XCT tests of 40mm concrete cubes under compression were conducted with a voxel resolution of 37.2μm at the University of Manchester and the details were reported in [7]. The 3D XCT images were cropped to 37.2mm, re-sampled to 0.1mm resolution and segmented into digital models, with one example shown in Fig.1. The digital models were then converted to 3D FE meshes using AVIZO [10] and Simpleware [11]. A generated mesh from Fig. 1 is shown in Fig. 2 with 4,422,638 tetrahedral elements and 5,932,268 nodes using targeted minimum edge length 0.1mm and maximum edge length 0.5mm. The mesh information was then exported into an ABAQUS input file. Zero-thickness 6-noded cohesive interface elements (COH3D6) were then pre-inserted into the 3D mesh to simulate complex fracture processes using an in-house MATLAB code augmented from an original one for homogeneous materials [3] by accounting for multi-phases and interfaces.

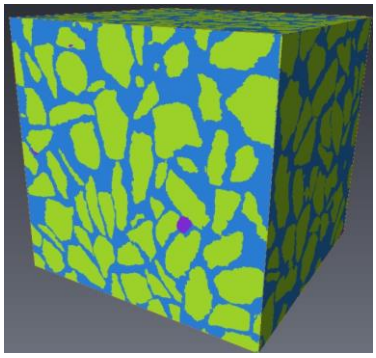


Figure 1. Segmented 3D image model

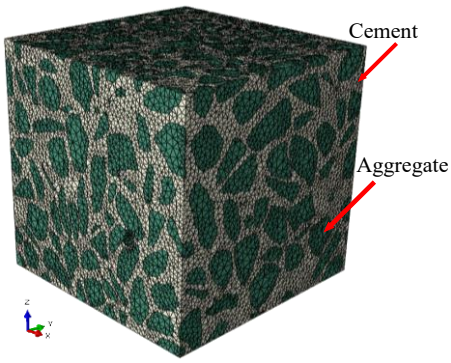


Figure 2. Generated 3D FE mesh

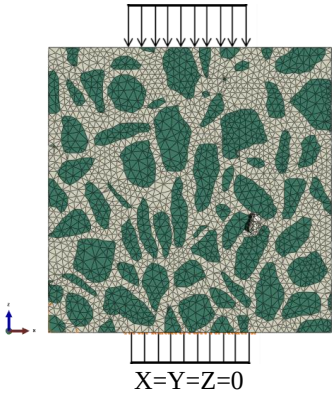


Figure 3. Boundary conditions

There are three sets of CIEs, namely, CIE_AGG within the aggregates, CIE_CEM within the cement paste, and CIE_INT on the aggregate-cement interfaces. In total, 2,939,571 CIEs were inserted into the mesh in Fig. 2. The solid elements of aggregates and cement were assumed to behave linear elastically. The Young's moduli (E) of aggregates and cement are 51 GPa and 13.6 GPa, respectively, measured by micro-indentation tests [7]. The density of aggregate and cement are 2500 kg/m^3 and 2200 kg/m^3, respectively. The cohesive strength and fracture energy for interfaces and cement are 3 MPa and 6 MPa, 0.03 N/mm and 0.06 N/mm, respectively. The shear fracture energy was assumed as 10 times the value in the normal direction. The linear tension/shear softening laws were used to model CIEs, with the quadratic nominal stress initiation criterion, energy based damage evolution, and mixed-mode BK-law fracture energy criterion.

The displacement-controlled loading scheme was used for all the simulations using the ABAQUS/Explicit solver with a step time of 0.01s, which was found to be sufficiently long to ensure the quasi-static loading condition. The modelled boundary conditions are shown in Fig. 3, where only a 17.5mm central square on the top and bottom surfaces was constrained respectively, as in the tests.

Validation against in-situ XCT tests

Fig. 4 compares satisfactorily the simulated force-displacement (F-d) curve with the in-situ test data that were calculated from a digital volume correlation (DVC) analysis of XCT images [7]. In particular, the predicted peak load 16.3kN is very close to the test value of

16.5kN. The post-peak softening stage which cannot be obtained from the test was also well captured.

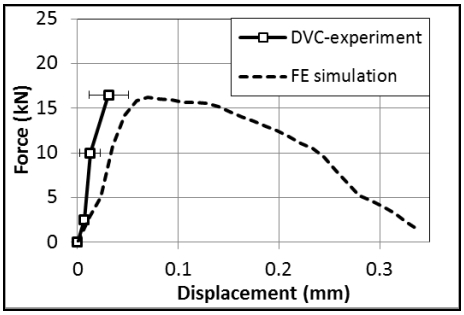
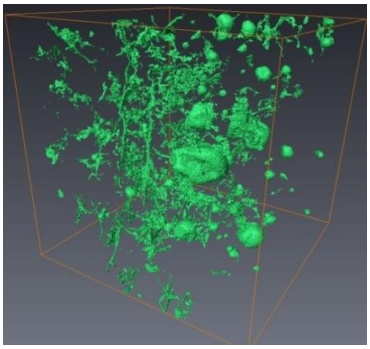
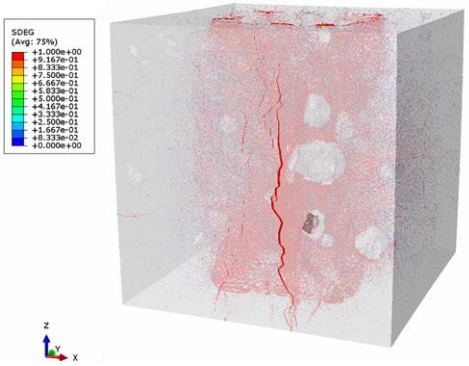


Figure 4. Comparison of *F-d* curves of XCT test and 3D FE simulation

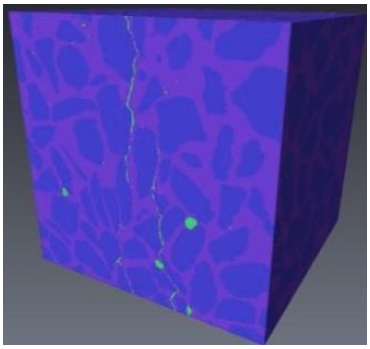
Fig. 5 compares the cracks (pores are also shown) in the segmented XCT image and the simulated cracks represented by CIEs with high damage index (*SDEG*>0.99) at the peak load, with Figs 5a and b showing the internal and Figs. 5c and d the external, respectively. A qualitative similarity can be seen, especially the crack paths on the surfaces.



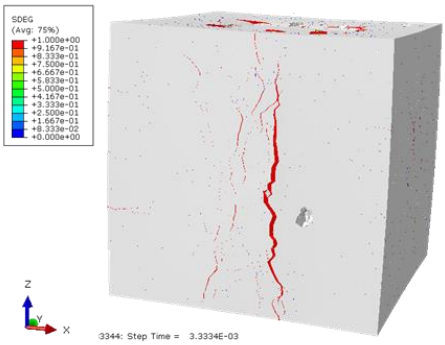
(a) Cracks and pores from in-situ test



(b) Pores with predicted cracks (*DSF*=10)



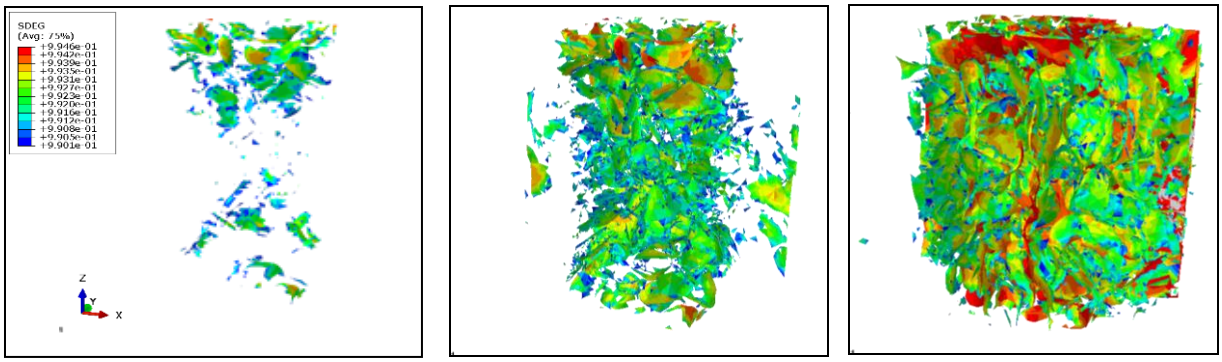
(c) XCT segmentation



(d) Predicted crack path (*DSF*=20)

Figure 5. Comparison of cracks at peak load (*DSF*=deformation scale factor)

Fig. 6 shows the detailed internal cracking process. It is clearly indicated that the damage initiated near the loaded surfaces, then propagated towards the centre in the loading direction, and finally expanded in the horizontal direction.

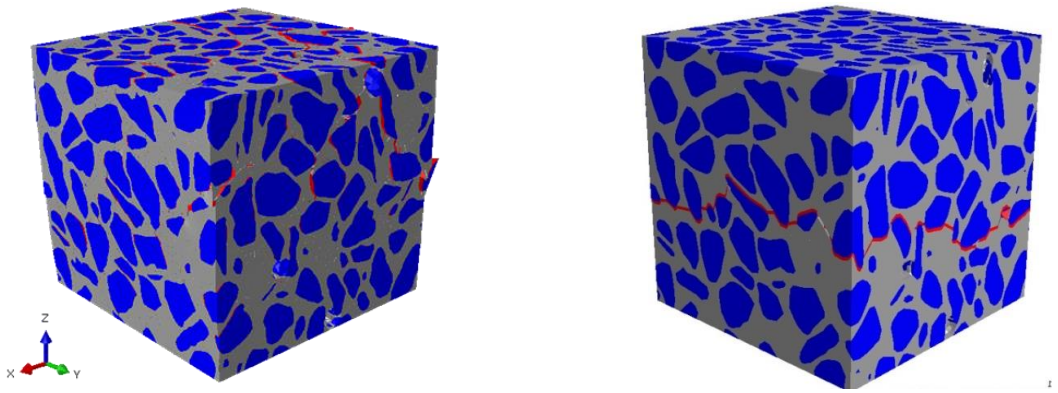


(a) d=0.020mm (b) d=0.033mm (c) d=0.073mm

Figure 6. Crack propagation process (*DSF*=10, *SDEG*>0.99)

Simulations of uniaxial compression and tension

The FE model in Fig. 3 was also simulated under the uniaxial compression and tension loading conditions. Fig. 7a shows that the distributed crack pattern under uniaxial compression is quite different from that under the previous Brazilian-like compressive loading case. In Fig. 7b, the 3D crack surfaces under uniaxial tension are linked to form a major crack path.



(a) uniaxial compression (b) uniaxial tension

Figure 7. Simulated crack patterns (*DSF*=10)

The predicted stress-strain curves from the 2D statistical modelling results [9] and the 3D simulation under uniaxial tension are compared in Fig. 8. It can be seen that the results in the initial elastic stage and the final cracked stage are close between 2D and 3D cases, but the predicted tensile strength of 5.0MPa in 3D is much higher than the mean value of 3.3MPa in 2D. The main reason is that the formation of 3D cracks needs higher fracture energy than in 2D, due to the resistance from randomly distributed aggregates and the larger crack surfaces providing higher normal tractions across the cracks in 3D [6].

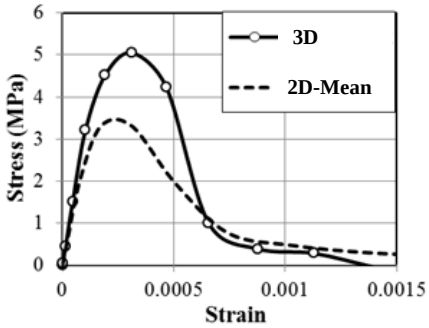


Figure 8. Stress-strain curves predicted by 3D simulation and 2D statistical analysis

Conclusions

In this study, a 3D meso-scale XCT-image based cohesive crack modelling technique is developed for complex fracture analysis of concrete. Direct comparisons of FE simulation results with in-situ XCT tests have been made with good quantitative agreement between the peak forces and satisfactory qualitative agreement between the crack patterns. The 3D simulation of uniaxial tension loading predicts a higher strength than the mean value of 2D statistical results. Although further works are needed, the combination of the in-situ XCT tests and image-based cohesive crack modelling proves very promising in studying the complicated 3D fracture mechanism in quasi-brittle composite materials like concrete.

Acknowledgements

The research is funded by the UK EPSRC grants (No. EP/J019763/1), the Fundamental Research Funds for the Central Universities in China (No. Z109021614 and 172210173), and Zhejiang Provincial Natural Science Foundation of China (No. LR14E080002).

References

[1] J.T. Oden, T. Belytschko, I. Babuska, T.J.R. Hughes, (2003) Research directions in computational mechanics, *Computer Methods in Applied Mechanics and Engineering*, **192**, 913-922.

[2] M.E. Kassner, S. Nemat-Nasser, Z. Suo, G. Bao, J.C. Barbour, L.C. Brinson, H. Espinosa, H. Gao, S. Granick, P. Gumbsch, K.-S. Kim, W. Knauss, L. Kubin, J. Langer, B.C. Larson, L. Mahadevan, A. Majumdar, S. Torquato, F. van Swol, (2005) New directions in mechanics, *Mechanics of Materials*, **37,** 231-259.

[3] Z.J. Yang, X.T. Su, J.F. Chen, G.H. Liu, (2009) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials, *International Journal of Solids and Structures*, **46,** 3222-3234.

[4] T.T. Nguyen, J. Yvonnet, M. Bornert, C. Chateau, (2016) Initiation and propagation of complex 3D networks of cracks in heterogeneous quasi-brittle materials: Direct comparison between in situ testing-microCT experiments and phase field simulations, *Journal of the Mechanics and Physics of Solids*, **95,** 320-350.

[5] X.F. Wang, Z.J. Yang, J.R. Yates, A.P. Jivkov, C. Zhang, (2015) Monte Carlo simulations of mesoscale fracture modelling of concrete with random aggregates and pores, *Construction and Building Materials*, **75,** 35-45.

[6] X.T. Su, Z.J. Yang, G.H. Liu, (2010) Monte Carlo simulation of complex cohesive fracture in random heterogeneous quasi-brittle materials: A 3D study, *International Journal of Solids and Structures*, **47,** 2336-2345.

[7] Z.J. Yang, W.Y. Ren, R. Sharma, S. McDonald, M. Mostafavi, Y. Vertyagina, T.J. Marrow, (2017) In-situ X-ray computed tomography characterisation of 3D fracture evolution and image-based numerical homogenisation of concrete, *Cement and Concrete Composites*, **75,** 74-83.

[8] Y.J. Huang, Z.J. Yang, W.Y. Ren, G.H. Liuand Ch Zhang (2015). 3D Meso-scale fracture modelling and validation of concrete based on in-situ X-ray computed tomography images using damage plasticity model, *International Journal of Solids and Structures*, **67-68**, 340-352.

[9] W.Y. Ren, Z.J. Yang, R. Sharma, C. Zhang, P.J. Withers, (2015) Two-dimensional X-ray CT image based meso-scale fracture modelling of concrete, *Engineering Fracture Mechanics*, **133,** 24-39.

[10] AVIZO, AVIZO User's Guide, 2013.

[11] Simpleware, ScanIP, +FE and +CAD Version 4.3 Reference Guide, Simpleware Ltd. , Exeter, UK, 2011.

Edge Effect on Eddy Current Detection for Subsurface Defects in Titanium Alloys

*Yibo Wang, Qian Bai, Wei Du, and †Bi Zhang

Key Laboratory for Precision and Non-traditional Machining Technology of Ministry of Education, Dalian University of Technology, China.

*Presenting author: 344729677@qq.com
†Corresponding author: 15140572053@qq.com

Abstract

Additive/Subtractive Hybrid Manufacturing (ASHM) offers an opportunity for on-line detection due to its layer by layer processes. Eddy current detection (ECD) is a well-established non-destructive method for inspection of surface/subsurface defects in electrically conductive materials. The ECD-enabled ASHM is a promising manufacturing technology to fabricate the parts with good surface condition and material quality. However, edge effect, as an inherent characteristic of ECD, limits the industrial application of ECD. In this study, a specimen with subsurface defects located in the edge area is fabricated. The effects of the edge on ECD signals from subsurface defects are investigated. An effective scanning path is proposed to eliminate the edge effect. A Finite Element (FE) model of ECD is established to analyze the edge effect on ECD. This study provides guidance to determination of the ECD parameters.

Keywords: Eddy current detection; Edge effect; Finite Element; Defects; Reactance

1. Introduction

Additive Manufacturing (AM) offers a great advantage of building parts with geometric complexities. However, defects, e.g. porosities, incomplete fusion holes and cracks, are inevitably introduced if process parameters are improperly chosen. In addition, the current AM methods provide a relatively poor surface finish as well as low dimensional accuracies due to the unstable melt pools. Additive/Subtractive Hybrid Manufacturing, integrating cutting processes into an additive manufacturing process to take advantage of both the simplex AM and subtractive processes, paves a way to enhancing dimensional accuracies as well as removing defects [1]. In laser-based AM processes, optical inspection methods have been investigated by several research groups [2, 3]. The optical methods can detect the abnormal melting pool straightforwardly, however, they provide only limited information on thermal stability and melt pool dimensions during the AM process, and then infer the existence of potential defects indirectly [4].

Eddy Current Detection (ECD), as a non-destructive technique, is widely used to detect surface/subsurface defects of metallic parts, and has a potential application in ASHM. Eddy currents are excited in a test-piece by a magnetic field at a certain excitation frequency [5]. If a defect exists, it perturbs the distribution of the induced eddy currents in the test-piece, and thus manifests localized magnetic anomalies. Therefore, a defect is detected by monitoring the impedance variation of the excitation coil. ECD can avoid the usage of a couplant. In addition, ECD permits high-speed inspection and can be used at elevated temperatures. With portable devices and instantaneous results of ECD, it can facilitate automated in-process inspection. With the limitation of the skin effect and edge effect, ECD is generally applied to detecting defects located away from the edges of a specimen. Besides, the specimen should

have a good surface finish for enhanced detection accuracies [6-10].

Edge effect is an inherent characteristic of ECD. When a detection probe is placed near the edge of a specimen, eddy current flow is perturbed by the edge, which generates mixed signals with that of the defects [10]. The edge effect of ECD has been studied by some investigators. Theodoulidis et al. proposed an analytical model to calculate the electromagnetic field of a cylindrical coil in the edge of a metal block [11]. Bowler et al. developed an expansion method using the truncated region eigenfunction to determine the tangential magnetic field and calculate the impedance of a coil near the edge of a metal plate [12]. Rao et al. combined a multi-frequency eddy current tester with an artificial neural network to eliminate the disturbing variables including the edge effect. They could evaluate the depth of the surface-breaking notches in austenitic stainless steel welds with a maximum deviation of 0.08 mm [13]. He et al. used pulsed the technique of eddy current imaging and frequency spectrum analysis to detect and evaluate the subsurface defects near the edge of a specimen [14]. However, fewer studies are found to cover different detection scanning paths for subsurface defects near the edge.

In order to detect the defects near the edge of a specimen, this paper investigates the effect of edge and defect for ECD signals. The study proposes a method for establishing the effective ECD scanning path to suppress the edge effect and validates the proposed method with ECD experiments and Finite Element (FE) simulations.

2. Experimental setup

2.1 Specimen preparation

In this study, a titanium alloy Ti-6Al-4V specimen with three artificial subsurface defects was fabricated for the ECD experiments, as shown in Fig. 1(a). In order to suppress noise signal of the rough surface, both top and side surfaces were machined. The dimension of the specimen was 40 mm in length, 12 mm in width and 12 mm in height. The subsurface defects are named as Hole #1, Hole #2 and Hole #3, which were prepared by drilling. The diameters of three holes were 2 mm, 1.5 mm and 1 mm, respectively. The depths of all the holes were 2.5 mm from the side surface, and 0.5 mm from the top surface. A coordinate system is built for the convenience to record and analyze data as shown in Fig. 1(b). On the top surface, the *x* axis coincides with the horizontal centerline of the specimen, and the *y* axis coincides with the centerline of Hole #3.

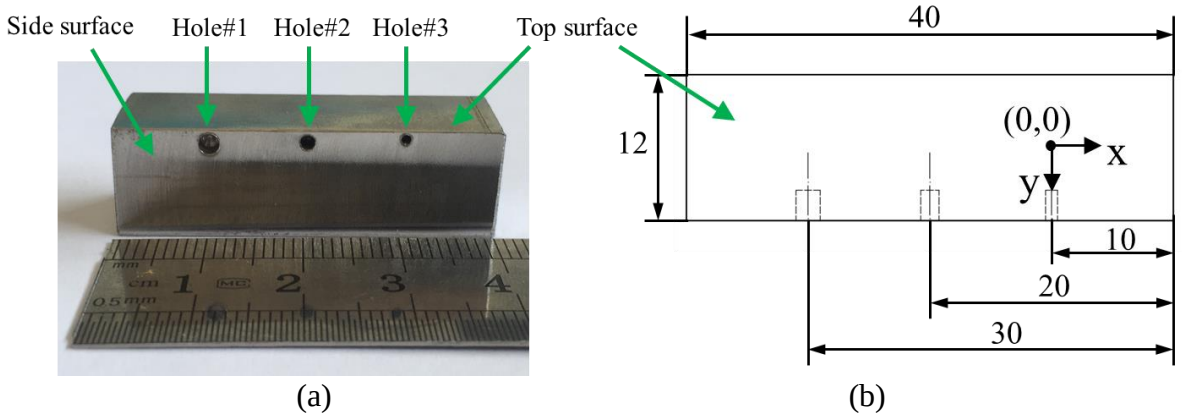
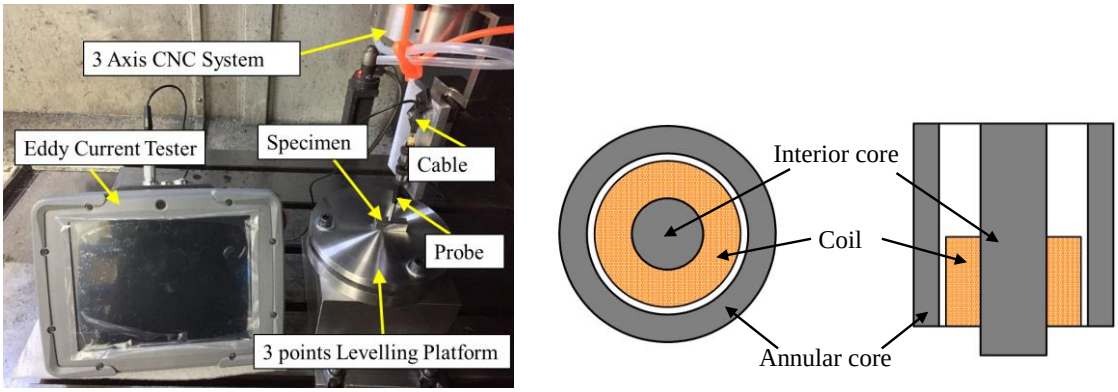


Fig. 1 (a) Specimen with artificial subsurface defects; (b) Schematic of the specimen in top view. (Unit: mm)

2.2 Experimental setup for eddy current detection

The experimental setup in Fig. 2(a) consists of an eddy current detection setup, a three-axis CNC system, and a three-point levelling platform. In Fig. 2(b), an absolute probe consisting of a copper coil, an interior ferrite core, and an annular ferrite core was used in the experiment. This configuration is able to concentrate magnetic lines around the probe for higher sensitivity and larger depth of detection [15]. The probe has a characteristic inductance L_0=119.96 μH and resistance R_0=5.6 Ω, both in air. Detailed parameters of the probe are listed in Table 1. In ECD, the probe was mounted on a three-axis motion table and moved along the *x* and *y* directions. This probe was connected to the eddy current detector to capture electromagnetic inspection data. Signals obtained from a defect-free section of the specimen were set to zero by automatically adjusting the output offset. Detection signals were transmitted to PC for further processing. Detection parameters used in this paper are presented in Table 2.



(a) (b)

Fig. 2 (a) ECD setup and (b) Schematic of probe configuration.

Table 1 Characteristics of the ECD probe

	Inner radius (mm)	Outer radius (mm)	Height (mm)	Turns
Coil	0.80	1.26	1.10	110
Interior core	-	0.76	4.32	-
Annular core	1.40	1.84	3.80	-

Table 2 Detection parameters for ECD

	Lift-off distance (mm)	Scanning speed (mm/min)	Excitation frequency (kHz)	Gain (dB)	Drive voltage (V)
Value	0.05	400	100	60	2.0

3. FE modeling for ECD

In order to study the edge effect on ECD signals, an FE model was established in the ANSYS Maxwell software. Eddy current was calculated based on Maxwell equations, which are defined as follows:

$$\nabla \times E = -\frac{\partial B}{\partial t}$$
$$\nabla \times H = J + \frac{\partial D}{\partial t}$$
$$\nabla \bullet D = \rho$$
$$\nabla \bullet B = 0$$

where ***E*** represents the electric field intensity; ***B*** the magnetic induction; ***J*** the electric current density; ***D*** the electric displacement; and $\boldsymbol{\rho}$ the volume charge density.

As shown in Figs. 3(a) and (b), the dimension of the defect in the FE model were set in accordance with the Hole #3. Materials of the coil and specimen were set as copper and Ti-6Al-4V, respectively, interior and annular cores were set as ferrite. The solution type was eddy current, and the excitation frequency was set as 100 kHz. Moreover, the simulation used the natural boundary condition as the default boundary condition, and the solution region was set as 500%. The probe model was set as inside subdivision, the specimen model was set as inside subdivision and skin depth subdivision [16]. The whole model had a mesh number of approximately 88,000. A sample result of the mesh operations is shown in Fig. 3(a).

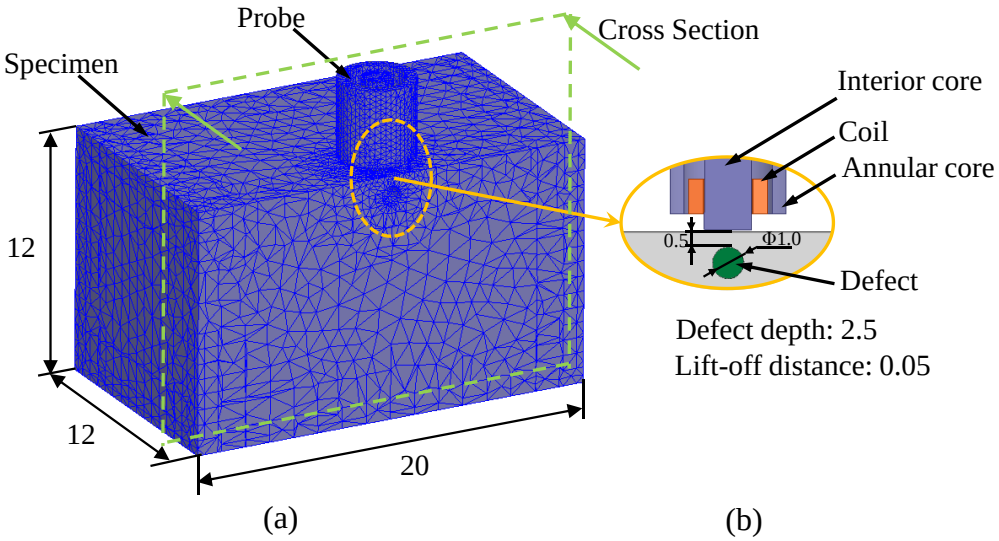


Fig. 3 (a) Meshing of the FE model. (b) Schematics of the cross section of the probe and the defect. (Unit in mm)

4. Results and discussion

4.1 Experiment results

As shown in Fig. 4(a), in top surface, Path 1 passes through the defect-free area between Hole #2 and Hole #3 along the *y* direction. Path 2 coincides with the centerline of Hole #3 along the *y* direction. ECD probe was moved along Path 1 and Path 2 in the *y* direction and the lift-off distance was 0.05 mm.

The reactance signals of Path 1 and Path 2 are shown in Figs. 4(b) and (c). It was difficult to identify the signal of Hole #3 as they were almost similar. Based on the reactance signal of Path 1, the boundary of the edge effect was determined, 3.5 mm from the edge. In Fig. 4(a), Path 3 is along the *x* direction, and the distance from Path 3 to the edge is 2 mm, which indicates that Path 3 locates in the edge effect area. The reactance signal of Path 3 is shown in

Fig. 4(d). The signal was strong enough to determine the defect locations. For these three defects, peak signals appear corresponding to the defect locations. Furthermore, the signal intensity increases with an increase in the defect size.

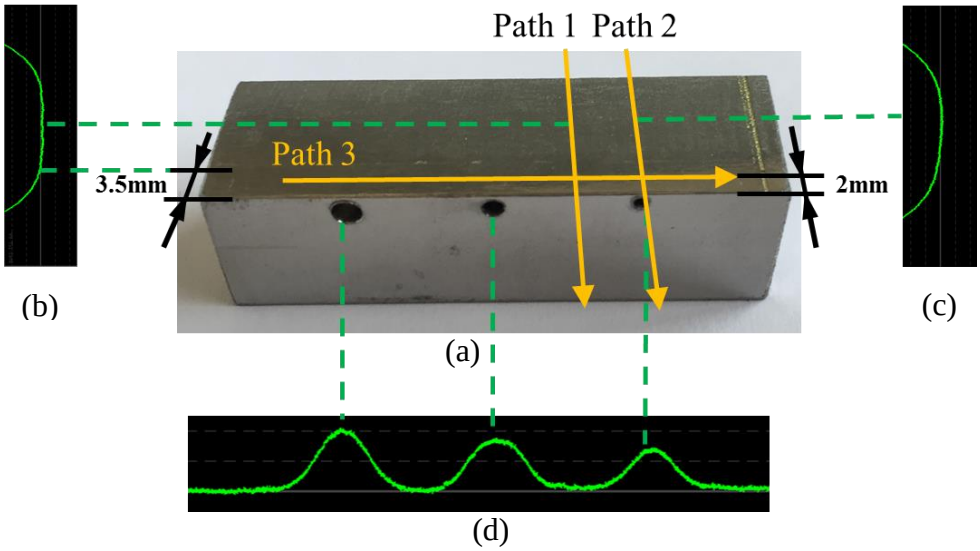


Fig. 4 (a) Specimen with subsurface defects; Screenshot of ECD reactance signal of (b) Path 1; (c) Path 2 and (d) Path 3.

4.2. FE simulation results

Figure 5 shows the normalized signals of reactance and resistance from the FE simulation along Path 1 and Path 2. The normalized factor X_0 and R_0 were the reactance and resistance when the probe was placed at the original point (0, 0). The solid lines and dashed lines represent the signals of Path 1 and Path 2, respectively. The horizontal axis represents the probe position in the *y* direction. Signals of two paths in each figure are very similar, which is in accordance with Figs. 4(b) and (c).

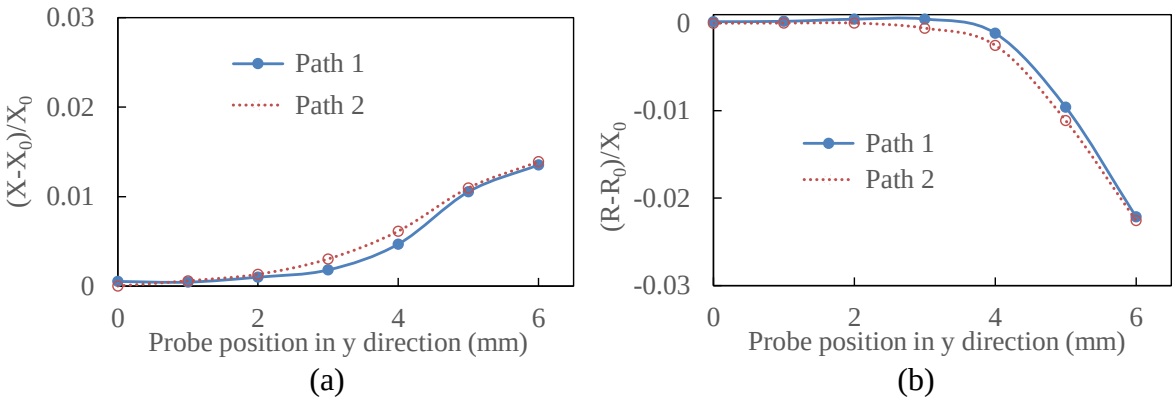
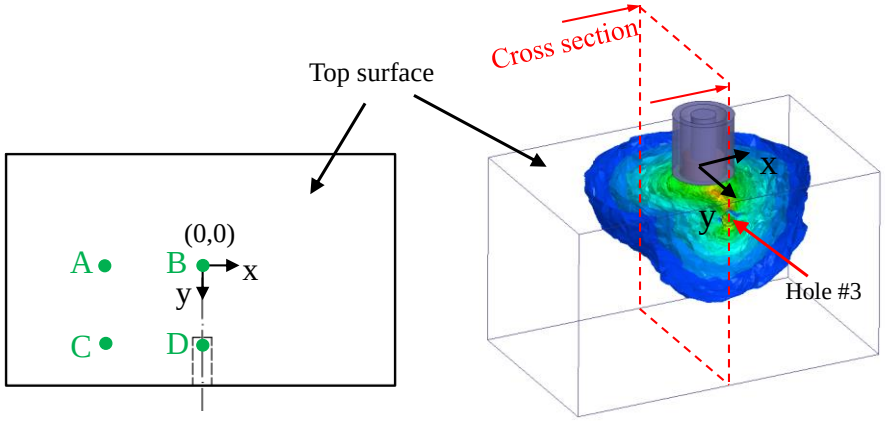


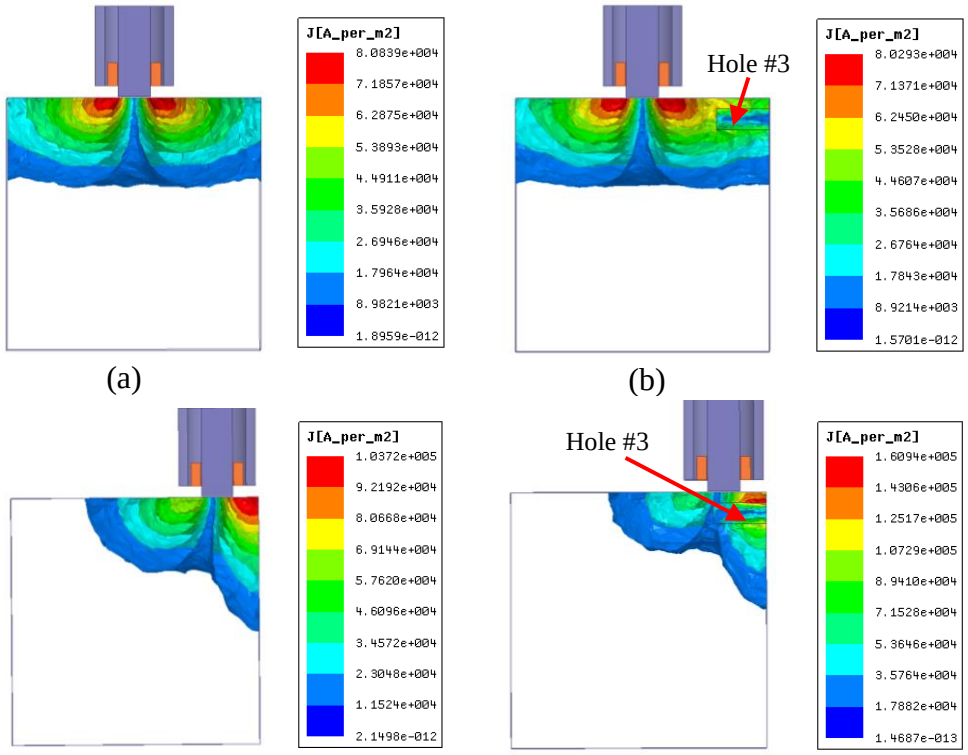
Fig. 5 Comparison of the normalized ECD Signals along Path 1 (solid lines) and Path 2 (dashed lines) simulated by FE (a) Reactance, (b) Resistance.

In order to analyze variations in the eddy current distributions along Path 1 and Path 2 shown in Fig. 6(a), Points A and C were chosen in Path 1, and Points B and D in Path 2. Corresponding to the original point defined in Fig. 1(b), the coordinates of Points A, B, C and D are (-5, 0), (0, 0), (-5, 4) and (0, 4), respectively. Fig. 6(b) shows contours of the eddy current distributions in the FE model.



(a) (b)

Fig. 6 (a) Locations of four points for case study; (b) Contours of eddy current distributions in FE model.



(a) (b)

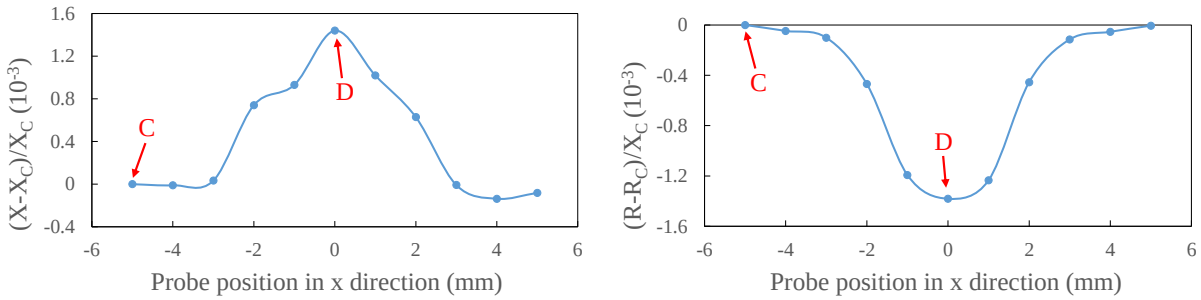
(c) (d)

Fig. 7 Eddy current distributions in cross section with probe on (a) Point A, (b) Point B, (c) Point C, (d) Point D.

The eddy current distributions for Points A, B, C and D are shown in Figs. 7(a), (b), (c) and (d), respectively. Comparing the eddy current distributions of Points A and C, it is found that when the probe approached the edge above the defect-free area along Path 1, eddy current distribution was only perturbed by the edge effect which influences the ECD signals correspondingly. When the probe approached the edge above Hole #3 along Path 2, as shown in Figs. 7(b) and (d), eddy current distribution was perturbed both by the edge and defect Hole #3. From the results of both the experiment (Figs. 4(b) and (c)) and the FE simulations (Figs. 5(a) and (b)), the signals of Path 2 (the path through defect Hole #3) were similar to those of Path 1 (the defect-free path). Therefore, the influence of the defect is less significant than that

of the edge effect. It is believed that if defects exist in the edge area of the specimen, the signal of defects is difficult to distinguish by ECD when the probe is moved from the central area to the edge area.

Figure 8 shows the normalized signals of reactance and resistance simulated by FE across Hole #3 along Path 3. The normalized factor X_C and R_C were the reactance and resistance when the probe was at Point C. Variations in the signal was observed, which is in accordance with the experimental result. Comparing the eddy current distributions of Point C and D (both located on Path 3), with the existence of Hole #3, the eddy current distribution of Point D was obviously different from that of Point C (as shown in Figs. 7(c) and (d)), which caused the variations in the signal. It is believed that when the probe was moved along Path 3, the edge effect on eddy current distributions can be considered as constant, and the influence of Hole #3 on eddy current distributions could be characterized by the signal variations. Therefore, the parallel motion of the probe along the edge of the specimen enabled ECD to detect defects in the edge area.



(a) (b)

Fig. 8 (a) Normalized reactance and (b) normalized resistance of the ECD Signals from FE simulation across Hole #3 along Path 3.

5. Conclusions

In order to detect the defects near the edge of a specimen, investigations are carried out to reveal the edge effect on ECD signals via experiments and FE simulation. It is found that the signal of defects is difficult to distinguish when the ECD probe is moved from the central area to the edge area. A scanning path parallel with the edge enables ECD to effectively detect the defects near the edge. The peak value of an ECD signal increases with an increase in the defect size even in the edge area.

Further study is recommended to exploiting the capability of the ECD method for ASHM. The study should focus on complex ASHM parts that contain defects of small sizes and different types (e.g., cracks, pores, inclusions, and unmelt powders) in a high temperature environment.

Acknowledgments

Financial supports from the National Natural Science Foundation of China (51605077), the Science Challenge Project (JCKY2016212A506-0101) and the Science Fund for Creative Research Groups of NSFC (51621064) are gratefully acknowledged.

References

[1] Du, W., Bai, Q., & Zhang, B. (2016). A Novel Method for Additive/Subtractive Hybrid Manufacturing of Metallic Parts. *Procedia Manufacturing*, *5*, 1018-1030.

[2] Doubenskaia, M., Pavlov, M., & Chivel, Y. (2010). Optical system for on-line monitoring and temperature control in selective laser melting technology. In *Key Engineering Materials* (Vol. 437, pp. 458-461). Trans Tech Publications.
[3] Wang, L., Felicelli, S. D., & Craig, J. E. (2007, August). Thermal modeling and experimental validation in the LENS TM process. In *18th Solid Freeform Fabrication Symposium. Austin, TX* (pp. 100-111).
[4] Kanko, J. A., Sibley, A. P., & Fraser, J. M. (2016). In situ morphology-based defect detection of selective laser melting through inline coherent imaging. *Journal of Materials Processing Technology*, *231*, 488-500.
[5] García-Martín, J., Gómez-Gil, J., & Vázquez-Sánchez, E. (2011). Non-destructive techniques based on eddy current testing. *Sensors*, *11*(3), 2525-2565.
[6] Janousek, L., Chen, Z., Yusa, N., & Miya, K. (2005). Excitation with phase shifted fields-enhancing evaluation of deep cracks in eddy-current testing. *NDT & E International*, *38*(6), 508-515.
[7] Mizukami, K., Mizutani, Y., Todoroki, A., & Suzuki, Y. (2015). Design of eddy current-based dielectric constant meter for defect detection in glass fiber reinforced plastics. *NDT & E International*, *74*, 24-32.
[8] Vacher, F., Alves, F., & Gilles-Pascaud, C. (2007). Eddy current nondestructive testing with giant magneto-impedance sensor. *NDT & E International*, *40*(6), 439-442.
[9] Tian, G. Y., & Sophian, A. (2005). Reduction of lift-off effects for pulsed eddy current NDT. *NDT & E International*, *38*(4), 319-324.
[10] Xu, P., Huang, S., & Zhao, W. (2011). A new differential eddy current testing sensor used for detecting crack extension direction. *NDT & E International*, *44*(4), 339-343.
[11] Theodoulidis, T., & Bowler, J. R. (2010). Interaction of an eddy-current coil with a right-angled conductive wedge. *IEEE Transactions on Magnetics*, *46*(4), 1034-1042.
[12] Bowler, J. R., & Theodoulidis, T. P. (2006). Coil impedance variation due to induced current at the edge of a conductive plate. *Journal of Physics D: Applied Physics*, *39*(13), 2862.
[13] Rao, B. P. C., Raj, B., Jayakumar, T., & Kalyanasundaram, P. (2002). An artificial neural network for eddy current testing of austenitic stainless steel welds. *NDT & E International*, *35*(6), 393-398.
[14] Theodoulidis, T. (2004). End effect modelling in eddy current tube testing with bobbin coils. *International Journal of Applied Electromagnetics and Mechanics*, *19*(1-4), 207-212.
[15] Liu, B., Huang, P., Zeng, X., & Li, Z. (2017). Hidden defect recognition based on the improved ensemble empirical decomposition method and pulsed eddy current testing. *NDT & E International*, *86*, 175-185.
[16] Li, Y., Sheng, X., Lian, M., & Wang, Y. (2016). Influence of tilt angle on eddy current displacement measurement. *Nondestructive Testing and Evaluation*, *31*(4), 289-302.

A coupled phase-field and finite element method to simulate the elasto-plastic deformation induced cementite dissolution in pearlitic rail steels

***Hu Chen[1, 2], Chi Zhang[1], †Lei Chen[2]**

[1]School of Materials Science and Engineering, Tsinghua University, China
[2]Department of Mechanical Engineering, Mississippi State University, USA

*Presenting author: chenhu14@mails.tsinghua.edu.cn
†Corresponding author: chen@me.msstate.edu

Abstract

A coupled phase-field and finite element method is originally proposed to investigate the cementite dissolution behavior in pearlitic steels subjected to cyclic deformation. The stress distribution and plastic strain accumulation are calculated by a rolling-sliding contact finite element model assisted with a plastic strain accumulation model. With the input parameters from the finite element model, an elasto-plastic phase-field model is then employed to simulate the real-time evolution of cementite volume fraction, microstructure morphology and carbon distribution for different rolling cycles and contact depths. Upon experimental validations, the proposed model predicts more accurate and realistic results than Sauvage's model. A three-stage dissolution kinetics is also revealed, which well explains an experimentally observed microstructure gradient along the depth direction. The proposed elasto-plastic phase-field model can be potentially extended to simulate cementite dissolution under various manufacturing or serving conditions, and even any stress-driven microstructure evolution containing cementite dissolution.

Keywords: phase-field; finite element; cyclic deformation; cementite dissolution; microstructure prediction

1. Introduction

Rails are subjected to increasing mechanical wear and rolling contact fatigue (RCF) due to fast train speed [1] and massive axle loads [2] in modern railway systems. Therefore, there are surging interests to figure out the mechanism of formation of those material failures to improve the properties of existing rail steels and to prolong service life of rails. Microstructure change in the surface layer of rail steels is considered to be related to the degradation of mechanical properties and thus microcrack formation in rails [3][4]. For example, the formation of White Etching Layer (WEL), a hard and brittle layer with considerable dissolution of cementite and grain refinement, on rail surface is usually believed to be favorable locations of crack initiation and be detrimental to the rail lifetime [5][6]. However, the microstructure changes, including pearlite deformation, grain refinement and especially cementite dissolution, during rolling-sliding contact loading remains elusive. Therefore, there is currently a critical need to understand the fundamental mechanism of microstructure evolution of rail steels under rolling-sliding contact deformation.

As reviewed by Gavriljuk [7], there are generally two interpretations of cementite dissolution under severe loadings. Gridnev et al. [8] proposed that the cementite dissolution is due to the higher binding energy between carbon interstitials and dislocations as compared to the binding energy between carbon atoms and iron atoms in cementite. On the other hand, Languillaume et al. [9] thought that the increase in free energy caused by the geometrical thinning of cementite

lamellae and the formation of slip steps during plastic deformation. Sauvage et al. [10] semi-quantitatively calculated cementite dissolution rate in cold-drawn pearlitic wires based on the Gibbs-Thomson effect and a diffusion-controlled dissolution process. Unfortunately, the predicted dissolution rates appeared to be underestimated in comparison with the experimental measurements. Further, the work was limited to cold-drawn pearlitic wires in which the mechanical deformation is applied in the fashion of only one or a few cycles. Pearlitic rail steels, however, have to endure thousands of rolling-sliding contact deformation cycles during service. More recently, Nematollahi et al. [11] suggested that in addition to the free energy increase mechanism proposed by Languillaume, an elastic strain induced thermodynamic driving force is possible according to first principle calculation, which may account for the underestimation of previous simulation. Hence combing the elastic and plastic effect may get a more reasonable result.

In the present study, an elasto-plastic phase-field model, integrating an elastic strain and plastic strain induced free energy increase in cementite and ferrite, is originally proposed to understand the fundamental mechanism of mechanical deformation induced cementite dissolution. The simulated cementite morphology agrees well with the experimental results.

2. Model description

In this study, a multiscale computational framework, in which an elasto-plastic phase-field model is coupled with the finite-element method, is proposed to model the cyclic mechanical deformation-driven cementite dissolution. The distributions of stress and strain as functions of rolling cycles and contact depths, at the macro-scale, are first calculated by a rolling-sliding contact finite element model. The accumulated strain is then fed into the phase-field model to simulate the real-time evolution of cementite volume fraction and microstructure morphology at the meso-scale.

2.1. Finite-element model

To improve the computation efficiency, a 2-D plane strain finite element model is employed to simulate the stress distribution in the rolling-sliding experiment. The penalty function methods are used to solve the contact and friction problem. Contact load and friction torque applied are 1500 N and 18 Nm respectively, from rolling-sliding experiments. The Young's modulus and the Poisson's ratio of the rail are set as 210 GPa and 0.3. The yield strength and tensile strength are 490 MPa and 900 MPa, respectively.

2.2. Plastic strain accumulation model

A plastic strain accumulation model proposed by Kapoor and Franklin [12] is used. With the shear stress distribution and material hardening behavior, the plastic strain accumulation as a function of rolling cycles and contact depths can be calculated. The shear strain increases with rolling cycles is given by [12]:

$$\Delta\varepsilon_p^{[z]} = C\left(\frac{\tau_{zx(\max)}^{[z]}}{k_{eff}^{[z]}} - 1\right) \tag{1}$$

where $\tau_{zx(\max)}^{[z]}$ is the maximum shear stress at the depth of z, $k_{eff}^{[z]}$ is the effective shear strength of material at the depth of z, and C has a constant value of 0.00237 [13].

The total accumulated shear strain is then calculated by:

$$\varepsilon_p^{[z]} = \sum_N \Delta\varepsilon_p^{[z]} \tag{2}$$

where N is the number of rolling cycles.

2.3. Phase-field model

Ferrite and cementite are described by a field variable ϕ, ranging from 0 in ferrite to 1 in cementite. The carbon distribution in the phases is described by the concentration field, c. Specifically, c^{α} and c^{Fe_3C} represent the carbon composition (in atomic ratio) of the two phases, respectively. Because of its accuracy of modeling stoichiometric compounds with sharp curvature of free energy function, Kim-Kim-Suzuki (KKS) model [14] is employed. The total free energy F_{total} of the system is given by:

$$F_{total} = \int_V \left[\frac{1}{\Omega_0} f_{local}(c,\phi,\varepsilon) + \frac{\kappa}{2}|\nabla\phi|^2 \right] dV \tag{3}$$

where Ω_0 is the molar volume of the system, and $f_{local}(c,\phi,\varepsilon)$ is the local free energy, which is related to the free energy of ferrite, f^{α}, and cementite, f^{Fe_3C}. The second gradient term partially accounts for the interfacial inhomogeneity between two phases and κ is a gradient coefficient.

The key novelty of the current phase-field model is to involve both the elastic strain and plastic strain contributions in f^{α} and f^{Fe_3C}:

$$f^{\alpha} = f^{\alpha}_{chem}(c^{\alpha}) + f^{\alpha}_{el}(\varepsilon_e) + f^{\alpha}_{pl}(\varepsilon_p) \tag{4}$$

$$f^{Fe_3C} = f^{Fe_3C}_{chem}(c^{Fe_3C}) + f^{Fe_3C}_{el}(\varepsilon_e) + f^{Fe_3C}_{pl}(\varepsilon_p) \tag{5}$$

where f^{α}_{chem} and $f^{Fe_3C}_{chem}$ are the chemical free energy of ferrite and cementite, respectively. The chemical free energy of ferrite follows the work of Loginova [15], which is based on regular solution model. The elastic part, $f^{\alpha}_{el}(\varepsilon_e)$ and $f^{Fe_3C}_{el}(\varepsilon_e)$, accounting for the lattice distortion of cementite and ferrite crystals, are obtained by fitting the total energy of bcc Fe and cementite as a function of elastic strain in Ref. [11] into a parabolic function.

The plastic part $f^{Fe_3C}_{pl}(\varepsilon_p)$ in cementite phase is then estimated using a linear approximation [10]:

$$f^{Fe_3C}_{pl}(\varepsilon_p) = \gamma V_m \frac{1+\ln(1+\varepsilon_p)}{t} \tag{6}$$

where γ is the ferrite/cementite interfacial energy without plastic deformation, V_m is the molar volume of cementite and t is the lamellar thickness of cementite, ε_p is the plastic strain accumulated during the rolling-sliding contact deformation calculated from the plastic strain accumulation model.

The spatial and temporal evolutions of composition and order parameters are governed by the Cahn–Hilliard equation [16] and Allen–Cahn equation [17], respectively:

$$\frac{\partial c}{\partial t} = \nabla \cdot \left[\frac{D(T)}{(f_{local})_{cc}} \nabla \frac{\partial f_{local}}{\partial c} \right] = \nabla \cdot \left[M(T) \nabla \frac{\partial f_{local}}{\partial c} \right] \tag{7}$$

$$\frac{\partial \phi}{\partial t} = -L\frac{\delta F}{\delta \phi} = \frac{L}{\Omega_0}\left[-\frac{\partial f_{local}}{\partial \phi} + \Omega_0 \kappa \nabla^2 \phi\right] \tag{8}$$

where $D(T)$ and $M(T)$ are chemical diffusivity and mobility, respectively, $(f_{local})_{cc}$ is the second derivative of f_{local}, and L is the interface mobility coefficient.

Both kinetic equations are solved in Fourier space [18] using the FFT algorithm in reduced units, with nondimensionalizing factors $E = 1.18 \times 10^{11}$ J/m^3 and $l = 10^{-9}$ m. A 2-D phase-field model with a size of 600 × 600 grid, representing a physical domain of 600 nm × 600 nm, is employed. Unless otherwise specified, the interface thickness (2λ) is 4 nm in the present simulation.

3. Model validation

The proposed elasto-plastic phase-field model is first validated by comparing the cementite dissolution amount with Sauvage's prediction and the existing experimental results in cold-drawn wires, and the results, i.e., the amount of dissolved cementite after cold drawing as a function of temperature are shown in Fig. 1. The parameters used in the validated case all follow those of Sauvage's work [10], in which the wire temperature is 200°C, the true logarithmic strain is 3.5 and the interlamellar spacing of pearlite is 90 nm. When the temperature decreases to 100°C, the dissolution amount from phase-field simulations, either with or without the elastic energy, agrees well with the experimentally measured amount, in contrast to Sauvage's model [10] which presents a clear (70%) underestimation of the dissolved cementite. As the ferrite/cementite interface thickness decreases to 1.3 nm, close to the grid size, the prediction of present phase-field model converges to Sauvage's model. Hence Sauvage's model is a special case of present model and the discrepancy between the above two models is primarily attributed to the difference in the description of the ferrite/cementite interface. In Sauvage's model, a sharp interface and a constant carbon concentration (25 at.%) in cementite were considered; whereas a diffuse interface proposed in the phase-field model presents a carbon concentration gradient at the cementite side of the ferrite/cementite interface, which is more consistent with the experimental results measured by 3DAP [19], as shown in the inserted picture in Fig. 1. The model with the elastic energy clearly outperforms the counterpart without the elastic contribution, when compared with the experimental measurements at the same strain level [20]. Consequently, we rationalize the necessity of the contribution by the elastic strain in the free energy, in addition to the plastic contribution. The phase-field model, integrating both elastic and plastic energy, is subsequently applied to the pearlitic steels subjected to the cyclic rolling-sliding contact deformation.

4. Application in rolling-sliding condition

The shear stress distribution of contact zone is shown in Fig. 2(a). It can be seen that the stress pattern is only existed in surface region of about 3 mm in thickness, so the use of a circular segment with height of 6mm in the rolling-sliding contact finite model would not affect the stress calculation of contact zone. An enlarge picture of the contact zone is displayed in Fig. 2(b). It can be observed that the maximum shear stress is located at the top surface with a value of 543 MPa. The shear stress which is large than the shear strength of material is within 200 μm region, so the maximum shear stress distribution at the corresponding depths is plotted in Fig. 2(c). The maximum shear stress decreased linearly to 400 MPa with the depth increasing to 50 μm, and then the maximum shear stress stayed unchanged in the next 25 μm, resulting from the position indicating by the dot circle in the inserted figure. After that, the maximum shear stress gradually declines to 300 MPa at depth of 200 μm.

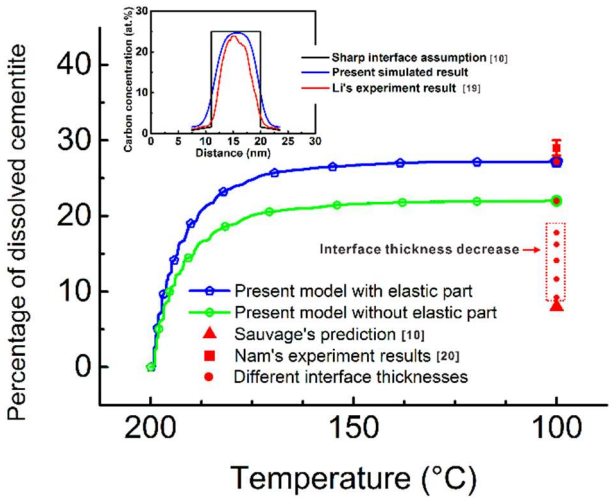


Fig. 1. Cementite dissolution behavior simulated by present phase-field model with ferrite/cementite interface thickness of 4 nm, by Sauvage's model [10] and of experiment results [19] in cold-drawn wires; and cementite dissolution behavior with different ferrite/cementite interface thicknesses (2 nm, 1.8 nm, 1.6 nm, 1.4 nm and 1.3 nm). Carbon concentration across ferrite/cementite interface based on sharp interface assumption [10], present diffusion interface assumption and experiment measurement [20] in inserted picture.

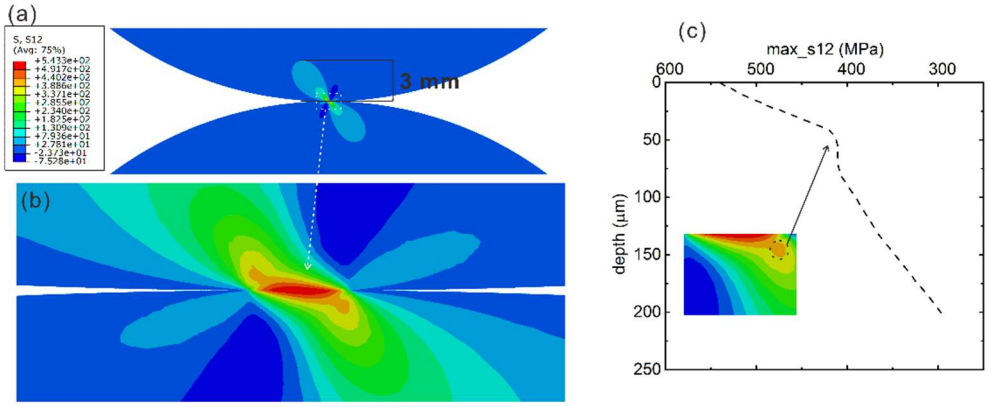


Fig. 2. (a) Shear stress distribution of sample under rolling-sliding loading; (b) magnified view of the vicinity of contact region; (c) maximum shear stress distribution with depth.

The accumulated strain as functions of contact depths and rolling cycles is shown in Fig. 3. In Fig. 3(a), it can be seen that at 2000 cycles, the accumulated strain at the top surface is 2.5, and it gradually decreases with the increase of depth, to almost 0 at the depth of 200 μm. With the increase of rolling cycles, the accumulated strain increased rapidly at the top surface reaching the critical value 11 at 14000 cycles. However, the strain accumulation speed is much slower at inner part as shown by the slope of the curves in Fig. 3(b). This is because that the maximum shear stress at inner part is not much larger than the yield strength of material and strain hardening even narrows the difference. After 14000 cycles, the accumulated strains reach their maximum values at all depths, i.e. maximum shear stress reaches an equilibrium with the

effective yield strength. The maximum shear strains at depth of 10 μm, 20 μm, 30 μm, 40 μm, 60 μm and 100 μm are 9.18, 7.13, 4.98, 3.21, 2.65 and 2.00 respectively.

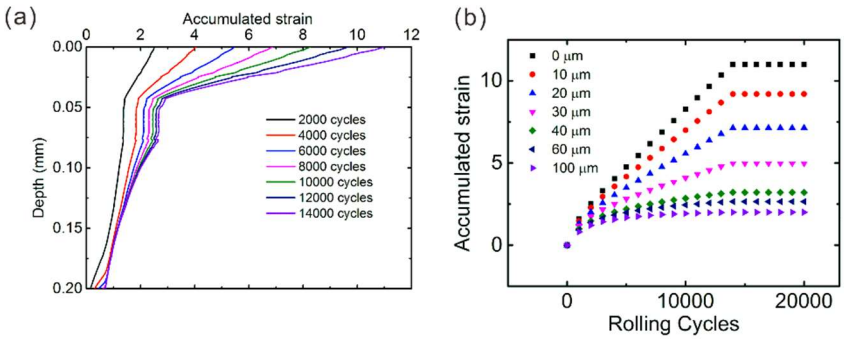


Fig. 3. Accumulated plastic shear strain as a function of (a) contact depth and of (b) rolling cycles at the contact depth of 0 μm,10 μm, 20 μm, 30 μm, 40 μm, 60 μm, and 100 μm.

The evolution of cementite morphology and carbon concentration is shown in Fig. 4. At the contact surface, i.e. at the depth of 0 μm, cementite shows a fastest dissolution rate and the thickness of cementite decreases to 2.7 nm and starts to show breakage or discontinuity at 34700 rolling cycles. After that cementite length witnesses a rapid shrinkage and only four small pieces of cementite left at 45000 cycles. After 60000 cycles, no cementite left in the phase field, ϕ, profile and carbon concentration shows a redistribution.

Although cementite is still almost fully dissolved at 10 μm below the contact surface after 90000 cycles, the dissolution speed is a little slower than that of the contact surface. The breakage of cementite appears at 50000 cycles with an average cementite thickness of 4.0 nm. At 51500 cycles, cementite lamellae are broken into small particles and the boundaries between cementite and ferrite starts to disappear.

At depth of 20 μm, there is still 3.6% volume fraction of cementite left after 120000 cycles, which may because that the driving force is not large enough to induce a full dissolution or the rolling cycles is not enough for the system to reach another quasi-equilibrium. At depth of 30 μm, however, only decrease in the thickness cementite lamellae is observed in the morphology profile, without any breakage showing up.

The thickness of cementite lamellae as a function of rolling cycles and contact depths is displayed in Fig. 5. The thickness is the average value calculated from the volume fraction of cementite, so it is underestimated after the emergence of cementite breakage. Because the trend of cementite dissolution is more important than the exact value in present study and the thickness of cementite is already very small after breakage, using the above average thickness is acceptable. It can be seen that in Fig. 5, before a huge drop to 1.6 nm during 30000 to 35000 cycles, the thickness of cementite lamellae decreases dramatically to 6.1 nm in the first 30000 cycles at the depth of 0 μm. The cementite dissolution behavior at depth of 10 μm is similar to that of the contact surface except that the huge drop comes up at 45000 cycles. Cementite lamellae both witness a complete dissolution in those two depths; however, at the depth of 20 μm, the dissolution speed is much slower and gradually decreases to 5.3 nm at 120000 cycles. For depths at more than 40 μm, the thickness of cementite lamellae shows obvious decrease from 18nm to 14.8 nm at depth 40 μm, to 15.3 at depth 60 μm and to 15.8 nm at depth 100 μm in the first 50000 cycles and then stays almost constant.

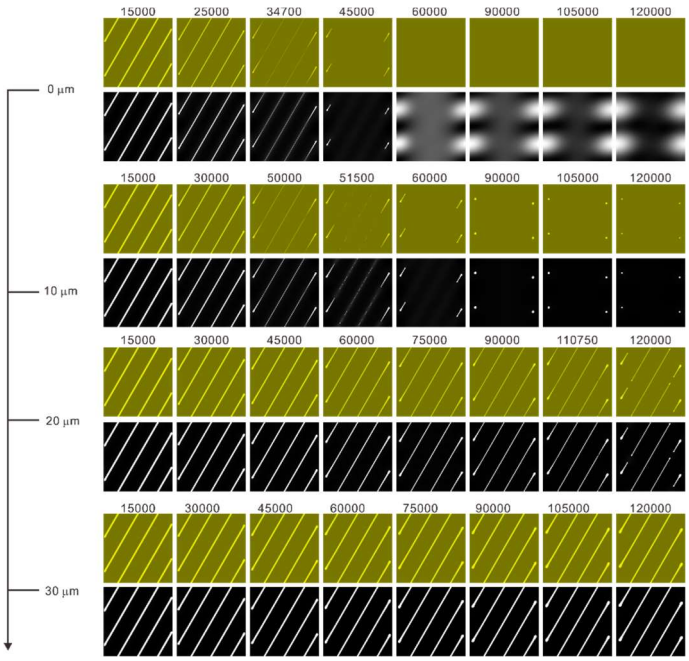


Fig. 4. Microstructure morphology (yellow) and carbon distribution (white-black) of pearlite after different rolling cycles and at different contact depths.

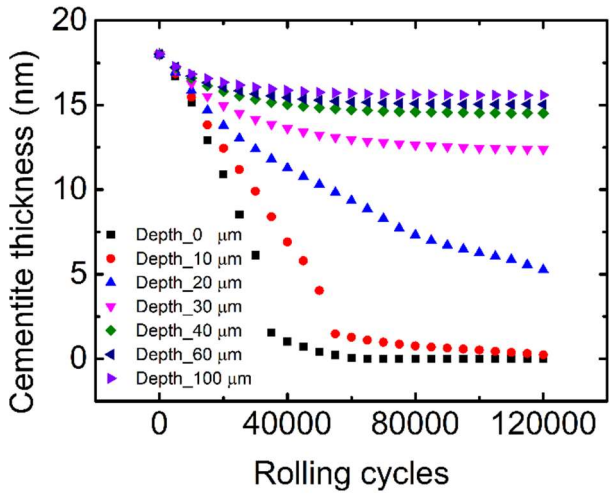


Fig. 5. Cementite lamellae thickness evolved with rolling cycles.

As revealed in Fig. 6(1), cementite lamellae show a three stages dissolution behavior. In stage I, cementite displays a linear dissolution rate due to the increase of free energy as indicated by line a–b, d–e and g–h. At the transition point between stage I and II, discontinuity or breakage starts to appear in cementite lamellae. The increase of interface area where the cementite breaks can significantly enhance the dissolution kinetics in stage II. In addition, the curvature at the surface of the breakage part is much larger than the plain interface between the prior ferrite and cementite lamellae, which can lead to a fast shortening of cementite lamellae due to the Gibbs-Thomson effect. Consequently, a steep dissolution behavior is observed in stage II. At the end of stage II, the interfaces between ferrite and cementite lamellae become vague and only some small particles located at prior cementite lamellae are left. In stage III, spheroidization of cementite and diffusion of carbon atoms are achieved.

It should be noted that the cementite dissolution behavior is strongly dependent on the depth. More importantly, due to the fast dissolution kinetics in stage II, the "transition zone" for different kinetic stages in the sample after certain rolling cycles is quite sharp. As illustrated in Fig. 6(2), after 120000 cycles, the volume fractions of cementite at depths of less than 17.5 μm are all well below 1.0%, showing an almost full dissolution behavior. However, at the depth of 20 μm, 3.6% cementite is present with an average lamellae thickness of 5.3 nm, indicating a sharp transition zone of less than 2.5 μm. It is worth noting that sharp boundaries between the WEL with complete cementite dissolution and the layer below are frequently observed experimentally under rail surfaces after service and are often used to support the heat induced cementite dissolution and WEL formation mechanism. In our simulation assuming the mechanical deformation-driven cementite dissolution mechanism, however, if the full dissolution regions are considered as the WEL, the sharp boundaries can also be realized due to the transition of the cementite dissolution kinetics.

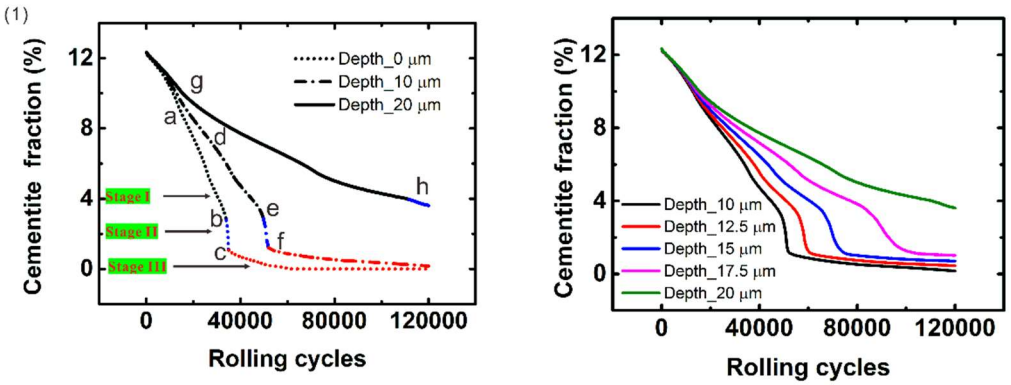


Fig. 6. (1) Cementite dissolution kinetics; and (2) cementite dissolution behavior across the transition zone, i.e. at the depth of 10 μm, 12.5 μm, 15 μm, 17.5 μm and 20 μm.

5. Conclusions

An elasto-plastic phase-field model is proposed to simulate mechanical deformation-driven cementite dissolution behavior of pearlitic rail steels under cyclic deformation. The proposed elasto-plastic phase-field model can well predict the evolution of cementite morphology and carbon concentration of pearlitic steel. A three stages dissolution kinetics of cementite lamellae is determined from the phase-field simulations. Because of the kinetics transition, a great microstructure gradient is predicted along the contact depth, which well explains an experimentally observed sharp microstructure transition along the depth direction. The proposed elasto-plastic phase-field model can be potentially extended to simulate cementite dissolution under various manufacturing or serving conditions, and even any stress-driven microstructure evolution containing cementite dissolution.

References

[1] H.H. Ding, Z.K. Fu, W.J. Wang, J. Guo, Q.Y. Liu, M.H. Zhu, Investigation on the effect of rotational speed on rolling wear and damage behaviors of wheel/rail materials, *Wear* 330 (2015) 563-570.
[2] Y. Zhou, S. Wang, T. Wang, Y. Xu, Z. Li, Field and laboratory investigation of the relationship between rail head check and wear in a heavy-haul railway, *Wear* 315 (2014) 68-77.
[3] J.E. Garnham, C.L. Davis, The role of deformed rail microstructure on rolling contact fatigue initiation, *Wear* 265 (2008) 1363-1372.
[4] R. Carroll, J. Beynon, Rolling contact fatigue of white etching layer: Part 1: Crack morphology, *Wear* 262 (2007) 1253-1266.

[5] J. Seo, S. Kwon, H. Jun, D. Lee, Numerical stress analysis and rolling contact fatigue of White Etching Layer on rail steel, *Int. J. Fatigue* 33 (2011) 203-211.
[6] H.W. Zhang, S. Ohsaki, S. Mitao, M. Ohnuma, K. Hono, Microstructural investigation of white etching layer on pearlite steel rail, *Mater. Sci. Eng. A* 421 (2006) 191-199.
[7] V. Gavriljuk, Decomposition of cementite in pearlitic steel due to plastic deformation, *Mater. Sci. Eng. A* 345 (2003) 81-89.
[8] V. Gridnev, V. Nemoshkalenko, Y. Meshkov, V. Gavrilyuk, V. Prokopenko, O. Razumov, Mössbauer effect in deformed Fe–C alloys, *physica status solidi (a)* 31 (1975) 201-210.
[9] J. Languillaume, G. Kapelski, B. Baudelet, Cementite dissolution in heavily cold drawn pearlitic steel wires, *Acta Mater.* 45 (1997) 1201-1212.
[10] X. Sauvage, J. Copreaux, F. Danoix, D. Blavette, Atomic-scale observation and modelling of cementite dissolution in heavily deformed pearlitic steels, *Phil. Mag. A* 80 (2000) 781-796.
[11] G.A. Nematollahi, J. von Pezold, J. Neugebauer, D. Raabe, Thermodynamics of carbon solubility in ferrite and vacancy formation in cementite in strained pearlite, *Acta Mater.* 61 (2013) 1773-1784.
[12] F. Franklin, I. Widiyarta, A. Kapoor, Computer simulation of wear and rolling contact fatigue, *Wear* 251 (2001) 949-955.
[13] A. Bower, Cyclic hardening properties of hard-drawn copper and rail steel, *Journal of the Mechanics and Physics of Solids* 37(4) (1989) 455-470.
[14] S.G. Kim, W.T. Kim, T. Suzuki, Phase-field model for binary alloys, *Phys. Rev. E* 60 (1999) 7186.
[15] I. Loginova, J. Ågren, G. Amberg, On the formation of Widmanstätten ferrite in binary Fe–C–phase-field approach, *Acta Mater.* 52 (2004) 4055-4063.
[16] J.W. Cahn, J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys.* 28 (1958) 258-267.
[17] S.M. Allen, J.W. Cahn, A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening, *Acta Metall.* 27 (1979) 1085-1095.
[18] S.Y. Hu, J. Murray, H. Weiland, Z.K. Liu, L.-Q. Chen, Thermodynamic description and growth kinetics of stoichiometric precipitates in the phase-field approach, *Calphad* 31 (2007) 303-312.
[19] Y. Li, P. Choi, C. Borchers, S. Westerkamp, S. Goto, D. Raabe, R. Kirchheim, Atomic-scale mechanisms of deformation-induced cementite decomposition in pearlite, *Acta Mater.* 59 (2011) 3965-3977.
[20] W.J. Nam, C.M. Bae, S.J. Oh, S.-J. Kwon, Effect of interlamellar spacing on cementite dissolution during wire drawing of pearlitic steel wires, *Scripta. Mater.* 42(5) (2000) 457-463.

Temperature variations and cooling efficiency of forced convective heat transfer of nanofluids in microchannel laminar flow

†*V.K. Sin and K.K. Teng

[1]Department of Electromechanical Engineering, University of Macau, Macau SAR, China.
*Presenting author: vksin@umac.mo
†Corresponding author: vksin@umac.mo

Abstract

Numerical study of heat transfer of nanofluids has been conducted to investigate forced convective cooling in rectangular microchannel. A micro-electro-mechanical system is modelled by building with two heating resistors embedded in a silicon substrate. Microchannel made from polydimethylsiloxane is planted in the substrate with a length of 17.5mm. Two cross section areas with high and width of 100μm x 100μm and 100μm x 1000μm are being used for the study. TiO_2/water nanofluid with different volume fractions are used as the coolant. Because of the not well-insulated micro-electro-mechanical system, heat generated by the heating resistors will be lost to the surrounding as well as carried away by the nanofluids through the microchannel. Temperature variations along the microchannel is calculated based on the concept of conjugate convective heat transfer. Force convective heat transfer is analyzed at low Reynolds number for different nanofluid concentrations. Numerical results are compared with experimental data available in the literature.

Keywords: Nanofluid, convective heat transfer, microchannel

Introduction

Due to scientific and technological progress and innovations, traditional heat exchanger technology has become difficult to meet the cooling requirement. It is necessary to develop a better heat transfer technology. Some researchers began to explore the nanotechnology and cooling of micro-electro-mechanical system (MEMS). Nanofluid is proposed to improve the cooling effect by increasing heat transfer coefficient [1]. Nanofluid is a fluid which consists of nanometer-sized particles dispersed in liquids. Nanoparticles are typically made of metals, oxides, carbides, or carbon nanotubes and typically base fluids were water, ethylene glycol and oil. Nanoparticles sizes typically on the order of 1 to 100 nm. Moreover, nanoparticle shapes typically have sphere, cylinder, platelet, blade and brick. Nanofluid is primarily used as coolant in heat transfer equipment such as microelectronics, engine cooling, chiller and heat exchanger.

In order to improve heat transfer technology, one of the effective methods is to add high conductivity solid particles in fluid because thermal conductivity of solid is much larger than liquid. As shown in Table 1.1 of [2], the thermal conductivity of copper is larger than thermal conductivity of water almost by 650 times. Moreover, thermal conductivity of Alumina also is more than that of common nonmetallic liquid by a large amount. Adding solid particles in fluid can greatly increases thermal conductivity of fluid; however it may lead to an increase in friction and viscosity if micrometer scale particles are added to the fluid. Along with the progress of science and technology, adding nanometer scale particles to fluid is found that can effectively improve thermal conductivity of fluid without increasing friction and viscosity.

Numerous literatures began to study about the nanofluid for increasing the heat transfer coefficient. Although many results have indicated that nanofluids can increase the convective heat transfer coefficient [3]-[6], some other results show that adding nanoparticles do not improve the cooling effect or increase the friction factor [7][8]. The possible explanations for the anomalous reasons were Brownian motion of the particles, molecular-level layering of the liquid at the liquid/particle interface, the nature of heat transport in the nanoparticles and the effects of nanoparticle clustering.

Nanofluids experiments have many limits. For example, adiabatic wall condition, constant temperature wall condition and supply constant heat flux are difficult to achieve. The limited amounts of sensors and low precision sensors will also affect the result of the experiment. Computational investigation could meet the requirements more easily. Experimental data and conditions such as temperature field, velocity field and energy distribution can be realized through computational investigation.

In order to further analyze heat transfer enhancement of nanofluids in rectangular microchannels flow, this numerical study is based on the experiment of [7] for comparison. The objective of present study is to analyze the nanofluids for heat transfer enhancement in different conditions such as different channel size, Reynolds number, and volume fraction of the nanoparticles. Temperature variation along the channel and heat carried away by the nanofluid will be investigated and the results will be compared with experimental data available in the literature.

Model Setting in Simulation

The experiment of [7] is modelled by COMSOL-Multphysics and is given in Fig. 1. To complete this complex modelling, the following settings which are close to [7] as possible are applied in the modelling. A base silicon substrate with a size of 45mm x 45mm x 525μm is built first. Two heating resistors with size of 2mm width, 30mm long and 0.1mm high are then embedded on the substrate. Two reservoirs with diameter of 2mm are connected to the inlet and outlet of the microchannel separately. Microchannel with length of 17.5mm and two different rectangular cross section areas of 100μm x 100μm and 100μm x 1000μm are used for study. For analysis of the experiment and simulation data, the model is marked with seven points which are the locations of experiment thermocouples of [7] as shown in Fig. 2. Point A and point G represent inlet and outlet position of microchannel, respectively. Point B to point F are locations for temperature measurement which has distance of 0 (inlet), 4.4, 8.8, 13.2, and 17.5mm (outlet) from the inlet, respectively. The entire microchannel and reservoirs are sheltered with a polydimethylsiloxane (PDMS) channel cover. Property of silicon substrate and PDMS channel cover is given in Table 1.

The simulation model is solved with finite element method for the governing partial differential equations. The more the amounts of grids are meshed, the more precise the result should be. Because computer operational capability is limited, most of the grids are set at channel for analyzing fluids heat transfer. The model is mainly meshed to triangular prism grids and tetrahedral prism grids. Moreover the heat transfer analysis is mainly needed for the acquisition data of channel surface, for example calculating the convective heat transfer coefficient is required the surface temperature and surface heat flux. Boundary layer meshing way is used in this model to mesh grids at the channel boundary. The total number of grids consist of around one million elements as shown in Fig. 3. In order to find out the amount of heat that is lost to the ambient air which surrounds the MEMS, an appropriate convective heat transfer coefficient of air has to be determined. It is found in [9] that boundary condition with heat transfer coefficient, h=12 W/m^2/K, at room temperature of twenty-five degrees Celsius is accurate enough for present simulation. This value of heat transfer coefficient is in the range of natural convection coefficient between 2 and 25 W/ m^2/K [2].

.

Table 1. Property of silicon substrate and PDMS cover

Material	Silicon substrate	PDMS cover
Density (kg/m^3)	2329	0.97
Thermal conductivity (W/m/K)	130	0.15
Specific heat (J/kg/K)	700	1460

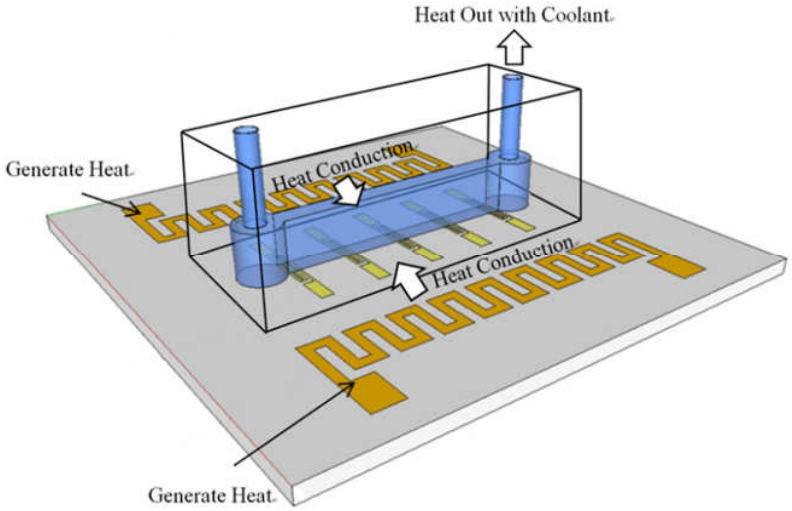


Figure 1. : Model setting of heat transfer in [7]

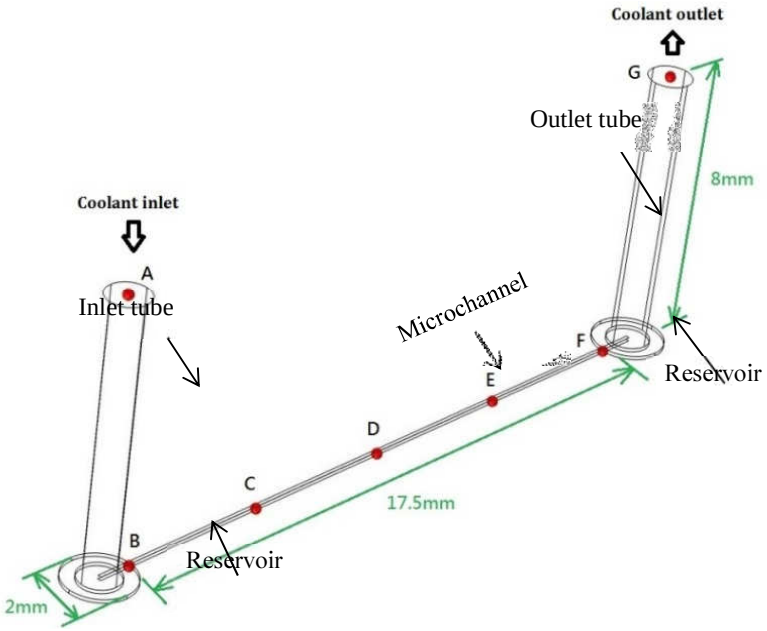


Figure 2. Locations of thermocouple and temperature simulation for comparison

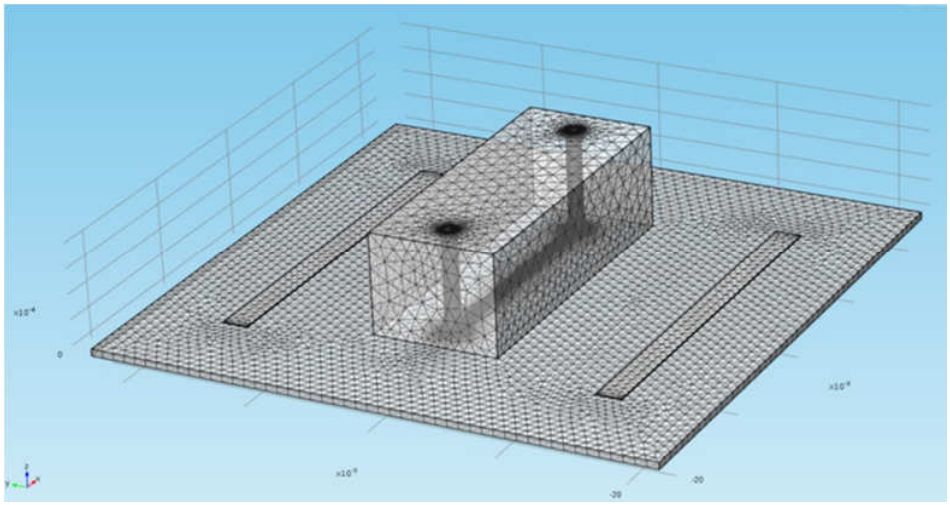


Figure 3. Typical mesh simulation model

Thermal and Transport properties of Nanofluids

Nanofluids in convective heat transfer has been researched only in past few decades, so nanofluid properties could not be explained by existing theories completely. In this study, nanofluid was assumed to be a new kind of fluid in the simulation. Effective thermal conductivity, specific heat capacity and dynamic viscosity of this new fluid were required in order to perform the simulation and analysis. Because there are many literatures indicated that the variation of dynamic viscosity can almost be ignored such as in [7], the dynamic viscosity was set the same as base fluid. Calculation of all these properties is given in [6] and the information is reproduced in Table 2.

Table 2. Properties of nanofluids

Volume fraction	0%	0.6%	1.2%	1.8%
Density (kg/m^3)	998.57	1017.96	1037.34	1056.73
Thermal conductivity (W/m/K)	0.6070	0.61644	0.62599	0.63564
Specific heat (J/kg/K)	4181.3	4094.24	4010.43	3929.70
Dynamic Viscosity (kg/m/s)	0.00084	0.00084	0.00084	0.00084

Results and Discussion

In order to analyze heat transfer enhancement of nanofluids flowing inside rectangular microchannels, the temperature of nanofluid along the microchannel is required for the analysis. The computational results and experimental results of [7] of nanofluid temperature along the channel with heating power 0.58Watt in different Reynolds numbers and different volume fractions for two cross section areas of 100μm×100μm and 100μm×1000μm are shown in Fig. 4 and Fig. 5, respectively. Reynolds number is defined as $Re = \frac{\rho u_m D_h}{\mu}$, where ρ is density of

nanofluid, u_m is mean velocity of nanofluid in microchannel, D_h is hydraulic diameter, and μ is dynamic viscosity of nanofluid.

In Fig 4, solid lines are for computational results while solid symbols for experimental data of [7]. For low Reynold number such as Re=6 shown in Fig 4(a), both simulation results and experimental results cannot clearly indicate the increase of temperature with increase of volume fractions. But for Re=10 and 18 as indicated in Fig. 4(b) and 4(c), both simulation and experimental results of temperature increase with increasing the volume fraction of nanofluid. In Fig. 4(c) the simulation results with volume fraction of 1.2% and 1.8% have very close temperature variations, so these two results collapsed on a single line. Moreover It is noted that while the simulation results find the nanofluid temperature increases gradually along the channel from inlet to outlet, the experimental results of [7] show the other way, nanofluid temperature decreases along the channel. According to the report of [7], this temperature decreases along the channel may due to the fact that some titanium dioxide nanoparticles may have agglomeration inside the channel at lower Reynolds number with cross section area of 100μm×100μm. The simulation results seem more reasonable than the experimental results of [7] as simulation results truly predict the increase of temperature as the fluid flows along the channel, this is because the fluid absorbs heat from the walls and make the temperature rise.

Figure 5 gives temperature result for cross section area of 100μm×1000μm. Although the simulation nanofluid temperatures do not increase with increase in the volume fraction of nanofluid for all Reynolds numbers being considered, the nanofluid with 1.8% volume fraction has temperature higher than that of deionized water. Adding nanoparticles to fluid does not contribute any regular variations to the experimental temperatures. Both simulation and experiment results of temperature increase as the fluid flows from inlet to outlet. It is also found that simulation results of temperature decrease with increase of Reynolds number.

Cooling efficiency (E) of nanofluid in microchannel represents the ratio of heat carried out by coolant (Q_f) over heat generated by the heating electrodes (Q_{chip}), which can be expressed as:

$$E = \frac{Q_f}{Q_{chip}} \times 100\% \tag{1}$$

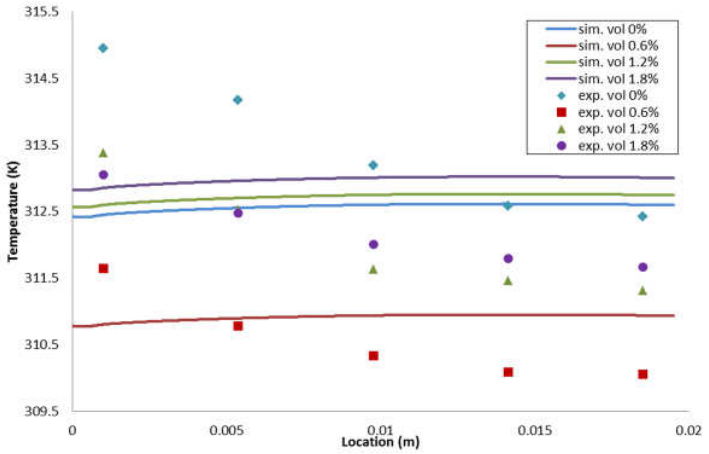
Where the heating power generated by electrodes is expressed as:

$$Q_{chip} = I * V$$

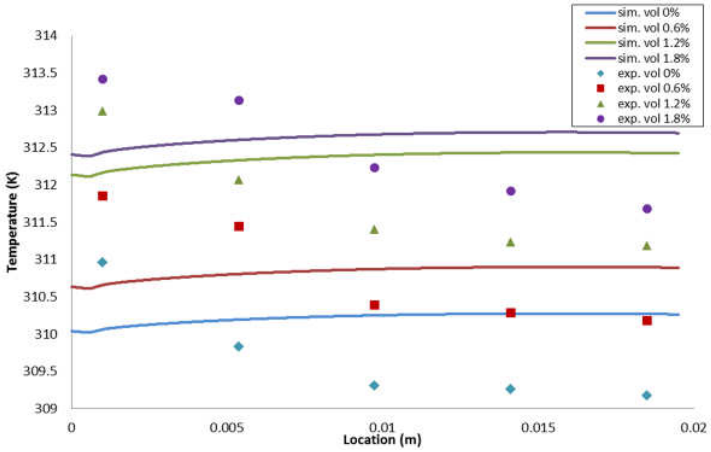
Where I is electric current and V is voltage. Heat carried out by the coolant can be expressed as

$$Q_f = \dot{m} * C_p * (T_{out} - T_{in})$$

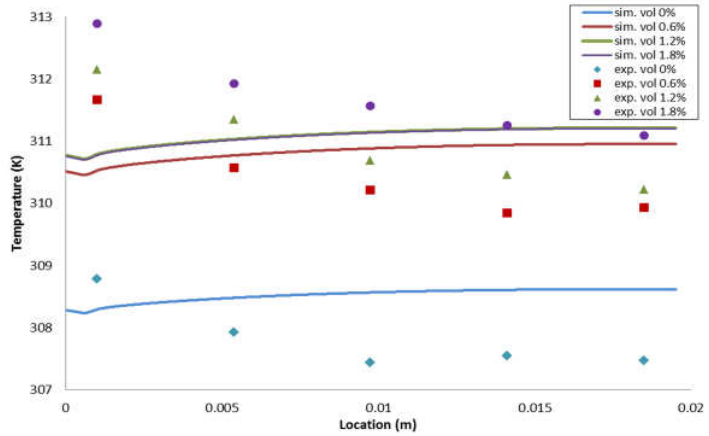
Where T_{out} and T_{in} are the outlet and inlet temperature of nanofluid respectively, $\dot{m}$ is the mass flow rate of nanofluid, and C_p is the specific heat capacity of nanofluid.



