



ICCM Proceedings

**Proceedings  
of the International Conference  
on Computational Methods**

**(Vol.2, 2015)**

6th ICCM2015, 14<sup>th</sup>-17<sup>th</sup> July 2015, Auckland, New Zealand

**Editors: G. R. Liu and Raj Das**

# ICCM2015

Proceedings of the International Conference on Computational  
Methods (Vol.2, 2015)

6th ICCM2015, 14th-17th July 2015, Auckland, New Zealand

Edited by

**G. R. Liu**

University of Cincinnati, USA

**Raj Das**

University of Auckland, New Zealand

Proceedings of the International Conference on Computational Methods, Vol.2,  
2015

This volume contains full papers accepted by the 6th ICCM2015, held on 14th-  
17th July 2015, at Auckland, New Zealand.

First Edition, December 2015

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 6th ICCM2015, 14th-17th July  
2015, Auckland, New Zealand*, Eds: G.R. Liu, Raj Das, ScienTech Publisher.

Note: The papers/data included in this volume are directly from the authors. The  
editors are not responsible of the inaccuracy, error, etc. Please discuss with the  
authors directly, if you have any questions.

Published by

Sciencetech Publisher LLC, USA

<http://www.sci-en-tech.com/>

## PREFACE

### Dear Colleagues

On behalf of the organising committees, we are delighted to welcome you to the 6th International Conference on Computational Methods (ICCM2015) at Auckland, New Zealand.

The ICCM2015 is an international conference that provides an international forum for the exchange of ideas on recent advances in areas related to computational methods, numerical modelling & simulation, as well as their applications in engineering and science. It will accommodate presentations on a wide range of topics to facilitate inter-disciplinary exchange of ideas in science, engineering and related disciplines, and foster various types of academic collaborations internationally. All papers accepted for publication in the proceedings have been peer reviewed. Papers may also be selected and invited to be developed into a full journal paper for publication in special issues of the journals.

The conference series originated in Singapore in 2004 by Professor GR Liu, followed by ICCM2007 in Hiroshima, Japan; ICCM2010 in Zhangjiajie, China; ICCM2012 in Golden Coast, Australia; and ICCM2014 in Cambridge, UK. This year, the ICCM2015 conference programme covers over 319 oral presentations in 59 technical sessions, including 2 Plenary talks, 9 Thematic Plenary talks, and a number of Keynote talks in technical sessions. These presentations cover a broad range of topics related to computational mechanics, including formulation theory, computational methods and techniques, modelling techniques and procedures, materials, deformation processing, materials removal processes, bio-system, bio-mechanics, processing of new and advanced materials, welding and joining, surface engineering and other related processes.

We would like to express my gratitude to all the members of the Local Organizing Committee, International Scientific and Organization Committee, Honorary Chairmen and Co-Chairmen, who have provided advices and guidance timely in planning and executing this conference. We also would like to use this opportunity to express our gratitude to the Faculty of Engineering of the University of Auckland, and to colleagues for their strong support and encouragement. Sincere thanks and appreciation go to some 100 international reviewers for their prompt review reports on the submitted papers. Our appreciation goes also to all the Mini-Symposium Organizers for their efforts and contributions in the organization.

**G.R. Liu**, University of Cincinnati, USA

**Raj Das**, University of Auckland, New Zealand

## ORGANIZATION COMMITTEES

**Conference Chair:** Raj Das (University of Auckland, New Zealand)

**Conference Co-Chairs:**

Anatoly Yagola  
(Lomonosoy Moscow State University)

Carlo Sansour  
(The University of Nottingham)

Hiroshi Okada  
(Tokyo University of Science)

Jinsong Leng  
(Harbin Institute of Technology)

Ming Li  
(Taiyuan University of Technology)

Sung-Kie Youn  
(Korea Advanced Institute of Science and Technology)

Weihong Zhang  
(Northwestern Polytechnical University)

Xi-Qiao Feng  
(Tsinghua University)

Yuantong Gu  
(Queensland University of Technology)

Zhongwei Guan  
(University of Liverpool)

**Honorary Chairs:**

Gui-Rong Liu (USA)

Mingwu Yuan (China)

**Local Organizing Committee:**

Debes Bhattacharyya (NZ)

Graeme Finch (NZ)

James Lim (NZ)

John Cater (NZ)

Justin Fernandez (NZ)

Mark Battley (NZ)

Martyn Nash (NZ)

Peng Cao (NZ)

Piaras Kelly (NZ)

Raj Das (NZ)

Richard Lin (NZ)

Simon Bickerton (NZ)

**International Scientific Committee:**

Alex Ng (Australia)  
Alexander Korsunsky (UK)  
Alessandro Reali (Italy)  
Amar Khennane (Australia)  
Andrei Kotousov (Australia)  
Bhushan L. Karihaloo (UK)  
Bozidar Sarler (Slovenia)  
Byeng Dong Youn (Korea)  
Chao Jiang (China)  
Chao Xu (China)  
Cheng Yuan (Singapore)  
Chengfeng Li (UK)  
Chin-Long Lee (New Zealand)  
Daya Reddy (South Africa)  
Ekkehard Ramm (Germany)  
Ernian Pan (USA)  
Feng Xiao (Japan)  
Francesco Berto (Italy)  
Francesco Mammoliti (Italy)  
Francesco Noto (Italy)  
George S. Dulikravich (USA)  
Grant Steven (Australia)  
Hengan Wu (China)  
Hiroshi Kanayama (Japan)  
Hongling Ye (China)  
J. N. Reddy (USA)  
Jan Blachut (UK)  
Jihua Guo (USA)  
Jizeng Wang (China)

Joe Petrolito (Australia)  
Jorge Ambrosio (Portugal)  
Julien Yvonnet (France)  
Karol Miller (Australia)  
Kun Zhou (Singapore)  
Leonid Antanovskii (Australia)  
Lian Shen (USA)  
Liangchi Zhang (Australia)  
Lihai Zhang (Australia)  
Maenghyo Cho (South Korea)  
Makoto Ohsaki (Japan)  
Marcelo Colaco (Brazil)  
Martin Veidt (Australia)  
Mike Xie (Australia)  
Mingwu Yuan (China)  
Moubin Liu (China)  
Muneo Hori (Japan)  
Nasr Ghoniem (USA)  
Nasser Hassan Sweilam (Egypt)  
Nguyen Xuan Hung (Vietnam)  
OhJoon Kwon (South Korea)  
Perumal Nithiarasu (UK)  
Qing Li (Australia)  
Qing Zhang (China)  
Qinghua Qin (Australia)  
Robert O. Richie (USA)  
Rooh Khurram (Saudi Arabia)  
Seiichi Koshizuka (Japan)  
Sergey Panin (Russia)

Shankar Kalyanasundaram (Australia)  
Sundararajan Natarajan (India)  
Sung-Kie Youn (South Korea)  
Tinh Quoc Bui (Germany)  
Tomasz Bednarz (Australia)  
Wei Gao (Australia)  
Weihong Zhang (China)  
Weiqiu Chen (China)  
Wingkong Chiu (Australia)  
Xi-Qiao Feng (China)  
Xiong Zhang (China)  
Xu Guo (China)  
Xu Han (China)  
Yusuf C. Toklu (Turkey)  
Ya-Pu Zhao (China)  
Yang-Yao Niu (Taiwan)  
Yijun Liu (USA)  
Yixiang Gan (Australia)  
Yongjin Yoon (Singapore)  
Yuantong Gu (Australia)  
Yue-Sheng Wang (China)  
Yuichi Tadano (Japan)  
Zbigniew Kowalewski (Poland)  
Zhan Kang (China)  
Zhenjun Yang (China)  
Zheng Zhong (China)  
Zheng-Ming Huang (China)  
Zhiqian Zhang (Singapore)  
Zhongwei Guan (UK)

## CONTENTS

Preface	3
Committees	4
Contents	5
Quick modal reanalysis for large modification of structural topology based on multiple condensation model <i>He Jian-jun, Chen Xiang-zi</i>	11
Back-analysis of initial stress fields in underground powerhouse using deformation data observed in field <i>Shouju Li, Zerun Tian</i>	16
On the solution of sine-Gordon solitons via localized kernel-based method <i>Marjan Uddin, Kamran and Amjad Ali</i>	24
Numerical Simulation of Infrared Staking Plastic for an Automotive Part <i>Hong- Seok Park, Trung-Thanh Nguyen</i>	35
An unsteady double diffusive natural convection in an inclined rectangular enclosure with different angle of magnetic field <i>Sabyasachi Mondal and Precious Sibanda</i>	42
Establishing effective criteria to link atomic and macro-scale simulations of dislocation nucleation in FCC metals <i>N. J. Burbery, R. Das, W.G. Ferguson, G. Po, N. Ghoniem</i>	62
Magneto hydrodynamic nanofluid flow and heat transfer along a permeable stretching surface with non uniform heat generation/absorption <i>Md. S. Ansari, R. Nandkeolyar, S. S. Motsa</i>	74
Parametric study of FRP-strengthened reinforced concrete panels under blast loads <i>X. Lin and Y.X. Zhang</i>	86
Frequency Domain Elastodynamic Solutions Using Iterative Coupling of FEM and BEM <i>Delfim Soares Jr., Kleber A. Gonçalves, and José Claudio de Faria Telles</i>	94
The self-assembly of self-interstitial-atoms caused by an electron beam <i>S. T. Nakagawa, K. Ohashi, and K. Sugita</i>	107
Cognition of Parameters' Role on Vertical Control Device for Aerodynamic Characteristics of Aircraft Using Data Mining <i>Kazuhisa Chiba, Taiga Omori, Yasuto Sunada, and Taro Imamura</i>	117
Evidence-theory-based analysis for structural-acoustic field with epistemic uncertainties <i>Jian Liu, Longxiang Xie, Xianfeng Man, and Yongchang Guo</i>	137
Towards wideband mechanical metamaterials: comparing nonlinear oscillator mechanisms <i>A. Banerjee, E.P. Calius, and R. Das</i>	156
Prediction of Human Elbow Joint Torque Based on Improved BP Neural Network <i>Gao yongsheng, Sun xiaoying, Fan jizhuang, Zhu yanhe, and Liu gangfeng</i>	168