(a) Re = 6

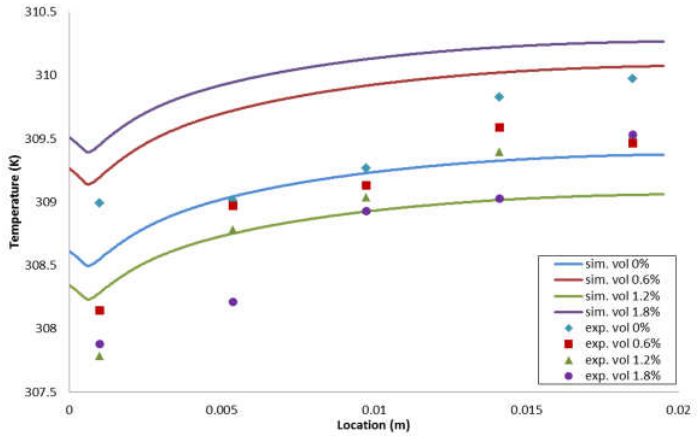


(b) Re=10

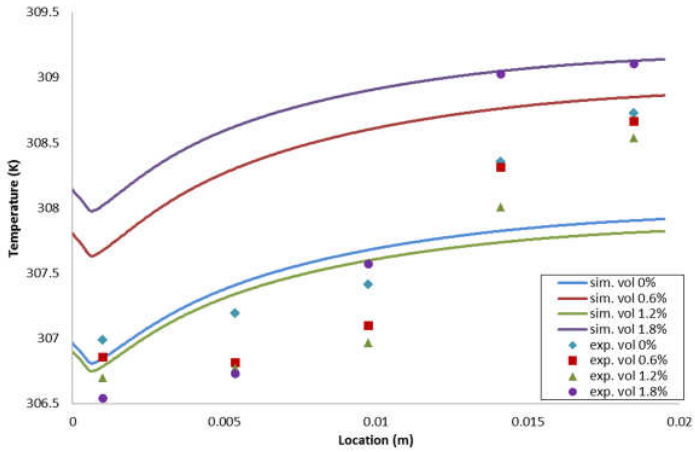


(c) Re = 18

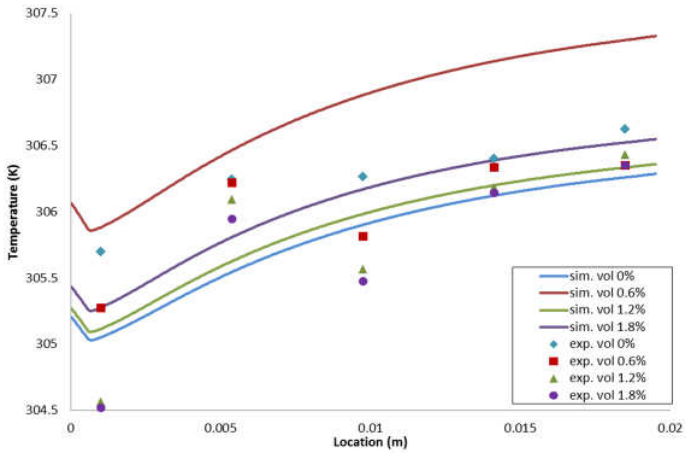
Figure 4. Temperature variations of nanofluid along the microchannel for different Reynolds Numbers (cross section=100μm×100μm, heating power=0.58Watt).



(a) Re = 6



(b) Re = 10



(c) Re = 18

Figure 5. Temperature variations of nanofluid along the microchannel for different Reynolds Numbers (cross section=100μm×1000μm, heating power=0.58Watt).

Figure 6 shows the numerical results of cooling efficiency versus Reynold number based on Eq. (1) for different combinations of cross section areas and heating powers. Also shown in the figure are the experimental results of [7] for comparison. Hollow symbols and solid symbols represent computational findings and experimental data respectively. In Figs. 5 (a) and 5(b) it is shown that simulation results of cooling efficiency is always lower than the experimental results for heating power = 0.58Watt, but in Fig. 5 (c) simulation results of cooling efficiency gradually become higher than experimental results, especially for higher Reynolds number. Moreover the cooling efficiency increases with increasing Reynolds number for all combinations of cross section areas and heating powers considered in this study. When comparing with Figs 5 (a) and 5(b) which different only in cross section area, it is clearly indicated that increase of cross section area of the microchannel always leads to increase of cooling efficiency.

Conclusions

Simulation of nanofluid heat transfer with software COMSOL Multiphysics was performed based on an existing experiment [7]. Simulation results show that nanofluid temperature increases along the microchannel. This finding is different from the experimental data for small cross section area of 100μm×100μm. It is also displayed that cooling efficiency increases with increasing cross sections area of microchannel.

Acknowledgment

The research is supported by the Research Committee of University of Macau under Multi-Year Research Grant MYRG2016-00074-FST. We would also like to thank Professor U. Lei from National Taiwan University for providing experimental data for comparison and giving this research so many valuable suggestions and comments

References

[1] Choi, S. U. S. and Eastman, J. A. (1995) Enhancing thermal conductivity of fluids with nanoparticles, *ASME International Mechanical Engineering Congress & Exposition*, 12-17.

[2] Cengel, Y. A. and Ghajar, A. J. (2015) *Heat and Mass Transfer Fundamentals and Applications*, 5th edn, McGraw Hill, USA

[3] Lie, J. and Kleinstreuer, C. (2008) Thermal performance of nanofluid flow in microchannels, *International Journal of Heat and Fluid Flow* **29**, 1221-1232.

[4] Choi, E. S., Brooks, J. S., Eaton, D. L., Al-Haik, M.S., Hussaini, M.Y., Garmestani, H., Li, D. and Dahmen, K. (2003) Enhancement of thermal and Electrical properties of carbon nanotubepolymer composites by magnetic field processing, *Journal of Applied Physics* **9**4, 6034-6039.

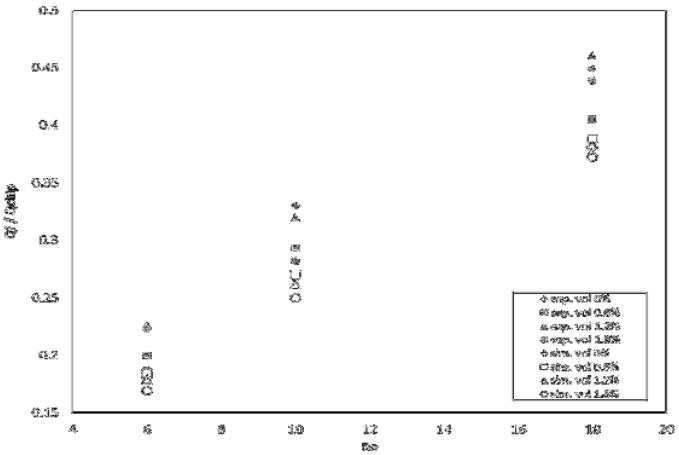
[5] Heris, S. Z., Etemad, S. G. and Esfahany M. N. (2006) Experimental investigation of oxide nanofluids laminar flow convective heat transfer, *International Communications in Heat and Mass Transfer* **33**, 529–535.

[6] Dalkilic, A. S., Kayaci, N., Celen, A., Tabatabaei, M., Yildiz, O., Daungthongsuk, W. and Wongwises, S. (2012) Forced Convective Heat Transfer of Nanofluids – A Review of the Recent Literature, *Current Nanoscience* **8**, 949-969.

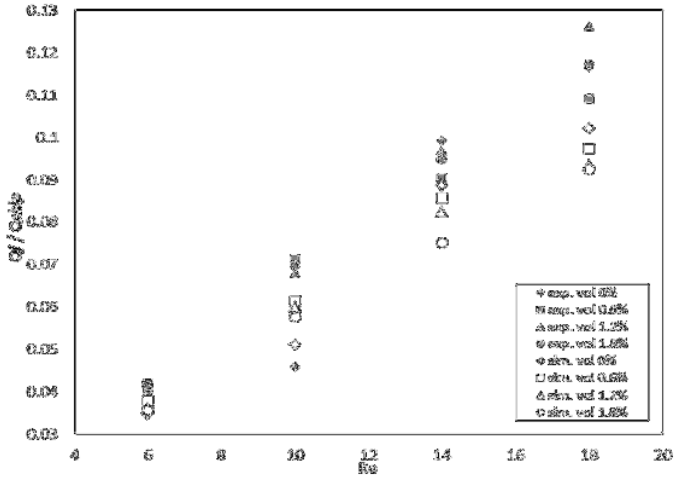
[7] Chen, Y. M., Experiment on Forced Convection of Nanofluids in Micro Channel, Master Thesis, National Taiwan University, Taiwan, 2013.

[8] Ho, C. J., Wei, L. C. and Li, Z. W., (2010) An experimental investigation of forced convective cooling performance of a microchannel heat sink with AL2O3/water nanofluid, *Applied Thermal Engineering* **30**, 96-103.

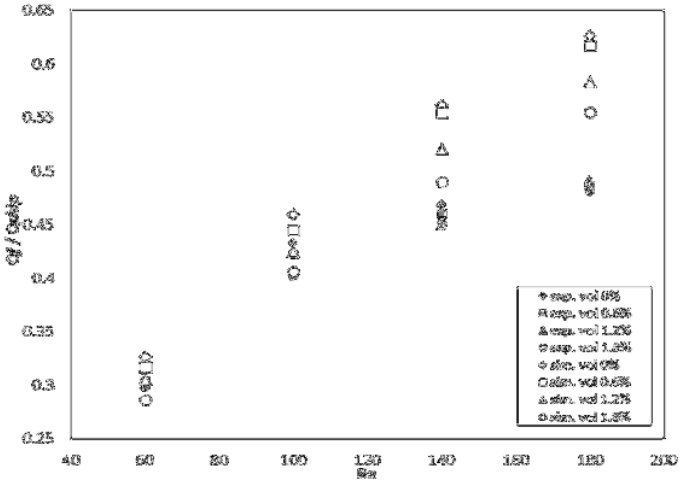
[9] Sin, V.K., Teng, K.K. and Deng, W.Y., Computational investigation of heat transfer of nanofluids in microchannel with improper insulation, *Lecture Notes in Computer Science 9786, The 16th International Conference on Computational Science and Its Applications (ICCSA 2016)*, Beijing, China, July 4–7, 2016, Proceedings, Part I, O. Gervasi et al. (Eds.): ICCSA 2016, Part I, LNCS 9786, 505–513, 2016**, ISBN:** 978-3-319-42084-4, Springer Publisher
http://link.springer.com/chapter/10.1007/978-3-319-42085-1_39/fulltext.html



(a) Cross section=100μm×1000μm, Heating power= 0.58Watt



(b) Cross section=100μm×100μm, Heating power= 0.58Watt



(c) Cross section=100μm×100μm, Heating power= 1.9Watt

Figure 6. Cooling Efficiency for different Reynolds numbers

Simulation driven development of a CFRPT gearbox housing

†Dr. T. Schneider[1], M. Kreutzmann[2], R. Rademacher[3], C. Dominé[4], H. Motte[5], *C. Tok[6]

[1]Technical Director ARRK/P+Z Engineering, Munich, Germany.
[2]Center of Competence Composite ARRK/P+Z Engineering, Munich, Germany.
[3]Powertrain department, ARRK/P+Z Engineering, Munich, Germany
[4]Low volume production department, ARRK/Shapers, France
[5]Low volume production department, ARRK/Shapers, France
[6]Crash department, ARRK/P+Z Engineering, Malaysia

*Presenting author: chyehock.tok@arrk-engineering.com
†Corresponding author: thomas.schneider@arrk-engineering.com

Abstract

To optimize the energy consumption of electric vehicles, the reduction of powertrain mass is important. Optimizing the weight of the typical two stage electric transmission is possible by substitution of the aluminum housing material with materials of lower density. Still, all requirements have to be fulfilled.

Carbon fiber reinforced thermoplastic materials provide a good combination of mechanical properties and fast manufacturing processes. The fulfilment of the necessary stiffness, especially at high temperatures of more than 100°C, is challenging for this group of materials. This resistance is important for durability and acoustics of the gearbox.

In this project, the housing material of an electric transmission, aluminum, is substituted by fiber reinforced thermoplastic material. An organo sheet (thermoplastic matrix) is over molded by a short fiber reinforced thermoplastic material. Aluminum inserts as bearing seats ensure to transmit the bearing loads into the organo sheet and reduce the deviation of the gear mesh. Additional injection molding ribs and UD-tapes ensure the stiffness requirements.

The morphological analysis investigated several concepts by using FE analyses and evaluated their feasibility. Topology optimizations, layer optimization and optimization of UD-Tape and injection molding ribs placement lead to load path optimized shape. Stamping and injection molding simulations were performed to optimize the manufacturability.

Close cooperation with the tool design ensures the manufacturability. The production of a prototype takes place in a three-phase process. In the first phase, the organo sheet and reinforcing UD-tapes are heated up, using an IR heating device and afterwards pressed in a stamping tool. Waterjet cutting produces the outline of the preform. During the final step, the organo sheet is heated again and over molded with the optimized rib geometry and all inserts to finalize the part.

Keywords: CFRPT, Gearbox housing material, hot pressing of organo sheet. Over molding of organo sheet

1. Leading motivation

Reduction of overall weight of the powertrain is one measures to reduce fuel consumption of vehicles. The weight of a gearbox housing can be reduced significantly by replacing the aluminum by a lighter material. Carbon fiber reinforced plastics with thermoset plastic matrices is suitable, as stiffness is one of the most important requirements for a gearbox housing. Unfortunately, the cost for this material is very high, compared to an aluminum cast part. Second, the production time is times 10-20 longer than the production of an aluminum casting part, what is a severe problem in mass production.

Injection molding and hot pressing of organo sheets are very cost effective. The process time is close to the process time in casting. Using organo sheets, carbon or glass fiber can be used in an orientated way in a plastic part, what provides stiffness to the part. In combination with thermoplastic over molding, any kind of geometry or functional detail can be added. The

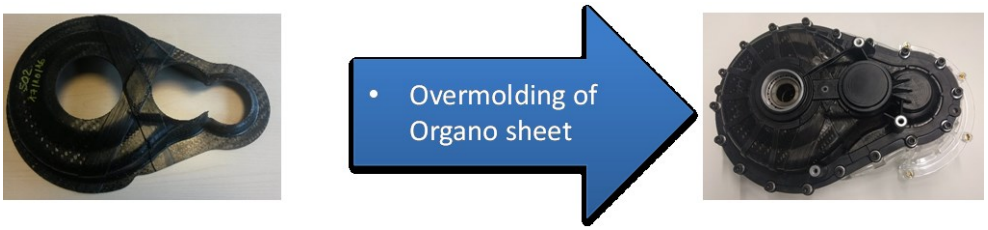


Figure 1: Over molding of organo sheet to introduce all relevant functions.

combination of both processes derive a part that has high stiffness and all necessary functional details.

As a demonstrator a gearbox housing used in an electric vehicle was chosen. The aim was to replace only the housing, all interior parts have been reused and work as proper as in the aluminum gearbox housing. The engineering approach used was simulation driven. Different types of simulations like finite element method, molding simulation, stamping simulation and optimization were used.

The companies ARRK P+Z, ARRK Shapers, ARRK SPG and ARRK UK were working together to show the ability of the ARRK product development group to develop the parts, to build the tools, to set up the production process and to produce in a small series production size. ARRK P+Z was responsible for the engineering, ARRK Shapers set up the production process, build the tools and the prototype.

2. Loads in the gearbox and simulation model

The gearbox used as basis of the project is sold by the company Getrag and used in the electric Smart. From the available data for the gearbox we derived the maximum torque of 130Nm at the input side for drive load case. With the given gear ratio, we derived an output torque of 1300Nm. For coast situation the torque depends mainly on the car weight and the car geometry. We derived a torque of 50Nm at the input shaft, which means a torque of 500Nm on the output shaft.
For the calculation of the torque load cases we imposed the loads on a simulation model. As drive and coast do not impose symmetrical forces into the gearbox housing, it is necessary to check both load cases in every step of the development. As the drive load case is more important for the optimization of the gearbox housing as the coast load case, the influence of

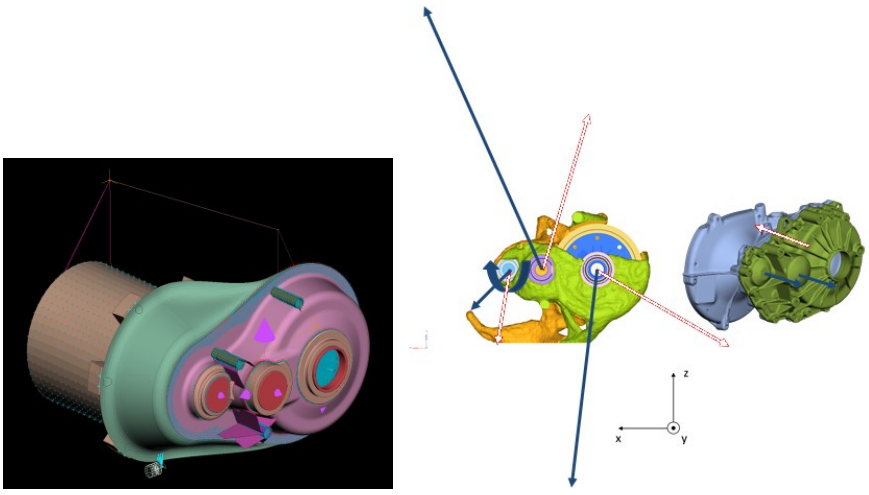


Figure 2: Left: Shell model of the gearbox (close to final design) with electric engine and fixation in the body. Model also includes internal parts like shafts and gears. Right: Qualitative radial and axial loads in the gearbox housing. Additionally a result of the topology optimization can be seen left hand.

the coast load case was not significant. We examined gravity load cases in x direction with 60g and z direction and 10g in x direction to simulate a crash situation and to examine the stability of the fixation of the gearbox to the car body.

To examine the impact of the torque load cases on the housing a model of the original gearbox with all the internal parts was build up. The geometry of the original gearbox was derived by disassembling the original gearbox and scanning the housing geometry.

3. Targets for the gearbox

The important performance targets for a gearbox are efficiency, lifetime and acoustics. The performance in acoustics is strongly coupled with the production of thermal energy, or the energy loss. A bad acoustic behavior, a short lifetime or generated thermal energy result in a low efficiency of the gearbox. In an ideal gearbox there would be no deformation of the housing by the transmitted loads. In such a gearbox, the tooth geometry could be designed for optimal efficiency [1].

Any displacement of bearings of the shafts leads to a change in position of the gears running on that shaft. Depending on the direction of the displacements, this leads to a misalignment of the gear mesh. The teeth of the gears compensate the deformation of the gearbox using e.g. crowning. The changes of the gears should be as small as possible to keep the efficiency of the gearbox as high as possible. Therefore the same stiffness values with the new housing material as in the original aluminum gearbox have to be reached. In terms of gear mesh, this means that the relative displacement between two shaft bearings have to be lower than in the aluminum gearbox. Axis deviation and axis inclination error of the intermediate and output shafts, have to be on the same level or below the values in the aluminum housing.

4. Material selection

The stiffness of the material depends on the kind of fibers and the matrix used. The glass fiber reinforced material showed about 50% of the stiffness of the carbon fiber reinforced material. Reaching the required stiffness was only possible by using carbon fiber reinforced material.
Equilibrated fabrics with 0°/90°/45°/-45°/-45°/45°/90°/0° stack were used and the thickness of the layers was optimized.
The organo sheet was available in thickness up to 5mm from the company TenCate. As thermoplastic molding material Grivory 40% fiber volume was chosen.

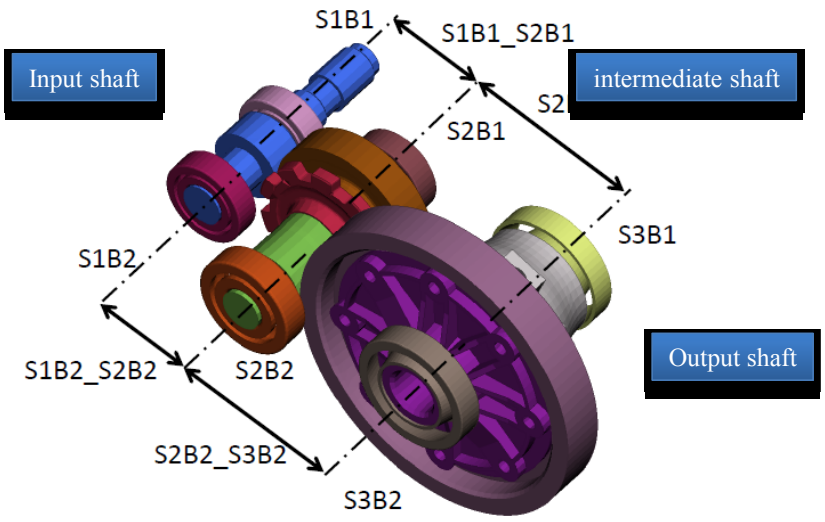


Figure 3: Declination errors occur between middle and output shaft of the gearbox.

5. Development of the gearbox

Concept phase

In the concept phase general feasibility was checked. The original gearbox housing was scanned and a FEA model was build up. Using this model we retrieved the following important results: Glass fiber showed a very low performance in stiffness, so we concentrated on carbon fiber for the organo sheet.

The housing part on the engine side was quiet stiff. Problems occurred from the part that we call the cover of the gearbox. Here, the most critical load case was the drive load case, and here it was the torsion of the gearbox that the cover could not prevent efficiently. The target value with the biggest impact on all decisions was the axis inclination between the intermediate and the output shaft.

The main parameters in the morphological box [2] for the gearbox were the following: Geometry, Loads, functional design, environmental topics, costs, maintenance. The three main concepts (five concepts were investigated) had the characteristics as can be seen in the figure 4. From the morphological analysis concept 1 and 3 were chosen.

Concept 1 offers the highest potential for weight saving and cost saving and the lowest production time. Concept 2 does not show the full potential of the thermoplastic fiber reinforced matrix material in combination with the carbon fibers. The metal link between the bearings covers all stiffness requirements. Concept 3 is a good compromise between concept 1 and 2 with cost and weight on a good level.

Design phase

Feasibility checks of concept 1 showed, that it is not possible to reach the necessary stiffness between intermediate and output shaft without using metal bearing seats. So we did not go on with concept 1 in the design phase and concentrated on concept 3.

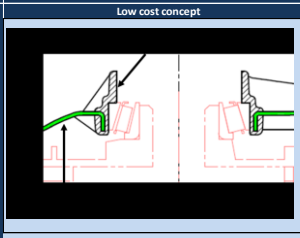
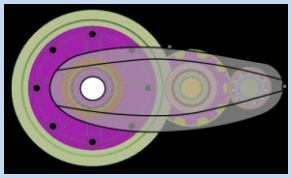
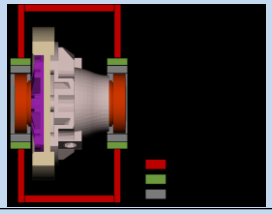
B6A1:B5A1:B4	1	2.1	3
Description	Low cost concept	Cage insert concept - GFK	Seperate bearing insert concpet
			
Details	- current gears, bearings, sealing - thermoplastic bearing seat - glass fiber organosheet - glued flanges (housing/cover) - no inserts at screws - center surface housing to cover - engine screwed to housing, centering by surface - simple air drain petcock with tube to avoid oil loss - air drain petcock also used for oil (re)fill - defined oil volume - smart lubrication concept - measuring oil temperature at oil	- current gears, bearings, sealing - connected insert for bearing seat - glass fiber organosheet - screwed flanges (hosuing/cover) - 2 inserts for each screw - dowel pin for housing/cover centering - engine screwed to housing, centering by surface - simple air drain petcock with tube to avoid oil loss - oil fill plug at maximaum oil level - smart lubrication concept - measuring oil temperature at oil drain plug	- current gears, bearings, sealing - metal insert for bearing seat - glass fiber organosheet - screwed flanges (hosuing/cover) - 2 inserts for each screw - dowel pin for housing/cover centering - engine screwed to housing, centering by surface - simple air drain petcock with tube to avoid oil loss - air drain petcock also used for oil (re)fill - smart lubrication concept - measuring oil temperature at oil drain plug - measuring bearing temperature at drilled holes

Figure 4: Final rating of the main three concepts investigated.

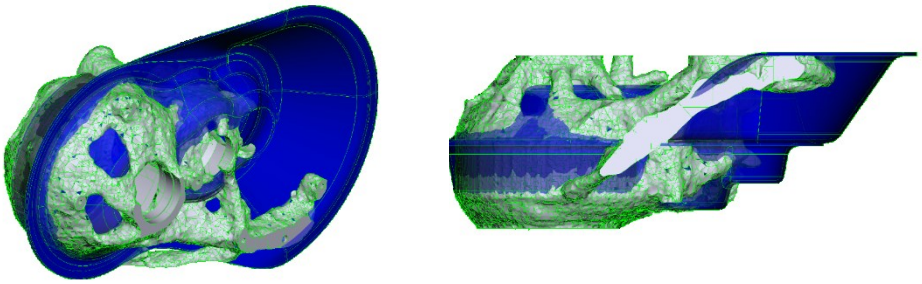


Figure 5: Results of topology optimization and derived geometry.

To derive a new housing geometry and a suitable fiber layout, optimization simulations were set up. The design space of the interior parts was designed due to the rules for gearbox housing design. In the topology optimization both load cases, drive and coast were used. Strain energy was optimized with weight, axis deviation and the axis inclination error as constraints. The main load paths for the stiffness performance of the gearbox housing were derived.

The optimization derived a geometry for both parts of the gearbox. This geometry was used to build up a simulation model for the detailed design. For all following simulations and optimization we concentrated on the cover. The constraints coming from production were already implemented in this design. These were mainly maximum steps in the cover, minimum radii and a maximum wall thickness of the organo sheet of 5mm.

After setting up this geometry, the layer thickness in the organo sheet was optimized. The optimization showed, that +- 45° layers should be the thickest ones. This is in good correlation with the fact, that the torsion in the housing is the main deformation causing the axis inclination error. As the stiffness derived from the optimization of the organo sheet was not satisfying, crossing UD Tapes and ribs were introduced into the geometry.

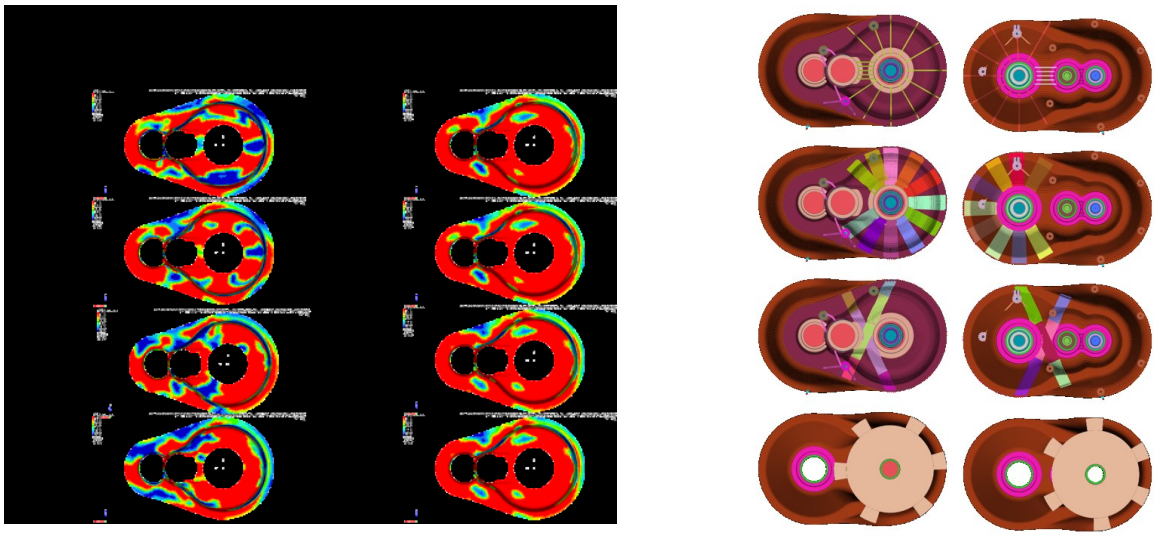


Figure 6: Left: Results of layer optimization. Right Different models for investigating crossing UD tapes.

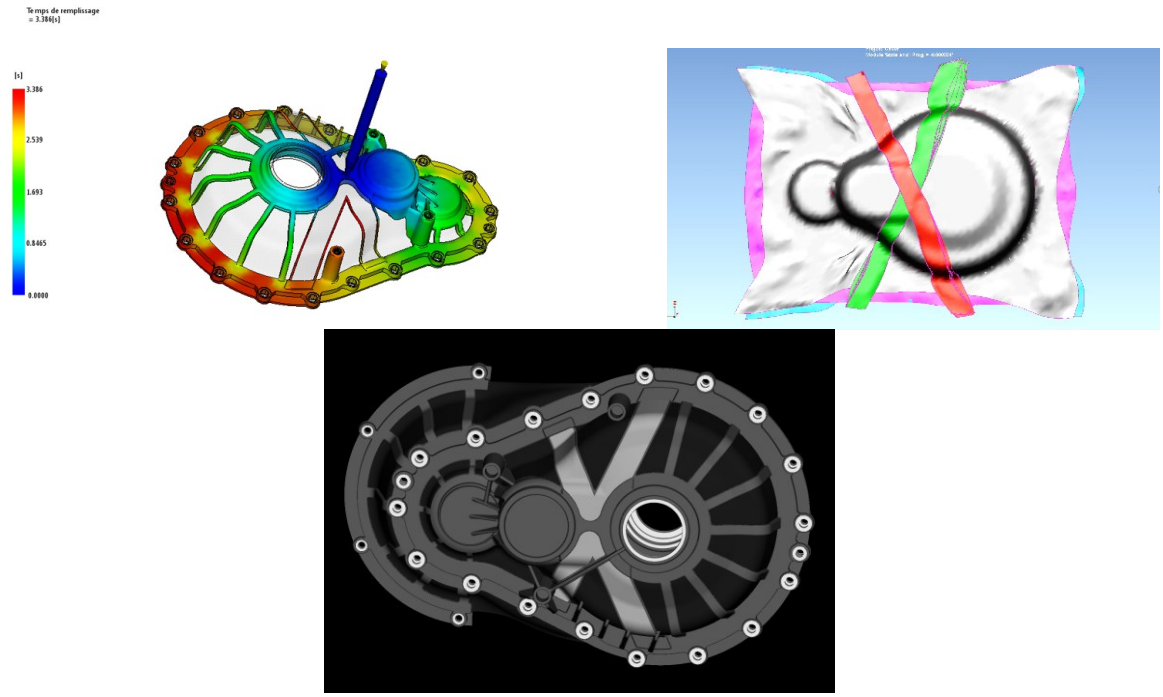


Figure 7: Top left: Simulation of over molding, top right: stamping simulation. Bottom: final design of the cover

Detailing phase:

The detailing phase added all details of functional fixing points and connections. For the connection of the two housing parts inserts were introduced. The flange, the centering pins and the functional faces have been detailed.

To make sure, not to have problems concerning the stamping process and to derive a starting cut for the raw organo sheet before stamping, a stamping simulation was performed with the help of the company ESI. To secure the overmoulding process a molding simulation was performed.

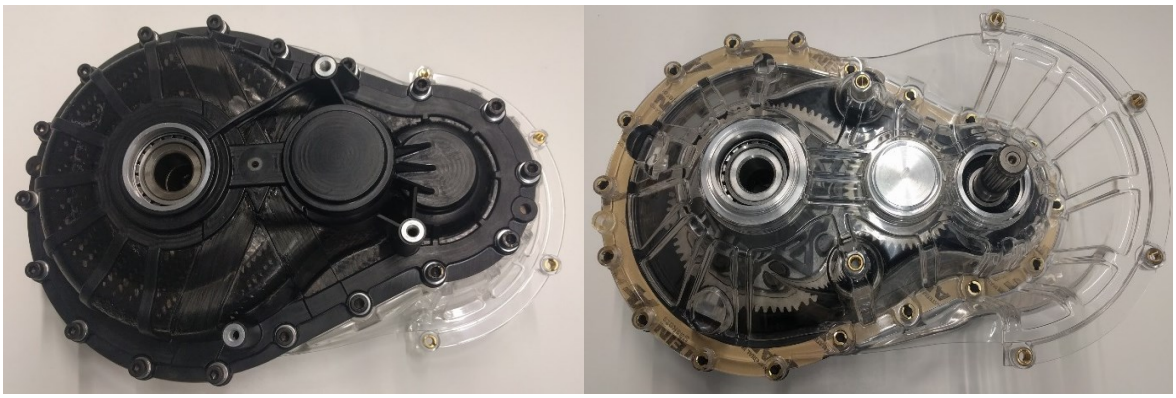


Figure 8: Finished and assembled prototype with PMMA part replacing the second housing part.

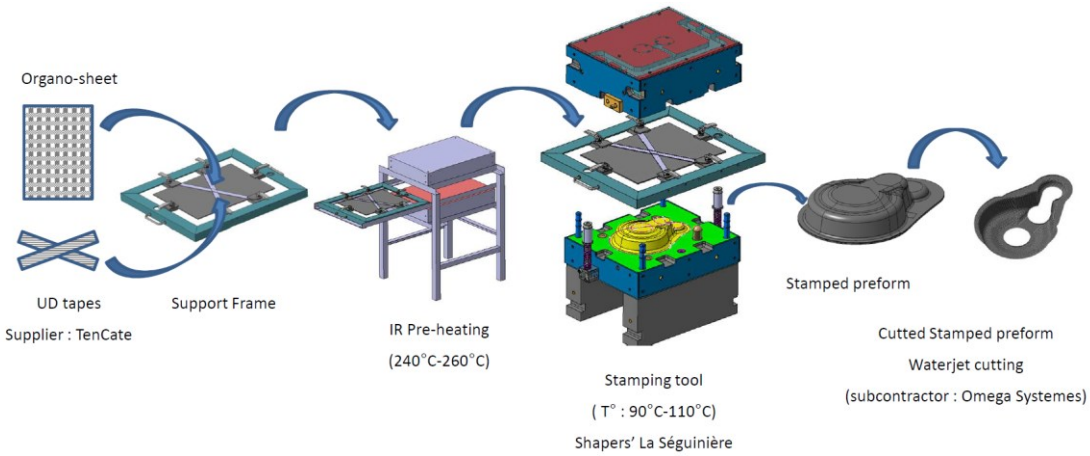


Figure 9: Production process for housing part.

6. Production of the prototype

The production of the prototype was done in four steps:

- Heating of the organo sheet together with the UD tapes using an IR heating device. Both components, organo sheet and UD tapes, are fixed in a frame that also keeps the UD tapes in the right position.
- Stamping of the heated organo sheet and UD tapes in a preheated stamping tool. After stamping, water cut to end shape.
- Overmoulding the pressed Organo sheet after heating the pressed Organo sheet. In the over molding step, the centering sleeves, the inserts for the screws and the bearing seats are inserted into the mold.
- Milling process of the flanges and the bearing seats to keep the required tolerances on the part.

7. Conclusion

Aluminum as material for a gearbox housing was successfully replaced by fiber reinforced thermoplastic material, CFRTP, with cross UD tapes stiffeners. After a milling step of the bearing seats, introduced in an over molding step, all tolerances are fulfilled. The performance of the gearbox cover is sufficient compared to the original aluminum gearbox. The results of the simulations show clearly that the torsional deformation of the gearbox is dominant in causing axis inclination errors. The feasibility of the production process and industrialization of the production was shown.

The weight achieved using the new material is 4kg compared to 5.8kg from the aluminum gearbox, what means about 30% saving. The cost of the cover was estimate between 50 and 80 Euro. The most expensive component is the organo sheet. The cycle time of pressing is about five seconds, the over molding is about 2 minutes. Water jet cutting can be removed by determining the stack outline in a way that after stamping no water jet cutting is necessary and the injection molding can take place at the same time.

The next step will be the production of the second half of the gearbox housing and stiffness measurements of the complete gearbox. Additionally organo sheet with PPA or PPS matrix will be introduced, what will lead to higher performance of the housing at high temperatures and reduced organo sheet thickness.

References

[1] Kissoft, (03/2016) User Manual. Bubikon, Switzerland
[2] Fritz Zwicky (1959): Morphologische Forschung. Winterthur, New. Glarus: Baeschlin, 1989

Analytical and experimental investigation on crack generated in diameter-enlargement section

†*Xia Zhu[1], Nagatoshi Okabe[1], Keiji Ogi[1] and Hiromichi Toyota[1]

[1]Department of Mechanical Engineering, Ehime University, Japan

*Presenting author: zhu.xia.mx@ehime-u.ac.jp

†Corresponding author: zhu.xia.mx@ehime-u.ac.jp

Abstract

We have proposed a new cold processing method to enlarge the diameter of a short section of a metal shaft using a combination of a cyclic bending load and an axial compressive load that is lower than the yield stress of the sample material. We call this cold processing method the diameter-enlargement working method, and refer to the enlarged section of the processed shaft as the diameter-enlargement section. The processing method easily produces large plastic deformation in the processed section under a low axial compressive load at room temperature. However, a crack is sometimes generated in the stepped section during processing. Therefore, we conducted processing experiments to clarify the crack generation conditions, and simulated the working process using the finite element method to investigate the behaviors of stress and strain during processing. Based on the experiments and analyses, we determined that a fatigue crack is generated because of cyclic axial normal stress in the root of stepped section at the axial-compressive loading side. The experimental and analytical fatigue strength was well described by the Coffin-Manson expression.

Keywords: Crack, Diameter-enlargement working method, Low cycle fatigue damage, Finite element method

Introduction

We have proposed a new cold processing method to enlarge the diameter of a short section of a metal shaft using a combination of a cyclic bending load and an axial compressive load that is lower than the yield stress of the sample material. We call this cold processing method the diameter-enlargement working method, and refer to the enlarged section of the processed shaft as the diameter-enlargement section [1-3]. The key features of this processing method are as follows: First, the diameter-enlargement deformation progresses easily under a low axial compressive load at room temperature through the Bauschinger effect (also referred to as the mechanical ratchet phenomenon) arising from alternate stresses caused by the cyclic bending load during processing. Second, although the processing causes large plastic deformation, the processed shaft exhibits little temperature increase. Finally, material wastage does not occur because no waste particles are generated as in cutting operations. We have previously clarified the influence of the processing conditions, such as the axial-compressive load, bending angle, and rotating speed, as well as the mechanical properties of the sample materials in terms of diameter-enlargement deformation behavior [4-7]. From this previous research, it was determined that the plastic deformation progresses as the rotating speed increases. However, crack generation occurs at the notch root near the diameter-enlargement section according to the processing conditions that causes the test specimen to break for reasons that are not yet clear. Therefore, this present study experimentally and analytically investigated the conditions of crack generation at the notch root. First, we conducted processing experiments to investigate the point in the process where crack initiation occurs, thereby evaluating the limiting number of rotations for crack initiation for different processing conditions. Second, we conducted low-cycle fatigue tests to obtain a repeated stress-strain curve, which was applied to elasto-plastic

Table 1 Mechanical properties of the sample material

Young's modulus E (GPa)	210
Yield stress σ_y (MPa)	580
Tensile strength σ_b (MPa)	735
Percentage reduction of area φ (%)	37.6

Table 2 Processing conditions

Axial-compressive force P (kN) (Normalized stress σ_c/σ_y)	46, 69, 91 (1.0, 1.5, 2.0)
Bending angle θ (degree)	2, 3
Rotating speed ω (rpm)	60
Radius of notch root r (mm)	2

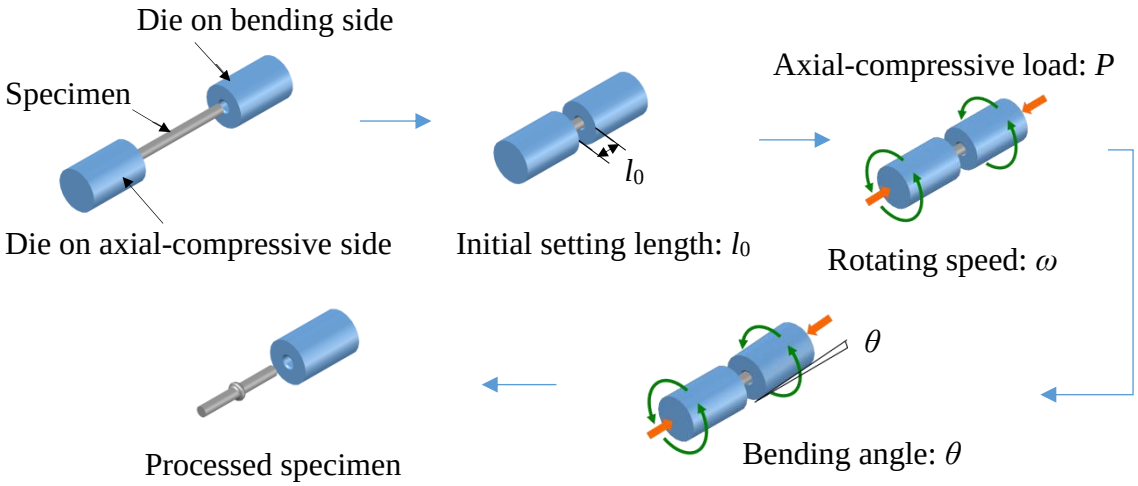


Fig. 1 Diameter-enlargement working procedures

numerical analyses. Using the finite element method (FEM), we simulated the stress and strain behaviors in the processed shaft in the vicinity of the diameter-enlargement section during processing, and calculated an elasto-plastic stress concentration factor, an elasto-plastic strain concentration factor, and a fatigue strength reduction factor. Finally, we evaluated the fatigue strength of the processed shaft using the Manson-Coffin expression.

Processing experiments and finite element analyses for calculating stresses

Experiments

Experiments were done by using the machine developed originally, Cold drawing steel SS400 (Japan Industrial Standard) was used for the experiments. Its material properties are described in Table 1. Table 2 lists the processing conditions, Fig. 1 illustrates the experimental procedures. First, a smooth specimen is coaxially placed between the bending side and the axial-compressive side dies with the distance defined as an initial setting length l_0, and an axial-compressive force P is loaded on the test specimen. Next, keeping the load P constant, the specimen is rotated with a rotating speed ω while simultaneously setting a bending angle θ on

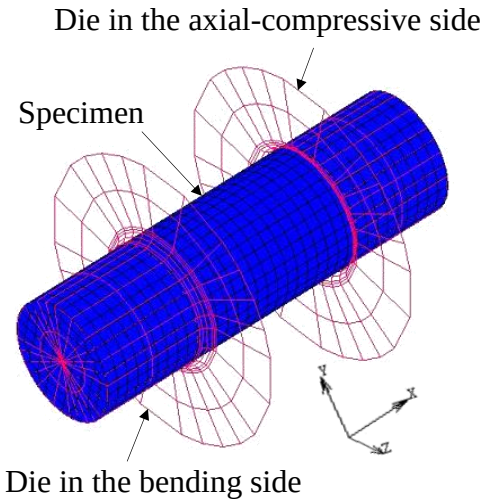


Fig. 2 Analysis model

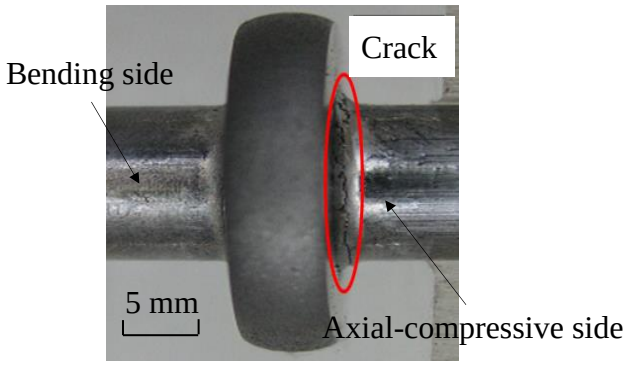


Fig. 3 Processed specimen exhibiting crack formation near the stepped section

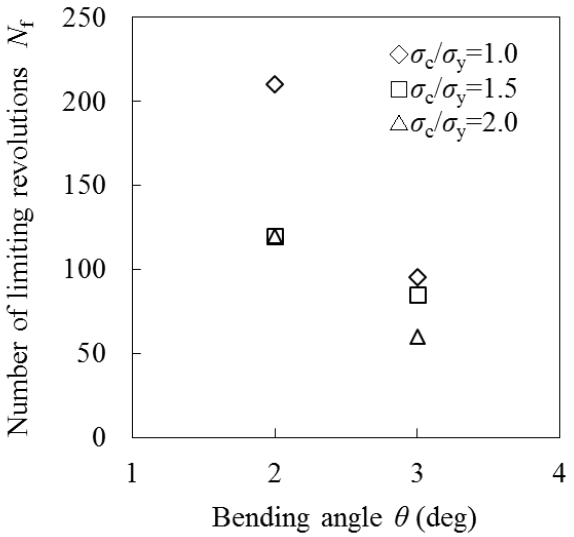


Fig. 4 Number of limiting revolutions under each processing condition

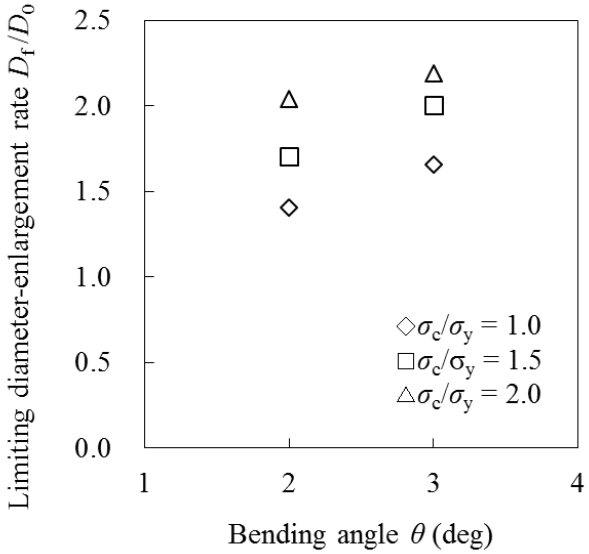


Fig. 5 Limiting diameter-enlargement rates under each processing condition

the specimen. When the diameter of the diameter-enlargement section reaches the target diameter, the application of the bending angle and rotation is stopped.

Analyses

Finite element analyses were carried out for simulating the processing. Fig. 2 shows a three-dimensional model with the same dimension as the specimen using eight node isoparametric elements. The specimen is modeled as a deformable body, whereas the dies are rigid bodies. Number of nodes is 6240 and number of elements is 5504. Type of contact interaction between specimen and die is touching. The friction force is calculated based on the shear friction rule. The coefficient of friction is assumed to be 0.15. A bending angle θ and a rotating speed ω are applied to the bending die, and an axial compressive force P and the same rotational speed ω are applied to the pressure die. The von Mises yield criterion, kinematic hardening law

considering the Bauschinger effect, and Prandtl-Reuss's flow rule are adopted. The relation between the actual stress σ and the plastic strain ε_p, obtained by the low-fatigue tests, is expressed by Eq. 1:

$$\sigma = K{\varepsilon_p}^n \tag{1}$$

where K is a material constant set at 1100, and n is the cycle strain hardening factor with a value of 0.25. FEM analyses were performed under the same loading conditions as the processing experiments using commercial nonlinear finite element software (MSC. Marc Mentat 2013.0.0).

Experimental and analytical results and discussion

Process limitations

A perpendicular crack in the axial direction occurs near a notch root on the axial-compressive load side, and the specimen breaks owing to compressive deformation when the rotating speed is increased. Fig. 3 shows an image of the resulting crack formation. We define the number of revolutions prior to when a crack occurs as the number of limiting revolutions N_f, and this is shown in Fig. 4 for the various processing conditions considered (Table 2). The diameter-enlargement rate at the time of crack generation is defined as the limiting diameter-enlargement rate D_f /D_0, and this is shown in Fig. 5 under the various processing conditions. The limiting diameter-enlargement rate increases with increasing bending angle and axial-compressive loading, but the number of limiting revolutions decreases. The main reason is that diameter-enlargement deformation advances fast because of the Bauschinger effect when the bending angle increases. And the alternating stress amplitude of tension-compression, loaded on the processed part, increases, so it becomes easy to generate the fatigue failure, number of limiting revolutions becomes small.

Behaviors of stresses and strains during processing

Figure 7 shows the axial stress behavior, the largest stress component near the notch root on the bending side is the axial normal stress σ_x. And Fig. 8 shows the strain distribution during processing, the axial strain ε_x the axial-compressive loading side changes from compression deformation to tensile deformation after 40 revolutions. On the other hand, the axial strain ε_x on the bending side represents compression deformation during the processing. Therefore, based on the results, a fatigue crack is generated because of the cyclic axial normal stress in the

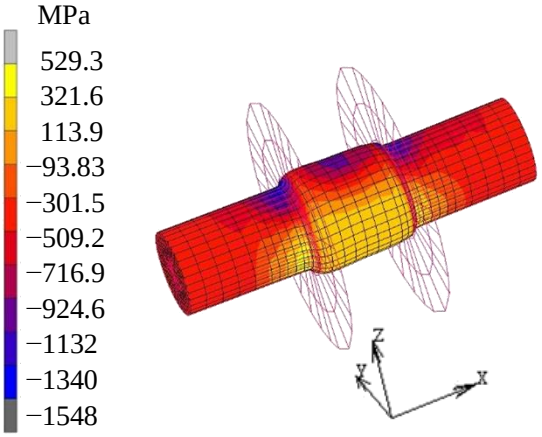


Fig. 7 Axial normal stress (σ_x) distribution under P = 46 kN, θ = 2 degree, and N = 20

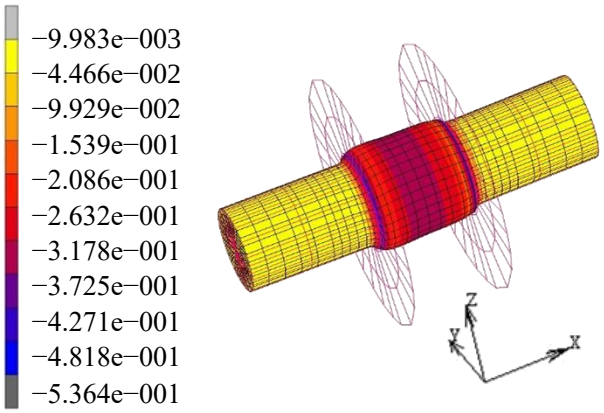


Fig. 8 Axial normal strain (ε_x) distribution under P = 46 kN, θ = 2 degree, and N = 20

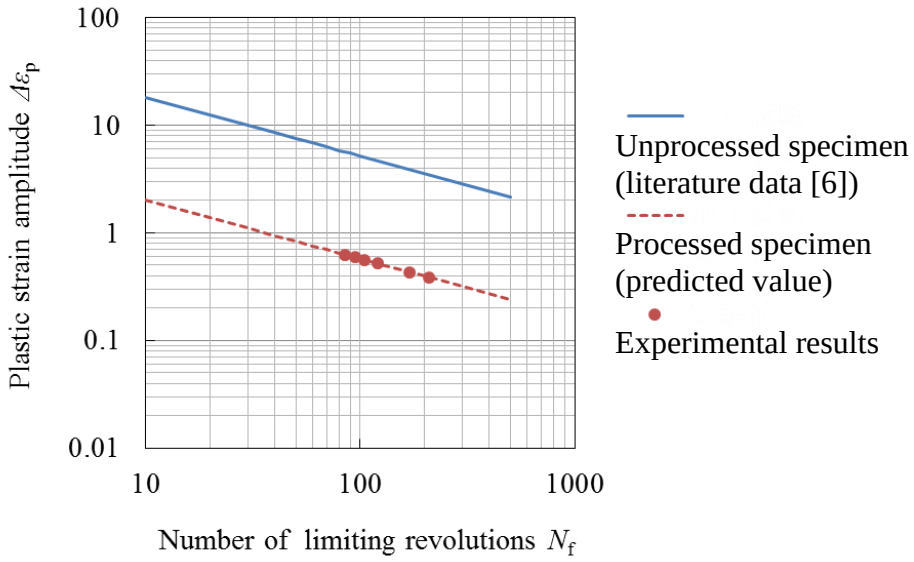


Fig. 9 Fatigue curve.

root of the axial-compressive loading side and not the bending side. The crack develops perpendicular to the specimen axis, and the test specimen breaks. This estimation is demonstrated by the experimental results as well.

Fatigue strength of the processed shaft

The solid line in Fig. 9 shows the fatigue strength of an unprocessed specimen, and the dashed line and filled symbols show the analytical and experimental results, respectively. These plastic strain amplitudes $\Delta\varepsilon_p$ can be expressed well by the Coffin-Manson law [8], as given by Eq. 2:

$$\Delta\varepsilon_p\left(N_f{}'k_f\right)^{K_p} = C_p \tag{2}$$

where k_p is the fatigue ductility index with a value of 0.54 [6], C_p is the fatigue ductility coefficient with a value of 62.2 [6]. k_f is the fatigue strength reduction factor with a value of 5.8, calculated from the elastic-plastic stress concentration coefficient and the elastic-plastic strain concentration factor based on FE analyses, and $N_f{}'$ is the fatigue strength of the processed specimen and calculated on the basis of the minor rule. The predicted values shown in Fig. 9 correspond very well to the experimental values.

Conclusions

The present study has clarified the mechanism of fatigue crack generation during the diameter-enlargement method processing, and evaluated the conditions of fatigue damage of a processed shaft using the fatigue strength reduction factor. We determined that a fatigue crack is generated because of cyclic axial normal stress in the notch root at the axial-compressive loading side. The experimental and analytical fatigue strength was well described by the Coffin-Manson expression.

References

[1] Iura, T., Mori, K., Yamamoto, K., Aono, Y., Kurita, K., Nakamura, O., and Sano, T., LOCAL COLLAR DEVELOPED BY METAL PLASTIC DEFORMATION, Proc. 6th ICTP, (1999)613-618.
[2] Okabe, N., Zhu, X., Iura, T., and Mori, K., Model Analysis for a Clod Working of Collar-forming in a Metal Bar with Local Superplastic Deformation by a New Conceptual Technology, Proc. 7th ICTP, (2002) 871-976.

[3] Iura, T., Okabe N., and Zhu, X., Development of Novel Plastic Working Process for Producing a Collar on a Round Shaft, Journal of the JSTP, Vol.44, No.514, (2003) 45-49.
[4] Iura, T., Okabe N., and Zhu, X., Generation of Bending Moment and Twist Torque and Temperature Increase during the Working Process for the Production of a Collar in a Round Bar, Journal of the JSTP, Vol.45, No.516, (2004) 35-39.
[5] Fatigue Properties and Fatigue Damage in a Round Shaft with a Collar Formed by a Novel Processing Method for Enlarging Diameter, Journal of the JSTP, Vol.45, No.516, (2004) 35-39.
[6] Iura, T., Okabe N., and Zhu, X., Influence of Strength Properties on Diameter Increase for Processed Material during the Process of Enlarging a Partial Diameter, Journal of the JSTP, Vol.46, No.531, (2005) 55-59.
[7] Xia Zhu, Nagatoshi Okabe, Keiji Ogi and Manabu Takahashi, Fatigue Strength of Shaft with Diameter Enlarged Partially by Cyclic Bending and Axial Compressive Loading, Applied Mechanics and Materials, Vols.217-219, (2012) 2346-2350.
[8] The Japan Society of Mechanical Engineers, JSME Data Book: Fatigue of Metals, 4, Low Cycle Fatigue Strength, Maruzen Publishers, (1983) 9-15, 147, 179.

A general way to construct a new optimal scheme with eighth-order convergence for nonlinear equations

†R. Behl[1], Changbum Chun[2], Ali Saleh Alshormani[3] and S.S. Motsa[1,4]

[1]Department of Mathematics, Statistics and Computer science, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa
[2]Department of Mathematics, Sungkyunkwan University, Suwon 16419, Republic of Korea
[3]Department of Mathematics, King Abdulaziz University, Jeddah, Saudi Arabia
[4]Mathematics Department, University of Swaziland, Private Bag 4, Kwaluseni, M201, Swaziland
†Corresponding author: ramanbehl87@yahoo.in

Abstract

In this paper, we present a new and interesting optimal scheme of order eight in a general way for solving nonlinear equations, numerically. The construction of the scheme is based on rational function approach. The beauty of the proposed is that it is capable to produce further new and interesting optimal schemes of order eight from every existing optimal fourth-order scheme whose first substep employs Newton's method. The theoretical and computational properties of the proposed scheme are fully investigated along with main theorem which establishes the order of convergence and asymptotic error constant. Several numerical examples are given and analyzed in detail to demonstrate faster convergence and high computational efficiency of the proposed methods.

Keywords: Nonlinear equations, Simple roots, Computational order of convergence, Newton's method.

Introduction

With the advancement of digital computer, advanced computer arithmetics and symbolic computation, a special attention has been paid to the development of optimal eighth-order iterative methods in the past two decades. The merit of these methods is that they converge fast towards a sought root. Moreover, we can reach our desired accuracy in a very small number of iterations.

The researchers from the world wide proposed a large number of optimal eighth-order methods [3, 4, 7, 8, 9, 10, 11, 15, 16, 19, 20, 21, 23, 24]. Most of them are extensions of Newton's method or Newton-like method or any particular existing method like Ostrowski's method, King's method, King-type method, etc. at the expense of additional functional evaluations or increased number of substeps of the original methods. However, there still are a few number of optimal schemes which are applicable to every existing iterative method of particular order to further obtain higher-order methods.

In the recent years, Sharma et al. [20] have given an optimal eighth-order scheme in a general way, which is applicable to every optimal fourth-order method whose first substep is Newton's to further extend eighth-order convergence. But, it should be noted that they provided the third substep in their scheme without any justification. Optimal schemes applicable to any fourth-order iterative scheme with full justification of the development are more interesting and challenging task in the field of numerical analysis.

For the construction of a new iterative scheme, it is quite often used to approximate the functions or derivatives of the involved function. In the available literature, we have several kinds of approximations for e.g. Functional approach, Sampling approach, Geometric approach, Weight function approach, Adomain approach, Composition approach and Rational function approach.

Every approach has some advantages and disadvantages because it's dependent on the problem under consideration. The choice of suitable approximation approach not only produce simple and interesting schemes but also can save considerable amount of computation. Rational function approach is one of the most important techniques in numerical analysis for approximating the function or to find the next approximation.

In general, the number of tangency conditions are equal to number of undetermined constants. Further, we will get an improved method with higher-order convergence as we increase the number of undetermined constants in the rational function (for the details, see Jarratt and Nudds [26]).

The principle aim of this study is to present a new and interesting optimal scheme of order eight in general way instead of like earlier study [3, 4, 7, 8, 9, 10, 11, 15, 16, 19, 20, 21, 23, 24], where researchers proposed some eighth-order extensions of some particular methods like Ostrowski' method or King's family or Ostrowski-type, etc. Our proposed scheme is applicable on every existing optimal fourth-order scheme (which can be chosen from the available literature) whose first substep employs Newton's method to produce further new and interesting optimal eighth-order scheme. We construct this scheme with the help of rational approximation approach. In order to check the effectiveness and validity of our study, we compare them with the existing methods of same order on a concrete variety of nonlinear functions. From the numerical experiments, it is observed that our proposed methods perform better than existing ones.

Development of eighth-order optimal schemes

This section is devoted to the construction of an optimal and interesting eighth-order scheme in general way. Therefore, we consider a general fourth-order scheme whose first substep is classical Newton's method in the following way:

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = \psi_4(x_n,\ y_n). \end{cases} \tag{1}$$

In order to obtain the next iteration and eighth-order convergence, we simply apply the classical Newton's method, which is given as follows:

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}. \tag{2}$$

Since, the above scheme uses five functional evaluations. So, this scheme can not be optimal in the sense of Kung-Traub conjecture [15]. However, we can reduce the number of functional evaluations by introducing a rational function $\eta(x)$, which is defined as follows:

$$\eta(x) = \eta(x_n) + \frac{(x - x_n) + \alpha_1}{\alpha_2(x - x_n)^2 + \alpha_3(x - x_n) + \alpha_4}, \tag{3}$$

where $\alpha_i (1 \leq i \leq 4)$ are the disposable parameters. We will determine the values of these parameters with the help of following tangency conditions

$$\eta(x_n) = f(x_n),\ \eta'(x_n) = f'(x_n),\ \eta(y_n) = f(y_n),\ \eta(z_n) = f(z_n). \tag{4}$$

With the assumption of one more tangency condition $\eta'(z_n) = f'(z_n)$, we will obtain the last substep in the following way

$$x_{n+1} = z_n - \frac{f(z_n)}{\eta'(z_n)}, \tag{5}$$

which no longer requires the evaluation of $f'(z_n)$.

Now, with the help of first two tangency conditions, we will yield

$$\alpha_1 = 0,\ \alpha_4 = \frac{1}{f'(x_n)}. \tag{6}$$

Again, by using the last two tangency conditions and the above values of α_1 and α_4, we will obtain the following two linear independent equations

$$\begin{aligned} \alpha_2(y_n - x_n) + \alpha_3 &= \frac{1}{(y_n - x_n)}\left[\frac{1}{f[y_n,\ x_n]} - \frac{1}{f'(x_n)}\right], \\ \alpha_2(z_n - x_n) + \alpha_3 &= \frac{1}{(z_n - x_n)}\left[\frac{1}{f[z_n,\ x_n]} - \frac{1}{f'(x_n)}\right], \end{aligned} \tag{7}$$

which further yield

$$\begin{aligned} \alpha_2 &= \frac{f(x_n)f[y_n,\ x_n]f[z_n,\ x_n] - f'(x_n)f[y_n,\ x_n](f(x_n) + f[z_n,\ x_n](x_n - z_n)) + f'(x_n)^2 f[z_n,\ x_n](x_n - z_n)}{f(x_n)f[y_n,\ x_n]f[z_n,\ x_n](x_n - z_n)(f(x_n) + f'(x_n)(z_n - x_n))}, \\ \alpha_3 &= \frac{f'(x_n)\left(\frac{f(x_n)(f[z_n,\,x_n] - f'(x_n))}{f'(x_n)^2 f[z_n,\,x_n](x_n - z_n)} + \frac{(f'(x_n) - f[y_n,\,x_n])(x_n - z_n)}{f(x_n)f[y_n,\,x_n]}\right)}{f(x_n) + f'(x_n)(z_n - x_n)}, \end{aligned} \tag{8}$$

where $f[y_n,\ x_n] = \frac{f(y_n) - f(x_n)}{y_n - x_n}$ and $f[z_n,\ x_n] = \frac{f(z_n) - f(x_n)}{z_n - x_n}$ are divided difference of order one. With the help of expression (3), we can easily obtain

$$\eta'(z_n) = \frac{\alpha_4 - (z_n - x_n)^2\alpha_2}{\left[(z_n - x_n)^2\alpha_2 + (z_n - x_n)\alpha_3 + \alpha_4\right]^2}. \tag{9}$$

Finally, by using the expressions (1), (5) and (9), we obtain

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = \psi_4(x_n,\ y_n), \\ x_{n+1} = z_n - \dfrac{f(z_n)\left[(z_n - x_n)^2\alpha_2 + (z_n - x_n)\alpha_3 + \alpha_4\right]^2}{\alpha_4 - (z_n - x_n)^2\alpha_2}, \end{cases} \tag{10}$$

where $\alpha_2,\ \alpha_3$ and a_4 are defined earlier in this section. The following Theorem 1 demonstrates three important things: first one is related to optimal eighth-order convergence without using any additional functional evaluations; second one is how a rational function $\eta(x)$ plays a vital role in the construction of iterative scheme (10) in a general way; third one is how a single coefficient B_1 in $\psi_4(x_n,\ y_n)$ contributes to its role in the construction of the desired asymptotic error constant.

Theorem 1 *Let $f : \mathbb{R} \to \mathbb{R}$ be a sufficiently differentiable function in an interval containing ξ, where ξ is a simple zero of the involved function. In addition, we assume that $\psi_4(x_n,\ y_n)$ is any optimal fourth-order scheme whose first sub step employs Newton's method. Moreover, we*

consider initial guess $x = x_0$ is sufficiently close to ξ for guaranteed convergence. Then, the proposed scheme (10) *has an optimal eighth-order convergence.*

Proof: Let us consider that $e_n = x_n - \xi$ be the error at nth term. The Taylor's series expansion of the function $f(x_n)$ and $f'(x_n)$ around $x = \xi$ with the assumption $f'(\xi) \neq 0$ leads us to:

$$f(x_n) = f'(\xi)[e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + c_5e_n^5 + c_6e_n^6 + c_7e_n^7 + c_8e_n^7 + O(e_n^9)] \tag{11}$$

and

$$f'(x_n) = f'(\xi)[1+2c_2e_n+3c_3e_n^2+4c_4e_n^3+5c_5e_n^4+6c_6e_n^5+7c_7e_n^6+8c_8e_n^7+9c_9e_n^8+O(e_n^9)], \tag{12}$$

respectively, where $c_k = \frac{f^{(k)}(\xi)}{k!f'(\xi)}$ for $k = 2,\ 3,\ \ldots,\ 8$.

With the help of above expressions (11) and (12) in the first substep, we obtain

$$\begin{aligned} y_n - \xi = {} & c_2e_n^2 + (2c_3 - 2c_2^2)e_n^3 + (4c_2^3 - 7c_3c_2 + 3c_4)e_n^4 + (20c_3c_2^2 - 8c_2^4 - 10c_4c_2 \\ & - 6c_3^2 + 4c_5)e_n^5 + \Big\{16c_2^5 - 52c_3c_2^3 + 28c_4c_2^2 + (33c_3^2 - 13c_5)c_2 - 17c_3c_4 + 5c_6\Big\}e_n^6 \\ & - 2\Big\{16c_2^6 - 64c_3c_2^4 + 36c_4c_2^3 + 9(7c_3^2 - 2c_5)c_2^2 + (8c_6 - 46c_3c_4)c_2 - 9c_3^3 + 6c_4^2 \\ & + 11c_3c_5 - 3c_7\Big\}e_n^7 + \Big\{64c_2^7 - 304c_3c_2^5 + 176c_4c_2^4 + (408c_3^2 - 92c_5)c_2^3 + (44c_6 \\ & - 348c_3c_4)c_2^2 + (118c_5c_3 - 135c_3^3 + 64c_4^2 - 19c_7)c_2 + 75c_3^2c_4 - 31c_4c_5 - 27c_3c_6 \\ & + 7c_8\Big\}e_n^8 + O(e_n^9). \end{aligned} \tag{13}$$

Again, we obtain the following expansion of $f(y_n)$ about a point $x = \xi$ with the help of Taylor series

$$\begin{aligned} f(y_n) = f'(\xi)\Big[& c_2e_n^2 + (2c_3 - 2c_2^2)e_n^3 + (5c_2^3 - 7c_3c_2 + 3c_4)e_n^4 - 2(6c_2^4 - 12c_3c_2^2 \\ & + 5c_4c_2 + 3c_3^2 - 2c_5)e_n^5 + \{28c_2^5 - 73c_3c_2^3 + 34c_4c_2^2 + (37c_3^2 - 13c_5)c_2 + 5c_6 \\ & - 17c_3c_4\}e_n^6 - 2\{32c_2^6 - 103c_3c_2^4 + 52c_4c_2^3 + (80c_3^2 - 22c_5)c_2^2 + (8c_6 - 52c_3c_4)c_2 \\ & - 9c_3^3 + 6c_4^2 + 11c_3c_5 - 3c_7\}e_n^7 + O(e_n^8)\Big]. \end{aligned} \tag{14}$$

By using the expression (11), (13) and (14), we have

$$\begin{aligned} \frac{f(y_n) - f(x_n)}{y_n - x_n} = {} & 1 + c_2e_n + (c_2^2 + c_3)e_n^2 + (3c_3c_2 - 2c_2^3 + c_4)e_n^3 + (4c_2^4 - 8c_3c_2^2 + 4c_4c_2 \\ & + 2c_3^2 + c_5)e_n^4 + \{20c_3c_2^3 - 8c_2^5 - 11c_4c_2^2 + (5c_5 - 9c_3^2)c_2 + 5c_3c_4 + c_6\}e_n^5 \\ & + \{16c_2^6 - 48c_3c_2^4 + 29c_4c_2^3 + (31c_3^2 - 14c_5)c_2^2 + 6(c_6 - 4c_3c_4)c_2 - 2c_3^3 \\ & + 3c_4^2 + 6c_3c_5 + c_7\}e_n^6 + O(e_n^7). \end{aligned} \tag{15}$$

Since, $\psi_4(x_n,\ y_n)$ is an optimal fourth-order scheme. So, it will satisfy the error equation of the following form

$$z_n - \xi = B_1e_n^4 + B_2e_n^5 + B_3e_n^6 + B_4e_n^7 + B_5e_n^8 + O(e_n^9), \tag{16}$$

where $B_1 \neq 0$.

Now, we can expand the function $f(z_n)$ about a point $z = \xi$ with the help of Taylor series expansion, which is given as follows

$$f(z_n) = f'(\xi)\left[B_1 e_n^4 + B_2 e_n^5 + B_3 e_n^6 + B_4 e_n^7 + (B_1^2 c_2 + B_5) e_n^8 + O(e_n^9)\right]. \tag{17}$$

By using the expression (11), (16) and (17), we obtain

$$\begin{aligned}\frac{f(z_n) - f(x_n)}{z_n - x_n} &= 1 + c_2 e_n + c_3 e_n^2 + c_4 e_n^3 + (B_1 c_2 + c_5) e_n^4 + (B_2 c_2 + B_1 c_3 + c_6) e_n^5 \\ &\quad + (B_3 c_2 + B_2 c_3 + B_1 c_4 + c_7) e_n^6 + (B_4 c_2 + B_3 c_3 + B_2 c_4 + B_1 c_5 \\ &\quad + c_8) e_n^7 + O(e_n^8).\end{aligned} \tag{18}$$

Now, with the help of expressions (11) – (18), we further obtain

$$\begin{aligned}\frac{f(z_n)\left[(z_n - x_n)^2 \alpha_2 + (z_n - x_n)\alpha_3 + \alpha_4\right]^2}{\alpha_4 - (z_n - x_n)^2 \alpha_2} &= B_1 e_n^4 + B_2 e_n^5 + B_3 e_n^6 + B_4 e_n^7 \\ &\quad - B_1 c_2 (B_1 + c_2^3 - 2 c_2 c_3 + c_4) e_n^8 + O(e_n^9).\end{aligned} \tag{19}$$

Finally, by inserting the expressions (16) and (19) in the last substep of the proposed scheme (10) and after some simplification, we obtain

$$e_{n+1} = B_1 c_2 (B_1 + c_2^3 - 2 c_2 c_3 + c_4) e_n^8 + O(e_n^9), \tag{20}$$

This completes the proof. □

Remark 2 *The above asymptotic error constant* (20) *reveals that the proposed scheme* (10) *attains an optimal eighth-order convergence in the sense of Kung-Traub conjecture. In addition, one generally expects that the asymptotic error constant of the proposed scheme* (10) *also contains some constants namely,* $c_2, c_3, c_4, c_5, c_6, c_7, c_8$ *and* B_1, B_2, B_3, B_4, B_5*. However, only* $B_1,\ c_2,\ c_3$ *and* c_4 *appears in the asymptotic error constant which can be seen in* (20)*. This simplicity clearly reflects that our current rational function approach with the tangency conditions which is used for the reduction of functional evaluations, plays a vital role in the development of an optimal eighth-order method.*

Numerical experiments

In this section, we will check the effectiveness and validity of our theoretical results which we have proposed in Section 2. For this purpose, we shall consider a concrete variety of nonlinear equations, which are given as follows:

$f_1(x) = e^x \sin(x) + \log(x^2 + 1)$;[5]	$\xi = 0$
$f_2(x) = x^6 - x^4 - x^3 - 1$; [23]	$\xi = 1.40360212487421664327913855768$
$f_3(x) = e^x - 4x^2$; [16]	$\xi = 0.71480591236277780613762220811 2$
$f_4(x) = \tan^{-1}(x) - x + 1$; [1]	$\xi = 2.13226772527288513162542069694$
$f_5(x) = e^{-x} + \cos(x)$; [19]	$\xi = 1.74613953040801241765070308895$
$f_6(x) = \log x$;[17]	$\xi = 1$

First of all, we shall verify the theoretical order of convergence of the proposed methods on the basis of the results obtained from $\left|\frac{x_{n+1} - x_n}{(x_n - x_{n-1})^8}\right|$ and computational order of convergence. In Table

1, we displayed the number of iteration indexes (n), approximated zeros (x_n), absolute residual error of the corresponding function $(|f(x_n)|)$, error in the consecutive iterations $|x_{n+1} - x_n|$, $\left|\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^8}\right|$, the asymptotic error constant $\eta = \lim_{n\to\infty}\left|\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^8}\right|$ and computational order of convergence (ρ). In order to calculate the computational order of convergence (ρ), we use the following formula:

$$\rho = \left|\frac{(x_{n+1}-x_n)/\eta}{(x_n-x_{n-1})}\right|,\ n = 1, 2, 3.$$

We calculate the computational order of convergence, asymptotic error constant and other constants up to several number of significant digits (minimum 1000 significant digits) to minimize the round off error. But due to the limited paper space, we display the value of x_n and ρ up to 15 and 6 significant digits, respectively. In addition, we also display $\left|\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^8}\right|$ and η up to 10 significant digits. Moreover, absolute residual error in the function $|f(x_n)|$ and error in the consecutive iterations $|x_{n+1} - x_n|$ are displayed up to 2 significant digits with/without exponent power which are mentioned in Tables 1. Furthermore, the approximated zeros up to 30 significant digits are also displayed in Table 1 although minimum 1000 significant digits are available with us.

For the computer programming, all computations have been performed using the programming package $Mathematica\ 9$ with multiple precision arithmetic. Further, the meaning of $a(-b)$ is $a \times 10^{(-b)}$ in the following Tables 1.

Now, we want to see the comparison of our methods with the other existing optimal methods of same order. Therefore, we consider some special cases of the proposed scheme in the following way

(i) Let us consider the well known fourth-order King's family [13]. By using King's family in the proposed scheme, we obtain a new optimal eighth-order extension of King's family, which is defined as follows:

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \left[\dfrac{f(x_n)+\beta f(y_n)}{f(x_n)+(\beta-2)f(y_n)}\right]\dfrac{f(y_n)}{f'(x_n)},\ \beta \in \mathbb{R}, \\ x_{n+1} = z_n - \dfrac{f(z_n)\left[(z_n-x_n)^2\alpha_2 + (z_n-x_n)\alpha_3 + \alpha_4\right]^2}{\alpha_4 - (z_n-x_n)^2\alpha_2}. \end{cases} \tag{21}$$

For a computational point of view, let us consider $\beta = 0$ in the above scheme, called by $(OM1)$.

(ii) Now, we shall choose another optimal family of fourth-order methods proposed by Chun in [5]. Then, we obtain another new optimal family of eighth-order methods, which is

described as follows:

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \left[\dfrac{\{f(x_n)\}^2}{\{f(x_n)\}^2 - 2f(x_n)f(y_n) + 2\beta\{f(y_n)\}^2}\right]\dfrac{f(y_n)}{f'(x_n)}, \quad \beta \in \mathbb{R}, \\ x_{n+1} = z_n - \dfrac{f(z_n)\left[(z_n - x_n)^2\alpha_2 + (z_n - x_n)\alpha_3 + \alpha_4\right]^2}{\alpha_4 - (z_n - x_n)^2\alpha_2}. \end{cases} \tag{22}$$

Let us choose $\beta = \frac{1}{4}$ in the above scheme for computational experiments, known by $(OM2)$.

(iii) Again, we consider another optimal family of fourth-order methods proposed by Behl et al. [1]. With the help of our proposed scheme (9), we obtain the following optimal family of eighth-order methods

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \dfrac{f(x_n)}{f'(x_n)}\left[\dfrac{(b_1^2 + b_1b_2 - b_2^2)f(x_n)f(y_n) - b_1(b_1 - b_2)\{f(x_n)\}^2}{\Big(b_1f(x_n) - b_2f(y_n)\Big)\Big((2b_1 - b_2)f(y_n) - (b_1 - b_2)f(x_n)\Big)}\right], \\ x_{n+1} = z_n - \dfrac{f(z_n)\left[(z_n - x_n)^2\alpha_2 + (z_n - x_n)\alpha_3 + \alpha_4\right]^2}{\alpha_4 - (z_n - x_n)^2\alpha_2}, \end{cases} \tag{23}$$

where $b_1,\ b_2 \in \mathbb{R}$ such that $b_1 \neq 0$ & b_2. For a computational point of view, let us consider $b_1 = 1$ and $b_2 = \frac{1}{10}$ in the above scheme, denoted by $(OM3)$.

In the similar way, we can choose any optimal fourth-order iterative method/family of iterative methods from available literature whose first substep employs Newton's method to further obtain optimal eighth-order iterative method/family of iterative methods.

Now, we will compare them on a concrete variety of nonlinear functions with the following optimal eighth-order methods

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \left[\dfrac{f(x_n)}{f(x_n) - 2f(y_n)}\right]\dfrac{f(y_n)}{f'(x_n)}, \\ x_{n+1} = z_n + \dfrac{f(x_n)f(z_n)\left(f(x_n) + 2f(z_n)\right)\left(f\left(y_n\right) + f(z_n)\right)}{f'(x_n)f\left(y_n\right)\left(2f(x_n)f\left(y_n\right) - f(x_n)^2 + f\left(y_n\right)^2\right)}, \end{cases} \tag{24}$$

$$\begin{cases} u_n = x_n + \alpha f(x_n),\ \alpha \in \mathbb{R}, \\ y_n = x_n - \dfrac{\alpha f(x_n) f(x_n)}{f(u_n) - f(x_n)}, \\ z_n = = y_n - \dfrac{f(y_n)}{-\frac{f(u_n)(x_n-y_n)}{\alpha f(x_n)(\alpha f(x_n)+x_n-y_n)} + \frac{\alpha f(x_n)+x_n-y_n}{\alpha(x_n-y_n)} - \frac{f(y_n)(\alpha f(x_n)+2x_n-2y_n)}{(x_n-y_n)(\alpha f(x_n)+x_n-y_n)}}, \\ x_{n+1} = z_n - \dfrac{f(z_n)(u_n-x_n)(u_n-y_n)(u_n-z_n)(x_n-y_n)(x_n-z_n)(y_n-z_n)}{a_1 - a_2 f(z_n)(u_n-x_n)(u_n-y_n)(x_n-y_n)}, \end{cases} \tag{25}$$

where $a_1 = f(y_n)(u_n-x_n)(u_n-z_n)^2(x_n-z_n)^2 + f(y_n)(u_n-x_n)(u_n-z_n)^2(x_n-z_n)^2 + (y_n-z_n)^2(f(u_n)(x_n-y_n)(x_n-z_n)^2 - f(x_n)(u_n-y_n)(u_n-z_n)^2),\ a_2 = (u_n(x_n+y_n-2z_n) + x_n(y_n-2z_n) + z_n(3z_n-2y_n))$,

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \left[\dfrac{2f(x_n)-f(y_n)}{2f(x_n)-5f(y_n)}\right]\dfrac{f(y_n)}{f'(x_n)}, \\ x_{n+1} = z_n - \dfrac{f(z_n)}{2f[z_n,\ x_n]-f'(x_n)}\left[1 + \dfrac{f(z_n)}{f(y_n)} + \left(\dfrac{f(y_n)}{f'(x_n)}\right)^3 - \dfrac{2f(z_n)}{f'(x_n)} \right. \\ \qquad \left. - \dfrac{31}{4}\left(\dfrac{f(y_n)}{f(x_n)}\right)^4 - \dfrac{3}{2}\left(\dfrac{f(y_n)}{f(x_n)}\right)^3 + \left(\dfrac{f(z_n)}{f(x_n)}\right)^2 + \left(\dfrac{f(z_n)}{f(y_n)}\right)^2\right], \end{cases} \tag{26}$$

$$\begin{cases} w_n = x_n + \beta f(x_n),\ \beta \in \mathbb{R} \\ y_n = x_n - \dfrac{\beta f(x_n) f(x_n)}{f(w_n) - f(x_n)}, \\ z_n = y_n - \dfrac{f(w_n) f(y_n)(y_n - x_n)}{(f(w_n) - f(y_n))(f(y_n) - f(x_n))}, \\ x_{n+1} = z_n - \dfrac{f(w_n) f(y_n)\left(\frac{f(x_n)(z_n-x_n)}{f(z_n)-f(x_n)} - x_n + y_n\right)}{(f(w_n) - f(z_n))(f(y_n) - f(z_n))} + \dfrac{f(y_n)(z_n - y_n)}{f(z_n) - f(y_n)}, \end{cases} \tag{27}$$

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \left[\dfrac{f(x_n) - f(y_n)}{f(x_n) - 2f(y_n)}\right]\dfrac{f(x_n)}{f'(x_n)}, \\ u_n = z_n - \left(\dfrac{f(x_n) - f(y_n)}{f(x_n) - 2f(y_n)} + \dfrac{f(z_n)}{2(f(y_n) - 2f(z_n))}\right)^2 \dfrac{f(z_n)}{f'(x_n)}, \\ x_{n+1} = u_n - \dfrac{3(b_2+b_3) f(z_n)(u_n - z_n)}{f'(x_n)(b_1(u_n - z_n) + b_2(y_n - x_n) + b_3(z_n - x_n))} \end{cases} \tag{28}$$

and

$$\begin{cases} y_n = x_n - \dfrac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \left[\dfrac{f(y_n)}{f(x_n) - 2f\,(y_n)}\right]\dfrac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = z_n - \left[\dfrac{6f(y_n)^4\{f(x_n) + 5f(y_n)\}}{f(x_n)^5}\right]\dfrac{f(z_n)}{f'(x_n)} \\ \qquad - \dfrac{f(x_n) + 31f(z_n)}{f(x_n) + 30f(z_n)}\left[\dfrac{f[y_n,\ x_n]f(z_n)}{f[z_n,\ x_n]f[y_n,\ z_n]}\right], \end{cases} \tag{29}$$

which were proposed by Džunić and Petković [9], Khattri and Steihaug [11] (for $\alpha = 1$), Soleymani et al. in [21], Kung and Traub [15] (for $\beta = 1$), Cordero et al. in [7] (for $b_1 = 1,\ b_2 = 1,\ b_3 = 2$) and Heydari et al. [10], respectively called by $DP,\ KS,\ SM,\ KT,\ CM$, and HM.

For comparisons of our proposed methods with the other existing ones, we experimented with the functions $f_i(x),\ i = 1, .., 6$. We have taken 500 equally spaced points $\{t_i\}_{i=0}^{500}$ in the interval $[-3, 3]$ for $f_i(x),\ i = 1, .., 5$ and in $[0.1, 6.1]$ for $f_6(x)$ as initial points for the methods. Notice that $f_2(x) = 0$ contains two solutions $\xi = 1.4036021248742166432791385576 8, \xi = -1$ in $[-3,\ 3]$, and the others only one solution.

If x_0 attempts a root with tolerance $\epsilon = 10^{-5}$ in 14 iterations we have decided it converged to the root, otherwise, it diverged. We have registered the total number of iterations required to converge to a root and also collected the CPU time in seconds required to run each method on all the points using Samsung desktop computer with Intel(R) Core(TM) i5-4590 CPU. We then computed the average number of iterations required per point and the number of points requiring more than 14 iterations.

We have averaged performance results for the methods in comparison in Tables 2-4 across the 6 test functions. Based on Table 2 we find that the minimum the number of divergent points on average is achieved by OM1 (5.67 out of 500 points) followed by KS (6 points), OM3 (11.3 points), DP (37.8 points) and OM2 (67.5 points). All the others have $150 - 286.3$ number of points requiring more than 14 iterations on average. We will remove these methods from further consideration, since these methods have more than 24 percent of divergence. In terms of CPU time (see Table 3), the fastest method is DP (0.973 seconds) closely followed by OM1 (1.158 seconds), KS (1.287 seconds) and OM3 (1.3 seconds). The slowest is OM2 (2.231 seconds), which will be removed from further discussion. Recall that although SM is the fastest of all the methods considered, it is no longer being considered now since it is one of the methods having more than 24 percent of initial points diverged. Consulting the average number of iterations per point on average (see Table 4), we find that OM1 is best (2.49) followed by KS (2.58) and OM3 (2.63). The worst is DP (3.30).

In view of our analysis of the results in Tables 2-4 given above, the best method overall is OM1.

Conclusions

In this paper we proposed a new optimal eighth-order family of methods based on rational function in a general way. Some of our methods have been compared to several existing methods of the same order. OM1, one of our methods, is found to be the best method based on 3 quantitative criteria (Divergence percent, CPU time, Average number of iterations per point), confirming that the proposed methods are highly efficient as compared to the existing methods.

Table 1: Convergence behavior of methods $OM1,\ OM2$ and $OM3$ on each test function

Cases	$f(x)$	n	x_n	$\lvert f(x_n)\rvert$	$\lvert x_{n+1}-x_n\rvert$	$\lvert\frac{x_{n+1}-x_n}{(x_n-x_{n-1})^8}\rvert$	η	ρ
$OM1$	f_1	0	0.5	1.0	$5.0(-1)$			
		1	0.00306695875782981	$3.1(-3)$	$3.1(-3)$	$8.247549737(-1)$	$1.980000000(+2)$	15.8377
		2	$1.48036410450262(-18)$	$5.1(-18)$	$5.1(-18)$	$1.891058911(+2)$		8.00794
		3	$4.56681645644905(-141)$	$4.6(-141)$	$4.6(-141)$	$1.980000000(+2)$		8.00000
$OM1$	f_2	0	1.5	2.0	$9.6(-2)$			
		1	1.40360330825001	$1.9(-5)$	$1.2(-6)$	$1.587178031(+2)$	$4.605587105(+2)$	8.45540
		2	1.40360212487422	$2.8(-44)$	$1.8(-45)$	$4.605524658(+2)$		8.00000
		3	1.40360212487422	$7.0(-355)$	$4.5(-356)$	$4.605587105(+2)$		8.00000
$OM2$	f_3	0	0.6	$3.8(-1)$	$1.1(-1)$			
		1	0.714806004989988	$3.4(-7)$	$9.3(-8)$	3.069175663	1.085366264	7.51976
		2	0.714805912362778	$2.2(-56)$	$5.9(-57)$	1.085365407		8.00000
		3	0.714805912362778	$5.7(-450)$	$1.6(-450)$	1.085366264		8.00000
$OM2$	f_4	0	2.4	$2.2(-1)$	$2.7(-1)$			
		1	2.13226772533188	$4.8(-11)$	$5.9(-11)$	$2.234686093(-6)$	$5.519129858(-6)$	8.68610
		2	2.13226772527289	$6.6(-88)$	$8.1(-88)$	$5.519129857(-6)$		8.00000
		3	2.13226772527289	$8.4(-703)$	$1.1(-702)$	$5.519129858(-6)$		8.00000
$OM3$	f_5	0	1.5	$2.9(-1)$	$2.5(-1)$			
		1	1.74613952980597	$7.0(-10)$	$6.0(-10)$	$4.468629204(-5)$	$1.786446252(-4)$	8.98850
		2	1.74613953040801	$3.6(-78)$	$3.1(-78)$	$1.786446246(-4)$		8.00000
		3	1.74613953040801	$1.7(-624)$	$1.5(-624)$	$1.786446252(-4)$		8.00000
$OM3$	f_6	0	0.5	$6.9(-1)$	$5.0(-1)$			
		1	0.999983241870036	$1.7(-5)$	$1.7(-5)$	$4.291231744(-3)$	$8.979552469(-4)$	5.74343
		2	1.00000000000000	$5.6(-42)$	$5.6(-42)$	$8.979882433(-4)$		8.00000
		3	1.00000000000000	$8.5(-334)$	$8.5(-334)$	$8.979552469(-4)$		8.00000

(It is straightforward to say that our methods not only converging very fast to the desired zero but also have the smaller asymptotic error constant which confirm the theoretical results.)

Table 2: Number of points requiring more than 14 iterations for each test function (1–6) to corresponding method and divergence percentage

Methods	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$f_6(x)$	average	Divergence Percentage
OM1	1	7	1	19	6	0	5.67	1.13%
OM2	17	26	12	27	21	302	67.5	13.5%
OM3	2	17	1	12	36	0	11.3	2.26%
DP	7	32	2	179	6	1	37.8	7.56%
KS	19	9	2	0	6	0	6	1.2%
SM	83	136	37	180	118	346	150	30%
KT	77	500	500	0	24	500	266.8	53.4%
CM	123	500	332	190	500	73	286.3	57.3%
HM	72	137	33	134	38	330	124	24.8%

Table 3: CPU time (in seconds) required for each test function (1–6) to corresponding method

Methods	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$f_6(x)$	average
OM1	2.153	0.578	0.593	1.077	1.280	1.264	1.158
OM2	4.228	1.388	1.139	1.856	1.950	2.824	2.231
OM3	2.215	0.686	0.671	1.108	1.607	1.513	1.3
DP	1.825	0.390	0.437	0.702	1.155	1.326	0.973
KS	2.356	0.656	0.687	1.014	0.827	2.184	1.287
SM	1.732	0.374	0.515	0.749	0.936	0.764	0.845
KT	4.337	2.215	3.183	3.447	2.683	3.479	3.224
CM	1.981	0.437	1.217	1.779	1.248	2.714	1.563
HM	1.810	0.406	0.499	0.827	1.092	0.670	0.884

Table 4: Average number of iterations per point for each test function (1–6) to corresponding method

Methods	$f_1(x)$	$f_2(x)$	$f_3(x)$	$f_4(x)$	$f_5(x)$	$f_6(x)$	average
OM1	2.43	3.10	2.45	2.72	2.51	1.71	2.49
OM2	4.92	6.34	4.90	4.47	3.90	9.88	5.74
OM3	2.42	3.42	2.44	2.48	3.16	1.88	2.63
DP	2.61	3.73	2.48	6.23	2.55	2.18	3.30
KS	3.04	3.97	2.74	1.64	1.66	2.42	2.58
SM	4.25	5.75	3.22	6.30	4.87	10.23	5.77
KT	6.42	14	14	6.37	5.88	14	10.11
CM	5.52	14	13.41	9.19	14	5.44	10.26
HM	4.01	5.82	3.24	5.35	3.02	9.80	5.21

References

[1] Behl, R., Kanwar, V., Sharma, K.K. (2013) Optimal equi-scaled families of Jarratt's method. *Int. J. Comput. Math.* **90**, 408–422.

[2] Behl, R., Cordero, A., Motsa, S.S., Torregrosa, J.R. (2015) Construction of fourth-order optimal families of iterative methods and their dynamics. *Appl. Math. Comput.* **271**, 89–101.

[3] Behl, R., Motsa, S.S. (2015) Geometric construction of eighth-order optimal families of ostrowski's method. *T. Sci. W. J.* **2015**, article ID 614612, 11 pages.

[4] Bi, W., Ren, H., Wu, Q. (2009) Three-step iterative methods with eighth-order convergence for solving nonlinear equations. *J. Comput. Appl. Math.* **255**, 105–112.

[5] Chun, C. (2007) Some variants of King's fourth-order family of methods for nonlinear equations. *Appl. Math. Comput.* **190**, 57–62.

[6] Chun, C. (2007) A family of composite fourth-order iterative methods for solving nonlinear equations. *Appl. Math. Comput.* **187**, 951–956.

[7] Cordero, A., Torregrosa, J.R., Vassileva, M.P. (2011) Three-step iterative methods with optimal eighth-order convergence. *J. Comput. Appl. Math.* **235**, 3189–3194.

[8] Cordero, A., Hueso, J.L., Martínez, E., Torregrosa, J.R. (2010) New modifications of Potra-Pták's method with optimal fourth and eighth order of convergence. *J. Comput. Appl. Math.* **234**, 2969–

2976.

[9] Dzŭnić J., Petković, M. (2012) A family of three-point methods of Ostrowski's type for solving nonlinear equations. *J. Appl. Math.* **2012**, doi : 10.1155/2012/425867.

[10] Heydari, M., Hosseini, S.M., Loghmani, G.B. (2011) On two new families of iterative methods for solving nonlinear equations with optimal order. *Appl. Anal. Disc. Math.* **5**, 93–109.

[11] Khattri, S.K., Steihaug, T. (2014) Algorithm for forming derivative-free optimal methods. *Numer. Algor.* **65**, 809–824.

[12] Khattri, S.K., Noor, M.A., Al-Saidc, E. (2011) Unifying fourth-order family of iterative methods. *Appl. Math. Lett.* **24**, 1295–1300.

[13] King, R.F. (1973) A family of fourth order methods for nonlinear equations. *SIAM J. Numer. Anal.* **10**, 876–879.

[14] Kou, J., Li, Y., Wang, X.(2017) A composite fourth-order iterative method for solving non-linear equations. *Appl. Math. Comput.* **184**, 471–475.

[15] Kung, H.T., Traub, J.F. (1974) Optimal order of one-point and multi-point iteration. *J. ACM* **21**, 643–651.

[16] Liu, L., Wang, X. (2010) Eighth-order methods with high efficiency index for solving nonlinear equations. *J. Comput. Appl. Math.* **215**, 3449–3454.

[17] Maheshwari, A.K. (2009) A fourth order iterative method for solving nonlinear equations. *Appl. Math. Comput.* **211**, 383–391.

[18] Petković, M.S., Neta, B., Petković, L.D., Džunić, J. (2012) Multipoint methods for solving nonlinear equations. *Academic Press*.

[19] Sharma, J.R., Guha, R.K., Gupta, P. (2013) Improved King's methods with optimal order of convergence based on rational approximations. *Appl. Math. Lett.* **26**, 473–480.

[20] Sharma, J.R., Arora, H. (2014) An efficient family of weighted-Newton methods with optimal eighth order convergence. *Appl. Math. Lett.* **29**, 1–6.

[21] Soleymani, F., Vanani, S.K., Khan, M., Sharifi, M. (2012) Some modifications of King's family with optimal eighth-order of convergence. *Math. Comput. Model.* **55**, 1373–1380.

[22] Soleymani, F., Sharma, R., Li, X., Tohidi, E. (2012) An optimized derivative-free form of the PotraPták method. *Math. Comput. Model.* **56**, 97–104.

[23] Thukral, R. (2010) A new eighth-order iterative method for solving nonlinear equations. *Appl. Math. Comput.* **217**, 222–229.

[24] Thukral, R., Petkovíc, M.S. (2010) A family of three point methods of optimal order for solving nonlinear equations. *J. Comput. Appl. Math.* **233**, 2278–2284.

[25] Traub, J.F. (1964) Iterative methods for the solution of equations. *Prentice-Hall, Englewood Cliffs*.

[26] Jarratt, P., Nudds, D. (1965) The use of rational functions in the iterative solution of equations on a digital computer. *The Comput. J.* **8(1)**, 62–65.

Effect of slit inclusions in drag reduction of Flow over square cylinders for low Reynolds number in the laminar regime.

***Rohit Bhattacharya[1], † Fausto Moreira-Izurieta[2],**

[1]School of Aerospace, Mechanical and Mechatronics Engineering. The University of Sydney, Australia.
[2] School of Aerospace, Mechanical and Mechatronics Engineering. The University of Sydney, Australia.

*Presenting author: rbha4656@uni.sydney.edu.au
†Corresponding author: andrew.mi8@gmail.com

Abstract

The flow over bluff bodies is separated compared to the flow over streamlined bodies. The investigation of the fluid flow over a circular cylinder with a streamwise slit has shown a reduction in the drag coefficient in the past for a very low Reynolds number. This work helps in understanding the fluid flow over bluff bodies in the laminar regime. An increase in the slit ratio is inversely proportional to the reduction in the drag coefficient resulting in a narrower wake which is a phenomenon seen in the turbulent regime, hence reducing the drag coefficient.

In this work two different approaches are used to simulate fluid flow over 2D cylinder of a square cross section and a comparison between the finite volume method and the Lattice Boltzmann Method is made. The width of the slits progressively increase from 10% all the way to 40% of the diameter of the cylinder. Reduction in the drag coefficients will be show for different values of Reynolds numbers as the width of the slit increases. The effect of slit inclusions on flow over cylinder for different values of Reynolds number is studied in further detail and discussed in this paper.

Keywords: CFD, Lattice-Boltzmann, Drag reduction, Flow over slit cylinder.

Introduction

Computational Fluid Dynamics (CFD) solves the Navier-strokes equations numerically for fluid flows using computers .The Lattice Boltzmann Method (LBM) is a promising method in simulation of flow in complex geometries [1].

The computational investigation of the fluid flow over a cylinder with a streamwise slit has been done before for cylinders of circular cross section where a reduction in drag coefficients have been seen for low Reynolds number (~10) [2]. This work shows a reduction of the drag coefficient by 7% and this was demonstrated for a slit/diameter ratio of 0.2 approximately. This work aims to highlight the effect of slit size on the flow over cylinders in laminar regime (Re=10). A comparative study between the finite volume method and the Lattice Boltzmann method will show flow patterns over the cylinders which will be used to measure the efficiency and accuracy of the two difference approaches.

CFD Simulation using Finite volume Analysis on ANSYS Fluent

The laminar model on ANSYS Fluent is used to carry out the 2D Simulations. A moving velocity to the wall is assigned which has the same speed of the fluid at the inlet of the computational fluid domain. Air is used as the fluid for the analysis that flows at a Reynolds number of 10. The following equation is used to calculate the Reynolds number.

$$Re = \frac{\rho V D}{\mu} \quad (1)$$

Here is the ρ density of the fluid which is 1.225Kg/m^3,V is the velocity of the fluid which is 0.1477 m/s and μ is the viscosity of the fluid which is 0.0181 (mPa.s).

The drag force over the cylinder is obtained using ANSYS and this is used to calculate the drag coefficient which is given by the following equation.

$$C_d = \frac{F_x}{0.5 \rho A V^2} \quad (2)$$

Here C_d is the drag coefficient, F_x is the drag force, ρ is the density of the fluid, A is the projected area and V is the velocity of the fluid.

Lattice Boltzmann Method (LBM)

The Lattice Boltzmann Method algorithm was built on Python .The Velocity of the flowing fluid is discretised first which is obtained from the discretized Boltzmann transport equation [2].

$$f_\alpha(\boldsymbol{r} + \boldsymbol{v}_\alpha \Delta t, t + \Delta t) = f_\alpha(\boldsymbol{r}, t) + C(f_\alpha, f_\alpha) \quad (3)$$

Where f is the distribution function for velocity, subscript α stands for the direction, $\boldsymbol{r}$ is the position vector for each lattice node, t is time and c is the lattice speed of sound. The D2Q9 model is a two dimensional lattice model with 9 discrete nodes assigned with velocity vectors, one at the centre and eight others surrounding it. The most critical point here is to specify the number of cells on the x and y axis as this defines the computational fluid domain to carry out simulations using the Boltzmann equations.

The following is the formula for the distribution function

$$f^{eq} = w_\alpha \left(1 + \frac{\boldsymbol{v}_\alpha \cdot \boldsymbol{u}}{c_s^2} + \frac{(\boldsymbol{v}_\alpha \cdot \boldsymbol{u})^2}{2c_s^4} - \frac{|\boldsymbol{u}|^2}{2c_s^2}\right) \quad (4)$$

Here w_α is the weight function in the α direction.

LBM Algorithm

The macroscopic variable is defined based on the distribution function. The information of the molecular number density (n) can be found using equation 5 , and momentum density ($n\boldsymbol{u}$) can be found using equation 6.

$$n = \sum_\alpha f_\alpha \quad (5)$$

$$\boldsymbol{nu} = \sum_{\alpha} \boldsymbol{f_\alpha v_\alpha} \tag{6}$$

The pressure distribution is given by the following equation.

$$\boldsymbol{p = nc_s^2} \tag{7}$$

Here c_s is the speed of sound as lattice constant. Using the complementary ideal gas equations ($pV = NRT, p = nRT$), It can be shown that$RT = c_s^2$. The values of w (weight factor) and c_s depend on the specific choice of the lattice velocity model. Table 1 summarises values of w and c_s for the chosen lattice model D2Q9 [4].

Table 1: Weight factors used for the lattice nodes

Model	v_α	w_α	c_s^2
D2Q9	(0,0)	16/36	
	(±1,0) ,(0,±1)	4/36	1/3
	(±1,±1)	1/36	

The computational fluid domain is a 2D channel of a 12D × 7D Cross section where the square cylinder has a side of D (=1μm). The square cylinder is located in the centre of the computational fluid domain for carrying out the iterations. Fig.1 shows the schematic of the computational fluid domain. In previous studies where computational analysis was performed using MATLAB where the simulation time was 21 seconds. The same simulation was carried out using Python in 15 seconds which also shows improved efficiency and accuracy.

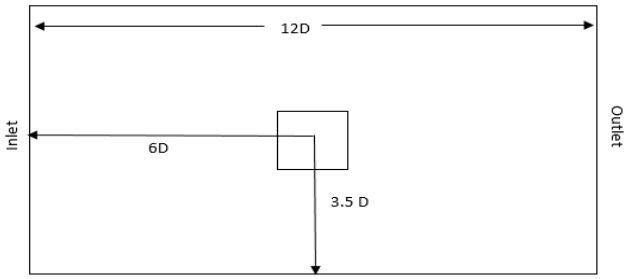


Figure 1: The schematic of the Computational fluid Domain

Table 2: Validation of numerical results obtained using CFD and LBM, against a variety of prior works in the literature at constant *Re*=10.

Literature C_d	C_d obtained by CFD	C_d obtained by LBM
7.29 (Durao et al) [5]	8.97	8.48
7.53 (Sohankar et al)[6]	8.97	8.48

The reported data indicate that the values of the drag coefficients obtained from LBM and CFD simulation are very close to prior works results. The drag coefficient results are especially close to that obtained by Durao [5] under Poiseuille flow conditions, suggesting the wall effects are not completely avoided in our work. Upon acceptable validation of our CFD model and LBM, further investigations are conducted for the slit cylinder. The boundary conditions remain the same, only the geometry in the computational fluid domain would change as a slit would be considered inside the cylinder.

Investigation of the effect of the slit

A slit is a narrow cut or opening in a bluff body. This section highlights the effect of the size of slits in the reduction of drag coefficient in bluff bodies. Cases where the slit to width (side of the square) ratio progressively increases by 10% up to 40% are considered and shown in Table 3.

Table 3: Drag coefficient and % reduction in comparison to slit free cylinder, calculated from CFD and LBM for various slit ratio and compared.

D↘ S/D 	CFD (ANSYS)		LBM (Python)		%Difference CFD vs. LBM
	C_d	%Reduction	C_d	%Reduction	
0	8.97	-	8.48	-	5.77
0.1	8.21	8.47	8.22	3.06	0.12
0.2	8.02	10.59	8.05	5.07	0.37
0.3	7.94	11.48	7.87	7.19	0.89
0.4	7.87	12.26	7.78	8.25	1.16

All the values obtained from ANSYS Fluent and LBM have been compared in terms of the percentage difference with respect to a slit-free cylinder. It can be seen from the results in the last column that the calculated drag coefficient from the two methods agree well and in most cases the difference is less than 2%.

Based on the results obtained we can conclude that both models are reliable and produce similar results for the flow field and drag coefficient. However there are some advantages in terms of algorithm simplicity and computational cost. The lattice Boltzmann method has a few advantages over the CFD method using ANSYS Fluent as outlined below

- The algorithm can be easily implemented on Python.
- Due to the regular lattice structure and because of the limited dynamic interaction that requires only one contact between each lattice node and its nearest neighbours for each iteration step.[7]
- Discretization of the macroscopic continuum equations is not needed. Hence, the LBM does not consider explicitly the distribution of pressure on interfaces of refined grids since the implicitly is already included in the computational scheme.[2]
- The computation time using LBM is considerably lower.

Conclusions

Numerical Investigation of fluid flow around cylinders of square cross sections were carried out with and without slits of varying sizes. The CFD method using ANSYS Fluent and the LBM Script using Python were used to simulate fluid flow and to calculate the drag coefficient. The numerical results agreed reasonably with available experiments at Reynolds number of 10 for a 2D cylinder. The incorporation of slits on such cylinders showed a considerable reduction on 12.3 % in drag coefficients.

References

[1] Liaw, K. (2005). *Simulation of flow around bluff bodies and bridge deck sections using CFD* (Doctoral dissertation, University of Nottingham).

[2] Bhattacharya, R., Moshfegh, A., & Jabbarzadeh, A. (2016). Effect of Slit Inclusions in Drag Reduction of Flow over Cylinders. In Applied Mechanics and Materials (Vol. 846, pp. 18-22). Trans Tech Publications.

[3] Dixon, A. G., Nijemeisland, M., & Stitt, E. H. (2006). Packed tubular reactor modeling and catalyst design using computational fluid dynamics. *Advances in Chemical Engineering*, *31*, 307-389.

[4] Zhang, H. (2008). Lattice Boltzmann method for solving the bioheat equation. Physics in medicine and biology, 53(3), N15.

[5] Durao, D. F. G., Heitor, M. V., & Pereira, J. C. F. (1988). Measurements of turbulent and periodic flows around a square cross-section cylinder. Experiments in Fluids, 6(5), 298-304.

[6] Sohankar, A., Norberg, C., & Davidson, L. (1998). Low-Reynolds-number flow around a square cylinder at incidence: study of blockage, onset of vortex shedding and outlet boundary condition. International journal for numerical methods in fluids, 26(1), 39-56.

[7] Abhijeet, T (2011). Introduction to the Lattice Boltzmann Method. 10th Indo-German Winter Academy 2011, IIT Kharagpur, India.

Failure analysis of laminated tubes under tension-torsion biaxial loading

†Jingmeng Weng [1], *Weidong Wen [1], Haitao Cui [1], Ying Xu [1], Yaoxia Huo [1]

[1] Jiangsu Province Key Laboratory of Aerospace Power System, Nanjing 210016, P.R. China
State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing 210016, P.R. China
College of Energy & Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China

*Presenting author: wjm19890606@126.com
†Corresponding author: jm_weng@nuaa.edu.cn

Abstract

An experimental and numerical study of laminated tubes with [45/70/-45/0/2]s under different combinations of tension-torsion biaxial loading is presented. The effect of biaxiality ratio on biaxial strength is discussed. Moreover, a progressive damage model to simulate the failure behaviour of laminated tubes under different combinations of tension-torsion biaxial loading is presented. The main advantage of the model is that it can simulate the failure behavior of laminated tubes under combined tension-torsion biaxial loading by using the experimental results of unidirectional flat specimens. The maximum error between predicted strength and experimental results is within 9%. The experimental and numerical results show that the axial load carrying capacity of tubular specimen decreases rapidly as biaxiality ratio increases.

Keywords: Laminated tubular specimen, Biaxial loading, Biaxial strength prediction model

Introduction

Due to their superior strength-to-weight and modulus-to-weight ratios, fiber reinforced composite materials are widely used in military and commercial applications, such as airplanes and motor vehicles. The majority of these structural components in service are subjected to biaxial or triaxial state of stress.

If the goal of biaxial testing is to generate a failure envelope in σ_1-σ_2 stress space, cruciform specimen is the most appropriate choice. Under biaxial loading, the distribution and repartition of stress is not constant over the cruciform specimen. Therefore, strain field monitoring techniques are required, such as strain gages or strain rosettes [1], high-speed stereo digital image correlation [2], infrared thermography[3], air-coupled guided waves [3], and digital image correlation [4, 5], etc. Test monitoring and numerical simulation [1-11] have demonstrated that it's almost impossible to eliminate the stress concentration in the milled zone and the outer fillet corner between two perpendicular arms.

Thin tubular specimens avoid problems associated with stress concentrations and free edges effects that are encountered with cruciform specimens, and a wide range of biaxial and triaxial stress space can be applied by subjecting tubular specimens to different combinations of internal/external pressure, torsion and axial load. It has been widely employed in investigations to study the failure behavior of tubular specimens under biaxial and triaxial loading [12].

Due to their microscopic heterogeneity, tubular specimens can fail in a variety of ways according to the structure of tubular specimen and the loading condition. The static and fatigue failure mechanisms of tubular specimens under multiaxial loading are almost researched experimentally through filament wound tubes and plain woven fabric tubes.

Kaddour A S [13] and Qi [14] studied the failure behavior of $\pm\theta$ filament wound tubes under various combinations of biaxial loads, matrix cracking was taken as the failure criterion. Fujii et al. [15-18] studied the static and fatigue failure behavior of plain woven fabric tubes under different combinations of tension-torsion loading, the failure mechanisms include delamination, matrix cracking and fiber breakage.

The aim of the present work is to examine the failure forms and failure strength of laminated tubes under different combinations of tension-torsion biaxial loading. Firstly, material properties of T700/epoxy are tested. Secondly, laminated tubes with $[45/70/-45/0_2]_s$ are tested under different combinations of tension-torsion biaxial loading. Thirdly, a progressive damage model is established to simulate the damage from initiation and propagation to the final catastrophic failure of tubular specimens, modified Hashin criteria are used to predict the strength of tubular specimens under different combinations of tension-torsion biaxial loading. Finally, some conclusions are drawn from the experimental and numerical studies.

Experimental investigation

Unidirectional (UD) T700/epoxy prepreg tape was used to manufacture all the unidirectional flat specimens and laminated tubular specimens.

Unidirectional flat specimens

According to ASTM D 3039-07 and ASTM D 3410-03, unidirectional flat specimens with off axis angles equal to 0°, 45° and 90° were prepared and tested on MTS 809 testing system. With the aid of strain rosettes, elasticity modulus and Poisson's ratio can be tested. Five specimens were tested for each material property. The results of elasticity modulus, Poisson's ratio and strength in each material principal direction are listed in Table 1.

Table 1. Material properties of T700/epoxy

	E_1	E_2	G_{12}	X_T	X_C	Y_T	Y_C	S_{12}	ν_{21}
	GPa	GPa	GPa	MPa	MPa	MPa	MPa	MPa	–
Average	106.5	6.77	3.23	1388.5	378.41	31.58	82.72	97.57	0.3476
Std. dev. (%)	5.26	4.06	5.46	5.09	5.05	2.24	9.78	7.19	9.33

Laminated tubular specimens

Some researchers [19, 20] tested laminated tubular specimens with different lay-ups under combined tension-torsion biaxial loading, such as $[90]_n$, $[0_F/90_{U,3}]$, $[0_F/90_{U,3}/0_F]$ and $[0/45/90/-45]_s$. For these lay-ups, there is a seam in the circumferential direction of all 90° plies in tubular specimen. The existence of the seam will lead to stress discontinuity. Theoretically and practically, 90° ply may never exist in laminated tubular specimens if no seam is required. Therefore, the stacking sequence of laminated tubular specimens is set as $[45/70/-45/0_2]_s$ in this paper. The geometry of laminated tubular specimen is shown in Fig. 1.

All tubular specimens were tested on MTS 809 testing system for different combinations of tension-torsion biaxial loading. Four specimens were tested for each test condition.

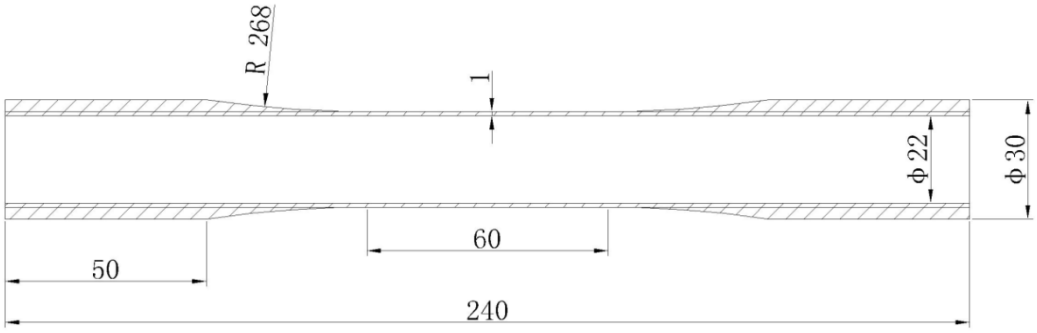


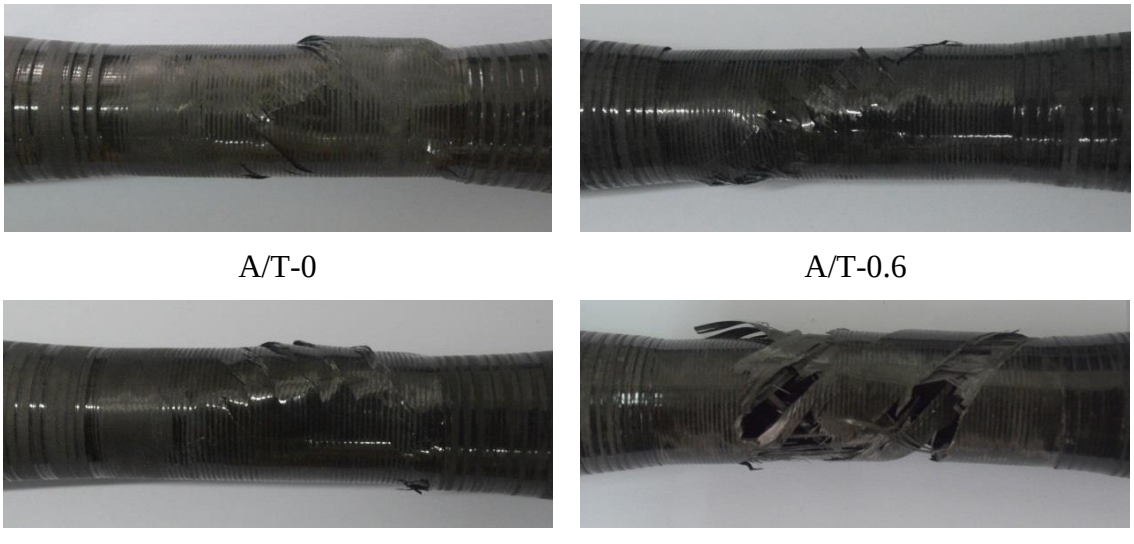
Figure 1. Geometry of tubular specimen (units in millimeters)

Experimental results

The final failure modes of tubular specimens under different combinations of tension-torsion biaxial loading are presented in Fig. 2. In appearance, the main fracture of every specimen lies in the middle of the gage section.

For every loading condition, matrix cracking is parallel to fibers, the fracture path of fiber is perpendicular to the direction of fiber in each layer, and delamination occurs around fiber breakage and matrix cracking. Even though all these kinds of failure modes occur in each biaxial loading condition, the damage degree of laminated tubes under different biaxiality ratio is different from each other, and the damage degree of laminated tubes becomes more and more serious along with the biaxiality ratio increases.

The biaxial strength of laminated tubes under different combinations of tension-torsion loading are listed in Table 2. It can be seen that as the biaxiality ratio decreases, the tensile strength decreases rapidly, while the torsional strength increases slowly. For instance, with A/T-2.2 as a standard, the tensile strength decreases by 64.46%, while the torsional strength only increases by 32.34% when biaxiality ratio equals to 0.6.



A/T-0 A/T-0.6

A/T-1.1 A/T-2.2

a. A/T is the biaxiality ratio (divide tension stress by torsion stress).

Figure 2. Failed tubular specimens under different combinations of tension-torsion loading

Table 2. Biaxial strength of laminated tubes under different different combinations of tension-torsion loading

Direction		A/T-2.2	A/T-1.1	A/T-0.6	A/T-0
Tensile strength	Experimental result (MPa)	332.48	190.79	118.18	-
	Std. dev. (%)	13.82	6.86	7.97	-
Torsional strength	Experimental result (MPa)	152.19	172.24	201.41	208.26
	Std. dev. (%)	15.39	9.27	5.12	7.76

Biaxial strength prediction model

Progressive damage model has been successfully utilized to study the failure behavior of composite materials under uniaxial static loading and uniaxial fatigue loading [21]. A typical progressive damage model comprises three major components: stress analysis, failure analysis and material property degradation rules. Progressive damage model can simulate the damage from initiation and propagation to the final catastrophic failure in detail. In this paper, the progressive damage model is extended for the failure analysis of laminated tubular specimens which is subjected to tension-torsion biaxial loading, and the flow chart is plotted in Fig. 3.

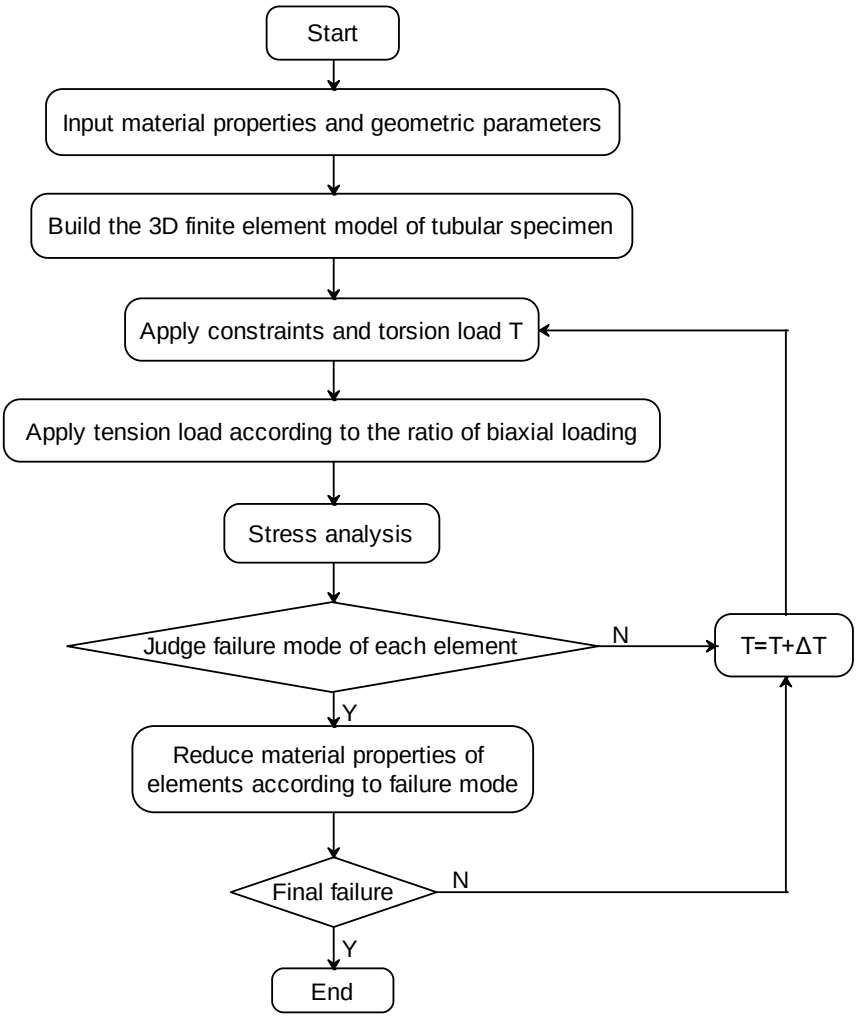


Figure 3. Flow chart of biaxial strength prediction model

Stress analysis

The first component of progressive damage model, the stress analysis, is explained in this section. For composite laminates or laminated tubular specimens without hole or any other cutouts, the stress field in the gauge section is homogeneous. However, the existence of end-tab in finite element model will cause significant stress concentrations near the end-tab. [22] According to failure criteria, these elements near the end-tab will be the first one to fail. This phenomenon is not consistent with experiment.

Based on above analysis, the tube was divided into three sections during the modeling process. As shown in Fig. 4, the middle one corresponds to the gauge section, and the other two sections correspond to the end-tab section. In the finite element model, all nodes on one end of the tube were all fixed, and all nodes on the other end of the tube were coupled with the node which lies on the axis of the tube. All tension and torsion loads were applied on this node which lies on the axis of the tube. During the progressive damage analysis, failure analysis and material property degradation were only applied on elements which belong to the gauge section. In this way, even though there are significant stress concentrations near the constraint region and the load region, the initial and final failure is caused by the homogeneous stress field in the gauge section.

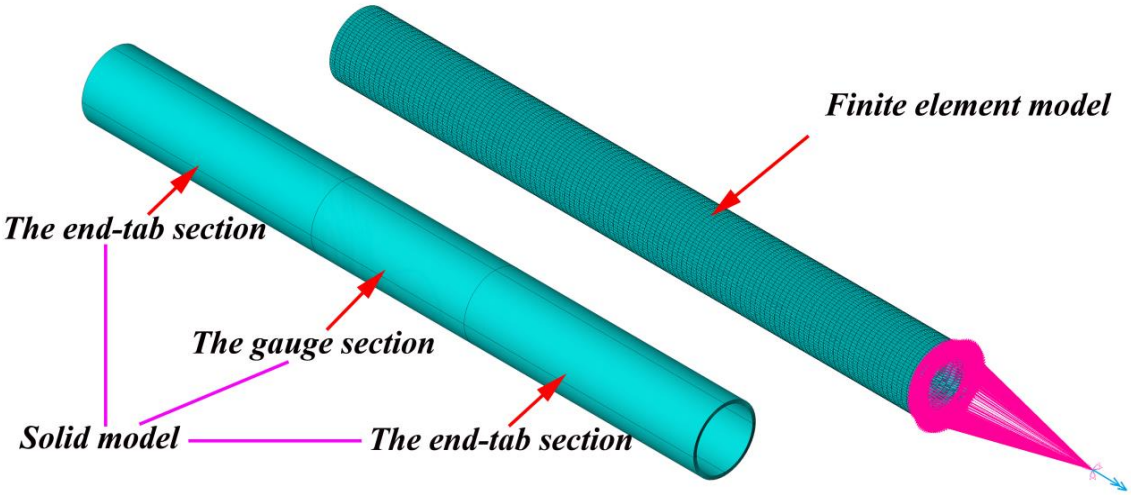


Figure 4. Solid model and finite element model of tubular specimen

Failure criteria

Due to their microscopic heterogeneity, composite materials present different failure modes under multiaxial state of stress. Fiber tensile/compressive failure, matrix tensile/compressive failure, tensile/compressive delamination failure and fiber-matrix shear failure are considered. Hashin [23] proposed a set of famous two-dimensional failure criteria for predicting the failure of composite materials. These criteria have been extensively applied in the progressive damage models. Modified three-dimensional Hashin failure criteria are used to predict the strength of tubular specimens under different combinations of tension-torsion biaxial loading. The specific expressions are given as follows:

(1) Fiber tensile failure ($\sigma_1 > 0$)

$$\left(\frac{\sigma_1}{X_T}\right)^2 + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 - r^2 \geq 0 \tag{1}$$

(2) Fiber compressive failure ($\sigma_1 < 0$)

$$\left(\frac{\sigma_1}{X_C}\right)^2 - r^2 \geq 0 \tag{2}$$

(3) Matrix tensile failure ($\sigma_2 > 0$)

$$\left(\frac{\sigma_2}{Y_T}\right)^2 + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \left(\frac{\sigma_{23}}{S_{23}}\right)^2 - r^2 \geq 0 \tag{3}$$

(4) Matrix compressive failure ($\sigma_2 < 0$)

$$\left(\frac{\sigma_2}{Y_C}\right)^2 + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 + \left(\frac{\sigma_{23}}{S_{23}}\right)^2 - r^2 \geq 0 \tag{4}$$

(5) Tensile delaminaion failure ($\sigma_3 > 0$)

$$\left(\frac{\sigma_3}{Z_T}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 + \left(\frac{\sigma_{23}}{S_{23}}\right)^2 - r^2 \geq 0 \tag{5}$$

(6) Compressive delamination failure ($\sigma_3 < 0$)

$$\left(\frac{\sigma_3}{Z_C}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 + \left(\frac{\sigma_{23}}{S_{23}}\right)^2 - r^2 \geq 0 \tag{6}$$

(7) Fiber-matrix shear failure ($\sigma_1 < 0$)

$$\left(\frac{\sigma_1}{X_C}\right)^2 + \left(\frac{\sigma_{13}}{S_{13}}\right)^2 + \left(\frac{\sigma_{12}}{S_{12}}\right)^2 - r^2 \geq 0 \tag{7}$$

Where σ_i (i=1,2,3) are the normal stress components in each material principal direction, σ_{ij} (i, j=1,2,3) are the shear stress components, X_T, X_C, Y_T, Y_C, Z_T and Z_C represent tensile and compressive strength in longitudinal, transverse and normal direction, G_{ij} (i, j=1,2,3) and S_{ij} (i, j =1,2,3) represent the initial shear modulus and shear strength in ij plane, r is damage threshold.

Material property degradation rules

As failure occurs, material properties of failed elements are degraded. Some of the failure modes are catastrophic and some of them are not. A complete set of sudden material property degradation rules for all failure modes are given in Table 3.

Table 3. Material property degradation rules

Modes of failue	E_1	E_2	E_2	G_{12}	G_{13}	G_{23}	v_{12}	v_{13}	v_{23}
Fiber tensile failure	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Fiber compressive failure	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
Matrix tensile failure	-	0.2	-	0.2	-	0.2	0.2	-	0.2
Matrix compressive failure	-	0.4	-	0.4	-	0.4	0.4	-	0.4
Tensile delaminaion failure	-	-	0	-	0	0	-	0	0
Compressive delamination failure	-	-	0	-	0	0	-	0	0
Fiber-matrix shear failure	-	-	-	0	0	-	0	-	-

Results and discussion

A summary of the biaxial strength of tubular specimens from experiments and the progressive damage model is presented in Fig. 5 and Table 4. The maximum error between predicted strength and experiments is within 9%.

Under tension-torsion biaxial loading, there will be shear stress in 0 °plies, and the shear load carrying capacity of fiber is less than the axial load carrying capacity. Therefore, axial load carrying capacity of tubular specimen decreases rapidly as biaxiality ratio increases.

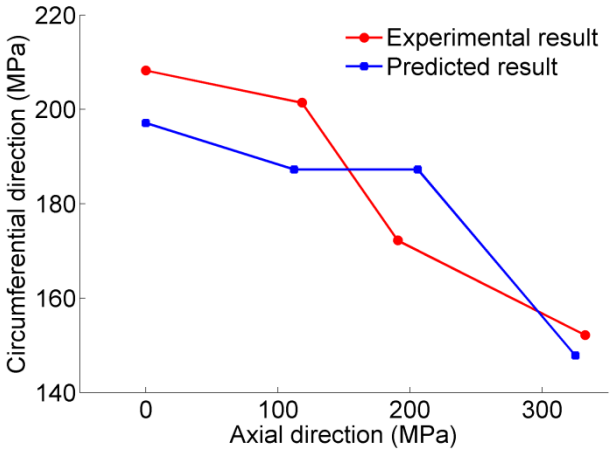


Figure 5. Biaxial strength of tubular specimens under tension-torsion loading

Table 4. Comparison of biaxial strength from the experiments and the model

Direction		A/T-2.2	A/T-1.1	A/T-0.6	A/T-0
Tensile strength	Experimental result (MPa)	332.48	190.79	118.18	-
	Predicted result (MPa)	325.26	206.00	112.36	-
	Error (%)	-2.17	7.97	-4.92	-
Torsional strength	Experimental result (MPa)	152.19	172.24	201.41	208.26
	Predicted result (MPa)	147.85	187.27	187.27	197.13
	Error (%)	-2.87	8.73	-8.51	-5.34

Conclusions

The present paper studied the biaxial strength of laminated tubes under tension-torsion biaxial loading experimentally and numerically. Firstly, the biaxial strength of laminated tubes with [45 °/70 °/-45 °/0 °2]s are tested under four different biaxial loading ratios. Secondly, a progressive damage model to simulate the failure behavior of laminated tubular specimens under different combinations of tension-torsion loading is proposed. The main advantage of the model is that it can simulate the failure behavior of laminated tubes under combined tension-torsion biaxial loading by using the experimental results of unidirectional flat specimens. The simulated tension/torsion strength show good agreement with experimental results. The experimental and numerical results show that the axial load carrying capacity decreases rapidly as the biaxiality ratio decreases. It should be noted that the progressive damage model proposed in this paper is a deterministic model, further research is needed to couple the defects for a more realistic simulation.

References

[1] Sun X S, Haris A, Tan V B C, et al., A multi-axial fatigue model for fiber-reinforced composite laminates based on Puck's criterion, *Journal of Composite Materials*, **46**, 2012, pp. 449-469.

[2] Busca D, Fazzini M, Lorrain B, et al., High-Speed Stereo Digital Image Correlation: Application to Biaxial Fatigue, *Strain*, **50**, 2014, pp. 417-427.

[3] Rheinfurth M, Schmidt F, Döring D, et al., Air-coupled guided waves combined with thermography for monitoring fatigue in biaxially loaded composite tubes, *Composites Science and Technology*, **71**, 2011, pp. 600-608.

[4] Smits A, Van Hemelrijck D, Philippidis T P, et al., Design of a cruciform specimen for biaxial testing of fibre reinforced composite laminates, *Composites Science and Technology*, **66**, 2006, pp. 964-975.

[5] Lecompte D, Smits A, Sol H, et al., Mixed numerical–experimental technique for orthotropic parameter identification using biaxial tensile tests on cruciform specimens, *International Journal of Solids and Structures*, **44**, 2007, pp. 1643-1656.

[6] Lamkanfi E, Van Paepegem W, Degrieck J, et al., Strain distribution in cruciform specimens subjected to biaxial loading conditions. Part 1: Two-dimensional versus three-dimensional finite element model, *Polymer Testing*, **29**, 2010, pp. 7-13.

[7] Lamkanfi E, Van Paepegem W, Degrieck J, et al., Strain distribution in cruciform specimens subjected to biaxial loading conditions. Part 2: Influence of geometrical discontinuities, *Polymer Testing*, **29**, 2010, pp. 132-138.

[8] Makris A, Vandenbergh T, Ramault C, et al., Shape optimisation of a biaxially loaded cruciform specimen, *Polymer Testing*, **29**, 2010, pp. 216-223.

[9] Serna Moreno M C, Curiel-Sosa J L, Navarro-Zafra J, et al., Crack propagation in a chopped glass-reinforced composite under biaxial testing by means of XFEM, *Composite Structures*, **119**, 2015, pp. 264-271.

[10] Serna Moreno M C, Martínez Vicente J L, López Cela J J, Failure strain and stress fields of a chopped glass-reinforced polyester under biaxial loading, *Composite Structures*, **103**, 2013, pp. 27-33.

[11] Smith E W, Pascoe K J, Biaxial fatigue of a glass-fibre reinforced composite,part 1: fatigue and fracture bahaviour, *Mechanical Engineering Publications*, 1989, pp. 367-396.

[12] Soden P D, Hinton M J, Kaddour A S, Biaxial test results for strength and deformation of a range of E-glass and carbon fibre reinforced composite laminates: failure exercise benchmark data, *Composites Science and Technology*, **62**, 2002, pp. 1489-1514.

[13] Kaddour A S, Soden P D, Hinton M J, Failure of ±55degree filament wound glassepoxy composite tubes under biaxial compression, *Journal of Composite Materials*, **32**, 1998, pp. 1618-1645.

[14] Dongtao Q, Guangxu C, Fatigue behavior of filament-wound glass fiber reinforced epoxy composite tubes under tension/torsion biaxial loading, *Polymer Composites*, **28**, 2007, pp. 116-123.

[15] Kawakami H, Fujii T J, Morita Y, Fatigue degradation and life prediction of glass fabric polymer composite under tension/torsion biaxial loadings, *Journal of Reinforced Plastics and Composites*, **15**, 1996, pp. 183-195.

[16] Fujii T, Lin F, Fatigue behavior of a plain-woven glass fabric laminate under tension/torsion biaxial loading, *Journal of Composite Materials*, **29**, 1995, pp. 573-590.

[17] Fujii T, Shiina T, Okubo K, Fatigue notch sensitivity of glass woven fabric composites having a circular hole under tension/torsion biaxial loading, *Journal of Composite Materials*, **28**, 1994, pp. 234-251.

[18] Fujii T, Amijima S, Lin F, Study on strength and nonlinear stress-strain response ofplain woven glass fiber laminates under biaxial loading, *Journal of Composite Materials*, **26**, 1992, pp. 2493-2510.

[19] Quaresimin M, Carraro P A, Damage initiation and evolution in glass/epoxy tubes subjected to combined tension-torsion fatigue loading, *International Journal of Fatigue*, **63**, 2014, pp. 25-35.

[20] Schmidt F, Rheinfurth M, Protz R, et al, Monitoring of multiaxial fatigue damage evolution in impacted composite tubes using non-destructive evaluation, *Composites Part A: Applied Science and Manufacturing*, **43**, 2012, pp. 537-546.

[21] Shokrieh M M, Lessard L B, Progressive fatigue damage modeling of composite materials, part I: Modeling, *Journal of Composite Materials*, **34**, 2000, pp. 1056-1080.

[22] Xiao Y, Kawai M, Hatta H, An integrated method for off-axis tension and compression testing of unidirectional composites, *Journal of Composite Materials*, **45**, 2011, pp. 657-669.

[23] Hashin Z, Failure Criteria for Unidirectional Fiber Composites, *Journal of Applied Mechanics*, **47**, 1980, pp. 329-334.

Research on failure strength of composites bolted joints under temperature condition

†Hongyan.Wang ¦, *Haitao Cui[1],Ying Xu[1] and Wenli Lv[1]

[1] Jiangsu Province Key Laboratory of Aerospace Power System, Nanjing 210016, P.R. China
State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing 210016, P.R. China
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China

*Presenting author: 15705185005@163.com

†Corresponding author: why03357750487@sina.com

Abstract

The method of predicting the strength of laminated composite bolted joints under temperature condition was derived, which was based on the progressive damage analysis method, and took into consideration the effects of temperatures on the composite materials properties and the total stresses. The simulation analysis of the failure modals and failure strengths of the bolted joints under 22℃、150℃、230℃ and 310℃ was conducted. The simulation results was compared with the experimental results , and the maximum error of predicted strength value was 6.32%, which shows that the method of predicting the strength of composite bolted joints and the experiment results on composite laminate bolted joints agrees well.

Keywords: Temperature condition; laminated composites; bolted joints; progressive damage; strength

Introduction

Composites are widely used in aeronautical structures, and it is difficult to avoid various connection problems in complex aviation structures. Bolt connection can withstand high load, easy loading and unloading, which has become the most important and widely

connected. However,the bolted connection structure will destroy the continuity of the composite fiber, resulting in the bearing capacity decreased, leading to premature failure of the structure, and thus the strength of the composite bolt connection structure has become a hot topic of concern to scholars. At the same time, the aviation structure often works in different temperature environment, when the resin-based composite structure in high temperature environment, the mechanical properties of the material will be reduced to varying degrees. Therefore, it is very important to study the mechanical properties of the resin-based composite bolts in the temperature environment.

Based on the above analysis. This article experimental study on the tensile strength and damage mechanism of the composite structure of composite laminates in four temperature environments was carried out. And based on the method of gradual accumulation of damage, considering the establishment of a method for forecasting temperatures of the composite tensile strength bolted joint, which method takes into account the influence of stress and temperature mechanical properties of the basic material. In order to improve the use of the forecasting method, the relationship between the basic mechanical properties and the temperature of the composite material in the range of 22℃~ 310℃was established. The results show that the simulation results are in good agreement with the experimental results. The results show that the simulation results are in good agreement with the experimental results.

Experimental approach

In order to compare the effects of different ratio of the width with hole pore and temperature on the failure mode of bolted connections. Two kinds of the ratio of the width with hole pore composit laminates of T300/BMP316, which were in the order of $[45/-45/90/0/-45/0/45/0/90/0]_{s,}$, were used for this investigation.The specimens were exposed to different temperature environments, and were numbered by T-E-W-Z where T was the tensile

strength specimen,E was the pore size,W was the test temperature, Z was different serial numbers of the same ratio of width with hole pore.

Table 1 Specimens parameters

Ratio of width with hole pore	pore size / mm	(length x width) / mm	temperature condition /℃	quantity / pieces
3	8	160x24	22	3
3	8	160x24	150	3
3	8	160x24	230	3
3	8	160x24	310	2
4	6	160x24	22	2
4	6	160x24	150	3
4	6	160x24	230	3
4	6	160x24	310	2

(a)T-8-RT-1

(b)T-8-150-1

(c)T-8-230-1

(d)T-8-310-1

Fig.1 Fracture of bolted joint

Taking the photograph of the damage test specimens, which the ratio of the width with hole pore is three,to analyze the failure mode,as shown in Fig.1.

Tensile failure, crushing failure,shear failure are the basic failure modes of the bolted joints at room temperature. It can be seen from Fig.2, that tensile failure is the main failure mode of the bolted joints at 150℃ and 230 ℃, the shear failure is the main failure mode at 310℃, and the extrusion deformation occurs at the edge of the hole, and the damage is extended to the joint end unit, indicating that failure mode of bolted connections at room temperature is the same with in a high temperature environment,but the ultimate failure mode at different temperatures are different, tensile failure mainly at 150 ℃ and 230℃, the main failure mode of the joint is shear failure at 310℃, which indicates that the high temperature has an important effect on the final failure mode of the bolt joint.

Example analysis

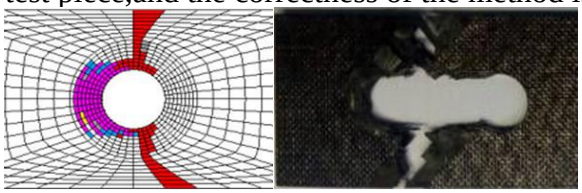
In this paper, the corresponding static tensile experiments were conducted at 22℃, 150℃, 230℃, 310℃, to valitate the prediction results. The results were compared shown in Table 2.

Table 2 Comparison of tension ultimate failure strength of bolted joints

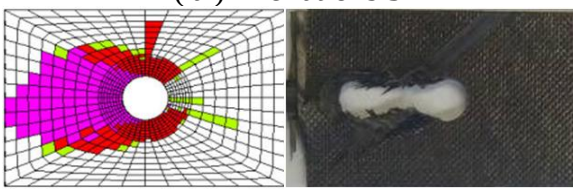
The ratio of the width with hole is three				The ratio of the width with hole is four			
temperature /℃	Experimental value /MPa	Predicted value /MPa	error%	temperature /℃	Experimental value /MPa	Predicted values /MPa	error/%
22	140.71	146	-3.76	22	153.38	158	-3.01
150	100.64	107	-6.32	150	146.84	149	-1.47
230	88.26	92	-4.24	230	141.66	146	-3.06
310	81.83	86	-5.10	310	90.42	95	-5.07

It can be seen from Table 2, that the maximum error of the predicted strength value is -6.32%,and the calculation results of each intensity are too large, which is a risky prediction.

This article only compares the damage fracture of failure mode simulation with the actual test damage fracture only of two different ratio of width with hole pore at 230℃,as show in Fig.3,From it can be seen the damage projection pattern of the bolt joint is consistent with the final failure result of the test piece,and the correctness of the method is proved.



(a) The ratio is 3



(b) The ratio is 4

Fig.2 Final failure results of experiment and simulation of jointed bolts with two different ratios under 230 °C

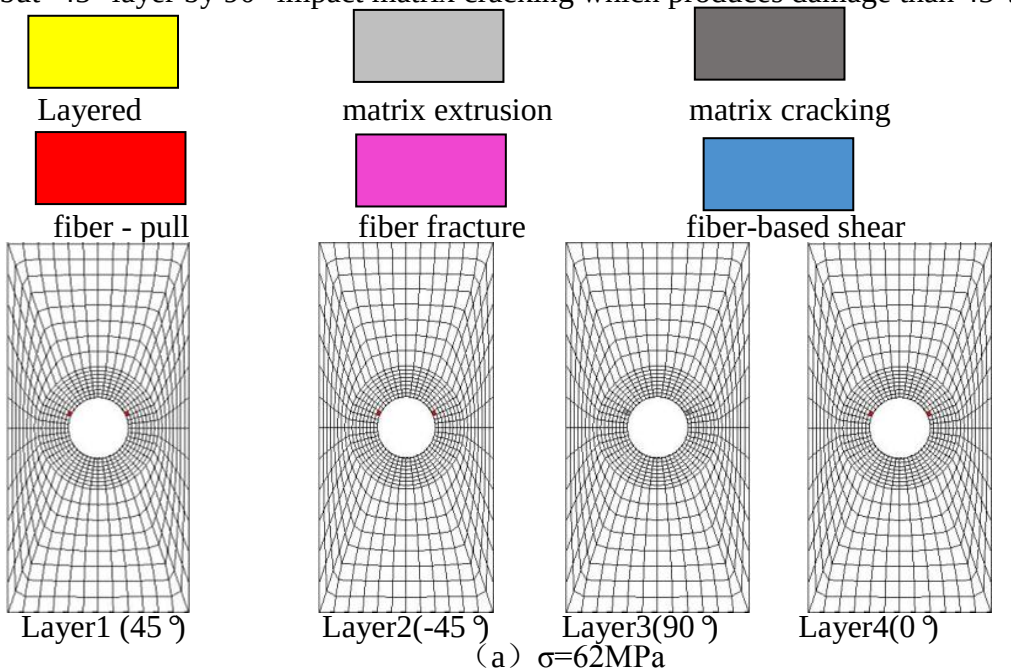
The study shows that, the injury in two groups of specimens are similar in Table 2, in order to save space, this article only take the bolted joint with the ratio of 3 at 230℃ as example to analyse the initernal damage propagation law and form of the bolted joint laminates,selecting the first four layers of the laminates(layer -45 °,0 °,45 °,90 °) as represent during the analysis process,as shown in Fig.3.

It can be seen from Fig. 3 (a) that when the load is increased to 62MPa under the uniform tensile load , the 0 °layer has the fiber breaking damage, and the damage position is at the maximum stress on the left and right sides of the connection hole. In this tensile load,the 90 °layer of the substrate cracking damage appeared more obvious expansion, 45 °and -45 °layer appeared obvious fiber breaking damage, mainly because the 45 °and -45 °layer is the matrix,And the fiber co-bearing, although the matrix failure to lose the carrying capacity, but because the fiber can still be carried to make the damage unit still has a certain carrying capacity, so with the increase in load, the original matrix cracking damage units have appeared fiber break damage. In addition, the occurrence of other substrate damage caused the redistribution of the stress field inside the laminates, which made the stress concentration more obvious, which further promoted the fracture of the bearing fiber. As the load increases further, the damage of each layer begins to extend toward the end of the joint and in the width of the plate. Fig.3(b) shows the specific damage forms and the expansion of the first four layers when the load reaches 92 MPa . At this time, the 45 °layer mainly occurs in the fiber fracture, matrix cracking and matrix cracking - fiber breakage form damage; -45 °layer occurred mainly in fiber fracture and matrix cracking, matrix cracking - fiber fracture, matrix cracking -fiber damage - fiber-based shear and other damage coexist damage; 90 °layer mainly occurred in the matrix cracking and fiber fracture, accompanied by fiber-based shear, matrix cracking – fiber fracture and other damage; 0 °layer mainly occurred

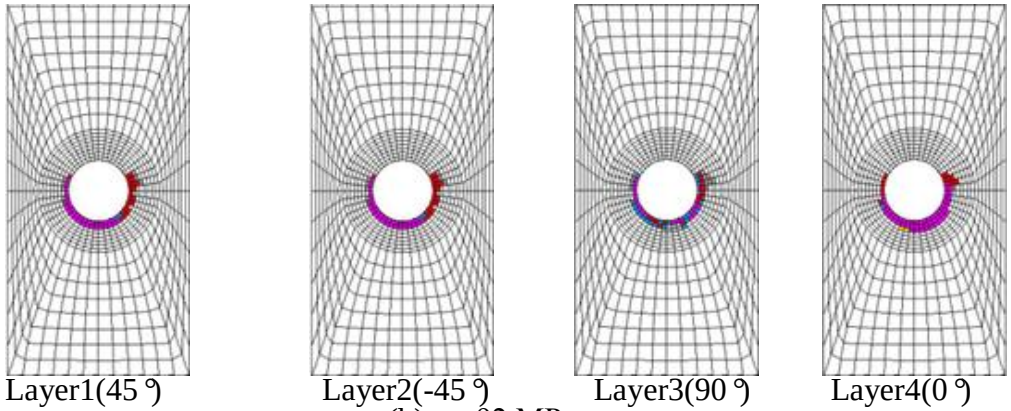
in the fiber fracture, fiber fracture - fiber-based shear damage, but also occurred a small amount of stratification and matrix cracking damage.The damage pattern of the -45 °layer unit is more diversified than that of the 45 °layer, mainly because the -45 °layer is adjacent to the 90 °layer. When the 90 °layer is damaged, the adjacent -45 °layer have a certain effect on the shear effect;

Fig.3(c) shows that when the tensile load increases to 137Mpa , the first four layers of damage to further expand the situation , then you can see the pavement damage are obvious expansion , mainly along the bolt extrusion direction ,the damage is already extended to the dense area of the dividing grid.

Fig.3(d) shows when the tensile load is increased 146Mpa specifically injury before breakage when four failure modes, it can be seen, 0 ° Layer, ± 45 °and 90 °layers have been extended to damage the plate width boundary,which indicates that the bolt failure has occurred in the final failure damage, the main form of damage is the tensile damage, which is consistent with the test results; 0 °layer occurred in the form of fiber breakage damage occurred at the same time in the joint extrusion;The damage of the 90 °layer is more, mainly due to the matrix cracking and matrix damage - fiber-based shear damage extended to the joint board wide boundary, also accompanied by a certain hierarchical fracturing and matrix - fiber-yl shear - fiber breaking pressure damage; 45 ° primary failure mode of the fiber layer is pulled off to the damage propagation in the sheet width at the boundary, while the side of the bolt holes pushing region in the 45 °direction breaking pressure significantly damage the fiber, also accompanied by some fiber damage shear; -45 ° layer when the final destruction occurs, and the damage in the form of 45 °substantially uniform layer, and with Vera off the main failure mode, in the -45 °direction hole pushing region also produces a more severe fiber damage off pressure, but -45 °layer by 90 °impact matrix cracking which produces damage than 45 °.



（a）σ=62MPa



(b)σ = 92 MPa

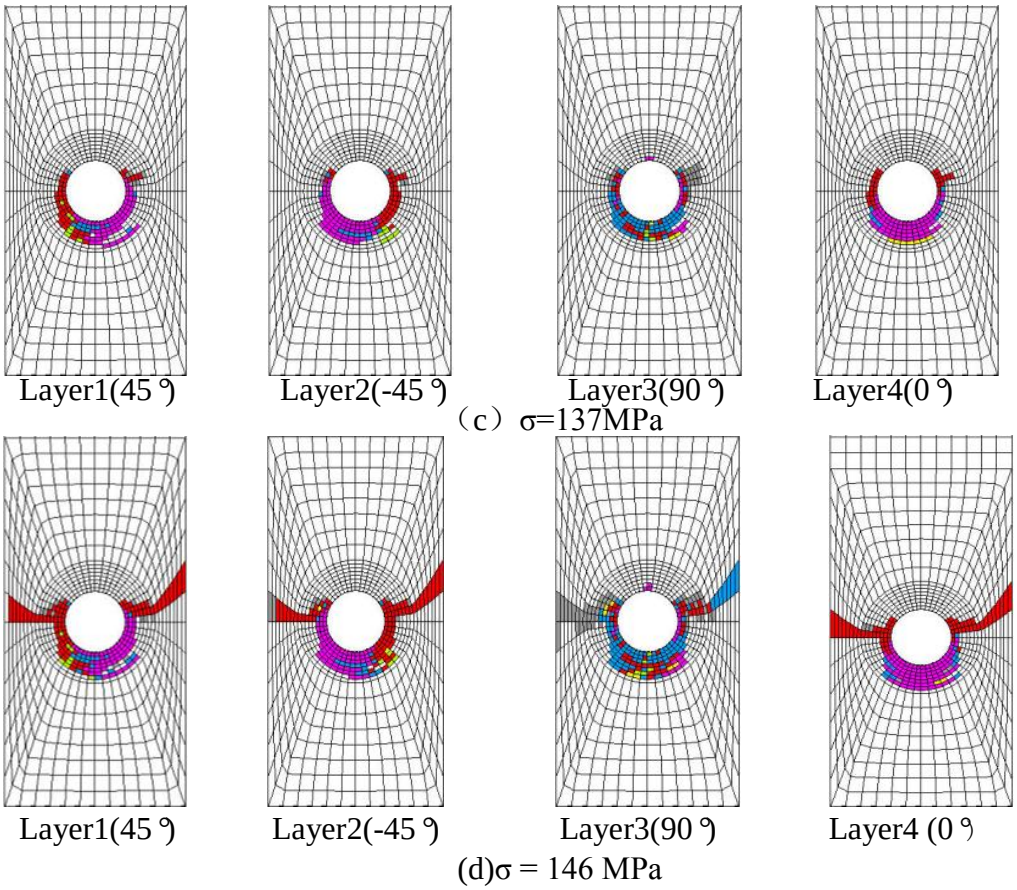


Fig.3 Progressive damage of composite bolted joint (230 °C)

Conclusion

The tensile strength of two kinds of bolt joints with different width ratio was studied at room temperature, 150 ℃ , 230 ℃ and 310 ℃.The experimental results show that with the same width ratio,the increase of the temperature of the bolt joint strength decreases gradually, and the higher the temperature, the greater the decrease. The decrease static load strength of bolt joint at 15℃， 230℃,and 310℃ is 4.26%， 7.64%,and 41.05%(ratio is 3);28.48%,37.28%and 41.84%(ratio is 4),indicating that the temperature plays a very important influence on the strength of the joint bolt; at the same temperature,the larger ration the strength value of the joint is smaller,and the smaller pore size, the more obvious the stress concentration at the joint, the lower the breaking strength.

In this paper, the experimental study on the tensile strength and damage mechanism of the composite structure of laminates with four kinds of temperature environment is carried out. Based on the method of progressive damage analysis method, the method of predicting the tensile strength of composite bolts with temperature environment is established. The effect of temperature on the stress and the basic mechanical properties of the material is taken into account. In order to improve the use of the forecasting method, the relationship between the basic mechanical properties and the temperature of the composite material in the range of 22 ℃ ~ 310 ℃ was established. The results show that the simulation results are in good agreement with the experimental results.

References

[1] ZHOU S , WANG ZQ , ZHOU JS , et al. Experimental and numerical investigation on bolted composite joint made by vacuum assisted resin injection [J]. Composites Part B: Engineering, 2013, 45 (1): 1620-1628.

[2] CAMANHO P , MATTHEWS F. A progressive damage model for mechanically fastening joints in composite laminates [J]. Journal of Composite Materials, 1999, 33 (24): 2248-2280.

[3] ZHU Y , CUI H , WEN W. Three dimensional uniform damage strength analysis of single-row multiple bolted joint in composite laminates [J]. Journal of Nanjing University of Aeronautics and Astronautics. 2012, 44 (4): 497 -502.
[4] WANG Z , ZHOU S , ZHANG J , et al. Progressive failure analysis of bolted single-lap composite joint based on extended finite element method [J]. Materials & Design. 2012,37: 582-588.
[5] Guo Zaiang , Wan small Peng , Li Bo . Consider Prediction Analysis strength delamination damage bolted composite [J]. Aeronautical Engineering Progress .2014, 5 (4): 534-540.
[6] GUO Ziang, WAN Xiaopeng, Li Bo. Strength Prediction and Damage Analysis of Bolted Composite Joints Based on a Delamination Damage Model [J]. Advances In Aeronautical Science And Engineering. 2014, 5 (4): 534-540.
[7] ATA Ş A , SOUTIS C. Strength prediction of bolted joints in CFRP composite laminates using cohesive zone elements [J]. Composites. Part B: Engineering, 2014, 58: 25-34.
[8] ZHANG J, LIU F, ZHAO L, et al. A progressive damage analysis based characteristic length method for multi-bolt composite joints [J]. Composite Structures, 2014, 108: 915-923.
[9] OLMEDO Á , SANTIUSTEC. On the prediction of bolted single-lap composite joints [J]. Composite Structures, 2012, 94 (6): 2110-2117.
[10] Zhu Yuanlin , Cui Haitao , Wen Weidong . Composite single-row multi-dimensional cumulative damage nail connection strength analysis [J]. Nanjing University of Aeronautics and Astronautics , 2012, 44 (4): 497-502.
[11] ZHU Yuanling, CUI Haitao, WEN Weidong, et al. Three dimensional integrity damage of single-row multiple bolted joint in composite laminates. Journal of Nanjing University of Aeronautics & Astronautics. 2012, 44 (4): 497-502.
[12] Wang Zhiqiang . Composite Laminates bolt connection performance analysis of [D]. Harbin: Harbin Engineering University, 2010. Only . WANG Zhiqiang ON the Analysis The Performance of Composite Laminate Bolt [D] Harbin:. Harbin Engineering University , 2010.
[13] Song Jian, Wen Weidong . Mechanical properties of resin-based composite laminates at different temperatures [J]. Journal of Aeronautical Dynamics, 2016,31 (4): 1006-1017.
[14] SONG Jian, WEN Weidong. Experiment on mechanical properties of resin matrix composites laminates under various temperatures [J]. Journal of Aerospace power. 2016,31 (4): 1006-1017
[15] Rupnowski P, Gentz M, Kumosa M. Mechanical response of a unidirectional graphite fibrous / polyimide composite as a function of temperature [J]. Composites Science and Technology, 2006, 66 (7): 1045-1055.

Simulation and experimental research on the slicing temperature of the sapphire with diamond wire

Xuerun Huang [1], Hui Huang [2], *Hua Guo [3]

[1,2,3]Institute of Manufacturing Engineering, Huaqiao University, Xiamen, 361021, Fujian Province, P. R. China

* Corresponding author: fax:+86-595-6162615; Email: guoh1214@hqu.edu.cn

[1]2461771302@qq.com, [2]huanghuihh@hotmail.com

Abstract

Temperature has an important influence on the wafer quality in the sapphire slicing process with fixed abrasive diamond wire saw. Temperature fields on the wafer surface during the slicing of sapphire ingot were studied with simulation and experiment. The effect of coolant on the slicing temperature field was explored. The simulation temperature field without coolant is good consistent with the experimental result. The maximum temperature located at the middle of slicing zone increase with the cutting depth. Coolant had a significant influence on the wafer temperature field. The narrow slicing kerf and wire movement have effect on the coolant supply, which cause the deviation of temperature field between the simulation result and experiment result.

Keywords: Slicing temperature, Wire saw; Sapphire wafer.

Introduction

Sapphire is classified as hard and brittle material and it has superior physical and chemical properties. Therefore, sapphire is often used for the optoelectronics research field [1]. During the manufacturing process of sapphire wafer, slicing is the first step and has important implications in the subsequently processing. The emergence of fixed diamond wire saw has significantly improved the slicing efficiency and wafer quality by generating smaller surface damage [2].

A massive source of heat was generated during the cutting process. A part of heat is carried away by the coolant while the rest is absorbed by the ingot and dispersed on it. Some studies have shown the relationship between temperature and surface quality of the wafers [3]-[4]. The uneven temperature field caused by slicing is severely detrimental to the processing quality. It could further increase the undesirable warping of wafers by the uneven thermal expansion of the sapphire ingot. To explore this issue, it is critical to research the slicing temperature in the cutting process. Therefore, in the past decade, researchers have paid mainly attention to the slicing temperature of brittle material. Sumeet Bhagavat conducted a finite element 2D model to analyze and synthesize temperature variation of silicon wafers during multi-wire saw slicing [5]. Lars Johnsen studied the heat transfer experimentally and computationally solved with the steady state 3D model by the Computational Fluid Dynamics (CFD) software during multi-wire sawing of silicon wafers [3]. Recently, Shinya Moriyama used the infrared camera to measure the temperature distribution of the sapphire during multi-wire sawing and measured the local dynamic temperature of the ingot by using K-type thermocouple [6]. Among the previous reports, simulation has emerged as an essential candidate because of its impressive convenience. Hence, it is no surprise that combination of simulation and experiment have been targeted as the best choice to study the single-wire

sawing temperature.

In this paper, the finite element method (FEM) is employed to simulate the temperature field of the sapphire with fixed abrasive diamond wire saw. The sawing tangential force was measured to calculate the sawing heat. The sawing temperature was also measured by infrared thermal imager camera to verify simulation results. The effect of coolant on the wafer temperature field was discussed.

FEM Modeling

FEM model

The commercial finite element package ANSYS® was employed to perform the FEM modeling of the sapphire in slicing process. The sapphire ingot was modeled as a cylinder with a diameter of 50.8 mm. The length of ingot was cut to 15 mm under the premise that the calculation accuracy is guaranteed. In full accordance with the actual experimental situation, the wafer thickness considered in the FEM model is 1 mm and the kerf width is 0.25 mm. The following assumption is made to obtain the simplified model: the wire is assumed to move straight through the slicing kerf with a constant gap distance. Although the wire is flexible and bended due to forces acted upon it in the slicing kerf in practice.

For more realistic results, the properties of the C-plane sapphire were initially introduced in the FEM model in accordance to the literature [7]-[8]. The main constitutive parameters are heat capacity, conductivity, and coefficient of thermal. The values of these parameters, as well as Poisson ratio, density and elasticity modulus of the modeled materials are summarized in Table 1. For simplification purposes, the parameters are equal to their values at normal temperature and under small changes in the surrounding water temperature.

Table 1. Material properties

Material	C-plane sapphire
Density (kg/m^3)	3980
Elasticity modulus (Pa)	$3.3251 \cdot 10^{11}$
Poisson ratio	0.28
Heat capacity (J/kg K)	757.3
Conductivity (W/m K)	49.97
Coefficient of thermal expansion (/K)	$5 \cdot 10^{-6}$

Boundary conditions

In a slicing system, mechanical energy is transformed into thermal energy. The heat quantity in the contact area is generated by the friction forces. The energy dissipated in the form of heat can generate temperature. The temperature of the ingot will change through heat conduction and convection when the heating disperse freely with movement of wire saw, thereby influencing the wafer temperature distribution. The temperature field of the wafer changes with cutting depth and time.

The objective of thermal analysis is to assess the temperature distribution in the wafer surface with different cases. Input parameters used in simulation are material property, heat flux, film coefficient and bulk temperature. Output parameters are the temperature profile of wafer surface.

This is a transient thermal analysis as a function of time with three boundary conditions:

(1) An initial temperature is equal to ambient temperature.
(2) A heat flux entering the contact zone of wire saw and slicing kerf.
(3) A heat transfer by convection applied to all the free surfaces of the ingot. And the heat transfer coefficient depends on experiment conditions because it varies with environment.

The unsteady heat exchange equation is as follows:

$$\frac{\partial t}{\partial \tau} = \frac{1}{C(t)\rho}\left[\frac{\partial}{\partial x}\left(\lambda\frac{\partial t}{\partial x}\right) + \frac{\partial}{\partial y}\left(\lambda\frac{\partial t}{\partial y}\right) + \frac{\partial}{\partial z}\left(\lambda\frac{\partial t}{\partial z}\right)\right] + \frac{q_n}{C(t)\rho} \quad (1)$$

Where τ is the time, t is the temperature, C is the specific heat capacity, ρ is the density, λ is the thermal conductivity, and q_n is the internal heat source. In this case, $q_n = 0$.

The initial condition was specified by setting the initial temperature to 23℃ unless otherwise mentioned. The initial temperature of the ingot is constant.

$$T(x, y, z, \tau) = 23\ ℃ \quad \text{at time } \tau = 0\ s \quad (2)$$

The determination of the heat transfer coefficients is essential. It is difficult to obtain by exact calculation because these coefficients depend on the location and the construction of the workpiece, the ingredient of coolant fluid, the flow rate of the coolant and, furthermore, air circulation. The convection coefficient of the ingot was determined in this case because the process of heat transfer by radiation is not significant. For the ingot, a convective heat condition simulating cooling was used due to the coolant splashing on it. A heat transfer coefficient for forced cooling by the coolant was applied. The forced convection coefficient for the surrounding of sapphire is h_1 = 1000 W/m^2℃ [9]. Despite the forced cooling, the sapphire ingot was also cooled by the surrounding air at the surface. The boundary condition of the ingot surface is considered free convection heat transfer with outside air. The natural convection coefficient for the front and back of sapphire ingot h_2 = 5 W/m^2℃ [5], the natural convection coefficient for the cutting planes of sapphire is h_3 = 1 W/m^2℃ [5].

Determination of heat flux

Heat flux during any machining can be defined as the power supplied for machining per unit area. Thus, heat flux q entering the kerf can be expressed as:

$$q = \frac{\varepsilon P}{V} \quad (3)$$

Where P is the consumed power for the wire slicing process, ε is the ratio of the power that transform to the heat and is absorbed by the sapphire. ε was defined as 0.667 according to the reference [4]. V is the total material removal rate being machined.

The consumed power P can be expressed as:

$$P = F_c v_c \quad (4)$$

Where F_c is the tangential force during slicing process, v_c is the velocity of the wire during slicing.

The total material removal rate can be expressed as:

$$V = Ldv_w \tag{5}$$

Where d is the width of the sawn kerf, v_w is the feed speed. L is the contact length between the workpiece and wire saw. For this case, the contact length L insists on changing with the increase of cutting depth due to the circular shape for the sapphire wafer. The total material removal rate increases with the increase of L, therefore, the amount of heat entering the workpiece also increases.

Take Eq. 4 and Eq. 5 to the Eq. 3, then the heat flux is described as:

$$q = \varepsilon \frac{F_c V_c}{LdV_w} \tag{6}$$

FEM meshing

With regards to the model meshing, eight-node 3D elements SOLID70 with each node having a single degree of freedom----temperature which used to model conduction heat transfer problems were considered. These elements were chosen because they show a faster convergence than the tetrahedral elements with respect to mesh refinement.

Element size is treated to be the most influencing constituent in the cutting process simulations. The reference [10] confirms that suitable element length should be chosen to achieve a balance between computational accuracy and calculation time.

The sizes of elements were designed. The length and the width of the elements in the wafer surface are both 2 mm, the three kinds of thickness of the wafer element is 0.3 mm, 0.5 mm, 0.7 mm respectively. For a better computational performance, a refined mesh should be employed on the contact area. Therefore, fine meshes were considered for the kerf. The kerf element is 0.05 mm, 0.08 mm and 0.125 mm respectively. The remainder elements were also extruded with 8 different sizes. However, it would be computationally expensive to use this fine mesh for the entire specimen. Hence, spacing ratio of remainder is proposed. This can create elements with gradual increased size to reduce grid number while develop a smooth connection between kerf and remainder. The spacing ratio of remainder that perpendicular to the wafer surface direction (z-direction) is 3 in all cases.

The sizes of grids were varied to conduct 16 groups test was shown in Table 2. FEM mesh-independence in slicing process simulations was certified as shown in Fig.1. From Fig. 1, it can be seen that for the more mesh grids, the deviation between simulated results becomes less and less. It can be concluded that the meshing method should be choose which temperature is insensitive to the element size with least calculation time.

Table 2. Variation of the sizes of grids

size(mm) / numbering	wafer surface	wafer	kerf	remainder	grid number
1	2x2	0.3	0.05	3	16000
2	2x2	0.5	0.08	3	13600
3	2x2	0.7	0.125	3	12000
4	2x2	0.3	0.05	5	12800
5	2x2	0.5	0.08	5	11600
6	2x2	0.7	0.125	5	8800
7	2x2	0.3	0.05	7	12000

8	2x2	0.5	0.08	7	9600
9	2x2	0.7	0.125	7	8000
10	2x2	0.3	0.05	10	10400
11	2x2	0.5	0.08	10	8000
12	2x2	0.7	0.125	10	6400
13	2x2	0.5	0.08	0.3	82400
14	2x2	0.5	0.08	0.5	52000
15	2x2	0.5	0.08	1	28800
16	2x2	0.5	0.08	2	17600

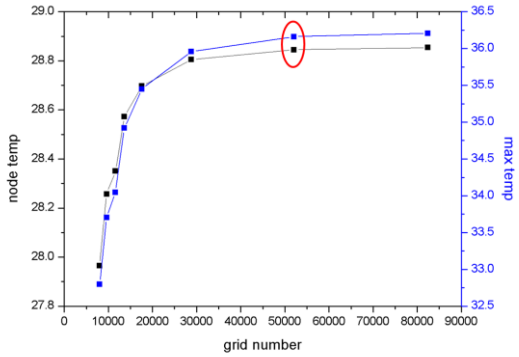


Figure 1. Profile of the grid independence of the model.

By comparison, the suitable result is chosen as shown the red circle in Fig.1. The length and the width of the elements in the wafer surface are both 2 mm, the thickness of the wafer element is 0.5 mm. A column of elements with width equal to a quarter of the thickness of the kerf is constructed for the part of kerf. The size of remainder element is 0.5 mm. Total number of elements and nodes were 25600 and 27456. Thus, it was concluded that the proposed grid resolution was sufficient. A better view of the model and the meshing distribution is presented in Fig. 2.

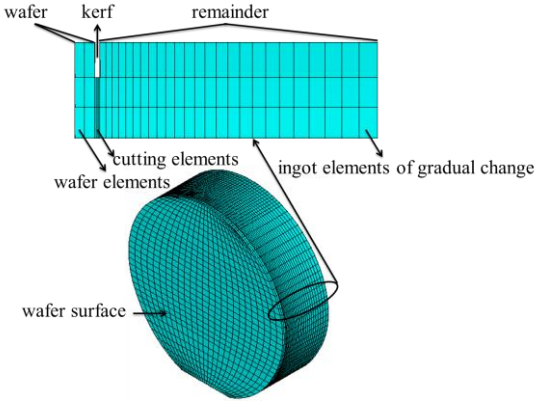
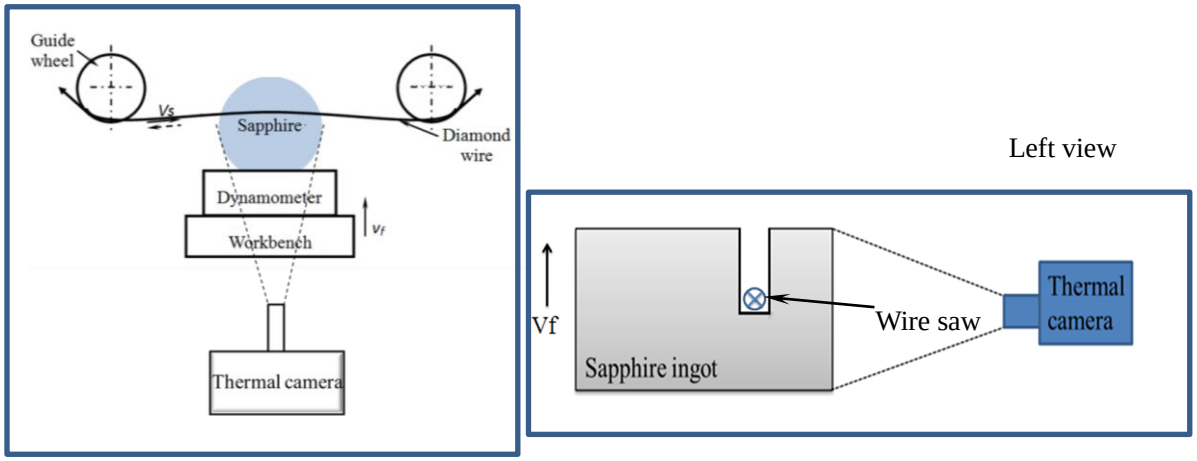


Figure 2. Finite element model with cut elements.

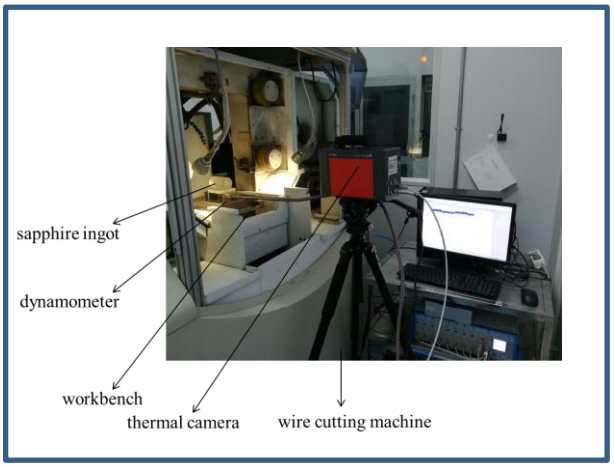
The wafer slicing is a continuous process, a method to express material removal in sawing was proposed by Toshiro [4]. The elements in sliced zone were named “cutting elements”. While the cutting elements were sliced, heat was transfer to them. After the cutting elements had been sliced completely, all of stiffness, heat conductivity and thermal capacity of them become zero. With the operation, the cutting elements were made exception from solution.

Experimental Conditions

Sawing experiments were conducted on JXQ-1201 single wire reciprocating wire cutting machine. The diamond wire saw produced by Asahi Company was 0.25 mm in diameter with 30 to 40 um diamond grits run reciprocally between the two rollers on the machine. Wire tension was 30 N. During this experiment, the wire speed was 5.83 m/s and the feed rate was 0.3 mm/min. Fig. 3 presents images of the apparatus. The specimen material was C-plane sapphire. The ingot was 50.8 mm in diameter and 100 mm in length.



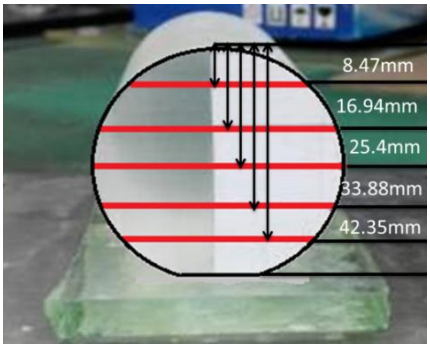
(a) Illustration of the wire slicing setup. **(b) Left view of wire slicing setup**



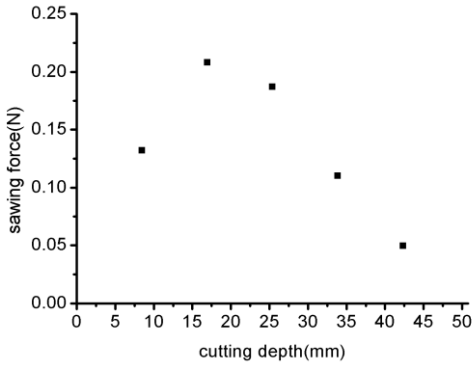
(c) Picture of wire slicing setup.

Figure 3. The wire slicing setup.

A piezoelectric dynamometer Kistler 9119AA2 was equipped to measure the sawing forces. The horizontal force measured by the dynamometer was acted as the tangential force due to the small wire deflection during slicing process. Five different cutting depths divided the wafer equally were selected to measure the sawing force signals. The cutting depths and the corresponding cutting lengths were shown in Fig. 4(a). Two sawing force signals at least at each position were recorded and the average value was taken as the final sawing force for this position. The sawing forces varied with the cutting depth of C-plane sapphire were shown in Fig. 4(b). With the increase of the cutting depth, the tangential force of the material was increase firstly and then decrease.



(a) Measuring Position marked with red line and crystal ingot outline indicated by black line.



(b) Sawing forces with different cutting depths.

Figure 4. Variation of the sawing forces with respect to the cutting depth of C-plane sapphire.

An Image IR 5325 thermal camera was used to determine the temperature distribution on the wafer surface at the same time of the force measurement. The thermal camera was mounted on the camera holder. The distance between the camera and the sapphire ingot was kept at 20 cm during the whole slicing process. The temperature distribution was measured at five cutting depths with and without fluid. In the experimental process, water-based cutting fluid was used, and the flow rate of cutting fluid was 2 L/min. Based on the literature [7], the emissivity of sapphire varies in the interval $0.53<\varepsilon<0.55$ for 273 K<T<313 K. In the present work, an emissivity of sapphire was set as 0.54. The wavelength of the thermal camera is $3.7\mu m<\lambda<4.8\mu m$. According to the preliminary experiment results, the measurement temperature ranged from -10℃ to 40℃ was used as thermal camera setting.

In order to stabilize environmental temperature, room temperature was maintained at 23℃. Tangential force, wire speed, feed speed and the corresponding cutting length were recorded to calculate heat generated by slicing process.

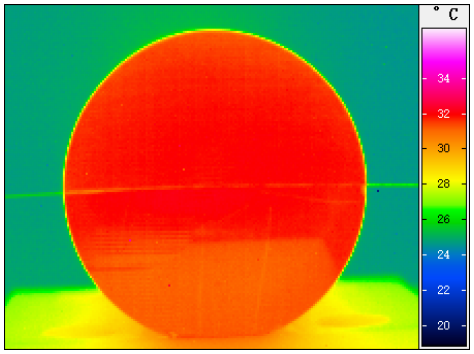
Results and Discussion

Experimental and simulation result of slicing temperature without coolants

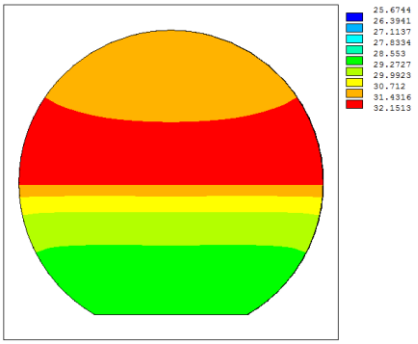
Fig. 5 shows the experimental and simulation result of slicing temperature without coolant when the cutting depth is 25 mm. The temperature field distribution from experiment is shown in Fig. 5(a). While cutting the ingot without coolant, the temperature field of sapphire wafer is more than 30℃, which is obviously higher than the ambient temperature. The

ambient temperature reaches at about 25℃, 2℃ higher than the initial temperature because the more heat was transfer to the surrounding air. The temperature of wire saw is lower than that on wafer surface which may due to the good thermal conductivity properties of metal wire. For the sapphire wafer, the temperature field below the wire saw is slightly less than that of other regions as shown in Fig. 5(a).

Simulation result is shown in Fig. 5(b). The high temperature area is concentrate on the region over the wire saw. In the wire saw cutting zone, the temperature remains constant. The temperature decrease slightly with the distance from the wire saw. The temperature below the wire saw is lower than that of other region, which is similar with the experimental results.



(a) Temperature field from measurement without coolant.



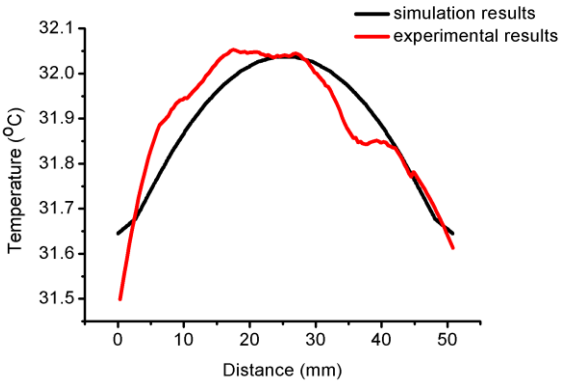
(b) Temperature field from simulation without coolant.

Figure 5. Results from simulations and measurements without coolant when the wire is half way through the ingot of dry cutting.

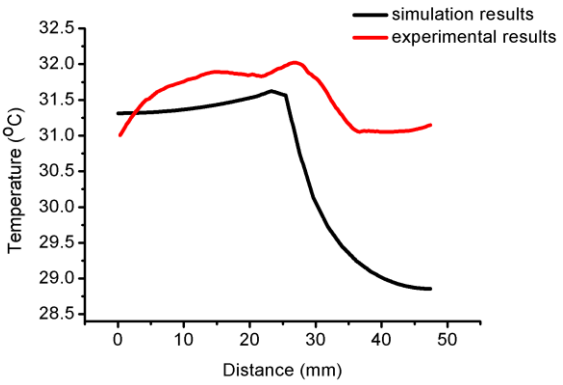
The temperature along the horizontal direction at the position of 25 mm cutting depth is shown in Fig. 6(a). The simulation temperature curve is similar with the experimental temperature curve. It is shown that the maximum temperature is about 32 ℃ as shown in Fig. 6(a). The simulation temperature profile is fairly axisymmetric, as shown the black curve in Fig 6(a). However, the experimental temperature profile (red curve) is a slightly different. The temperature peak is deviate slightly between measured temperature and calculated temperature. It is indicated that the temperature near the exit region is higher than that near the enter region..

The temperature along the vertical direction is shown in Fig. 6(b). It is found that the simulation temperature varies greatly with the vertical position. The temperature is increase slightly along the vertical direction and the highest temperatures are reached at the wafer center that is the position of wire saw. The maximum temperature is 31.5 ℃ then it decreases to 28.7℃ rapidly over distance. The measurement temperature curve variation along the

vertical direction is similar with the simulation results. The temperature is increase slightly firstly and then decrease sharply, as shown the red curve in Fig. 6(b). However, the variation range of temperature value for the experiment is smaller than the case for the simulation, as shown in Fig. 6(b).