Solution of computational acoustics and wave propagation problems using a high order, high resolution coupled compact difference scheme <i>Jitenjaya Pradhan, Y. G. Bhumkar and S. D. Dhandole</i>	175
New model and method simulating effective protection of space vehicles from high-velocity debris <i>A.V. Gerasimov, and Yu. F. Khristenko</i>	186
Development of a Discontinuous Galerkin Method for Supersonic Flow Simulations on Hybrid Mesh <i>SU Peng-hui, ZHANG Liang</i>	193
Mode decomposition method for non-classically damped structures using acceleration responses <i>J.-S. Hwang, S.-H. Shin, and H. Kim</i>	199
An Effective Three Dimensional MMALE Method for Compressible Fluid Dynamics <i>Shudao Zhang, Zupeng Jia, Jun Liu, Jun Xiong, and Haibing Zhou</i>	211
Modelling interaction of incompressible fluids and deformable particles with the Material Point Method <i>Rachel Gelet, Giang Nguyen, Pierre Rognon</i>	216
Hydroelastic Analysis in Frequency Domain and Time Domain <i>Frank Lin, Michael Johnson, Zhenhong Wang, and Nigel White</i>	233
A comparison of two computational methods (RANS and LES) in analysing of ventilation flow through a room fitted with a two-sided windcatcher for free flow condition <i>A. Niktash and B.P. Huynh</i>	255
Mass-redistributed method in the evaluation of eigenfrequency of solid systems <i>Eric Li, ZC He, Xu Xu</i>	263
Viscoelastic Response of Composite Laminated Shells Based on Efficient Higher-Order Theory <i>S-N Nguyen, J. Lee and M. Cho</i>	273
A better way for managing all of the physical sciences under a single unified theory of analytical integration <i>Mike Mikalajunas</i>	279
A Framework for Multiscale Modeling of Warfighter Blast Injury Protection <i>Raj K. Gupta and Andrzej J. Przekwas</i>	324
Validation of a Fast Transient Solver based on the Projection Method <i>Darrin W. Stephens, Chris Sideroff, and Aleksandar Jemcov</i>	331
Novel multimedia architecture design pattern using Audio joiner prototype with delay remover <i>Sachin M. Narangale and G. N. Shinde</i>	351
Innovative development of the RP assisted customized surgical guides in various surgeries <i>Sandeep W. Dahake, Abhaykumar M. Kuthe, Mahesh B. Mawale, Ashutosh D. Bagde</i>	356
Higher-order two-point efficient family of Halley type methods for simple roots <i>Ramandeep Behl and S. S. Motsa</i>	362
Numerical Study of Interaction Between Waves and Floating-Body by MPS Method <i>Y.L. Zhang, Z.Y. Tang, D.C. Wan</i>	374
Development of Microsystems Analysis (Î¼sys) Software Using Hybrid Finite Elements and Direct Solution of Coupled Equations <i>A. Roychowdhury, Kunal D. Patil, Arup Nandy, C.S. Jog, Rudra Pratap and G.K. Ananthasuresh</i>	384

Osteoporotic bone fracture healing under the locking compression plate system <i>L. Zhang, S. Miramini, P. Mendis, M. Richardson</i>	398
Chord rotation demand for Effective Catenary Action of RC Beams under Gravitational Monotonic Loadings <i>Meng-Hao Tsai</i>	403
Evaluation of the Singular Stress Field of the Crack Tip of the Infinite Elastic Plate <i>J.Y. Liu, G. Lin, Z.C.Sun, P. Zhang and Y.Y. Wang</i>	413
Application of Least Squares Stress Stabilization Method on Nodal Integration of RPIM with Tetrahedral Background Cells <i>M.M. Yavuz and B. Kanber</i>	418
Imaging and characterizing damages in metallic plates using Lamb waves <i>C.T. Ng</i>	438
A Non-intrusive Polynomial Chaos For Uncertainty Quantification On Numerical Simulation Of Flows Around Cylinder <i>Yanjin Wang, Shudao Zhang</i>	445
Multi-Resolution MPS Method for 2D Free Surface Flows <i>Z.Y. Tang, Y.L. Zhang and D.C. Wan</i>	449
Numerical Investigation of Influence of Eccentricity on the Hydrodynamics of a Ship Maneuvering into a Lock <i>Qingjie Meng, Decheng Wan and Wenhua Huang</i>	461
SSRLS based Enhanced Impulsive Noise OFDM Suppressor in AWGN Channel <i>Alina Mirza, Sumrin Mehak Kabir, and Shahzad Amin Sheikh</i>	471
Spatial scaling issues in constitutive modelling of geomaterials <i>C.T. Nguyen, G.D. Nguyen, H.H.Bui, and V.P. Nguyen</i>	478
Multiphase model for thermal activity in a multistory residential home <i>C. De Costa and Satoru Ushijima</i>	492
Numerical simulation of flows around jellyfish in a current <i>T. Inomoto, K. Matsuno, M. Yamakawa, S. Asao and S. Ishihara</i>	503
A reliability optimization allocation method considering differentiation of functions <i>X.J. Yi, Y.H. La, H.P. Dong, and P. Hou</i>	515
Acoustic resonance scattering of Bessel beam by elastic spheroids in water <i>Z.X. Gong, W. Li, and Y.B. Chai</i>	526
A Highly sensitive mass sensor using high-mode resonant cantilever with step change in thickness <i>Renjing Gao, Mingli Li, Jian Zhao</i>	535
F-bar aided edge-based smoothed finite element method with tetrahedral elements for large deformation analysis of nearly incompressible materials <i>Y. Onishi, K. Amaya</i>	543
A multi-level meshless method based on an implicit use of the method of fundamental solutions <i>C. Gáspár</i>	547
Bird strike on an engine primary compressor at high rotational speed: Numerical simulations and parametric study <i>Jia Huang, Yulong Li, Zhixue Zhang, Jun Liu, Zhongbin Tang</i>	559



Crashworthiness simulation of multi-cell tubes under oblique impact loads <i>Jianguang Fang, Guangyong Sun, Qing Li</i>	570
Topology optimization of anisotropic constrained damping structures <i>L. Zhao, Q. Deng, S.Y. Liu, M.J. Su, and Y.H. Xu</i>	580
Free Surface Flow Simulation Using Moving-Grid Finite-Volume Method <i>Sadanori Ishihara, Kenichi Matsuno, Masashi Yamakawa, Takeshi Inotomo, and Shinichi Asao</i>	588
An effective improved algorithm for Finite Particle Method <i>Yang Yang, Xu Fei, Li Xiao-ting, Wang Lu</i>	597
Multiscale modeling of multiple-cracking fracture behaviour of engineered cementitious composites <i>Ting Huang, Y.X. Zhang, and Chunhui Yang</i>	606
Numerical study on the thermal characteristics in a Twin Swirl pulverized coal Combustor <i>Yinli LIU, Hao TANG, Yongfen WANG</i>	611
The inadequacy of elastic properties from tensile tests for Lamb wave analysis <i>W.H. Ong, W.K. Chiu, N. Rajic and C. Rosalie</i>	617
Prediction on the onset of global failure of irregular honeycombs under compression <i>Youming Chen, Raj Das, and Mark Battley</i>	624
Development of an anatomically-based SPH model for cranial ballistic injury <i>E.E. Kwon, M.R. Singh, R.D. Vallabh, R. Das, J.W. Fernandez and M.C. Taylor</i>	634
Cementless total hip arthroplasty with and without screw fixation: a randomized study of 78 hips <i>Zhi-qi Zhu, Peng Li, Ming Hou, De-gui Wang, Xue-yang Lin</i>	647
Global sensitivity analysis for structural models by sparse grid integration <i>Zhou Changcong, Zhang Feng, and Wang Wenxuan</i>	656
Preliminary work on the potential of extending structural health monitoring concepts for healing assessment <i>W.K. Chiu, W.H. Ong, M. Russ and Z. Chiu</i>	668
Finite element analysis of different chipbreaker types for turning tool process <i>Dyi-Cheng Chen, Ci-Syong You, Chia-An Tu</i>	675
Study of the characteristic of droplet transfer in laser-MIG hybrid welding based on the phase matching control of laser pulse and arc waveform <i>G. Song, J. Wang, and L.M. Liu</i>	679
Efficient Global Optimization Applied to Multi-Objective Design Optimization of Lift Creating Cylinder Using Plasma Actuators <i>Masahiro Kanazaki, Takashi Matsuno, Kengo Maeda and Mitsuhiro Kawazoe</i>	687
The Failure Behavior of Composite Honeycomb Sandwich Structure with Stringer Reinforcement and Interfacial Debonding <i>Ruixiang Bai, S.H. Bao, X.Y. Liu, and X.H. Liang</i>	696
Seismic earth pressure on a multi-story underground structure <i>Z.Y. Chen, W. Chen, and H. Fan</i>	702
Stress analysis of functionally graded plate under different gradient direction <i>M. Amirpour, R. Das, and E.I. Saavedra Flores</i>	712