(a) Temperature along the saw wire from the wire entrance to the wire exit.



(b) Temperature on wafer surface perpendicular to the saw wire.

Figure 6. Results from simulations and measurements without coolant when the wire is half way through the ingot of dry cutting.

The maximum temperatures obtained by simulation and measurement for each measured depth is shown in Fig. 7. It is found that the experimental results are very consistent with the simulation results. The maximum temperature increases linearly from 26℃ to 35℃ with the cutting depth from 8.47mm to 42.35 mm. Heat accumulation is the main cause of the temperature increase. At position of cutting depth 42.35 mm, the simulation temperature is slightly small than the experiment results, which may due to the effect of fixture of sapphire ingot.

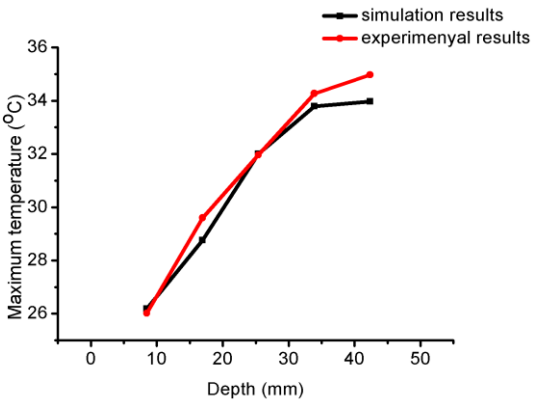
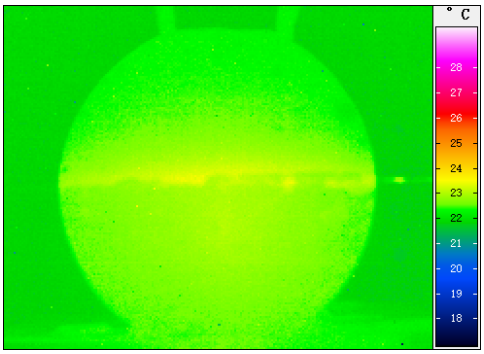


Figure 7. Results from simulation and measurement of temperatures when the wire is through the five cutting depth that divide the ingot equally.

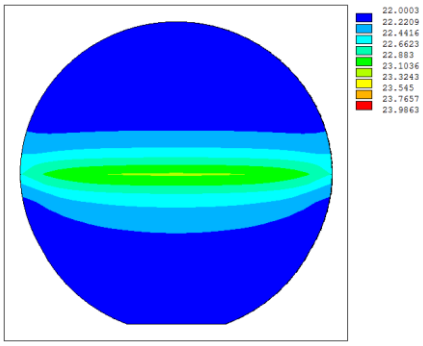
Experimental and simulation result of slicing temperature with coolants

Fig. 8 shows the experimental result and simulation result of slicing temperature with coolant when the cutting depth is 25 mm. The experimental temperature field of sapphire wafer and ambient has little variation as shown in Fig. 8(a). The experimental temperature of sapphire wafer ranged from 22℃ to 23℃ for the wire slicing with coolant, which is much lower than the case without coolant. Ambient temperature remains at 22℃ as shown in Fig. 8(a), which indicates that the more heat was taken by the coolant. The wire saw is the bright yellow line in the picture; it indicates the zone in which heat is generated more than other region. However, the temperature of wire saw is only about 24℃.

Simulation temperature of sapphire wafer is shown in Fig. 8(b). The high temperature zone is concentrate on the position of wire saw, which is similar with the experimental result. The temperature in the middle zone is higher than that in the two sides as shown in Fig. 8(b). The temperature upper and below the wire saw position is almost equal to the initial temperature. The whole temperature field of simulation result is almost consistent with the case of experimental result.



(a) Temperature field from measurement with coolant.



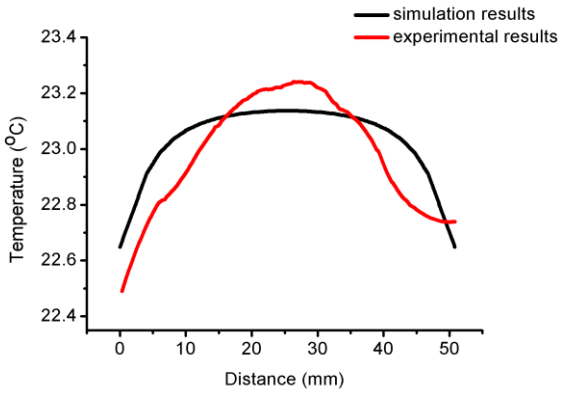
(b) Temperature field from simulation with coolant.

Figure 8. Temperature field from simulation and experiment with coolant when the wire is half way through the ingot.

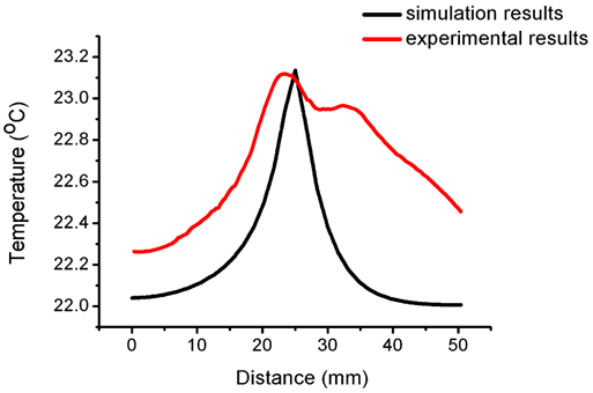
Fig. 9(a) is the temperature curves along the horizontal direction at the position of 25 mm cutting depth. Both simulation temperature curve and experimental temperature curve show that the temperature in the middle position is higher than that in the two sides position. The experimental temperature is higher than the simulation temperature in the middle position,

and is lower than simulation result in two sides. The simulation temperature profile is fairly axisymmetric, as shown the black curve in Fig 9(a). The maximum temperature located in the middle of the temperature curve is about 23.1℃ for simulation and 23.25℃ for experiment respectively. It is similar with the experimental results presented by Shinya [Shinya (2015)], which carried out the sapphire slicing with CeO_2 abrasive in water.

Fig. 9(b) shows the temperature curves of sapphire wafer perpendicular to the wire saw. The maximal temperatures are in the central area of the wafer and the values are similar. Unlike the temperature curve without coolant, the temperature increases sharply to the highest value along the vertical direction, and then decreases sharply as shown in Fig. 9(b). The profile of simulation result is more steeply than that of experimental result. The measured temperature profile has two peaks, as shown the red line in the Fig. ((b), which may due to the influence of the grits insert surface around the cutting zone.



(a) Temperature along the saw wire from the wire entrance to the wire exit.



(b) Temperature on wafer surface perpendicular to the saw wire.

Figure 9. Temperature profile results of simulation and experiment with coolant when the wire is half way through the ingot.

The maximum temperature with coolant obtained by simulation and measurement for each measured depth is shown in Fig. 10. It is found that the experimental temperature increases from 23.1℃ with the slicing depth until to 23.24℃ at the middle position and decreases slightly to 23.19℃ as shown the red line in the Fig. 10. However, the simulation temperature is keep at 23.1℃ except 24.1℃ at the position of slicing depth 16.94 mm.

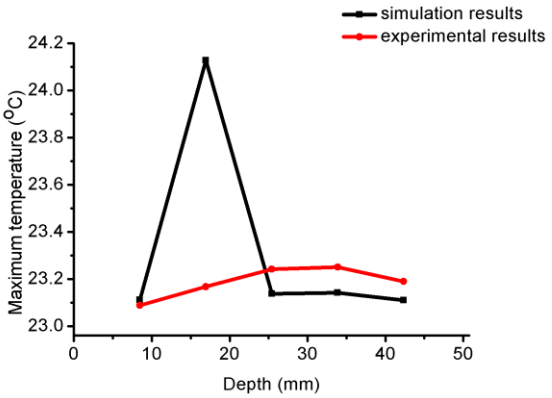


Figure 10. Results from simulation and measurement of temperatures when the wire is through the five cutting depth that divide the ingot equally.

Discussions

In the sapphire slicing process, the heat is generated in the slicing zone and transfers to the workpiece, wire saw and surrounding medium such as surrounding air, coolant flow. The final temperature field of sapphire wafer is the result of the heat transferred to sapphire ingot and the coolant condition. A good coincidence between the simulation result and experimental result without coolant shown that heat source model used in the FEM model is consistent with experimental practice. The coolant has significant effect on the temperature field of sapphire wafer, The coolant drastically reduce the slicing maximum temperature from 32℃ to 23.1℃ which is proved not only the experimental results but also the simulation result, as compared with Fig. 5 and Fig. 8. Taking into account Eq. 6, ε is represents the fraction of the power that is transforms to the heat and absorbed by the sapphire. This dimensionless parameter is significantly influenced by the cooling condition. The effect of coolant should be discussed in detail because there are mismatch between the simulation results and experimental results. In the FEM model, the heat transfer coefficient for forced cooling by the coolant was uniformly acted on the workpiece. However, it is not coincide with the actual conditions.

A narrow kerf was formed through the wire slicing process. The narrow slicing kerf has significant effect on the coolant supply. The coolant was supplied to the machining zone by the cooling tube mounted two side of workpiece. The enough coolant was provided to kerf and transferred the machining heat in initial stage of slicing. With the increase of the slicing depth, the length of slicing arc increases as shown in Fig. 4(a). It is difficult for the enough coolant to enter the whole slicing zone, especially in the middle of slicing zone. Therefore, the temperature in the middle region is higher than that in the two sides as shown in Fig. 9(a). With the increase of slicing zone, the coolant provided to the middle region reduces and the slicing temperature elevates. When the slicing depth is more than half of wafer diameter, the slicing zone reduce with the increase of slicing depth. The coolant condition is improved; therefore, the slicing temperature remains stable, as shown in Fig. 10.

Wire movement also has effect on the coolant supply. The wire saw is covered by a thin layer of coolant when it enters into the slicing zone, which pushes coolant into the slicing kerf. It can be found the profiles of numerical results are symmetric, but it was found that the temperature peak is deviate slightly between measured temperature and simulated temperature. The temperature near the exit is higher than that far away from the exit in the actual slicing process in the horizontal direction, as shown in Fig. 9(a). More cold air was taken into the slicing zone by the wire movement, so temperature peak deviation was also observed in Fig. 6(a).

According to the above discussion, it is found that numerical simulation and experimental measurement on the sapphire slicing process with different cooling conditions were conducted, the simulated results are found to be acceptable. The uneven coolant condition should be taken into account to improve the simulation accuracy.

Conclusions

A FEM model was successfully established to simulate the temperature field of sapphire wafer in the slicing process. The simulation result is good consistent with the experimental result in the condition without coolant. However, the maximum temperature with coolant obtained by simulation is lower than that obtained by experiment due to the coolant supply method. No matter with or without coolant, the maximum temperature located in the middle of slicing zone. The maximum temperature without coolant increases linearly from 26℃ to 35℃ with the cutting depth. Coolant significantly reduces the temperature field of sapphire wafer. The maximum temperature with coolant increases with the cutting depth with coolant due to the effect of narrow slicing kerf on the coolant supply.

Acknowledgment

Xuerun Huang, Hui Huang and Hua Guo would like to acknowledge the financial supports from the National Natural Science Foundation of China (51375179, U1305241).

References

[1] MS Akselrod, FJ Bruni. (2012). Modern trends in crystal growth and new applications of sapphire. Journal of Crystal Growth, **360(1)**:134–135.
[2] C. H. Chung, G. D. Tsay, M. H. Tsai. (2014). Distribution of diamond grains infixed abrasive wire sawing process. International Journal of Advanced Manufacturing Technology, 73(9-12):1485–1494.
[3] Johnsen, L., Olsen, J. E., Bergstrom, T. et al. (2012). Heat Transfer During Multiwire Sawing of Silicon Wafers. Journal of Thermal Science and Engineering Applications, 1–3.
[4] Toshiro, Yamada. (2004). Warpage Analysis of Silicon Wafer in Ingot Slicing by Wire-Saw Machine. Materials Processing and Design: Modeling, Simulation and Applications, 1459–1463.
[5] Sumeet, Bhagavat. and Imin, Kao. (2008). A finite element analysis of temperature variation in silicon wafers during wiresaw slicing. Machine Tools ＆ Manufacture , **48**: 95–106.
[6] Shinya, M. (2015). Study on the slicing of sapphire using a wire tool. Advanced Materials Research, 333–337.
[7] Wakaki, M., Kudo, K. and Shibuya, T. (2007). Physical Properties and Data of Optical Materials. Optical Science and engineering, 383–384.
[8] Dobrovinskaya, E. R., Lytvynov, L. A. and Pishchik, V. (2008). Sapphire：Material, Manufacturing, Applications. Springer Science and Business Media, 12–54.
[9] Incropera, F. P., and Dewitt, D. P. (1996). Fundamentals of Heat and Mass Transfer. Wiley & Sons, Hoboken, NJ.
[10] Ravindra, Ambati. and Yuan, Huang. (2011). FEM mesh-dependence in cutting process simulations. The International Journal of Advanced Manufacturing Technology, 313–323.

Application of anisotropic Hyperelastic constitutive model in finite element analysis of flexible deformable bump

†Ying.Xiang[1], *Haitao Cui[1] and Hongjian Zhang[1]

[1] Jiangsu Province Key Laboratory of Aerospace Power System, Nanjing 210016, P.R. China
State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing 210016, P.R. China
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China

*Presenting author:15705186098@163.com

†Corresponding author:1455466692@qq.com

Abstract

The three-dimensional nonlinear bump model is established by using ABAQUS software. The anisotropic hyperelastic model established by the strain energy density function , which is used to simulate the property of cord - rubber composites in the flexible bump. The main direction of the material can simulate the geometrical and material nonlinearity of the cord - rubber composite. By modeling the bump structure as a whole, the model is simplified and the calculation efficiency and calculation accuracy are improved. The deformation height and inner stress of the bump structure under different inflation pressures are obtained by three-dimensional nonlinear analysis, the large deformation of cord - rubber composite is also taken into account in the analysis.

Keywords: bump; anisotropic hyperelastic; cord - rubber composites; nonlinear analysis; deformation

Introduction

The cord - rubber composite is a flexible composite material formed by the effective combination of the cord and the rubber . It is widely used in tires, inflatable springs, air bags and conveyors. The rubber is a material with a low elastic modulus and a high elongation, while the cord is a material with a high modulus of elasticity and a low elongation. The composite exhibits a high toughness and a high strength by the cooperation of the substrate and the fibers high strength characteristics. Because of the obvious anisotropy and nonlinearity of the cord rubber material during the bearing, the anisotropic hyperelastic constitutive model[1]-[4] can well characterize the characteristics of the cord - rubber composite , and has good engineering practicability

In this paper, ABAQUS is used to establish the finite element model of the bump. By the UANISOHYPER_INV interface in the ABAQUS do secondary development of the constitutive model, called the written UANISOHYPER_INV subroutine as cord - rubber composite constitutive equations during the finite element calculation. This can more accurately simulate the deformation behavior of the cord - rubber composite structure.

1. Anisotropic hyperelastic constitutive model

Anisotropic elastic constitutive model is based on the continuum mechanics of fiber reinforced composite material[5]-[7]. the fiber composite material having a direction, the strain energy function can be expressed as the function of the right Cauchy-Green deformation tensor $\mathbf{C}=\mathbf{F}^T\mathbf{F}$ and the invariant I_i which is related to the original fiber directional vector $\mathbf{a}_0$. Generally have the following form, Here $\mathbf{F}$ is the deformation gradient tensor

$$W(\mathbf{C},\mathbf{a}_0)=W_{iso}(I_1,I_2,I_3)+W_{triso}(I_4) \qquad 1\text{-}1$$

Where W_{iso} is the isotropic hyperelasticity of the the rubber, W_{triso} is the anisotropic strain energy due to the reinforcement of the cord; the strain tensor invariant I_1,I_2,I_3 characterizes

the isotropic property of the material, which I_4 is related to the direction of the material. Where the invariants are given by

$$I_1 = \mathrm{tr}(\mathbf{C});\ I_2 = \frac{1}{2}[(\mathrm{tr}\mathbf{C})^2 - \mathrm{tr}(\mathbf{C})^2]$$
$$I_3 = \det(\mathbf{C});\ I_4 = a_0 \cdot C \cdot a_0 \qquad 1\text{-}2$$

Where $\mathbf{c} = \begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix}$, λ_i represents a ratio of three principal directions of elongation material ($\lambda_i = 1 + \varepsilon_i$, ε_i represents the strain of principal directions).

In an ideal conditions, the rubber can be regarded as incompressible material, The energy W_{iso} stored in the matrix is given by the Mooney-Rivlin model[8]-[10] form

$$W_{iso} = \sum_{ij=0}^{ij=n} C_{ij}(I_1 - 3)^i (I_2 - 3)^j$$

The Mooney-Rivlin two-parameter strain energy function , which is widely used in ABAQUS finite element software , can be obtained by the first two terms of the equation in the case of sufficiently small deformation

$W_{iso} = C_{10}(I_1 - 3) + C_{01}(I_2 - 3)$

Where C_{10}、C_{10} are the material parameters, the unit is Mpa .

The strain energy of the cord is generally considered to be related to the elongation of the cord and equal to the square of the cord draw ratio. Therefore, a polynomial is defined to represent the strain energy of the cord. When the cord is compressed, Can be ignored. The strain energy function of the cord is as follows

$$W_{triso} = \begin{cases} k_1(I_4 - 1)^2 + k_2(I_4 - 1)^3 + k_3(I_4 - 1)^4, I_4 > 1 \\ 0 \end{cases}$$

Where k_1、k_2、k_3 are material parameters, in units of Mpa .

In order to obtain a stress-strain constitutive relation of the form, according to the formula 1-1 derived composite material the second Piola- Kirchhoff stress tensor $\mathbf{S}$ is

$$\mathbf{S} = 2\frac{\partial W(\mathbf{C}, \mathbf{a}_0)}{\partial \mathbf{C}} = 2\sum_{m=1}^{4} \frac{\partial W(\mathbf{C}, \mathbf{a}_0)}{\partial I_m} \frac{\partial I_m}{\partial \mathbf{C}}$$

Where the invariant I_m pairs of right Cauchy - Green deformation tensor partial derivatives are as follows:

$$\frac{\partial I_1}{\partial \mathbf{C}} = \mathbf{I} \qquad \frac{\partial I_2}{\partial \mathbf{C}} = I_1\mathbf{I} - \mathbf{C}$$
$$\frac{\partial I_3}{\partial \mathbf{C}} = I_2\mathbf{I} - I_1\mathbf{C} + \mathbf{C}^2 \qquad \frac{\partial I_4}{\partial \mathbf{C}} = \mathbf{a}_0 \otimes \mathbf{a}_0$$

Cauchy stress tensor

$$\boldsymbol{\sigma} = J^{-1}\mathbf{F}\mathbf{S}\mathbf{F}^T = \frac{2}{J}\left[(W_1 + W_2 I_1 + W_3 I_2)\mathbf{B} - (W_2 + W_3 I_1)\mathbf{B}^2 + W_3\mathbf{B}^3 + I_4 W_4 \mathbf{a} \otimes \mathbf{a}\right]$$

Which W_i represents the partial derivative of the invariant, $\mathbf{B}$ represents the left Cauchy-Green deformation tensor, and the post-deformation cord direction $\mathbf{a} = \frac{1}{\lambda_F}\mathbf{F}\mathbf{a}_0$,(λ_F representing the elongation ratio of the cord)

$$\begin{cases}\sigma_1 = 2\left[(W_1 + W_2 I_1)\lambda_1^{\,2} - W_2\lambda_1^{\,4} + W_4\lambda_1^{\,2}\cos^2\alpha\right] \\ \sigma_2 = 2\left[(W_1 + W_2 I_1)\lambda_2^{\,2} - W_2\lambda_2^{\,4} + W_4\lambda_2^{\,2}\sin^2\alpha\right] \\ \sigma_3 = 2\left[(W_1 + W_2 I_1)\lambda_3^{\,2} - W_2\lambda_3^{\,4}\right]\end{cases}$$

The material parameters in the constitutive model can be fitted by uniaxial tensile experiments

(1) Get the rubber matrix material parameters C_{10} and C_{01} according rubber uniaxial stretching test data

(2) Get the cord material parameters k_1、k_2 and k_3 according uniaxially stretched in the direction of the cord fitting experimental data

The tensile test data is fitted by the least squares method to obtain the material parameters in the constitutive model as shown in the following table

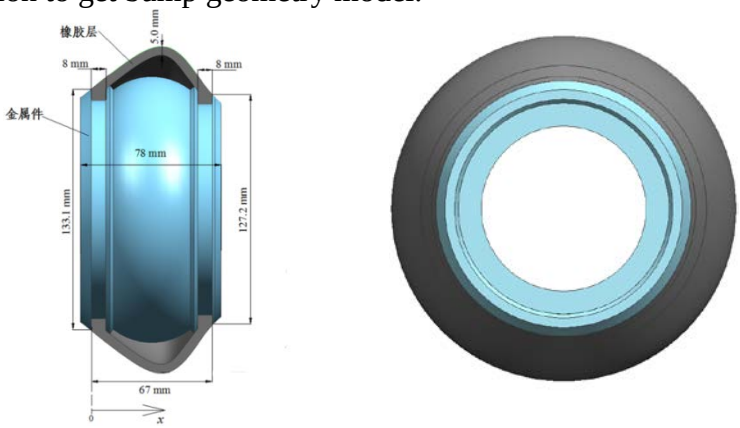
Table 1.Material parameters

C_{10}	C_{01}	k_1	k_2	k_3
-0.89	1.78	13.43	-17.88	14.45

2. Three-dimensional finite element model of bump

2.1 Three-dimensional structure of the bump

The actual geometry of bump is complicated, get the airbag outside the cavity when set up the finite element model, as show in the figure(a) . the airbag on both sides of the metal chamber is determined by applying a fixed boundary conditions . the airbag section around the central axis of rotation to get bump geometry model.

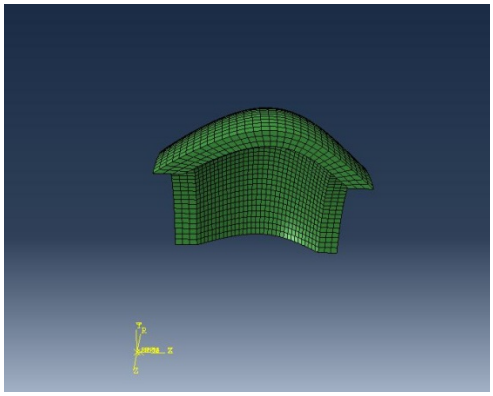


Principal view　　Left view

Figure(a)

2.2 Mesh and element type

The bump structure is divided into three-layer. And the bump are constructed by different material layers. The unit of the model is C3D8 , a total of 3840 units 6412 nodes, as show in the figure(b).



Figure(b)

2.3 Grid division and cell type

Due to the use of cord - rubber composites in the drum structure, the drum structure is divided into three-layer units when the meshes are divided, and the drums are constructed by different material layers. Three-dimensional model using 8- node entity unit, a total of 3840 units 6412 nodes.

2.4 Material model

The rubber layer use the Mooney-Rivlin material model, cord - rubber composite material layer using an anisotropic elastic constitutive material models, When defining the cord - rubber composite layer, the direction of the material can be assigned according to the actual laying angle of the cord.

2.5 Boundary conditions and load conditions

Considered the bump and the metal cavity fixed constraint , set the completely fixed constraint of two sides of bump, regard the inflation pressure as a uniform load for the inner surface of the bump in the simulation analysis.

3. Calculation result and analysis

Table 2 is the maximum deformation height of bump structure under the different loads and different cord angle, figure (c) is a maximum deformation height of bump structure under different load and the same cord angle; figure(d) is the maximum deformation height of bump structure under different cord angle and the same load.

Table 2. the maximum deformation height of bump

load/**Pa** / angle	0.1	0.3	0.5	0.7	0.9
15°	83.09	84.42	85.78	87.18	88.60
30°	83.21	84.77	86.36	88.01	89.70
45°	83.43	85.38	87.40	89.55	91.77
60°	83.72	86.19	88.82	91.78	94.72

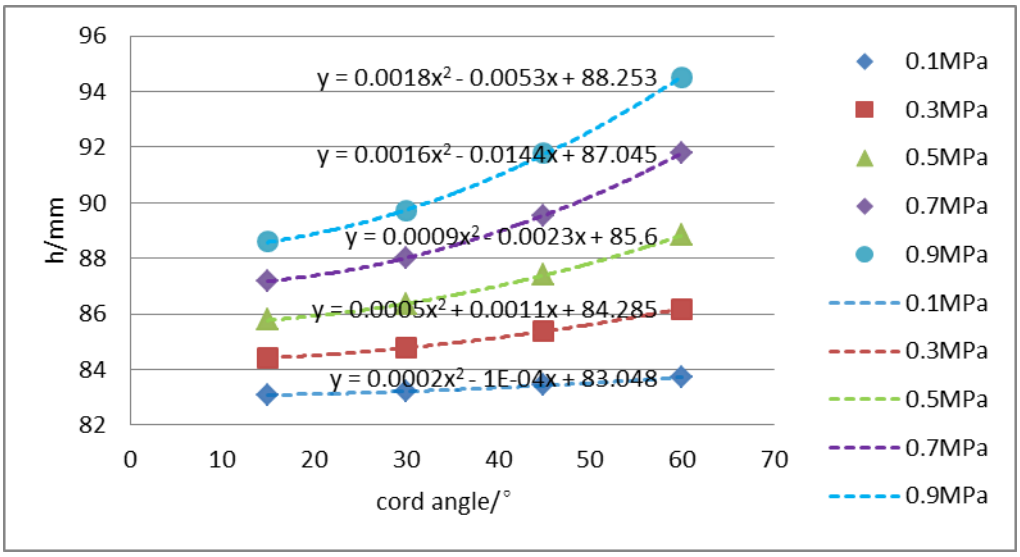


Figure (c)

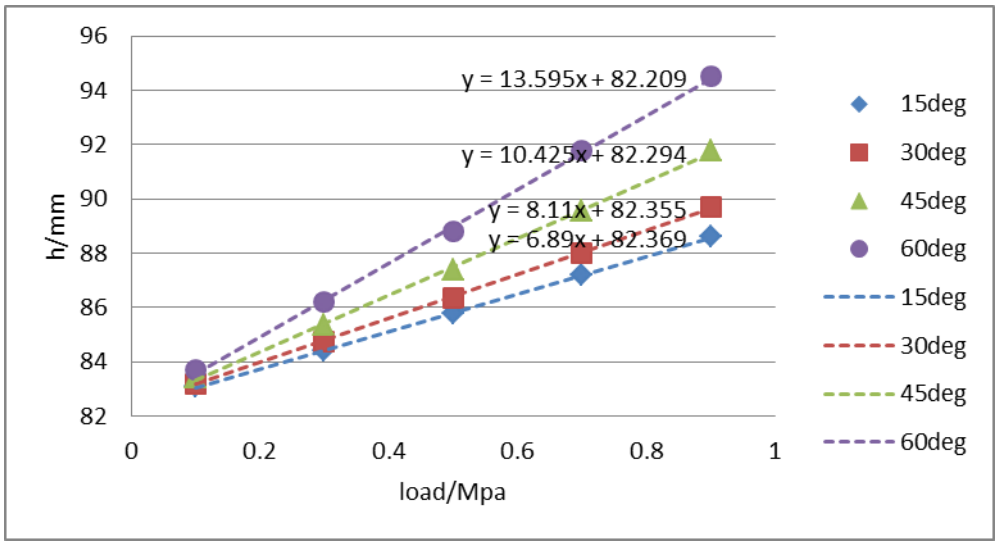


Figure (d)

It can be seen from figure(c) that the maximum deformation height of the bump increases by the increase of the uniform load in the bump inner surface at the same cord angle, and the maximum deformation height of the bump under different loads is approximately in a straight line , It can be shown that the maximum deformation height of the bump increases linearly with the increase of the uniform load when the cord angle is unchanged. As can be seen from figure(d) , with the uniform distribution load, The maximum deformation height of the bump is also increased by the increase of the cord angle, the maximum deformation height of the bump is approximately in a quadratic curve. In figure(c) and figure(d) , the curve function expression acquired by the least squares method fit.

4.Conclusions

The cord-rubber composite material is a nonlinear material. The anisotropic hyperelastic constitutive model can simulate the deformation behavior of the cord-rubber accurately, and provide the analytical basis for the design and optimization of the bump structure in the subsequent finite element analysis.

References

[1] Peng XQ,Guo ZY and Moran B. An anisotropic hyperelastic constitu-tive model with fiber-matrix shear interaction for the human annulusfibrosus. Journal of Applied Mechanics , 2006，73：815-824.
[2] Guo Guodong, Peng Xiongqi and Zhao Ning An anisotropic hyperelastic constitu-tive model with shear interaction for cord-rubber tire composities. Chinese Journal of Theoretical and Applied Mechanics , 2013, 45(3):451-455.
[3] Sun Shulei,Mao Jianliang and Peng Xiongqi. An Anisotropic Hyperelastic Constitutive Model With Fibre Bending Stiffness for Cord-Rubber Composites. Applied Mathematics and Mechanics, 2014, 35(5):471-477.
[4] Huang Xiaoshuang, Peng Xiongqi and Zhang Bichao. An anisotropic visco-hyperelastic constitutive model for cord-rubber composites. Chinese Journal of Theoretical and Applied Mechanics,2016, 48(1):140-145.
[5] Spencer A J M. Continuum Theory of the Mechanics of Fibre-Reinforced Composites. Journal of Applied Mechanics, 1984, 53(1):233.
[6] Clark SK. Theory of the elastic net applied to cord-rubber compos-ites. Rubber Chemistry and Technology, 1983, 56(2): 372-389.
[7] Ishikawa S, Tokuda A, Kotera H. Numerical simulation for fibre reinforced rubber. Journal of Computational Science and Technology, 2008, 2(4): 587-596.
[8] L. R. G. Treloar. The elasticity of a network of long-chain molecules I/II. Trans Faraday Soc, 1943, 42(4):83-94.
[9] Boyce M C, Arruda E M. Constitutive Models of Rubber Elasticity: A Review[J]. Rubber Chemistry & Technology, 2000, 73(3):504-523.
[10] Rivlin R S. Large Elastic Deformations of Isotropic Materials. V. The Problem of Flexure[J]. Proceedings of the Royal Society A, 1949, 195(1043):463-473.

Micromechanical Studies of Strain Rate Dependent Compressive Strength in Brittle Polycrystalline Materials

Hao Jiang, Zongyue Fan and *†Bo Li

Department of Mechanical and Aerospace Engineering, Case Western Reserve University, Cleveland, OH 44106, USA.

*Presenting author: bxl295@case.edu
†Corresponding author: bxl295@case.edu

Abstract

We propose a micromechanical computational framework for the high fidelity prediction of failure mechanisms in brittle polycrystalline materials. A three-dimensional direct numerical simulation of polycrystalline structures is constructed to explicitly account for the microstructural features, such as grain sizes, grain orientations, and grain boundary misorientations, by using the finite element method. In particular, grain boundaries are represented by a thin layer of elements with non-zero misorientation angles. The Eigen-fracture algorithm is employed to predict the crack propagation in the grain structure including intergranular and transgranular fractures. In the Eigen-fracture approach, an equivalent energy release rate is defined at the finite elements to evaluate the local failure state by comparing to the critical energy release rate, which varies at the grain boundaries and the interior of grains. Moreover, the constitutive model is considered as functions of the local microstructure features. As a result, the anisotropic response of brittle polycrystalline materials and the interaction between the fracture and topological defects in the microstructure under general loading conditions are explicitly modeled. Finally, the compressive dynamic response of hexagonal SiC with equiaxed grain structures is studied at different strain rates by using the proposed computational framework. The predicted compressive strength as well as the strain rate dependence of SiC agrees well with measurements in Split Hopkins Pressure Bar (SHPB) experiments.

Keywords: Brittle fracture Polycrystalline; Eigen-fracture; Anisotropic; Dynamic compression; Grain boundary

Introduction

Modern structural ceramics and ceramic composites have been considered as high-performance materials due to the high-temperature stability, high hardness, low density and superior elastic moduli and strength. These exceptional thermal and mechanical properties in ceramics lead to a wide range of applications, such as lightweight armors for personal protection and high performance turbine blades. Nevertheless, unlike metal or polymers, the failure of ceramics is usually sudden and catastrophic. This feature in failure mechanisms of brittle polycrystalline materials including monolithic ceramics and ceramic composites makes it challenging to accurately predict their resistance to extreme loading conditions, especially impact loading at high strain rates. In addition, the ultimate strength of ceramic has been observed to be strain-rate sensitive ([1]–[4]). While it has been generally recognized that the rate-dependent behavior is related to the intrinsic microstructures and flaws (e.g., crystal structure, grain orientation, grain boundaries, micro-voids, second phase particle and stacking faults), evolution of the microstructure, in particular, dynamic interaction between the crack and topological defects in the microstructure, at high strain rates has not been addressed quantitatively due to the lack of fundamental understanding and the limitations in experimental diagnostics. Sarva and Nasser

studied strain rate sensitivity of Silicon Carbide under compression[1]. They found that, beside of the compressive strength, the fragment sizes are also rate sensitive. Smaller fragment sizes were obtained at higher strain rates, which indicates more micro-cracks took place. Wang and Ramesh performed experimental studies on hot-pressed SiC under high strain rate compression in 2004 [2]. By analyzing the failure process, their results show that the subcritical micro crack propagates and coalesces before macro crack splits. Recent studies on the dynamic behavior of AlN ([3], [4]) indicates that transgranular fracture becomes more common at high strain rates. Chan and Ravichandran performed similar studies on AlN [5], which show that ceramics exhibit an increase in shear strength with increasing confinement pressure and a strain-rate-sensitive material model was developed with fitting to experimental data. To date, the experimental investigations listed in the above literature and others characterize the rate-dependent dynamic response of brittle polycrystalline materials at the macro-scale. However, the underlying micromechanisms are only perceived in a qualitative manner. There is very few, if any, quantitative analysis at the microstructure level to explain the rate sensitivity of the compressive strength in brittle materials by experiments.

Alternatively, computational models and numerical simulations have been utilized to predict the strain rate dependent failure mechanisms in brittle materials. The wing-crack array model ([1], [2], [6]) is commonly adopted to explain the mechanism of strain rate effect. The wing-crack array model utilizes the local stress intensity at the preexisting crack as the failure criterion. The rate-dependent influence is included by considering inertia effects at the crack tip. The final strength of the material is related to the initial flaw size and the flaw distribution. However, rate-dependent material coefficients are introduced in the model and need to be calibrated to experiments. On the other hand, topological defects in the microstructure, such as grain boundaries and grain orientations are not explicitly accounted for. The intergranular fractures are often modeled using a cohesive surface approach within finite element formulation ([7]–[10]).The cohesive method provides a phenomenological framework to describe the traction–displacement relation of the crack surface. The fracture characteristics and anisotropic properties of the grain boundaries are embedded in the cohesive elements. When the crack surfaces are fully separated, a contact algorithm can be employed to deal with the friction between fractured interfaces. Similar to wing crack model, an intrinsic flaw length is introduced into the model. In addition, a characteristic relaxation time is derived to explain the rate dependent character of the failure strength ([11], [12]). Nittur and others[9] investigated the dynamic fragmentation of ceramic under compression. They have found that the material remains largely intact when peak compressive strength is achieved, but shows a catastrophic increase in accumulated damage after that point. Sfantos and Aliabadi[10] also studied the intergranular fracture in brittle material under compression. They found that the internal friction of the material becomes important in cases of compressive localized pressures over cracked surfaces. As this internal friction increases, crack propagation was slowed down while crack branching appeared faster. For most polycrystalline related cohesive models, only the intergranular fracture is included ([7]–[10]). It is computationally consuming for modeling transgranular fracture since each element surface/edge require a cohesive element. The fracture path is constrained to follow element interfaces that rely on the mesh spacing and orientation. It has been stated out that the cohesive zone models suffer from mesh-dependence and the lack of strict conservation and convergence properties [13]. Most recently, the phase field model becomes popular to simulate the fracture process ([14]–[16]) In this method, the discontinuities caused by crack is approximated by a phase-field. The phase field is independent of the displacement field and can be calculated implicitly. An order parameter is introduced along the crack to accommodate the material transition from the undamaged to the damaged state. The phase field model for quasi-static brittle fracture can be derived from the variational formulation

for Griffith's type of fracture models [17]. Therefore, less user-defined parameters and fracture criteria are required. However, the material properties of the damaged material need to be estimated and the order parameter function through the fracture thickness needs to be calibrated. The polycrystalline phase field model developed by Clayton and Knap is able to capture the changes of crack paths and bulk material properties by changing the grain boundary properties. However, to predict the strain rate sensitivity in the dynamic response of brittle materials, rate-dependent phase field model needs to be further developed.

Another variational approach to fracture is deducted using Eigen-fracture scheme based on element erosion ([18],[19]). Similar to phase field model, the discontinuities are approximated using an Eigen-deformation field, which is widely used in mechanics to describe deformation modes that cost no local energy. Instead of considering the energy minimization globally by the phase field method, the Eigen-fracture approach regards fracture as a dissipative process with dissipation located at the crack front [20]. In this approach, the crack front velocity obeys a kinetic law in terms of the local driving force, and the entire trajectories of the system, including the crack paths, follow as minimizers of energy-dissipation functional [21]. In this regard, the propagation of the crack can be related to the combination and competition of various energy dissipation pathways in the materials microstructure. Different types of defects, preexisting micro cracks, grain boundaries and etc., can all be considered as potential energy sinks to minimize energy when stress flow passing by.

In this work, a micromechanical computational framework based on the Eigen-fracture scheme is presented for the high fidelity prediction of failure mechanisms in brittle polycrystalline materials. A direct numerical simulation of the polycrystalline structure explicitly accounts for the microstructural features, such as grain sizes, grain orientations, and grain boundary misorientations, by using the finite element method. Furthermore, it furnishes an effective manner to directly accommodate other types of defects, including porosity and second phase particles. An equivalent energy release rate is defined at the finite elements to evaluate the local failure state by comparing to the critical energy release rate, which varies at the grain boundaries and the interior of grains. Since the calculation of the equivalent energy release rate is carried out within a local neighborhood of an element, it has the effect of eliminating spurious mesh-dependent artifacts. The Eigen-fracture scheme is known to properly converge to Griffith fracture in the limit of vanishingly small mesh sizes. The anisotropic characters of the polycrystalline structure are obtained by using a crystal elasticity constitutive model and grain boundary properties related to the misorientation angles. In order to validate the proposed computational framework, the compressive dynamic response of hexagonal SiC with equiaxed grain structures is studied numerically at different strain rates and compared to experimental measurements.

General Framework

In this section, the general framework for analyzing the brittle polycrystalline material response at the grain scale will be presented based on the finite element method. Since this work focuses on the crack propagation in the polycrystalline microstructure (6H SiC in particular), the crystal elasticity model is integrated to describe the anisotropic constitutive relationship ([22], [23]). The study of failure mechanisms is carried out by using the Eigen-fracture model ([18], [24]), which is an energy-based variational fracture principle to the generalized Griffith's model.

Governing equations

Given initial and boundary conditions, the dynamic deformation problem can be analyzed using finite element method by seeking solutions of the governing equations of conservation of mass

and momentum. Consider a body initially occupying a reference configuration Ω_0 in a d-dimensional Euclidean space with boundary Γ. Suppose the boundary can be decomposed into the essential boundary Γ_u and natural boundary Γ_t, such that $\Gamma = \Gamma_u \cup \Gamma_t$ and $\Gamma_u \cap \Gamma_t = \emptyset$. The differential form of conservation equations can be written, respectively, in the Lagrangian configuration as

$$\rho = \frac{\rho_0}{J} \quad \text{in } \Omega_0, \tag{1}$$

$$\nabla \cdot \boldsymbol{P} + \rho_0 \boldsymbol{B} = \rho_0 \boldsymbol{A} \qquad \text{in } \Omega_0, \tag{2}$$

where ρ_0 and ρ are the density in the reference and deformed configuration, respectively, $\boldsymbol{F}$ the deformation gradient, J the Jacobian defined as $J = \det(\boldsymbol{F})$, $\boldsymbol{P}$ the first Piola-Kirchhoff stress tensor, $\boldsymbol{B}$ the body force density per unit undeformed volume and $\boldsymbol{A}$ the acceleration. The first Piola-Kirchhoff stress tensor $\boldsymbol{P}$ and the Cauchy stress tensor $\boldsymbol{\sigma}$ is related as

$$\boldsymbol{P} = J\boldsymbol{\sigma}\boldsymbol{F}^{-T}, \tag{3}$$

The essential and natural boundary conditions are given by

$$\boldsymbol{u} = \bar{\boldsymbol{u}} \qquad \text{on } \Gamma_u, \tag{4}$$

$$\boldsymbol{P} \cdot \boldsymbol{N} = \bar{\boldsymbol{t}} \quad \text{on } \Gamma_t, \tag{5}$$

where $\bar{\boldsymbol{u}}, \bar{\boldsymbol{t}}$ and $\boldsymbol{N}$ are the prescribed displacement on the essential boundary, the external traction on the natural boundary and the unit outward normal to Γ_t, respectively. The displacement is denoted by $\boldsymbol{u}$, which is related to the deformation gradient as

$$\boldsymbol{F} = \nabla \boldsymbol{u} + \mathbf{I}. \tag{6}$$

The weak form associated with the governing equations follows from the principle of virtual work as,

$$\int_{\Omega_0} [\boldsymbol{P} : \nabla(\delta \boldsymbol{u}) - \rho_0(\boldsymbol{B} - \boldsymbol{A}) \cdot \delta \boldsymbol{u}] dV - \int_{\Gamma_t} \bar{\boldsymbol{t}} \cdot \delta \boldsymbol{u} dS = 0. \tag{7}$$

In this framework, the above weak form is discretized by finite elements and the system of discretized equations is solved by the Newmark's algorithm ([25], [26]).

Constitutive model

We consider a perfectly brittle material. The constitutive behavior can be described by the simplest hyperelastic material model, for which

$$\boldsymbol{P} = \frac{\partial W(\boldsymbol{F})}{\partial \boldsymbol{F}}, \tag{8}$$

where $W(\boldsymbol{F})$ is the strain energy density per unit volume. Within a finite time increment in the Newmark's method, the response of the brittle polycrystalline material can be approximated as a small strain problem. In specific, the logarithmic strain is adopted which is defined as [???],

$$\boldsymbol{\varepsilon} = \frac{1}{2} \ln(\boldsymbol{F}^T \boldsymbol{F}). \tag{9}$$

For a high-fidelity prediction of the dynamic response of the materials under extreme loading conditions, such as high pressure and high strain rates, the strain energy density can be divided into the volumetric and isochoric parts,

$$W(\boldsymbol{\varepsilon}) = W^{vol}(J) + W^{iso}(\boldsymbol{\varepsilon}^{iso}). \tag{10}$$

Particularly, the quadratic equation of state and general hook's law are employed to model the volumetric and isochoric responses, respectively, i.e.,

$$W^{vol}(J) = \frac{1}{2} K (J - 1)^2, \tag{11}$$

$$W^{iso}\left(\boldsymbol{\varepsilon}^{iso}\right) = \frac{1}{2} \boldsymbol{\varepsilon}^{iso} : \boldsymbol{C}^{iso} : \boldsymbol{\varepsilon}^{iso}, \tag{12}$$

where K is the bulk modulus and $\boldsymbol{\varepsilon}^{iso}$ is the deviatoric part of the small strain tensor,

$$\boldsymbol{\varepsilon}^{iso} = \boldsymbol{\varepsilon} - \frac{1}{3}\mathrm{tr}(\boldsymbol{\varepsilon})\mathbf{I}\,. \tag{13}$$

Thus, the Cauchy stress tensor $\boldsymbol{\sigma}$ can be derived from Coleman's relations as:

$$\boldsymbol{\sigma} = J^{-1}\boldsymbol{P}\boldsymbol{F}^{T} = K(J-1) + \boldsymbol{C}^{iso}:\boldsymbol{\varepsilon}^{iso}. \tag{14}$$

The constitutive model in the simulations of deformation and failure in brittle materials for a finite size sample is usually homogeneous, and often isotropic. This approximation is fairly adequate as long as the structure has characteristic dimensions much larger than the microstructure dimensions of the materials, i.e. grain sizes. However, this continuum approach breaks down at the microscopic level when dealing with crack initiation and propagation in the microstructure, especially distinguished inter- and transgranular fractures. Indeed, the local stress concentration induced by the microstructure features such as crystal structure and orientations may have a strong influence on the crack growth and eventually affect the ultimate strength of the material. In this framework, the polycrystalline structure is directly simulated and discretized by finite elements. The grain structure is represented by the definition of the grain orientation and misorientation angles at each element. In specific, the grain boundaries consist of elements with non-zero misorientation angles as a transition zone. Figure 1 shows a typical finite element mesh for a polycrystalline structure in 2D.

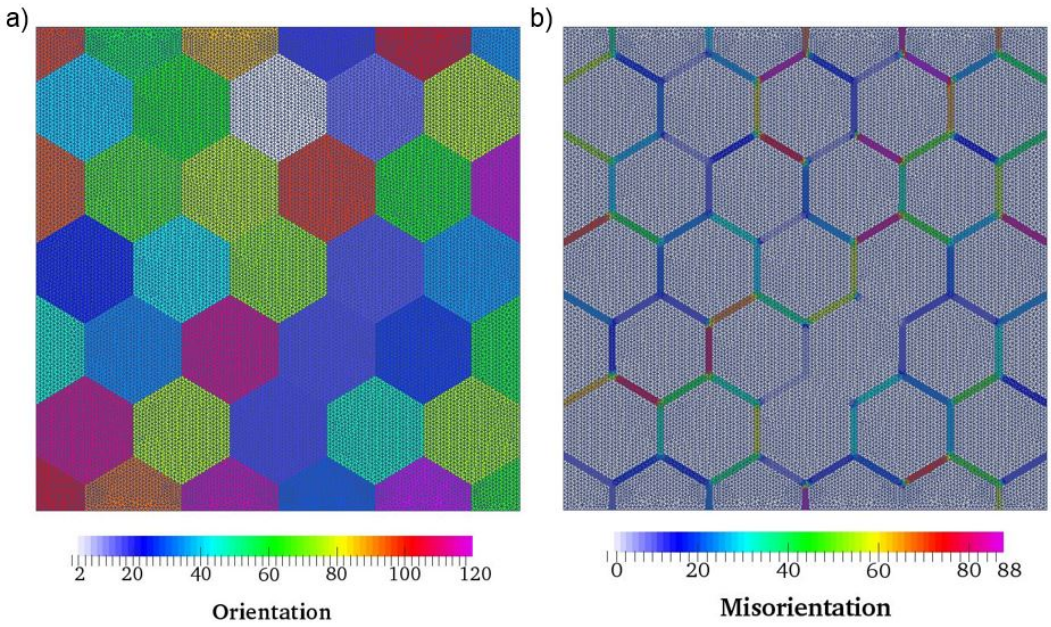


Figure 1. Typical 2D mesh for a polycrystalline structure: a) Grain orientation b) Grain boundary misorientation.

Therefore, a crystallographic constitutive model is employed in the current study to model the anisotropic behavior in the individual grains. The number of independent material parameters in the elastic moduli tensor $\boldsymbol{C}^{iso}$ is subject to the symmetry system of the crystal structure. For the hexagonal structure of SiC, the stiffness tensor can be written in terms of five independent parameters with matrix notation:

$$\boldsymbol{C}=\begin{bmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ & C_{11} & C_{13} & 0 & 0 & 0 \\ & & C_{33} & 0 & 0 & 0 \\ & & & C_{44} & 0 & 0 \\ & sym & & & C_{44} & 0 \\ & & & & & \frac{C_{11}-C_{12}}{2} \end{bmatrix}. \tag{15}$$

Table 1. The measured elastic constant in a unit of GPa of 6H-SiC at room temperature:

C_{11}	C_{33}	C_{44}	C_{12}	C_{13}
501	553	163	111	52

In order to compute the isochoric stress and strain energy, several elastic constants need to be modified:

$$C_{11}^{iso}=\frac{7C_{11}-2C_{12}-4C_{13}-C_{33}}{9},\ C_{12}^{iso}=\frac{-2C_{11}+7C_{12}-4C_{13}-C_{33}}{9},$$
$$C_{13}^{iso}=\frac{-2C_{11}-2C_{12}+5C_{13}-C_{33}}{9},\ C_{33}^{iso}=\frac{C_{11}+C_{12}+2C_{13}-4C_{33}}{9}. \tag{16}$$

Evaluation of the stresses at each element is performed in the local crystal coordinates instead of the global Cartesian coordinate system. The grain orientation is defined as the angle between the local crystal coordinates and global reference coordinate system, which is denoted by the proper Euler angles (α,β,γ) with respect to rotation axe z-x'-z". In three dimensions, the grain orientations can also be represented by a 3×3 orthonormal rotation matrix $\boldsymbol{R}$, with $\boldsymbol{R}^T\boldsymbol{R}=\mathbf{I}$, i.e.,

$$\boldsymbol{R}=\begin{bmatrix} \cos\alpha\cos\gamma-\cos\beta\sin\alpha\sin\gamma & -\cos\alpha\sin\gamma-\cos\beta\cos\gamma\sin\alpha & \sin\alpha\sin\beta \\ \cos\gamma\sin\alpha+\cos\alpha\cos\beta\sin\gamma & \cos\alpha\cos\beta\cos\gamma-\sin\alpha\sin\gamma & -\cos\alpha\sin\beta \\ \sin\beta\sin\gamma & \cos\gamma\sin\beta & \cos\beta \end{bmatrix}. \tag{17}$$

For non-zero grain orientations, a rotation of the constitutive relation is required. The coordinate transformations of fourth order stiffness tensor can be written in the following tensor notation,

$$C'_{pqrs}=R_{pi}R_{qj}R_{rk}R_{sl}C_{ijkl}, \tag{18}$$

and in the matrix notation as

$$\boldsymbol{C}'=\boldsymbol{D}\boldsymbol{C}\boldsymbol{D}^T, \tag{19}$$

where the 6×6 transformation matrix $\boldsymbol{D}$ related to the $\boldsymbol{R}$ as

$$\boldsymbol{D}=\begin{bmatrix} R_{11}^2 & R_{12}^2 & R_{13}^2 & 2R_{12}R_{13} & 2R_{13}R_{11} & 2R_{11}R_{12} \\ R_{21}^2 & R_{22}^2 & R_{23}^2 & 2R_{22}R_{23} & 2R_{23}R_{21} & 2R_{21}R_{22} \\ R_{31}^2 & R_{32}^2 & R_{33}^2 & 2R_{32}R_{33} & 2R_{33}R_{31} & 2R_{31}R_{32} \\ R_{21}R_{31} & R_{22}R_{32} & R_{23}R_{33} & R_{22}R_{33}+R_{23}R_{32} & R_{21}R_{33}+R_{23}R_{31} & R_{22}R_{31}+R_{21}R_{32} \\ R_{31}R_{11} & R_{32}R_{12} & R_{33}R_{13} & R_{13}R_{32}+R_{12}R_{33} & R_{13}R_{31}+R_{11}R_{33} & R_{11}R_{32}+R_{12}R_{31} \\ R_{11}R_{21} & R_{12}R_{22} & R_{13}R_{23} & R_{12}R_{23}+R_{13}R_{22} & R_{13}R_{21}+R_{11}R_{23} & R_{11}R_{22}+R_{12}R_{21} \end{bmatrix}. \tag{20}$$

Fracture model: Eigen-fracture approach

Finally, we present the Eigen-fracture approach in our framework for tracking the inter- and transgranular fractures in brittle polycrystalline materials. Consider an elastic body occupying a domain $\Omega \subset \mathbb{R}^n$, $n \geq 2$. The boundary of the body consists of an exterior boundary Γ, corresponding to the boundary of the uncracked body, and a collection of cracks jointly defining a crack set $|A|$. To this end, the energy-dissipation functional of a perfectly brittle material is given by

$$F(\boldsymbol{u}, A, t) = \int_{\Omega \backslash A} W\big(\boldsymbol{\varepsilon}(\boldsymbol{u})\big) dV - \int_{\Gamma_t} \bar{\boldsymbol{t}} \cdot \boldsymbol{u} dS + G_c |A|, \tag{21}$$

where Ω\A denotes the domain of the body with the crack set excluded, G_c is the critical energy release rate, and $|A|$ denotes the area of the crack set. Due to the irreversibility of the fracture, the crack set A must be increasing monotonically over time. Thus, the evolution of the crack growth and the corresponding stress field equilibrium can be sought by minimizing $F(\boldsymbol{u}, A, t)$ at all times with respect to both the displacement field $\boldsymbol{u}$ and the crack set A. Nevertheless, a mathematical description of the crack surface is difficult to be derived in engineering applications, especially for three-dimensional problems. With the help of the eigen-deformation field $\boldsymbol{\varepsilon}^*$, which describes the crack set occurring in the material as $\{\boldsymbol{\varepsilon}^* \neq 0\}$, the crack-tracking problem in perfectly brittle materials can be simplified as the minimization of the action [19]

$$F_\epsilon(\boldsymbol{u}, \boldsymbol{\varepsilon}^*, t) = \int_\Omega W(\boldsymbol{\varepsilon}(\boldsymbol{u}) - \boldsymbol{\varepsilon}^*) dV - \int_{\Gamma_t} \bar{\boldsymbol{t}} \cdot \boldsymbol{u} dS + G_c \frac{\|C_\epsilon\|}{2\epsilon}, \tag{22}$$

where ϵ is a small parameter that defines an $\epsilon-$neigborhood of the crack set, B_ϵ, and $||C_\epsilon||$ is the volume of B_ϵ, as shown in Figure 2.

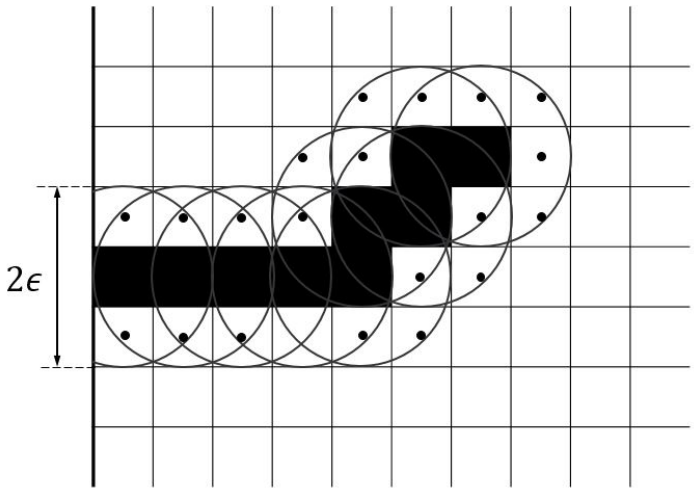


Figure 2.Visulization of the crack ϵ-neighborhood of the crack set $|A|$. Elements in the crack ϵ-neighborhood are marked with black dots.

The stationary of the simplified action in Equation (22) with respect to $\boldsymbol{\varepsilon}^*$ leads to the definition of an effective energy release rate at each finite element, Ω_e, i.e.,

$$G_e = \frac{\alpha\epsilon}{\|C_\epsilon\|} \sum_{\Omega_e \in B_\epsilon} V_e W(\varepsilon). \tag{23}$$

Thus, the energy release rate attendant to the failure of an element is estimated by a local energy averaging procedure in B_ϵ, and elements are failed when the effective energy release rate exceeds the critical energy release rate G_c as a minimizer of $F_\epsilon(\boldsymbol{u}, \boldsymbol{\varepsilon}^*, t)$, i.e., $G_e \geq G_c$. The

calculation of the effective energy release rate is carried out within a local neighborhood of the element and requires no explicit representation of the crack. For linear elasticity, the Eigen-fracture scheme is known to properly converge to the solution of Griffith fracture theory in the limit of vanishingly small mesh sizes [18]. In addition, the local neighborhood averaging of the energy has the effect of eliminating spurious mesh-dependent artifacts. While the minimization problem states the fracture is produced once the averaged local strain energy density exceeds a critical value, it is necessary to further investigate which part of the strain energy density contributes to the crack propagation under complex stress states.[27]. For instance, when an element is under hydrostatic tension, its total strain energy density may be used in the failure criterion to calculate the effective energy release rate. But when an element is under hydrostatic compression, only the isochoric strain energy is dissipated due to new fracture surface generation.

On the other hand, in the Griffith theory of brittle fracture, the critical energy release rate can be related to the surface energy density of potential fracture surfaces. It has been generally recognized that the fracture resistance of brittle materials heavily relies on the distribution of topological defects in the microstructure including grain boundaries, porosity and second phase particles, under dynamic loading conditions. In another word, the surface energy density is no longer a single value for a specific type of material but should be considered as a function of the local microstructure features. Consequently, by integrating a microstructure-informed critical energy release rate in the Eigen-fracture approach at each element, it allows us to explicitly model the interactions between the fracture and topological defects in the microstructure. In this work, only the grain boundaries are explicitly accounted for, where the critical energy release rate can be written as $G_c^{gb} = 2\gamma_{gb}$, with γ_{gb} the surface energy density of the grain boundary or grain boundary energy. In Reed and Shockley's dislocation model [28], a simple formula of grain boundary energy is derived for a 2-D cubic elastic material with small misorientation angles $\theta < 45^{\circ}$. In this formula, the grain boundary energy increases as the misorientation angle increases. However, in reality, the grain boundary properties such as energy, mobility, and diffusivity etc. are much more complicated, especially for a 3-D problem. Abrupt changes or even discontinuity may occur at critical misorientation angles. Therefore, in this work, G_c^{gb} is defined in Equation (24) as a power law function of its misorientation angle for the sake of simplicity. A more precise study will be deployed in the future using MD simulations.

$$G_c(\theta) = G_{c0}\left[1 - \left(\frac{\theta}{\theta_0}\right)^m\right], \tag{24}$$

where G_{c0} is the critical energy release rate of a single crystal, θ_0 is the reference misorientation angle, m is the exponential for grain boundary weakening. The misorientation angle θ can be calculated using the orientations of two neighboring grains,

$$\theta = \min\left|\cos^{-1}\left\{\frac{tr(\boldsymbol{R}_B\boldsymbol{R}_A^{-1})-1}{2}\right\}\right|. \tag{25}$$

where $\boldsymbol{R}_A$ and $\boldsymbol{R}_B$ are the rotation matrices for two neighboring grains A and B.

Numerical Simulations and Results

The proposed micromechanical computational framework is utilized to understand the rate dependence of the compressive strength of brittle polycrystalline materials. Particularly, the dynamic response of 6H-SiC under compressive loading in split Hopkins pressure bar experiments[29] is studied numerically in this work.

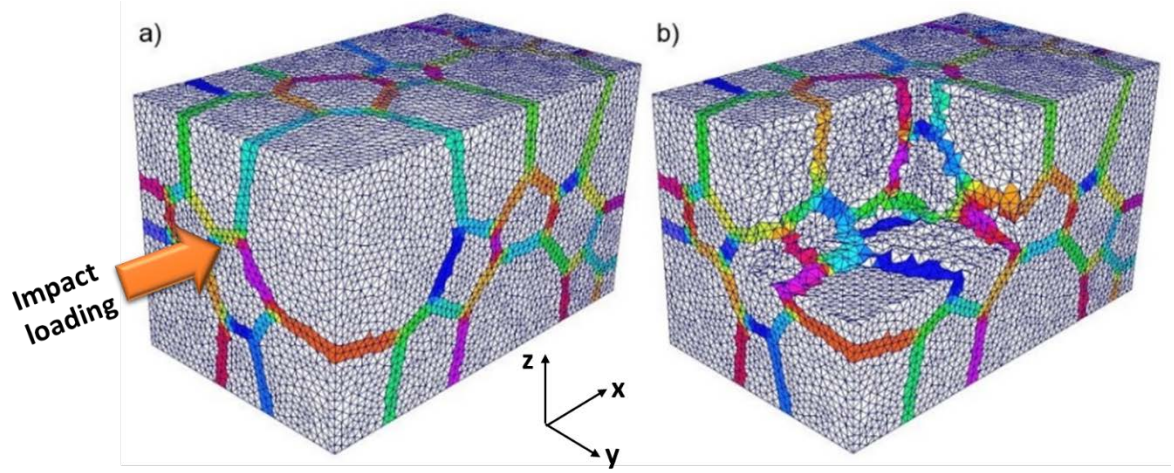


Figure 3. a) Surface meshes of the specimen and b) Interior meshes: Inner grain elements (grey) is enclosed by grain boundary elements (colored). Note that the color denotes the misorientation angles.

Figure 3 shows a typical 3D mesh of the polycrystalline brittle material in our simulations. The boundary conditions are illustrated in Figure 3 (a). The nodes on the front surface are assigned with displacements in x-direction following a linearly increasing velocity up to a constant value, while the nodes on the back surface can only move in y-z plane. The polycrystalline specimen consists of 45 equiaxed grains, which aggregate a cuboid with a dimension of 1000 μm × 600 μm × 600 μm. The average volume for an individual grain, V_{grain} is about 8×10^6 μm^3. The average grain size can be estimated as $d = \sqrt[3]{V_{grain}} = 200\,\mu m$. The computational mesh comprises 76,142 nodes and 427,264 tetrahedra. The grain boundaries are represented by two thin layers of elements adjacent to the interior surfaces. All the meshes on the interior surfaces are conformal and their nodes are shared by neighboring grain boundary elements. The average volume ratio between grain boundary and inter-grain elements is about 29.5%. The material properties and model parameters used in the simulations are listed in Table 2.

Table 2. Parameters used in simulation

Parameter	Value	Definition
K	231	Bulk modulus [GPa]
μ	192	Shear modulus [GPa]
ρ	3.21	Density [g/cm^3]
ν	0.16	Poisson's ratio
G_c	50	Critical energy release rate for single crystal [J/m^2]
θ_0	180	Relative misorientation angle in G_c^{gb} function [degree]
m	0.5	Exponential for grain boundary weakening
ϵ	$1.5h_e$	Epsilon neighbor size, where h_e is the element size

A series of snapshots in Figure 4 illustrates the fracture evolution at strain rate 3300 s^{-1}. The black translucent structure denotes the grain boundaries while the reddish elements represent the fracture zone. As shown in Figure 4(a), the cracks first initiate at grain boundaries on the front and back surfaces of the specimen. As the stress increases, weak grain boundaries inside of the specimen start to fail. While the intergranular cracks propagate into the material, the transition from intergranular fracture to transgranular fracture occurs as demonstrated in Figure

4(b). The transgranular fracture grows along the loading direction and creates longitudinal splits on the peripheral surfaces in Figure 4(c), which agrees well with the experimental observations [2]. The transgranular fractures connect the spatially dispersed grain boundary cracks into a network. As a result, the initially integrated material are divided into several pillars. In the end, the divided material comminutes one by one very rapidly until all the material fails as the deposited energy accumulates, Figure 4(d).

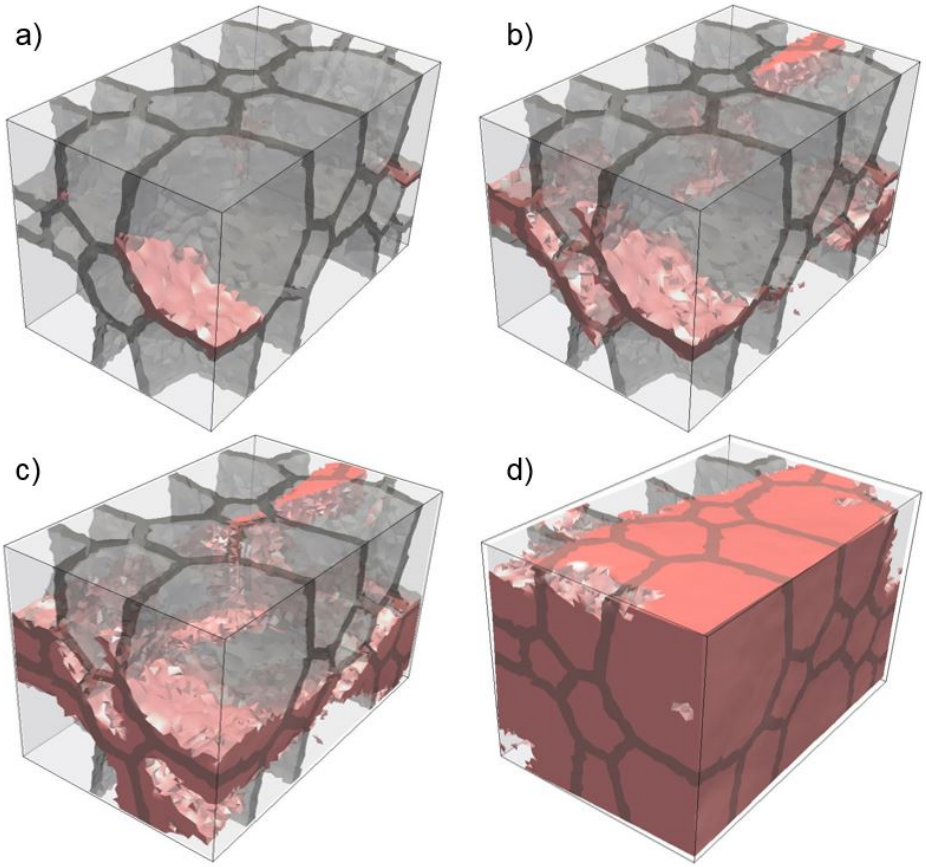


Figure 4. Evolution for cracks under uniaxial compression: a) crack surfaces initiation at grain boundaries, b) crack propagation, c) interaction between cracks and d) comminution.

The corresponding stress history of SiC under compressive loading at strain rate 3300 s^{-1} is shown in Figure 5. The stress is calculated as the homogenized one over the entire specimen alone the loading axis. As seen in the figure, the predicted compressive strength of SiC at strain rate 3300 s^{-1} by the direct numerical simulation of the polycrystalline structure is 3.45 GPa. After the stress reaches the peak, it dramatically drops and then becomes stable. The residual stress after the drop is due to the resistance to deformation in the comminuted material. Comparing Figure 4 to Figure 5, it is interesting to note that the stress continues to build up after the initial grain boundary cracks. The strengthening gradually slows down as the transgranular fracture propagates. When individual cracks coalesce with each other into a network, the stress approaches its maximum level. Afterward, the stress suffers a sudden decrease due to the material comminution.

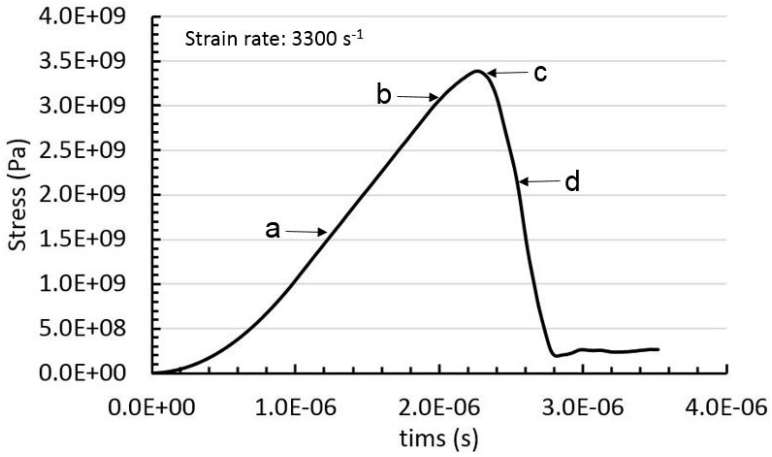


Figure 5. Stress history for strain rate at 3300 s^{-1}. The "a", "b", "c" and "d" locations pointed out in the stress curve correspond with the series of fracture images shown in Figure 4.

The same mesh and material configuration were exercised under dynamic compressive loading conditions at five different strain rates, ranging from O(10) to O(10^4) s^{-1}. The compressive strengths at various strain rates are plotted in Figure 6. The nonlinearity of strain rate dependence is illustrated in the figure. It is evident that there is a critical turning point between the strain rate 100 s^{-1} and 1000 s^{-1}. Before the turning point, the compressive strength increases slightly as the strain rate increases. As the strain rate goes beyond 1000 s^{-1}, a dramatic jump in the compressive strength is developed. The critical turning point predicted by our model matches well with the experimental measurements by Sarva and Nasser[1]. It is worth mentioning that neither the constitutive model nor the fracture model is defined to be rate dependent in our framework. The strain rate dependence of the compressive strength is a natural output of the model and results from the selection of the most effective energy dissipation pathways (i.e. intergranular or transgranular fracture) in the microstructure automatically by the algorithm. Therefore, the micromechanical computational framework allows us to develop a fundamental understanding of the rate-dependent fracture properties without introducing rate-dependent material parameters that need to be calibrated to experiments.

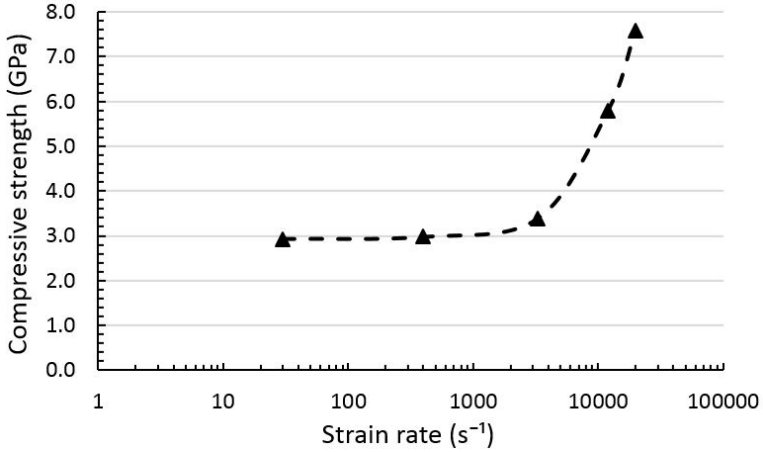


Figure 6. Strain rate dependent compressive strength of polycrystalline SiC predicted by the model.

Conclusions and Future Work

We have developed a three-dimensional micromechanical computational framework for the direct numerical simulation (DNS) of failure in brittle polycrystalline structures based on the finite element method and Eigen-fracture approach. The interaction of crack with the topological defects, such as grain boundaries, in the microstructure is explicitly modeled by considering the equivalent energy release rate as a function of the microstructural features, in specific, the grain boundary misorientation angles. The anisotropic dynamic response of polycrystalline structures is predicted by using the crystal elasticity model with local material properties related to the grain orientation and misorientations. The numerical model is validated in the example of unconfined dynamic compression tests of 6H-SiC at different strain rates. Without introducing any rate-dependent model parameters, the computational framework successfully predicts the rate sensitivity in the compressive strength of brittle polycrystalline structures. Our results show the rate dependence is intimately related to the competition and combination of intergranular and transgranular fractures in the microstructure. The selection of the optimal energy dissipation pathways in the microstructure in the case of high energy density deposition in a short time determines the ultimate strength of the brittle polycrystalline materials. It is evident that the proposed computational framework enables an automatic selection of the most effective energy dissipation pathways by integrating topological defect dependent local energy release rates in the DNS of the polycrystalline microstructure. Further investigation is necessary in order to quantify the correlation between the dynamic strength and the comparison of trans- and intergranular fractures at the microscale.

References

[1] S. Sarva and S. Nemat-nasser, “Dynamic compressive strength of silicon carbide under uniaxial compression,” vol. 317, pp. 140–144, 2001.

[2] H. Wang and K. T. Ramesh, “Dynamic strength and fragmentation of hot-pressed silicon carbide under uniaxial compression,” *Acta Mater.*, vol. 52, no. 2, pp. 355–367, 2004.

[3] G. Hu, C. Q. Chen, K. T. Ramesh, and J. W. Mccauley, “Mechanisms of Dynamic Deformation and Dynamic Failure in Aluminum Nitride,” vol. 60, no. June, pp. 3480–3490, 2012.

[4] G. Hu, K. T. Ramesh, B. Cao, and J. W. Mccauley, “The Compressive Failure of Aluminum Nitride Considered as a Model Advanced Ceramic,” vol. 59, no. June, pp. 1076–1093, 2012.

[5] W. Chen and G. Ravichandran, “Failure mode transition in ceramics under dynamic multiaxial compression,” *Int. J. Fract.*, vol. 101, no. 1–2, pp. 141–159, 2000.

[6] S. Nemat-Nasser and H. Deng, “Strain-rate effect on brittle failure in compression,” *Acta Metall. Mater.*, vol. 42, no. 3, pp. 1013–1024, 1994.

[7] H. D. Espinosa and P. D. Zavattieri, *A grain level model for the study of failure initiation and evolution in polycrystalline brittle materials. Part I: Theory and numerical implementation*, vol. 35, no. 3–6. 2003.

[8] S. Maiti, K. Rangaswamy, and P. H. Geubelle, “Mesoscale analysis of dynamic fragmentation of ceramics under tension,” *Acta Mater.*, vol. 53, no. 3, pp. 823–834, 2005.

[9] P. G. Nittur, S. Maiti, and P. H. Geubelle, “Grain-level analysis of dynamic fragmentation of ceramics under multi-axial compression,” *J. Mech. Phys. Solids*, vol. 56, no. 3, pp. 993–1017, 2008.

[10] G. K. Sfantos and M. H. Aliabadi, “Aboundary cohesive grain element formulation formodelling intergranularmicrofracture in polycrystalline brittle,” no. July 2006, pp. 1590–1626, 2012.

[11] G. Ruiz, M. Ortiz, and A. Pandolÿ, “Three-dimensional ÿnite-element simulation of the dynamic Brazilian tests on concrete cylinders,” no. January 1999, pp. 963–994, 2000.

[12] G. T. Camacho and M. Ortiz, “Computational modelling of impact damage in brittle materials,” *Int. J. Solids Struct.*, vol. 33, no. 20–22, pp. 2899–2938, 1996.

[13] M. Negri, “Numerische Mathematik A finite element approximation of the Griffith ’ s model,” pp. 653–687, 2003.

[14] C. M. Landis, “A phase-field description of dynamic brittle fracture by,” no. May, 2011.

[15] J. D. C. J. Knap, “A geometrically nonlinear phase field theory of brittle fracture,” pp. 139–148, 2014.

[16] J. D. Clayton and J. Knap, “Phase field modeling of directional fracture in anisotropic polycrystals,” *Comput. Mater. Sci.*, vol. 98, no. February, pp. 158–169, 2015.

[17] B. Bourdin, G. A. Francfort, and J. Marigo, *The Variational Approach to Fracture*. 2008.

[18] B. Schmidt, F. Fraternali, and M. Ortiz, “Eigenfracture: an eigendeformation approach to variational

fracture ∗," vol. 7, no. 3, pp. 1237–1266, 2009.
[19] A. Pandolfi and M. Ortiz, "An eigenerosion approach to brittle fracture," *Int. J. Numer. Methods Eng.*, vol. 92, no. 8, pp. 694–714, 2012.
[20] C. J. Larsen, M. Ortiz, and C. L. Richardson, "Fracture Paths from Front Kinetics : Relaxation and Rate Independence," vol. 193, pp. 539–583, 2009.
[21] A. Mielke and M. Ortiz, "A CLASS OF MINIMUM PRINCIPLES FOR CHARACTERIZING THE TRAJECTORIES AND THE RELAXATION OF DISSIPATIVE SYSTEMS," vol. 14, pp. 494–516, 2008.
[22] J. M. J. Den Toonder, J. a W. Van Dommelen, and F. P. T. Baaijens, "The relation between single crystal elasticity and the effective elastic behaviour of polycrystalline materials: theory, measurement and computation," *Model. Simul. Mater. Sci. Eng.*, vol. 7, no. 6, pp. 909–928, 2000.
[23] A. Fallahi, A. Ataee, G. Tial, and C. Stresses, "Effects of Crystal Orientation on Stress Distribution Near the Triple Junction in a Tricrystal γ -TiAl," *Mater. Sci. Technol.*, pp. 1–13, 2008.
[24] B. Li, A. Pandolfi, and M. Ortiz, "Mechanics of Materials Material-point erosion simulation of dynamic fragmentation of metals," 2014.
[25] T. Belytschko, "An overview of semidiscretization and time integration procedures," in *Computational Methods for Transient Analysis*, North-Holland, Amsterdam, 1983, pp. 1–65.
[26] T. J. R. Hughes, "Analysis of transient algorithms with particular reference to stability behavior," in *Computational Methods for Transient Analysis*, North-Holland, Amsterdam, 1983, pp. 67–155.
[27] Q. M. Li, "Strain energy density failure criterion," vol. 38, pp. 6997–7013, 2001.
[28] W. T. R. W.Shockley, "Dislocation Models of Crystal Grain Boundary," *Phys. Rev.*, vol. 78, no. 3, pp. 275–289, 1950.
[29] K. T. Ramesh, "High Strain R 33.1," *Handb. Exp. Solid Mech.*, p. 874, 2008.

Static mechanical properties test of 3 dimensional and 4 direction braid angle composite materials

†XuYibing[1], *LuLuLiu[1] and ZhaoZhenHua[1]

[1] Nanjing University of Aeronautics and Astronautics, Nanjing 210016, P.R. China,

*Presenting author: 18168028717@163.com

†Corresponding author: 2275892215@qq.com

Abstract

As the advantage of composite materials of high strength, low weight ratio, widely used in the field of aerospace. The static mechanical experiments of 3D braided composites with different braid angles and different thicknesses are carried out to study the effects of different braid angles and thicknesses on the microstructure, mechanical properties and failure modes of the materials. Due to the anisotropy of 3D braided composites and their in-plane properties were tested, the longitudinal tensile test and the transverse tensile test were carried out respectively ,from the experiment ,the static elastic modulus, Poisson's ratio, ultimate strength are obtained.

Keywords: composite materials, static mechanical

Introduction

Laminated composite material due to the presence of pure base region easily brought interlayer delamination[1], cracking and damage propagation speed sensitive , in the thickness direction of weakness low stiffness and strength mechanical properties ,three-dimensionally braided composite as it contains no delamination the overall three-dimensional network structure to solve the problem in line with laminate materials[2].Three-dimensional braided composite material due to its series of advantages , general concern engineering sector , an important structural materials in aviation, aerospace , and has been in terms of bio-medical has also been applied[3].At present,there are studies on the elastic properties of woven composites[4];Avva et al[5] lists the test process, test data and compression test results of three-dimensional braided composites;Li Suhong[6],Liu Qian[7] Considering the influence of braiding angle on the tensile properties ,Yang Zhaokun[8] found that under the premise of the knitting angle , the tensile modulus increased with the increase of fiber volume content

In this paper,the static mechanical experiments of 3D braided composites with different braid angles and different thicknesses were carried out to study the effects of different braid angles and thickness on the microstructure, mechanical properties and failure modes of the composites. As a result of the anisotropy of 3D braided composites, the in-plane properties were tested and subjected to longitudinal tensile tests and transverse tensile tests, respectively. In order to obtain data effectively, the T -strain gauge and the extensometer are used for the test at the same time.

Test equipment introduce

MTS793-10T tension and fatigue test system, the maximum tension 100KN , with dynamic / static tensile (compression, bending), fracture mechanics, low fatigue, high fatigue and high temperature and other conditions of the test capacity. Static tension and compression testing at MTS793-10T performed on the material type hydraulic servo fatigue testing machine, the

major components of the test system includes a load frame, an oil pump, the cooling system, as in FIG1.

（a）Load framework （b）hydraulic system （c）cooling system

Fig1. MTS-793# Hydraulic servo tension and compression and composite fatigue test system

In the fatigue test, the test system records the cyclic load and displacement response values by load and displacement sensors, and in order to record the response values of the tensile strain strain,634.12F-24Positive Strain of Test Section of Static and Dynamic Stretch Tensile at Normal Temperature. The data collected for the measurement of T-type bi-directional BE120-3BB resistive strain gauges are used to calculate the Poisson's ratio.

Three-dimensional four-way braided composite materials is commissioned by the Tianjin University of Technology Institute of composite materials processing, using 12K T700 carbon fiber braid molding, TDE86 epoxy resin for curing. The specimen size and die size, is first processed to FIG.2, illustrated 380mm*180mm*2mm or 4mm plate, and then cutting the test piece according to the material specifications. Tensile test pieces are cut by the standard of GBT 1447-2005[9], see FIG.3.

Test specimen introduce

Fig2 Three-dimensional four-way composite board

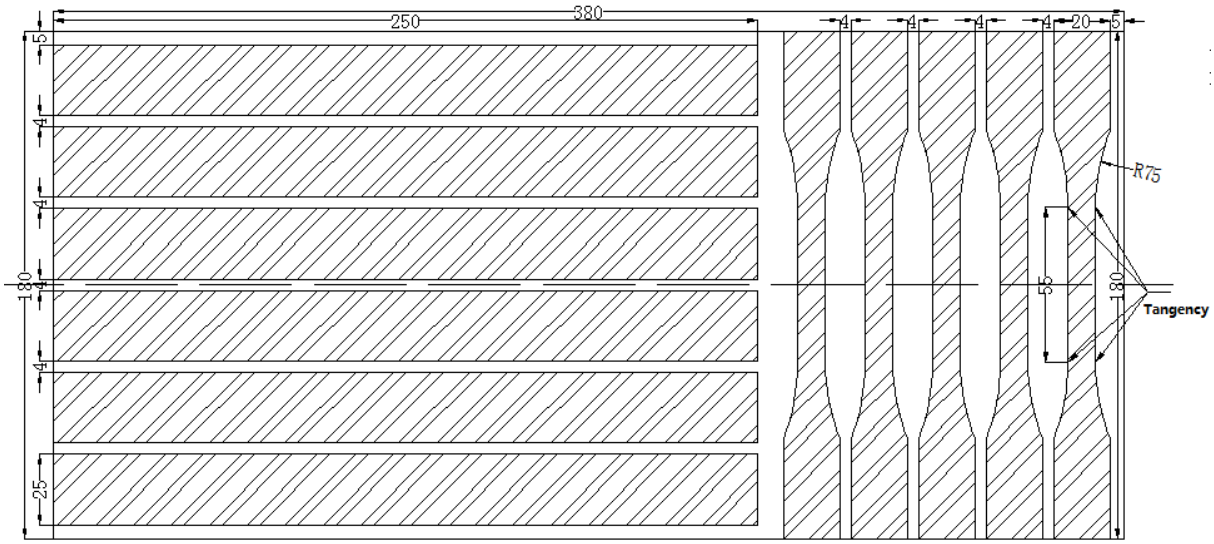
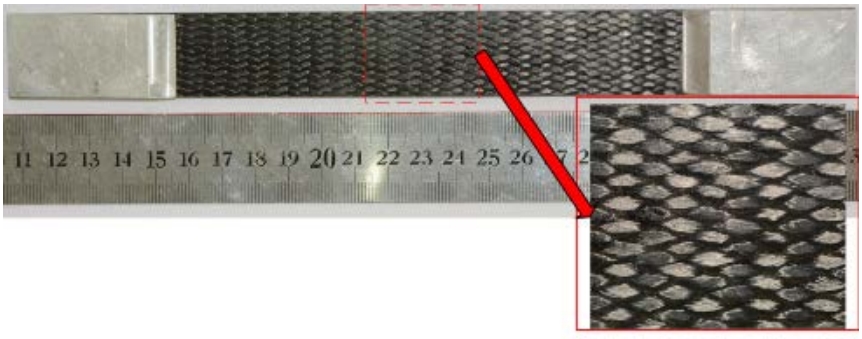
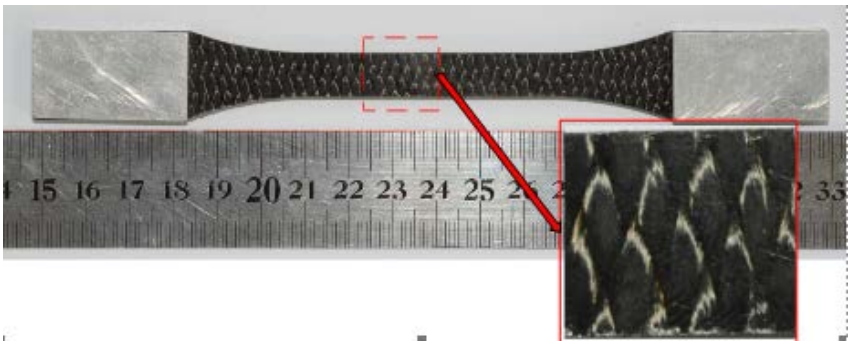
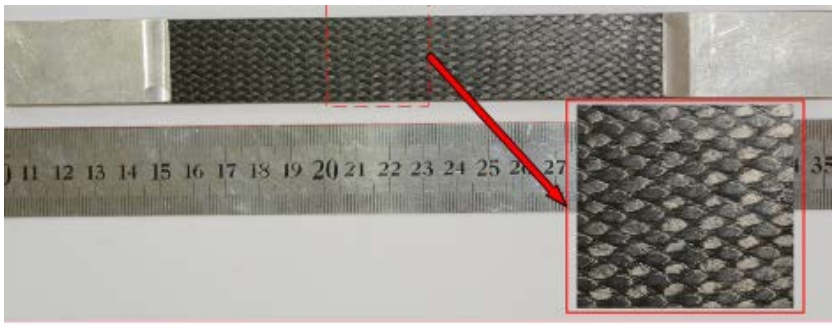


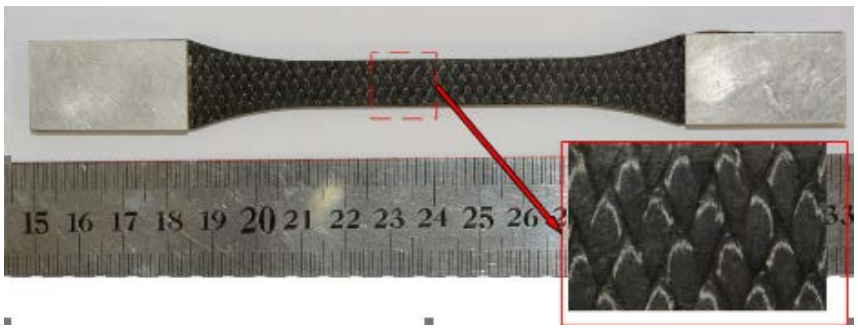
Fig3 Tensile test piece cutting scheme



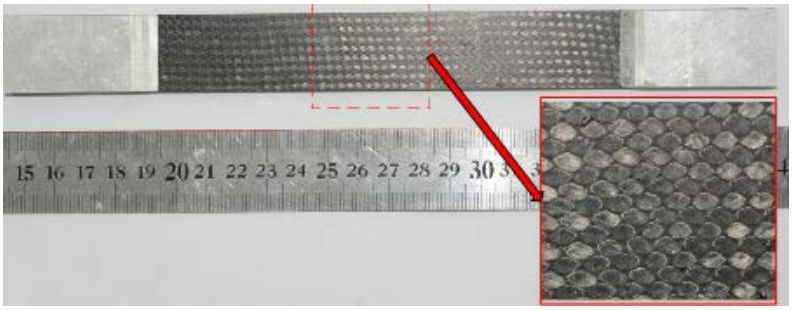


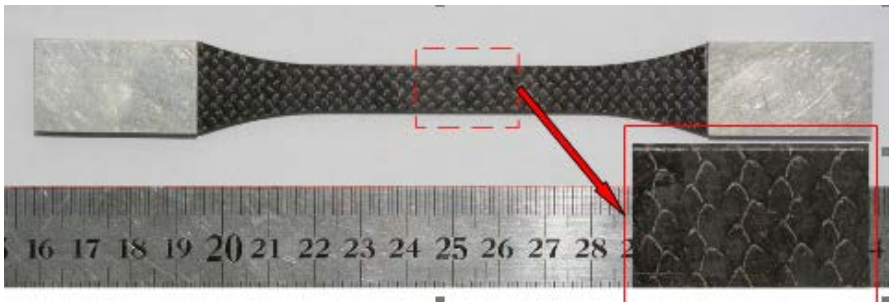
(a) longitudinal tensile 20° braiding angle test specimen (b) transverse tensile 20° braiding angle test specimen





(c) longitudinal tensile 30° braiding angle test specim (d) transverse tensile 30° braiding angle test specimen





(e) longitudinal tensile 45° braiding angle test specimen (f) transverse tensile 45° braiding angle test specimen

Fig4 tensile test specimen

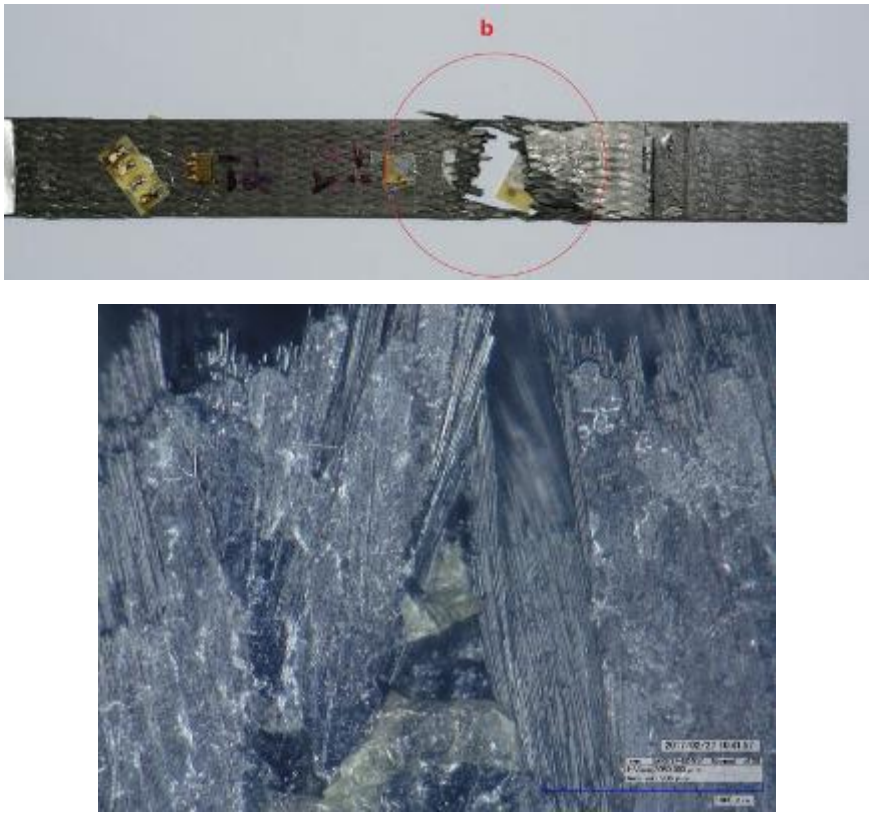
As a result of the quasi-static test, the strain rate effect can be neglected, initially set in the MTS793-10T equipment, a certain degree of tension, combined with the length of the test piece parallel to the length of the strain rate in the order of $10\text{-}4S^{-1}$, in line with static test Requirements, the test pieces of the clamping and loading force needs to be strictly aligned, the installation process to adjust the clamping position, to maintain the vertical tensile test pieces, to avoid the tilt caused by the shear stress. As shown in Fig.5(b), the extensometer is fixedly fixed in the tensile section of the test piece and the strain data is measured.