Stochastic analysis of a radial-inflow turbine in the presence of parametric uncertainties <i>A. Zou, E. Sauret, J.-C. Chassaing, S. C. Saha, and Y.T. Gu</i>	725
Aerodynamic Analysis of the Airfoil of a Vawt by Using 2D CFD Modelling <i>Luis F. Garcia, Julian E. Jaramillo and Jorge L. Chacón</i>	736
Structural Mechanics Optimization of the AISHa Ion Source <i>F. Noto, L. Andò, L. Celona, F. Chines, G. Ciavola, G. Costa, S. Gammino, O. Leonardi, S. Marletta, D. Nicolosi, G. Torrisi</i>	749
Entropically Damped Artificial Compressibility for SPH <i>Prabhu Ramachandran, Kunal Puri</i>	762
Numerical Investigation of Aerosol Particle Transport and Deposition in Realistic Lung Airway <i>Mohammad S. Islam, Suvash C. Saha, Emilie Sauret, Y.T. Gu, Zoran D. Ristovski</i>	766
The influence of the parameter h in Homotopy analysis method for initial value problems <i>Z. Wang, L. Zou, and Y.P. Qin</i>	775
Test of the GEM Front Tracker for the SBS Spectrometer at Jlab <i>F. Mammoliti, V. Bellini, M. Capogni, E. Cisbani, P. Musico, F. Noto, R. Perrino, L.Re, S. Spinali, M.C. Sutera</i>	787
Probabilistic Structural Analysis of the Composite Crew Module - Substructuring with High Resolution Grid <i>Shantaram S. Pai, Vinod K. Nagpal, and Ian Miller</i>	794
Development of a simplified computational model to study cranial backscatter using SPH <i>E.E. Kwon, A. Malhotra, R. Das, J.W. Fernandez and M.C. Taylor</i>	809
Development of a fully coupled flow-geomechanics simulator for flow in saturated porous media <i>Chao Zhang, Sadiq J. Zarrouk and Rosalind Archer</i>	816
Adaptive combined DE/FE algorithm for analyzing impact fracture problem <i>Wei Xu and Jiro Sakamoto</i>	830
Scaled Boundary Finite Element Analysis of three-dimensional crack configurations in laminate structures <i>Sascha Hell, Wilfried Becker</i>	837
An Investigation of Nano-Particle Deposition in Cylindrical Tubes Under Laminar Condition Using Lagrangian Transport Model <i>M. Babaei, P. Talebizadeh, K. Inthavong, G. Ahmadi, Z. Ristovski, H. Rahimzadeh, R. Brown</i>	847
Parametric reliability sensitivity analysis using failure probability ratio function <i>Wei Pengfei</i>	853
The Clenshaw-Curtis quadrature in isogeometric analysis <i>X.F. Shi, P. Xi, Y.W. Song and X.M. Cai</i>	866
Fluid Models and Parameter Sensitivities: Computations and Applications <i>L. Davis, M. Neda, Fran Pahlevani, and J. Waters</i>	884
A Simple lumped mass-damper idealization for dam-reservoir-foundation system for seismic analysis <i>D.K. Paul, A. Banerjee, R.N. Dubey, M.K. Alam-Chowdhury, A. P. Singh</i>	912

Ball's motion, sliding friction and internal load distribution in a high-speed ball bearing subjected to a combined radial, thrust, and moment load <i>Mário César Ricci</i>	921
Damage prediction of carbon-epoxy composites under shear and three point bending loads using the finite element method <i>K.M. Karumbaiah, R. Das, and S. Campbell</i>	939
A computational approach for the modelling of rolling shear cracks in cross-laminated timber structures <i>E.I. Saavedra Flores, K. Saavedra, and R. Das</i>	951
An Implicit Algorithm for Finite Volume - Finite Element Coupling <i>Davor Cokljat, Aleksandar Jemcov, and Joseph P. Maruszewski</i>	956
Finite element modelling of deformation behavior in incremental sheet forming of aluminum alloy <i>C.P. Jiang and T.H. Huang</i>	971
A ghost particle-based coupling approach for combined finite-discrete element method <i>Hu Chen, Y.X. Zhang and Mengyan Zang</i>	976
Improved edge-based smoothed finite element method (IES-FEM) for Mid-frequency acoustic analysis <i>Z.C. He, G.Y. Li, Eric Li, G.Y. Zhang, Xin Nie</i>	986
High Resolution Numerical Simulation for corner-turning in LX-17 <i>C. Wang, X.Q. Liu</i>	995
Three Dimensional High Order Parallel Investigations on Underwater Explosion <i>C. Wang, J. X. Ding, W. J. Wang and T. Li</i>	1002
Authors Index	1013

# Quick modal reanalysis for large modification of structural topology based on multiple condensation model

He Jian-jun, Chen Xiang-zi

(School of Automobile and Mechanical Engineering, Changsha University of Science and Technology, Changsha 410004, China)

**Abstract:** The paper focuses on the modal reanalysis for large topological modified structure based on the multiple dynamic condensation model. Firstly, basing on multiple dynamic condensation model of original structure and topological modified structures, with combining the independent mass orthogonalization and Rayleigh-Ritz analysis, the proposed method was generated. Finally, two numerical examples were used to show that the presented method can provide quite efficient and high quality approximation result for modal reanalysis even when large modifications of the structural topology are made.

**Key words:** modal reanalysis; large topological modified structure; multiple dynamic condensation; independent mass orthogonalization

## 1. Introduction

In many large-scale and complex structural systems, linear eigenvalue analysis is very important and useful for predicting modal response. However, in order to grasp the dynamic characteristic of huge structure, we usually directly calculate the Eigen-problem by LQ algorithm or subspace method. Therefore, the computational cost may be too time consuming. Modal reanalysis [1] provides efficient and quick numerical procedures for calculating the eigenvalues and eigenvectors due to modifying the multiple properties of the original structure, without having to repeat the dynamic equation problem several times. Thus, the aim of reanalysis is to reduce lots of computational cost in optimization design. For example, in structural topology optimization design with lots of iterations, the direct FEA analysis must be repeated for each modification of the design variables. Therefore, the computational cost may be too time consuming especially for large structure.

However, the modal reanalysis based on multiple dynamic condensation model are rarely studied in the corresponding reanalysis literature [2-8]. so in the paper, we focus on the structural modal reanalysis for large modifications by using multiple condensation model. The aim is to further reduce the global CPU time than the general modal reanalysis with the non-condensation model. Thus the proposed reanalysis method could be an high-efficiency method for large modifications in optimization design.

As for the large modifications, the computational effort is significantly reduced by the proposed approach. The reanalysis procedure is easy to implement and realize. Moreover, the approximate eigensolution of modified structure still has high accuracy even if the original structure and modified structure are reduced model. Section 2 briefly introduces the Eigen-equations for modal reanalysis problem based on reduced model. Section 3 describes the proposed method and its steps. Section 4 discusses the efficiency of the precision and the CPU time of the proposed method by two

numerical examples. Section 5 summarizes our conclusions.

## 2. modal reanalysis formulations for multiple condensation model

### 2.1 Eigenproblem of original structure based on multiple condensation model

In finite element analysis, the natural vibration of undamped original structures with  $m$  DOF leads to a general algebraic eigenproblem

$$\mathbf{K}_m^0 \boldsymbol{\Psi}_{mi}^0 = \lambda_i^0 \mathbf{M}_m^0 \boldsymbol{\Psi}_{mi}^0 \quad (1)$$

where  $\mathbf{K}_m^0$ ,  $\mathbf{M}_m^0$ ,  $\lambda_i^0$  and  $\boldsymbol{\Psi}_{mi}^0$  are the stiffness matrix, the mass matrix, the  $i$ th eigenvalue and  $i$ th normalized eigenvector of the original structure, respectively.

By using the dynamic condensation method in literature [9], the three times condensation model for the original structure is obtained. Due to the limit of space, the reduced process is omitted here. The full iterative process is expressed as the literature [9].

Through the three time dynamic condensation process, finally, the eigenproblem of reduced model for original structure can be expressed as

$$\mathbf{K}_{R_0}^{(i+1)} (\boldsymbol{\Psi}_{11}^0)^{(i+1)} = \mathbf{M}_{R_0}^{(i+1)} (\boldsymbol{\Psi}_{11}^0)^{(i+1)} (\mathbf{A}_{11}^0)^{(i+1)} \quad (2)$$

### 2.2 Eigenproblem of reduced model for topologically modified structure with added DOF

Similarly, the vibration eigenproblem of the complete model for topologically modified structure with added  $n$  DOF is expressed as

$$\mathbf{K}_{m+n} \boldsymbol{\Psi}_i = \lambda_i \mathbf{M}_{m+n} \boldsymbol{\Psi}_i \quad (3)$$

where  $\mathbf{K}_{m+n}$  and  $\mathbf{M}_{m+n}$  are the stiffness and mass matrix of the topologically modified structure, respectively; and the  $i$ th corresponding eigenpair are denoted as  $\lambda_i$  and  $\boldsymbol{\Psi}_i$ .

Basing the dynamic condensation method and its implementation for topologically modified structure, similarly, the stiffness matrix  $\mathbf{K}_R^{(i)}$  and mass matrix  $\mathbf{M}_R^{(i)}$  of topologically modified structure could be obtained as the original structure.

$$\mathbf{K}_R^{(i)} = \begin{bmatrix} \mathbf{K}_{R_0}^{(i)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \Delta \mathbf{K}_{R_{mm}} & \Delta \mathbf{K}_{R_{mn}} \\ \Delta \mathbf{K}_{R_{nm}} & \Delta \mathbf{K}_{R_{nn}} \end{bmatrix}, \quad \mathbf{M}_R^{(i)} = \begin{bmatrix} \mathbf{M}_{R_0}^{(i)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \Delta \mathbf{M}_{R_{mm}} & \Delta \mathbf{M}_{R_{mn}} \\ \Delta \mathbf{M}_{R_{nm}} & \Delta \mathbf{M}_{R_{nn}} \end{bmatrix} \quad (4)$$

## 3. Modal reanalysis of the multiple condensation model

The proposed modal reanalysis method is similar to the method in the reference [10]. Due to the limit of space, the complete modal reanalysis process is omitted here.