(a) (b)

Fig 5 Speicemen Clamped way

Test results

In order to compare the influence of different braiding angles on the static mechanical properties, a quasi - static tensile test was carried out on the test pieces with a strain rate of 10-4 S -1 and 20 °, 30 ° and 45 ° thickness of 4 mm . A static tensile test was carried out on the longitudinal cut test piece until the test piece was broken and the test procedure was observed and the data recorded. The stress-strain curves are obtained from the strain collected by the extensometer and the force collected by the MTS , as shown in the following figure.



（a） （b）

Fig 6 20° longitudinal tensile speicemen damage morphology

（a） （b）

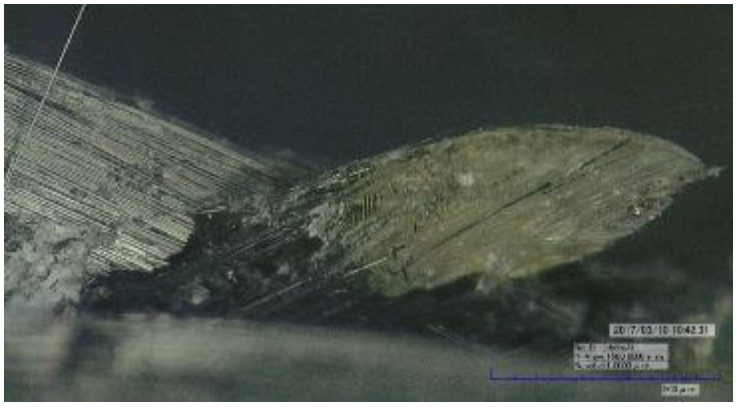
Fig 7 30° longitudinal tensile speicemen damage morphology

（a） （b）

Fig 8 45° longitudinal tensile speicemen damage morphology

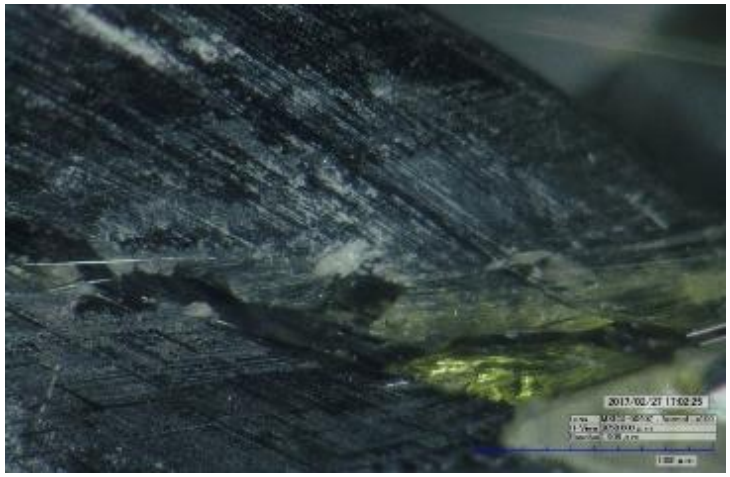
During the tensile test of the longitudinal test piece, due to the different modulus of elasticity of the matrix and the fiber bundle, the load and displacement of the matrix and the fiber bundle are different, resulting in the initial separation of the matrix from the fiber bundle, The crack. When the load reaches a certain value, the specimen began to issue a slight sound, then the matrix and fiber bundles gradually separated. As the load increases, the force bearing from the matrix fiber complex to the fiber bundle, fiber bundle to bear the main longitudinal tension, until a certain value, fiber bundles break[10], then the material was completely pulled off. From the analysis of the tensile section of the figure, you can see the fracture is basically flat fracture, can explain the fiber bundle at the last moment is almost simultaneously pulled off.

In order to compare the difference between the transverse cutting and the longitudinal cutting test piece and the mechanical properties, the test results were carried out on the transverse cutting test pieces. The experimental results are shown below.



（a） （b）

Fig 9 20° transverse tensile speicemen damage morphology



（a） （b）

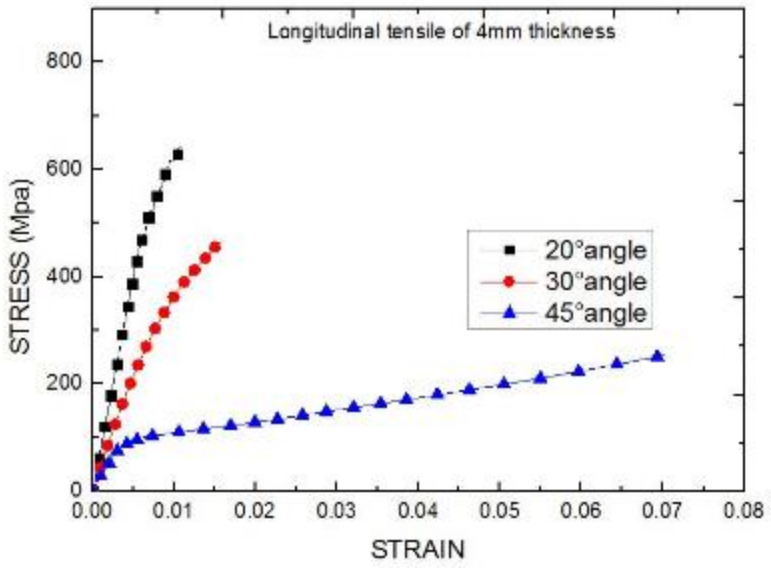
Fig 10 30°transverse tensile speicemen damage morphology

（a） （b）

Fig 11 45° transverse tensile speicemen damage morphology

Horizontal cutting test pieces of the fiber can not afford the basic tensile force, the specimen surface a small amount of cracks, which is the fiber bundle and resin interface caused by the phenomenon of localized debonding. Mainly for the fiber slip, interface debonding, matrix yield cracking. Showing a clear brittle failure characteristics. The transverse tensile strength at this time is determined by the properties of the fiber bundle and the matrix. With the increase of the braid angle, the longitudinal force of the fiber is increased and the tensile strength is increased.

Using the origin to export 4mm tensile test pieces of the horizontal, longitudinal tensile test data, as shown below:



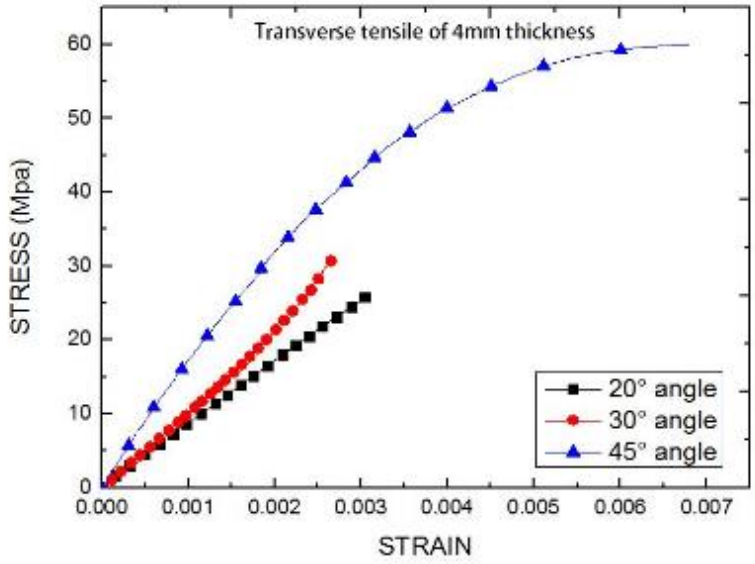


Fig 12 stress-strain curve of longitudinal tensile in 4mm Fig 13 stress-strain curve of transverse tensile in4mm

The longitudinal tensile and transverse tensile tests of the three - dimensional four -direction braided composites with thickness of 4mm were used to compare the static mechanical properties. The following conclusions can be drawn:

1) When the thickness and fiber volume fraction are constant, the elastic modulus and the ultimate strength of the longitudinal tensile test piece decrease with the increase of the knitting angle. This is because the larger the knitting angle is, The greater the angle between the direction of force, the smaller the component force that the test piece can bear in the longitudinal direction, so the longitudinal tensile mechanical properties will decrease;

2) When the thickness and fiber volume fraction are constant, the elastic modulus and the ultimate strength of the transverse tensile test piece increase with the increase of the knitting angle. This is because the larger the knitting angle is, the longitudinal yarn and the force The greater the angle of the direction, the test piece in the horizontal capacity to bear the greater the force, so the lateral tensile mechanical properties will increase.

3) The longitudinal mechanical properties of the test piece are better than those in the transverse mechanical properties, and the elastic modulus and ultimate strength of the former are higher than those of the latter (except for the 45 ° compression test, the longitudinal and lateral results are similar) The mechanical properties are mainly determined by the strength of the fiber bundle. The transverse mechanical properties are mainly determined by the strength of the matrix, and the strength of the fiber bundle is obviously higher than that of the matrix.

4) From the transverse stretching and longitudinal stretching can be seen, 45 ° braided composite material, because the fiber can withstand the absorption of the load, solid show better toughness, and 20 °, 30 ° prepared fiber, fiber by Force is not broken in the form of fiber bundle fracture and fiber and matrix fracture, almost no yield stage, showing brittleness.

In order to understand the effect of thickness on mechanical properties, a quasi-static tensile test with a thickness of 2 mm was carried out on the base of 4 mm test piece . The test results are shown below.

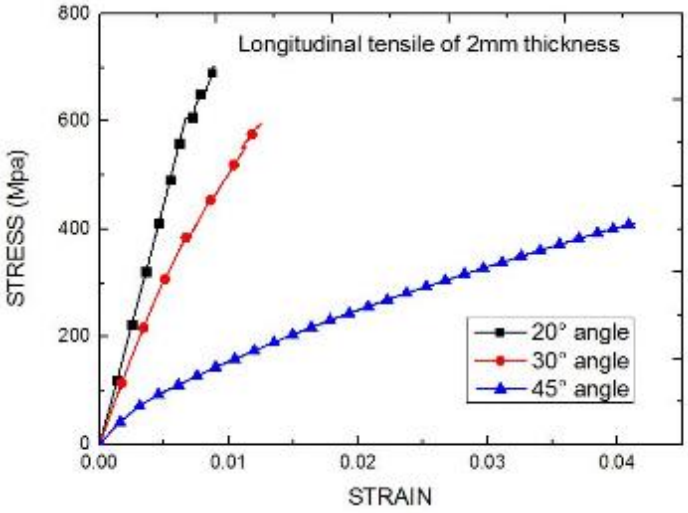


Fig 14 stress-strain curve of longitudinal tensile in 2mm

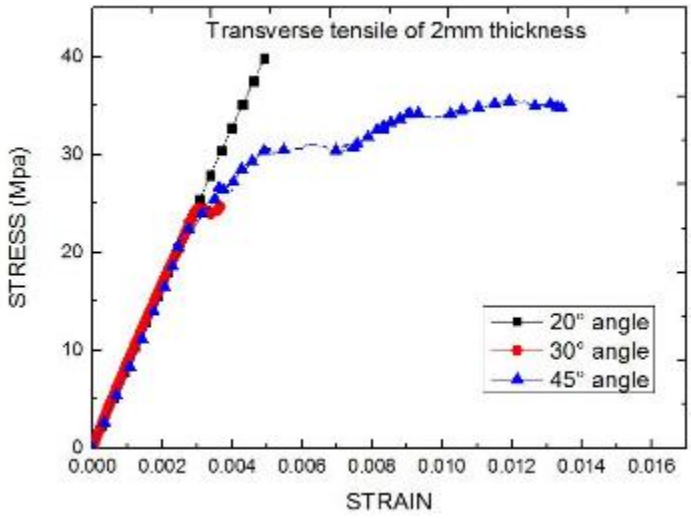


Fig 15 stress-strain curve of transverse tensile in 2mm

According to ASTM_D3039M—2014, the calculation of elastic modulus is based on strain 0.001, and strain 0.003 is the straight slope of the endpoint.

T -strain gauges were attached to the test pieces , and the longitudinal and transverse strain values were measured, and then connected to the strain gauge to obtain the corresponding test pieces Poisson's ratio.

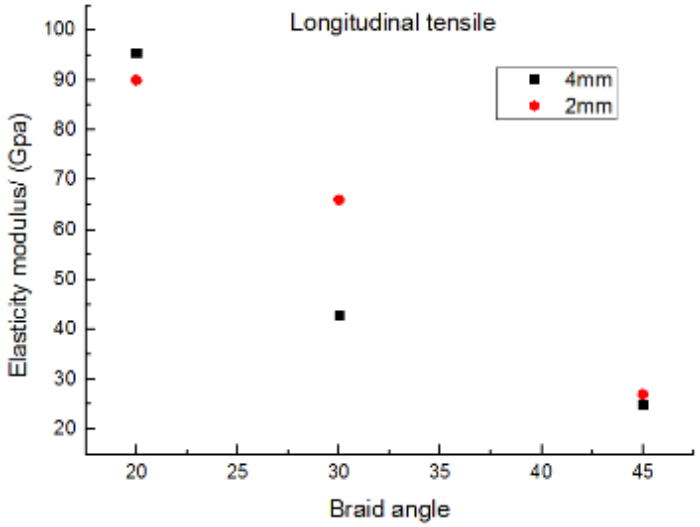
Table 1 Longitudinal tension mechanical parameter

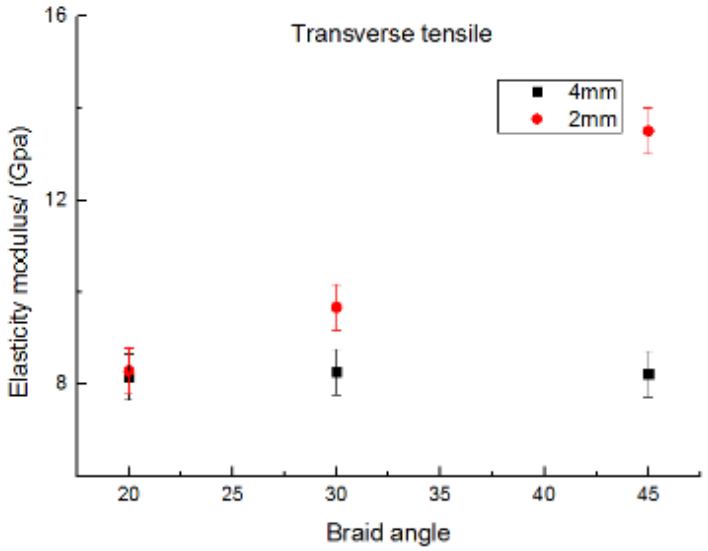
thickness	braid angle	elasticity modulus/Gpa	ultimate strength/Mpa	Poisson's ratio
2mm	20	90.66	741.70	0.71
	30	66.27	592.36	0.89
	45	27.35	409.93	1.00

	20	95.6	821.5	0.88
4mm	30	43.25	469.3	1.11
	45	25.4	258.8	0.55

Table 2 Transverse tensile mechanical parameter

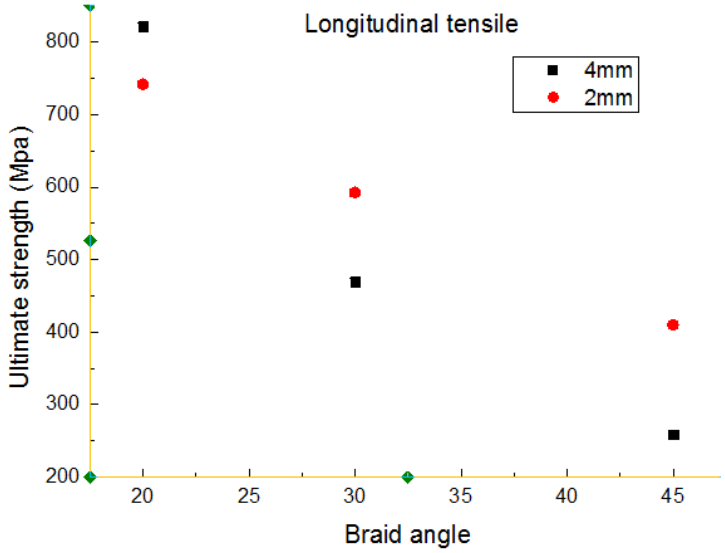
thickness	braid angle	elasticity modulus/Gpa	ultimate strength/Mpa	Poisson's ratio
	20	8.16	38.71	0.07
2mm	30	8.27	40.24	0.12
	45	8.23	47.60	0.27
	20	8.29	25.60	0.069
4mm	30	9.67	32.85	0.16
	45	13.50	56.62	0.43

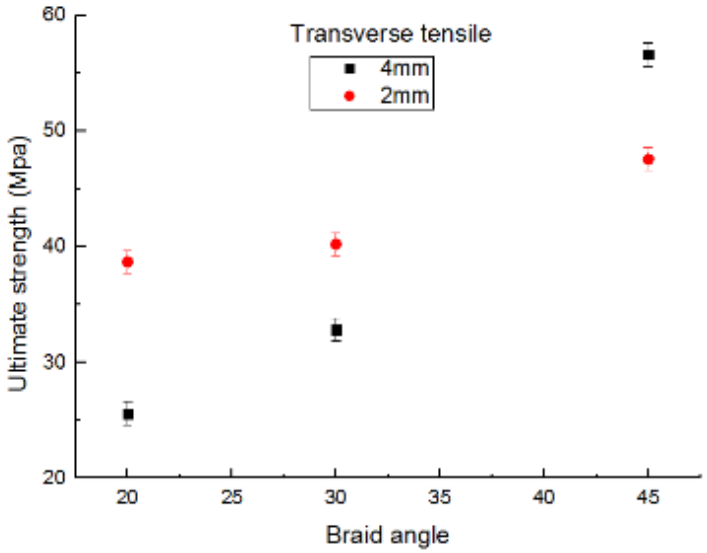




(a) (b)

Fig 16 Elastic modulus along with braiding angle





(a) (b)

Fig 17 Ultimate strength along with braiding angle

Then,conclusions can be drawn:

1)4mm and 2mm different thickness of the quasi - static tensile modulus is basically flat, indicating that the thickness of the static tensile modulus of the mechanical properties of little impact;

2)The transverse tensile strength of 2mm is not obvious at the angle of 38~48Mpa . When the thickness increases to 4mm , the ultimate strength increases with the angle.

3)Comparison of the longitudinal and transverse tensile strength of 2mm and 4mm with the bending angle of the trend. Compared with the 4mm test piece, the tensile strength of 2mm is less sensitive to the direction of braided angle, and 4mm is more sensitive.

References

[1] Wu Delong , Shen Huairong . Mechanical properties of textile composite materials [M]. Changsha : National University of Defense Technology Press , 1998

[2]Li Jialu,Xiao Lihua .Dimensional braided structure to multiple impact on the performance of the composite [J]. Composite Materials Journal,1996, (3): 71-75.

[3]Tao Xiaoming,Xian Xingjuan ,crested Hoon.Textile Structural Composites [M].Academic Press,2001, 1~9.

[4]Sun HY, Qiao X. Prediction of mechanical properties of three-dimensionally braided composites [J]. Composite Science and Technology,1997,57:623-630.

[5]Avva VS,Sadler RL,Shivakumar KN,et al.Mechanical properties of 3-D braided composite panels fabricated form using RTM techniques [D].1996,1838-1847

[6]Li Suhong etc, Affect the three-dimensional braided structure parameters on tensile properties of the composites.FRP/composites, 2014 (07):on 69-73

[7]Liu Qian,Li Jialu , Li Xueming .Affect the three-dimensional braiding process parameters on tensile properties of the composite [J].Aerospace Materials & Technology,2000,1:55-58.

[8]Institute of Materials Science and Engineering, Central South University, Changsha 410083, China Measurement and Analysis of Mechanical Properties of Braided Structure Composites [J]. FRP/composites,2002,3:11-14.

[9] nationwide fiber-reinforced composite material Standardization Technical Committee Secretariat . Fiber reinforced composites (FRP) Standard Series [M]. Beijing : China Standard Press , 1986:230-235.

[10]Meng Songhe , Tian Xiaoxiao . 3D woven carbon / epoxy composite mechanical properties testing and failure mechanism [J] Composite Materials, 2012

Research on US-MRI fusion scheme based on non-negative matrix factorization and dual modality contrast agent

† M. Hou [1], * M. Y. Yang [1], X. T. Hu [1], N. N. Wang[1], W. H. Yu [2]

[1] School of Computer Science and Technology, Jiangsu Normal University, Xuzhou 221116, China.
[2] Jiangsu Key Laboratory of Education Big Data Science and Engineering, Xuzhou 221116,China.

†Corresponding author: 469954990@qq.com

Abstract

In order to realize the image fusion and improve the quality of the integration, by the aid of DMCA(dual modality contrast agent), a novel fusion method between US(ultrasound imaging) and MRI(magnetic resonance imaging) is put forward in the article. Due to US's strong speckle noise, it is an enormous challenge to fuse US with any other modality images. Under the circumstances, DMCA is used in both US and MRI to strengthen the most important information region of interest. Then, because of Rayleigh distribution of ultrasound imaging, an self-adaptive weighted non-negative matrix factorization(SWNMF) scheme is utilized to complete the fusion process. In view of the above-mentioned method, the multiple group comparison tests indicate that US-MRI fusion may be a remarkable method for gaining high-quality fusion image.

Keywords: US, MRI, fusion, DMCA, SWNMF

1. Introduction

US has been widely used in clinical diagnosis for its advantages of real-time and low-cost. Nowdays with the use of contrast agents, the quality of ultrasound imaging has been greatly improved [1,2]. However, the quality of US is mediocre compared its contrast with that of MRI or computed tomography (CT), because US is reflected very strongly when passing from gas to tissue, and vice versa.

MRI is another used commonly imaging modality with ideal soft-tissue contrast and high spatial resolution; besides particularly MRI can provide functional information needed by the clinic. Magnetic iron oxide nanoparticles owning superparamagnetic property can be used as a effective contrast agent for MRI to enhance its contrast [1]. One critical defect of MRI can't offer real-time motion-related images.

In a word, no single imaging modality holds all the merits satisfying all kinds of clinical needs, different imaging modalities have their respective advantages and disadvantages in clinical practice. In many circumstances, US and MRI are complimentary. Accordingly, it is ideal to fuse US with MRI. Owning to US's strong noise, it is a great challenge to fuse US with any other imaging modalities.

We have done prophase work on MRI-US registration and fusion based on DMCA, and have yielded some definite results [3~6]. The DMCA mentioned in the article is the dual-modality contrast agent, holding both US and MRI contrast function property. Microbubbles can be used as a effective contrast agent for US. Superparamagnetic iron oxide nanoparticles (SPIO) can be used as a powerful contrast agent for MRI. The combination of microbubbles and SPIO, DMCA, can be used as the contrast agent for both US and MRI because the DMCA can eliminate the defects of magnetic nanoparticles or microbubbles, respectively. Nowadays, effort to improve algorithms for medical image processing has seen very little progress [7~10].

The key contribution of this article is the introduction of the mentioned above DMCA to US-MRI fusion. Using DMCA prepared by Yang et al [1], this paper carries on the US-MRI fusion based on SWNMF, and arrivals at a conclusion that with the use of DMCA and SWNMF, fusion method performs well. The remainder of the paper is organized as follows: US-MRI fusion scheme (SWNMF) is described in Section 2. Section 3 provides several groups of comparison experiments, and analyzes the experimental results, while Section 4 concludes our paper.

2. Fusion method (SWNMF)

The noise distribution of ultrasonic imaging approximately conforms to Rayleigh distribution, because of which, SWNMF based on Rayleigh distribution is introduced as follows. Rayleigh distribution is defined as:

$$p_x(x) = \frac{2x}{a}\exp(-\frac{x^2}{a}) \ , \quad \text{s.t} \quad x \geq 0 \ . \tag{1}$$

For (1) a is a constant satisfied $a > 0$. To simplify the derivation process, a is set equal to 1 in the following section. If we now set $\frac{dp_x(x)}{dx} = 0$, we obtain $x = \frac{1}{\sqrt{2}}$. When x is set equal to $\frac{1}{\sqrt{2}}$, $p_x(x)$ reaches maximum value. According to Rayleigh distribution of the speckle noise of US, a new NMF is proposed as follows:

$$p(A_{ij} \mid U,V) = (A_{ij} - (UV)_{ij} + \frac{1}{\sqrt{2}})\exp(-(A_{ij} - (UV)_{ij} + \frac{1}{\sqrt{2}})^2) \tag{2}$$

For (2) A represents a given non-negative $n \times m$ matrix. This matrix is then approximately factorized into an $n \times r$ matrix U and an $r \times m$ matrix V.
Here we set

$$p(A \mid U,V) = \prod_{ij} p(A_{ij} \mid U,V) \tag{3}$$

then the maximum likelihood solution is the minimization of the loss function as follows

$$L(U,V) = \sum_{i,j} \ \{ \ (A_{ij} - (UV)_{ij} + \frac{1}{\sqrt{2}})^2 - \log(A_{ij} - (UV)_{ij} + \frac{1}{\sqrt{2}}) \ \} \tag{4}$$

To simplify (4), we add $\frac{1}{\sqrt{2}}$ to all elements of matrix A, and then (4) is written as

$$L(U,V) = \sum_{i,j} \ \{ \ (A_{ij} - (UV)_{ij})^2 - \log(A_{ij} - (UV)_{ij}) \ \} \tag{5}$$

We will derive below algorithm for the problem of WNMF which minimize the following weighted cost functions: the Weighted Rayleigh Distance

$$L_W(A \| UV) = \frac{1}{2}\sum_{i,j} W \circ \{(A-UV)^2 - \log(A-UV)\} \tag{6}$$

where $W = \{W_{ij}\} > 0$ is a nonnegative weight matrix, and $X \circ Y$ is the Hadamard product (or element by element product) of the matrices X and Y. We only derive the updating rule for V since that of U can be infer in a similar fasion. We can consider the partial cost function for a single column of A, V and W, which we denote by a, v and w, respectively:

$$F(v) = F_w(a,Uv) = \frac{1}{2}\sum_i w_i \{(a_i - [Uv]_i)^2 - \log(a_i - [Uv]_i) \ \} \tag{7}$$

$$= \frac{1}{2}\sum_i w_i (a_i - [Uv])^2 + \frac{1}{2}\sum_i w_i \{-\log(a_i - [Uv]_i)\}$$
$$= \frac{1}{2}(a-Uv)^T D_W (a-Uv) + \frac{1}{2} w^T \{-\log(a-Uv)\}$$
$$= F_1(v) + F_2(v)$$

Where $D_W = diag(w)$ (8)

$$F_1(v) = \frac{1}{2}(a-Uv)^T D_w (a-Uv) \quad (9)$$

$$F_2(v) = \frac{1}{2} w^T \{-\log(a-Uv)\} \quad (10)$$

The first and two order derivatives of (9) and (10) can be obtained as (11)、(12) 、(13) and (14), respectively.

$$\nabla_v F_1(v^k) = -U^T D_w (a - Uv^k) \quad (11)$$

$$\nabla_v^2 F_1(v^k) = U^T D_w U \quad (12)$$

$$\nabla_v F_2(v^k) = \begin{pmatrix} w^T \dfrac{U_1}{a-Uv^k} \\ \vdots \\ w^T \dfrac{U_i}{a-Uv^k} \\ \vdots \\ w^T \dfrac{U_r}{a-Uv^k} \end{pmatrix}_{r\times 1} \quad (13)$$

$$\nabla_v^2 F_2(v^k) = \begin{pmatrix} w^T\left(\dfrac{U_1 \circ U_1}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_1 \circ U_i}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_1 \circ U_r}{(a-Uv^k)\circ(a-Uv^k)}\right) \\ \vdots & & \vdots & & \vdots \\ w^T\left(\dfrac{U_i \circ U_1}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_i \circ U_k}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_i \circ U_r}{(a-Uv^k)\circ(a-Uv^k)}\right) \\ \vdots & & \vdots & & \vdots \\ w^T\left(\dfrac{U_r \circ U_1}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_r \circ U_k}{(a-Uv^k)\circ(a-Uv^k)}\right) & \cdots & w^T\left(\dfrac{U_r \circ U_r}{(a-Uv^k)\circ(a-Uv^k)}\right) \end{pmatrix}_{r\times r} \quad (14)$$

By iterating over (7) a number of partial differential iterations until the specified threshold is satisfied, we can obtain (15).

$$v^{k+1} = v^k - v^k \circ \frac{\nabla_v F_1(v^k) + \nabla_v F_2(v^k)}{(\nabla_v^2 F_1(v^k) + \nabla_v^2 F_2(v^k))v^k} \quad (15)$$

Where $\dfrac{X}{Y}$ is the Hadamard division (or element by element division) of the matrix X and Y.

Before fusion, the fused ultrasound and MRI are converted into column vectors ($k\times 1$, where $k = n\times m$), respectively. Then the above two column vectors are then constructed into a $k\times 2$ matrix A, and the matrix is then approximately factorized into a $k\times 1$ column vector U and a 1×2 row vector V. Finally, when the iteration terminates, the $k\times 1$ vector U are then transformed into $n\times m$ matrix, namely gray level matrix of fusion image.

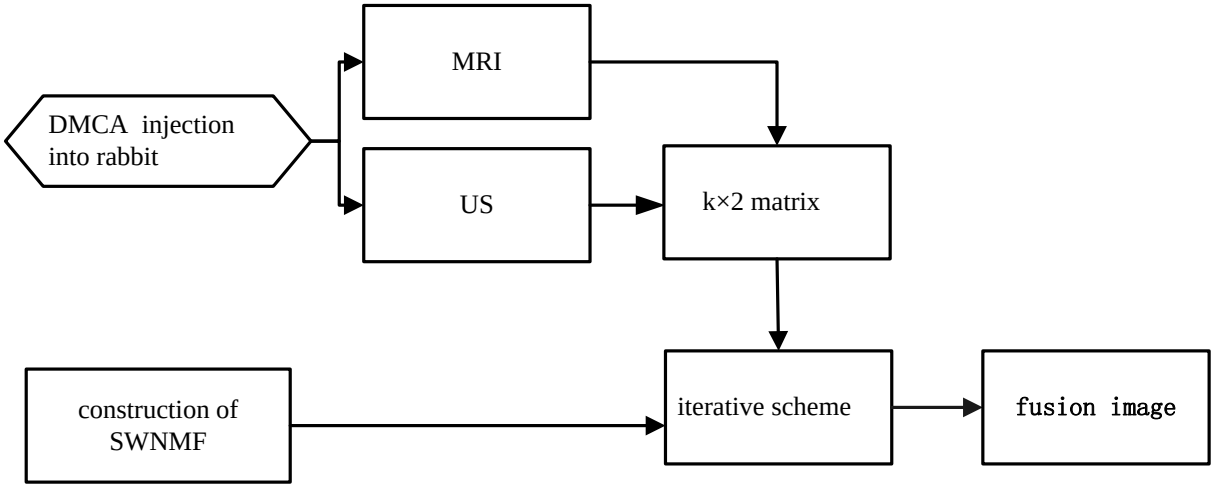


Figure 1. Flow chart of the proposed MRI-US fusion based on DMCA and SWNMF

In the process of image fusion, with the increasing of iterations, the fusion image pixel gray value is also real-time dynamic change result in that the gray distribution of each region of the image changes as well. So the weight corresponding to each region of the image should also change accordingly, and a fixed weight matrix does not reflect this dynamic process. In this paper, a iterative strategy is proposed, the weight matrix is dynamically adjusted according to the latest iterative results. The weighted coefficients of weight matrix reflect the importance of the corresponding pixels. When the iteration result has reached the target threshold, the adjustment of the corresponding weight coefficients is stopped. On the contrary, the weight coefficients are adjusted in the direction of the improvement of specific indicators.

3. Experiments

DMCA mentioned above is obtained from Jiangsu Key Laboratory for Biomaterials and Devices. DMCA can negatively strengthen T2-weighted (T2*WI) imaging signal. It can also enhance ultrasound backscattering echo intensity and positively increase the contrast and brightness of US.

Figure 2 include US and MRI fusion images without DMCA and with DMCA, respectively. As it can be intuitively observed from the perspective of fusion, Fig. 2 (c) has significant merits over Fig. 2 (c'). For better quantitative assessment, the four evaluation indexes AG (average gradient), EI (edge intensity), EN(Entropy) and SF(spatial frequency) are introduced. After fusion, performance evaluation with the use of DMCA had better improvement than that without DMCA. For example, AG, EI, EN and SF rises up from 2.0464 to 3.7630, from 22.5399 to 39.4881, from 6.5306 to 6.9099 and from 5.6550 to 10.2735 respectively. In short, the qualitative and quantitative analyses indicates that US-MRI fusion based on the above-mentioned method is effective.

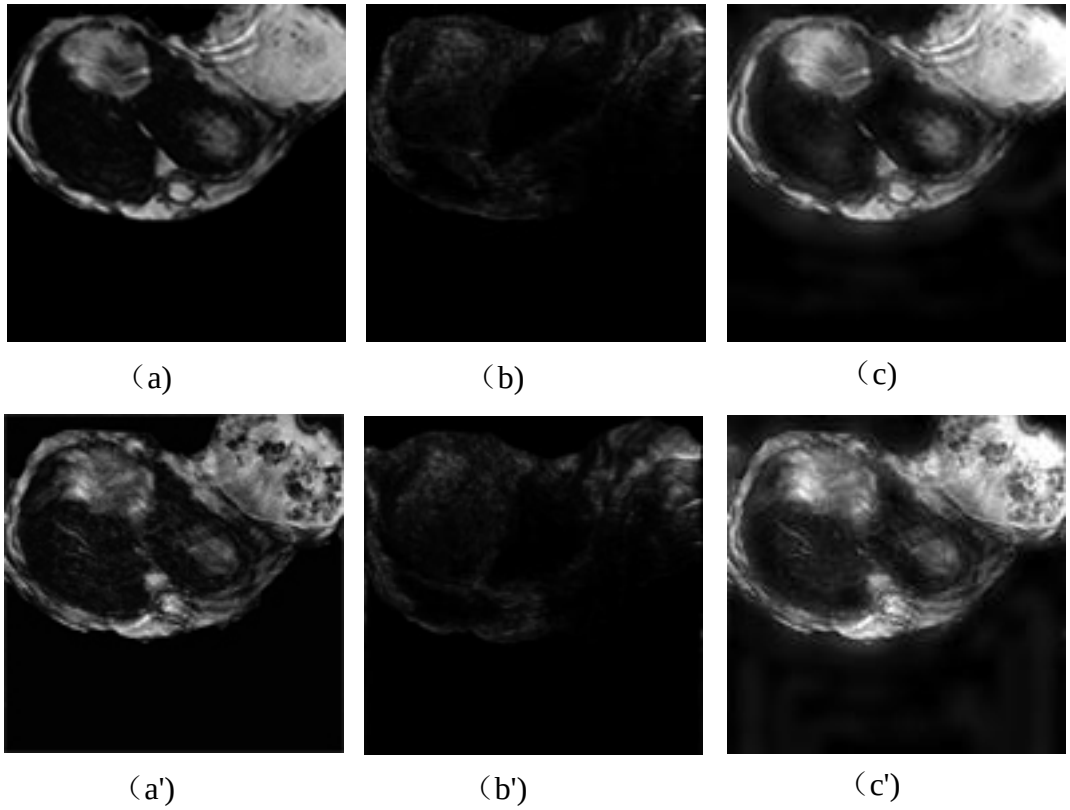


Figure 2. Fusion result comparison

(a)~(c) are MRI、US and fusion image without DMCA, respectively; (a')~(c') are MRI、US and fusion image with DMCA, respectively.

Table 1 Fusion result comparison

method category	evaluation indexes			
	AG	EI	EN	SF
the proposed method (without DMCA)	2.0464	22.5399	6.5306	5.6550
the proposed method (with DMCA)	3.7630	39.4881	6.9099	10.2735

4. Conclusions

As a new dual-modality contrast agent, DMCA is introduced into the field of medical image fusion. DMCA can enhance the texture details of medical imaging, which result in contrast improvement of important organization in the fusion image. Our fusion results are encouraging, However, they are still at preliminary stage. Further in vivo studies including toxicological and pathological studies will be necessary before our methods could be implemented in clinical applications.

References

[1] Yang, F. , Li, Y. X., Chen, Z. P., Zhang, Y., Wu, J. and Gu, N. (2009) Superparamagnetic iron oxide nanoparticle-embedded encapsulated microbubbles as dual contrast agents of magnetic resonance and ultrasound imaging, *Biomaterials* **30**, 3882–3890.

[2] Cai, X. W., Yang, F. and Gu, N. (2012) Applications of Magnetic Microbubbles for Theranostics, *Theranostics* **2**, 103–112.

[3] Hou, M., Luo, S. H., Yang, F., Ruan, X. B. and Gu, N. (2012) MRI-ultrasound registration based on dual modality contrast agent (in Chinese), *Chinese Science Bulletin* **57 (22)**, 2119-2126.

[4] Hou, M., Chen, C. X., Tang, D. L., Luo, S. H., Yang, F. and Gu, N. (2015) Magnetic microbubble-mediated ultrasound-MRI registration based on robust optical flow model, *BioMedical Engineering OnLine* **14(Suppl 1)**, S14.

[5] Hou, M., Luo, S. H., Yang, F., Ruan, X. B. and Gu, N. (2013) MRI- ultrasound fusion based on magnetic microbubble contrast agent (in Chinese), *Scientia Sinica (Informationis)* **43 (11)**, 1445-1459.

[6] Xiong, J., Luo, Y. and Tang, G. (2008) An Improved Optical Flow Method for Image Registration with Large-scale Movements, *Acta Automatica Sin* **34**, 760–764.

[7] Jung, E.M., Schreyer, A.G., Schacherer, D., Menzel, C., Farkas, S., Loss, M., Feuerbach, S., Zorger, N. and Fellner, C. (2009) New real-time image fusion technique for characterization of tumor vascularization and tumor perfusion of liver tumors with contrast-enhanced ultrasound, spiral CT or MRI: First results, *Clinical Hemorheology and Microcirculation* **43**, 57–69.

[8] Lindseth, F., Ommedal, S., Bang, J., Unsgard, G. and Nagelhus Hernes, T. A. (2001) Image fusion of ultrasound and MRI as an aid for assessing anatomical shifts and improving overview and interpretation in ultrasound-guided neurosurgery, *International Congress Series* **1230** 254–260.

[9] Schlaier, J. R., Warnat, J., Dorenbeck, U., Proescholdt, M., Schebesch, K. M. and Brawanski, A. (2004) Image fusion of MR images and real-time ultrasonography: evaluation of fusion accuracy combining two commercial instruments, a neuronavigation system and a ultrasound system, *Acta Neurochir (Wien)*, **146**, 271–277.

[10] Stone, M. (2005) A guide to analysing tongue motion from ultrasound images, *Clinical Linguistics and Phonetics* **19 (6-7)**, 455–502.

Impact of carotid bifurcation geometry on atherosclerotic formation: A hemodynamic study

Xin-Ke Yao[1†*], Yan Cai[1†], Jie Gao[2†], Ge-Lin Xu[2‡], Zhi-Yong Li[1,3‡]

[1]School of Biological Science & Medical Engineering, Southeast University, Nanjing 210096, P.R. China
[2]Department of Neurology, Jinling Hospital, Medical School of Nanjing University, Nanjing 210002, P.R. China
[3]School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4001, Australia
*Presenting author: yaofe6881@sina.cn

† These authors contributed equally to this article.

‡Corresponding author: gelinxu@nju.edu.cn; zylicam@gmail.com

Abstract:

Background: Recent studies have demonstrated major variations in carotid bifurcation geometry, in favor of the concept that individual vascular anatomy may play a role in the development of atherosclerosis. To test these assumptions, the present study aimed to investigate the impact of wide variations on carotid artery hemodynamics by computational fluid dynamic (CFD) simulations.
Method: In the present work, six groups of three-dimensional synthetic models of carotid bifurcation with different morphological parameters were established by ANSYS. The geometric variations included internal carotid artery (ICA) /common carotid artery (CCA) diameter ratio, sinus/CCA diameter ratio, external carotid artery (ECA)/CCA diameter ratio, ICA angle, tortuosity and planarity. Pulsatile blood flow through a model of the carotid artery bifurcation was simulated using a finite volume numerical method. The changes of inlet velocity during the cardiac cycle were taken into consideration while outlet pressure was a constant value. The simulated results were visualized in TECPLOT for subsequent analysis.
Result: The temporal average value of wall shear stress (AWSS) on the sinus wall was obtained from the hemodynamic simulations. Multiple regression analysis revealed a significant ($P<0.001$) relationship between AWSS and both ICA angle ($\beta=0.48$) and sinus/CCA diameter ratio ($\beta=-0.48$). Larger ICA angles generally increase the AWSS on the sinus wall, hence lowering the risk of plaque build-up. In contrast, high sinus/CCA diameter ratio was found to decrease AWSS on the sinus for geometries with the baseline ICA angle.
Conclusion: Compared to benchmark model, the changes of carotid bifurcation geometry can lead to apparent differences in hemodynamics. It may therefore be reasonable to consider certain geometric features to be surrogate markers of low AWSS, which is mainly related to atherosclerotic formation.
Keywords: Atherosclerosis, Hemodynamics, Carotid Bifurcation, Wall Shear Stress.

Introduction

As one of the most dangerous diseases, atherosclerosis of large arteries causes one fourth of stroke cases worldwide [1]. Atherosclerotic plaques are prone to appearing near arterial bifurcations and bends, which leads to an acknowledged concept that the unique vascular anatomy of each person plays an important role in the development of this disease. The ability to understand the underlying mechanism of geometry effects in atherosclerotic plaque formation would be of great clinical value for patient prognosis [2]- [3].

While there is some controversy over the most indicative geometry markers related to the hemodynamic environment and subsequent atherosclerosis formation, many clinical studies strongly suggest the relationship between geometry of carotid bifurcation and disturbed flow. Lee et al. [4] modeled 50 carotid bifurcation geometries derived from magnetic resonance imaging (MRI) of 25 young adults and found that disturbed flow is sensitive to both proximal area ratio and bifurcation tortuosity. Markl et al. [5] analyzed the in vivo distribution of absolute WSS in the carotid bifurcation to evaluate its dependence on geometry. Time-resolved 3D blood flow in this research was achieved with flow-sensitive 4D MRI in 64 normal carotid bifurcations and 17 carotid arteries with moderate stenosis. Common to these studies is the assumption that secondary effects of stenosis on geometry can be ignored. To systematically examine the influence of geometry on the blood flow pattern, a Y-shaped model was established by Nguyen et al. to quantify the risk of atherogenesis associated with different bifurcation angle and the out-of-plane angle. However, they excluded many other geometry factors, such as the diameter ratios and tortuosity.

The objective of this work is to test the impact of wide variations of carotid bifurcation geometry on hemodynamics by control variate method. Six geometric variations, including internal carotid artery（ICA）/common carotid artery (CCA) diameter ratio, sinus/CCA diameter ratio, external carotid artery (ECA)/CCA diameter ratio, ICA angle, tortuosity and planarity, , are designed to establish 30 synthetic models. Computational fluid dynamic (CFD) simulations are performed with the same initial and boundary conditions. The numerical results of flow patterns and wall shear stress (WSS) on the sinus wall are compared between different synthetic models. In addition, the significance analysis is carried out between the temporal average value of wall shear stress (AWSS) and certain geometric variations.

Materials and Method

Geometry of the models

The human CCA bifurcates into the ECA and the ICA at approximately the level of the fourth cervical vertebra. Bharadvaj et al. [6] digitized film angiograms of patients to develop the Y-shaped model of the average human carotid bifurcation (Y-AHCB), which then has been widely used in experimental and numerical investigations of hemodynamics in the carotid bifurcation [7]- [13]. However, much arteriograms and specimen figures revealed the fact that most ICA is bent inward [14]- [16]. To optimize it, the tuning-fork-sharped model of average human carotid bifurcation (TF-AHCB) was proposed by observation and statistical analysis of

specimens from 74 cadavers [17]. In the present study, the benchmark model was established according to the TF-shaped model (Figure 1).

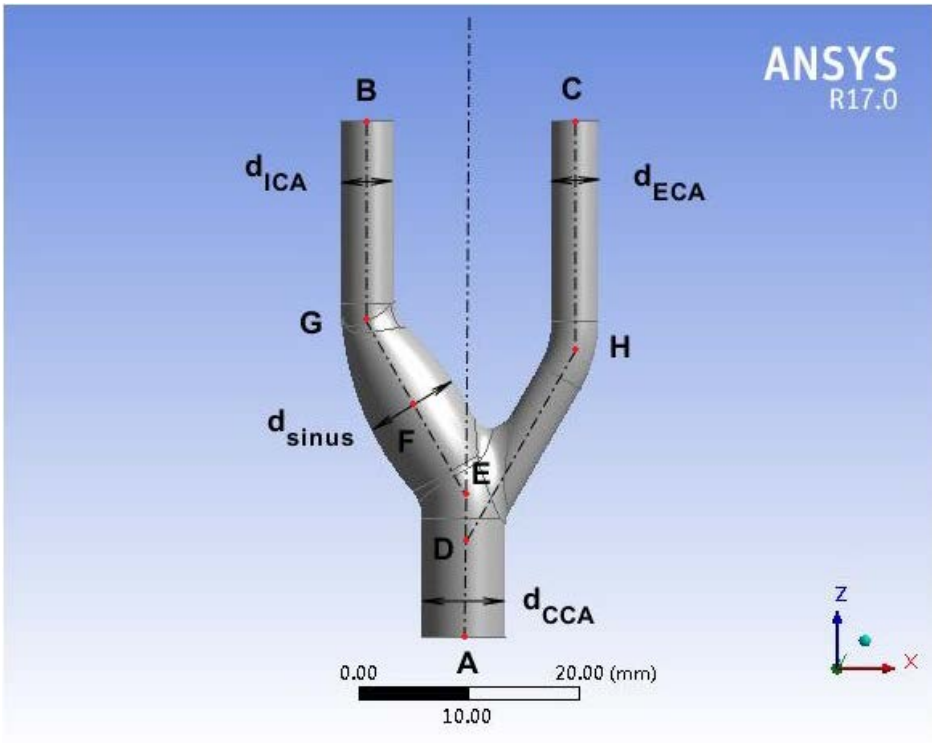


Figure 1. Typical carotid artery geometry of the benchmark model (the tuning-fork-sharped model of average human carotid bifurcation, TF-AHCB model).

The geometric values of the benchmark model are presented in Table 1. Specifically, the arteries were modeled by cylindrical pipes through the operation of sweep. A specific graphics revolved around the center line of ICA to form the sinus region. The CCA was modeled by a cylinder with length of 13mm. The ICA outlet and ECA outlet were at the same horizontal plane in benchmark model. That is, both ICA and ECA outlet had a perpendicular distance of 47mm to CCA inlet. The diameter of ICA was 4.87mm while ECA had a diameter of 4.27mm. The carotid sinus, whose maximum diameter was 8.60mm, had the same center line as the ICA. The maximum cross section of sinus intersected the center line at F, which was the midpoint of sinus axis. In order to smooth the interface, a fixed radius blend of 3mm was applied in present work.

The ICA angle was defined as the angle between the center line of sinus and of CCA, while planarity angle was defined as the angle of the ICA with the CCA-ECA plane. The ratio between the diameter of distal ICA and CCA was called ICA/CCA diameter ratio. Similarly, the ratio between the diameter of distal ECA and CCA was called ECA/CCA diameter ratio. Sinus/CCA diameter ratio was calculated as the maximum diameter of ICA sinus divided by CCA diameter. Although vessel tortuosity was usually defined as a ratio between Euclidean distance and the length of the centerline, it was represented as the angle between the axis of distal ICA and sinus in this study. Hence, larger tortuosity angle means straighter ICA.

Table 1. Dimensions of the benchmark model (mm)

CCA			ICA				ECA		
AD	AE	d_{CCA}	EG	GB	d_{sinus}	d_{ICA}	DH	HC	d_{ECA}
8.5	13	7.58	17.486	18.856	8.60	4.87	18.205	21	4.27

Diameter ratios varied based on the measurements taken by Ding et al, where unit dimensions of ICA, ECA, sinus and CCA diameter have been discussed. According to the data list in Table 1, the baseline ICA/CCA, sinus/CCA, ECA/CCA diameter ratio (0.64, 1.13, and 0.56, respectively) can be calculated. The average bifurcation angle was 63.6° in a control group of older subjects presented by Thomas et al. [18]. Therefore, the baseline ICA angle and ECA angle were both 30° in current study. Because distal ICA was parallel to CCA and distal ECA in benchmark model, baseline tortuosity angle was doubtlessly 150°. The variation of ICA angle ranged from 10° to 50° and tortuosity angle from 120° to 160°. In addition, planarity angle was varied from 0 to 10° based on the discovery that the average off-plane angle was 7.0° for young adults versus 8.5° for another age group.

Computational fluid dynamics

Tetrahedral-element meshes were generated by ICEM CFD, Version 17.0 (ANSYS Inc.) using a maximum element size of 0.5mm and a minimum size of 0.2mm, which was demonstrated to be sufficient for resolving WSS [19].

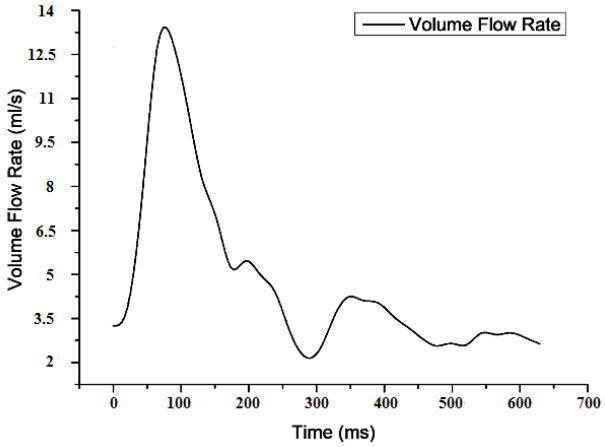


Figure 2. Inlet volume flow rate in one cardiac cycle

Blood was assumed to be a Newtonian fluid in large blood vessels as the size of blood cells there is small compared to the diameter of the tubes [20]. The viscosity of blood is

approximately four times that of water and its density ranges from 1020 to 1150 kg/m^3. Therefore, blood was modeled to have a density of 1056 kg/m^3 and a viscosity of 0.0035 Pa·s in present work. The vessel wall was assumed to be rigid without displacement.

The CCA inflow boundary condition was a time-dependent volume flow waveform of a healthy volunteer (Figure 2). According to the inlet volume flow rate, the mean blood velocity at the inlet can be obtained by following equation:

$$v_{maen}(t) = \frac{q(t)}{\pi(\frac{d}{2})^2} \tag{1}$$

where d is the diameter of the inlet, t is time, q is the volume flow rate and v_{mean} is the mean velocity. The pressure at the ICA and ECA outlets was set to be equal to 75mmHg. Transient CFD simulation was processed with Fluent, Version 17.0 (ANSYS Inc.) for one cardiac cycles with a time step of 0.00628 seconds. In the CFD simulation, the convergence criterion was satisfied when the residual of continuity was less than 10^{-4}.

The main hemodynamic parameter studied in present work was AWSS. It was defined as:

$$\text{AWSS} = \left|\frac{1}{T}\int_0^T \overrightarrow{WSS}\, dt\right| \tag{2}$$

Additionally, the median values of parameter were used for further analysis. The hemodynamic results were obtained by Fluent and Tecplot 360, Version 2014 (Tecplot Inc.).

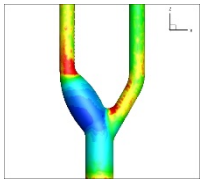
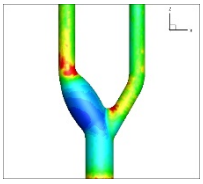
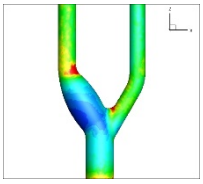
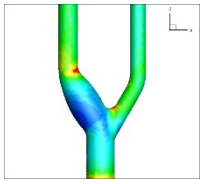
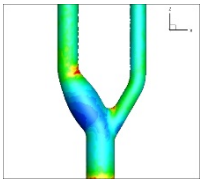
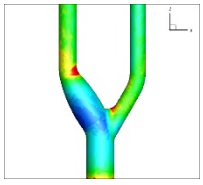
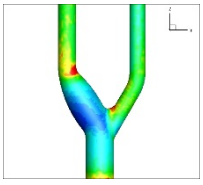
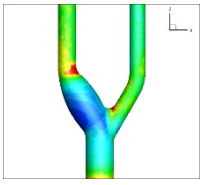
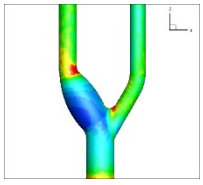
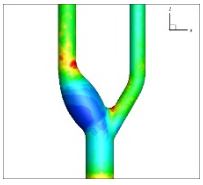
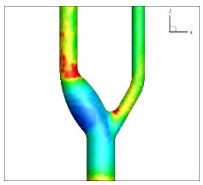
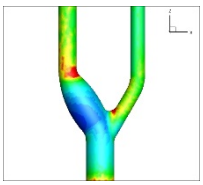
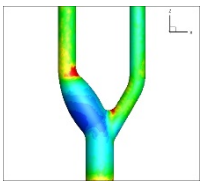
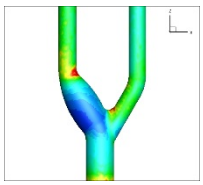
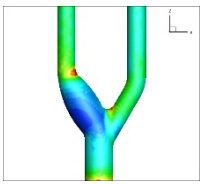
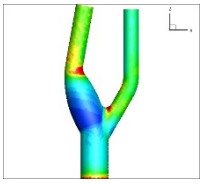
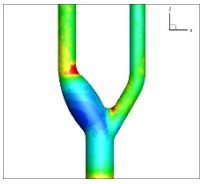
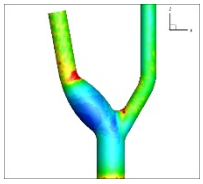
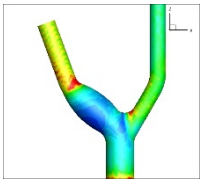
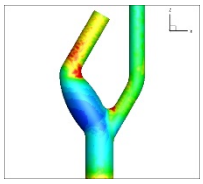
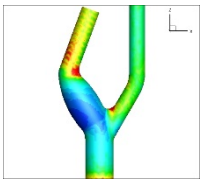
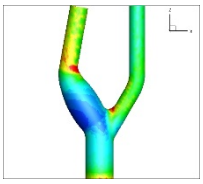
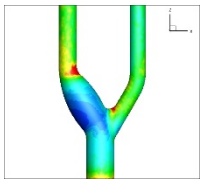
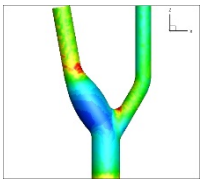
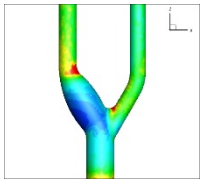
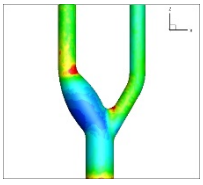
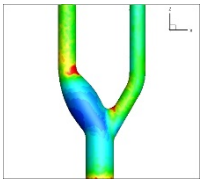
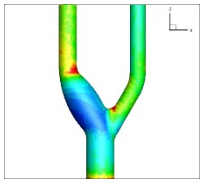
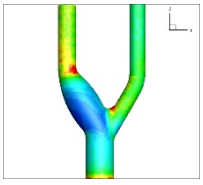
Multiple linear regression was used to quantify the relationship between AWSS and predictor variables. The overall quality of the regression was assessed by Pearson's correlation coefficient, adjusted by the number of independent predictors (R^2_{adj}). The relative contributions of the geometric variables were identified by the standardized regression coefficients (β). Statistical analyses were carried out by SPSS 19.0.

Results

Simulations and analysis were performed for a wide range of diameter ration and angles. The changes of instantaneous WSS at systolic peak with different geometry were shown in Figure 3. The region indicated by dark blue color has WSS lower than the critical value of 0.4 Pa [21] while red color represents WSS larger than 15 Pa. The dark blue region only appears on the outer wall of the carotid sinus, which indicates the low WSS on the sinus wall as a stimulative factor of intima-media thickness. Furthermore, the WSS is lower at the entry of sinus and higher at the exit of sinus. There is a red region at the terminal of the sinus in all synthetic models. When the ICA/CCA diameter ratio increases, this red region becomes contractible most dramatically.

To analyze the relationship of AWSS and geometry variations more directly, we perform the curves of the median of sinus AWSS in different synthetic models (Figure 4). The median of sinus AWSS changes linearly according to the parameters. It decreased when sinus/CCA and ECA/CCA diameter ratio increased. In contrast, the rise of ICA/CCA diameter ratio can lead

to high AWSS on the sinus wall, which indicated the hindering effect of narrow downstream. AWSS increased with the enlargement of ICA angle. This trend became gentle gradually when

ICA/CCA=0.54 	**ICA/CCA=0.59** 	**ICA/CCA=0.64** 	**ICA/CCA=0.69** 	**ICA/CCA=0.74** 
Sinus/CCA=1.03 	**Sinus/CCA=1.08** 	**Sinus/CCA=1.13** 	**Sinus/CCA=1.18** 	**Sinus/CCA=1.23** 
ECA/CCA=0.46 	**ECA/CCA=0.51** 	**ECA/CCA=0.56** 	**ECA/CCA=0.61** 	**ECA/CCA=0.66** 
ICA angle=10 ° 	**ICA angle=20 °** 	**ICA angle=30 °** 	**ICA angle=40 °** 	**ICA angle=50 °** 
Tortuosity=120 ° 	**Tortuosity=130 °** 	**Tortuosity=140 °** 	**Tortuosity=150 °** 	**Tortuosity=160 °** 
Planarity=0 ° 	**Planarity=2.5 °** 	**Planarity=5 °** 	**Planarity=7.5 °** 	**Planarity=10 °** 

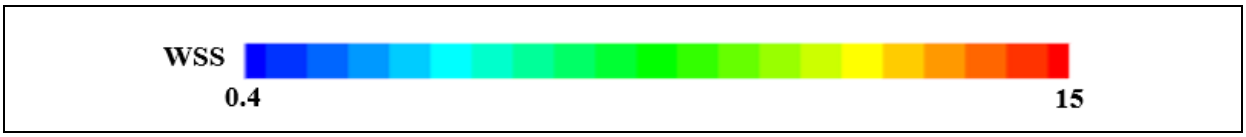


Figure 3. WSS at systolic peak with different geometry

ICA angle was larger than 30°. Tortuosity angle had a similar effect on AWSS. While it should be noticed that the larger tortuosity angle was, the straighter the ICA was. Hence, bent ICA was more liable to have plaques on the sinus wall. With the baseline ICA angle of 30°, higher off-plane angle results in lower AWSS, which was consistent with the notion suggested by Nguyen et al. [22] previously that planarity angle is a major contributor to vascular disease when ICA angle is larger than 25°.

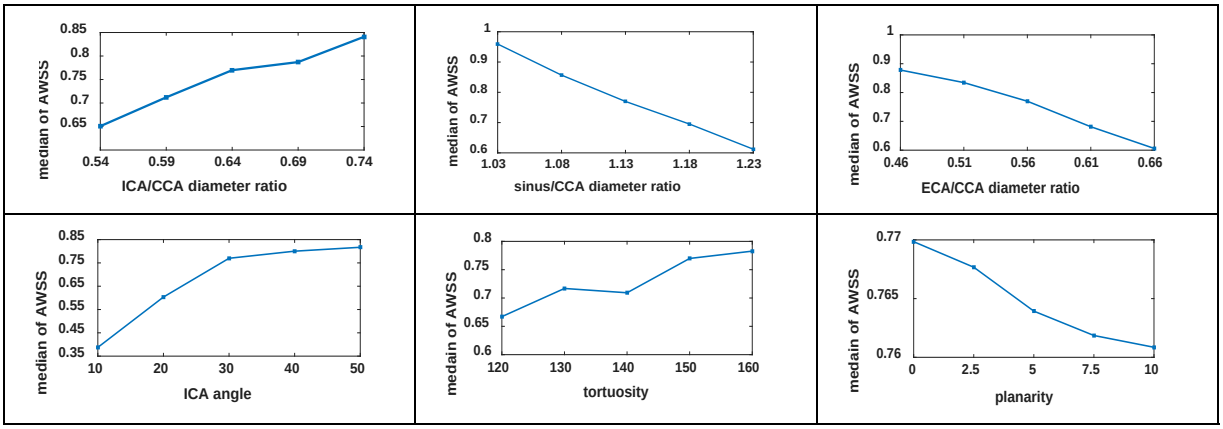


Figure 4. The relationship between AWSS and geometrical variations

Table 2. Multiple linear regressions of AWSS on the sinus wall with geometric variations

		Median of AWSS
Model Quality	R^2_{adj}	0.634
	P	<0.001
ICA/CCA Diameter Ratio	β	0.254
	P	0.030
Sinus/CCA Diameter Ratio	β	-0.478
	P	<0.001
ECA/CCA Diameter Ratio	β	-0.389
	P	0.002
ICA Angle	β	0.483
	P	<0.001
Tortuosity	β	0.147
	P	0.197

Planarity	β	0.072
	P	0.552

As plaque would preferentially form around the carotid artery bifurcation (extracranial site) and the carotid artery siphon (intracranial site), carotid sinus was chosen as the particular area to be analyzed. As summarized in Table 2, multiple regressions revealed that the combination of ICA angle and sinus/CCA diameter ratio may be far stronger predictors of AWSS on the sinus wall ($P<0.001$). However, β for tortuosity and planarity was not significant ($P>0.1$). The median of AWSS had a positive correlation with ICA angle (β=0.483) and a negative correlation with sinus/CCA diameter ratio (β=-0.478).

Discussion

It is acknowledged that an individual's vascular anatomy or local hemodynamics will affect the progression of atherosclerosis disease. Most studies were carried out to explore the impact of geometry on stenosis through patient-derived carotid arteries. Thomas et al. [18] researched on 25 young adults and 25 older subjects and pointed out that there is a complex interrelationship among vascular geometry, local hemodynamics, vascular aging, and atherosclerosis. Interindividual variations of carotid bifurcations rise dramatically due to aging or early atherosclerosis. Besides, 178 samples were analysed by Phan et al. [23] to demonstrate that carotid anatomy and geometry, such as ICA angle and ICA radius at the bifurcation, may enhance the risk of stenosis independent of traditional vascular risk factors like age and smoker. Multiple factors will change simultaneously in individual carotid arteries. To examine the effects of each variation systematically, six groups of ideal models with different morphologic parameters were established in this study. There is only one variation changing in each group. The present investigation has demonstrated wide geometric variations in the exposure to low WSS.

It is noteworthy that low WSS was invariably focal on the sinus wall in all synthetic models. Studied of hemodynamics have demonstrated that the pulsatile ‘‘disturbed’’ flow (with low and oscillating shear stress) patterns are one of the dominant flow features for atherosclerosis formation [24]. Therefore, there is no doubt that the risk of atherosclerotic disease is directly associated with the low WSS region on the carotid sinus. The median value of AWSS on the particular portion of sinus wall was calculated to evaluate the sinus hemodynamic level in this study. It was found that the main factors leading to atherosclerosis formation were ICA angle and sinus/CCA diameter ratio [25]. The important role for branch angle has been suggested in previous model studies [26], which is consistent with our finding here. Moreover, Karino and Goldsmith [25] insisted the relative importance of diameter ratio versus angle on vortex formation by calculating a T-junctions model with wide variations.

As mentioned in Methods, a few assumptions were made to simplify the analysis. For example, the arterial wall was considered as a rigid one with no slip which was contrary to the elasticity of human vessels. Moreover, flow entering the vasculature after passing a straight tube with a cross sectional area the same as inlet region will lead to a fully-developed entry

flow different from reality. These limitations make it reasonable to believe that only a minor effect on the resulting carotid bifurcation flow dynamics has been analysed in current study.

Aside from the modeling assumptions, a major shortcoming was the deficiency of other hemodynamic parameters. The present work concentrated on the WSS distribution of the carotid arteries. The impact of wall shear stress gradient (WSSG) and oscillatory shear index (OSI) on atherosclerosis formation was ignored. The continued perfection of hemodynamic analysis will be performed in further study.

The following directions may be further investigated in the future. First, more 3-dimensional geometric parameters can be analysed. Most morphologic variations suggested in present work, such as diameter ratio, angle and even tortuosity, were confined to 2-dimensional space which is incongruent with realistic carotid bifurcation anatomy. Then, patient-specific models based on imaging of carotid artery will be attached to verify the consequence of current study. In addition, significant displacement caused by blood flow should be taken into consideration in CFD simulation. The elasticity of arterial wall, in turn, will make a difference in flow pattern. This interaction can be coupled in future study.

Conclusion

In summary, the present study investigated the impact of wide variations on carotid artery hemodynamics by computational fluid dynamic (CFD) simulations. Geometric variation, including diameter ratio, ICA angle, tortuosity and planarity, can result in significant changes of WSS distribution. Among them, ICA angle and sinus/CCA diameter ratio were most relevant with AWSS on the sinus wall. That is, these two parameters had remarkable effect on plaque formation. The certain geometric features suggested in this work, in the sense, can be used as surrogate markers of low AWSS, and even as predictors of atherosclerotic formation.

Acknowledgements

This research is supported by the National Basic Research Program of China (973 program) (No. 2013CB733800), the National Nature Science Foundation of China (No. 11422222, 11302050, 81571143), the Fundamental Research Funds for the Central Universities, the National Demonstration Center for Experimental Biomedical Engineering Education (Southeast University), and ARC (FT140101152).

References

[1] Ruan, L., Chen, W., Srinivasan, S. R., Sun, M., Wang, H., Toprak, A., & Berenson, G. S. (2009). Correlates of common carotid artery lumen diameter in black and white younger adults. *Stroke*, *40*(3), 702-707.

[2] Ku, D. N., Giddens, D. P., Zarins, C. K., & Glagov, S. (1985). Pulsatile flow and atherosclerosis in the human carotid bifurcation. Positive correlation between plaque location and low oscillating shear stress. *Arteriosclerosis, thrombosis, and vascular biology*, *5*(3), 293-302.

[3] Younis, H. F., Kaazempur-Mofrad, M. R., Chan, R. C., Isasi, A. G., Hinton, D. P., Chau, A. H., ... & Kamm, R. D. (2004). Hemodynamics and wall mechanics in human carotid bifurcation and its consequences for atherogenesis: investigation of inter-individual variation. *Biomechanics and modeling in mechanobiology*, *3*(1), 17-32.

[4] Lee, S. W., Antiga, L., Spence, J. D., & Steinman, D. A. (2008). Geometry of the carotid bifurcation predicts its exposure to disturbed flow. *Stroke*,*39*(8), 2341-2347.

[5] Markl, M., Wegent, F., Zech, T., Bauer, S., Strecker, C., Schumacher, M., ... & Harloff, A. (2010). In vivo wall shear stress distribution in the carotid artery.*Circulation: Cardiovascular Imaging*, *3*(6), 647-655.

[6] Bharadvaj, B. K., Mabon, R. F., & Giddens, D. P. (1982). Steady flow in a model of the human carotid bifurcation. Part I—flow visualization. *Journal of biomechanics*, *15*(5), 349-362.

[7] Perktold, K., Resch, M., & Florian, H. (1991). Pulsatile non-Newtonian flow characteristics in a three-dimensional human carotid bifurcation model. *J Biomech Eng*, *113*(4), 464-475.

[8] Perktold, K., Thurner, E., & Kenner, T. (1994). Flow and stress characteristics in rigid walled and compliant carotid artery bifurcation models.*Medical and Biological Engineering and Computing*, *32*(1), 19-26.

[9] Perktold, K., & Rappitsch, G. (1995). Computer simulation of local blood flow and vessel mechanics in a compliant carotid artery bifurcation model. *Journal of biomechanics*, *28*(7), 845-856.

[10] Lee, D., & Chiu, J. J. (1996). Intimal thickening under shear in a carotid bifurcation—a numerical study. *Journal of biomechanics*, *29*(1), 1-11.

[11] Gijsen, F. J. H., Palmen, D. E. M., Van der Beek, M. H. E., Van de Vosse, F. N., Van Dongen, M. E. H., & Janssen, J. D. (1996). Analysis of the axial flow field in stenosed carotid artery bifurcation models—LDA experiments. *Journal of biomechanics*, *29*(11), 1483-1489.

[12] Rindt, C. C. M., & Steenhoven, A. V. (1996). Unsteady flow in a rigid 3-D model of the carotid artery bifurcation. *TRANSACTIONS-AMERICAN SOCIETY OF MECHANICAL ENGINEERS JOURNAL OF BIOMECHANICAL ENGINEERING*, *118*, 90-96.

[13] Ma, P., Li, X., & Ku, D. N. (1997). Convective mass transfer at the carotid bifurcation. *Journal of biomechanics*, *30*(6), 565-571.

[14] DeBakey, M. E., Lawrie, G. M., & Glaeser, D. H. (1985). Patterns of atherosclerosis and their surgical significance. *Annals of surgery*, *201*(2), 115.

[15] Salzar, R. S., Thubrikar, M. J., & Eppink, R. T. (1995). Pressure-induced mechanical stress in the carotid artery bifurcation: a possible correlation to atherosclerosis. *Journal of biomechanics*, *28*(11), 1333-1340.

[16] Nicholls, S. C., Phillips, D. J., Primozich, J. F., Lawrence, R. L., Kohler, T. R., Rudd, T. G., & Strandness, D. E. (1989). Diagnostic significance of flow separation in the carotid bulb. *Stroke*, *20*(2), 175-182.

[17] Ding, Z., Wang, K., Li, J., & Cong, X. (2001). Flow field and oscillatory shear stress in a tuning-fork-shaped model of the average human carotid bifurcation. *Journal of Biomechanics*, *34*(12), 1555-1562.

[18] Thomas, J. B., Antiga, L., Che, S. L., Milner, J. S., Steinman, D. A. H., Spence, J. D., ... & Steinman, D. A. (2005). Variation in the carotid bifurcation geometry of young versus older adults. *Stroke*, *36*(11), 2450-2456.

[19] Moyle, K. R., Antiga, L., & Steinman, D. A. (2006). Inlet conditions for image-based CFD models of the carotid bifurcation: is it reasonable to assume fully developed flow?. *Journal of biomechanical engineering*, *128*(3), 371-379.

[20] Ku, D. N. (1997). Blood flow in arteries. *Annual review of fluid mechanics*,*29*(1), 399-434.

[21] Malek, A. M., Alper, S. L., & Izumo, S. (1999). Hemodynamic shear stress and its role in atherosclerosis. *Jama*, *282*(21), 2035-2042.

[22] Phan, T. G., Beare, R. J., Jolley, D., Das, G., Ren, M., Wong, K., ... & Srikanth, V. (2012). Carotid artery anatomy and geometry as risk factors for carotid atherosclerotic disease. Stroke, 43(6), 1596-1601.

[23] Nguyen, K. T., Clark, C. D., Chancellor, T. J., & Papavassiliou, D. V. (2008). Carotid geometry effects on blood flow and on risk for vascular disease. *Journal of biomechanics*, *41*(1), 11-19.

[24] Ku, D. N., & Giddens, D. P. (1983). Pulsatile flow in a model carotid bifurcation. *Arteriosclerosis, Thrombosis, and Vascular Biology*, *3*(1), 31-39.

[25] Karino, T., Kwong, H. H., & Goldsmith, H. L. (1979). Particle flow behaviour in models of branching vessels: I. Vortices in 90 degrees T-junctions.*Biorheology*, *16*(3), 231-248.

[26] Friedman, M. H., O'brien, V., & Ehrlich, L. W. (1975). Calculations of pulsatile flow through a branch: implications for the hemodynamics of atherogenesis. *Circulation Research*, *36*(2), 277-285.

Research on Pipe Vibration Analysis and Optimal Supporter Arrangement

***J.J. Liu[1,2,3], Ding Yu[1,2,3], and †Long Yu[1,2,3]**

[1]State Key Laboratory of Ocean Engineering, Shanghai Jiaotong University, Shanghai 200240, China;
[2]Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration(CISSE), Shanghai 200240, P.R. China;
[3]School of Naval Architecture, Ocean & Civil Engineering, Shanghai Jiaotong University, Shanghai 200240, China;
*Presenting author: 15855702880@163.com
†Corresponding author: yulone@sjtu.edu.cn

Abstract

The discharge pipe line is one of the important devices for cutter suction dredgers, while the pipes connecting to cutter, pump and valves guarantee the high dredging and transportation efficiency. The flow mixed with water and soil with different density is highly turbulent and unsteady, which induces vibrations along the pipe and also cause abrasion inside the pipe especially at bends and T branches. The flow exciting force induce the pipe stress and deformation while the pipe deformation will influence the flow pattern. Due to the large diameter of discharge pipe and lots of bends and T branches, the Fluid Structure Interaction (FSI) should be introduced for pipe dynamic analysis and supporter design. The uncertainty of the mixed flow makes it not possible to design or test all the working conditions. Therefore, the CFD technique has been employed for numerical analysis. The typical pipe lines and their supporters with specified exciting source are investigated. Modal analysis and stress results give proof for optimal supporter design and pipe line.

Keywords: discharge pipe line, FSI, modal analysis, CFD

1.Introduction

The pipe line is a convenient way of transport. With the rapid development of economic, discharge pipe line has been used in more and more fields, such as biomedical, aerospace engineering, automobile manufacturing, civil engineering, electronics industry, and especially in the naval architecture and marine engineering(Levitan et al., 1991). The discharge pipe line plays a very important role in the ships, which are always concerned as large and complex systems. For example, the hydraulic pipe systems have great effects on the starting, reversing, transmission and other action of the ships, which can reduce the pressure on the ship power station(Dai Xueliang et al.,2000). In addition to the hydraulic pipes, ships also contain a variety of discharge pipe lines, such as the pipe lines of cold water, hot water, sewage, black water and oil. They are all hidden in the corner to maintain the ship's normal operation.

There are very complex pipe line systems in the cabin of the large ships. In the process of voyage, the various pipe lines take the task to discharge the oil, water, gas and other fluids of the whole ship. Working in a high temperature and high pressure environment, once the pipe line systems are damaged, they will pose a serious threat to the security of the ships. For example, if the pipe lines supplying fuel to the main engine leaks, it will not only cause fire and explosion, but also make the main engine break down. In some serious cases, ships may extremely lose power, which will even lead to serious shipwrecks(Liang Chunyu,2013). Therefore, reasonable and effective measures must be taken to reduce the pipe vibration

To analyze vibration and stability of the pipe lines, people mainly study the influence of the flow velocity on the dynamic characteristics of the pipeline system. Ignoring the fluid compressibility, people usually pay attention to the interaction between the flow and the

pipeline. At the end of the nineteenth century, people began to study the problem of pipeline vibration, but the study was once paused for a long time. From the fifties of the twentieth century, people began to systematically study the vibration and stability of the pipeline, and since then, the articles around this field is endless(Liu Zhongqun,2001). In the field of vibration and stability of the pipeline, Chen and Paidoussis (1984) have done a representative research work. When the flow is non-stick, incompressible and steady, if neglecting the gravity, structural damping and external force on pipe, equations of the bending and free vibration of the straight pipe line is given by:

$$EI\frac{\partial^4 u_y}{\partial z^4}+MW^2\frac{\partial^2 u_y}{\partial z^2}+2MW\frac{\partial^2 u_y}{\partial z\partial t}+(M+m)\frac{\partial^2 u_y}{\partial t^2}=0 \tag{1}$$

In which EI is the bending stiffness of the pipe, M is the linear density of the fluid, m is the line density of the pipe, W is the average flow rate of the fluid, u_y is the displacement of the transverse vibration of the pipe, z is the axial coordinate of the pipe, and t is the time variable.

Paidoussis and Issid(1980) proposed a more general equation based on the above equation, which takes into account the axial load of tension and compression of the pipe, the gravity and the damping of the pipe. The form of the equation is given by:

$$\begin{aligned}&E^*I\frac{\partial^5 u_y}{\partial z^4\partial t}+EI\frac{\partial^4 u_y}{\partial z^4}+\{MW^2-T+PA(1-2\mu\delta)-[(M+m)g-M\frac{\partial W}{\partial t}]\\&-(L-z)\}\frac{\partial^2 u_y}{\partial z^2}+2MW\frac{\partial^2 u_y}{\partial z\partial t}+(M+m)g\frac{\partial u_y}{\partial z}+C\frac{\partial u_y}{\partial t}+(M+m)\frac{\partial^2 u_y}{\partial t^2}=0\end{aligned} \tag{2}$$

In which E^*I is the internal resistance coefficient of the material, C is the viscous damping coefficient of the support, δ (0 or 1) is the factor that indicates whether the end of the pipe can be moved, μ is the Poisson ratio, P is the average pressure in the pipe, T is the axially load on the end of the pipe.

Paidoussis(1987) gave an overview of the vibration of the discharge pipeline system, and introduced the bifurcation behavior of the discharge pipeline system and showed many obtained research results. He stated that there were two kinds of instabilities: divergence instability and flutter instability. The kinds of instabilities that firstly appeared depended on the supporting situation. His review was mainly for the linear problem. And over the past decades, people had much promising work done in the nonlinear vibration of discharge pipeline system, finding new phenomenon never found in the linear range.

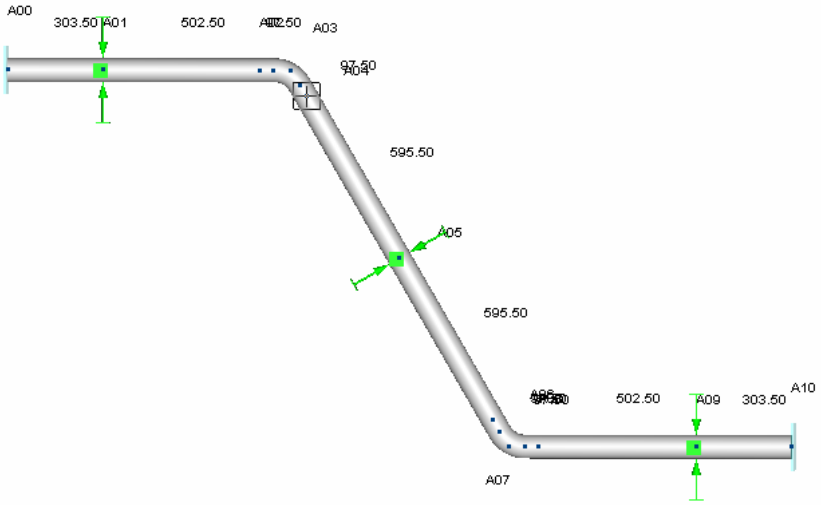
In this paper, the finite element analysis method is used to study the static structure and natural vibration characteristics of the pipeline. Models of pipeline are established both in Autopipe and Ansys for numerical simulation. After calculation by the softwares, the stress, strain and deformation of the pipeline can be obtained and compared. And in the dynamic simulation, the vibration response of the pipeline can also be calculated. Finally, various damping measures and the most optimal supporter arrangement can be taken to ensure the safety of the discharge pipeline according to the simulation results

2.Analysis by Autopipe

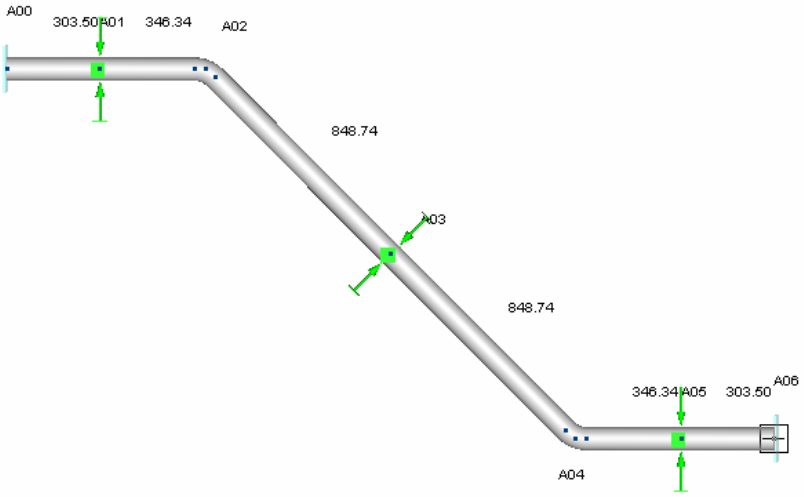
2.1 Model of pipeline

The establishment of pipeline model in Autopipe consists of three steps: defining pipe

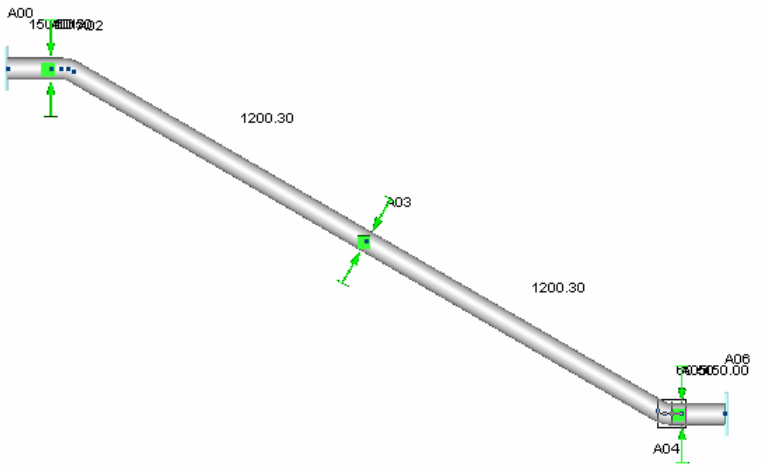
properties, connecting pipes and adding pumps, valves and supports. The outside diameter of the pipeline is 76mm, the wall thickness is 2.5mm, and the density is 7850 kg/m^3, Young's modulus is 200 GPa, Poisson's ratio is 0.3.Both ends of the pipeline are pumps, so both ends are set to fixed. Guide brackets are installed in the support position and all the gaps of the guide brackets are set to 0 to ensure a better support. Different models are established with the angle of the corner A= 120°, 135°, 150°, and in each degree the distance between two brackets also varies from L=1600mm, 1900mm to L=2200mm. (Difference L means difference distance between the fixed ends and the brackets).Some of the models is shown in Figure 1.



（a）A=120°, L=1600mm



（b）A=135°, L=1600mm

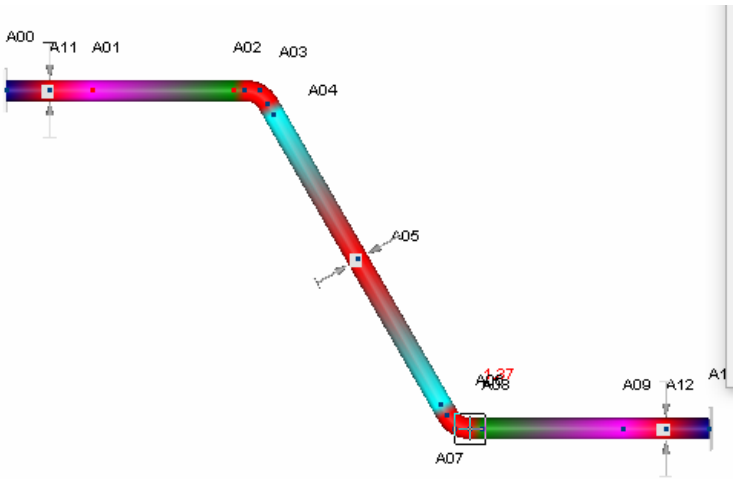


（c）A=150°, L=1900mm

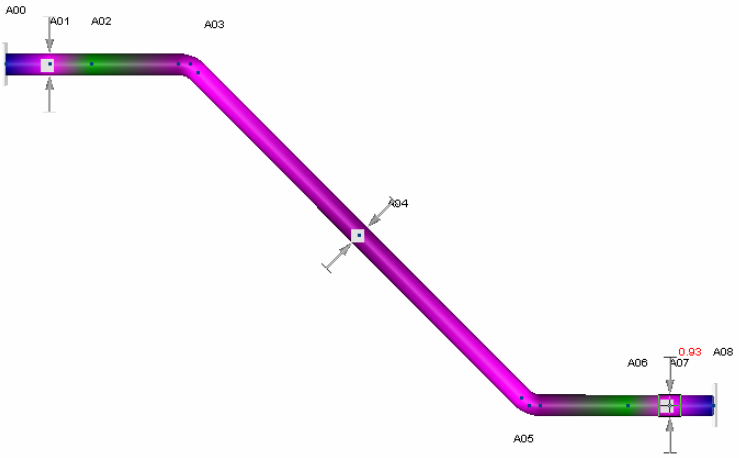
Figure 1. Models established by Autopipe

2.2 Static Analysis

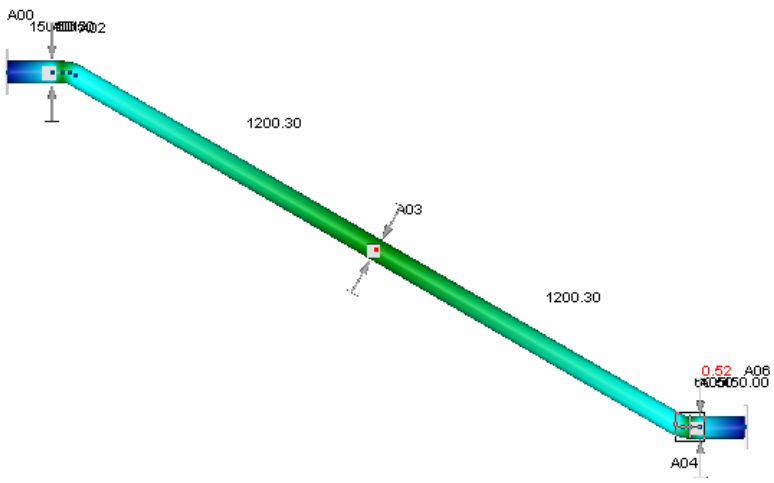
After simulation and calculation in Autopipe, the static stress of the pipeline will be obtained. Figure 2 are the static stress of some models.



（a）A=120°, L=1600mm



（b）A=135°, L=1600mm

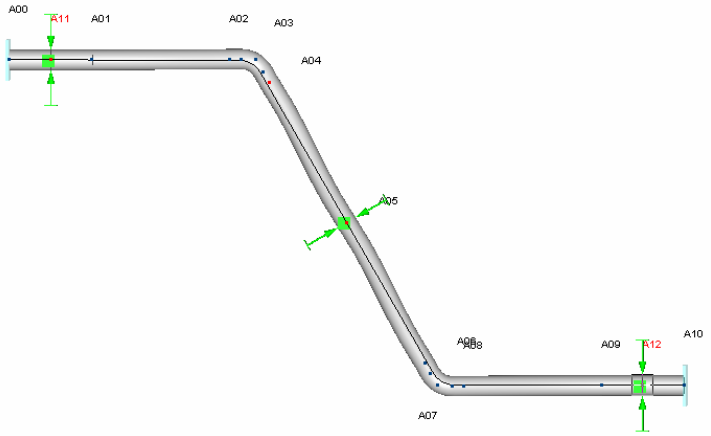


（c）A=150°, L=1900mm

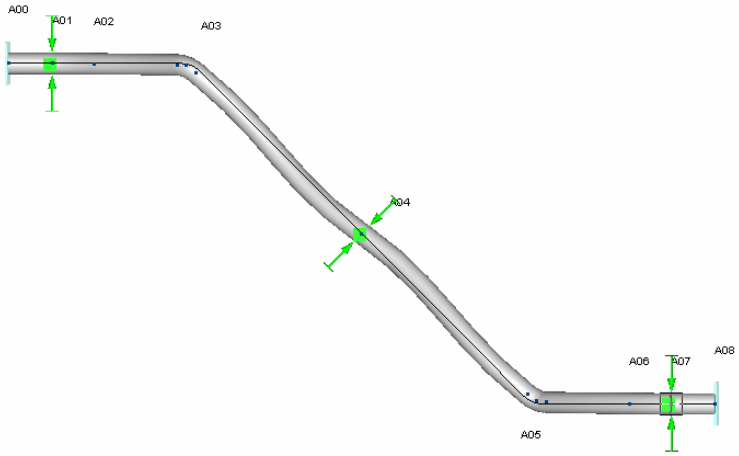
Figure 2. Static stress

2.3 Modal analysis

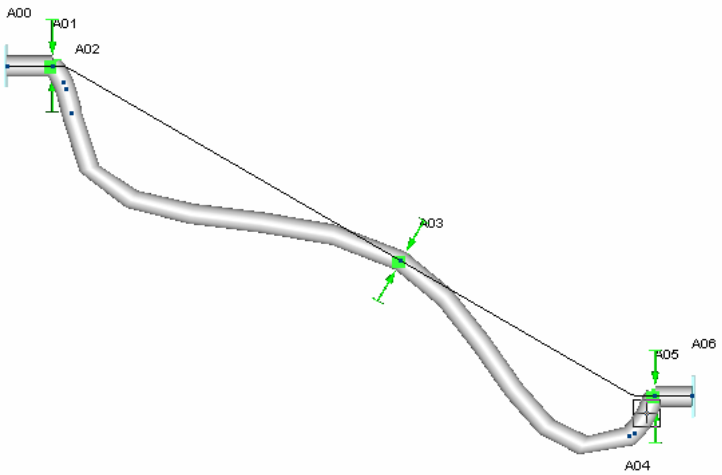
The modal of vibration of some modals is shown in Figure 3.



（a）A=120°, L=1600mm



（b）A=135°, L=1600mm



（c）A=150°, L=1900mm

Figure 3. Modal of vibration

2.4 brief summary

From the static results, it can be concluded as follows.

(1)There is only sustained stress rather than swelling stress in all different cases.

(2)In the case of A=120 degrees and 135 degrees.

a. The max stress appears at the brackets.

b. The closer the brackets to the corner, the smaller the sustained stress and the sustained stress ratio.

(3)In the case of A=150 degrees

a. The max stress appears at the both fixed end when no brackets are set.

b. When brackets are 50mm or 100mm to the fixed end, the max stress appears at the midpoint. And the closer the brackets to the corner, the smaller the sustained stress and the sustained stress ratio.

c. When brackets are 150mm to the fixed end, the max stress appears at the corner.

From the dynamic results, it can be concluded as follows.

(1)As the angle increases, the modal frequency also increases;

(2)The first-order modal frequency are far greater than 50Hz;

(3)If concerning about the sustained stress only, A=120 degrees is better than others.

3.Analysis by Ansys

3.1 model of pipeline

The model built in Ansys is similar to the model in Autopipe. The outside diameter of the pipeline is 76mm, the wall thickness is 2.5mm, and both ends of the pipeline are set to fixed. The angle of the corner is initially set as 120 degrees, and the length of the three straight pipeline is 0.806m, 1.191m and 0.806m. Considering of meshing, the radius of the corner can not be too small, so it’s set to 0.0931m. And after meshing, the total number of the elements are 4448 and the nodes are 8960.

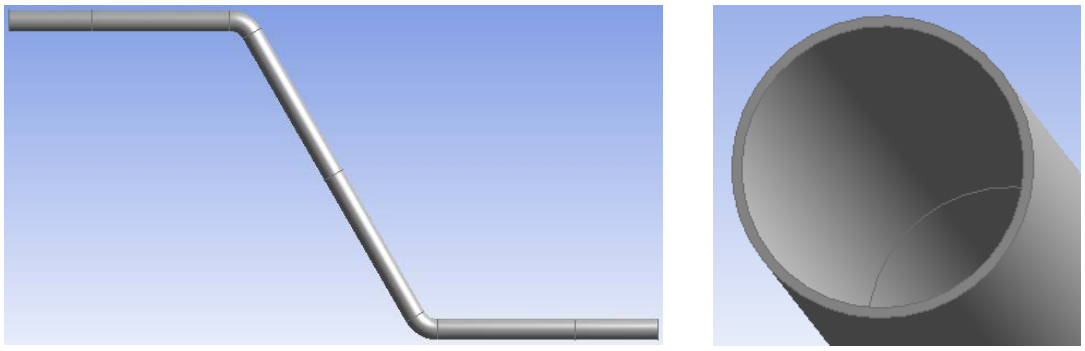


Figure 4. Model of pipeline in Ansys

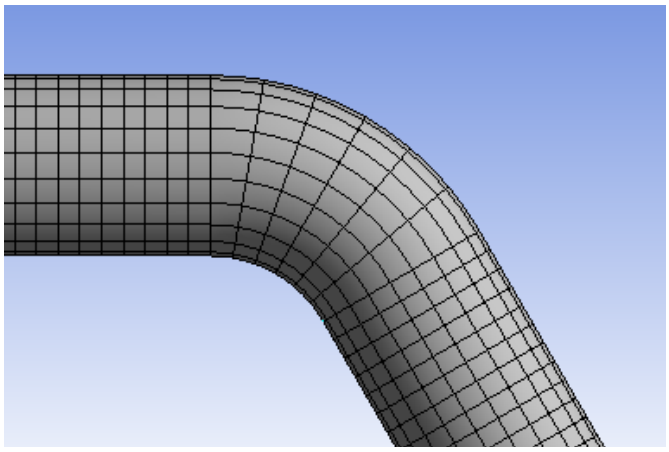
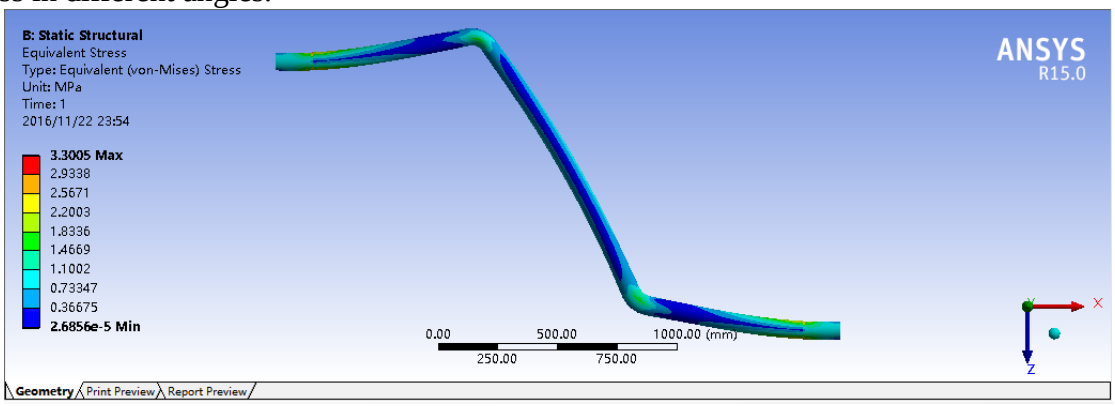


Figure 5 Mesh of the corner

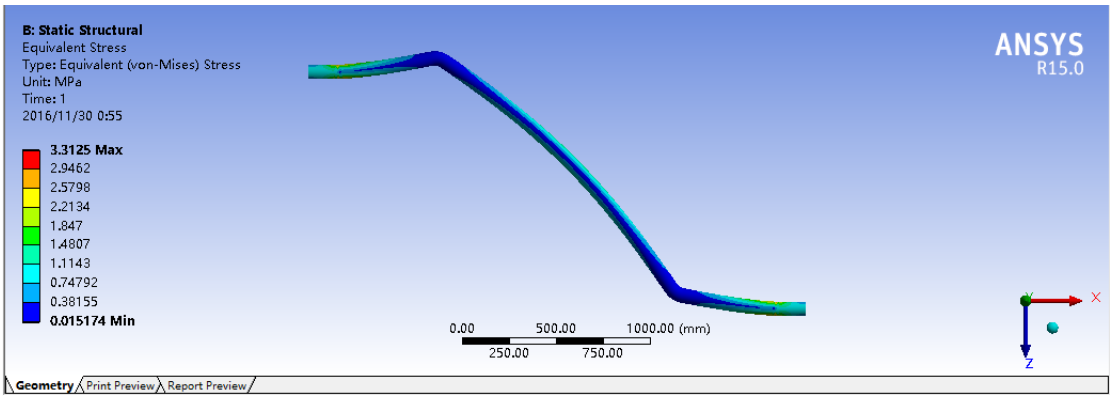
3.2 Static analysis

3.2.1 Influence of angle changing

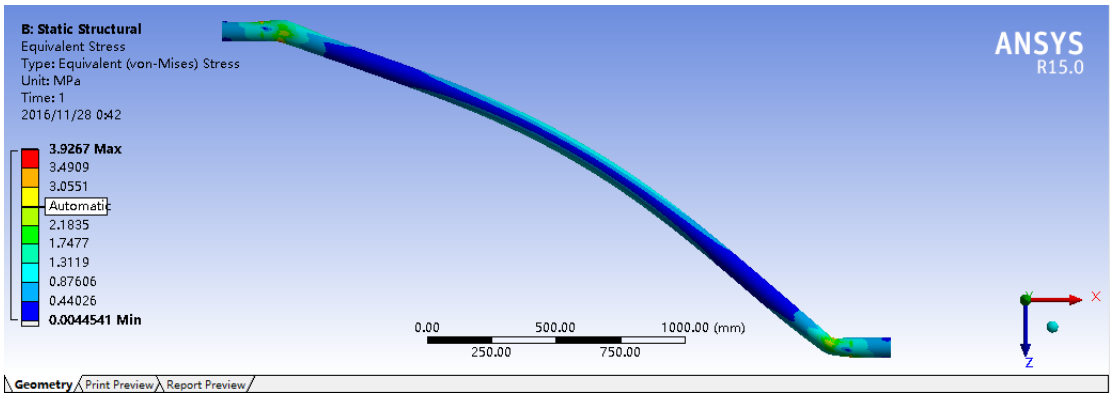
Keep the distance between the brackets and fixed endings to 155.50mm, change the angle of the corner, and obtain the stress of the pipeline in different angles. Figure 6 shows the pipeline stress in different angles.



（a）A=120°



（b）A=135°



（c）A=150°

Figure 6. Static stress in different angles

The max stress, strain and deformation in different cases are showed in the Table 1.

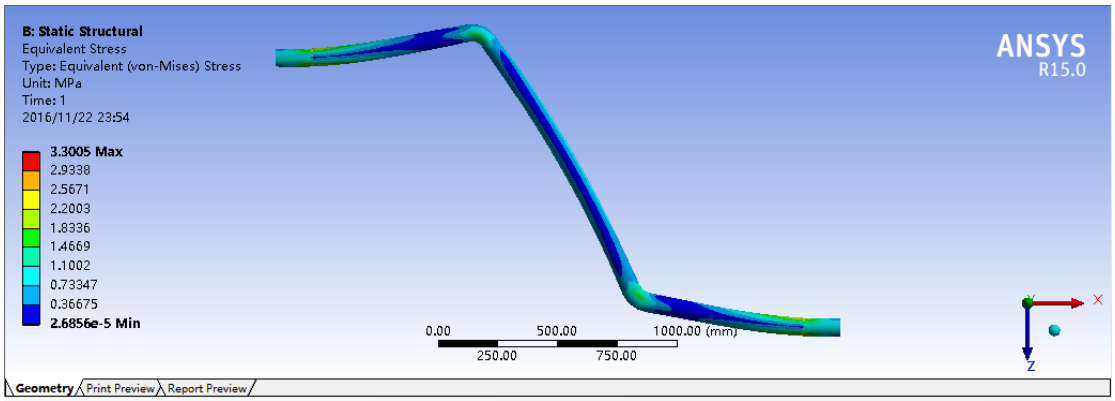
Table 1. Max stress, strain and deformation of the pipeline

Angle of the corner/°	120	135	150
Stress/MPa	3.3005	3.3125	3.9267
Strain/$\times 10^{-5}$	1.668	1.6723	1.9641
Deformation/$\times 10^{-2}$mm	7.7511	9.2919	9.0469

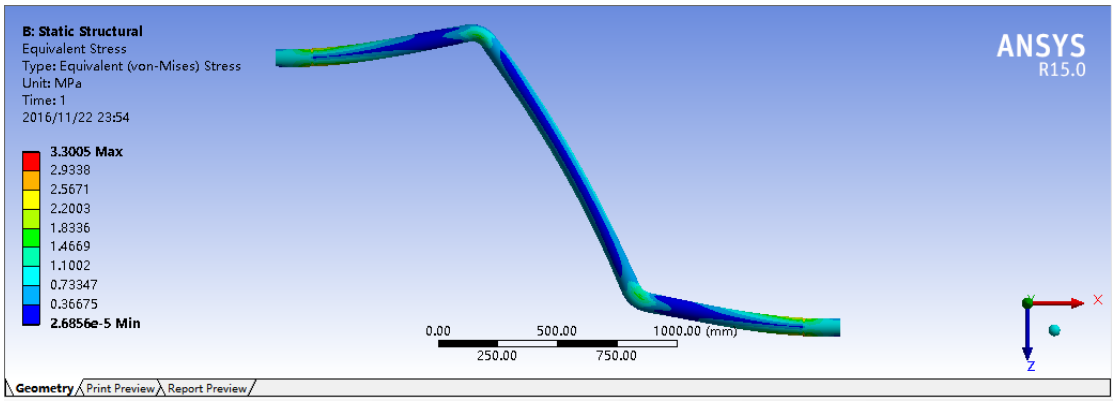
From the Figure 6 and Table 1, it can be concluded that when the angle of the corner increases, the max stress and strain also increase. So the angle of the corner can not be too large so that the structure would not be damaged by the large stress and strain. In addition, when the angle increases 15 degrees (from 120 to 135), the stress increases 0.012MPa, the strain increases 0.0043e-5; when the angle increases 15 degrees (from 135 to 150), the stress increases 0.6142MPa, the strain increases 0.2918e-5. So it can be concluded that when the angle becomes larger, the effect posed by the angle changing on strain changing will be more obvious.

3.2.2 Influence of the brackets position

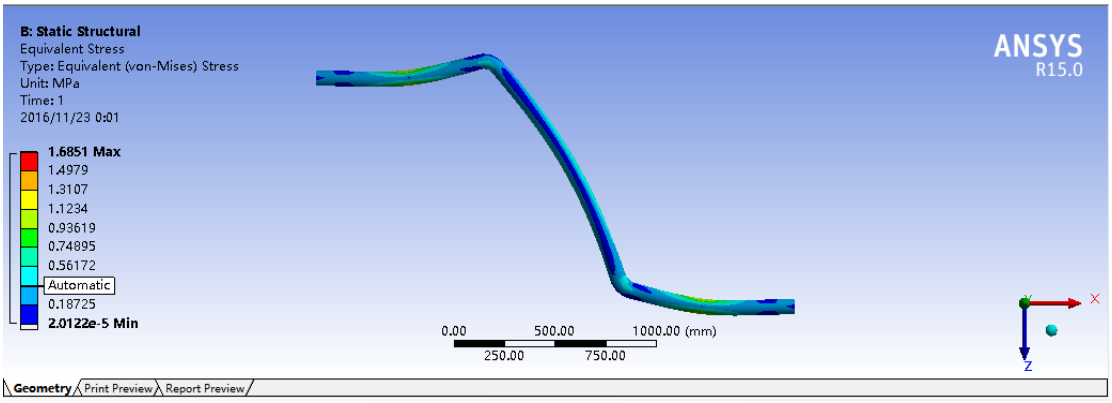
Change the distance between the brackets and the fixed ends, keep the angle of the corner A=120 degrees (the performance of the pipeline is better when A=120 degrees according to above analysis), and analyze the dynamic characteristics. The pipeline stress is shown as in Figure 7.



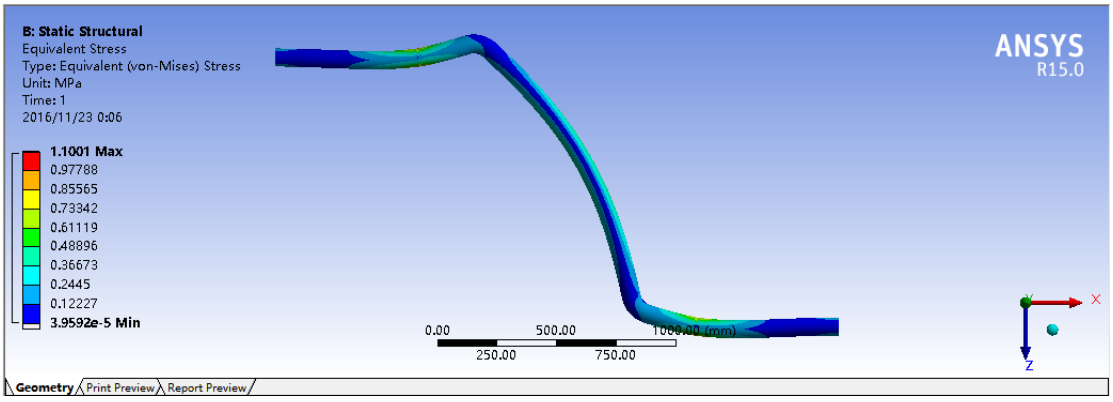
（a）155.50mm to the fixed end



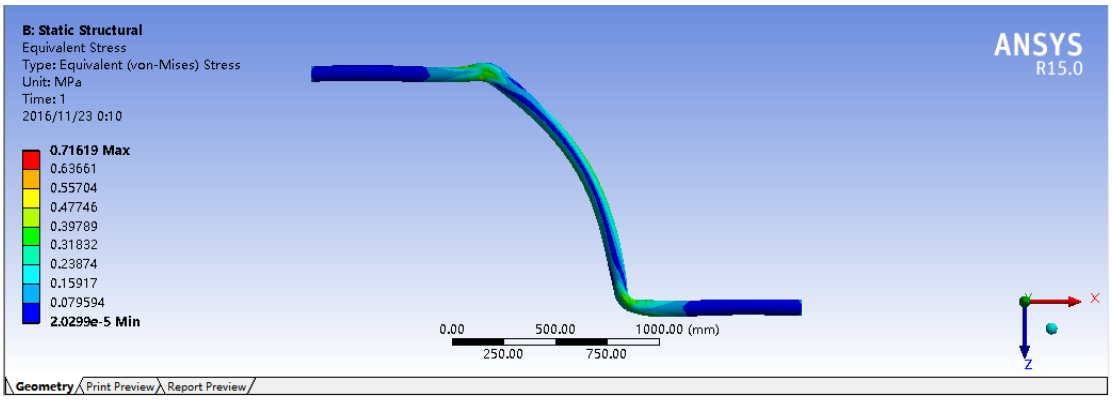
（b）305.50mm to the fixed end



（c）455.50mm to the fixed end



（d）605.50mm to the fixed end



（e）755.50mm to the fixed end

Figure 6. static stress in different distance

The max stress, strain and deformation in different cases are showed in the table4-2.

Table 2. Max stress, strain and deformation of the pipeline

Distance to the fixed end/mm	155.50	305.50	455.50	605.50	755.50
Stress/MPa	3.3005	2.4304	1.6851	1.1001	0.71619
Strain/$\times 10^{-6}$	16.68	12.277	8.514e	5.5015	3.6249
Deformation/$\times 10^{-2}$mm	7.7511	4.8394	2.8164	1.5596	0.85126

From the Figure 6 and Table 2 above, it can be concluded that when the distance between the guide brackets and the fixed ends increases (the guide brackets become closer to the corner), the max stress and strain of the structure decreases. As a result, it would be safer for the pipeline structure. So in the actual project, it's a good choice to shorten the distance between the guide brackets and the corners to ensure the safety of the pipeline structure.

3.3 modal analysis

3.3.1 Influence of angle changing

Change the angle of the corner, keep the distance between the brackets and the fixed ends, and analyze the dynamic characteristics. The results of the first six modals is shown in Table 3 and Table 4.

Table 3. Frequency of vibration

Angle of the corner/°	120	135	150
first-order frequency/Hz	60.781	55.079	55.79
second-order frequency/Hz	61.78	56.168	56.758
third-order frequency/Hz	105.08	124.63	161.48
forth-order frequency/Hz	183.15	169.86	162.93
fifth-order frequency/Hz	249.32	224.79	319.32
sixth-order frequency/Hz	484.08	279.27	321.2

Table 4. Modal of vibration

Angle of the corner/°	120	135	150
first-order modal/mm	13.426	13.809	14.043
second-order modal/mm	13.002	13.762	14.077
third-order modal/mm	16.062	13.847	13.736
forth-order modal/mm	17.114	12.802	13.549
fifth-order modal/mm	14.555	16.224	13.508
sixth-order modal/mm	16.71	15.662	13.8

From the modal analysis, it can be concluded that the vibration frequency of each modal has no fixed relationship with the angle. The vibration modal changes with the angle, but the difference is not so significant. So there is little meaning to make the angle a criterion for program evaluation. In addition, the external input frequency is 50Hz, so the chosen frequency should be away from 50Hz to avoid resonance. If only concerning about the first and second modal, the model of 120 degrees it the best.

3.3.2 Influence of the brackets position

Change the distance between the brackets, and still keep the angle A=120 degrees, the results of the first six modals is shown in Table 5 and Table 6.

Table 5. Frequency of vibration

Angle of the corner/°	120	135	150
first-order frequency/Hz	60.781	75.436	95.568
second-order frequency/Hz	61.78	77.341	98.298
third-order frequency/Hz	105.08	135.16	185.11
forth-order frequency/Hz	183.15	197.99	227.64
fifth-order frequency/Hz	249.32	285.21	345.65
sixth-order frequency/Hz	484.08	495.2	497.89

Table 6. Modal of vibration

Angle of the corner/°	120	135	150
first-order modal/mm	13.426	14.344	15.574
second-order modal/mm	13.002	13.923	15.373
third-order modal/mm	16.062	16.582	16.965
forth-order modal/mm	17.114	16.768	15.86
fifth-order modal/mm	14.555	16.484	18.817
sixth-order modal/mm	16.71	18.076	18.124

From the modal analysis, it can be concluded that the vibration frequency of each modal would increase when the distance between brackets and fixed ends increases. In this cases, the external input frequency is 50Hz, so the chosen frequency should be away from 50Hz to avoid resonance. If only concerning about the first and second modal, it's best to set the distance between brackets and fixed ends to 755.50mm. If the input frequency changes, the optimal distance may also change.

4.Comparison of calculation results in Autopipe and Ansys

Since there are two different calculation software, there may be deviations in the calculation results. In order to learn more about the deviations, two sets of models are chosen to compare the results.

Model 1: angle A=120 degrees, both ends are fixed, no guide brackets;

Model 2: angle A=120 degrees, both ends are fixed, guide brackets are 303.5mm to the fixed ends.

Calculation results of the two models are shown in Table 7 and Table 8.

Table 7. Comparison of the Model 1

	Autopipe	Ansys
Stress/MPa	3.9	3.7003
first-order frequency/Hz	46.1753	52.827
second-order frequency/Hz	46.3340	53.471
third-order frequency/Hz	70.6685	89.773

forth-order frequency/Hz	148.2002	175.56
fifth-order frequency/Hz	177.5147	228.88
sixth-order frequency/Hz	465.9268	422.14

Table 8. Comparison of the Model 2

	Autopipe	Ansys
Stress/MPa	2.3	2.4304
first-order frequency/Hz	64.5740	75.436
second-order frequency/Hz	65.8979	77.341
third-order frequency/Hz	98.9831	135.16
forth-order frequency/Hz	159.1330	197.99
fifth-order frequency/Hz	194.4324	285.27
sixth-order frequency/Hz	678.3102	495.2

From the two tables above , it can be concluded that:

(1) The calculation results of the static analysis are similar in Autopipe and Ansys. If based on the Autopipe results, the error between them is as follows:
Model 1:Δ1=（3.9-3.7003）/3.9=5.12%
Model 2:Δ2=（2.4304-2.3）/2.3=5.67%.
The error stabilized at around 5%, which is acceptable in the engineering practices, so the static results are ideally matched.

(2) The calculation results of modal analysis are different in Autopipe and Ansys. In the first five modal, all the results in Ansys is significantly greater than those in Autopipe. But in the sixth modal, the results in Autopipe increase rapidly, and become greater than that in Ansys. If based on the Autopipe results, the error between them is as follows:
First modal, Model 1: Δ11=（52.827-46.1753）/46.1753=14.41%
First modal, Model 2:Δ12=（75.436-64.574）/64.574=16.82%
Sixth modal, Model 1: Δ61=(465.9268-422.14)/465.9268=9.40%
Sixth modal, Model 2: Δ62=（678.3102-495.2）/678.3102=27.00%
The error in first modal is around 15%, and the error in varies with models. Both results are not good.

During the modal analysis, both Autopipe and Ansys use the same equation:

$$([K]-w_i^2[M])\{\phi_i\}=0 \tag{3}$$

In which K is the stiffness matrix, ϕ_i modal matrix, W_i is the vibration frequency matrix, and M is the mass matrix.

There is no difference between the two softwares in the calculation principle, and the density Young's modulus, Poisson's ratio and shear modulus of the pipeline are the same. So if only concerning the finite element method, the two results should be consistent, but in fact the results have large difference. The reason for the large errors may be as follows.

(1) In the calculation, Autopipe uses a prestressed modal analysis, while Ansys does not consider prestressing, which results in a different stiffness matrix.

(2) The two models established in Autopipe and Ansys may not be exactly the same. Maybe there are some difference in the corner or the fixed ends

5.Initial analysis with FSI method

When the fluid flows in the pipeline, it induces pipe stress and deformation while the pipe deformation will influence the flow pattern at the same time(Yu Meng,2007), which is solved by FSI method. It is used in this paper to get the fluid pressure in the mid-section of the pipeline, the shear stress in the pipeline wall and the normal stress in the pipeline wall. The simulation results are as shown in Figure 7, Figure 8 and Figure 9.

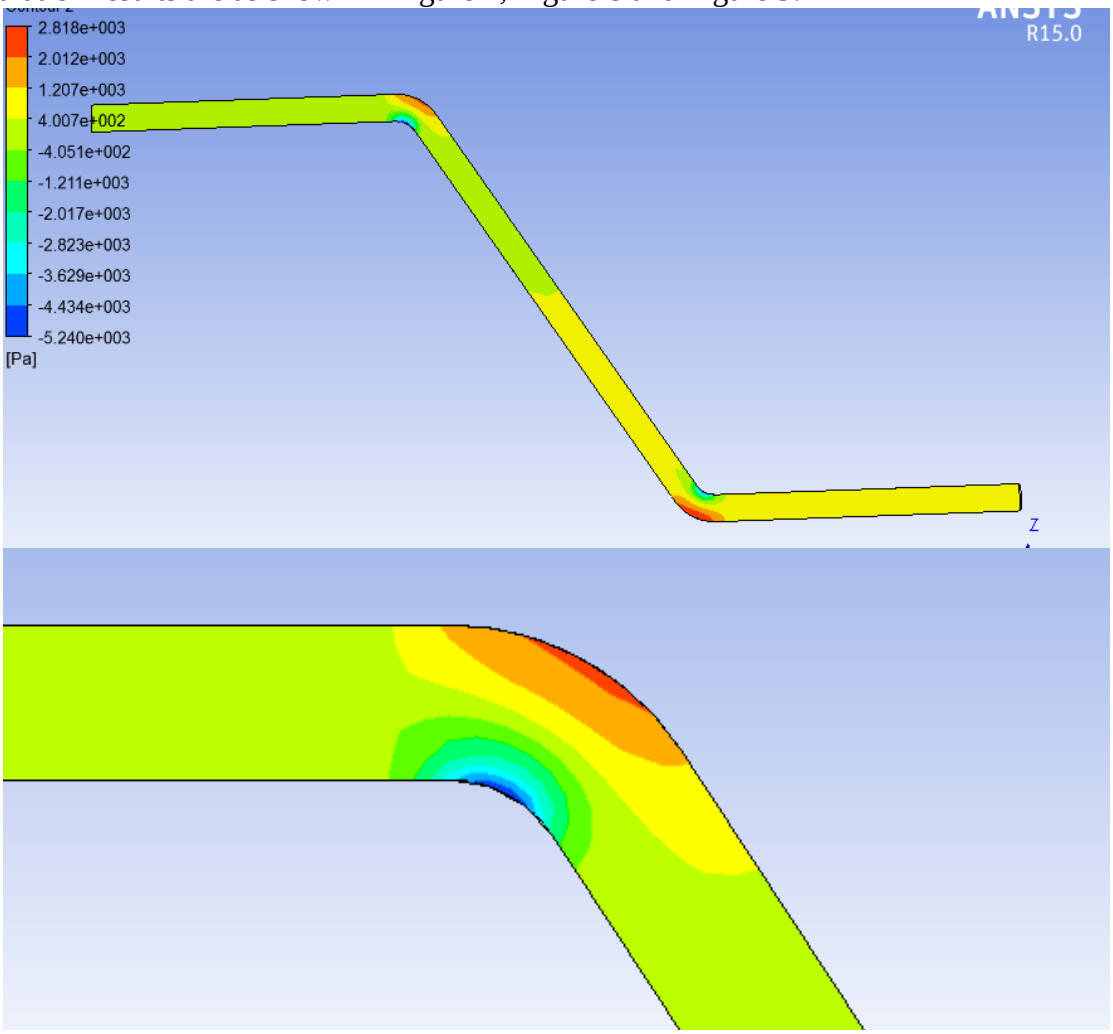
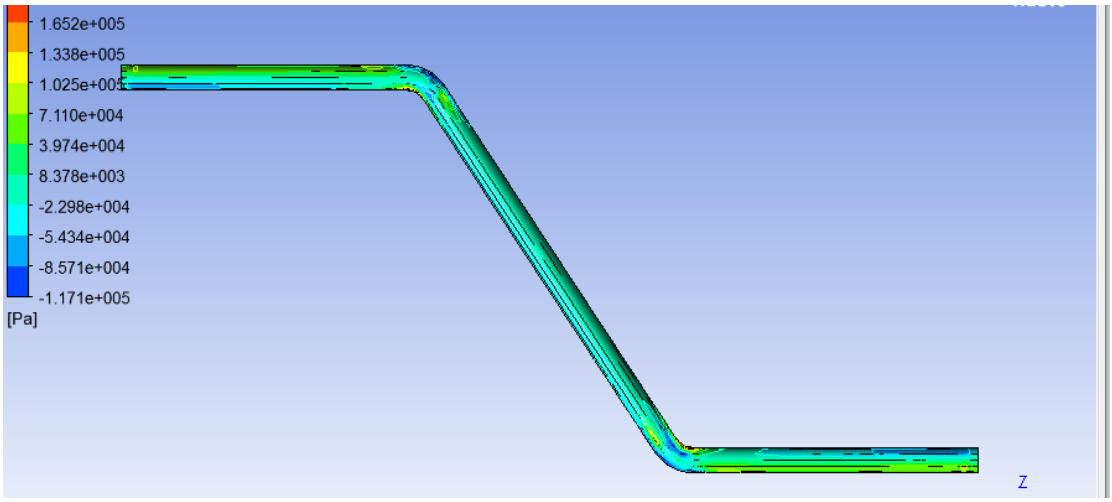


Figure 7. Fluid pressure in the mid-section of the pipeline



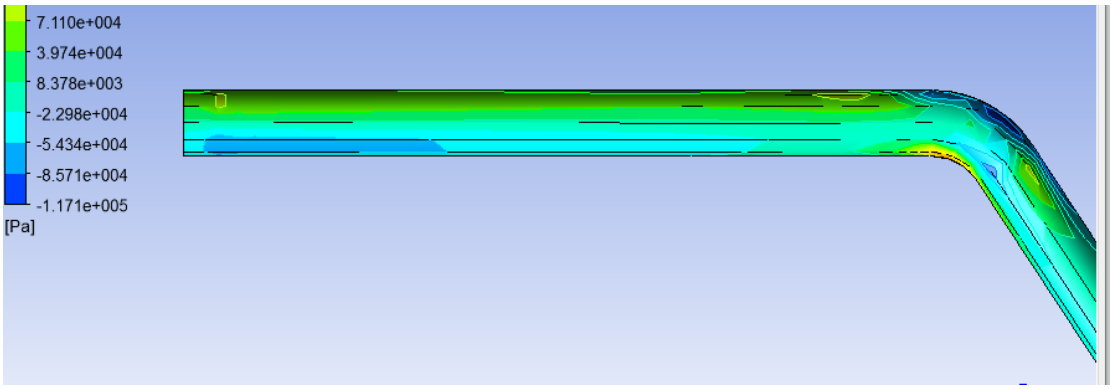


Figure 8. Shear stress in the pipeline wall

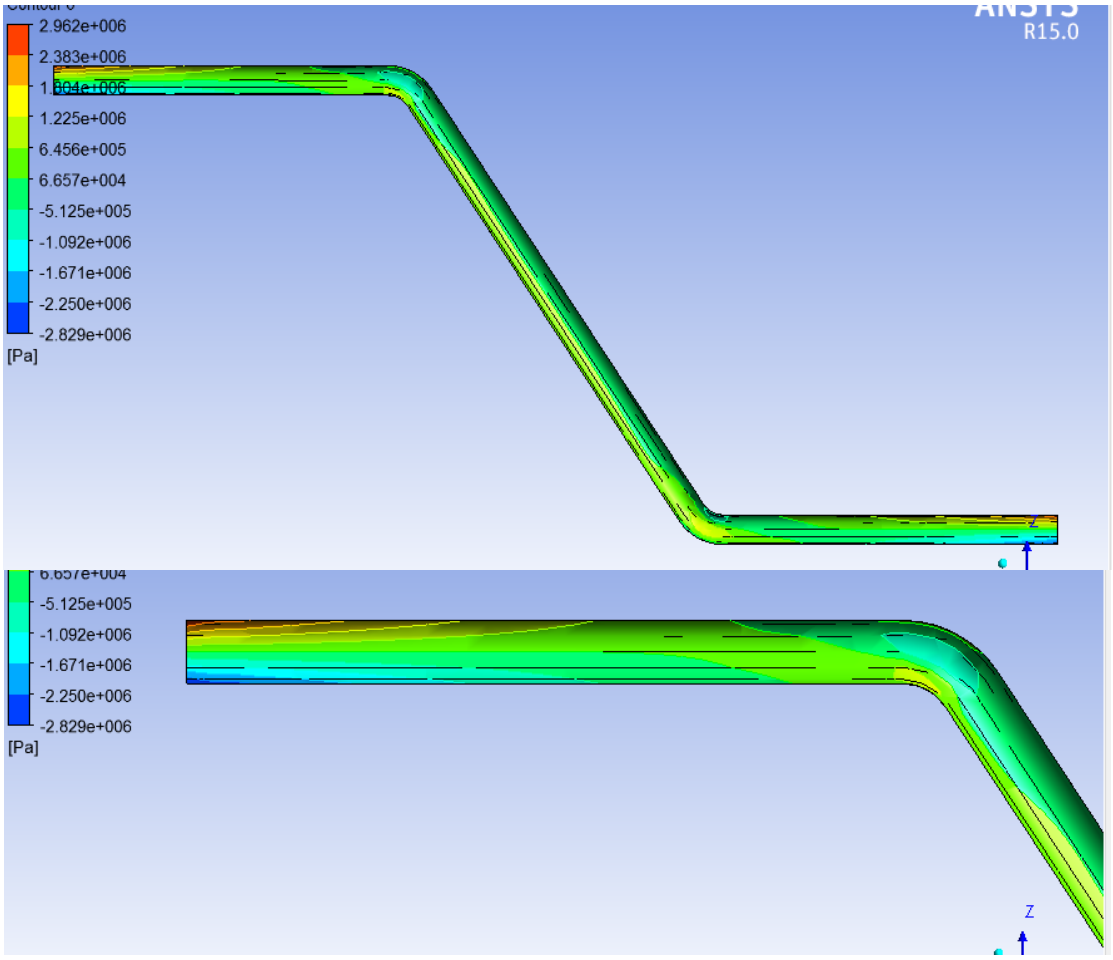


Figure 9. Normal stress in the pipeline wall

6.Conclution

Base on the simulation in Autopipe and Ansys, there are some suggestions:

(1) When the angle of the bend increases, the max stress and strain of the structure will also increase. The stress and strain will be more sensitive to the angle change. Large angle of the corner should be avoid so that the structure would not be damaged.

(2) When the distance between the guide brackets and the fixed ends increases (the guides bracket become closer to the corner), the max stress and strain of the structure decreases. As a result, it would be safer for the pipeline structure. So it's a good choice to shorten the distance between the guide brackets and the corners to decrease the stress and strain to ensure the safety of the pipeline structure.

(3) When the distance between the guide brackets and the fixed ends increases (the guides brackets become closer to the corner), the vibration frequency of each modal also increase. When the frequency of external excitation is close to the resonance area, it's a good idea to adjust the position of the brackets to change the natural frequency to avoid resonance

Of course there are still many deficiencies**:**

(1) The result in this paper is only the simulation results. The comparison with the corresponding experiment results or on-site measured data should be added in further research.
(2) There may be some problems with the results of Fluid and Structure Interaction method. Maybe when using the software, some parameters are not consistent with the reality. Those parameters need to be further corrected.
(3) Further analyses are needed to explain the different results in Autopipe and Ansys.

Reference

[1] M.L.Levitan,K.C.Mehta and W.P.Vann et al. Field measurement of pressure on the Texas Tech Building, Journal of Wind Engineering and Industrial Aerodynamics, 1991, 38(4): 227-234.
[2]Dai Xueliang, Liu Ping. Application and Precautions of Hydraulic Pipeline in Ship. Jiangsu Ship. 2000 (3) :27-29
[3] Liang Chunyu.Study of Vibration of the Ship Engine Fuel Oil Pipeline System. Power Machinery and engineering. Harbin Engineering University.2013.05.
[4] Liu Zhongqun, Sun Yudong, Wu Yousheng. Current Situation and Trends on the Study of Coupled Fluid - Structure Vibration and Sound Propagation of Pipeline Systems. China Ship Scientific Research Center,Wuxi 214082,China. Ship Mechanics,2001.
[5] Chen.S.S, Paidoussis.M.P, Au‐Yang.M.K, et al. Flow‐Induced Vibration of Circular Cylindrical Structures by S.S. Chen, M.P. Paidoussis, and M.K. Au‐Yang[J]. Journal of the Acoustical Society of America, 1984, 75(1):299-299.
[6] Paidoussis.M.P, Issid.N.T, M.Tsui. Parametric Resonance Oscillations of Flexible Slender Cylinders in Harmonically Perturbed Axial Flow. Department of Mechanical Engineering, Journal of Applied Mechanics, 1980.47:709-714.
[7] Paidoussis MP. Flow-induced Instabilities of Cylindrical Structures. ASME. Appl. Mech. Rev. 1987.40(2):163-175. doi:10.1115/1.3149530.
[8] Song Xueguan, Cai Lin, Zhang Hua. Ansys Fluid-Solid Coupling Analysis and Engineering Examples [M]. Beijing: China Water Resources and Hydropower Press, 2012.
[9] Xie Longhan, Zhao Xinyu. Ansys Cfx Fuid Analysis and Aimulation [M]. Beijing: Electronic Industry Press, 2013.
[10] Yu Meng. Analvsis on Characteristics of Fluid-structure Interaction for Fluid Conveying Pipes by ANSYS. China Ship Development and Design Center, Wuhan 430064,China.2007.10.

The Displacement Response Study for Different Graded Parameter Functionally Graded Materials Based on Strip Element Method

[†]J.H. Tian，L.L. Ma

（CAE Analysis Room for Engineering Application ,School of Mechatronic Engineering ,Xi'an Technological University, Xi'an 710021 China）

[†]Correspondence to: J.H .Tian, School of Mechatronic Engineering , Xi'an Technological University, Xi'an 710021 China.

[†]Corresponding author: 540673737@qq.com

Abstract

In this paper, the displacement response of functionally graded material structure under mechanical loading was studied by using strip element method. Established a parameterized model for describing the functional gradient material physical properties, the displacement expression of strip element method was deduced, using the virtual work principle to get the control equations of functionally graded materials plane problem, using the modal superposition method to complete the solution of the equation. The accuracy of the calculation method was verified by comparison with the result of finite element method. The research shows that the results of the strip element method are consistent with the finite element method in the displacement response of the functionally graded materials under mechanical loading. With the increase of the gradient coefficient, the ceramic content is increased, and the material stiffness is also increased, then the structure response caused by external loading is decreased. With the increase of the gradient coefficient, the influence of the gradient coefficient to the structure displacement response is decreased gradually.

Keyword: Displacement Response; Functionally Graded Materials; Strip Element Method.

1 Introduction

Functional gradient material is a kind of polyphase material, which continuously controls the distribution of components in the preparation of materials so as to meet different requirements of different parts of the structure [1][2]. At the same time due to material and components of continuous changes in the structure, there is no macro interface, avoid the traditional composite interface place due to the performance of the mutation and the damage was caused by the occurrence of failure [3]. Functional gradient materials can be designed to change the distribution of material components and bring the development of composite materials to a new level.

In recent years, experts and scholars had conducted in-depth study on the mechanical problems of functionally graded materials. The free vibration of functional gradient materials under mechanical loading and temperature loading was analyzed using Peano-Baker series method by Liu Wuxiang[4]. The dynamic characteristic analytical solution of the quadrilateral and rhombic function gradient plate was obtained by using the beam function group [5]. Reddy [6] adopted the first-order shear deformation theory to study the structural response of the functional gradient plate structure under the thermo-machine coupling. Using the first-order shear deformation theory, Thai [7] had studied the problem of bending and free

vibration of functionally graded materials. Liew [8] studied the vibration response of functionally graded material plates under the coupling of mechanical load and electromagnetic load based on the first-order shear deformation theory using the finite element method. Zenkour [9] used first-order shear deformation theory and classical elastic theory to study the structural bending deformation of the sandwich structure under mechanical loading. Ferreira [10] used three-order shear deformation theory and meshless method to study the structure response of functional gradient materials under static load. Reddy [11] studied the deformation of functional gradient rectangular plate based on the three-order shear deformation theory. Brischetto [12] studied the analytic solution of the structural displacement response of functionally graded material plates under transverse loading. Ray [13] used the finite element method to study the structural response of the functional gradient fiber piezoelectric materials in exponential distribution. Kulikov [14] adopted surface sampling method to solve the problem of three-dimensional thermal stress of functional gradient materials.

Due to physical performance parameters of functionally gradient materials with the coordinate changing continuously, the constant coefficient differential equation turns into the variable coefficient differential equation while the theory of elastic body deformation is derived, it is difficult to solving such problem by analytical method. The numerical method is used to solve complex problem while the large computation is in process. In recent years, the researchers have proposed a semi-numerical and semi-analytic method, that is called the strip element method [15]. It uses the merit of analytical method and numerical method for widely used.

2 Strip element method theory

The strip element method is a series of strip element separated by the solution domain as shown in fig.1. The displacement function is the continuous function of x, and the displacement of the element is obtained through the line displacement interpolation.

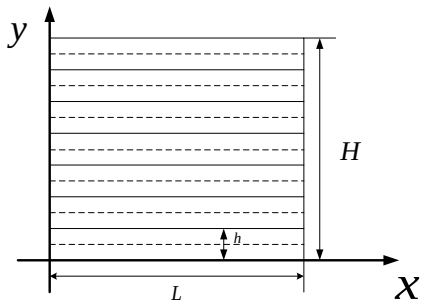


Fig.1 strip element method for solving model

It is assumed that the displacement functions of functionally graded materials under external loads are

$$U = N(y)V(x)\exp(-i\omega t) \tag{1}$$

where $N(y)$ represents shape function for displacement interpolation; $V(x)$ represents displacement function; ω represents circle frequency; t represents time.

The expression of $N(y)$ is

$$N(y)=\left[(1-\frac{3y}{h}+2\frac{y^2}{h^2})\mathrm{I}\quad 4(\frac{y}{h}-4\frac{y^2}{h^2})\mathrm{I}\quad(-\frac{y}{h}+2\frac{y^2}{h^2})\mathrm{I}\right] \tag{2}$$

where h represents element thickness; I represent 2 order element matrix.
According to Kausel 's equation, the equilibrium equation of the system is

$$\rho\ddot{U}-L^T\sigma=0 \tag{3}$$

where U represents displacement vector; L represent differential operator matrix; σ represents stress vector and $\sigma=[\sigma_x\quad\sigma_y\quad\tau_{xy}]^T$; superscript T denotes transpose.

Where the expression for L is

$$L=\begin{bmatrix}\frac{\partial}{\partial x} & 0\\ 0 & \frac{\partial}{\partial y}\\ \frac{\partial}{\partial y} & \frac{\partial}{\partial x}\end{bmatrix} \tag{4}$$

The displacement function is introduced into (3), given

$$W=\rho\ddot{U}-L^T\sigma \tag{5}$$

Since the internal displacement function of the element is interpolated by the nodal displacement, the residual value must be exist in the element, then $W\neq 0$
By the principle of virtual work, there is

$$\delta V(x)^T F=\delta V(x)^T S+\int_0^h \delta U^T W dy \tag{6}$$

Where F represents external mechanical load; S represent nodal stress vector.
Given an external load as

$$F=\overline{F}\exp(-i\omega t) \tag{7}$$

where $\overline{F}$ represents external load amplitude.

The element stress is

$$S^T=[R_x^T\big|_{y=0}\quad R_x^T\Big|_{y=\frac{h}{2}}\quad R_x^T\big|_{y=h}] \tag{8}$$

where stress R_x is

$$R_x=L_x^T cLU \tag{9}$$

where $L_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$; c is the elastic constant matrix $c = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$

The displacement expression is substituted into the virtual work equation and we obtain

$$\overline{F} = -A_2 \frac{\partial^2 V(x)}{\partial x^2} + A_1 \frac{\partial V(x)}{\partial x} + A_0 V(x) - \omega^2 M V(x) \tag{10}$$

where A_0 , A_1 and A_2 represent coefficient matrix. M represents mass matrix.

It can be seen from the equation (10) that the governing equation obtained by element discretization is the ordinary differential equation of x, which reduces the difficulty of the solution, it also reflects the idea of y direction discretization.

The coefficient matrix is

$$A_0 = \frac{1}{3h} \begin{bmatrix} 7D_{xx} & -8D_{xx} & D_{xx} \\ -8D_{xx} & 16D_{xx} & -8D_{xx} \\ D_{xx} & -8D_{xx} & 7D_{xx} \end{bmatrix} \tag{11}$$

$$A_1 = \frac{1}{3} \begin{bmatrix} 3(D_{xy} - D'_{xy}) & -4D_{xy} & D_{xy} \\ 4D_{xy} & 0 & -4D_{xy} \\ -D_{xy} & 4D_{xy} & 3(D_{xy} - D'_{xy}) \end{bmatrix} \tag{12}$$

$$A_2 = \frac{h}{30} \begin{bmatrix} 4D_{yy} & 2D_{yy} & -D_{yy} \\ 2D_{yy} & 16D_{yy} & 2D_{yy} \\ -D_{yy} & 2D_{yy} & 4D_{yy} \end{bmatrix} \tag{13}$$

where

$D_{xx} = \begin{bmatrix} c_{11} & c_{13} \\ c_{13} & c_{33} \end{bmatrix}$; $D_{yy} = \begin{bmatrix} c_{33} & c_{23} \\ c_{23} & c_{22} \end{bmatrix}$; $D_{xy} = \frac{1}{2} \begin{bmatrix} 2c_{13} & c_{33} + c_{12} \\ c_{33} + c_{12} & 2c_{23} \end{bmatrix}$; $D'_{xy} = \begin{bmatrix} c_{13} & c_{12} \\ c_{33} & c_{23} \end{bmatrix}$

The expression for the mass matrix is

$$M = \int_0^h \rho N(y)^T N(y) dy \tag{14}$$

Assemble all the elements in the domain of the problem, and the differential equation is obtained in the whole domain.

$$\overline{F_t} = [-A_{2t} \frac{\partial V_t(x)}{\partial x^2} + A_{1t} \frac{\partial V_t(x)}{\partial x} + A_{0t} V_t(x) - \omega^2 M_t V_t(x)] \tag{15}$$

where A_{0t}, A_{1t} and A_{2t} are the coefficients matrix for the element assembly in the solving domain; M_t is the mass assembly matrix; $\overline{F_t}$ is the external mechanical load; V_t (x) is the displacement function in the solving domain.

Assumed the displacement function form is

$$V_t(x) = d_t \exp(ikx) \tag{16}$$

Assumed external mechanical load form is

$$\overline{F_t} = P_t \exp(ikx) \tag{17}$$

The equation (16), (17) substitution into equation (15), we obtain

$$P_t = [k^2 A_{2t} + ikA_{1t} + A_{0t} - \omega^2 M_t]d_t \tag{18}$$

P_t represents the load acting on the nodal line, while $P_t = 0$ the equation (18) is transformed into a homogeneous equation, and the eigenvalue equation for the *k* is obtained.

$$[k^2 A_{2t} + ikA_{1t} + A_{0t} - \omega^2 M_t]d_t = 0 \tag{19}$$

Equation (19) can be written

$$\left(\begin{bmatrix} 0 & \mathrm{I} \\ \omega^2 M_t - \mathrm{A}_{0t} & -iA_{1t} \end{bmatrix} - k\begin{bmatrix} \mathrm{I} & 0 \\ 0 & \mathrm{A}_{2t} \end{bmatrix}\right)\begin{Bmatrix} d_t \\ kd_t \end{Bmatrix} = 0 \tag{20}$$

The 2M (M=6N-2(N-1)=4N-2) eigenvalue is obtained by solving the equation (20), and the eigenvectors corresponding to the *j*th eigenvalues are represented as ϕ_j

$$\phi_j = \begin{bmatrix} \phi_{1,j} & \phi_{2,j} \cdots \phi_{\mathrm{M},j} \end{bmatrix} \tag{21}$$

By means of the modal superposition method, the solution of the equation can be expressed as

$$V_t = \sum_{j=1}^{2\mathrm{M}} C_j \phi_j \exp(ik_j x) = G(x)C \tag{22}$$

where

$$G(x) = \begin{bmatrix} \phi_{1,1}X_1 & \phi_{1,2}X_2 & \cdots & \phi_{1,2M}X_j \\ \phi_{2,1}X_1 & \phi_{2,2}X_2 & \cdots & \phi_{2,2M}X_j \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M,1}X_1 & \phi_{M,2}X_2 & \cdots & \phi_{M,2M}X_j \end{bmatrix} \tag{23}$$

$$X_j = \exp(ik_j x) \tag{24}$$

Equation (22) is a fundamental solution of the equation. There are 2M constants C in this fundamental solution, so it is necessary to determine the unknown coefficients by the boundary conditions.

Since equation (22) is the fundamental solution in the problem domain, the displacement at any point satisfies the equation (22), and the displacement is satisfied at the right boundary of the solution.

$$V_{bt}^{R} = \sum_{j=1}^{2M} C_j \phi_j \exp(ik_j x^R) = G(x^R)C \tag{25}$$

That can be written

$$G(x^R)C = V_{bt}^{R} \tag{26}$$

The same way, on the left side

$$G(x^L)C = V_{bt}^{L} \tag{27}$$

The equation (26) and (27) are assembled and sorted, and the constant C expression can be obtained

$$C = G_d^{-1} V_{bt}^{T} = \begin{bmatrix} G(x^R) \\ G(x^L) \end{bmatrix}^{-1} V_{bt}^{T} \tag{28}$$

Where V_{bt} is the displacement at the boundary

$$V_{bt} = \begin{bmatrix} V_{bt}^{R} & V_{bt}^{L} \end{bmatrix} \tag{29}$$

By equation (23), $G(x^R)$ can be expressed as

$$G(x^R) = \begin{bmatrix} \phi_{11} X_1^R & \phi_{12} X_2^R & \cdots & \phi_{1L} X_j^R \\ \phi_{21} X_1^R & \phi_{22} X_2^R & \cdots & \phi_{2L} X_j^R \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M1} X_1^R & \phi_{M2} X_2^R & \cdots & \phi_{ML} X_j^R \end{bmatrix} \tag{30}$$

$$X_j^R = \exp(ik_j x^R) \tag{31}$$

The superscript *R* represents the right boundary.

Similarly, the left boundary condition has the same form as the displacement matrix at the right boundary. The difference between the two forms is that the *x* coordinate value at the right boundary is changed to the *x* coordinate value at the left boundary.

$$X_j^L = \exp(ik_j x^L) \tag{32}$$

3 Application of the boundary conditions

The stress vector of the internal element can be expressed as

$$R_x = D_{xx} \frac{\partial U}{\partial x} + D_{xy}^{'} \frac{\partial U}{\partial y} \tag{33}$$

The displacement function is substituted

$$R = R_1 V(x) + R_2 \frac{\partial V(x)}{\partial x} \tag{34}$$

Where R_1 and R_2 are element coefficient matrices

R_1 is

$$R_1 = \frac{1}{h}\begin{bmatrix} D_{yy} & -4D_{yy} & 3D_{yy} \\ -D_{yy} & 0 & D_{yy} \\ -3D_{yy} & 4D_{yy} & -D_{yy} \end{bmatrix} \tag{35}$$

R_2 is

$$R_2 = \begin{bmatrix} D_{yx}' & 0 & 0 \\ 0 & D_{yx}' & 0 \\ 0 & 0 & D_{yx}' \end{bmatrix} \tag{36}$$

Calculate the R_1 and R_2 for all elements, and obtain the overall stress vector R_t as follows

$$R_t = R_{1t} V_t(x) + R_{2t} \frac{\partial V(x)}{\partial x} \tag{37}$$

where :R_{1t}、R_{2t} is the assemble matrix of the element coefficient matrix.

The displacement is simplified by using the left and right boundaries

$$\frac{\partial V(x)}{\partial x} = \sum_{j=1}^{2M} ik_j \exp(ik_j x) \cdot c_j = G' \cdot C = G' \cdot G_d^{-1} V_{bt} \tag{38}$$

The stress vector of the inner nodal line is equal to the average value of the stress vector of the adjacent left and right elements. We obtains

$$R_{bt} = KV_{bt} \tag{39}$$

In the equation, $R_{bt} = \begin{bmatrix} R_{bt}^R & R_{bt}^L \end{bmatrix}$ is an external load acting on the left and right boundaries, where the stiffness matrix K is

$$K = \begin{bmatrix} R_{1t} & 0 \\ 0 & R_{2t} \end{bmatrix} + \begin{bmatrix} R_{2t} G^{R\prime} G_d^{-1} \\ R_{2t} G^{L\prime} G_d^{-1} \end{bmatrix} \tag{40}$$

In equation (40), $G^{R\prime}$ and $G^{L\prime}$ have the same expressions as G^R and G^L. The difference is that G^R and G^L in X_j^R and X_j^L are replaced by $X_j^{R\prime} = ik_j \exp(ik_j x^R)$ and

$X_j^{L'} = ik_j \exp(ik_j x^L)$ respectively.

Equation (39) reflects the relationship between the stress boundary condition and displacement, the stress and displacement boundary conditions are transformed into displacement boundary conditions by equation (39), a constant *c* can be obtained by substituting the boundary displacement into equation (28), from equation (22) and (1) the displacement can be obtained in the solving problem.

4 Establishment of gradient parameter model

The functional gradient materials studied in this paper are composited by metal and ceramic materials. The metal material is $1Cr_{18}Ni_{11}Nb$ and the ceramic material is Si_3N_4. The physical properties of the two materials are shown in table 1.

Tab.1 physical properties of $1Cr_{18}Ni_{11}Nb$ and Si_3N_4 materials

material	Modulus of elasticity / MPa	Poisson's ratio
$1Cr_{18}Ni_{11}Nb$	2.38×10^5	0.3177
Si_3N_4	3.22×10^5	0.2400

Because functionally graded materials are continuously changed by different materials according to the design requirements, the physical performance parameters of the materials are expressed as a function of the volume fraction of the material, the physical properties of the materials and the content of components, it denotes as

$$\begin{aligned} E(y) &= (E_C - E_M)Q_C + E_M \\ \nu(y) &= (\nu_C - \nu_M)Q_C + \nu_M \\ \rho(y) &= (\rho_C - \rho_M)Q_C + \rho_M \end{aligned} \tag{41}$$

where $E(y)$, $\nu(y)$ and $\rho(y)$ are the elastic modulus, Poisson's ratio and density of FGM plates, respectively, E_C, ν_C and ρ_C are the elastic modulus, Poisson's ratio and density of ceramic materials, respectively. E_M, ν_M, ρ_M are the elastic modulus, Poisson's ratio and density of metal materials, respectively. Q_C is the volume fraction of the ceramic.

To ensure the continuity of the material, the sum of the metal volume fraction and the ceramic volume fraction at any location of the material is 1.

$$Q_C + Q_M = 1 \tag{42}$$

where, Q_M is the volume fraction of metals.

The volume fraction change function of the functionally graded metal is

$$Q_{\mathrm{M}} = (1 - a(1 - \frac{y}{H}) + b(1 - \frac{y}{H})^{c})^{p} \tag{43}$$

where y is the vertical position and $0 \le y \le H$; H is FGM plate thickness; $p \ge 0$ is the gradient parameters.

By equation (43), it can be seen that when p=0, Q_{M}=1, and Q_{C}=0. Substituting that into (41) we have $E(y) = E_{\mathrm{M}}$, $\nu(y) = \nu_{\mathrm{M}}, \rho(y) = \rho_{\mathrm{M}}$ and functionally graded material degenerates into pure metal homogeneous material, When p tends to infinity, the Q_{C} tends to 1, and the equation (41) can be obtained $E(y) = E_{\mathrm{C}}, \nu(y) = \nu_{\mathrm{C}}, \rho(y) = \rho_{\mathrm{C}}$, and the material is reduced to a pure ceramic homogeneous material.

When a=1 and b=0 are functionally graded materials, the volume fraction of metals varies as shown in Figure 2

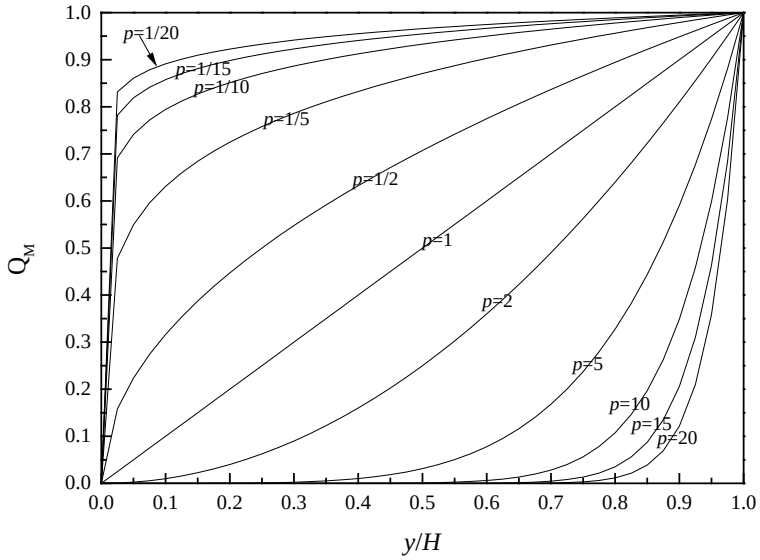


Fig. 2 law of change of volume fraction of metal with thickness under different gradient parameters at a=1 and b=0

The figure 2 can be shown, when y = 0, metal volume fraction is 0, this position is pure ceramic material, when $y = H$, metal volume fraction is equal to 1, this position is the pure metal homogeneous material. So when a = 1, b = 0, the transition form of materials is a continuous transition of ceramic to metal. When the gradient parameter p = 1, the change of the volume fraction of the functionally graded material is continuous linear change. When the gradient parameter is lesser than 1, the change of the metal volume fraction decreases. When the gradient parameter is greater than 1, the metal volume fraction of the functional gradient material is slower in the early, then the volume fraction is accelerated with the thickness increased.

5 computational model

5.1 Validation of strip element method

The model of figure 3 is solved by using the strip element method. In figure 3, L = 100mm, H

= 100mm, the lower boundary is fixed, q = 100N/mm uniform load is applied on the upper boundary. The solution domain is divided into 20 elements and the material properties in each element are metal material. The metallic material property is shown in table 1. The displacement at y = 10mm, y = 50mm, and y = 90mm is shown in figure 4 by strip element method and finite element method.

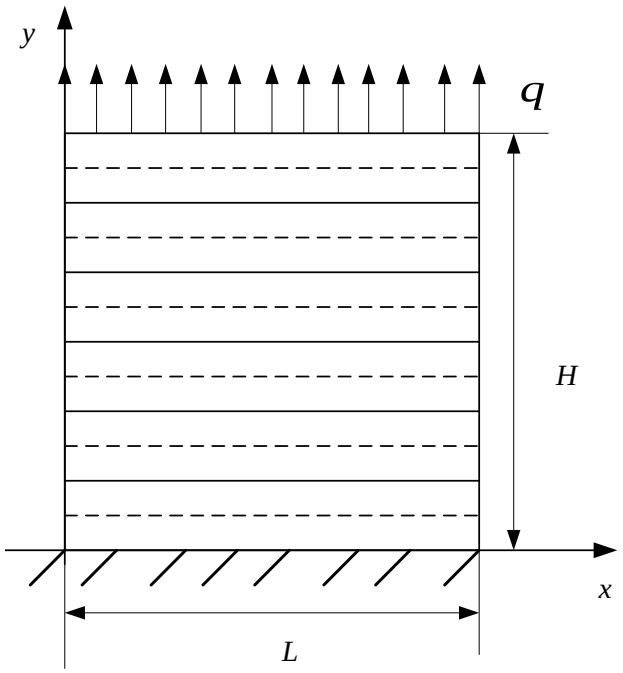


Fig. 3 physical model of plane problem

From fig.4 we can see that the displacement obtained by strip element method and finite element method in *y*-direction at different position is almost the same. From y = 10mm, y = 50mm, to y = 90mm, the displacement in *y*-direction is increased. The accuracy of the strip element method is verified.

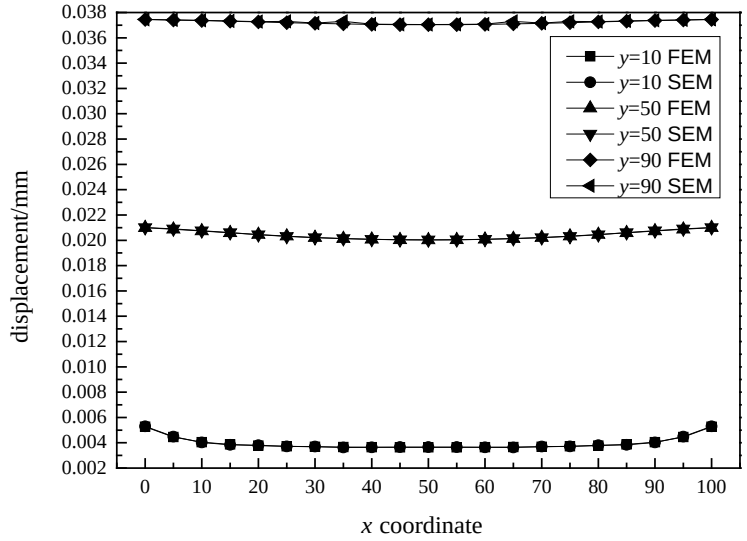


Fig. 4 the displacement in *y*-direction in different position by FEM and SEM

5.2 The influence of element number for the results

In the calculation process, because of the discrete of the solving domain, the size of the element has an influence on the accuracy of the displacement solution. Figure 5 and figure 6 are the *x*-direction displacement and the *y*-direction displacement at *y*=10mm using 10 elements, 20 elements and 50 elements, respectively. It can be known that displacement results have small difference when the element number is 10 and 20, while when the element number is 20 and 50, the result difference is not obvious, so in this calculation the element number is set to 20 for ensuring that precision and calculation speed.

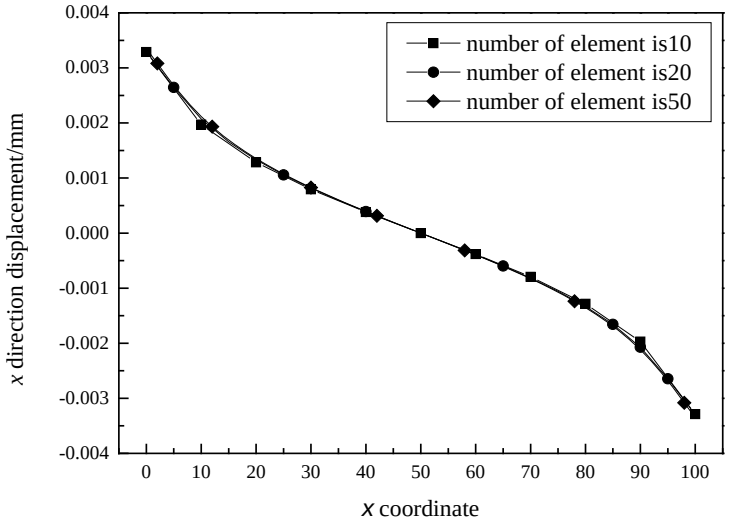


Fig. 5 calculation results of *x* direction displacement at *y*=10mm

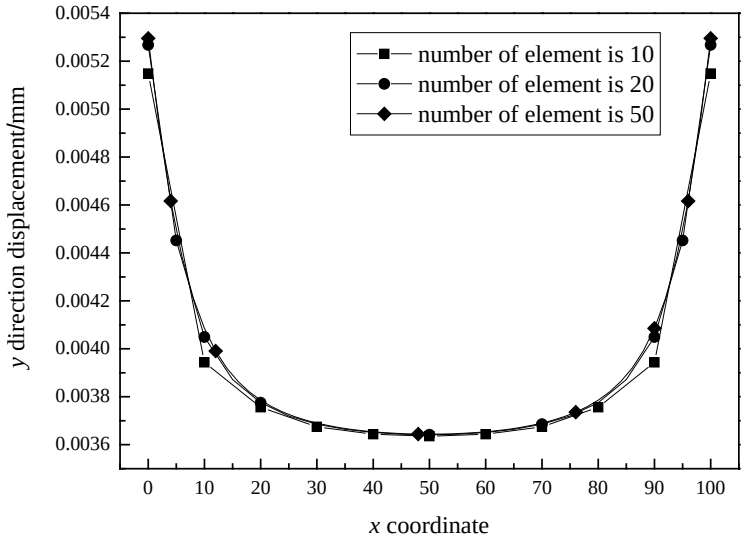


Fig. 6 calculation results of *y* direction displacement at *y*=10mm

5.3 Structural response analysis of functionally graded materials

The displacements with different gradient parameters are shown in figure 7 at y = 10 mm in the *x* direction. It can be seen from figure 7, with the increase of gradient coefficient, the

displacement response decreases, and this is due to metal volume fraction decreases in the functionally graded materials, ceramic volume fraction increases, causing the structure stiffness increased. In addition, the displacement in the *x* direction at *y* = 10mm has nonlinear characteristics.

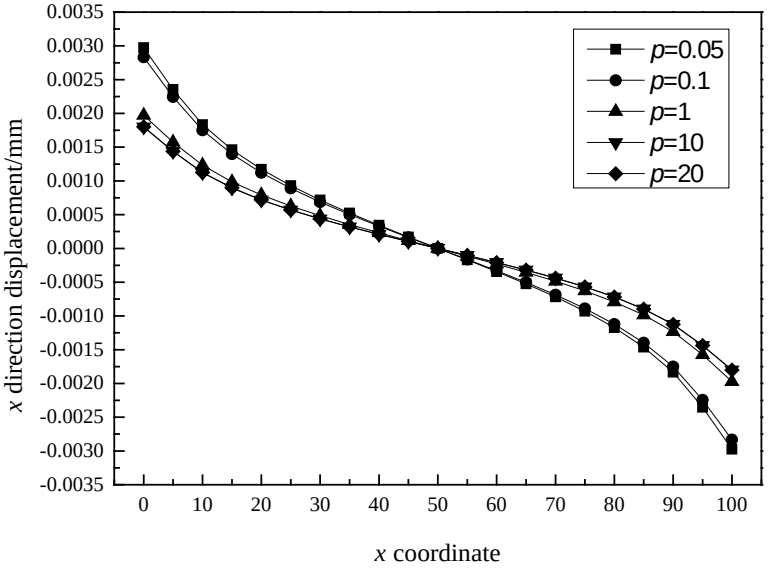


Fig. 7 *x* direction displacement of different gradient parameters at *y*=10mm

The displacements with different gradient parameters are shown in figure 8 at *y* = 10 mm in the *x* direction. It can be seen from figure 8, the displacement in *y* direction at *y* = 10mm with the increase of *x* coordinate, has the characteristics of first increases, then decreases, then increases. As the gradient parameter increases, the displacement of the same position decreases. As the gradient coefficient increases, the influence of the gradient coefficient for the displacement response decreases gradually.

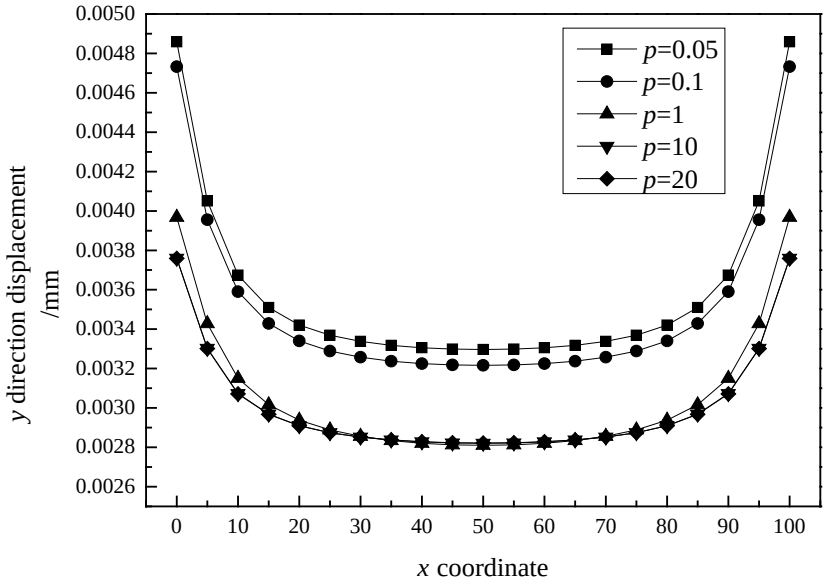


Fig. 8 y directional displacement of different gradient parameters at *y*=10mm

x direction of displacement as shown in figure 9 and 10 at *y* = 50 mm and *y* = 90mm. They has the same law at *y* = 50 mm and *y* = 90mm. The displacement absolute value of *x* direction with respect to *x* = 50mm is symmetrical, and with the increase of gradient coefficient, *x* direction displacement is reduced, this is due to the increase of the gradient coefficient, the ceramics content of the material keeps increasing causing to the structure stiffness increased.

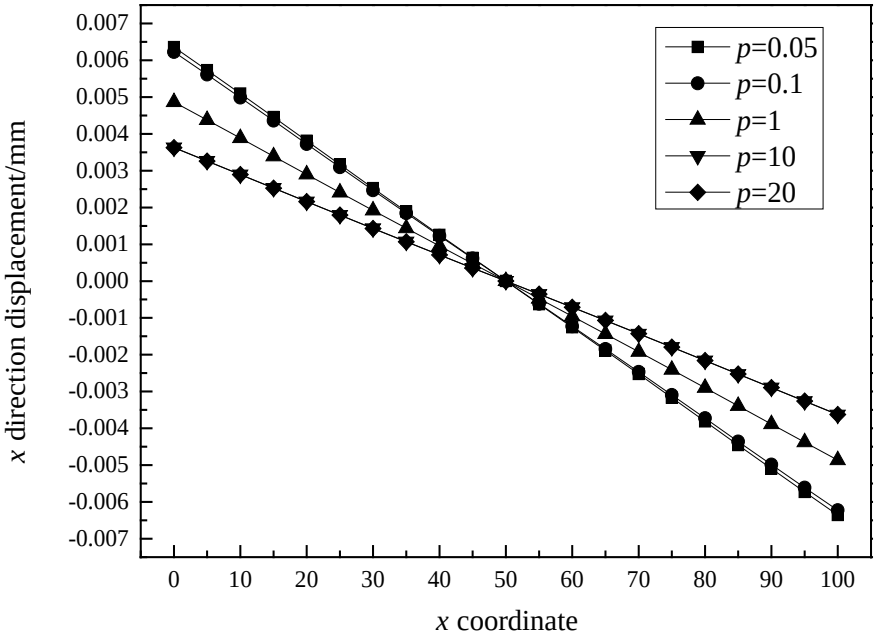


Fig. 9 *x* directional displacement of different gradient parameters at *y*=50mm

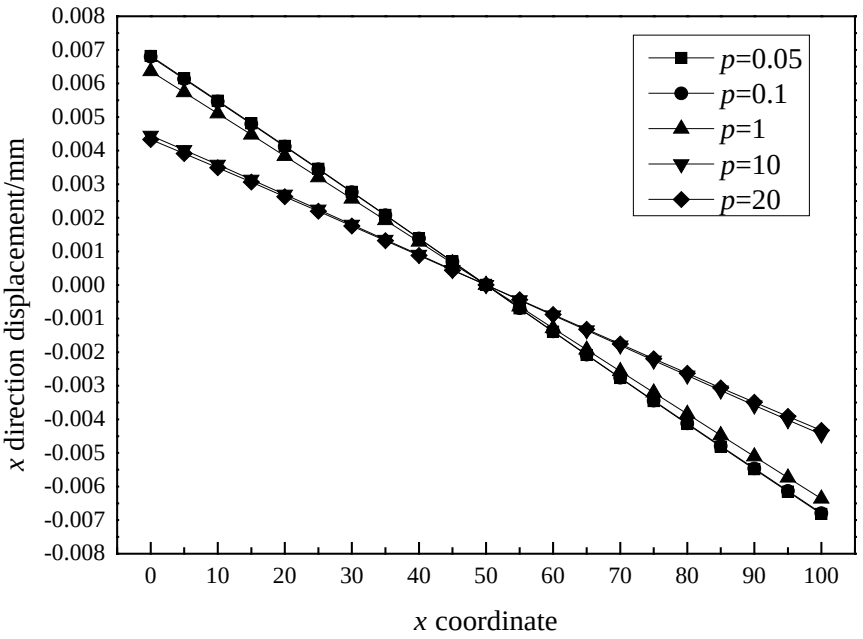


Figure 10 *x* direction displacement for different gradient parameters at *y*=90mm

y direction of displacement as shown in figure 11 and 12 at *y* = 50 mm and *y* = 90mm. It can be seen that the displacement in *y* direction has class parabolic distribution, and as the change

of spatial coordinates, the change curve approximation to straight line. With the increase of gradient coefficient, structural stiffness is gradually increased, the displacement response under external loading is reduced.

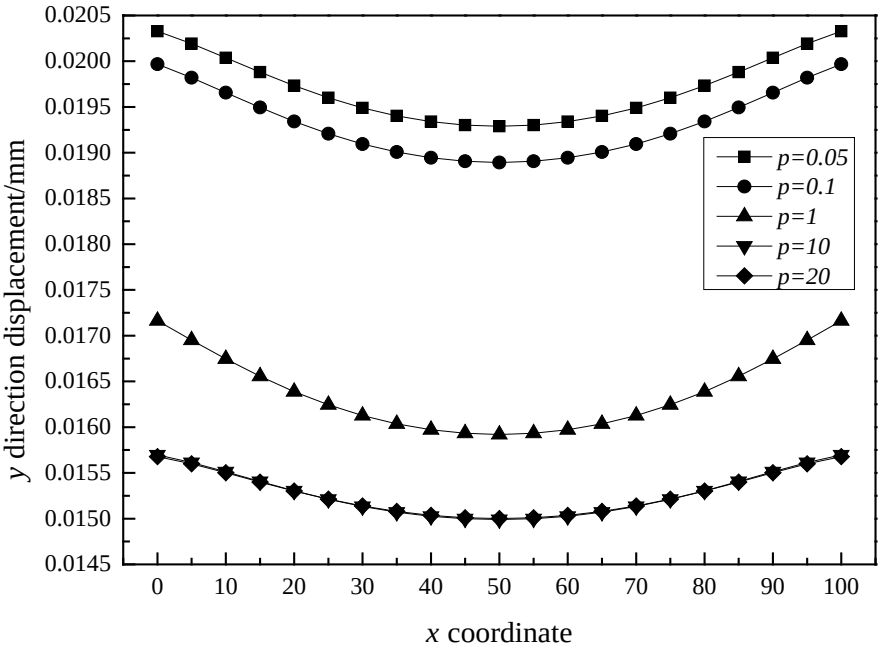


Fig. 11 *y* direction displacement of different gradient parameters at *y*=50mm

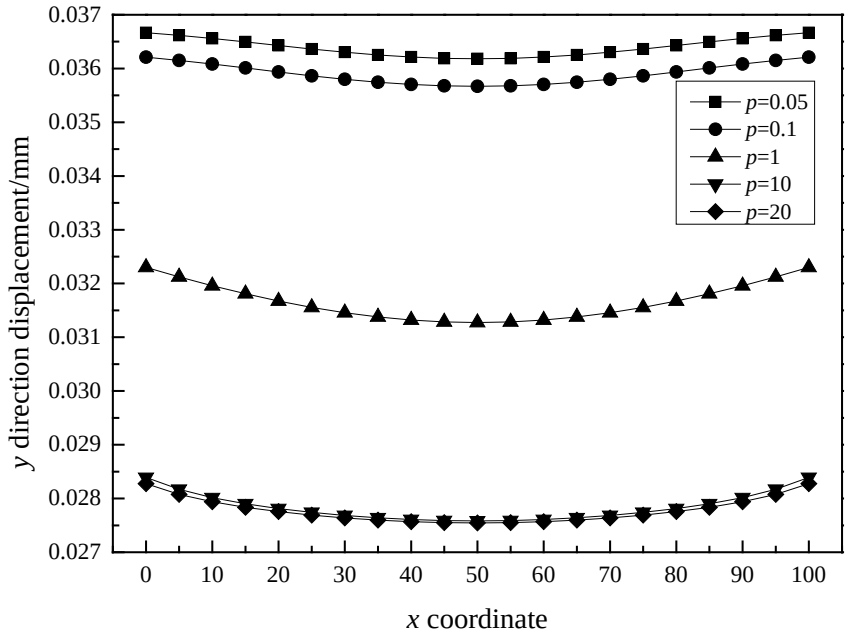


Figure 12 *x* direction displacement for different gradient parameters at *y*=90mm

6 Conclusion

The parameters model of the functional gradient materials is presented in this paper. The displacement response of considering different parameterized model is investigated in

mechanics loading using the strip element method, some useful results have been obtained.

1) The displacement response result is almost the same by using two methods, so to verify the validation of the strip element method.

2) With the ceramic volume fraction increases, the metal volume fraction decreases in the functionally graded materials, the structure stiffness is increasing, displacement response decreases.

3) As the gradient coefficient increases, the influence of the gradient coefficient on the structure displacement response decreases gradually.

Acknowledgements

The present paper is financially supported by The National Natural Science Foundation of China (11302159), Young Talent fund of University Association for Science and Technology in Shaanxi , China(No. 20170519), and the Principal Foundation of Xi'an Technological University(No.XAGDXJJ16009).

Reference

[1] Birman V. Modeling and Analysis of Functionally Graded Materials and Structures[J]. Applied Mechanics Reviews, 2007, 60(5): 195-216.

[2] Zhong Zhen, Wu Linzhi, Chen Weiqiu. Progress in the Study on Mechanics Problems of Functionally Graded Materials and Structures[J]. Advances in Mechanics, 2010, 40(5): 528-541.

[3] Cao Zhiyuan. Unified Expression of Natural Frequency Solutions for Functionally Graded Composite Rectangular Plates Under Various Boundary Conditions[J]. Acta Materiae Compositae Sinica, 2005, 22(5): 172-177.

[4] Liu Wuxiang. Two-dimensional Thermoelasticity Vibration Analysis of Functionally Graded Plate for Arbitrary Graded Distribution[J]. Journal of Jiangsu University(Natural Science Edition), 2014, 35(3): 284-289.

[5] CAO Zhiyuan. Analytical Solution of Dynamic Character for Functionally Graded Material Parallelogrammic Plates with Various Boundary Conditions[J]. Chinese Quarterly of Mechanics, 2006, 27(2): 255-261.

[6] Reddy J N. Analysis of Functionally Graded Plates[J]. International Journal for Numerical Methods in Engineering, 2010, 47(47): 663-684.

[7] Thai H T, Choi D H. A Simple First-Order Shear Deformation Theory for the Bending and Free Vibration Analysis of Functionally Graded Plates[J]. Composite Structures, 2013, 101(15): 332-340.

[8] Liew K M, He X Q, Kitipornchai S. Finite Element Method for the Feedback Control of FGM shells in the Frequency Domain Via Piezoelectric Sensors and Actuators[J]. Computer Methods in Applied Mechanics & Engineering, 2004, 193(3-5): 257-273.

[9] Zenkour A M. A Comprehensive Analysis of Functionally Graded Sandwich Plates: Part 1-Deflection and Stresses[J]. International Journal of Solids & Structures, 2005, 42(18): 5224-5242.

[10] Ferreira A J M, Batra R C, Roque C M C, et al. Static Analysis of Functionally Graded Plates Using Third-Order Shear Deformation Theory and a Meshless Method[J]. Composite Structures, 2005, 69(4): 449-457.

[11] Reddy J N, Chin C D. Thermomechanical Analysis of Functionally Graded Cylinders And Plates[J]. Journal of Thermal Stresses, 1998, 21(6): 593-626.

[12] Brischetto S, Carrera E. Advanced Mixed Theories for Bending Analysis of Functionally Graded Plates[J]. Computers & Structures, 2010, 88(23-24): 1474-1483.

[13] Ray M C, Sachade H M. Finite Element Analysis of Smart Functionally Graded Plates[J]. International Journal of Solids & Structures, 2006, 43(18-19): 5468-5484.

[14] Kulikov G M, Plotnikova S V. A Sampling Surfaces Method and its Implementation for 3D Thermal Stress Analysis of Functionally Graded Plates[J]. Composite Structures, 2015, 120(10): 315-325.

[15] Liu G R, Achenbach J D. A Strip Element Method for Stress Analysis of Anisotropic Linearly Elastic Solids[J]. Journal of Applied Mechanics, 1994, 61(2): 270-277.

[16] Tian Jianhui, Han Xu , Sun Xiaowei. Transient Thermal Response of Functionally Graded Material Plates Based on the Hybrid Numerical Method[J]. Chinese Journal of Solid Mechanics, 2008, 29(4): 396-401.

Genetic Algorithms In Electoral Districting

†*Maw-kae Hor[1,2], Yi-Fan Peng[2,3], Ying-Che Hung[4], and Cheng-Yuan Tang[4]

[1]School of Informatics, Kainan University, Taiwan.
[2]Department of Computer Science, National Chengchi University, Taiwan
[3] Center of GIS, RCHSS, Academia Sinica, Taiwan
[4] Department of Information Management, Huafan University, Taiwan

*Presenting author: mhor@mail.knu.edu.tw
†Corresponding author: mhor@mail.knu.edu.tw

Abstract

In this paper, we present a mechanism that uses the concepts of genetic algorithms for electoral districting. Principles of genetic algorithms are adopted in finding the better districting solutions. The geometric information as well as the attribute data provided through the geographic information systems are used in selection, crossover, or mutation emulations. We also design various quantitative measurements to evaluation the performance (fitness) of our mechanism. We have applied this mechanism to CEC's result of Taipei City and the results show that our mechanism works.

Keywords: electoral districting, genetic algorithm, geographic information system

1. Introduction

Electoral districting is one of the traditional problems in political science for centuries. It may influence the election results directly. Electoral redistricting is normally required due to electoral regulation changes or due to new census results.

One of the crucial problems in electoral districting is to find a fair districting result efficiently. Traditionally, electoral districting was done manually and the process was time consuming and required huge amount of resources. Most important, manual districting normally introduce controversies among political parties.

In this paper, we present a mechanism that can be used to improve an existing electoral districting result. One can choose the electoral districting result announced by the Central Election Commission or developed in our early results. And then apply the principles of genetic algorithm to search for a new and better electoral districting result. In the searching processes, the concept of selection, crossover, and mutation operators are emulated for finding the better solutions. The geometric information as well as the attribute data provided through the geographic information systems are used in these operators. The fitness function or the performance evaluation function can also be defined using these data.

In order to obtain a fair electoral districting result, the election outcome is not considered in developing the entire mechanism. Only the general districting principles are included when making the decisions. However, one can evaluate the preference of different electoral districting results by the election outcomes afterward and chose a favorite one. The following three general principles are considered in our mechanism: (1) population equality: the populations in every district are close enough; (2) region contiguity: the units in the same district must be connected; and (3) shape compactness: the shape of each district shall be good.

2. Related Works

One of the most critical issues in solving the electoral districting problem is to find a fair method in order to avoid the Gerrymandering phenomenon [1] and satisfy the preset requirements.

Figure 1. Gerrymandering phenomenon

Electoral districting is one of the most important problems in political sciences. Due to a recent change of local legislative election regulations [2], many researches have been reported recently. Hsieh [9] and Pan [12] discussed the mechanism and standards. Yu [13] presented more complete discussions on the regulation changes as well as the underline legal issues. Others gave simulated studies in predicting the election outcomes [10]. Rectangle method by Harris [4] and linear programming method by Hess [6], modified by Helbig [5] are the early districting methods that we can trace. Kaiser's [11] model in evaluating population equality probably is among the earliest papers used computers to simulate electoral districting. Others may deal with finding a solution from a particular initial configuration but did not explore too many choices.

We have proposed various mechanisms in solving the electoral districting problems in the last decade [7][8]. The first method we proposed is the bricklayer method that uses only vertical and horizontal lines in the districting processes. Then we developed the two-partitioning method that uses the knowledge of computational geometry as well as the concepts adopted from dynamic programming. A weighted Voronoi diagram mechanism is also proposed in finding an initial districting result that reduced the post-processing time. However, there are two major issues have to be answered in solving the electoral districting problems. The first one is getting a solution that satisfies all the preset requirements. The second one is can we improve a given districting result without violating the preset requirements.

3. Methods

In this paper, we use the principles of evolutionary algorithms in finding better solutions for a given districting result. Evolutionary algorithms adopt the concepts of keeping the better members among the new generations. This can normally be described in the following diagram.

In general, there are three operators can be used in the evolutionary process, namely, reproduction, crossover, and mutation. Evolving through these operators may generate huge amount of members in the next generations. Only the better or the best members will be survived through the life competitions. One normally mimic these competitions using a

selection mechanism in cooperate with some evaluation functions. The members receive the higher scores will be selected and kept.

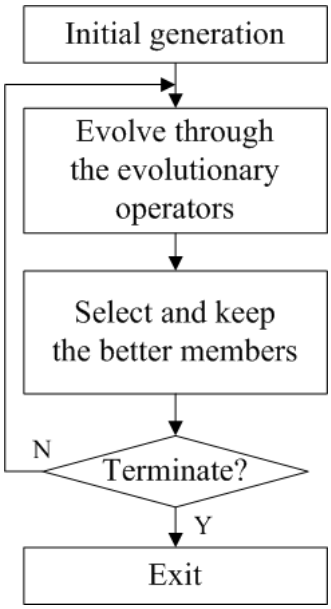


Figure 2. Concept of the evolution flow.

However, in some applications, producing huge amount of next generations and select the better members may not be practical due to the limited computation resources. One can search for the better members among the possible new generations before the actual evolving process take places. This can prevent unnecessary producing the dummy generations or members that will not be kept in any future references. We use a modified flow chart in our studies under this assumption. This is the actual evolutionary flow diagram as depicted in Figure. 2. In this diagram, the selection is implemented using heuristic searches so that the algorithm always selects the better results to be evolved. Researchers may argue that whether one shall check for better solutions or not before the actual evolving process. Nevertheless, there are two aspects must be considered in designing a system evolving through genetic algorithms, namely, how to evolve into the next generation and how to evaluate the result is better or not. The first aspect is generally accomplished through the combination of selection and actual evolving. The second aspect is generally practiced through the comparison of the fitness functions. We will discuss these issues in the following sections.

3.1 The Selection Considerations

Selecting the target for evolving may or may not include the evaluation issues before and after the evolutions. In the contrast, the evaluation process generally does not consider the selection issues in the evaluating mechanisms. If the evaluation issues are not included in the selection considerations, random search is generally used in most of the similar applications. However, random search suffers slow convergent speed and is not suitable in our studies. We shall integrate heuristic search into the evolutionary processes in our studies.