## 4. Numerical examples

For the purpose of validation verifications, the eigenvalues of the large topologically modified structure are computed using the direct method (LQ algorithm) and the proposed approximate reanalysis method, respectively. Define the relative error of the eigenvalues by using the criterion:

$$ree = (|\lambda_{ai} - \lambda_{ei}| / \lambda_{ei}) \times 100\% \tag{5}$$

where the exact eigenvalue  $\lambda_{ei}$  is directly calculated by using equation (3); basing on the multiple condensation model in equation (4), and the approximate eigenvalue  $\lambda_{ai}$  is computed by using the approximate modal reanalysis method from the literature [10]. **Example 1** Considering a square spatial prismatic truss structure as shown in Fig.1, with its parameters given by: elasticity modulus is  $E=1.0 \times 10^{11}$ Pa, cross-section area of each rod is  $A=0.0025\text{m}^2$ , mass density is  $\rho=8.90 \times 10^3 \text{ kg/m}^3$ . Considering large topological modification by adding 20 new nodes and 65 members as shown in Fig.2. The four lower-order eigenvalues, which resulted from those three methods, are shown in the Table 1.

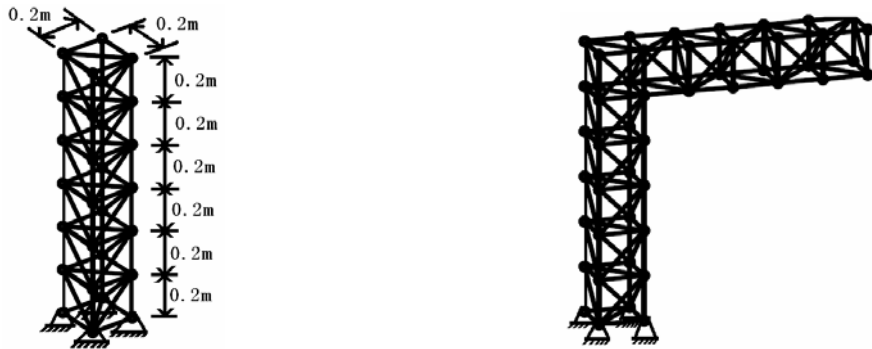


Figure.1. Initial design of spatial truss structure      Figure. 2. Topologically modified structure  
 Table 1. The comparisons of the four lower-order eigenvalues for large modified structure

Mode	The direct method	Literature [11]'s method	Ree (%)	The proposed method	Ree (%)
1	1.15E+4	1.41E+4	22.82	1.15E+4	3.27E-3
2	2.27E+4	3.95E+4	73.86	2.27E+4	1.04E-1
3	9.57E+4	9.73E+4	1.72	9.64E+4	7.39E-1
4	1.28E+6	9.31E+5	625.70	1.29E+6	6.49E-1

Here, the comparisons of computational cost are as follow: the average computational time of exact computation by using the direct method is 0.1508 second; and the average computational time by using the proposed reanalysis method is 0.009514 second. So the saving of computational time is nearly 93.7 percentages.

**Example 2:** Considering an original rectangle bending plate structure, shown in Fig.3, with its material and structural parameters given by: elasticity modulus is  $E=2.1 \times 10^{11}$ Pa, the thickness of plate is  $t=0.01\text{m}$ , mass density is  $\rho=7.80 \times 10^3 \text{ kg/m}^3$ ,

the Poisson's ratio is equal to 0.3, the length and width of plate are 1m and 0.3m, respectively. The original plate is fixed at left, right boundary and discretized into 44 nodes and 30 square plate elements. Considering large topological modifications as shown in Fig.4. The modified plate structure has 76 nodes and 54 elements, with its material parameters given as follow: elasticity modulus is  $E=1.0 \times 10^{11}$  Pa, mass density is  $\rho=8.90 \times 10^3$  kg/m<sup>3</sup>, the Poisson's ratio is  $\mu=0.3$ . The sizes of the modified structure are transformed to 1.2m $\times$ 0.84m.

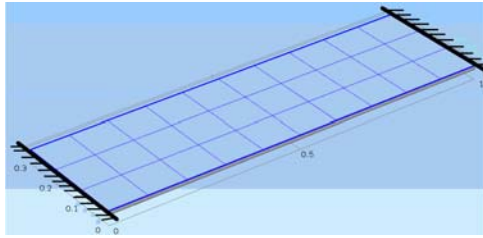


Figure.3. The original bending plate structure

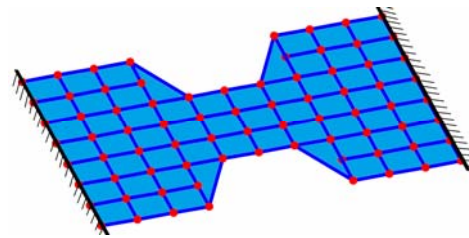


Figure.4. The modified structure

The six lower-order eigenvalues, which resulted from those three methods, are shown in the Table 2.

Table 2. The comparisons of the six lower-order eigenvalues for large modified structure

Mode	The direct method	Literature [11]'s method	Ree (%)	The proposed method	Ree (%)
1	2.86E+4	3.23E+4	12.8	2.86E+4	4.86E-2
2	8.35E+4	8.49E+4	1.64	8.40E+4	4.70E-1
3	1.20E+5	1.52E+5	25.89	1.23E+5	1.83
4	1.61E+5	2.05E+5	27.22	1.66E+5	3.33
5	3.00E+5	4.76E+5	59.21	3.10E+5	3.87
6	4.63E+5	9.08E+5	96.11	4.81E+5	3.94

Here, the comparisons of computational cost are as follow: the average computational time of exact computation by using the direct method is 0.1630 second; and the average computational time by using the proposed reanalysis method is 0.01357 second. So the saving of computational time is nearly 91.7 percentages.

## 5. Conclusion

For large changes in structural topological and parameter modifications with added DOF, an improved modal reanalysis method for the finite element system based on multiple condensation model of the original structure and the modified structure is proposed. By comparing to the method from literature [9], the presented method has higher accuracy and efficiency. From the approximate results of numerical examples and the comparisons of average computational cost, the proposed method not only has high approximation, but also can reduce vast amount of computational efforts than the direct calculation method, so it is much efficient than the direct method. In spite of the complete model or reduced model, the proposed modal method can be used with a general reanalysis for multiple condensation model when large changes of structure

topology and parameters are made, and it may be one of the highest efficient and exact modal reanalysis methods for large topologically modified structure so far.

### **Acknowledgements**

The authors would like to thank for the supports by Natural Science Foundation of China under grant 51305048, and Scientific Research Fund of Hunan Provincial Education Department under grant 11C0045.

### **Reference**

- [1] U. Kirsch, *Reanalysis of Structures—A Unified Approach for Linear, Nonlinear, Static and Dynamic Systems*, Springer, Dordrecht, 2008
- [2] Oded Amir a, Ole Sigmund, Boyan S. Lazarov, et al. Efficient reanalysis techniques for robust topology optimization [J]. *Comput. Methods Appl. Mech. Engrg.*, 2012, 245-246: 217-231
- [3] Liang Ma, Su Huan Chen , Guang Wei Meng. Combined approximation for reanalysis of complex eigenvalues [J]. *Computers and Structures*, 2009, 87: 502-506
- [4] P. Cacciola, G. Muscolino. Reanalysis techniques in stochastic analysis of linear structures under stationary multi-correlated input [J]. *Probabilistic Engineering Mechanics*, 2011, 26: 92-100
- [5] F. Massa, T. Tison, B. Lallemand. Structural modal reanalysis methods using homotopy perturbation and projection techniques [J]. *Comput. Methods Appl. Mech. Engrg.*, 2011, 200: 2971-2982
- [6] Mitsuhiro Kashiwagi. A numerical method for eigensolution of locally modified systems based on the inverse power method [J]. *Finite Elements in Analysis and Design*, 2009, 45: 113-120
- [7] F. Massa, B. Lallemand, T. Tison. Multi-level homotopy perturbation and projection techniques for the reanalysis of quadratic eigenvalue problems: The application of stability analysis [J]. *Mechanical Systems and Signal Processing*, 2015, 52-53: 88-104
- [8] Guoqiang Gao, Hu Wang, Guangyao Li. An adaptive time-based global method for dynamic reanalysis [J]. *Struct Multidisc Optim*, 2013, 48: 355-365
- [9] Zu-Qing Qu, Wenjun Chang. Dynamic condensation method for viscously damped vibration systems in engineering [J]. *Engineering Structures*, 2000, 22(11): 1426-1432
- [10] He Jianjun, Jiang Jiesheng, Kang Xingwu. New method of dynamical reanalysis for large topologically modified structures [J]. *Zhendong Gongcheng Xuebao*, 2007, 20(4): 407-411
- [11] Huang C, Chen S H, Liu Z S. Structural modal reanalysis for topological modifications of finite element systems [J]. *Engineering Structures*, 2000, 22(4): 304-310



## **Back-analysis of initial stress fields in underground powerhouse using deformation data observed in field**

**Shouju Li, Zerun Tian**

State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian 116024, China

\* lishouju@dlut.edu.cn

### **Abstract**

Based on in-situ observed deformation data in underground powerhouse constructions and the response surface approach, back analysis procedure for estimating geotechnical parameters of rock mass is proposed. The relationship between unknown geotechnical parameters of rock mass and deformations of some observing point is approached by the response surface procedure. The excavation process is divided into 13 operation steps for simulation of the stress redistribution due to excavation. The variations of the stress state and deformation field in the rock mass due to the excavation is simulated through a finite element analysis. By defining objective function for back analysis procedure to make the difference between computed deformations and observed ones minimize, the inverse problem of parameter identification is changed into optimization problem. By using optimization algorithm, the stress ratio and elastic modulus are determined. The practical application results show that the forecasted deformations by finite element method agree well with observed ones in-situ. The effectiveness of proposed back analysis procedure for estimating geotechnical parameters of rock mass is validated.