Recall that genetic algorithms normally use three operations, namely, reproduction, crossover, and mutation, for evolving into the next generation. We use the general term “selection” to represent these operations in electoral districting since they all involve selecting a target village to be evolved. In the evolving process of electoral districting, one can simply select a target village, from another district, to be merged into an existing district or can also consider find a village, in another district, to be exchanged with a village currently in an existing

district. We shall refer the first strategy as “simple selection” and the second strategy as “trading”.

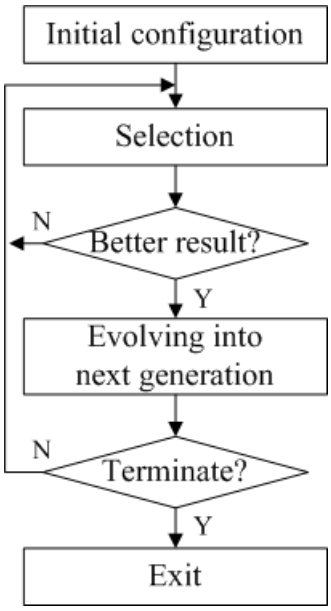


Figure 3. Evolution flow diagram.

Note that a simple selection is an operation of selecting a village v from a district D_j and merging it into another district D_i. If we use superscripts to indicate the time stamps and subscripts to indicate the districts, then we can write the districts after the selection operation, D^{t+1}, in terms of the districts before the selection operation, D^t, as follows:

$$\begin{aligned} D_i^{t+1} &= D_i^t + \{v\} \\ D_j^{t+1} &= D_j^t - \{v\} \end{aligned} \tag{1}$$

where the operator “+” and “-“ indicate “union” and “subtraction” of the set operations, respectively. Similarly, the trading operation can be written as

$$\begin{aligned} D_i^{t+1} &= D_i^t - \{v_i\} + \{v_j\} \\ D_j^{t+1} &= D_j^t + \{v_i\} - \{v_j\} \end{aligned} \tag{2}$$

where $v_i \in D_i^t$ and $v_j \in D_j^t$.

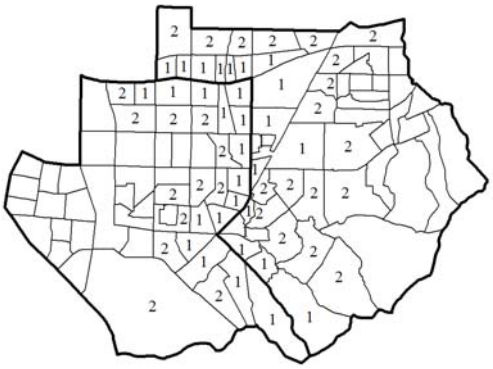


Figure 4. Districting diagram

Either simple selection or trading, the principle of population equality and region contiguity must be satisfied. In our system, we do want to select a better target for evolving into the next generation. Hence, both strategies were considered.

Refer to Figure 4, the districting diagram. There are two adjacent districts in this figure. The villages on the boarder are labelled "1" and the villages next to them are labelled "2". (These labels indicate that the village is either 1 village away from the boarder or 2 villages away from the boarder, etc.) In our studies, crossover is considered as a trading between the "1-villages" that belong to different districts and mutation is considered as a trading when a district wants a "2-village" from another district and the associated connecting "1-village(s)" must be traded together with this "2-village".

Algorithms for selection

A naïve algorithm can use only the distance function to select the target to be evolved. However, we are searching for a better target for evolving in the sense of "easy" producing a good next generation. Hence, in addition to the distance function, we also consider the population issue and geometry issues when searching for a better village in the selection processes. The following function, f_s, is used in selecting a target village v and add it to a particular district D:

$$f_s = \alpha\, f_d(D,v) + \beta\, f_P(D,v) \tag{3}$$

where $f_d(D,v)$ is the distance evaluation function of a village v and a district D and $f_P(D,v)$ is the population evaluation function, after v is added to D. Note that α and β are the weighting coefficients satisfying $\alpha + \beta = 1$.

3.2 The Evaluation Principles

Three evaluation functions are considered in our studies. They reflect the issued related to distance, population, as well as geometry aspects and shall be discussed in this section.

Distance evaluation function

In the simple selection or trading operation considerations, as the distance between D and v increases, v is less likely to be selected. Thus $f_d(D,v)$ is inverse proportion to the distance, $d(D,v)$, between them.

Population evaluation function

For a given district, D_i, the population evaluation function of it, $f_P(D_i)$, reflects the population deviation of D_i to the average district population, P_a. We can write the related evaluation function as

$$f_P(D_i) = \frac{\left|P_a - P(D_i)\right|}{P_a} \quad \text{or} \quad f_P(D_i) = \left|1 - P(D_i)/P_a\right|. \tag{4}$$

And the total population evaluation function, $f_P(D)$, of a particular district method for the given county, D, can be expressed as the sum of the individual functions. Hence, we have

$$f_P(D) = \sum_i f_P(D_i). \tag{5}$$

4. Results

We use Taipei City as an example in illustrating our mechanism. Taipei city has 12 counties consists of 488 villages that will be districting into eight districts. The population of Taipei

city is around 2.6 million, hence, the average population of each district is 0.32 million. For simplicity, we choose the initial configuration to be the following diagram, Figure 5, where nine counties are districted into eight district and 3 counties remain undissolved.

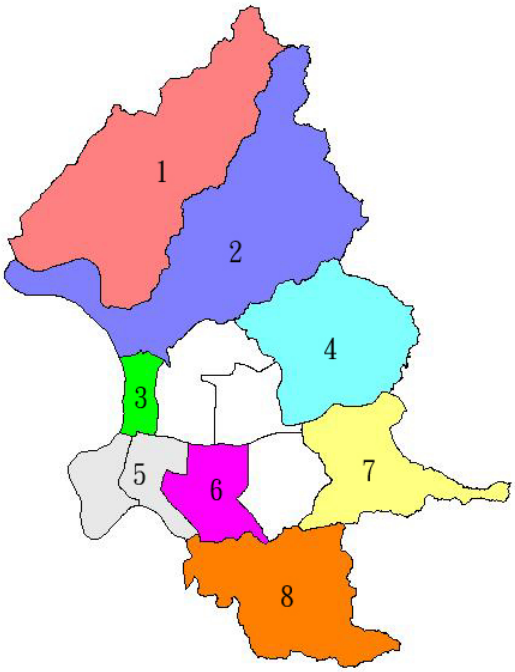


Figure 5. Taipei city, initial districting configuration

After applying our method, one can see that, for instance, some villages in district 4 have been "immigrated" to district 2, whereas some villages in district 2 have been immigrated to district 1. Figure 6 shows the districted results and the populations are listed in Table 1. As indicated in Table 1, one can see that the population difference is all within the preset requirement, 15%.

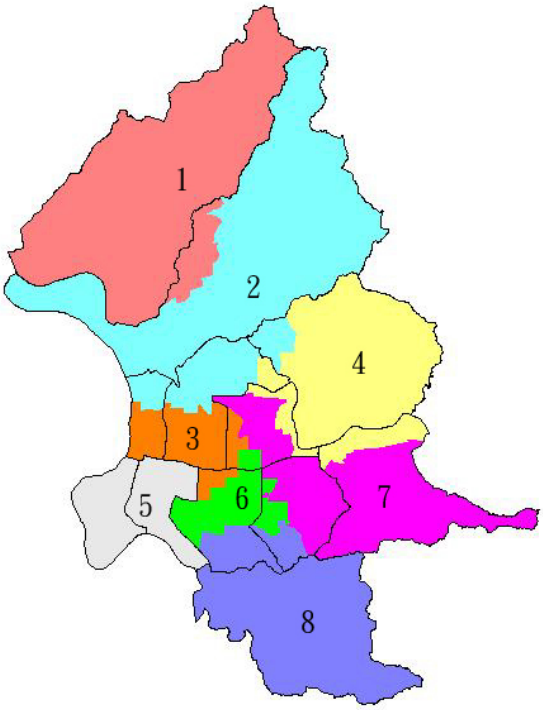


Figure 6. Taipei city, final districting configuration

Table 1. Taipei city, populations for the final districting configuration

District ID	Including counties	District population	Population error (%)
1	Baitou, Shilin	326,240	-0.7
2	Shilin, Zhongshan, Datong, Neihu	324,135	-1.3
3	Datong, Zhongshan, Daan, Songshan	326,004	-0.7
4	Neihu, Nangang, Zhongshan, Songshan	323,369	-1.5
5	Zhongchen, Wanhua	359,865	9.6
6	Daan, Xinyi, Songshan	323,538	-1.5
7	Nangang, Xinyi, Songshan	322,139	-1.9
8	Wenshan, Daan, Xinyi	321,848	-2.0

Conclusion

We have proposed a genetic algorithm that uses the concept of evolution for solving the electoral districting problem. The mechanism has been applied to district the Taipei City with successful results.

References

[1] http://en.wikipedia.org/
[2] Central Election Commission, *Electoral Districting Rules for the Seventh Legislator Elections*. Taipei: Central Election Commission, 2005. (in Chinese).
[3] Corman, T. H., Charless, E.L., Ronald, L.R., and Clifford, S., "Introduction to Algorithms." Second Edition. MIT Press, 2001.
[4] Harris Jr., Curtis C, "A Scientific Method of Districting." *Behavioral Science* Vol 9, No. 3, pp. 219-225, 1964.
[5] Helbig, Robert E., Patrick K. Orr, and Robert R. Roediger, "Political Redistricting by Computer." *Communications of the ACM*, Vol 15, No. 8, pp. 735-741, 1972.
[6] Hess, S. W., J. B. Weaver, H. J. Siegfeldt, J. N. Whelan, and P. A. Zitlau, "Nonpartisan Political Redistricting by Computer." *Operations Research*, Vol 13, No. 6, pp. 998-1006, 1965.
[7] Hor, M, "Intelligent Electoral Districting Mechanism," International Conference of Machine Learning and Cybernetics (ICMLC 2010), Qingdao, China, July 2010.
[8] Hor M, Lee C, Liu K, *et al*., Analysis of Electoral Redistricting. *The 10th TAAI Conference on Artificial Intelligence and Applications*, Kaohsiung 2005. (in Chinese).
[9] Hsieh S, Decision Making and Influence Factors of Electoral Districting of the Seventh Legislator Elections. *Annual Assembly and Conference*, TPSA, Taipei 2007. (in Chinese).
[10] Hsu J, "Simulation of the Political Impact of Single District Two Votes Model." *Journal of New Century Think Tank*, 17: 6–15, 2002. (in Chinese).
[11] Kaiser, Henry F., "A Measure of the Population Quality of Legislative Apportionment." *American Political Science Review*, Vol 62, No. 1, pp. 208-215, 1968.
[12] Pan C, "Electoral System Reformation: from Electoral Districting Point of View." *Journal of Fu Hsing Kang*, 72: 237–268, 2001. (in Chinese).
[13] Yu C, Liu Y, and Wang Y, "On the Jurisprudence of Electoral Districting Problems after Amendments to the Constitution." Research Report, Ministry of Internal Affair, 2006. (in Chinese).

Research on motion response of wind turbine installation vessel in navigation

***Yanan Huang[1,2], Fengsheng Sun[1,2], Dong Lu[1],Xiaoli Yu[1] ,Yanying Wang[2]and Yongze Xu[3]**

[1] Navigation and Naval Architecture College, Dalian Ocean University, Dalian, China

[2] School of Naval Architecture, Dalian University of Technology, Dalian, China

[3] CD-adapco, xian, china

*Presenting author: 14537370@qq.com

Abstract

The proposal of this investigation is motion response of wind turbine installation vessel in navigation. The topic is studied based on the combinations of theoretical analysis and numerical simulation. It is established a three dimensional numerical pool, the finite volume method is adopted to solve the Navier-Stokes equations, to the actual sea condition of wind turbine installation of numerical analysis of the motion response of the vessel, makes comparison and validation on combining experimental wind wave motion of the ship installation numerical prediction is established. By incident boundary ahead on the way to create the incident wave spectrum functions are defined, at the very back of the computational domain to define momentum source to wave damping, water quality in the area of the wave point on the vertical direction of vibration attenuation finally reached zero, step by step to avoid the influence of reflection wave. When solving the Navier-Stokes equations using Reynolds time-averaged (RANS) numerical simulation method, assuming that turbulent flow field in the change from an average and a pulsating quantity of based on grid control unit to solve mass conservation, momentum conservation equation, energy conservation and VOF volume control, solving the free surface and the flow field near the structure. In order to ensure the accuracy of the calculation, a 2-order time difference scheme is adopted, and the HRIC difference scheme is adopted in the space with the combination of windward difference and downwind difference. In addition, various meshing formats and mesh size sensitivity are verified.

More systematic and innovative conclusions are drawn, which provides a theoretical basis and technical support for prediction of wave motion response for wind turbine installation vessel.

Keywords: wind turbine installation vessel, computation overset grids, wave motion response, VOF

Introduction

With the development of marine strategy, the great attention has been paid to the process of offshoring and large-scale of offshore fans in recent years, they have been higher demand of the equipment for the transport capacity, operation depth and lifting power. Self-propelled marine wind turbine installation vessel is specially adapted for offshore wind farm development and maintenance. The characteristics of Strong force, high operation efficiency

and good economic benefits have gradually made it become the first choice for installation and maintenance of large offshore fans[1][2].

The research on offshore wind turbine installation vessel mostly focuses on the structural strength analysis[3] of wind turbine installation vessel under operating conditions, and also on the performance of wind turbine installation vessel[4] under lifting conditions, the application of AQWA software[5] to study the performance of wind turbine installation vessel under self-propelled condition as well.

In the course of the wind farm, all the windmill blades and motor equipment placed on the deck, causing ship full loaded with the position of the center of gravity under the condition of a ballast condition at the same time increase a lot, in the process of navigation to leg back to the main hull, the position of the center of gravity further increased, so the motion performance of wind turbine installation vessel is great difference from other general ship. At the same time, the ordinary merchant is usually based on the motion analysis of infinite depth assumption, the target ship in the wind farm area navigation with shallow water depth, only tens of meters with lowly speed, so it is necessary to study the motion characteristics of the navigation process.

There are some advantages of Overset Grids.one is easier to perform and automate parametric studies. With a single set of grids, many different configurations can be computed. Grid quality is not affected by changing position or orientation of bodies; Boundary conditions easier to set and so on. The other is easier to handle relative motion of bodies, arbitrary motion can be handled, paths can cross, the tangential motion at close proximity can be handled as well. It can be easy transition to other multiphase models such as VOF to Lagrangian and vice-versa, Fluid film to VOF and vice versa Eulerian or Lagrangian multiphase models within VOF phases.

Numerical Equation

In order to analyze the wave performance of wind turbine installation vessels, the calculation of movement response of ships in waves based on wave potential flow theory[6]. Firstly, the velocity potential of the flow field is solved based on the Laplace equation and boundary conditions. Secondly, the hydrodynamic pressure distribution of the floating body is solved according to Bernoulli equation. Thirdly, the surface integral of the wet surface of the object is subjected to the wave force of the floating body.

$$F_{\mathrm{j}}^{(1)}(\omega) = \iint_S \; p^{(1)}\, n_k ds = \left(f_{0k} + f_{7k} + \sum_{j=1}^{6} T_{kj}\, x_j\right) e^{-i\omega t} \qquad (1)$$

Where: f_{0k}—The incoming power;

f_{7k}—The diffraction force;

T_{kj}—The radial force in the k direction in which the floating body is subjected to the j degree of freedom at unit velocity;

$p^{(1)}$—The pressure distribution;

n_k—The normal direction of the object.

Using generalized wave force, additive mass and damping in the frequency domain, the generalized wave force, additive mass and delay function are obtained by transformation in the time domain. The delay function is expressed as:

$$K_{ij}(t) = \frac{2}{\pi}\int_0^\infty \lambda_{ij}(\omega)\cos(\omega t)d\omega \quad (2)$$

In the equation: λ—the damping matrix of the floating body in the frequency domain.
The additional mass expression of floating body in time domain is expressed.

$$m_{ij}(t) = \mu_{ij}(\omega_0) + \frac{1}{\omega_0}\int_0^\infty K_{ij}(t)\sin(\omega_0 t)\,dt \quad (3)$$

In the formula: μ— the additional mass matrix of floating body in the frequency domain.
Express it as:

$$F_{\omega i}(t) = \int_0^t h_i^1(t-\tau)\eta(\tau)d\tau \quad (4)$$

In the formula: $F_{\omega i}(t)$ — the first wave force of the floating body in the i direction;
$\eta(\tau)$ — wave time calendar;
$h_i^1(\omega)$ —frequency domain wave force Fourier transform.
The motion characteristics of the ship are solved according to the floating time domain equation (5).

$$M + m(t)\ddot{x}(t) + \int_{-\infty}^t K(t-\tau)\times\dot{x}(t)\}d\tau + Cx(t) = F(t) \quad (5)$$

In the formula: M — the generalized mass matrix of floating body;
$m(t)$— additional mass matrix of floating body;
$K(t-\tau)$ — system delay function matrix;
C — the static water recovery coefficient matrix of floating body;
$F(t)$ —generalized force matrix for floating bodies.

Numerical analysis

The principal dimensions:
length: 133. 1 m width: 39. 2 m
depth: 9. 8 m Draft: 5. 6 m
operating depth: 40 m design speed: 12 kn
The numerical simulation of wind turbine installation vessel has chosen cycle 20s, wave Angle from 0°,45°,90°,135° to180° as well.
VOF-approach is suitable, when the grid is fine enough to resolve the interface between two immiscible fluids. VOF considers a single effective fluid whose properties vary according to volume fraction of individual fluids.
The Integral calculation is based on a nested grid (Overset Cell), which is surrounded by a grid of the hull, and the rest is a background grid shown in Figure 1, Figure9, Figure 10, Figure 13 and Figure 14.
In this paper, the VOF algorithm is adopted to show the volume fraction of the air and water in the middle section of the ship, showing the position of free liquid surface in Figure 2, Figure 11, Figure 12, Figure 15 and Figure 16.

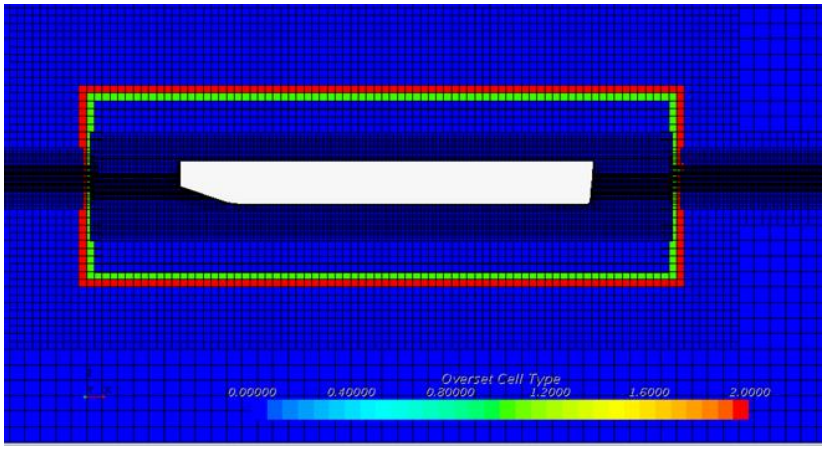


Figure 1. Overset Cell schematic with 90°

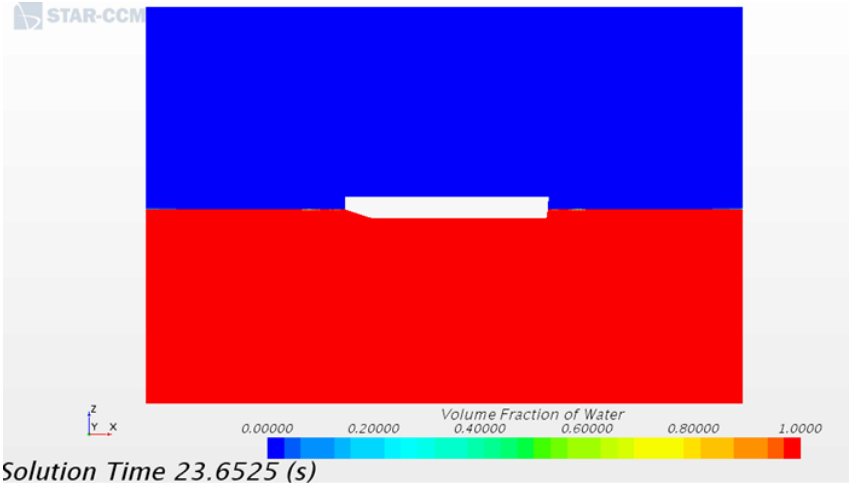


Figure 2. VOF algorithm schematic with 90°

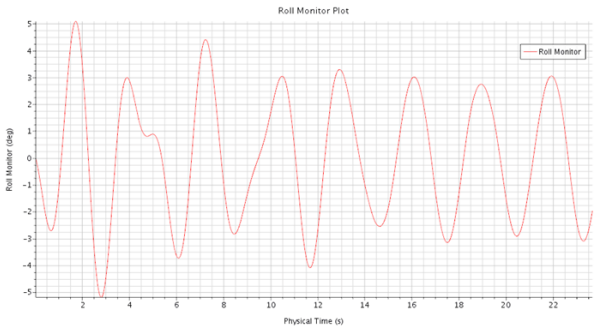


Figure 3. roll monitor plot with 90°

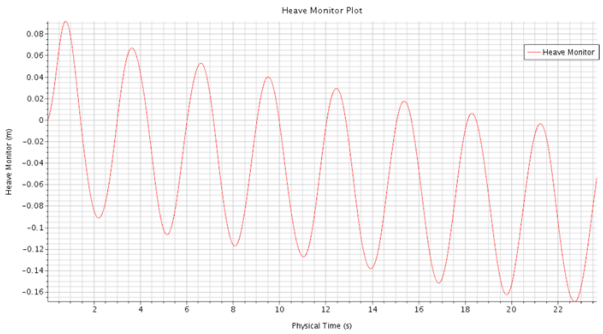


Figure 4. heave monitor plot with 90°

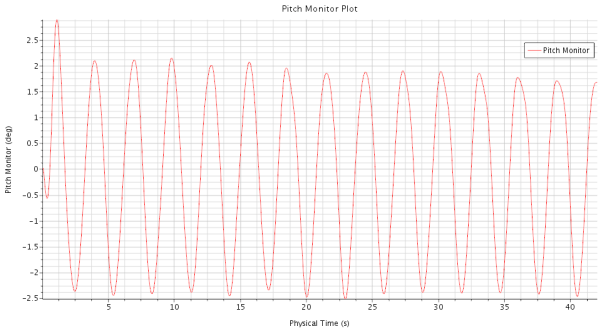


Figure 5. pitch monitor plot with 0°

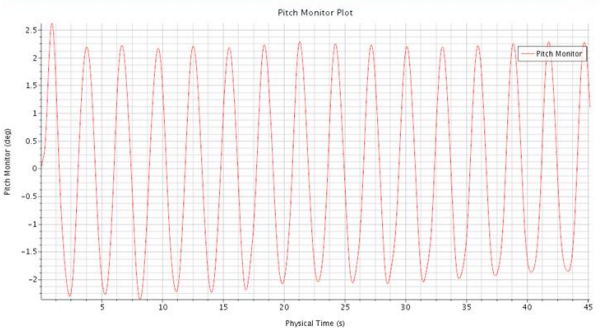


Figure 6. pitch monitor plot with 180°

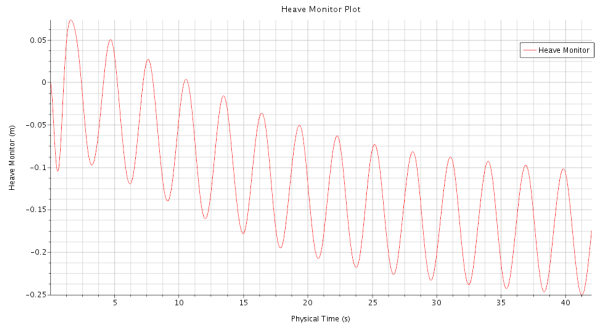


Figure 7. heave monitor plot with 0°

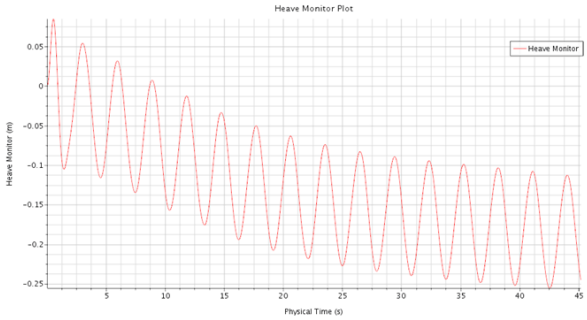


Figure 8. heave monitor plot with 180°

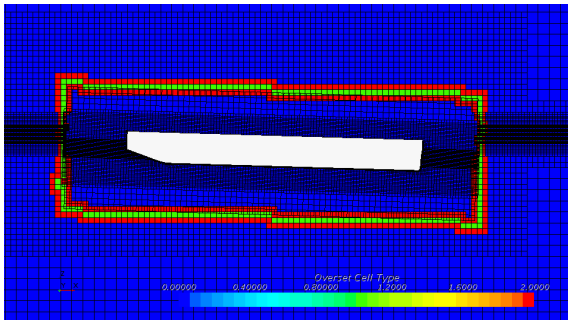


Figure 9. Overset Cell schematic with 0°

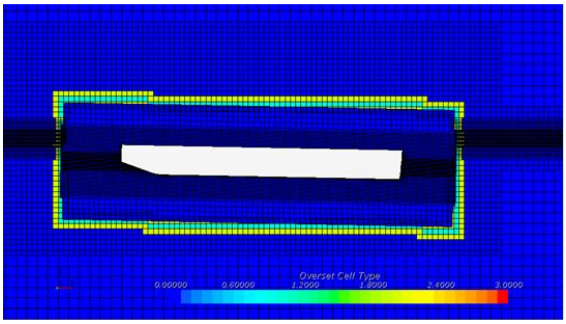


Figure 10. Overset Cell schematic with 180°

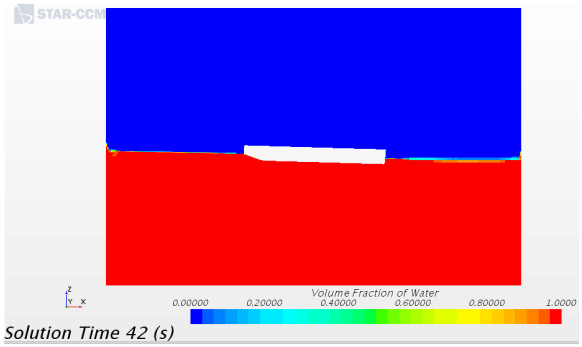


Figure 11. VOF schematic with 0°

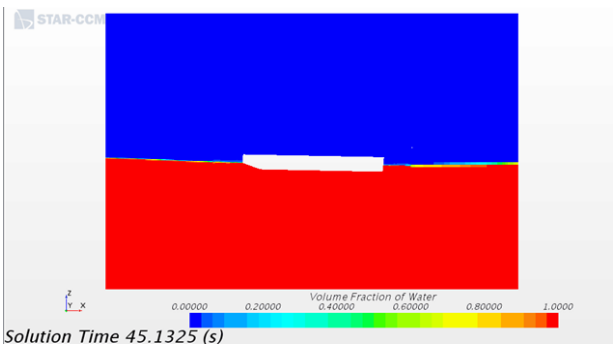


Figure 12. VOF schematic with 180°

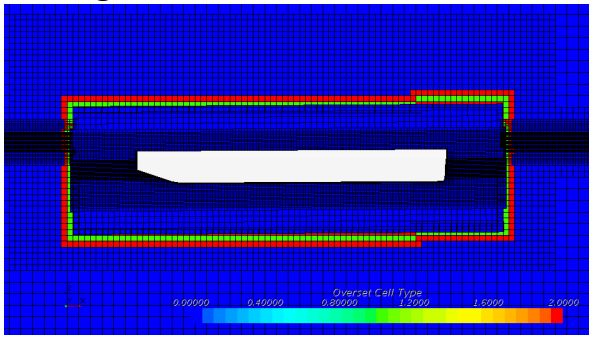


Figure 13. Overset Cell with 45°

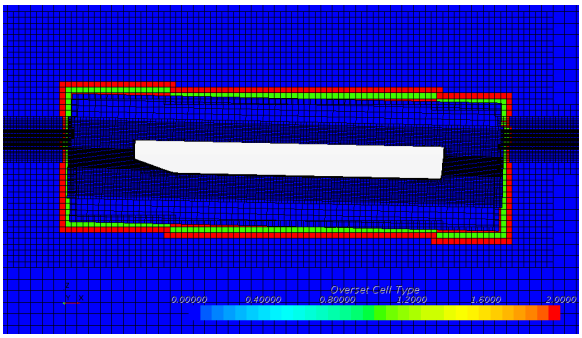


Figure 14. Overset Cell with 135°

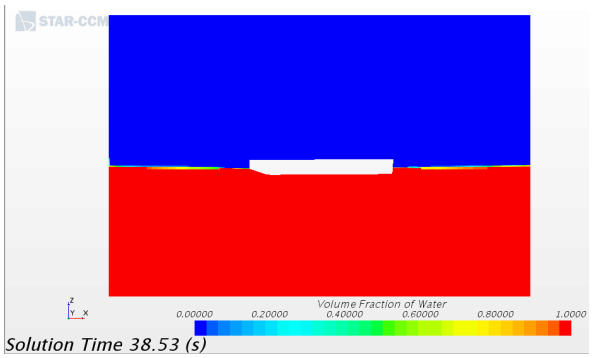


Figure 15. VOF schematic with 45°

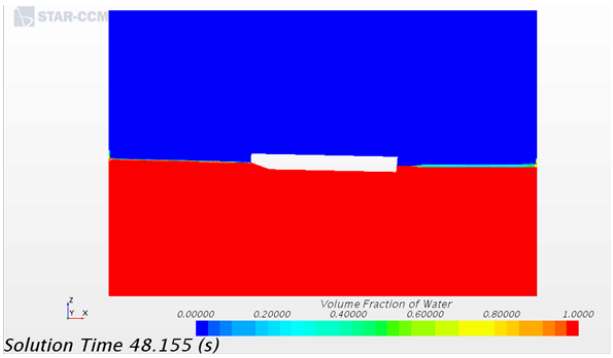


Figure 16. VOF schematic with 135°

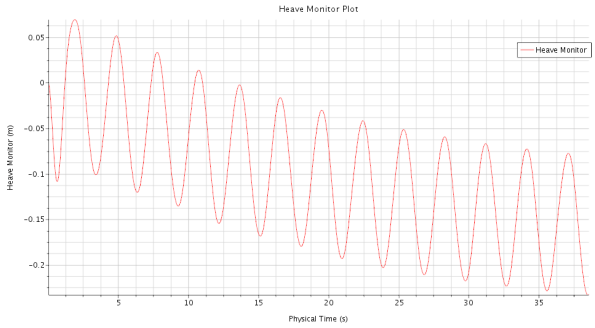


Figure 17. heave monitor plot with 0°

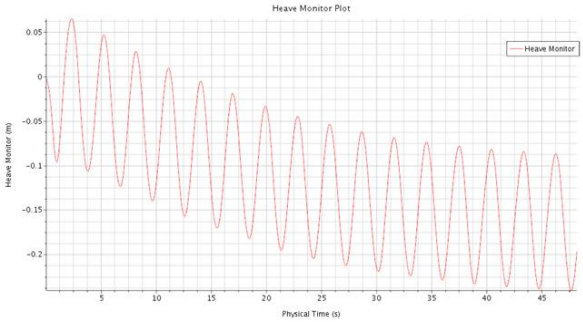


Figure 18. heave monitor plot with 180°

Conclusions

The topic are shown the change of the ship's pitch and the calculation time with 0° and 180°, showing that the period of the ship's pitch is consistent with the period of the incident wave, and presents the periodic change, just as Figure 5 and Figure 6 are shown.

The figures show the heave of the hull with the time of calculation, which shows that the period of the ship's heaving and sinking are consistent with the period of the incident wave, and presents periodic changes. Just as Figure4, Figure7 and Figure 8, Figure17 and Figure 18 are shown.

Traditional theory of potential flow is usually combined with the far field boundary conditions using the boundary element method is used to solve the wave potential function and stream function, the algorithm can more quickly and efficiently solving the wave diffraction and radiation problems of structure, but to the breaking of waves and the splash phenomenon such as the lack of effective means to solve. Through CFD discrete finite Volume method is used to solve the Navier-Stokes equations, for a large scale such as waves of two intersecting interface using Eulerian multiphase flow VOF (Volume of Fluid) method for processing. It is able to directly capture wave propagation phenomena, such as crushing and splash in the process of model can be directly obtained also ship hydrodynamics problems, and full dimensions can be more accurate to calculate and forecast problems such as the motion response of wind turbine installation vessel.

References

[1] Ding jinhong, tan jiahua. (2009) Overview of special installation ships for offshore wind power, *China Offshore platform* **5**, 6–11.

[2] He yanping, Yang qi, du pengfei, etc. (2009)Transportation, installation and maintenance of offshore wind turbine units, *SHIP &OCEAN ENGINEERING* **4**, 136-139.
[3] Yao zhenqiu, ma yimeng, han qiang, et al. (2014)Finite element analysis of the structural strength of the whole ship of Marine fan hoisting operation, *Ship mechanics* **16**:136 -146.
[4] Tu yan, xu rong, Yang jun, et al. (2014) Analysis of motion response frequency domain under the self-promotion of superlarge fan installation platform , *SHIP &OCEAN ENGINEERING* **3**,148-152.
[5] ZHANG Li-jun，ZHAO Zhi-jian, WANG Wen-hua , WANG Yan-ying. (2015) Analysis on Motion Response of Wind Turbine Installation Vessel in Navigation, *SHIP &OCEAN ENGINEERING* **8**, 28-31.
[6] YATES M L, BENOIT M. (2015) Accuracy and efficiency of two numerical methods of solving the potential flow problem for highly nonlinear and dispersive water waves, *International Journal for Numerical Methods in Fluids* **10**,610-640.

A state space boundary element method with analytical nearly singular integral formulation

†Changzheng Cheng[1], *Zhilin Han[1], Meng Wu[2], Zhongrong Niu[1], Hongyu Sheng[1]

[1] School of Civil Engineering, Hefei University of Technology, Hefei 230009, PR China;
[2] School of Mathematics, Hefei University of Technology, Hefei 230009, PR China

*Presenting author: han_zhilin@mail.hfut.edu.cn
†Corresponding author: changzheng.cheng@hfut.edu.cn

Abstract:
A new method named the state space boundary element method (SSBEM) is established, in which the problem domain is divided into two parts. One is the boundary element domain which includes the interested inner point and the other is the state space domain. The boundary integral equation and state space equation are combined together by the interfacial continuity condition to form the system equation of the SSBEM. The SSBEM synthesizes both advantages of the boundary element method and state space method, while it will get inaccurate result when it is used to evaluate the mechanical quantity of the point very close to the boundary element, because the Gaussian's quadrature fails to calculate the nearly singular integral in the boundary integral equation. The analytical formulation developed by part of the authors before is introduced to deal with the nearly singular integral. Thus, the SSBEM can yield accurate physical quantities for the point very close to the boundary element. The SSBEM result can well agree with the one from the finite element method (FEM), while the discretized element is much less than the one in the FEM. Meanwhile, the SSBEM can analyze very thin coating while the FEM fails due to the limitation of Boolean operation tolerance.

Keywords: state space method, boundary element method, nearly singular integral, analytical formulation

1. Introduction

The state space method (SSM) is widely used in the analysis of laminated structure because of its simple recursive formulation [1]. According to different discretization method of in-plane variables, the SSM can be summarized as four types, which are the finite strip SSM [2]-[4], the finite element SSM [5]-[7], the differential quadrature SSM [8] and the meshless SSM [9]-[11]. Since only in-plane interfaces are meshed, the SSM can decrease the element number when they are compared with the finite element method (FEM). However, when an interested inner point is designated to be analyzed, the structure has to be delaminated into more sub-layers to ensure that the interior point is exactly located on the interface which undoubtedly increases the computational cost.

The boundary element method (BEM) is based on the classical integral equation formulation

of boundary value problem [12]-[15]. The unknown physical quantity at any interior point can be directly evaluated by the boundary integral equation. The experience has shown that the BEM has advantages over other numerical methods for reducing the calculation amount. However, the sub-domain method has to be introduced when it comes to the laminated structure, which will increase the redundant interface quantity in the BEM.

By coupling two independent methods, their respective merits can be enlarged and shortcomings can be overcome. There are some direct coupling techniques, such as the BEM-BEM method [16][17], the FEM-BEM method [18]-[20], fast multipole BEM-SSM method [21], etc. There is also an iterative coupling technique [22]-[24] developed in the FEM-BEM method, which provides a novel calculative strategy for the coupled method. Herein, the BEM-SSM coupling method named the state space boundary element method (SSBEM) is developed basing on the direct coupling technique, which can efficiently analyze the laminated structure and easily yield the physical quantity of the inner point without increasing the amount of calculation.

It should be pointed out that the integral kernel in the boundary integral equation will present nearly singularity when the distance from the source point to the field point is close to but not equal to zero. The accuracy of SSBEM, such as the evaluation of the near boundary inner point, depends on the accurate evaluation of nearly singular integral significantly. The analytical formulation for the nearly singular integral proposed by part of the authors before [25] is introduced to the SSBEM, which can be used to calculate the physical quantity for near boundary interior point accurately.

2. Boundary integral equation with analytical nearly singular integral formulation

The thermal elasticity problem is considered here, which can be degenerated to the elasticity problem if the temperature loading is set as zero. The displacement boundary integral equation with respect to the source point $\boldsymbol{y}$ for two-dimensional thermal elasticity problem can be written as

$$C_{ij}(\boldsymbol{y})u_j(\boldsymbol{y}) = \int_\Gamma [U_{ij}^* t_j(\boldsymbol{x}) - T_{ij}^* u_j(\boldsymbol{x})]\mathrm{d}\Gamma + \int_\Gamma [R_i^* T(\boldsymbol{x}) - Q_i^* \partial T(\boldsymbol{x})/\partial n]\mathrm{d}\Gamma \tag{1}$$

where $\boldsymbol{x}$ is the field point, $i, j = 1, 2$ respectively denote the direction along *x*- and *y*-axis in the rectangular coordinate system. $C_{ij}(\boldsymbol{y})$ is the coefficient determined by the local geometry at point $\boldsymbol{y}$. $u_j(\boldsymbol{x})$ and $t_j(\boldsymbol{x})$ are, respectively, the displacement and traction on the boundary Γ of the structure. $T(\boldsymbol{x})$ and $\partial T(\boldsymbol{x})/\partial n$ are the temperature and temperature gradient at point $\boldsymbol{x}$, respectively. U_{ij}^* and T_{ij}^* are the fundamental solutions for governing equation of elasticity, for the plane strain problem which can be written as

$$U_{ij}^{*}=\frac{1}{8\pi(1-\nu)G}[(3-4\nu)\ln r\delta_{ij}-\frac{r_i r_j}{r^2}] \tag{2a}$$

$$T_{ij}^{*}=\frac{1}{4\pi(1-\nu)}[(1-2\nu)(\frac{r_i n_j-r_j n_i-r_n\delta_{ij}}{r^2})-\frac{2r_i r_j r_n}{r^4}] \tag{2b}$$

where δ_{ij} is the Kronecker symbol, G is the shear modulus and ν is the Poisson's ratio. $r_i=x_i-y_i$ and $r=(r_i r_i)^{1/2}$ is the distance between the field point and source point. R_i^{*} and Q_i^{*} in Eq. (1) are the fundamental solutions related to the thermal elasticity governing equation, which can be written as

$$R_i^{*}=\frac{\alpha(1+\nu)}{4\pi(1-\nu)}[(\ln\frac{1}{r}-\frac{1}{2})n_i-\frac{r_i r_n}{r^2}] \tag{3a}$$

$$Q_i^{*}=\frac{\alpha(1+\nu)}{4\pi(1-\nu)}(\ln\frac{1}{r}-\frac{1}{2})r_i \tag{3b}$$

where α is the thermal expansion coefficient.

By discretizing Eq.(1) along the boundary, the boundary element method system equation can be yielded like

$$\boldsymbol{Hu}=\boldsymbol{Gt} \tag{4}$$

where $\boldsymbol{H}$ and $\boldsymbol{G}$ are coefficient matrices, $\boldsymbol{u}$ is nodal displacement vector and $\boldsymbol{t}$ is nodal traction vector.

After the unknown displacement and traction on the boundary Γ being solved from Eq. (4), the displacement at any interior point can be evaluated by the displacement boundary integral equation

$$u_i(\boldsymbol{y})=\int_\Gamma[U_{ij}^{*}t_j(\boldsymbol{x})-T_{ij}^{*}u_j(\boldsymbol{x})]\mathrm{d}\Gamma+\int_\Gamma[R_i^{*}T(\boldsymbol{x})-Q_i^{*}\,\partial T(\boldsymbol{x})/\partial n]\mathrm{d}\Gamma \tag{5}$$

By taking the derivative of Eq. (5) and introducing the strain-stress relationship, the stress at any interior point can be expressed by the stress boundary integral equation

$$\sigma_{ij}(\boldsymbol{y})=\int_\Gamma[W_{ijk}^{*}t_k(\boldsymbol{x})-S_{ijk}^{*}u_k(\boldsymbol{x})]\mathrm{d}\Gamma+\int_\Gamma[R_{ij}^{*}T(\boldsymbol{x})-Q_{ij}^{*}\partial T(\boldsymbol{x})/\partial n]\mathrm{d}\Gamma-\sigma_{ij}^{0}(\boldsymbol{y}) \tag{6}$$

where $i,j,k=1,2$, $\sigma_{ij}^{0}(\boldsymbol{y})=2G(1+\nu)\alpha\cdot\Delta T(\boldsymbol{y})\delta_{ij}/(1-2\nu)$ and the integral kernels are listed as follows

$$W_{ijk}^{*}=\frac{1}{4\pi(1-\nu)}[(1-2\nu)\frac{r_i\delta_{jk}+r_j\delta_{ik}-r_k\delta_{ij}}{r^2}+\frac{2r_i r_j r_k}{r^4}] \tag{7a}$$

$$S_{ijk}^{*}=\frac{G}{2\pi(1-\nu)}[\frac{(1-2\nu)(n_i\delta_{jk}+n_j\delta_{ki})-(1-4\nu)n_k\delta_{ij}}{r^2}-\frac{4r_n r_i r_j r_k}{r^6}+\frac{2(1-2\nu)(r_n r_k\delta_{ij}+r_i r_j n_k)+2\nu(r_n r_j\delta_{jk}+r_n r_j\delta_{ik}+n_i n_j r_k+n_j r_i r_k}{r^4}] \tag{7b}$$

$$R_{ij}^{*}=\frac{G\alpha(1+\nu)}{2\pi(1-\nu)}[\frac{r_n\delta_{ij}/(1-2\nu)+n_i r_j+n_j r_i}{r^2}-\frac{2r_i r_j r_n}{r^4}] \tag{7c}$$

$$Q_{ij}^{*}=\frac{G\alpha(1+\nu)}{2\pi(1-\nu)}[\frac{r_i r_j}{r^2}-\frac{1}{1-2\nu}(\ln\frac{1}{r}-\frac{1+2\nu}{2})\delta_{ij}] \tag{7d}$$

It can be observed that the integral kernels shown in Eqs. (2,3) and Eq. (7) contain nearly singular items of order $O(1/r^2)$, $O(1/r^4)$ and $O(1/r^6)$. When the source point is close to the integral element, the implementation with Gaussian's quadrature will lead to significant inaccuracy in obtaining the results of nearly singular integral.

The integrals in Eqs. (1,5,6) can be concluded as three kinds of forms as follows

$$I_1=\int_{-1}^{1}\frac{P_1(\xi)}{R}\mathrm{d}\xi,\quad I_2=\int_{-1}^{1}\frac{P_2(\xi)}{R^2}\mathrm{d}\xi,\quad I_3=\int_{-1}^{1}\frac{P_3(\xi)}{R^4}\mathrm{d}\xi \tag{8}$$

where $R=r^2$, $P_1(\xi)$, $P_2(\xi)$ and $P_3(\xi)$ are the polynomial functions of local coordinate ξ. By means of integration by parts, Eq. (8) is transformed into the analytical formulation [25] as follows

$$\begin{aligned}I_1=&\{\frac{2}{\delta}P_1 g-\frac{1}{a}P_1'[(\frac{R'}{\delta})g-\frac{1}{2}\ln(\delta^2+R'^2)]+\frac{\delta}{4a^2}P_1''[(\frac{R'}{\delta})^2 g-\frac{R'}{\delta}+g]-\\&\frac{\delta^2}{8a^3}P_1'''\ (K_0+K_2-\frac{2aR}{\delta^2})\}\ \Big|_{\xi=-1}^{1}-\frac{1}{2a}\int_{-1}^{1}P_1''\ln(\delta^2+R'^2)\,\mathrm{d}\xi\\&+\frac{\delta^2}{8a^3}\int_{-1}^{1}P_1^{(4)}(K_0+K_2-\frac{2aR}{\delta^2})\ \mathrm{d}\xi\end{aligned} \tag{9a}$$

$$\begin{aligned}I_2=&\{\frac{1}{\delta^2}\frac{R'}{R}P_2+\frac{2}{\delta^3}(2aP_2-R'P_2')g+\frac{1}{2a\delta}P_2''[(\frac{R'}{\delta})^2 g+g-\frac{R'}{\delta}]\\&-\frac{1}{4a^2}P_2'''(K_0+K_2-\frac{2aR}{\delta^2})+\frac{\delta}{8a^3}P_2^{(4)}(K_1+\frac{1}{3}K_3)\}\ \Big|_{\xi=-1}^{1}\\&-\frac{1}{4a^2}\int_{-1}^{1}P_2^{(4)}[\frac{2aR}{\delta^2}+\frac{1}{6}(\frac{R'}{\delta})^2+\frac{1}{3}\ln(\delta^2+R'^2)-\frac{2}{3}\ln\delta\,]\mathrm{d}\xi\\&-\frac{\delta}{8a^3}\int_{-1}^{1}P_2^{(5)}(K_1+\frac{1}{3}K_3)\mathrm{d}\xi\end{aligned} \tag{9b}$$

$$\begin{aligned}I_3=&\{\frac{R'}{2\delta^2R^2}P_3+\frac{R'}{2\delta^4R}(6aP_3-R'P_3')+\frac{1}{\delta^5}(12a^2P_3+\frac{3}{2}R'^2P_3''-6aR'P_3')g\\&+\frac{1}{2\delta^3}P_3''(g-\frac{R'}{\delta})+\frac{1}{4a\delta^2}P_3'''[\frac{2aR}{\delta^2}+\frac{1}{2}(\frac{R'}{\delta})^2-\frac{R'}{\delta}g-(\frac{R'}{\delta})^3 g]\\&+\frac{1}{32a^2\delta}P_3^{(4)}[(\frac{R'}{\delta})^4 g+2(\frac{R'}{\delta})^2 g+g-\frac{1}{3}(\frac{R'}{\delta})^3-\frac{R'}{\delta}\ ]\}\Big|_{\xi=-1}^{1}\\&-\frac{1}{4a\delta^2}\int_{-1}^{1}P_3^{(4)}[\frac{1}{2}(\frac{R'}{\delta})^2+\frac{2aR}{\delta^2}]\mathrm{d}\xi-\frac{1}{8\delta a^2}\int_{-1}^{1}P_3^{(5)}(K_1+K_3)\mathrm{d}\xi\end{aligned} \tag{9c}$$

where the prime and bracket in the superscript denote the partial derivatives to ξ, $K_n\,(n=0,1,2,3)$ are functions with respect to g, $g=\mathrm{arctg}(R'/\delta)$, $\delta=\sqrt{4ac-b^2}$, a, b and c are the expanded coefficients of r^2 in the local coordinate ξ, i.e.,

$$r^2=a\xi^2+b\xi+c \tag{10}$$

When the polynomial function whose order is lower than four is adopted, Eq.(9) is the analytical formulation.

3. State space equation with precise integral method

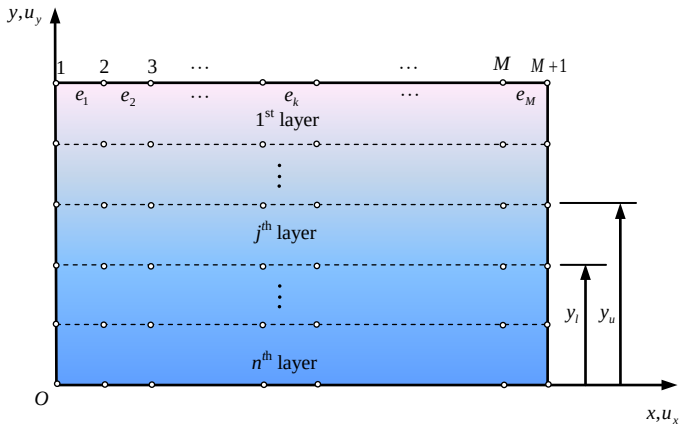


Fig.1 Discretization for laminated structure in the state space method

A laminated structure composed by n layers is considered here, see Fig.1, in which each layer could have different elastic modulus E and different thermal expansion coefficients α. On the basis of Hellinger-Reissner variational principle, the equilibrium equation and strain-displacement relation at any elastic layer shown in Fig. 1 can be equivalently written as

$$\iint_{\Omega} \delta \boldsymbol{u}^{\mathrm{T}} \left[\boldsymbol{E}(\nabla)\boldsymbol{\sigma} + \boldsymbol{f} \right] \mathrm{d}\Omega = 0 \tag{11a}$$

$$\iint_{\Omega} \delta \boldsymbol{\sigma}^{\mathrm{T}} \left[\boldsymbol{\varepsilon} - \boldsymbol{E}^{\mathrm{T}}(\nabla)\boldsymbol{u} \right] \mathrm{d}\Omega = 0 \tag{11b}$$

where $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ denote strain and stress vectors, respectively, $\boldsymbol{f}$ is the body force. $\boldsymbol{E}(\nabla)$ is the differential operator

$$\boldsymbol{E}(\nabla) = \begin{bmatrix} \partial/\partial x & 0 & \partial/\partial y \\ 0 & \partial/\partial y & \partial/\partial x \end{bmatrix} \tag{12}$$

For the thermal elasticity problem, the stress-strain relation can be written as

$$\boldsymbol{\varepsilon} = \boldsymbol{S}\boldsymbol{\sigma} + \boldsymbol{J} \tag{13}$$

where $\boldsymbol{J} = [\alpha \cdot \Delta T \ \alpha \cdot \Delta T \ 0]^{\mathrm{T}}$, ΔT is the temperature loading. The flexibility matrix $\boldsymbol{S}$ can be expressed as

$$\boldsymbol{S} = \begin{bmatrix} S_1 & S_2 & 0 \\ S_2 & S_1 & 0 \\ 0 & 0 & S_3 \end{bmatrix} \tag{14}$$

where $S_1 = 1/E$, $S_2 = -\nu/E$, $S_3 = 2(1+\nu)/E$.

After introducing the stress-strain relationship, Eq.(11) can be transformed into

$$\iint_{\Omega}\begin{Bmatrix}\delta u_x\\ \delta u_y\\ \delta\tau_{xy}\\ \delta\sigma_{yy}\end{Bmatrix}^{\mathrm{T}}\left(\begin{bmatrix}0&0&1&0\\0&0&0&1\\1&0&0&0\\0&1&0&0\end{bmatrix}\frac{\partial}{\partial y}\begin{Bmatrix}u_x\\ u_y\\ \sigma_{xy}\\ \sigma_{yy}\end{Bmatrix}-\begin{bmatrix}0&0&0&0\\0&0&-\partial/\partial x&0\\0&-\partial/\partial x&S_3&0\\0&0&0&S_1\end{bmatrix}\begin{Bmatrix}u_x\\ u_y\\ \sigma_{xy}\\ \sigma_{yy}\end{Bmatrix}+\begin{bmatrix}\partial/\partial x\\0\\0\\-S_2\end{bmatrix}\sigma_{xx}+\begin{bmatrix}0\\0\\0\\-\alpha\end{bmatrix}\Delta T\right)\mathrm{d}\Omega=0 \tag{15a}$$

$$\iint_{\Omega}\{\delta\sigma_{xx}\}^{\mathrm{T}}\left(S_1\sigma_{xx}-\begin{bmatrix}\partial/\partial x&0&0&-S_2\end{bmatrix}\begin{Bmatrix}u_x\\ u_y\\ \sigma_{xy}\\ \sigma_{yy}\end{Bmatrix}+\alpha\cdot\Delta T\right)\mathrm{d}\Omega=0 \tag{15b}$$

where u_x and u_y are the displacements along the $x-$axis and $y-$axis, respectively. The relationship between the traction and stress on the boundary or at the interface can be written as

$$t_x=\sigma_{xy},\quad t_y=\sigma_{yy} \tag{16}$$

In the state space method, only the horizontal interfaces are discretized and it is assumed that M linear elements are meshed on each interface as shown in Fig. 1. By integrating along the *x*-coordinate and considering Eq.(16), Eq.(15) can be written as

$$\int_y\begin{Bmatrix}\delta\boldsymbol{u}\\ \delta\boldsymbol{t}\end{Bmatrix}^{\mathrm{T}}\left(\boldsymbol{A}\frac{\mathrm{d}}{\mathrm{d}y}\begin{Bmatrix}\boldsymbol{u}\\ \boldsymbol{t}\end{Bmatrix}-\boldsymbol{B}\begin{Bmatrix}\boldsymbol{u}\\ \boldsymbol{t}\end{Bmatrix}+\boldsymbol{C}\boldsymbol{s}+\boldsymbol{M}\boldsymbol{T}\right)\mathrm{d}y=0 \tag{17a}$$

$$\int_y\{\delta\boldsymbol{s}\}^{\mathrm{T}}\left(\boldsymbol{D}\boldsymbol{s}-\boldsymbol{E}\begin{Bmatrix}\boldsymbol{u}\\ \boldsymbol{t}\end{Bmatrix}+\boldsymbol{N}\boldsymbol{T}\right)\mathrm{d}y=0 \tag{17b}$$

where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$, $\boldsymbol{M}$ and $\boldsymbol{N}$ are the coefficient matrices deduced from Eq.(15). The vector $\boldsymbol{s}$ and $\boldsymbol{T}$ are respectively composed of nodal stress σ_{xx} and nodal temperature variation.

By eliminating $\boldsymbol{s}$ from Eq. (17) and utilizing the variational principle, the state equation at any layer can be written as

$$\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}y}=\boldsymbol{W}\boldsymbol{R}+\boldsymbol{B} \tag{18}$$

where the state variable $\boldsymbol{R}=\{\boldsymbol{u}^{\mathrm{T}}\ \boldsymbol{t}^{\mathrm{T}}\}^{\mathrm{T}}$, $\boldsymbol{W}=\boldsymbol{A}^{-1}(\boldsymbol{B}-\boldsymbol{C}\boldsymbol{D}^{-1}\boldsymbol{E})$, and the vector $\boldsymbol{B}=\boldsymbol{A}^{-1}(\boldsymbol{C}\boldsymbol{D}^{-1}\boldsymbol{N}-\boldsymbol{M})\boldsymbol{T}$, which is the term related to the temperature variation.

The precise integral method [26] is introduced to solve Eq. (18). Let's take the jth layer in Fig. 1 for example. If the jth layer is sufficiently thin, $\boldsymbol{R}_j(y)$ and $\boldsymbol{B}_j(y)$ in Eq. (18) can be approximated by $[\boldsymbol{R}_j(y_u)+\boldsymbol{R}_j(y_l)]/2$ and $[\boldsymbol{B}_j(y_u)+\boldsymbol{B}_j(y_l)]/2$, respectively, where y_u and

y_l are *y*-coordinates of the upper interface and lower interface of the jth layer, respectively. For the layer which is not thin enough, it can be divided into $K_j = 2^k$ sub-layers with uniform thickness and each sub-layer can be approximately calculated as a thin layer. For the *i*th sub-layer in the jth layer, Eq. (18) can be written in the integrated form as follows

$$\int_{(i-1)\Delta_j}^{i\Delta_j} \mathrm{d}\boldsymbol{R}_j^{(i)}(y) = \int_{(i-1)\Delta_j}^{i\Delta_j} \boldsymbol{W}_j \boldsymbol{R}_j^{(i)}(y)\mathrm{d}y + \int_{(i-1)\Delta_j}^{i\Delta_j} \boldsymbol{B}_j^{(i)}(y)\mathrm{d}y \tag{19}$$

where Δ_j is the thickness of each sub-layer. By utilizing the approximation in the *i*th sub-layer which is thin enough, Eq. (19) can be transformed into

$$(\boldsymbol{I}-\boldsymbol{O})\boldsymbol{R}_j^{(i)}(i\Delta_j) = (\boldsymbol{I}+\boldsymbol{O})\boldsymbol{R}_j^{(i)}((i-1)\Delta_j) + \boldsymbol{U}_j^{(i)} \tag{20}$$

where $\boldsymbol{O} = \Delta_j \boldsymbol{W}_j / 2$, $\boldsymbol{U}_j^{(i)} = \Delta_j[\boldsymbol{B}_j(i\Delta_j) + \boldsymbol{B}_j((i-1)\Delta_j)]/2$ and $\boldsymbol{I}$ is a unit matrix. By applying Eq. (20) to each sub-layer and using the continuity condition between the interface of two adjacent sub-layers, i.e.,

$$\boldsymbol{R}_j^{(i)}[(i-1)\Delta_j] = \boldsymbol{R}_j^{(i-1)}[(i-1)\Delta_j] \tag{21}$$

one can get

$$(\boldsymbol{I}-\boldsymbol{O})^{K_j}\boldsymbol{R}_j(y_l) = (\boldsymbol{I}+\boldsymbol{O})^{K_j}\boldsymbol{R}_j(0) + \sum_{i=1}^{K_j}(\boldsymbol{I}+\boldsymbol{O})^{K_j-i}(\boldsymbol{I}-\boldsymbol{O})^{i-1}\boldsymbol{U}_j^{(i)} \tag{22}$$

The matrices $(\boldsymbol{I}-\boldsymbol{O})^{K_j}$ and $(\boldsymbol{I}+\boldsymbol{O})^{K_j}$ can be calculated by using the iteration process. The calculation of the last item $\boldsymbol{U}_j = \sum_{i=1}^{K_j}(\boldsymbol{I}+\boldsymbol{O})^{K_j-i}(\boldsymbol{I}-\boldsymbol{O})^{i-1}\boldsymbol{U}_j^{(i)}$ in Eq. (22) need to use the identity $(\boldsymbol{I}-\boldsymbol{O})(\boldsymbol{I}+\boldsymbol{O}) = (\boldsymbol{I}+\boldsymbol{O})(\boldsymbol{I}-\boldsymbol{O})$ which can reduce the iteration times from $K_j = 2^k$ to k. Then, Eq. (22) is transformed into the equation as follows

$$\boldsymbol{P}_j \begin{Bmatrix} \boldsymbol{u}_j(y_l) \\ \boldsymbol{t}_j(y_l) \end{Bmatrix} = \boldsymbol{Q}_j \begin{Bmatrix} \boldsymbol{u}_j(y_u) \\ \boldsymbol{t}_j(y_u) \end{Bmatrix} + \boldsymbol{U}_j \tag{23}$$

where $\boldsymbol{u}_j(y_l)$, $\boldsymbol{t}_j(y_l)$ and $\boldsymbol{u}_j(y_u)$, $\boldsymbol{t}_j(y_u)$ denote the displacement and traction on the lower and upper surface of the jth layer, respectively.

4. State space boundary element method

Let's take a three-layered laminated structure shown in Fig. 2 for instance to explain the establishment of the state space boundary element method. The laminated structure is divided into two parts, in which the 1st layer is the BEM domain and the others are in the range of the

SSM domain.

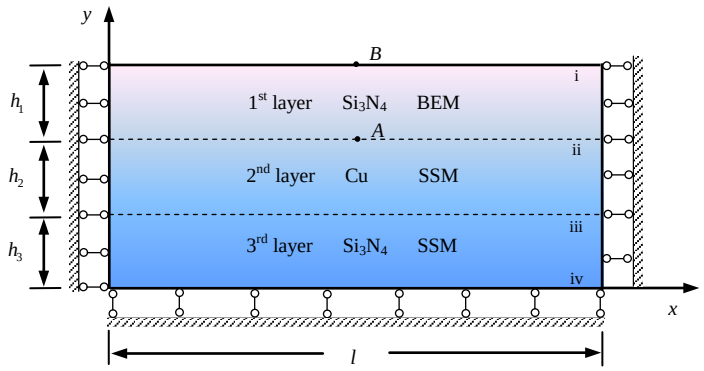


Fig.2 BEM and SSM domain in the state space boundary element method

In Fig. 2, the notation i and iv denote the upper and lower edge, respectively. ii and iii denote the interface between two adjacent layers, respectively. For the BEM domain, the linear equation system can be formed on the basis of Eq.(4) and written as

$$\begin{bmatrix}\boldsymbol{H}_1^{\mathrm{B}} & \boldsymbol{H}_2^{\mathrm{B}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{OB}}^{\mathrm{B}}\\ \boldsymbol{u}_{\mathrm{ii}}^{\mathrm{B}}\end{Bmatrix}=\begin{bmatrix}\boldsymbol{G}_1^{\mathrm{B}} & \boldsymbol{G}_2^{\mathrm{B}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{t}_{\mathrm{OB}}^{\mathrm{B}}\\ \boldsymbol{t}_{\mathrm{ii}}^{\mathrm{B}}\end{Bmatrix} \tag{24}$$

where the superscript B means that the corresponding component is in the BEM domain, the subscript OB denotes the outer boundary of the BEM domain. On basis of Eq.(23), the thermal elasticity state space equation for the second and third layer can be respectively expressed as

$$\begin{bmatrix}\boldsymbol{Q}_1^{\mathrm{S}} & \boldsymbol{Q}_2^{\mathrm{S}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{ii}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{ii}}^{\mathrm{S}}\end{Bmatrix}=\begin{bmatrix}\boldsymbol{P}_1^{\mathrm{S}} & \boldsymbol{P}_2^{\mathrm{S}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{iii}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{iii}}^{\mathrm{S}}\end{Bmatrix}-\boldsymbol{U}_2 \tag{25a}$$

$$\begin{bmatrix}\boldsymbol{Q}_3^{\mathrm{S}} & \boldsymbol{Q}_4^{\mathrm{S}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{iii}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{iii}}^{\mathrm{S}}\end{Bmatrix}=\begin{bmatrix}\boldsymbol{P}_3^{\mathrm{S}} & \boldsymbol{P}_4^{\mathrm{S}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{iv}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{iv}}^{\mathrm{S}}\end{Bmatrix}-\boldsymbol{U}_3 \tag{25b}$$

where the superscript S means that the corresponding parameter is in the SSM domain.

By utilizing the continuity conditions at the interface between the BEM domain and SSM domain

$$\boldsymbol{u}_{\mathrm{ii}}^{\mathrm{B}}=\boldsymbol{u}_{\mathrm{ii}}^{\mathrm{S}},\quad \boldsymbol{t}_{\mathrm{ii}}^{\mathrm{B}}=\boldsymbol{t}_{\mathrm{ii}}^{\mathrm{S}} \tag{26}$$

to combine Eq.(24) and Eq.(25), the state space boundary element equation of a laminated structure can be finally written as follows

$$\begin{bmatrix}\boldsymbol{H}_1^{\mathrm{B}} & \boldsymbol{H}_2^{\mathrm{B}} & -\boldsymbol{G}_2^{\mathrm{B}} & 0 & 0 & 0 & 0\\ 0 & \boldsymbol{Q}_1^{\mathrm{S}} & \boldsymbol{Q}_2^{\mathrm{S}} & -\boldsymbol{P}_1^{\mathrm{S}} & -\boldsymbol{P}_2^{\mathrm{S}} & 0 & 0\\ 0 & 0 & 0 & \boldsymbol{Q}_3^{\mathrm{S}} & \boldsymbol{Q}_4^{\mathrm{S}} & -\boldsymbol{P}_3^{\mathrm{S}} & -\boldsymbol{P}_4^{\mathrm{S}}\end{bmatrix}\begin{Bmatrix}\boldsymbol{u}_{\mathrm{OB}}^{\mathrm{B}}\\ \boldsymbol{u}_{\mathrm{ii}}^{\mathrm{B}}\\ \boldsymbol{t}_{\mathrm{ii}}^{\mathrm{B}}\\ \boldsymbol{u}_{\mathrm{iii}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{iii}}^{\mathrm{S}}\\ \boldsymbol{u}_{\mathrm{iv}}^{\mathrm{S}}\\ \boldsymbol{t}_{\mathrm{iv}}^{\mathrm{S}}\end{Bmatrix}=\boldsymbol{G}_1^{\mathrm{B}}\begin{Bmatrix}\boldsymbol{t}_{\mathrm{OB}}^{\mathrm{B}}\\ 0\\ 0\end{Bmatrix}-\begin{Bmatrix}0\\ \boldsymbol{U}_2\\ \boldsymbol{U}_3\end{Bmatrix} \tag{27}$$

The unknown physical quantity on the outer boundary at BEM domain and the one on all interfaces can be yielded after Eq.(27) being solved. Then, the displacement and stress of near boundary interior point in the BEM domain can be accurately evaluated by introducing the analytical formulation to the boundary integral equation.

5. Numerical examples

5.1 A sandwich structure

Let's take a three-layered sandwich structure shown in Fig. 2 into consideration. The first and third layer are set as Si_3N_4, whose elastic modulus is 3.04×10^{11}Pa, Poisson's ratio is 0.27 and thermal expansion coefficient is $2.8\times10^{-6}/^{\circ}$C. The second layer is set as Cu, whose elastic modulus is 1.08×10^{11}Pa, Poisson's ratio is 0.33 and thermal expansion coefficient is $1.64\times10^{-5}/^{\circ}$C. The length and thickness of the structure respectively are $l=1$m and $w=0.01$m. $h_2=h_3=0.5$m and h_1 is kept decreasing. It will lead to the nearly singular integral in the SSBEM when h_1 is small. The elasticity problem and thermal elasticity problem are respectively considered here. In the elasticity problem, the uniform pressure $p=1\times10^{8}\,\mathrm{N/m^2}$ is loaded on the upper surface of the first layer. In the thermal elasticity problem, the temperature loading $\Delta T=1\,^{\circ}$C is subjected.

The physical quantities of interfacial point $A(0.5\mathrm{m},1\mathrm{m})$ with different h_1 are calculated by three different methods. The first one is the conventional SSBEM abbreviated as CSSBEM, in which the nearly singular integral is calculated with Gaussian's quadrature. The second one is the analytical SSBEM abbreviated as ASSBEM, in which the analytical formulation is applied to evaluate the nearly singular integral. The third one is the finite element method, which is introduced to provide the reference result.

The vertical edges are restrained along the *x*-axis and the bottom edge is restrained along the *y*-axis. For the CSSBEM and ASSBEM, 20 elements with each length 0.05m are discretized on each interface and horizontal edge. The left and right vertical edges in the BEM domain are discretized respectively by 2 linear elements and the whole structure is meshed with 84 linear elements in total. For the FEM, 8-node quadratic plane element is implemented in the FEM software ANSYS (Version 10.0). In order to harmonize the element size in vertical and horizontal direction in the FEM, the element number thus increases dramatically as the 1st layer's height h_1 decreases. Due to the limitation of Boolean operation tolerance in ANSYS, the FEM can only model the cases when $h_1\geq5.0\mathrm{e}\text{-}4\mathrm{m}$. Meanwhile, there are totally 26000 elements and 78481 nodes when $h_1=5.0\mathrm{e}\text{-}4\mathrm{m}$ in the FEM model.

For the elasticity problem, the comparison of quantities at point A by three different methods is listed in Table 1, from which it can be concluded that the results by the ASSBEM and CSSBEM have a good agreement with the one by the FEM. The results of displacement and stress by the ASSBEM keep steady from $h_1 = 0.05\text{m}$ to $h_1 = 5.0\text{e-}8\text{m}$. However, the results by the CSSBEM deteriorate when $h_1 = 5.0\text{e-}5\text{m}$.

Table 1 Displacement and stress at point *A* in elasticity problem

h_1(m)	u_y ($\times 10^{-2}$mm)			σ_{xx} (MPa)		
	ASSBEM	CSSBEM	FEM	ASSBEM	CSSBEM	FEM
5.0e-02	-0.565	-0.565	-0.565	-27.00	-27.00	-27.00
2.5e-02	-0.565	-0.565	-0.565	-27.00	-26.99	-27.00
5.0e-03	-0.565	-0.565	-0.565	-27.00	-27.03	-27.00
5.0e-04	-0.565	-0.565	-0.565	-27.00	-27.04	-27.00
5.0e-05	-0.565	-0.566	/	-27.00	-30.23	/
5.0e-06	-0.565	-0.743	/	-27.00	/	/
5.0e-07	-0.565	/	/	-27.00	/	/
5.0e-08	-0.565	/	/	-27.00	/	/

For the thermal elasticity problem, the physical quantities at interfacial point A calculated by three methods are compared in Table 2. Due to the computational limitation, the ANSYS can only analyze the structure for $h_1 \geq 5.0\text{e-}4\text{m}$. It can be observed from Table 2 that the result from the ASSBEM keeps steady from $h_1 = 0.05\text{m}$ to $h_1 = 5.0\text{e-}9\text{m}$, however, the CSSBEM can only obtain accurate results when $h_1 \geq 5.0\text{e-}5\text{m}$.

Table 2 Displacement and stress at point A in thermal elasticity problem

h_1(m)	u_y ($\times 10^{-2}$mm)			σ_{xx} (MPa)		
	ASSBEM	CSSBEM	FEM	ASSBEM	CSSBEM	FEM
5.0e-02	1.2684	1.2684	1.2684	-85.12	-85.12	-85.12
2.5e-02	1.2684	1.2684	1.2684	-85.12	-85.12	-85.12
5.0e-03	1.2684	1.2684	1.2684	-85.12	-85.12	-85.12
5.0e-04	1.2684	1.2684	1.2684	-85.12	-85.12	-85.12
5.0e-05	1.2684	1.2687	/	-85.12	-85.11	/
5.0e-06	1.2684	1.3358	/	-85.12	-82.59	/
5.0e-07	1.2684	/	/	-85.12	/	/

5.0e-08	1.2684	/	/	-85.12	/	/
5.0e-09	1.2684	/	/	-85.12	/	/

To investigate the evaluation ability of nearly singular integral, h_1 is set as 0.5m, and let y-coordinate of the inner point approach to the boundary.

For the elasticity problem, the displacement u_y and stress σ_{xx} at different inner point (0.5m, y) obtained by the ASSBEM, CSSBEM and FEM are listed in Table 3, where the displacement and stress at boundary point $B(0.5\text{m},1.5\text{m})$ are listed at the bottom line. The physical quantity of the interior point will converge towards the one at boundary point B when it is approaching to the boundary. For the CSSBEM, the stress is invalid when $y=1.475\text{m}$ and the vertical displacement loses the accuracy when $y=1.495\text{m}$. On the contrary, the results by ASSBEM approximately equal to the boundary value when the inner point is very close to the boundary.

Table 3 Displacement and stress at near boundary inner point in elasticity problem

y(m)	u_y ($\times10^{-2}$mm)			σ_{xx} (MPa)		
	ASSBEM	CSSBEM	FEM	ASSBEM	CSSBEM	FEM
1.400000	-0.687	-0.687	-0.687	-27.00	-27.00	-27.00
1.450000	-0.702	-0.702	-0.702	-26.99	-26.99	-27.00
1.475000	-0.710	-0.709	-0.710	-27.00	-39.81	-27.00
1.490000	-0.715	-0.727	-0.714	-27.00	/	-27.00
1.495000	-0.716	-0.731	-0.716	-27.00	/	-27.00
1.497500	-0.717	-0.596	/	-27.00	/	/
1.499500	-0.717	/	/	-27.00	/	/
1.499750	-0.717	/	/	-27.00	/	/
1.499950	-0.718	/	/	-27.00	/	/
1.499995	-0.718	/	/	-27.65	/	/
1.500000	-0.718	-0.718	-0.718	-27.00	-27.00	-27.00

For the thermal elasticity problem, u_y and σ_{xx} of different inner point (0.5m, y) are listed in Table 4, from which it can be observed that the thermal displacement by the CSSBEM is invalid when y is smaller than 1.495m, and the thermal stress loses accuracy when y is smaller than 1.45m. On the contrary, the results by the ASSBEM are accurate and stable because they

converge to the values of boundary point $B(0.5\text{m},1.5\text{m})$.

Table 4 Displacement and stress at near boundary inner point in thermal elasticity problem

y(m)	u_y ($\times 10^{-2}$mm)			σ_{xx} (MPa)		
	ASSBEM	CSSBEM	FEM	ASSBEM	CSSBEM	FEM
1.400000	1.4106	1.4106	1.4106	-85.12	-85.12	-85.12
1.450000	1.4284	1.4284	1.4284	-85.14	-85.14	-85.12
1.475000	1.4373	1.4364	1.4373	-85.12	-59.42	-85.12
1.490000	1.4426	1.4681	1.4426	-85.13	/	-85.12
1.495000	1.4444	1.4742	1.4444	-85.12	/	-85.12
1.497500	1.4453	1.2018	/	-85.11	/	/
1.499500	1.4460	/	/	-85.11	/	/
1.499750	1.4461	/	/	-85.11	/	/
1.499950	1.4462	/	/	-85.11	/	/
1.499995	1.4462	/	/	-83.77	/	/
1.500000	1.4462	1.4462	1.4462	-85.12	-85.12	-85.12

5.2 A functionally graded material structure

The functionally graded material (FGM) is delaminated as an 11-layered structure shown in Fig. 1, except that the 1st layer is controlled by the BEM, the other 10 layers are set as the SSM domain. The length, height and thickness of each layer are set as 1m, 0.1m and 0.01m, respectively. The elastic modulus varies as the function $E(i)=2^{4(1.1-y_u^i)}\times 10^{11}\,\text{Pa}$ and the thermal expansion coefficient of each layer varies as $\alpha(i)=i\times 10^{-7}/^{\circ}\text{C}$, where i denotes the ith layer. The Poisson's ratio of each layer is set as 0.3. The vertical edges are restrained along the x-axis and the bottom edge is restrained along the y-axis. The whole structure is subjected to $1\,^{\circ}\text{C}$ temperature increase. The contour plot of displacement u_y by the ASSBEM and FEM are compared in Fig. 3.

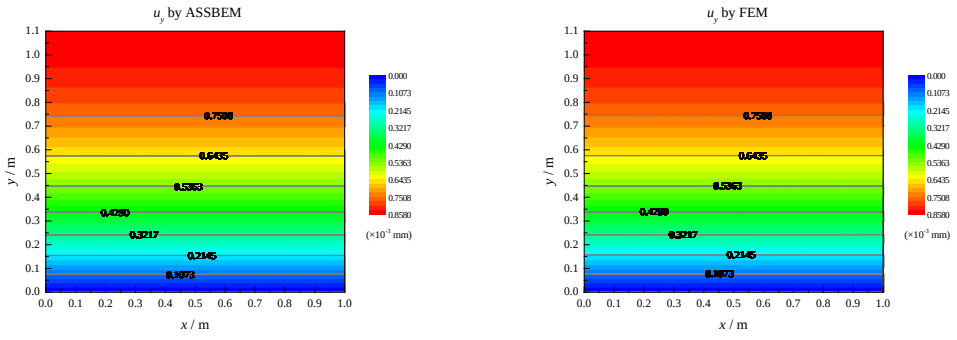


Fig.3 Comparison of contour plot of displacement u_y by ASSBEM and FEM

It can be observed from Fig. 3 that the results by the ASSBEM agree well with the ones by the FEM. The structure can freely expand along the *y*-axis according to the boundary condition. It can be deduced that the maximum vertical displacement should appear on the upper edge. Definitely, the maximum vertical displacement occurs on the upper edge in Fig. 3, which is 0.858×10^{-3} mm.

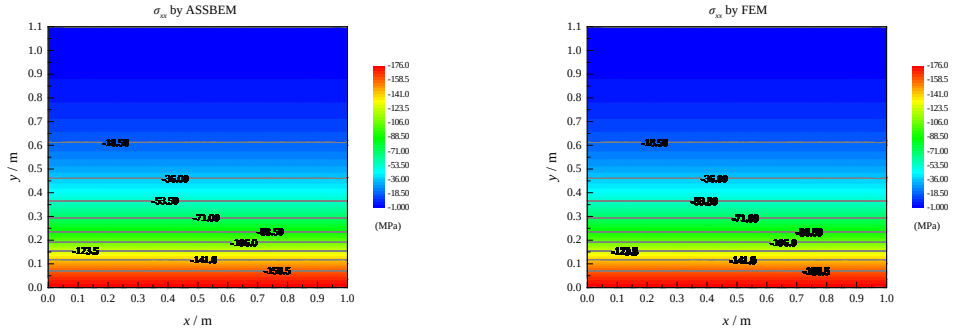


Fig.4 Comparison of contour plot of stress σ_{xx} by ASSBEM and FEM

The contour plot of stress σ_{xx} obtained by ASSBEM is compared with the one by FEM in Fig. 4, from which it can be observed that the results by these two methods have a good agreement with each other. It also can be found that the absolute σ_{xx} decreases along *y*-axis since the thermal expansion coefficient gets smaller in this direction. The absolute minimum σ_{xx} occurs on the upper edge which is 1 MPa, while the absolute maximum σ_{xx} appears on the bottom which is 176 MPa.

6. Conclusions

The state space boundary element method is established by coupling the state space method and the boundary element method, in which the analytical formulation of nearly singular integral is introduced. It is especially suitable for calculating the mechanical quantity of the near boundary inner point in the laminated structure. The accuracy of the present method is verified by the finite element method. The present results for the displacement and stress at near boundary inner point are approaching to the boundary values when the ANSYS fails to provide the reference results.

Acknowledgments

This work was supported by the National Natural Science Foundation of China, China (No.11372094).

References

[1] Fan, J. R., Ye, J. Q. (1990) An exact solution for the statics and dynamics of laminated thick plates with orthotropic layers, *International Journal of Solids and Structures* **26**, 655-662.

[2] Attallah, K. M. Z., Ye, J. Q., Sheng, H. Y. (2007) Three-dimensional finite strip analysis of laminated panels, *Computers & Structures* **85**, 1769-1781.

[3] Attallah, K. M. Z., Yu, M., Ye, J. Q. (2014) Three-dimensional state space spline finite strip analysis of angle-plied laminates, *Composites Part B: Engineering* **66**, 25-35.

[4] Khezri, M., Gharib, M., Bradford, M. A., Vrcelj, Z. (2015) Analysis of thick and orthotropic rectangular laminated composite plates using a state-space-based generalised RKP-FSM, *Composite Structures* **133**, 691-706.

[5] Sheng, H. Y., Ye, J. Q. (2002) A semi-analytical finite element for laminated composite plates, Composite Structures **57**, 117-123.

[6] Sheng, H. Y., Ye, J. Q. (2002) A state space finite element for laminated composite plates, *Computer Methods in Applied Mechanics and Engineering* **191**, 4259-4276.

[7] Ye, J. Q., Sheng, H. Y., Qin, Q. H. (2004) A state space finite element for laminated composites with free edges and subjected to transverse and in-plane loads, *Computers & structures* **82**, 1131-1141.

[8] Jodaei, A., Jalal, M., Yas, M. H. (2012) Free vibration analysis of functionally graded annular plates by state-space based differential quadrature method and comparative modeling by ANN, *Composites Part B: Engineering* **43**, 340-353.

[9] Wu, C. P., Jiang, R. Y. (2012) A state space differential reproducing kernel method for the 3D analysis of FGM sandwich circular hollow cylinders with combinations of simply-supported and clamped edges, *Composite Structures* **94**, 3401-3420.

[10] Sheng, H. Y., Ye, J. Q. (2005) State space solution for axisymmetric bending of angle-ply laminated cylinder with clamped edges, *Composite structures* **68**, 119-128.

[11] Khezri, M., Gharib, M., Bradford, M. A., Uy, B. (2015) A state space augmented generalised RKPM for three-dimensional analysis of thick and laminated composite plates, *Composite structures* **158**, 225-239.

[12] Gu, J., Zhang, J., Li, G. (2012) Isogeometric analysis in BIE for 3-D potential problem, *Engineering Analysis with Boundary Elements* **36**, 858-865.

[13] Simpson, R. N., Bordas, S. P. A., Trevelyan, J., Rabczuk, T. (2012) A two-dimensional isogeometric boundary element method for elastostatic analysis, *Computer Methods in Applied Mechanics and Engineering* **209**, 87-100.

[14] Zheng, B., Gao, X., Zhang, C. (2016) Radial integration BEM for vibration analysis of two-and three-dimensional elasticity structures, *Applied Mathematics and Computation* **277**, 111-126.

[15] Tsiatas, G. C., Yiotis, A. J. (2013) A BEM-based meshless solution to buckling and vibration problems of orthotropic plates, *Engineering Analysis with Boundary Elements* **37**, 579-584.

[16] Zhang, Y. M., Gu, Y., Chen, J. T. (2011) Stress analysis for multilayered coating systems using semi-analytical BEM with geometric non-linearlyities, *Computational Mechanics* **47**, 493-504.

[17] Cheng, C., Han, Z., Yao, S., Niu, Z., Recho, N. (2014) Analysis of heat flux singularity at 2D notch tip by singularity analysis method combined with boundary element technique, *Engineering Analysis with Boundary Elements* **46**, 1-9.

[18] Aour, B., Rahmani, O., Nait-Abdelaziz, M. (2007) A coupled FEM/BEM approach and its accuracy for solving crack problems in fracture mechanics, *International Journal of Solids and Structures* **44**, 2523-2539.

[19] Padrón, L. A., Aznárez, J. J., Maeso, O. (2007) BEM-FEM coupling model for the dynamic analysis of piles and pile groups, *Engineering Analysis with Boundary Elements* **31**, 473-484.

[20] Bonnet, G., Seghir, A., Corfdir, A. (2009) Coupling BEM with FEM by a direct computation of the boundary stiffness matrix, *Computer Methods in Applied Mechanics and Engineering* **198**, 2439-2445.

[21] Benedetti, I., Aliabadi, M. H., Milazzo, A. (2010) A fast BEM for the analysis of damaged structures with bonded piezoelectric sensors, *Computer Methods in Applied Mechanics and Engineering* **199**,

490-501.

[22] Lin, C. C., Lawton, E. C., Caliendo, J. A., Anderson, L. R. (1996) An iterative finite element-boundary element algorithm, *Computers & Structures* **59**, 899-909.

[23] Boumaiza, D., Aour, B. (2014) On the efficiency of the iterative coupling FEM–BEM for solving the elasto-plastic problems, *Engineering Structures* **72**, 12-25.

[24] Njiwa, R. K., Niane, N. T., Frey, J., Schwartz, M., Bristiel, P. (2015) A coupling strategy of FEM and BEM for the solution of a 3D industrial crack problem, *International Journal for Computational Methods in Engineering Science and Mechanics* **16**, 112-120.

[25] Niu, Z. R., Cheng, C. Z., Zhou, H. L., Hu, Z. J. (2007) Analytic formulations for calculating nearly singular integrals in two-dimensional BEM, *Engineering Analysis with Boundary Elements* **31**, 949-964.

[26] Zhong, W. X., Williams, F. W. (1994) A precise time step integration method, *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science* **208**, 427-430.

Analytical Computation for Turbulence and Enstrophy in the Taylor Green Vortex Model

Kaveh Fardipour[1], †Kamyar Mansour[1, 2]

[1] Department of Aerospace Engineering, Amirkabir University of Technology, 424 Hafez Ave, Tehran, Iran.
[2] Flow Research and Engineering Company, USA.

†Corresponding author: mansour@aut.ac.ir

Abstract

Using a classic example proposed by G. I. Taylor, we reconsider through the use of computer algebra, the mathematical analysis of a fundamental process in turbulent flow, namely: How do large scale eddies evolve into smaller scale ones to the point where they are effectively absorbed by viscosity? The explicit symbolic series solution of this problem, even for cleverly chosen special cases, requires daunting algebra, and so numerical methods have become quite popular. Yet an algebraic approach can provide substantial insight, especially if it can be pursued with modest human effort. The specific example we use dates to 1937 when Taylor and Green first published a method for explicitly computing successive approximations to formulas describing the three-dimensional evolution over time of what is now called a Taylor-Green vortex. With the aid of a computer algebra system, we have duplicated Taylor and Green's efforts and obtained more detailed time-series results. We have extended their approximation of the energy dissipation from order 5 in time to order 8.

Keywords: nonlinear equations, symbolic calculations, vortices

Introduction

The fundamental dynamical mechanism involved in homogeneous three dimensional turbulent flows is the enhancement of vorticity by vortex-line stretching and the consequent production of small-scale eddies. This process controls the turbulent-energy dynamics and hence the global structure and evolution of the flow. A model of this process is given by the Taylor-Green vortex ref [1], which is perhaps the simplest system in which to study the generation of small scales and the resulting turbulence. In the past this flow has been used to study such questions as: (i) enhancement of vorticity by vortex line stretching (ii) approach to isotropy of the small scales ref [2] (iii) possible singular behaviour of solutions of the Euler equations ref [3] and there are other issues of this flow discussed in ref [4-7].

Series Derivation and Computer Extension

The flow discussed at length in this paper is that for which the initial flow has two-dimensional streamlines, but the flow is three-dimensional for all $t > 0$.