**Keywords:** back analysis procedure, geotechnical parameters of rock mass, response surface approach, in-situ observed deformation data, underground powerhouse construction, stress ratio, elastic modulus.

### **Introduction**

Rock layer is composed of rock block, joint, fracture, fault and porosity. In ideal condition, rock at depth is only subjected to stress resulting from the weight of the overlying strata, and the stress ratio, ratio between horizontal and vertical stresses, is only related to Poisson ratio. Because of the influence of structural movement of the earth's crust, there exists drape, fracture and dislocation in rock mass, and the initial stress field in rock mass is changed. The stress ratio is not satisfied to classical elastic theory. In some sites, the horizontal stress is larger than vertical stress. So, how to estimate stress field in rock mass has increase intensive interests both from the point of engineering application and from scientific investigation. Cai developed a novel method to back-calculate rock mass strength parameters from acoustic emission monitoring data in combination with FEM stress analysis[Cai M., Morioka H., Kaiser P.K., 2007]. Loui examined the existing roadway support systems in seven underground manganese mines through numerical modelling employing theoretically derived in situ stresses to see whether such an approach is reasonable in a practical mining situation[Loui J.P., Jhanwar J.C., Sheorey P.R., 2007]. Kruschwitz evaluated the successful and reliable application of the complex resistivity method for the detection, specification and monitoring of the excavation damaged zone[Kruschwitz S., Yaramanci U., 2004]. Martino conducted extensive rock mechanics research, including work to understand the character and extent of excavation damage. Martino pointed out that damage exists around underground openings and that the damage develops from the energy imparted to the rock by the excavation method and by redistribution of the in situ stress field around the excavated openings[Martino J.B., Chandler N.A.,2004]. Kontogianni thought this induced deformation is not due to effects such as nearby excavations, changes in the hydrological conditions, etc., and to tertiary creep; its distribution along the tunnel axis seems to depend on the potential of host sections to accommodate additional stresses

from neighboring deformation source sections, and it may lead to a progressive, domino-type failure [Kontogianni V.A., Stiros S.C., 2004]. Ishida pointed out that a conventional elastic theory predicts that the maximum compression occurs just behind a stress relaxed region, and gradually decreases as a function of the distance from the chamber wall; thus, this theory cannot be applied to the stress redistribution of a heterogeneous jointed rock mass [Ishida T., Uchita Y., 2000]. Hatzor performed some numerical stress analysis and revealed that in the case of very large span openings, tensile fracture of intact rock may be responsible for instabilities, which may lead to global failure [Hatzor Y.H., Talesnick M., Tsesarsky M., 2002]. Farias used numerical analyses with the finite element method to simulate the full 3-D stages that characterize an NATM tunnel excavation. Some relevant techniques for settlement control were investigated and their relative importance was stated based on the numerical results [Farias M. M.de, Moraes Junior A.H., Assis A.P. de, 2004]. Based on field investigations, Sapigni used two numerical models (FEM and DEM codes) to investigate the overall stability of the excavation and to predict the expected deformation caused by each excavation phase. The measurements of actual deformations, by multi-base extensometer data, are reasonably close to those predicted through the numerical approaches [Sapigni M., Barbera G.La, Ghirotti M., 2003]. Cai used a coupled numerical method to study AE at the Kannagawa underground powerhouse cavern in Japan. Two codes, Fast Lagrangian Analysis of Continua, a finite difference code and Particle Flow Code, a distinct element code, are coupled [Cai M., Kaiser P.K., Morioka H., Minami M., 2007]. The aim of the paper is to propose a back analysis procedure for estimating geotechnical parameters of rock mass, and validate the effectiveness of proposed parameter estimation procedure through comparing field observed deformations with those predicted by numerical approaches.

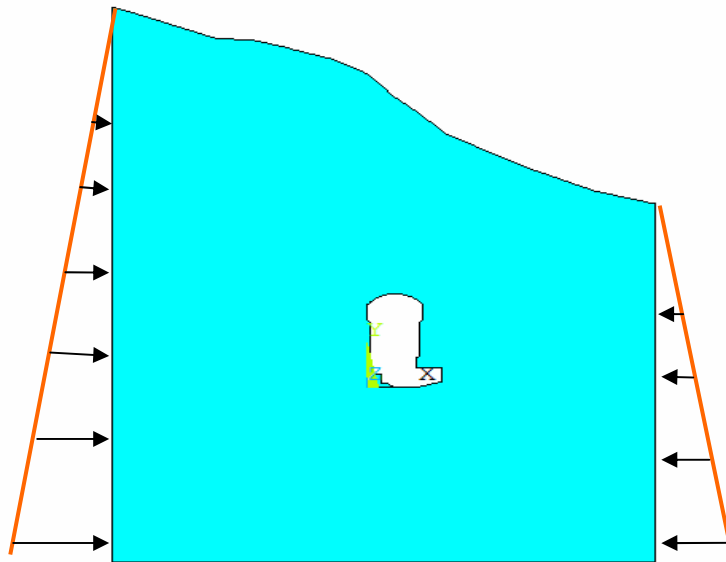
### **Numerical simulations for excavation process of underground powerhouse**

The Baishan hydropower station was constructed on the second Songhua River in the Jilin Province, northeast China, between 1975 and 1983. The concrete gravity arch dam is 149.5m high, and 676.5 m long, as shown in Figure 1. The power plant has an installed capacity of 1700MW. In order to increase the generating ability of the Baishan hydropower station, the pumped storage power station with an output of 300MW was constructed. The underground powerhouse of the pumped storage power station is located in the left bank of river. The underground powerhouses of the first and the second periods of power stations are located in the right and left banks of river, respectively, as shown in Figure 1.



**Figure 1. Baishan Hydropower station.**

The powerhouse cavern is located approximately 120m below ground surface, as shown in Figure 2. The underground cavern's dimensions are 24.5m wide, 50m high and 69m long, as shown in Figure 3. It was excavated by dividing 6 operation layers. In order to monitor the rock displacement and rock stress during excavation period, the rock displacement and the stress of rock anchor were measured. Continuous displacement monitoring on the cavern support helped designers to evaluate the support appropriateness. The finite element method is well suited to solving problems involving heterogeneous or non-linear material properties, since each element explicitly models the response of its contained material. However, finite elements are not well suited to modelling infinite boundaries, such as occur in underground excavation problems. One technique for handling infinite boundaries is to discretize beyond the zone of influence of the excavation and to apply appropriate boundary conditions to the outer edges. Another approach has been to develop elements for which one edge extends to infinity i.e. so-called 'infinity' finite elements. In practice, efficient pre- and post-processors allow the user to perform parametric analyses and assess the influence of approximated far-field boundary conditions.



**Figure 2. FEM model of underground powerhouse and rock mass**

Based on elastic theory, classical formula for calculating initial ground stresses in underground rock mass can be expressed as follows:

$$\sigma_v = \rho gh \quad (1)$$

$$\sigma_h = k\sigma_v \quad (2)$$

$$k = \frac{\mu}{1-\mu} \quad (3)$$

Where  $\sigma_v$  is vertical stress,  $\sigma_h$  is horizontal stress,  $\mu$  Poisson ratio,  $h$  is depth,  $k$  is stress ratio. Most in-situ observation results reveal that the horizontal stress calculated by above equations is not reasonable.

During the construction of the underground powerhouse of the first period (on the right bank of river ) of power station, geotechnical parameters in underground powerhouse on right bank of river were measured in-situ, and listed in Table 1. These data supply references for estimating geotechnical parameters of rock mass in underground powerhouse on left bank of river.

**Table 1: Geotechnical parameters measured in-situ in underground powerhouse on right bank of river**

Parameters	$E/\text{GPa}$	$\mu$	$\sigma_v/\text{MPa}$	$\sigma_h/\text{MPa}$	$k$
No.1	54.1	0.19	14.60	14.06	0.96
No.2	57.0	0.13	7.47	9.45	1.26
No.3	62.0	0.13	7.52	11.35	1.50

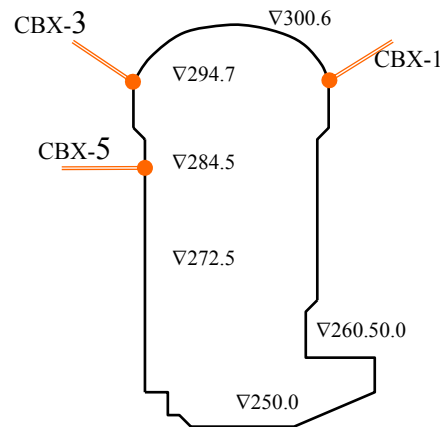
In order to simulate the excavation process of cavern, the initial horizontal stress in rock mass is supposed to linear distribution with depth, as shown in Figure2. On the left boundary of FEM model, the horizontal load is defined as:

$$\sigma_{hl} = k_e \rho g z_l \quad (4)$$

Where  $z_l$  is the depth from ground surface corresponding to left boundary of FEM model.  $k_e$  is the identified stress ratio and will be determined by back analysis procedure. On the right boundary of FEM model, the horizontal load is defined as:

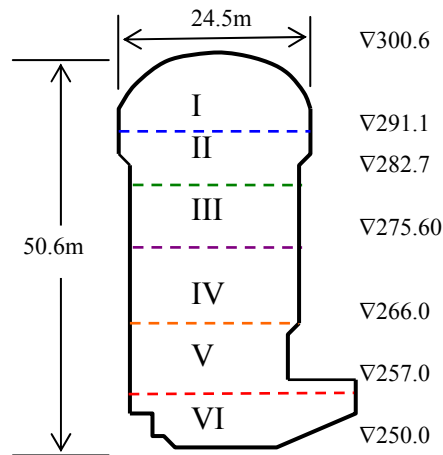
$$\sigma_{hr} = k_e \rho g z_r \quad (5)$$

Where  $z_r$  is the depth from ground surface corresponding to right boundary of FEM model.

**Figure 3. Displacement measuring profile.**

The excavation process is divided into six operation steps for simulation of the stress and deformation redistributions due to excavation, as shown in Figure 5. The variations due to the excavation of the stress state and deformation field in the rock mass can be evaluated through a finite element analysis.

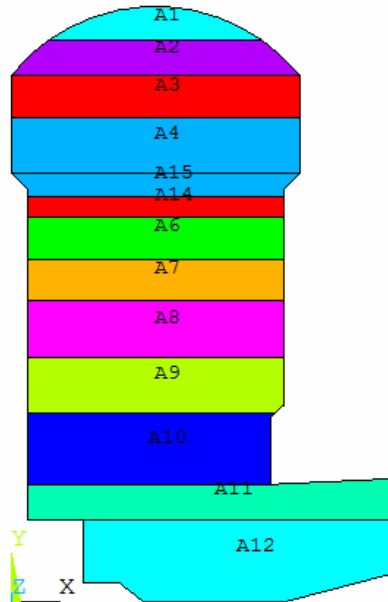
In-situ monitoring during the excavation and at longer intervals after the underground cavern is completed should be regarded as an integral part of the design not only for checking the structural safety and the applied design model but also for verifying the basic conception of the response of the rock mass to tunnelling and the effectiveness of the structural support. In order to monitor the displacements of powerhouse cavern during excavation process, the three sliding micrometers were installed as shown in Figure 3. Figure 4 shows sequences for cavern excavation. Table 2 lists scheduling for cavern excavation operation. ST denotes starting time. ET denotes ending time. Subareas of cavern excavation for FEM simulation is depicted in Figure 5.



**Figure 4. Sequences for cavern excavation**

**Table 2. Scheduling for cavern excavation operation**

No.	I	II	III	IV	V	VI
ST	03-04-01	13-08-17	03-11-27	03-12-28	04-01-14	04-02-14
ET	03-08-16	13-11-26	03-12-23	04-01-13	04-02-11	04-03-11



**Figure 5. Subareas of cavern excavation for FEM simulation**

According to Figure 4 and Table 2, subregions of cavern excavation for FEM simulation are divided into 13 parts, as shown in Figure5.

**Response surface procedures for back analysis of geotechnical parameters**

Based on the response surface method, the relationship between unknown geotechnical parameters of rock mass and displacement of some observing point is approached as:

$$s_k(\bar{x}) = a + \sum_{i=1}^2 b_i \bar{x}_i + \sum_{i=1}^2 c_i \bar{x}_i^2 \quad (6)$$

Where  $s_k(\bar{x})$  is displacement of some observing point  $k$ ,  $a$ ,  $b_i$ , and  $c_i$  are unknown coefficients,  $\bar{x}$  is unknown geotechnical parameter vector after dimensionless procedure.

$$\bar{x} = \{\bar{x}_1, \bar{x}_2\}^T = \{\bar{E}, \bar{k}\}^T \quad (7)$$

$$\bar{E} = \frac{E}{E_I} \quad (8)$$

$$\bar{k} = \frac{k}{k_I} \quad (9)$$

Where  $E_I$  and  $k_I$  denote initial evaluating values of geotechnical parameter according to prior to information induced from observed data in-situ in underground powerhouse of right bank of dam,  $E_I=40\text{MPa}$ ,  $k_I=1.0$ .

Taking the first observing point as an example, the right items of following equations can be calculated using finite element method under the given parameter combinations .

$$s_1^1(\bar{x}) = s(\bar{E}, \bar{k}) \quad (10)$$

$$s_1^2(\bar{x}) = s(\bar{E} + \Delta E, \bar{k}) \quad (11)$$

$$s_1^3(\bar{x}) = s(\bar{E} - \Delta E, \bar{k}) \quad (12)$$

$$s_1^4(\bar{x}) = s(\bar{E}, \bar{k} + \Delta k) \quad (13)$$

$$s_1^5(\bar{x}) = s(\bar{E}, \bar{k} - \Delta k) \quad (14)$$

Where  $\Delta E$  is increment of elastic modulus after dimensionless,  $\Delta E=0.2$ .  $\Delta k$  is increment of stress ratio,  $\Delta k =0.2$ .  $s_1^i$  denotes computed displacement of the first observing point under  $i$ -the parameter combination, which is computed by using finite element method. There exist 5 unknown coefficients and 5 equations. So, the 5 unknown coefficients in response surface functions about first observing point are determined by solving linear equation set with Excel software. The unknown coefficients in response surface functions for other observing point may be deduced by analogy. There are 3 observing points located on the surface of the cavern, as shown in Figure 3. The 8 excavation steps for every observing point are simulated using FEM. So, total 24 series of response surface functions will be determined. And 120 coefficients of response surface function are calculated.

### Estimating Geotechnical parameters of rock mass using optimization algorithm

The objective function of estimating geotechnical parameters of rock mass is defined as Root Mean Square ( RMS ) :

$$\min J = \sqrt{\frac{1}{N * M} \sum_{k=1}^{k=N * M} (s_k(\bar{x}) - s_k^m)^2} \quad (15)$$

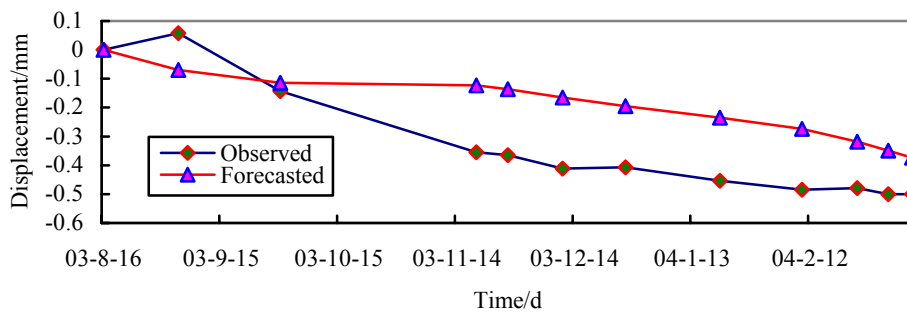
Where  $J$  is objective function,  $s_k^m$  is the observed displacement for the- $k$  observing point,  $N$  is the number of observing point,  $k=3$ ,  $M$  is the observing time for every point,  $M=8$ . In order to estimate geotechnical parameter vector  $x$ , some methods based on the gradient present convergence rates of first order are commonly used. Newton's methods guarantee convergence rates of second order; however, they present the inconvenience of calculating the Hessian and its inverse. The Gauss-Newton method is used as a way to achieve convergence rates closed to second order with lower

computation efforts when compared to Newton’s methods. According to observed deformations induced from cavern excavation and response surface functions, the geotechnical parameters of rock mass are identified, as listed in Table 3.

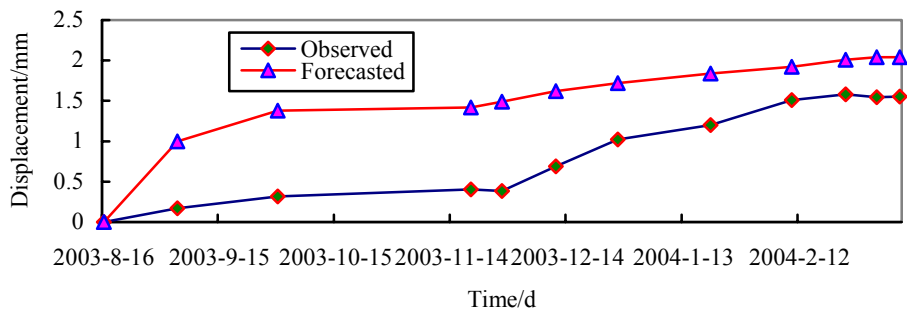
**Table 3. Identified geotechnical parameters of rock mass**

Geotechnical parameters	$E/\text{GPa}$	$k$
Initial estimated value	40.0	1.0
Estimated value by back analysis	56.0	1.2

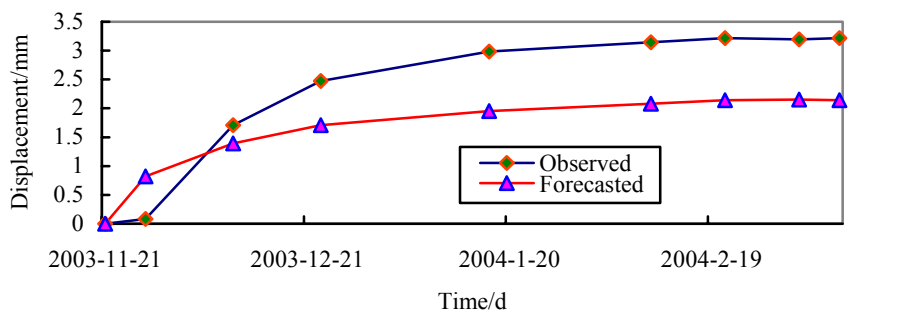
The deformations of different observing point induced from cavern excavation are further simulated by using identified geotechnical parameters of rock mass. Comparisons between observed and forecasted deformations in different excavation step for 3 observing point are shown in Figure 6, 7 and 8.