This section is an initial statement and development of Taylor's model [1]. The initial flow of an incompressible fluid is chosen as represented by the following equations:

$$u = A\cos(ax)\sin(by)\sin(cz)$$
$$v = B\sin(ax)\cos(by)\sin(cz)$$
$$w = C\sin(ax)\sin(by)\cos(cz)$$

The Taylor-Green vortex is that three-dimensional flow that develops from the single-Fourier mode initial condition .Using the fact that ρ, the density, is constant for an incompressible fluid, and the equation of continuity:

$$\rho_t + (\rho u)_x + (\rho v)_y + (\rho w)_z = 0 \tag{1}$$

We can derive the following consistency equations for our model:

$$u_x + v_y + w_z = 0$$
$$Aa + Bb + Cc = 0$$

The equations of motion are:

$$-u_t = uu_x + vu_y + wu_z + \frac{P_x}{\rho}$$
$$-v_t = uv_x + vv_y + wv_z + \frac{P_y}{\rho} \tag{2}$$
$$-w_t = uw_x + vw_y + ww_z + \frac{P_z}{\rho}$$

$$-\frac{1}{\rho}\nabla^2 P = u_x^2 + v_y^2 + w_z^2 + 2(v_x u_y + w_y v_z + u_z w_x) \tag{3}$$

It is possible to solve these equations (1), (2) and (3) as a regular expansion in the time

$$u = \sum_{i=0}^{\infty} U_i t^i$$
$$v = \sum_{i=0}^{\infty} V_i t^i \tag{4}$$
$$w = \sum_{i=0}^{\infty} W_i t^i$$
$$(\frac{P}{\rho}) = \sum_{i=0}^{\infty} (\frac{P}{\rho})_i t^i$$

We assume A=a=b=c=1 and B= -1, which leads to:

$$U_0 = \cos(x)\sin(y)\sin(z)$$
$$V_0 = -\sin(x)\cos(y)\sin(z)$$
$$W_0 = 0$$

Then substituting (4) in the equations (1), (2) and (3) and equating like powers of time t yields this sequence of successive linear equations, together with boundary conditions we obtain:

$$U_1 = -\frac{1}{8}\sin(2x)\cos(2z) \qquad V_1 = -\frac{1}{8}\sin(2y)\cos(2z)$$
$$W_1 = \frac{1}{8}\sin(2z)\cos(2x) + \frac{1}{8}\sin(2z)\cos(2y)$$

$$(\frac{P}{\rho})_1 = \frac{1}{44}\sin(x)\sin(3y)\sin(z)-\frac{1}{44}\sin(3x)\sin(y)\sin(z)$$

$$U_2 = -\frac{15}{352}\sin(3y)\sin(z)\cos(x)-\frac{5}{176}\sin(y)\sin(z)\cos(3x)+\frac{1}{32}\sin(y)\sin(3z)\cos(x)$$

$$-\frac{1}{32}\cos(x)\sin(y)\sin(z)+\frac{1}{32}\sin(y)\sin(3z)\cos(3x)$$

$$\mathrm{V}_2 = \frac{15}{352}\sin(z)\sin(3x)\cos(y)+\frac{5}{176}\sin(z)\sin(x)\cos(3y)-\frac{1}{32}\sin(3z)\sin(x)\cos(y)$$

$$+\frac{1}{32}\sin(x)\cos(y)\sin(z)-\frac{1}{32}\sin(3z)\sin(x)\cos(3y)$$

$$\mathrm{W}_2 = -\frac{15}{352}\sin(x)\sin(3y)\cos(z)+\frac{15}{352}\sin(3x)\sin(y)\cos(z)-\frac{1}{32}\sin(3x)\sin(y)\cos(3z)$$

$$+\frac{1}{32}\sin(x)\sin(3y)\cos(3z)$$

$$(\frac{P}{\rho})_2 = \frac{1}{128}\cos(2x)\cos(2y)\cos(4z)+\frac{21}{1408}\cos(2x)\cos(2y)+\frac{21}{1408}\cos(2y)+\frac{21}{1408}\cos(2x)+$$

$$\frac{1}{8448}\cos(2x)\cos(4y)\cos(2z)-\frac{1}{1408}\cos(2x)\cos(4y)-\frac{1}{220}\cos(4y)\cos(2z)$$

$$-\frac{5}{264}\cos(2x)\cos(2y)\cos(2z)+\frac{1}{8448}\cos(4x)\cos(2y)\cos(2z)-\frac{1}{220}\cos(2z)\cos(4x)$$

$$-\frac{1}{1408}\cos(4x)\cos(2y)-\frac{17}{2816}\cos(2y)\cos(2z)-\frac{17}{2816}\cos(2x)\cos(2z)+\frac{1}{128}\cos(4z)$$

$$+\frac{31}{2816}\cos(4x)+\frac{31}{2816}\cos(4y)$$

$$U_3 = \frac{1}{384}\cos(2y)\sin(4x)\cos(4z)-\frac{19}{3168}\cos(2y)\cos(2z)\sin(2x)-\frac{1}{198}\cos(2y)\cos(2z)\sin(4x)$$

$$+\frac{49}{12672}\sin(2x)\cos(2z)\cos(4y)-\frac{1}{192}\sin(4x)\cos(4z)-\frac{1}{192}\sin(2x)\cos(4z)-\frac{1}{192}\sin(2x)\cos(4y)$$

$$+\frac{27}{3520}\cos(2z)\sin(4x)+\frac{1}{384}\cos(2y)\sin(4x)+\frac{43}{4224}\sin(2x)\cos(2z)+\frac{1}{96}\cos(2y)\sin(2x)\cos(4z)$$

$$\mathrm{V}_3 = \frac{1}{384}\cos(2x)\sin(4y)\cos(4z)-\frac{19}{3168}\cos(2x)\cos(2z)\sin(2y)-\frac{1}{198}\cos(2x)\cos(2z)\sin(4y)$$

$$+\frac{49}{12672}\sin(2y)\cos(2z)\cos(4x)-\frac{1}{192}\sin(4y)\cos(4z)-\frac{1}{192}\sin(2y)\cos(4z)-\frac{1}{192}\sin(2y)\cos(4x)$$

$$+\frac{27}{3520}\cos(2z)\sin(4y)+\frac{1}{384}\cos(2x)\sin(4y)+\frac{43}{4224}\sin(2y)\cos(2z)+\frac{1}{96}\cos(2x)\sin(2y)\cos(4z)$$

$$\mathrm{W}_3 = -\frac{1}{384}\cos(2y)\cos(4x)\sin(4z)+\frac{19}{1584}\cos(2x)\cos(2y)\sin(2z)+\frac{79}{12672}\cos(4x)\cos(2y)\sin(2z)$$

$$+\frac{79}{12672}\cos(2x)\cos(4y)\sin(2z)+\frac{1}{192}\cos(4x)\sin(4z)\frac{1}{384}\cos(2x)\sin(4z)-\frac{27}{1760}\sin(2z)\cos(4x)$$

$$-\frac{43}{4224}\sin(2z)\cos(2x)-\frac{1}{96}\cos(2x)\cos(2y)\sin(4z)-\frac{1}{384}\cos(2x)\cos(4y)\sin(4z)+\frac{1}{192}\cos(4y)\sin(4z)$$

$$+\frac{1}{384}\cos(2y)\sin(4z)-\frac{27}{1760}\cos(4y)\sin(2z)-\frac{43}{4224}\sin(2z)\cos(2y)$$

$$(\frac{P}{\rho})_3 = \frac{1}{640}\sin(x)\sin(3y)\sin(5z) - \frac{1}{640}\sin(3x)\sin(y)\sin(5z) + \frac{1}{31680}\sin(3x)\sin(5y)\sin(z)$$
$$-\frac{1}{31680}\sin(5x)\sin(3y)\sin(z) - \frac{1}{544896}\sin(3x)\sin(5y)\sin(3z) + \frac{689}{300960}\sin(3x)\sin(y)\sin(3z)$$
$$-\frac{689}{300960}\sin(x)\sin(3y)\sin(3z) + \frac{1}{544896}\sin(5x)\sin(3y)\sin(3z) + \frac{97}{246400}\sin(x)\sin(5y)\sin(3z)$$
$$-\frac{1229}{190080}\sin(x)\sin(5y)\sin(z) - \frac{4337}{696960}\sin(x)\sin(3y)\sin(z) + \frac{1229}{190080}\sin(5x)\sin(y)\sin(z)$$
$$-\frac{97}{246400}\sin(5x)\sin(y)\sin(3z) + \frac{4337}{696960}\sin(3x)\sin(y)\sin(z)$$

The mean rate of energy dissipation is:

$$\bar{W} = \mu(\overline{\xi^2 + \eta^2 + \zeta^2}) = \mu\overline{\omega^2},$$

where ξ, η and ζ are components of vorticity ω = curl V, where V is the velocity vector and its x, y and z components are respectively u, v and was defined previously. For enstrophy

$$\overline{\omega^2} = \frac{3}{4} + \frac{5t^2}{64} + \frac{25t^4}{4224} - \frac{52439t^6}{188334080} + \frac{729080006648669t^8}{11942691907829760000}$$

Symbolic Computations by Different Software

For symbolic computation of the solution of the Taylor-Green vortex, we developed three different codes for three different symbolic computation software, namely, Maple, Maxima and Mathematica. Our first attempt was a direct computation with Maple without any considerable optimization of computational routines. This code was very slow and consumed a large amount of memory. Later we developed a new code using Maxima. As it was mentioned by Fateman [11] for his own Macsyma code, using Poisson series and its related commands in Maxima was a very important factor for increasing speed of computations and decreasing memory consumption but it did not have a good performance at higher powers of time and the programed crashed when number of terms and memory usage increased. We also developed a Mathematica code for solution of Taylor-Green vortex. This code showed the best performance between all of our codes did not have any of aforementioned problems of other software.

Conclusions

The time series expansion of the Euler equation for the unsteady three dimensional flow of an inviscid, incompressible, of Taylor' model, are obtained to 8 terms exactly by means of symbolic calculations and compared well with results of ref [11]. For this symbolic computations we used Maple, Maxima and Mathematica software and by comparing the performance of our codes in these three different software, Mathematica showed the best performance for symbolic computation of Taylor-Green vortex.

References

[1] G.I. Taylor, A.E. Green, Mechanism of the production of small eddies from large ones, Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 158 (1937) 499-521.
[2] S.A. Orszag, Numerical simulation of the Taylor-Green vortex, in: R. Glowinski, J.L. Lions (Eds.), Computing Methods in Applied Sciences and Engineering Part 2. Springer, Berlin, Heidelberg, 1974. pp. 50-64.
[3] S.A. Orszag, Design of large hydrodynamic codes, in: J.M. Ortega (Eds.), Computer Science and Scientific Computing. Academic Press, New York, 1976. pp. 191-204.

[4] S.A. Orszag, Statistical theory of turbulence, in: R. Balian, J.L. Peube (Eds.), Fluid Dynamics Les Houches Summer School 1973, Gordon and Breach, London, 1977, pp. 235-374.
[5] P.G. Saffman, Lectures on homogeneous turbulence, in: N.J. Zabusky (Eds.), Topics in Non-Linear Physics, Springer, Berlin, 1968, pp. 485-614.
[6] M.E. Brachet, D.I. Meiron, S.A. Orszag, B.G. Nickel, R.H. Morf, U. Frisch, Small-scale structure of the Taylor–Green vortex, J. Fluid Mech., 130 (1983), 411-452.
[7] M.E. Brachet, D.I. Meiron, S.A. Orszag, B.G. Nickel, R.H. Morf, U. Frisch, The Taylor-Green vortex and fully developed turbulence, Journal of Statistical Physics, 34.5 (1984), 1049-1063.
[8] G.A. Baker, Essentials of Pade Approximants, Academic Press, New York, 1975.
[9] K. Mansour, Laminar flow through a slowly rotating straight pipe, J. Fluid Mech. 150 (1985), 1-21.
[10] K. Mansour, Using Stokes expansion for natural convection inside a two-dimensional cavity, Fluid Dynamics Research, 12.1 (1993), 1-33.
[11] R.J. Fateman, Symbolic computation of turbulence and energy dissipation in the Taylor vortex model, International Journal of Modern Physics, C 9.03 (1998), 509-525.

Optimized Compact Scheme with High Order of Accuracy Using Maximum Norm

Kaveh Fardipour [1], † Kamyar Mansour [1]
[1] Amirkabir University of Technology, 424 Hafez Ave, Tehran, Iran.
†Corresponding author: mansour@aut.ac.ir

Abstract
Compact finite difference schemes have high order of accuracy and high numerical resolution. Because of such characteristics, these schemes are suitable for numerical simulation of multi-scale phenomena, like direct numerical simulation, large eddy simulation and computational aeroacoustics. In this paper, we intend to optimize pentadiagonal compact schemes with a seven point stencil, using maximum error norm with different error threshold. The second goal of this paper is to increase the order of accuracy of the optimized schemes, because optimized schemes with higher order of accuracy resolve large scales better than the lower order ones. Finally, by using numerical experiments, we investigate the order of accuracy and numerical resolution of the new optimized compact schemes.

Keywords: Compact finite difference scheme, Optimized scheme, Numerical dispersion, Maximum norm, High resolution, Direct numerical simulation, Computational aeroacoustics.

Introduction

Multi-scale phenomena are common in many physical and engineering fields like aerodynamics and astrophysics. Turbulence is an important of multi-scale phenomena and direct numerical simulation (DNS) and large eddy simulation (LES) are useful tools to study this problem and understand its underlying physical mechanisms. Another example of such phenomena is aerodynamically generated noise and computational aeroacoustics (CAA). Multi-scale phenomena contain a wide range of time- and length-scales and the largest scale can be orders of magnitude larger than the smallest scales, for example in aeroacoustics problems amplitude of acoustic waves are very small in comparison to mean flow and sound intensity can be five or six order of magnitude smaller [18].

To correctly resolve small scales in numerical simulation of multi-scale phenomena, numerical scheme should have high order of accuracy and high numerical resolution. Another issue in numerical simulation of multi-scale phenomena marginally resolved scales because error from such scales may affect whole computation. To resolve these scales more accurately, the numerical resolution at higher wavenumbers should be increased and optimization of numerical error is a useful method to achieve this goal, but it comes at the expense the order of accuracy of the scheme. Compact schemes are popular and widely used high order and high resolution numerical schemes. Application of compact schemes in computational fluid dynamics (CFD) can be traced back to 1970s [1-4]. Lele [8] proposed a series of compact schemes for derivatives, filtering and interpolation and by using Fourier analysis, showed their spectral-like resolution for a certain range of wavenumbers. Lele also proposed equating modified wavenumber and wavenumber at certain

wavenumbers as a method to optimize dispersion error of compact schemes. Tam and Webb [9] proposed the dispersion relation preserving (DRP) scheme and they used an integrated error to form a minimization problem to optimize the dispersion error of an explicit finite difference scheme with a seven point stencil. Kim and Lee [10] used DRP idea to a seven point pentadiagonal compact scheme with second and fourth order of accuracy and later they [11] developed a series of optimized boundary stencils for their optimized compact scheme. Zhuang and Chen [15] developed an optimized fourth order upwind DRP scheme for CAA applications. Bogey and Bailly [19] and Tam [18] extended the stencil size of DRP scheme from seven to nine, eleven, thirteen and fifteen points and the increase in the size of stencil resulted in more numerical resolution at higher wavenumbers.

Chu and Fan [13] and Mahesh [14] independently introduced combined compact schemes. This new family of compact schemes compute the first and second derivatives together and has better numerical resolution. Lui and Lele [16] used Lele [] optimization method to develop a sixth order pentadiagonal compact scheme. Ashcroft and Zhang [17] optimized a five point, fourth order prefactored compact scheme with two independent bidiagonal matrices. Jordan [21] introduced a new method for optimization of compact scheme. Instead of using a Fourier analysis this new method uses a composite template and unlike the Fourier transform, this template analysis the resolution of the whole matrix, including the boundary schemes, and gives a sets of pseudo-wavenumbers for each point. Jordan used these sets and a least square optimization to determine the coefficients of the new optimized scheme and later used the same method to optimize a pentadiagonal compact scheme [26].

Kim [20] introduced a new fourth order optimized pentadiagonal compact scheme and used a combination of polynomials and trigonometric series to formulate extrapolation function for optimization of new boundary schemes with better resolution. Later, Kim and Sandberg [27], Haeri and Kim [28] and Turner et al. [33] introduced new methods for optimization of boundary schemes for fourth order optimized pentadiagonal compact schemes and efficient parallel computation using domain decomposition. Liu et al. [22] used a sequential quadratic programming (SQP) to optimize a fourth order pentadiagonal compact scheme and its boundary schemes. Following Holberg [7], Venutelli [24] used minimization of the group velocity error to optimize fourth and sixth order staggered pentadiagonal schemes. Zhou and Zhang [25] introduced an optimized prefactored compact scheme for the second derivative. Yu et al. [31] developed a fifth order optimized version of combined compact scheme [13-14] to solve advection equation.

Recently, Zhang and Yao [29] used simulated annealing algorithm [6] to optimize zeroth order explicit finite difference scheme with a specified error threshold for maximum norm of error. They showed by this new method one can optimize error more flexibly and increase the cutoff wavenumber. Cunha and Redonnet [30] investigated effect of reduced order of accuracy on performance of DRP schemes. They showed the reduction of the order of accuracy may cause error in larger scales and therefore is harmful when such scales are important in final analysis. To solve this problem they proposed the usage of optimized schemes with higher order of accuracy and

later, they proposed a new approach for optimization of explicit finite difference schemes and two optimized scheme with sixth and eighth order of accuracy [32].

In this paper we intend to optimize compact pentadiagonal finite difference schemes by using maximum error norm and increase their order of accuracy to six and eight. In the second section of paper, we introduce the compact pentadiagonal finite difference scheme and analyze its numerical dispersion error by using Fourier analysis. Later we introduce optimization procedure by using maximum error norm and use this procedure to determine the coefficients of the optimized scheme. In the third section we use numerical experiments to assess the order of accuracy and numerical resolution of the optimized schemes.

Numerical Scheme

A compact pentadiagonal scheme with a seven point stencil is a linear combination as,

$$\beta f'_{i-2} + \alpha f'_{i-1} + f'_i + \alpha f'_{i+1} + \beta f'_{i+2} = \frac{a_1}{h}(f_{i+1} - f_{i-1}) + \frac{a_2}{h}(f_{i+2} - f_{i-2}) + \frac{a_3}{h}(f_{i+3} - f_{i-3}). \quad (1)$$

Tridiagonal and explicit approximation can be achieved in Eq. (1) by setting $\beta = 0$ and $\beta = \alpha = 0$ respectively. By matching Taylor series coefficients of different truncation error terms, one can determine the relations between coefficients of Eq. (1) these relations are:

$$2\alpha + 2\beta = 2a_1 + 4a_2 + 6a_3 - 1, \quad (2)$$

$$\alpha + 4\beta = \frac{1}{3}a_1 + \frac{8}{3}a_2 + 9a_3, \quad (3)$$

$$\frac{1}{12}\alpha + \frac{4}{3}\beta = \frac{1}{60}a_1 + \frac{8}{15}a_2 + \frac{81}{20}a_3, \quad (4)$$

$$\frac{1}{360}\alpha + \frac{8}{45}\beta = \frac{1}{2520}a_1 + \frac{16}{315}a_2 + \frac{243}{280}a_3, \quad (5)$$

$$\frac{1}{20160}\alpha + \frac{4}{315}\beta = \frac{1}{181440}a_1 + \frac{8}{2835}a_2 + \frac{243}{2240}a_3. \quad (6)$$

Solving Eqs. (2)-(6) gives a unique set of coefficients for a tenth-order pentadiagonal compact finite difference scheme.

Error analysis by Fourier transformation

Fourier analysis is an effective tool for analyzing and comparing numerical error of different finite difference and its application for such analysis is extensively described in [8]. Fourier analysis is also an effective method to quantify numerical resolution of different numerical schemes and its results can be used for error optimization [8, 9].
Fourier transformation and its inverse have the following form

$$\tilde{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\, e^{-Ikx} dx, \quad (7)$$

$$f(x) = \int_{-\infty}^{\infty} \tilde{f}(k)\, e^{-Ikx} dk. \tag{8}$$

$\tilde{f}(k)$ is the Fourier transformation of $f(x)$, k is the wavenumber and $I = \sqrt{-1}$. By taking Fourier transformation of Eq. (1), modified wavenumber is obtained as

$$w' = 2\frac{a_1 \sin(w) + a_2 \sin(2w) + a_3 \sin(3w)}{1 + 2\alpha \cos(w) + 2\beta \cos(2w)}, \tag{9}$$

where $w = k\Delta x$ and $w' = k'\Delta x$ are scaled wavenumber and scaled modified wavenumber, respectively. To measure the numerical error, $|w' - w|$ and $\frac{|w'-w|}{w}$ can be defined as the absolute and relative measures of error, respectively [8, 29].

Optimization procedure

Although compact schemes have good numerical resolution in low wavenumber range, but their resolution deteriorates as wavenumber increases. One can use an optimization procedure to increase the numerical resolution of the numerical scheme. To achieve this goal, the scaled modified wavenumber should a give a better approximation of scaled wavenumber at high wavenumbers. Tam and Webb suggested using an integrated error to form a minimization problem for dispersion error and solving this problem will lead to reduced dispersion error in the integration range.

Lele [8] proposed an alternative method for optimization of dispersion error. In this method instead of using integrated error, w' sets to be equal to w at certain wavenumbers, this will lead to new equations as

$$2a_1 \sin(w_j) + 2a_2 \sin(2w_j) + 2a_3 \sin(3w_j) - 2\alpha w_j \cos(w_j) - 2\beta w_j \cos(2w_j) = w_j. \tag{10}$$

By solving these new equations, one can determine some or all of coefficients in Eq. (1) and desired goal for optimization can be achieved by try and error. Zhang and Yao [29] used this method and a simulated annealing algorithm [6] to determine the coefficients of the optimized finite difference scheme by the error tolerance threshold, ε, as the maximum norm of the dispersion error.

In this paper, we used Lele's method [8] for optimization of the numerical error. We found try and error is a fast and easy way to optimize the compact finite difference scheme by maximum norm of error and therefore we did not use any other extra algorithm for optimization, like the usage of simulated annealing algorithm by Zhang and Yao [29]. Optimization of numerical comes at expense of the order of accuracy, because equations like Eq. (10) should be used to determine some of coefficients in Eq. (1) instead of using Eqs. (2)-(6). Cunha and Redonnet [30] showed this reduction of the order of accuracy can result in considerable error in lower wavenumber range and therefore optimized schemes with increased order of accuracy may be needed. Table 1 to Table 3

show coefficients of optimized fourth, sixth and eighth order scheme with different error thresholds, and we determined these coefficients by using Eqs. (2)-(5) and Eq. (10).

Table 1 Coeffients of fourth order optimized scheme with different error treshold

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
α	0.6082258801733940	0.5799409399256260	0.5562135732927600
β	0.1119841895552910	0.0913025587073633	0.0767799831080962
a_1	0.6226500993288040	0.6488542583945110	0.6691663401223610
a_2	0.2832201917744050	0.2514903896275570	0.2253610769453380
a_3	0.0103731956170232	0.0064694869944544	0.0043683541292728

Table 2 Coeffients of sixth order optimized scheme with different error treshold

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
α	0.5763671534480410	0.5538290120198530	0.5365894666122440
β	0.0873786847225768	0.0746214733595378	0.0659778177265754
a_1	0.6523994549163750	0.6717832954024240	0.6847527593400610
a_2	0.2476871227189510	0.2224843288821100	0.2043581271773250
a_3	0.0053240459387800	0.0038995107375816	0.0030327568813695

Table 3 Coeffients of eighth order optimized scheme with different error treshold

	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$
α	0.5367930568824080	0.5241025757235620	0.5157437214861430
β	0.0647172227529633	0.0596410302894250	0.0562974885944574
a_1	0.6868707168185950	0.6942734974945880	0.6991494957997490
a_2	0.2031640938486790	0.1911504383516390	0.1832373896735490
a_3	0.0027704583731389	0.0023897439383735	0.0021389783112510

Figure 1 to Figure 3 shows the absolute error of w' with different error thresholds for fourth, sixth and eighth order optimized schemes, respectively. We also plot the absolute error of w' for the original tenth order compact scheme. By comparing plots in different figures, we can conclude each optimized scheme has better resolution than the original scheme in higher wavenumbers.

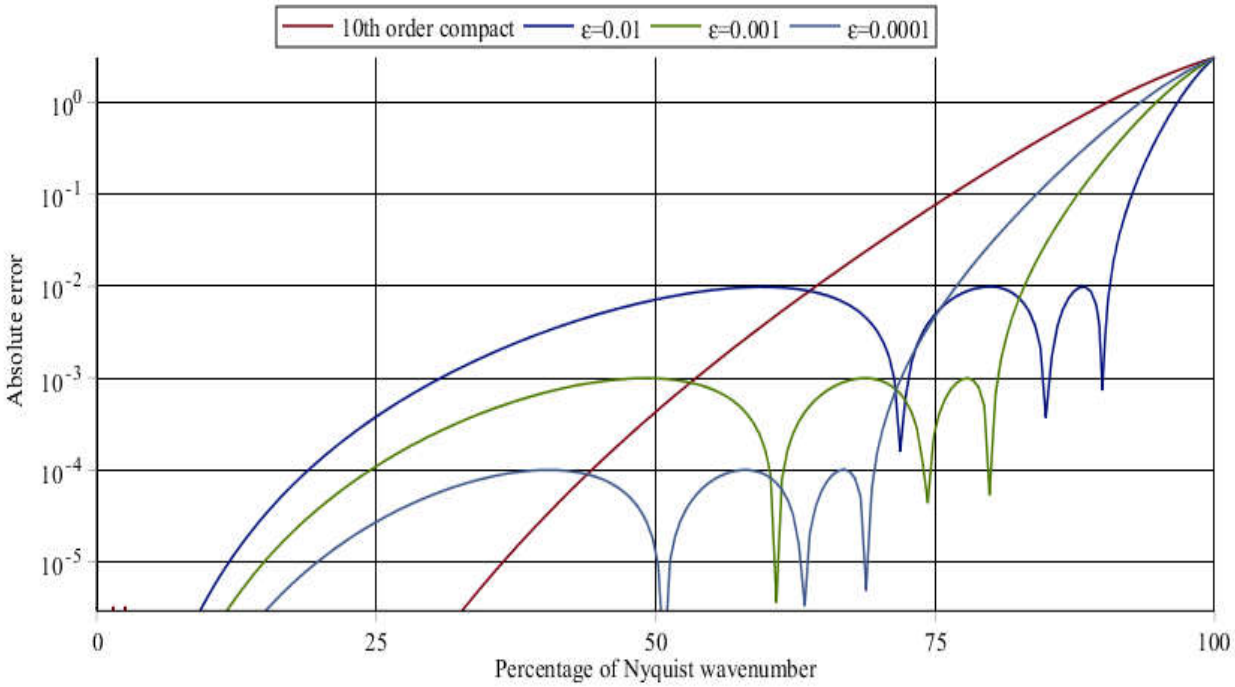


Figure 1 absolute error of w' for optimized fourth order schemes with different error threshold.

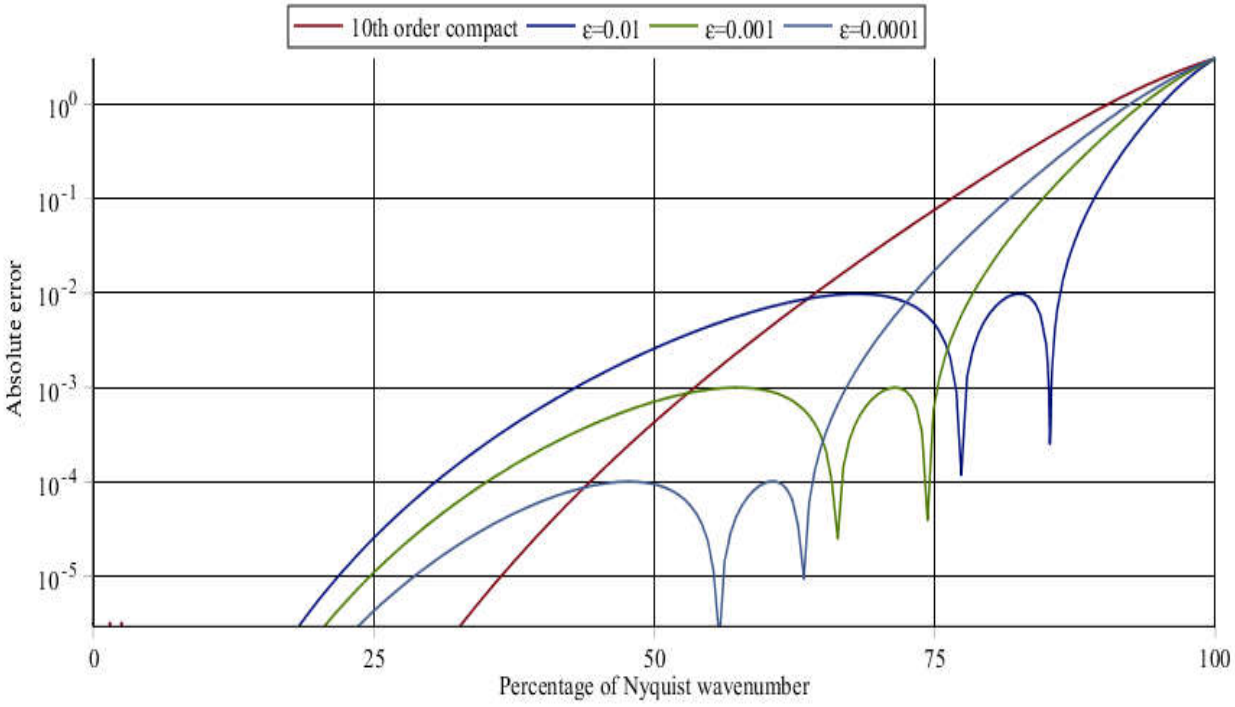


Figure 2 absolute error of w' for optimized sixth order schemes with different error threshold.

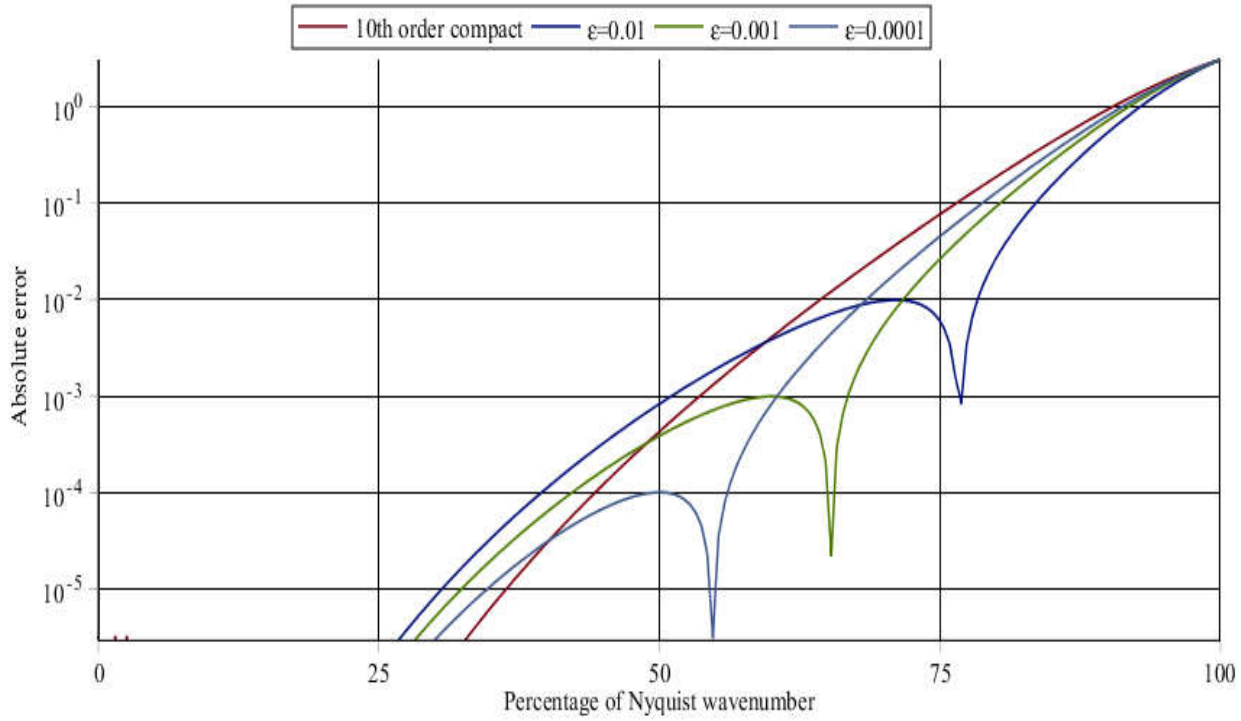


Figure 3 absolute error of w' for optimized eighth order schemes with different error threshold.

Numerical results

In this section we use numerical experiments to test the order of accuracy. To achieve this goal, we numerically solve a wave equation as

$$\frac{\partial u}{\partial t}+\frac{\partial u}{\partial x}=0, \qquad -1\le x\le 1. \tag{11}$$

The boundary conditions of this equation are periodic, and its initial condition is

$$u(t=0)=\sin(2\pi x), \tag{12}$$

which leads to an exact solution as

$$u(t)=\sin(2\pi(x-t)). \tag{13}$$

We use a forth order ten stage strong stability preserving (SSP) Runge-Kutta method [23] for time integration. The CFL number of fourth, sixth and eighth order schemes are 0.4, 0.05 and 0.01, respectively. The reason behind CFL number reduction for sixth and eighth order schemes is to reduce the effect time discretization error.

.

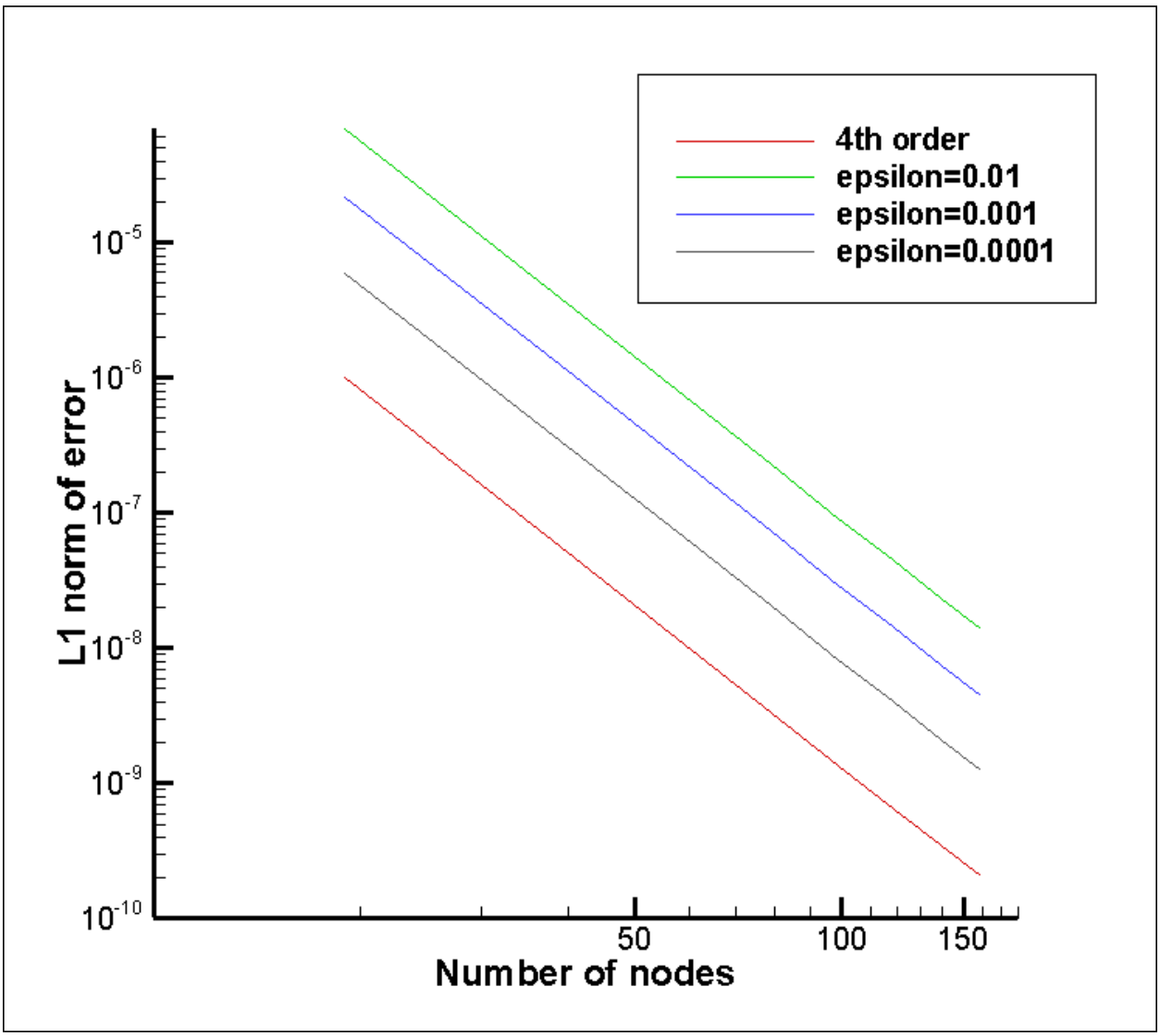


Figure 4 Convergence of L1 norm of error for fourth order optimized schemes with different error thresholds.

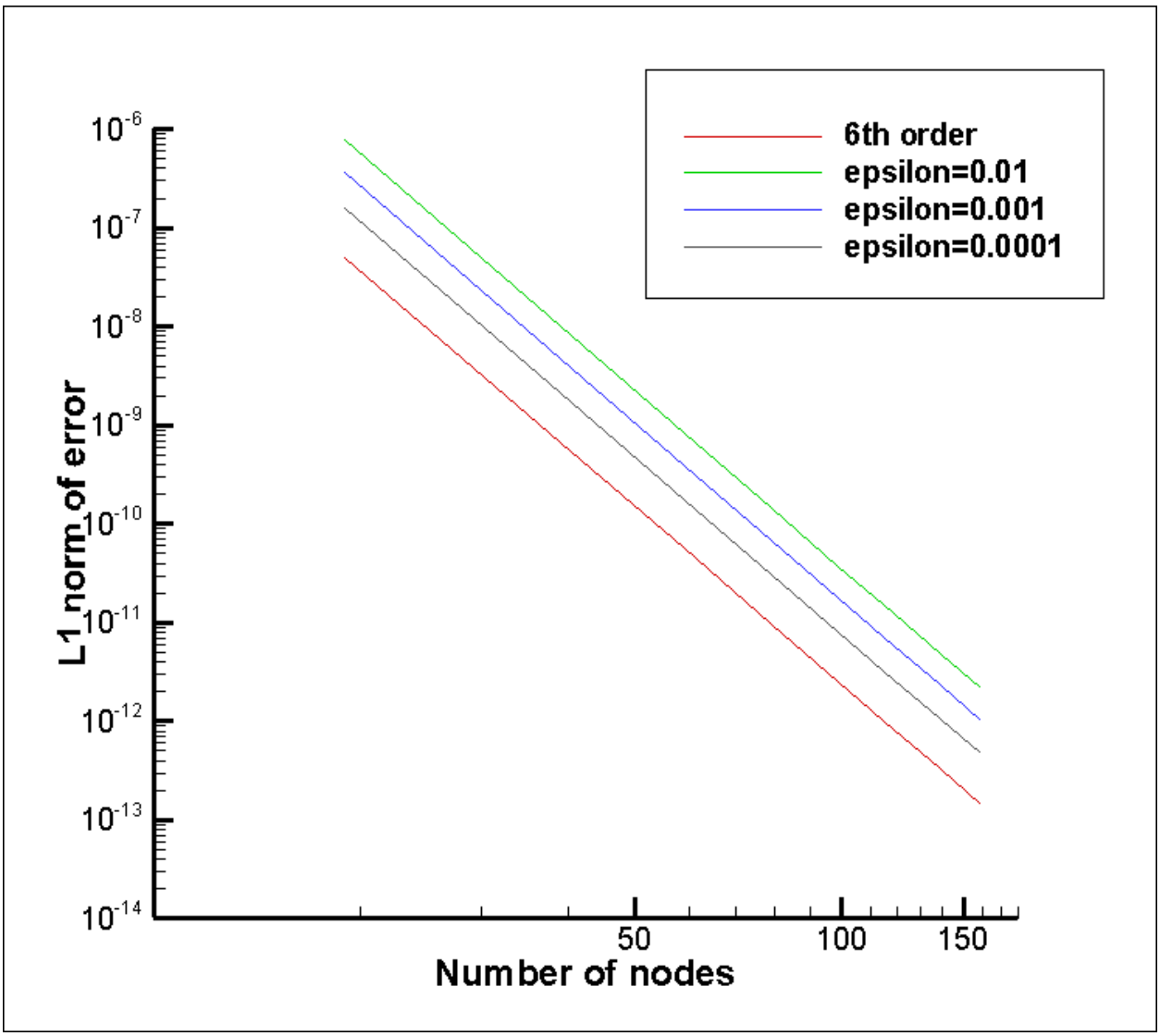


Figure 5 Convergence of L1 norm of error for sixth order optimized schemes with different error thresholds.

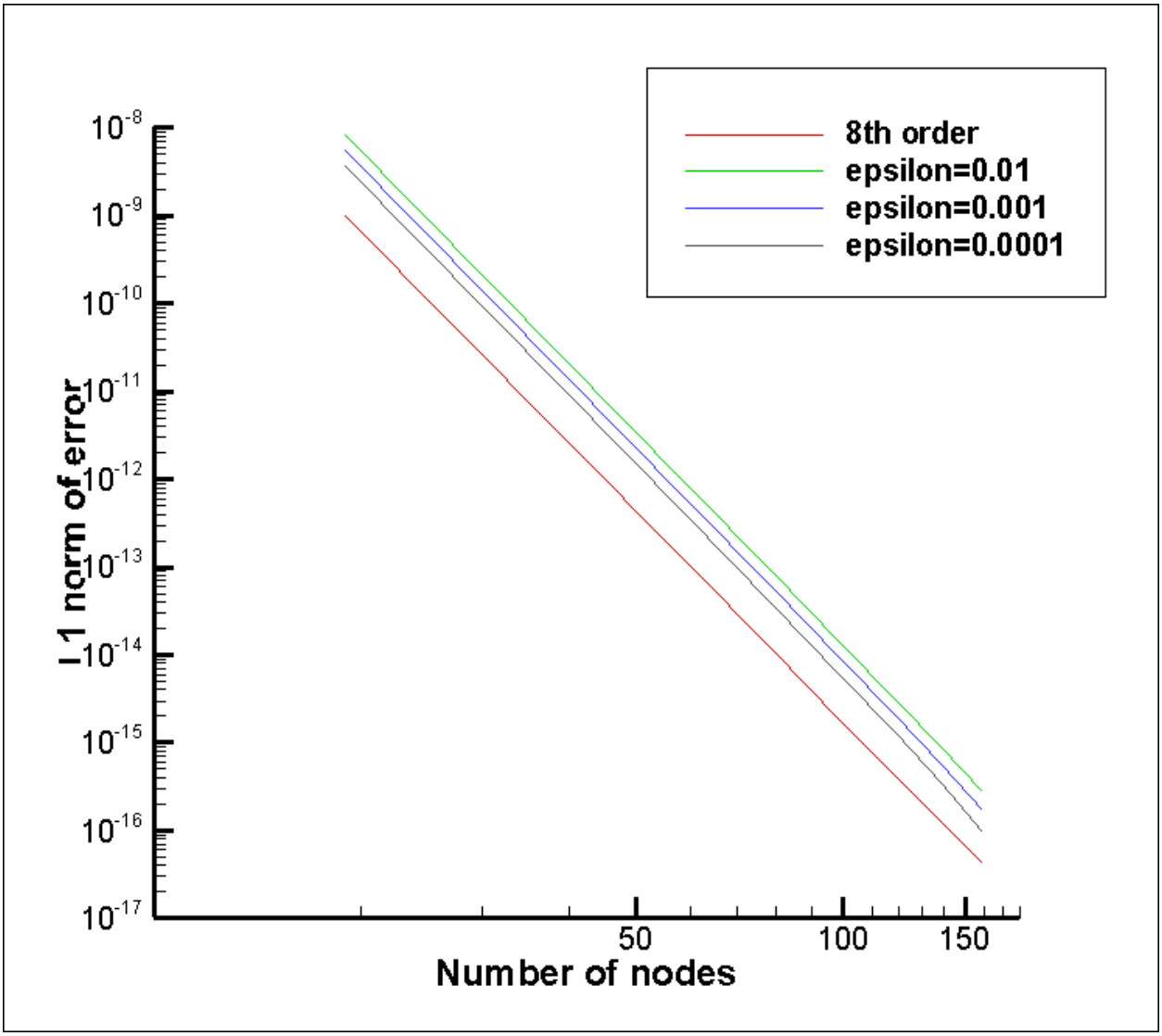


Figure 6 Convergence of L1 norm of error for eighth order optimized schemes with different error thresholds.

Figure 4 to Figure 6 show results of convergence test by solving Eq. (11) for optimized fourth, sixth and eighth order schemes by different error threshold. By considering these plots, we can say all of optimized schemes achieved their nominal order of accuracy and in group of optimized schemes with same order of accuracy, schemes with smaller error threshold have lower error.

Conclusion

In this paper we optimized compact pentadiagonal finite difference schemes by using maximum error norm and different error thresholds. The optimized schemes have fourth, sixth and eighth order of accuracy. We used numerical experiments to investigate the order of accuracy and numerical resolution of optimized schemes. Based upon results of these experiments we can say

all schemes achieved their designed order of accuracy and schemes with smaller error threshold have better numerical resolution and lower error in convergence tests.

References

[1] Hirsh, R. S. (1975). Higher order accurate difference solutions of fluid mechanics problems by a compact differencing technique. *Journal of computational physics*, **19**(1), 90-109.

[2] Adam, Y. (1975). A hermitian finite difference method for the solution of parabolic equations. *Computers & Mathematics with Applications*, **1**(3-4), 393-406.

[3] Adam, Y. (1977). Highly accurate compact implicit methods and boundary conditions. *Journal of Computational Physics*, **24**(1), 10-22.

[4] Rubin, S. G., & Khosla, P. K. (1977). Polynomial interpolation methods for viscous flow calculations. *Journal of Computational Physics*, **24**(3), 217-244.

[5] Vichnevetsky, R., & Bowles, J. B. (1982). *Fourier analysis of numerical approximations of hyperbolic equations*. Society for Industrial and Applied Mathematics.

[6] Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, **220**(4598), 671-680.

[7] Holberg, O. (1987). Computational aspects of the choice of operator and sampling interval for numerical differentiation in large-scale simulation of wave phenomena. *Geophysical prospecting*, **35**(6), 629-655.

[8] Lele, S. K. (1992). Compact finite difference schemes with spectral-like resolution. *Journal of computational physics*, **103**(1), 16-42.

[9] Tam, C. K., & Webb, J. C. (1993). Dispersion-relation-preserving finite difference schemes for computational acoustics. *Journal of computational physics*, **107**(2), 262-281.

[10] Kim, J. W., & Lee, D. J. (1996). Optimized compact finite difference schemes with maximum resolution. *AIAA journal*, **34**(5), 887-893.

[11] Kim, J. W., & Lee, D. J. (1997). Implementation of boundary conditions for optimized high-order compact schemes. *Journal of Computational Acoustics*, **5**(02), 177-191.

[12] Baeder, L. T. T. J. D. (1997). Uniformly accurate compact difference schemes. *AIAA paper* 1997-2093.

[13] Chu, P. C., & Fan, C. (1998). A three-point combined compact difference scheme. *Journal of Computational Physics*, **140**(2), 370-399.

[14] Mahesh, K. (1998). A family of high order finite difference schemes with good spectral resolution. *Journal of Computational Physics*, **145**(1), 332-358.

[15] Zhuang, M., & Chen, R. F. (1998). Optimized upwind dispersion-relation-preserving finite difference scheme for computational aeroacoustics. *AIAA journal*, *36*(11), 2146-2148.

[16] Lui, C., & Lele, S. (2001). Direct numerical simulation of spatially developing, compressible, turbulent mixing layers. In *39th Aerospace Sciences Meeting and Exhibit* (p. 291).

[17] Ashcroft, G., & Zhang, X. (2003). Optimized prefactored compact schemes. *Journal of computational physics*, **190**(2), 459-477.

[18] Tam, C. K. (2004). Computational aeroacoustics: an overview of computational challenges and applications. *International Journal of Computational Fluid Dynamics*, **18**(6), 547-567.

[19] Bogey, C., & Bailly, C. (2004). A family of low dispersive and low dissipative explicit schemes for flow and noise computations. *Journal of Computational physics*, **194**(1), 194-214.

[20] Kim, J. W. (2007). Optimised boundary compact finite difference schemes for computational aeroacoustics. *Journal of Computational Physics*, **225**(1), 995-1019.

[21] Jordan, S. A. (2007). The spatial resolution properties of composite compact finite differencing. *Journal of Computational Physics*, **221**(2), 558-576.

[22] Liu, Z., Huang, Q., Zhao, Z., & Yuan, J. (2008). Optimized compact finite difference schemes with high accuracy and maximum resolution. *International journal of aeroacoustics*, **7**(2), 123-146.

[23] KETCHESON, D.I. (2008). HIGHLY EFFICIENT STRONG STABILITY-PRESERVING RUNGE–KUTTA METHODS WITH LOW-STORAGE IMPLEMENTATIONS. *SIAM Journal of Scientific Computation*, **30**(4), 2113–2136.

[24] Venutelli, M. (2011). Holberg's optimisation for high-order compact finite difference staggered schemes. *International Journal of Computational Fluid Dynamics*, **25**(5), 287-296.
[25] Zhou, H., & Zhang, G. (2011). Prefactored optimized compact finite-difference schemes for second spatial derivatives. *Geophysics*, **76**(5), WB87-WB95.
[26] Jordan, S. A. (2011). Optimization, resolution and application of composite compact finite difference templates. *Applied Numerical Mathematics*, **61**(1), 108-130.
[27] Kim, J. W., & Sandberg, R. D. (2012). Efficient parallel computing with a compact finite difference scheme. *Computers & Fluids*, **58**, 70-87.
[28] Haeri, S., & Kim, J. W. (2013). Implementation of genetic algorithm for the optimization of boundary characteristics of compact finite difference schemes, *Submitted to Computer Methods in Applied Mathematics and Engineering*.
[29] Zhang, J. H., & Yao, Z. X. (2013). Optimized explicit finite-difference schemes for spatial derivatives using maximum norm. *Journal of Computational Physics*, **250**, 511-526.
[30] Cunha, G., & Redonnet, S. (2014). On the effective accuracy of spectral-like optimized finite-difference schemes for computational aeroacoustics. *Journal of Computational Physics*, **263**, 222-232.
[31] Yu, C. H., Wang, D., He, Z., & Pähtz, T. (2015). An optimized dispersion–relation-preserving combined compact difference scheme to solve advection equations. *Journal of Computational Physics*, **300**, 92-115.
[32] Cunha, G., & Redonnet, S. (2015). A novel optimization technique for explicit finite-difference schemes with application to AeroAcoustics. *International Journal for Numerical Methods in Fluids*, **78**(4), 189-216.
[33] Turner, J. M., Haeri, S., & Kim, J. W. (2016). Improving the boundary efficiency of a compact finite difference scheme through optimising its composite template. *Computers & Fluids*, **138**, 9-25.

An optimal eighth-order scheme for multiple zeros of univariate functions

Ramandeep Behl [a], Fiza Zafar [b], Ali Saleh Alshomrani [c], †Moin-ud-Din Junjua [b] and Nusrat Yasmin [b]

[a] School of Mathematics, Statistics and Computer Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

[b] Centre for Advanced Studies in Pure and Applied Mathematics, Bahauddin Zakariya University, Multan 60800, Pakistan

[c] Department of Mathematics, King Abdulaziz University, Jeddah 21577, Saudi Arabia

†Corresponding author: moin_junjua@yahoo.com

Abstract

We present an optimal eighth-order scheme which will work for multiple zeros with multiplicity $(m \geq 1)$, for the first time. Earlier, the maximum convergence order of multi-point iterative schemes was six for multiple zeros in the available literature. So, the main contribution of this study is to present a new higher-order and as well as optimal scheme for multiple zeros for the first time. In addition, we present an extensive convergence analysis with the main theorem which confirms theoretically eighth-order convergence of the presented scheme. Moreover, we consider several real life problems which contain simple as well as multiple zeros in order to comparison with the existing robust iterative schemes. Finally, we conclude on the basis of obtained numerical results that the proposed iterative methods perform far better than the existing methods in terms of residual error, computational order of convergence and difference between the two consecutive iterations.

Keywords: Nonlinear equations, Kung-Traub conjecture, multiple zeros, efficiency index, optimal iterative methods.

Introduction

In the earlier years, it was very tough to construct a higher-order optimal multi-point scheme for multiple zeros of the involved function f with multiplicity $(m \geq 1)$. One of the main reason was the lengthy and complicated calculation which was quite tough or consume a lots of time to solve. Nowadays, with the advancement of digital computer, advanced computer arithmetics and symbolic computation, the construction of higher-order optimal multi-point methods become more vital and popular in this field. Because, the calculation of error equations of iterative methods and asymptotic error constant term for multiple zeros become easier now than the earlier time. However, still there is a need of hard work in order to construct higher-order optimal schemes.

Several scholars from worldwide like Li et al. [1] in (2009), Sharma and Sharma [2] and Li et al. [3] in (2010), Zhou et al. [4] in (2011), Sharifi et al. [5] in (2012), Soleymani et al. [6], Soleymani and Babajee [7], Liu and Zhou [8] and Zhou et al. [9] in (2013), Thukral [10] in (2014), Behl et al. [11] and Hueso et al. [12] in (2015) and Behl et al. [13] in (2016) have presented optimal fourth-order methods for multiple zeros in last two-three decades. In addition, Li et al. [3] (expect two of them are optimal) and Neta [14] presented non-optimal fourth-order iterative methods. Most of the above listed methods are the extension or modification of modified Newton's method (also known as Rall's method [20]) or Newton like method at the expense of additional functional evaluations or increase the substep of the

original methods.

In the last two decades, many researchers from worldwide have tried to develop an optimal scheme whose convergence order should be greater than four (for multiple zeros with multiplicity $m \geq 1$ of univariate function). But, none of them have succeeded in this direction till date. However, some scholars have attained maximum sixth-order convergence in the case of multiple zeros which can be find in the available literature. There are only three multi-point iterative schemes with sixth-order convergence for multiple zeros till date, according to our best knowledge (which were proposed in the recent years). First one was proposed by Thukral [15] and other two were presented by Geum et al. [16, 17]. The details can be seen as follow:

In 2013, Thukral [15] presented a multi-point iterative method with sixth-order convergence, which is given by

$$\begin{aligned} y_n &= x_n - m\frac{f(x_n)}{f'(x_n)}, \\ z_n &= x_n - m\frac{f(x_n)}{f'(x_n)}\sum_{i=1}^{3} i\left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{i}{m}}, \\ x_{n+1} &= z_n - m\frac{f(x_n)}{f'(x_n)}\left(\frac{f(z_n)}{f(x_n)}\right)^{\frac{1}{m}}\left[\sum_{i=1}^{3} i\left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{i}{m}}\right]^2. \end{aligned} \tag{1.1}$$

In 2015, Geum et al. [16], have given the following two-point sixth-order iterative scheme:

$$\begin{aligned} y_n &= x_n - m\cdot\frac{f(x_n)}{f'(x_n)}, m>1, \\ x_{n+1} &= y_n - Q(p_n, s_n)\cdot\frac{f(y_n)}{f'(y_n)}, \end{aligned} \tag{1.2}$$

where, $p_n = \sqrt[m]{\frac{f(y_n)}{f(x_n)}}, s_n = \sqrt[m-1]{\frac{f'(y_n)}{f'(x_n)}}$ and $Q:\mathbb{C}^2 \to \mathbb{C}$ is holomorphic function in the neighborhood of origin $(0,0)$.

In 2016, Geum et al. [17], have again proposed a three-point iterative scheme with sixth-order convergence for multiple zeros. The proposed scheme was based on weight function approach, which can be seen in the following expression:

$$\begin{aligned} y_n &= x_n - m\cdot\frac{f(x_n)}{f'(x_n)}, m\geq 1, \\ w_n &= y_n - m\cdot G(p_n)\cdot\frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= w_n - m\cdot K(p_n, t_n)\cdot\frac{f(x_n)}{f'(x_n)}, \end{aligned} \tag{1.3}$$

where, $p_n = \sqrt[m]{\frac{f(y_n)}{f(x_n)}}$ and $t_n = \sqrt[m]{\frac{f(w_n)}{f(x_n)}}$. The weight functions $G:\mathbb{C}\to\mathbb{C}$ is analytic in a neighborhood of 0 and $K:\mathbb{C}^2\to\mathbb{C}$ is holomorphic in a neighborhood of $(0,0)$.

All of the above three schemes (1.1), (1.2) and (1.3) require four functional evaluations in order to produce sixth-order convergence with the efficiency index $6^{\frac{1}{4}} = 1.5650$. So, none of them is optimal scheme according to the classical Kung-Traub's conjecture [18]. In addition, the above expression (1.2) has one more drawback that it does not work for simple zeros (i.e. $m=1$). Moreover, there does not exist any optimal scheme whose convergence order is greater than four in the case of multiple zeros according to our best knowledge. So, we need optimal eighth-order schemes which will work for multiple zeros ($m>1$) as well as for simple zeros ($m=1$) because they have better efficiency index than fourth and sixth-order methods. Furthermore, these schemes also require a small number of iterations in order to obtain desired accuracy as compare to fourth and sixth-order methods.

Motivated and inspired by this, we present an optimal scheme with eighth-order convergence, which will work for multiple zeros with multiplicity $m\geq 1$, for the first time. The proposed scheme requires four functional evaluations in order to reach eighth-order convergence with the efficiency index $8^{\frac{1}{4}} = 1.6817$, which is higher than the efficiency index of any of the existing methods for multiple zeros in the available literature (also of the recent sixth-order schemes proposed by Thukral [15] and Geum et al. [16, 17]). The rest of the paper is organized as follows. Section 2 provides the methodology and convergence analysis for the proposed optimal eighth order scheme. In Section 3, some special cases of the new scheme are considered. Section 4 is devoted to numerical experiments and comparisons of different multiple zero finders using some real life problems. Finally, conclusions are given in Section 5.

Construction of optimal scheme with eighth-order convergence

This section is devoted to the main contribution of this study and convergence analysis of the proposed scheme with main theorem. Here, we consider the following proposed by Wang and Liu [19]

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)}, n\geq 0,$$

$$z_n = y_n - \frac{f(x_n)}{f'(x_n)} G\left(\frac{f(y_n)}{f(x_n)}\right),$$

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(x_n)}\left[H\left(\frac{f(y_n)}{f(x_n)}\right) + W\left(\frac{f(z_n)}{f(y_n)}\right) V\left(\frac{f(y_n)}{f(x_n)}\right)\right], \tag{2.1}$$

Where $G, H, V, W:\mathbb{R}\to\mathbb{R}$ are the weight functions and sufficiently differentiable in the neighborhood of origin. The above scheme is an optimal eighth-order scheme for only simple zeros.

Now, we want to extend this scheme for multiple zeros with multiplicity $m\geq 1$. So, we will

rewrite the above expression (2.1) in simpler form with some modifications in second and third substep, in the following way:

$$y_n = x_n - m\cdot\frac{f(x_n)}{f'(x_n)}, n \geq 0,$$

$$z_n = y_n - m\cdot u\cdot\frac{f(x_n)}{f'(x_n)}\frac{1+\beta u}{1+(\beta-2)u}, \beta \in \mathbb{R}$$

$$x_{n+1} = z_n - u\cdot v\cdot\frac{f(x_n)}{f'(x_n)}[\alpha_1 + (1+\alpha_2 v)P_f(u)], \tag{2.2}$$

where $\alpha_1, \alpha_2 \in \mathbb{R}$ are two free disposable parameters and the weight function $P_f: \mathbb{C} \to \mathbb{C}$ is an analytic function in a neighborhood of (0) with $u = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}}, v = \left(\frac{f(z_n)}{f(y_n)}\right)^{\frac{1}{m}}$ are disposable parameters. It is worthy to note that we will obtain well known King's family of fourth-order iterative methods for $m = 1$ with the help of first two substep. In addition, we can obtain an optimal eighth-order scheme for simple zeros as special case of Wang and Liu's scheme for $m = 1$.

In the next Theorem 2.1, we demonstrate that the order of convergence of the proposed scheme will reach at optimal eight without using additional functional evaluations. It is interesting to observe that how P_f and disposable parameters $(\alpha_i, i = 1,2)$ contributes their role in the construction of the desired eighth-order convergence (for the details please see the Theorem 2.1).

Theorem 2.1 *Let us consider $x = \xi$ (say) be a multiple zero with multiplicity $m \geq 1$ of the involved function f. In addition, we assume that $f: \mathbb{C} \to \mathbb{C}$ be an analytic function in the region enclosing a multiple zero ξ. The proposed scheme defined by (2.2) has an optimal eighth-order convergence, when it satisfies the following expressions*

$$\begin{aligned} &\alpha_1 = \frac{m}{2}, \alpha_2 = 2, P(0) = \frac{m}{2}, P'(0) = 2m, P''(0) = 2m(5-2\beta), \\ &P'''(0) = 12m(\beta^2 - 6\beta + 6), \end{aligned} \tag{2.3}$$

where $\beta \in \mathsf{R}$.

Proof. Let us assume that $e_n = x_n - \xi$ be the error at nth step. Now, expand $f(x_n)$ and $f'(x_n)$ about $x = \xi$ by the Taylor's series expansion (with the help of *Mathematica 11*), we have

$$f(x_n) = \frac{f^{(m)}(\xi)}{m!}e_n^m(1 + c_1e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + c_5e_n^5 + c_6e_n^6 + c_7e_n^7 + c_8e_n^8 + O(e_n^9)) \tag{2.4}$$

and

$$f'(x_n) = \frac{f^{(m)}(\xi)}{m!} e_n^{m-1}(m + c_1(m+1)e_n + c_2(m+2)e_n^2 + c_3(m+3)e_n^3 + c_4(m+4)e_n^4 + c_5(m+5)e_n^5$$

$$+ c_6(m+6)e_n^6 + c_7(m+7)e_n^7 + c_8(m+8)e_n^8 + O(e_n^9)), \tag{2.5}$$

respectively, where $c_k = \frac{m!}{(m+k)!} \frac{f^{(m+k)}(\xi)}{f^{(m)}(\xi)}, k = 1,2,3,\ldots 8$.

By inserting the above expressions (2.4) and (2.5), in the first substep of (2.2), we will yield

$$y_n - \xi = \frac{c_1 e_n^2}{m} + \frac{(2mc_2 - (m+1)c_1^2)e_n^3}{m^2} + \sum_{k=0}^{4} A_k e_n^{k+4} + O(e_n^9), \tag{2.6}$$

where $A_k = A_k(m, c_1, c_2, \ldots, c_8)$ are given in terms of $m, c_1, c_2, c_3, \ldots, c_8$ with explicitly written two coefficients $A_0 = \frac{1}{m^3}\{3c_3m^2 + c_1^3(m+1)^2 - c_1c_2m(3m+4)\}$ and $A_1 = -\frac{1}{m^4}\{c_1^4(m+1)^3 - 2c_2c_1^2m(2m^2+5m+3) + 2c_3c_1m^2(2m+3) + 2m^2(c_2^2(m+2) - 2c_4m)\}$, etc.

With the help of Taylor's series expansion and expression (2.6), we have

$$f(y_n) = f^{(m)}(\xi)e_n^{2m}\left[\frac{(\frac{c_1}{m})^m}{m!} + \frac{(2mc_2 - (m+1)c_1^2)(\frac{c_1}{m})^m e_n}{m!c_1} + \sum_{k=0}^{6} \overline{A}_k e_n^{k+2} + O(e_n^9)\right]. \tag{2.7}$$

By using the expressions (2.4) and (2.7), we get

$$u = \left(\frac{f(y_n)}{f(x_n)}\right)^{\frac{1}{m}} = \frac{c_1 e_n}{m} + \frac{(2mc_2 - (m+2)c_1^2)e_n^2}{m^2} + \tau_1 e_n^3 + \tau_2 e_n^4 + \tau_3 e_n^5 + O(e_n^6), \tag{2.8}$$

where

$$\tau_1 = \frac{1}{2m^3}[c_1^3(2m^2+7m+7) + 6c_3m^2 - 2c_1c_2m(3m+7)],$$

$$\tau_2 = -\frac{1}{6m^4}[c_1^4(6m^3+29m^2+51m+34) - 6c_2c_1^2m(4m^2+16m+17) + 12c_1c_3m^2(2m+5)$$

$$+12m^2(c_2^2(m+3) - 2c_4m)]$$

and

$$\tau_3 = \frac{1}{24m^5}[-24m^3(c_2c_3(5m+17)-5c_5m)+12c_3c_1^2m^2(10m^2+43m+49)+$$
$$12c_1m^2\{c_2^2(10m^2+47m+53)-2c_4m(5m+13)\}-4c_2c_1^3m(30m^3+163m^2+306m+209)+.$$
$$c_1^5(24m^4+146m^3+355m^2+418m+209)]$$

Now, insert the expressions (2.6) – (2.8) in the second substep of scheme (2.2), we obtain

$$z_n - \xi = \frac{(4\beta+m+1)c_1^3-2mc_1c_2}{2m^3}e_n^4+\sum_{j=0}^{4}B_je_n^{j+4}+O(e_n^9), \tag{2.9}$$

where $B_j = B_j(m,c_1,c_2,\ldots,c_8)$ are given in terms of $m,c_1,c_2,c_3,\ldots,c_8$ with explicitly written three coefficients $B_0 = -\frac{1}{6m^4}\{c_1^4(12\beta^2+36\beta+7m^2+12(3\beta+1)m+5)+12c_1c_3m^2+12c_2^2m^2-24c_2c_1^2m(3\beta+m+1)\}, B_1 = \frac{1}{24m^5}\{48\beta^3+144\beta^2+264\beta+46m^3+(288\beta+101)m^2+2m(96\beta^2+252\beta+37)+19\}$

and

$$B_2 = -\{12c_1^2c_3m^2(36\beta+13m+11)+(37-168c_2c_3m^3+4c_1^3c_2m(96\beta^2+252\beta+53m^2+$$
$$18(14\beta+5)m)+12c_1m^2(c_2^2(48\beta+17m+19)-6c_4m)\},$$

etc.

Again with the help of above expression (2.9) and the Taylor's series expansion, we have

$$f(z_n) = f^{(m)}(\xi)e_n^{4m}\left[\frac{2^{-m}\left(\frac{(4\beta+m+1)c_1^3-2mc_1c_2}{m^3}\right)^m}{m!}-\frac{\left(2^{-m}\left(\frac{(4\beta+m+1)c_1^3-2mc_1c_2}{m^3}\right)^{m-1}\theta_0\right)}{3(m^3m!)}e_n\right.$$

$$\left.+\sum_{j=0}^{7}\overline{B}_je_n^{j+1}+O(e_n^9)\right], \tag{2.10}$$

where
$\theta_0 = \{12\beta^2+36\beta+7m^2+12m(3\beta+1)+5\}c_1^4+12m^2c_1c_3+12m^2c_2^2-24m(3\beta+m+1)c_1^2c_2.$

By using the above expressions (2.7) and (2.10), we further obtain

$$v = \left(\frac{f(z_n)}{f(y_n)}\right)^{\frac{1}{m}} = \frac{c_1^2(4\beta+m+1)-2mc_2}{2m^2}e_n^2+\theta_1e_n^3+\theta_2e_n^4+\theta_3e_n^5+O(e_n^6), \tag{2.11}$$

Where $\theta_1 = -\frac{1}{3m^3}\{c_1^3(6\beta^2+12\beta+2m^2+3m(4\beta+1)+1)+6m^2c_3-6mc_1c_2(4\beta+$

$m+1)\},\theta_2 ==$
$\frac{1}{24m^4}[-12c_2c_1^2m(24\beta^2+36\beta+6m^2+m(40\beta+7)-1)+24c_1c_3m^2(12\beta+3m+2)+12m^2(c_2^2(16\beta+3m+3)-6mc_4)+c_1^4\{48\beta^3+96\beta^2+72\beta+18m^3+(144\beta+25)m^2+6(24\beta^2+36\beta-1)m-13\}]$ and $\theta_3=-\frac{1}{60m^5}[120m^3(c_2c_3(12\beta+2m+1)-2mc_5)+60c_3c_1^2m^2(18\beta^2+24\beta+4m^2+28\beta m+3m-3)+60c_1m^2\{c_2^2(24\beta^2+24\beta+4m^2+32\beta m+3m-3)-2c_4m(8\beta+2m+1)\}-20c_2c_1^3m(48\beta^3+72\beta^2+36\beta+12m^3+(108\beta+11)m^2)+18(7\beta^2+8\beta-1)m-17)+c_1^5\{3(40\beta^4+80\beta^3+40\beta^2-21)+48m^4+10(48\beta+5)m^3+15(48\beta^2+56\beta-7)m^2+10(48\beta^3+72\beta^2+36\beta-17)m\}]$.

Since it is clear from the expression (2.8) that u is of order e_n. Therefore, we can expand weight function $P_f(u)$ in the neighborhood of origin by Taylor's series expansion up to third-order terms as follows:

$$P_f(u)=P(0)+P^{'}(0)u+\frac{1}{2!}P^{''}(0)u^2+\frac{1}{3!}P^{'''}(0)u^3. \tag{2.12}$$

By using the expressions (2.4)–(2.12) in the last substep of proposed scheme (2.2), we have

$$e_{n+1}=\frac{c_1(c_1^2(4\beta+m+1)-2mc_2)(m-\alpha_1-\alpha_2P(0))}{2m^4}e_n^4+\sum_{i=1}^{4}E_ie_n^{i+4}+O(e_n^9), \tag{2.13}$$

where $E_i=E_i(m,\beta,\alpha_1,\alpha_2,P(0),P^{'}(0),P^{''}(0),P^{'''}(0),c_1,c_2,\ldots,c_8)$.

It is straightforward to say from the above expression (2.13) that we can easily obtain at least fifth-order convergence, when we will choose the following value of α_1

$$\alpha_1=m-P(0). \tag{2.14}$$

With the help of the above expression (2.14) and $E_1=0$, we obtain

$$2m-P^{'}(0)=0, \tag{2.15}$$

which further yield

$$P^{'}(0)=2m. \tag{2.16}$$

Again, inserting the above expressions (2.14) and (2.16) in $E_2=0$, we have

$$m-\alpha_2P(0)=0,-\alpha_2P(0)(4\beta+m+1)+m(m+11)-P^{''}(0)=0, \tag{2.17}$$

the above two independent expressions, which further leads us

$$P(0)=\frac{m}{\alpha_2},\quad P^{''}(0)=2m(5-2\beta). \tag{2.18}$$

Now, by using the above expressions (2.14), (2.16), (2.18) and $E_3=0$, we obtain

$$\alpha_2 - 2 = 0, \quad 12m(\beta^2 - 2\beta + m + 7) - 6\alpha_2 m(4\beta + m + 1) - P'''(0) = 0, \tag{2.19}$$

which further yield

$$\alpha_2 = 2, \quad P'''(0) = 12m(\beta^2 - 6\beta + 6). \tag{2.20}$$

Finally, by substituting the above expressions (2.14), (2.16), (2.18) and (2.20) in the expression (2.13), we obtain the following optimal asymptotic error constant term

$$e_{n+1} = \frac{c_1((4\beta + m + 1)c_1^2 - 2mc_2)}{24m^7}[(347 - 24\beta^3 + 288\beta^2 - 492\beta + 7m^2 + 18m(2\beta + 1))c_1^4 + 12m^2 c_1 c_3$$

$$+12m^2 c_2^2 - 12m(6\beta + 2m + 3)c_1^2 c_2]e_n^8 + O(e_n^9). \tag{2.21}$$

The above asymptotic error constant (2.21) reveals that the proposed scheme (2.2) reaches at optimal eighth-order convergence by using only four functional evaluations (viz. $f(x_n), f'(x_n), f(y_n)$ and $f'(z_n)$) per iteration. This completes the proof.

Some special cases of weight function

In this section, we will discuss some special cases of our proposed scheme (2.2) by assigning different kind of weight functions P_f. In this regard, please see following cases, where we have mentioned some different kind of members of the proposed scheme:

Case A: Let us describe the following weight function directly from the proposed Theorem 2.1

$$P_f(u) = \frac{m}{2}\left(1 + 4u + (10 - 4\beta)u^2 + 4(\beta^2 - 6\beta + 6)u^3\right). \tag{3.1}$$

Thus, the corresponding optimal eighth-order iterative scheme is given by

$$\begin{aligned} y_n &= x_n - m \cdot \frac{f(x_n)}{f'(x_n)} \\ z_n &= y_n - m \cdot u \cdot \frac{f(x_n)}{f'(x_n)} \frac{1 + \beta u}{1 + (\beta - 2)u}, \\ x_{n+1} &= z_n - \frac{m}{2} \cdot u \cdot v \cdot \frac{f(x_n)}{f'(x_n)}\left[1 + (2v + 1)(4(\beta^2 - 6\beta + 6)u^3 + (10 - 4\beta)u^2 + 4u + 1)\right]. \end{aligned} \tag{3.2}$$

Case B: Now, we suggest rational weight function satisfying the conditions (2.3) as follows.

$$P_f(u) = -\frac{m((8\beta + 2)u^2 - (2\beta^2 - 4\beta - 8)u - 2\beta + 5)}{2(2(\beta^2 - 6\beta + 6)u + 2\beta - 5)}, \tag{3.3}$$

which further yields

$$\begin{aligned} y_n &= x_n - m\cdot\frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - m\cdot u\cdot\frac{f(x_n)}{f'(x_n)}\frac{1+\beta u}{1+(\beta-2)u}, \\ x_{n+1} &= z_n + \frac{m}{2}\cdot u\cdot v\cdot\frac{f(x_n)}{f'(x_n)}\left[1-\frac{(2v+1)(2u(2\beta-\beta^2+4\beta u+u+4)-2\beta+5)}{2\beta+2(\beta^2-6\beta+6)u-5}\right], \end{aligned} \tag{3.4}$$

is a new optimal eighth-order scheme.

Case C: let us consider another rational weight function which satisfies the conditions of (2.3), is given by

$$P_f(u) = \frac{m(2(\beta^2+2\beta+2)u+2\beta+3)}{2((8\beta+2)u^2+2(\beta^2-2\beta-4)u+2\beta+3)}. \tag{3.5}$$

By using the above expression, we obtain the following optimal eighth-order scheme:

$$\begin{aligned} y_n &= x_n - m\cdot\frac{f(x_n)}{f'(x_n)}, \\ z_n &= y_n - m\cdot u\cdot\frac{f(x_n)}{f'(x_n)}\frac{1+\beta u}{1+(\beta-2)u}, \\ x_{n+1} &= z_n - m\cdot u\cdot v\cdot\frac{f(x_n)}{f'(x_n)}\left[\frac{(4\beta+1)u^2+2u(\beta^2+(\beta^2+2\beta+2)v-1)+(2\beta+3)(v+1)}{(8\beta+2)u^2+2(\beta^2-2\beta-4)u+2\beta+3}\right]. \end{aligned} \tag{3.6}$$

In the similar fashion, we can develop several new and interesting optimal schemes with eighth-order convergence for multiple zeros by just assigning different values to β or considering new weight functions which satisfy the conditions of Theorem 2.1.

Numerical experiments

This section is devoted to demonstrate the efficiency, effectiveness and convergence behavior of the presented scheme. In this regards, we consider some of the special cases of the proposed scheme namely, expression (3.2) $\left(\text{for } \beta=\frac{1}{2}\right)$, expression (3.4) $\left(\text{for } \beta=\frac{1}{3}\right)$ and expression (3.6) (for $\beta=0$), denoted by $(M1), (M2)$ and $(M3)$, respectively. In addition, we choose a total number of five test problems for comparison: first one is eigen value problem; second one is Van der Waals equation which state the behavior of real gas; third one again is related to chemical reactor problem but for simple zeros; last two are standard test functions, which can be seen in the examples 4.1–4.5.

Now, we want to compare our methods with other existing robust methods of same order on the basis of difference between two consecutive iterations, computational order of convergence ρ and residual errors in the function. Unfortunately, there is no optimal eighth-order iterative methods for multiple zeros available in the literature in order to comparison. So, we have chosen sixth-order iterative methods for the comparison which is the highest-order till date for multiple zeros.

Therefore, we compare the proposed methods with the family of two-point sixth-order methods, which were presented by Guem et al. in [16], out of them we consider the following expression:

$$\begin{aligned} y_n &= x_n - m\frac{f(x_n)}{f'(x_n)}, m > 1, \\ x_{n+1} &= y_n - m\left[1 + 2(m-1)(p_n - s_n) - 4p_n s_n + s_n^2\right]\cdot\frac{f(y_n)}{f'(y_n)}, \end{aligned} \tag{4.1}$$

called by $(GM1)$.

Finally, we compare them with another non-optimal scheme with sixth-order convergence based on weight function approach proposed by the same authors Guem et al. [17], out of them we chose the following expression:

$$\begin{aligned} y_n &= x_n - m\frac{f(x_n)}{f'(x_n)}, m \geq 1, \\ w_n &= x_n - m\left[1 + p_n + 2p_n^2\right]\cdot\frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= x_n - m\left[1 + p_n + 2p_n^2 + (1+2p_n)t_n\right]\cdot\frac{f(x_n)}{f'(x_n)}, \end{aligned} \tag{4.2}$$

denoted by $(GM2)$.

In Tables 1–2, we display the number of iteration indexes (n), error in the consecutive iterations $|x_{n+1} - x_n|$, computational order of convergence (ρ) (we used the formula given by Cordero and Torregrosa [24] in order to calculate ρ) and absolute residual error of the corresponding function $(|f(x_n)|)$. We did our calculations with several number of significant digits (minimum 3000 significant digits) to minimize the round off error.

As we mentioned in the above paragraph that we calculate the values of all the constants and functional residuals up to several number of significant digits but due to the limited paper space, we display the value of errors in the consecutive iterations $|x_{n+1} - x_n|$ and absolute residual errors in the function $|f(x_n)|$ up to 2 significant digits with exponent power in Tables 1–2. Moreover, computational order of convergence is up to 5 significant digits. Finally, we display the values of approximated zeros up to 30 significant digits in the examples.

For the computer programming, all computations have been performed using the programming package *Mathematica* 11 with multiple precision arithmetic. Further, the meaning of $a(\pm b)$ is $a \times 10^{(\pm b)}$ in the following Tables 1–2.

Example 4.1. Eigen value problem:

One of the toughest and challenging task of linear algebra is concern with the eigen values of a large square matrix. Further, finding the zeros of characteristic equation of square matrix greater than 4 is another big challenge. So, we consider the following 9×9 matrix

$$A=\frac{1}{8}\begin{bmatrix} -12 & 0 & 0 & 19 & -19 & 76 & -19 & 18 & 437 \\ -64 & 24 & 0 & -24 & 24 & 64 & -8 & 32 & 376 \\ -16 & 0 & 24 & 4 & -4 & 16 & -4 & 8 & 92 \\ -40 & 0 & 0 & -10 & 50 & 40 & 2 & 20 & 242 \\ -4 & 0 & 0 & -1 & 41 & 4 & 1 & 2 & 25 \\ -40 & 0 & 0 & 18 & -18 & 104 & -18 & 20 & 462 \\ -84 & 0 & 0 & -29 & 29 & 84 & 21 & 42 & 501 \\ 16 & 0 & 0 & -4 & 4 & -16 & 4 & 16 & -92 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 24 \end{bmatrix}.$$

The corresponding characteristic polynomial of the above matrix (A) is given as follows:

$$f_2(x) = x^9 - 29x^8 + 349x^7 - 2261x^6 + 8455x^5 - 17663x^4 + 15927x^3 + 6993x^2 - 24732x + 12960. \quad (4.3)$$

The above function has one multiple zero at $\xi = 3$ of multiplicity 4 with initial approximation $x_0 = 3.1$.

Example 4.2. Van der Waals equation of state

$$\left(P + \frac{a_1 n^2}{V^2}\right)(V - na_2) = nRT, \quad (4.4)$$

explains the behavior of a real gas by introducing in the ideal gas equations two parameters, a_1 and a_2, specific for each gas. The determination of the volume V of the gas in terms of the remaining parameters requires the solution of a nonlinear equation in V.

$$PV^3 - (na_2P + nRT)V^2 + a_1n^2V - a_1a_2n^2 = 0. \quad (4.5)$$

Given the constants a_1 and a_2 of a particular gas, one can find values for n, P and T, such that this equation has a three real zeros. By using the particular values, we obtain the following nonlinear function

$$f_2(x) = x^3 - 5.22x^2 + 9.0825x - 5.2675. \quad (4.6)$$

have three zeros and out of them one is a multiple zero $\xi = 1.75$ of multiplicity of order two and other one simple zero $\xi = 1.72$. However, our desired zero is $\xi = 1.75$. We considered initial guess $x_0 = 1.8$ for this problem.

Example 4.3. Fractional conversion in a chemical reactor:

Let us consider the following expression (for the details of this problem please see [26])

$$f_3(x) = \frac{x}{1-x} - 5\log\left[\frac{0.4(1-x)}{0.4-0.5x}\right] + 4.45977, \tag{4.7}$$

In the above expression x represents the fractional conversion of species A in a chemical reactor. Since, there will be no physical meaning of above fractional conversion if x is less than zero or greater than one. In this sense, x is bounded in the region $0 \le x \le 1$. In addition, our required zero to this problem is $\xi = 0.757396246253753879459641297929$. Moreover, it is interesting to note that the above expression will be undefined in the region $0.8 \le x \le 1$ which is very close to our desired zero. Furthermore, there are some other properties to this function which make the solution more difficult. The derivative of the above expression will be very close to zero in the region $0 \le x \le 0.5$ and there is an infeasible solution for $x = 1.098$. So, we consider the initial approximation $x_0 = 0.76$.

Example 4.4. Let us consider the following standard nonlinear test function from Behl et al. [13]

$$f_4(x) = \left(-\sqrt{1-x^2} + x + \cos\left(\frac{\pi x}{2}\right) + 1\right)^3 \tag{4.8}$$

The above function has a multiple zero at $\xi = -0.728584046444826716712333102423$ of multiplicity 3 with initial guess $x_0 = -0.69$

Example 4.5. We assume another standard test problem from Petkovíc et al. [21], which is defined by

$$f_5(x) = -\frac{x^4}{12} + \frac{x^2}{2} + x + e^x(x-3) + \sin(x) + 3, \tag{4.9}$$

This function f_5 has multiple zero at $\xi = 0$ of multiplicity 3. We will start with the initial approximation $x_0 = 0.6$ for this problem.

Table 1. Difference between two consecutive iterations (i.e. $|x_{n+1} - x_n|$) of different iteration functions.

$f_i(x)$	n	$GM1$	$GM2$	$M1$	$M2$	$M3$
$f_1(x)$	1	5.2(−3)	1.3(−2)	2.8(−2)	2.8(−2)	2.9(−2)
	2	2.5(−11)	2.4(−13)	4.0(−15)	2.8(−15)	1.9(−15)
	3	5.0(−33)	9.8(−78)	8.2(−118)	3.1(−119)	9.6(−121)
	ρ	2.9610	5.9962	7.9925	7.9939	7.9948

$f_2(x)$	1	8.1(−4)	1.1(−3)	2.6(−4)	2.1(−4)	1.6(−4)
	2	4.5(−12)	4.7(−11)	3.6(−19)	3.0(−20)	1.3(−21)
	3	1.6(−61)	5.0(−55)	6.1(−138)	5.2(−147)	3.1(−158)
	ρ	5.9908	5.9840	7.9977	7.9987	8.0512
$f_3(x)$	1	*	5.4(−9)	5.1(−12)	1.1(−13)	2.2(−14)
	2	*	4.6(−43)	1.2(−81)	6.4(−97)	5.2(−103)
	3	*	1.9(−247)	1.5(−638)	1.1(−762)	4.9(−812)
	ρ	**	6.0000	8.0000	8.0000	8.0512
$f_4(x)$	1	1.6(−8)	2.7(−8)	1.0(−10)	8.3(−11)	6.5(−11)
	2	2.5(−46)	9.2(−45)	7.0(−79)	1.0(−79)	9.7(−81)
	3	3.5(−273)	1.3(−263)	3.5(−624)	5.6(−631)	2.6(−639)
	ρ	6.0000	6.0000	8.0000	8.0000	8.0000
$f_5(x)$	1	5.8(−8)	2.0(−6)	4.4(−8)	4.3(−8)	4.3(−8)
	2	1.2(−47)	5.2(−38)	5.6(−64)	4.6(−64)	3.5(−64)
	3	9.4(−286)	1.4(−227)	4.2(−511)	6.9(−512)	6.0(−513)
	ρ	6.0000	6.0000	8.0000	8.0000	8.0000

(* means method is not working for simple zero ($m = 1$), ** means COC (ρ) can't be calculated for this method.)

Table 2. Comparison based on residual error (i.e. $|f(x_n)|$) of different iteration

functions.

$f_i(x)$	n	$GM1$	$GM2$	$M1$	$M2$	$M3$
$f_1(x)$	1	5.8(−12)	2.3(−6)	5.4(−5)	5.4(−5)	5.4(−5)
	2	2.9(−41)	2.5(−49)	2.1(−56)	5.0(−57)	1.1(−57)
	3	4.9(−128)	7.5(−307)	3.7(−467)	7.4(−473)	6.8(−479)
$f_2(x)$	1	2.0(−8)	3.4(−8)	2.0(−9)	1.4(−9)	7.9(−10)
	2	6.0(−25)	6.6(−23)	4.0(−39)	2.7(−41)	5.4(−44)
	3	7.5(−124)	7.5(−111)	1.1(−276)	8.1(−295)	2.9(−317)
$f_3(x)$	1		4.3(−7)	4.1(−10)	8.5(−12)	1.8(−12)
	2		3.7(−41)	9.9(−80)	5.1(−95)	4.2(−101)
	3		1.5(−245)	1.2(−636)	8.4(−761)	3.9(−810)
$f_4(x)$	1	1.1(−23)	5.2(−23)	2.7(−30)	1.5(−30)	6.9(−31)
	2	4.2(−127)	2.0(−132)	8.6(−235)	2.7(−237)	2.4(−240)
	3	1.1(−817)	6.0(−789)	1.1(−1870)	4.4(−1891)	4.7(−1916)
$f_5(x)$	1	3.3(−23)	1.4(−18)	1.4(−23)	1.4(−23)	1.3(−23)
	2	2.9(−142)	2.3(−113)	2.9(−191)	1.6(−191)	6.9(−192)
	3	1.4(−856)	4.7(−682)	1.2(−1532)	5.4(−1535)	3.6(−1538)

Conclusions

In this paper, we present an optimal eighth-order iterative scheme for finding multiple zeros

of the involved function f with multiplicity $m \geq 1$, for the first time. An extensive convergence analysis is done which confirms theoretically eighth-order convergence of the proposed scheme. In addition, the proposed scheme is optimal in the sense of classical Kung-Traub conjecture. The beauty of the proposed methods is that they have not only smaller errors difference between two consecutive iterations and minimum residual errors corresponding to the considered test functions f_i. But, they also demonstrate the stable computational order of convergence as compared to the other listed methods. Further, the computational efficiency index of the proposed schemes is $E = \sqrt[4]{8} \approx 1.682$ which is better than the efficiency index of classical Newton's method $E = \sqrt[2]{2} \approx 1.414$ and also the schemes proposed by Thukral [15] and Guem et al. [16, 17], $E = \sqrt[4]{6} \approx 1.565$. Moreover, we can obtain several new optimal and interesting iterative methods of order eight by considering different types of weight functions and assigning different values to disposable parameter β. Finally, on accounts of the numerical results obtained, it can be concluded that our proposed methods are highly efficient and perform better than the existing methods for multiple zeros.

References

[1] Li, S., Liao, X. and Cheng L. (2009) A new fourth-order iterative method for finding multiple roots of nonlinear equations, *Applied Mathematics and Computation* **215** 1288–1292.
[2] Sharma, J. R., Sharma, R. (2010) Modified Jarratt method for computing multiple roots, *Applied Mathematics and Computation* **217** 878–881.
[3] S.G. Li, L.Z. Cheng, B. Neta, Some fourth-order nonlinear solvers with closed formulae for multiple roots, Comput. Math. Appl. 59 (2010) 126–135.
[4] X. Zhou, X. Chen, Y. Song, Constructing higher-order methods for obtaining the muliplte roots of nonlinear equations, J. Comput. Math. Appl. 235 (2011) 4199–4206.
[5] M. Sharifi, D.K.R. Babajee, F. Soleymani, Finding the solution of nonlinear equations by a class of optimal methods, Comput. Math. Appl. 63 (2012) 764–774.
[6] F. Soleymani, D.K.R. Babajee, T. Lofti, On a numerical technique forfinding multiple zeros and its dynamic, J. Egypt. Math. Soc. 21 (2013) 346–353.
[7] F. Soleymani, D.K.R. Babajee, Computing multiple zeros using a class of quartically convergent methods, Alex. Eng. J. 52 (2013) 531–541.
[8] B. Liu, X. Zhou, A new family of fourth-order methods for multiple roots of nonlinear equations, Non. Anal. Model. Cont. 18(2) (2013) 143–152.
[9] X. Zhou, X. Chen, Y. Song, Families of third and fourth order methods for multiple roots of nonlinear equations, Appl. Math. Comput. 219 (2013) 6030–6038.
[10] R. Thukral, A new family of fourth-order iterative methods for solving nonlinear equations with multiple roots, J. Numer. Math. Stoch. 6 (1) (2014) 37–44.
[11] R. Behl, A. Cordero, S.S. Motsa, J.R. Torregrosa, On developing fourth-order optimal families of methods for multiple roots and their dynamics, Appl. Math. Comput. 265(15) (2015) 520–532.
[12] J.L. Hueso, E. Martínez, C. Teruel, Determination of multiple roots of nonlinear equations and applications, J. Math. Chem. 53 (2015) 880-892.
[13] R. Behl, A. Cordero, S.S. Motsa, J.R. Torregrosa, V. Kanwar, An optimal fourth-order family of methods for multiple roots and its dynamics, Numer. Algor. 71 (4) (2016) 775–796.
[14] B. Neta, Extension of Murakami's high-order non-linear solver to multiple roots, Int. J. Comput. Math. 87(5) (2010) 1023–1031.
[15] R. Thukral, Introduction to higher-order iterative methods for finding multiple roots of nonlinear equations, J. Math. 2013 (2013) Article ID 404635, 3 pages http://dx.doi.org/10.1155/2013/404635.
[16] Y.H. Geum, Y.I. Kim, B. Neta, A class of two-point sixth-order multiple-zero finders of modified double-Newton type and their dynamics, Appl. Math. Comput. 270 (2015) 387–400 .
[17] Y.H. Geum, Y.I. Kim, B. Neta, A sixth-order family of three-point modified Newton-like multiple-root finders and the dynamics behind their extraneous fixed points, Appl. Math. Comput. 283 (2016) 120–140.
[18] H.T. Kung, J.F. Traub, Optimal order of one-point and multipoint iteration, J. Assoc. Comput. Mach. 21 (1974) 643–651.
[19] X. Wang, L. Liu, New eighth-order iterative methods for solving nonlinear equations, J. Comput. Appl. Math. 234 (2010) 1611–1620.
[20] J.F. Traub, Iterative methods for the solution of equations, Prentice-Hall, Englewood Cliffs, 1964.

[21] M.S. Petkovi c′ , B. Neta, L.D. Petkovi c′ , J. Džuni c′ , Multipoint methods for solving nonlinear equations, Academic Press, (2013)
[22] G.V. Balaji, J.D. Seader, Application of interval Newton's method to chemical engineering problems, Rel. Comput. 1 (3) (1995) 215–223.
[23] M. Shacham, An improved memory method for the solution of a nonlinear equation, Chem. Eng. Sci. 44 (7) (1989) 1495–1501.
[24] A. Cordero, J.R.Torregrosa, Variants of Newton's method using fifth-order quadrature formulas, Appl. Math. Comput. 190 (1) (2007) 686–698 .
[25] A. M. Ostrowski, Solution of equations and systems of equations, Academic Press, New York (1960).
[26] M. Shacham, Numerical solution of constrained nonlinear algebraic equations, Int. J. Numer. Method Eng. 23 (1986), 1455–1481.

Authors Index