**Figure 6. Comparison between observed and forecasted deformations in different excavation step for CBX-1**



**Figure 7. Comparison between observed deformations and forecasted ones in different excavation step for CBX-3**



## Figure 8. Comparison between observed and forecasted deformations in different excavation step for CBX-5

### Conclusions

The back analysis procedure for estimating geotechnical parameters in rock mass is presented to demonstrate how the response surface method, finite element method and optimization algorithm should be used as the tools to understand ground stress field and mechanical parameters in rock mass. The observed deformation data of cavern during excavation process supply foundations for solving inverse problem of parameter identification. The effectiveness of proposed inversion procedure is validated by comparing the difference between predicted deformations and observed values in situ.

### References

- Cai M., Morioka H., Kaiser P.K., (2007) Back-analysis of rock mass strength parameters using AE monitoring data, *International Journal of Rock Mechanics & Mining Sciences*, **44**(4),538-549.
- Cai M., Kaiser P.K., Morioka H., Minami M., (2007) FLAC/PFC coupled numerical simulation of AE in large-scale underground excavations, *International Journal of Rock Mechanics & Mining Sciences*, **44** (4), 550-564.
- Farias M. M.de, Moraes Junior A.H., Assis A.P. de, (2004) Displacement control in tunnels excavated by the NATM: 3-D numerical simulations”, *Tunnelling and Underground Space Technology*,**19** (3) 283-293.
- Hatzor Y.H., Talesnick M., Tsesarsky M., (2002) Continuous and discontinuous stability analysis of the bell-shaped caverns at Bet Guvrin, Israel”, *International Journal of Rock Mechanics & Mining Sciences*, **39** (7), 867-886.
- Ishida T., Uchita Y., (2000) Strain monitoring of borehole diameter changes in heterogeneous jointed wall rock with chamber excavation; estimation of stress redistribution, *Engineering Geology* ,**56** (1), 63-74.
- Kruschwitz S., Yaramanci U., (2004) Detection and characterization of the disturbed rock zone in clay stone with the complex resistivity method, *Journal of Applied Geophysics* ,**57** (1) ,63-79.
- Loui J.P., Jhanwar J.C., Sheorey P.R., (2007) Assessment of roadway support adequacy in some Indian manganese mines using theoretical in situ stress estimates, *International Journal of Rock Mechanics & Mining Sciences*, **44**(1) ,148-155.
- Martino J.B., Chandler N.A.,(2004) Excavation-induced damage studies at the Underground Research Laboratory, *International Journal of Rock Mechanics & Mining Sciences*, **41** (8) 1413-1426.
- Kontogianni V.A., Stiros S.C.,(2004) Induced deformation during tunnel excavation: Evidence from geodetic monitoring”, *Engineering Geology*, **79** (1), 115-126.
- Sapigni M., Barbera G.La, Ghirotti M., (2003) Engineering geological characterization and comparison of predicted and measured deformations of a cavern in the Italian Alps, *Engineering Geology* ,**69** (1) ,47-62.



# On the solution of sine-Gordon solitons via localized kernel-based Method

Marjan Uddin<sup>1</sup>, Kamran and Amjad Ali

<sup>1</sup>Department of Basics Sciences, University of Engineering and Technology, Peshawar

\*Presenting author: Kamran

†Corresponding author: Marjan Uddin

## Abstract

In this work, a local kernel based numerical scheme is constructed for numerical solution of sine-Gordon equation in circular domain. The global kernel method resulted the dense differentiation matrices and hence difficult to apply for problem with large amount of data points. The present numerical scheme is local with sparse differentiation matrices, consequently capable of removing the deficiency of ill-conditioning in the global kernel method.

**Keywords:** Local kernel based scheme; two-dimensional sine-Gordon equation

## 1. Introduction

The kernels (RBFs) was first used for solving partial differential equations by Kansa in the year 1990 [Kansa (1990), fluid dynamics-I; Kansa (1990), fluid dynamics-II]. In this original work the fluid mechanics problems were solved by approximating the derivatives by the derivative of MQ kernel functions directly. The differentiation matrices obtained in this method are unsymmetric as well as dense. The dense linear system in the global kernel method solved by Gaussian elimination methods needed  $O(N^3)$  floating point operations. Due to the high resolution for large amount of data points it becomes difficult to solve the problem with global kernel based method. Many robust numerical approximation methods have been developed to overcome this difficulty some of them are the transforms based methods and the multi-pole approaches [Greengard and Strain (1991); Cherrie et al. (2002); Gumerov and Duraiswami (2007)], the domain decomposition methods [Beatson et al. (2001); Kansa and Hon (2000); Li, and Hon (2004)], the partition of unity methods [Wendland (2002)], the greedy algorithms [Hon et al. (2003); Schaback and Wendland (2000); Ling and Schaback (2008)], the multilevel methods [Fasshauer (1999)], and the use of locally supported kernel functions [Wendland (1995); Floater and Iske (1996)]. An other alternative approach to overcome this difficulty was developed by Tolstykh [Tolstykh (2000)], here local kernel interpolants in small domains centered around each node is used to form differentiation weights. This idea has been used to construct various types of local kernel based approximate methods and has been applied successfully to a wide range of problems. These include convection-diffusion [Chandhini and Sanyasiraju (2007); Stevens et al (2009); Sarler and Vertnik (2006); Sarra (2012)], incompressible NavierStokes [Chinchapatnam et al (2009); Shan et al

(2008); Shu et al (2003)], elliptic equations [Tolstykh and Shirobokov (2003); Wright and Fornberg (2006)] and [Wong et al (1999); Xiao and McCarthy (2003); Brown et al (2005)]. In the present work we used the same idea to construct local kernel based numerical scheme for simulating two-dimensional sine-Gordon equation.

The sine-Gordon equation in two space dimension is

$$s_{\tau\tau} = s_{\xi\xi} + s_{\eta\eta} - \mu(\xi, \eta) \sin s, \text{ where } \tau > 0 \text{ and } (\xi, \eta) \in \Omega, \quad (1)$$

with associated initial conditions

$$s(\xi, \eta, 0) = h_1(\xi, \eta), \quad s_\tau(\xi, \eta, 0) = h_2(\xi, \eta), \quad (2)$$

and with boundary condition

$$\beta s = (\xi, \eta, \tau), \quad (3)$$

In science and engineering we always need some robust numerical scheme to solve soliton type equations for large scaled data points in irregular domain for example the sine-Gordon type solitons. Many robust numerical scheme have been developed by many researchers over the years to approximate the sine-Gordon equation. For example the finite difference scheme [Guo et al (1986)], The leapfrog scheme [Christiansen and Lomdahl (1981)], the finite-elements approach [Argyris et al (1991)]. A predictor-corrector scheme [Khaliq, A. Q. M. et al. (2000)], and a split cosine scheme [Sheng, Q. et al. (2005)]. Bratsos [Bratsos (2007)] used a three-time level fourth-order explicit finite-difference scheme for solving sine-Gordon equation. In this work we used local kernel based numerical scheme to approximate the solution of 2d sine-Gordon equation.

## 2. Description of the method

In multivariate scattered data interpolation, we always need to recover an unknown function  $s : R^d \rightarrow R$  from a given set of  $N$  function values  $\{s(\xi_1), s(\xi_2), \dots, s(\xi_N)\} \subset R$ . Where the scattered centers  $\xi_1, \xi_2, \dots, \xi_N \in \Omega$  and  $\Omega \subset R^d$  is arbitrary shaped domain and the centers can be chosen anywhere in the domain. In the local kernel based approximation method, at each center  $\xi_i \in \Omega$ , the local interpolant takes the form

$$s(\xi_i, \tau) = \sum_{\xi_j \in \Omega_i} a_j(\tau) \kappa(\|\xi_i - \xi_j\|), \quad (4)$$

where  $a^i = [a_1, \dots, a_n]$  is a vector of expansion coefficients,  $\kappa: \Omega \times \Omega \rightarrow R$  is a radial kernel defined by  $\kappa(\xi, \xi_j) = \kappa(r_j)$  with  $r_j = \|\xi - \xi_j\|$  and  $\Omega_j \subset \Omega$  is a local domain corresponding to center  $\xi_i$  contains  $n < N$  centers. The corresponding  $N$  number of  $n \times n$  linear systems are given as,

$$s^i = \Lambda^i a^i, i = 1, 2, \dots, N, \quad (5)$$

Where the entries of the matrix  $\Lambda^i$  are  $\{\kappa(\|\xi_k - \xi_j\|)\}^i, k, j \in \Omega_i$ , the matrix  $\Lambda^i$  is called the interpolation matrix, and each system have to be solved for the expansion coefficients. Now to approximate the differential operator  $Ls(\xi, \tau)$ , we have

$$Ls(\xi_i, \tau) = \sum_{\xi_j \in \Omega_i} a_j(\tau) L\kappa(\|\xi_i - \xi_j\|), \quad (6)$$

The expression in (6) may be given in matrix form,

$$Ls(\xi_i, \tau) = \delta^i \bullet a^i, \quad (7)$$

Where  $a^i$  is the  $n \times 1$  vector of expansion coefficients, and  $\delta^i$  is the  $1 \times n$  vector with entries

$$\delta^i = L\kappa(\|\xi_i - \xi_j\|), \xi_j \in \Omega_i. \quad (8)$$

To eliminate the expansion coefficients, we have from equation (5)

$$a^i = (\Lambda^i)^{-1} s^i, \quad (9)$$

we substitute the values of  $a^i$  from (9) in (7) to get,

$$Ls(\xi_i, \tau) = \delta^i (\Lambda^i)^{-1} s^i = \sigma^i s^i, \quad (10)$$

where,

$$\sigma^i = \delta^i (\Lambda^i)^{-1}, \quad (11)$$

is the weight corresponding to center  $\xi_i$ . Hence for all centers locations, we have

$$Ls = \sum s, \quad (12)$$

where,  $\sum$  is  $N \times N$  sparse differentiation matrix, each row of the matrix  $\sum$  contains  $n$  non-zeros elements. After spatial local RBF approximation, we obtained the following system of ODEs

$$\frac{\partial s}{\partial \tau} = F(s). \quad (13)$$

Time integration can be carried out using any ODE solver like ode15s, ode113, ode45 etc from Matlab. In general, ode45 is the best function to apply as a first try for most problems. A good ODE solver will automatically select a reasonable time step  $\delta\tau$  and detect stiffness of the ODE system. For this ODE computation we have used Runge-Kutta method of order four.

### 3. Stability of the local meshless numerical scheme:

In the present local meshless method of lines our numerical scheme is given by

$$s_\tau = \sum s, \quad (14)$$

here the time-dependent PDE is transformed into a system of ODEs in time. The method of lines refers to the idea of solving the coupled system of ODEs by a finite difference method in  $\tau$  (e.g. Runge-Kutta, etc.) The numerical stability of the method of lines is investigated by a rule of thumb. *The method of lines is stable if the eigenvalues of the (linearized) spatial discretization operator, scaled by  $\delta\tau$ , lie in the stability region of the time-discretization operator* [Trefethen and Bau (1997)]. The stability region is a part of a complex plane consisting of those eigenvalues for which the technique produces a bounded solution. In the present meshless method of lines our numerical scheme is given in (13). We can investigate the stable and unstable eigenvalue spectrum for the given model by computing the eigenvalues of the matrix  $\sum$ , scaled by  $\delta\tau$ .

### 4. Choosing a good value of shape parameter:

A variety of kernel functions are available in the literature. In our computation we used the multiquadrics kernel functions,  $\phi(r) = \sqrt{1 + \varepsilon^2 r^2}$ . As usual these RBFs contain a shape parameter and the solution accuracy greatly depends on this parameter. There exist some strategies for the optimization of the shape parameter [Hardy (1971); Franke (1982); Carlson and Foley (1991); Foley (1994); Rippa (1999); Trahan and Wyatt (2003); Fasshauer and Zhang (2007); Scheuerer (2011)]. A condition number may be used to quantify the sensitivity to perturbations of a linear system, and to estimate the accuracy of a computed solution. The conditioning results require that in order for the system matrix to be well conditioned that the shape parameter and minimum separation distance be large. Obviously, both situations cannot occur at the same time. This observation has been referred to as the uncertainty principle [Schaback (1995)]. Incorporating this idea the smallest errors occur when the condition number  $\nu$  of the system matrix is approximately kept in the range  $10^{13} < \nu < 10^{15}$  in our computations. The system matrix is decomposed as  $\mathbf{A}, \mathbf{E}, \mathbf{B} = \text{svd}(\Lambda^i)$ . Here svd is the

singular value decomposition of the interpolation matrix  $\Lambda^i$ .  $\mathbf{A}$ ,  $\mathbf{B}$  are  $n \times n$  orthogonal matrices and  $\mathbf{E}$  is  $n \times n$  diagonal matrix contains the  $n$  singular values of  $\Lambda^i$ , and  $\nu = \|\Lambda^i\| \|(\Lambda^i)^{-1}\| = \max(E)/\min(E)$  is the condition number of the matrix  $\Lambda^i$ . When an acceptable value of shape parameter is returned by the above algorithm, then the svd is used to compute  $(\Lambda^i)^{-1} = (AEB^T)^{-1} = BE^{-1}A^T$  (see [Trefethen and Bau (1997)]). Note that for orthogonal matrices the inverse of the matrix is equal to its transpose. Consequently, we can compute the weights  $\sigma^i$  in (11).

## 5. Application of the method

In this section we apply the method described above to solve the two-dimensional sine-Gordon equation. We considered various types of initial solutions in the form of circular, ring solitons, interaction of two and four circular ring solitons. The two-dimensional sine-Gordon equation has been transformed into a system of two partial differential equations given by

$$s_\tau = p, \quad p_\tau = s_{\xi\xi} + s_{\eta\eta} - \mu(\xi, \eta) \sin s, \quad \tau > 0,$$

with the boundary condition as  $\nabla s \bullet q = 0$ , and  $\nabla p \bullet q = 0$ , and with the initial conditions

$s(\xi, \eta, 0) = h_1(\xi, \eta)$ ,  $p(\xi, \eta, 0) = h_2(\xi, \eta)$ , respectively and where  $q$  is a unit normal vector.

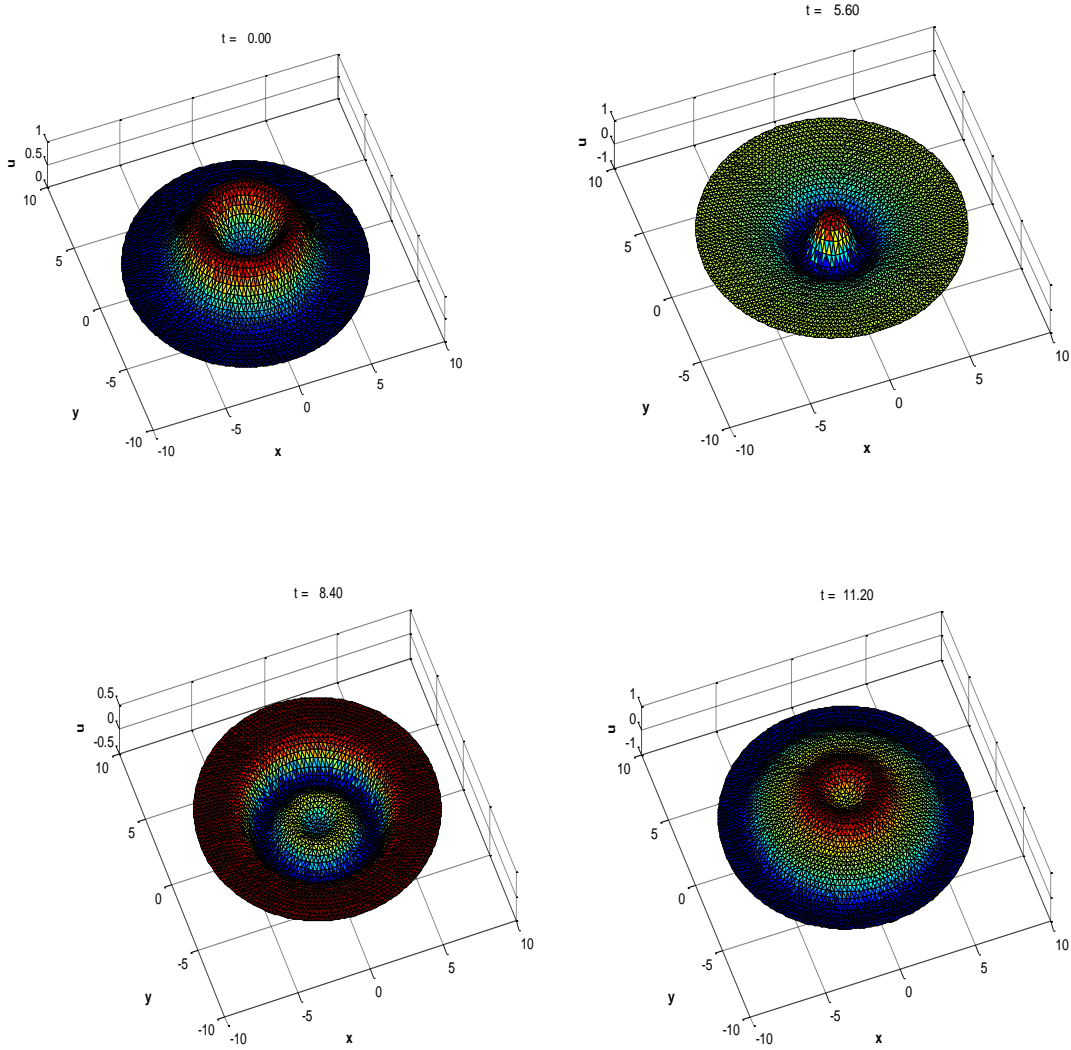
### 5.1. Circular solitons

We apply the proposed method for the case when  $\mu(\xi, \eta) = 1$ , the initial solution is taken as circular solitons [Argyris et al (1991)] given by

$$h_1(\xi, \eta) = 4 \tan^{-1} \exp[3 - \sqrt{\xi^2 + \eta^2}], \quad (16)$$

$$h_2(\xi, \eta) = 0, \quad (17)$$

the problem is solved in the circular domain of radius  $r = 8$  with  $N = 3000$  uniformly distributed interpolation nodes. We select  $n = 200$  points in each local domain  $\Omega_i$  corresponding to each node  $i = 1, 2, 3, \dots, N \in \Omega$ . The time integration is carried out with Runge-Kutta method of order 4 with time step  $\delta t = 0.005$ . The results are obtained by the present numerical method in terms of  $\sin(s/2)$ , where  $s$  is the approximate solution of the given model obtained with present local method. The obtained results at different times are shown in Figure 1.



**Figure 1: Circular soliton: approximate solution in the form of  $\sin(s/2)$  at  $\tau = 0, 5.6, 8.5,$  and  $11.2,$  in the domain  $\Omega$  with radius  $r = 8, N = 3000.$**

### 5.2. Two solitons collision

Here we consider the interaction of two expanding solitons for the choice  $\mu(\xi, \eta) = 1$  and with the initial solutions

$$h_1(\xi, \eta) = \sum_{i=1}^2 f_i(\xi, \eta), \quad h_2(\xi, \eta) = \sum_{i=1}^2 g_i(\xi, \eta) \quad (18)$$

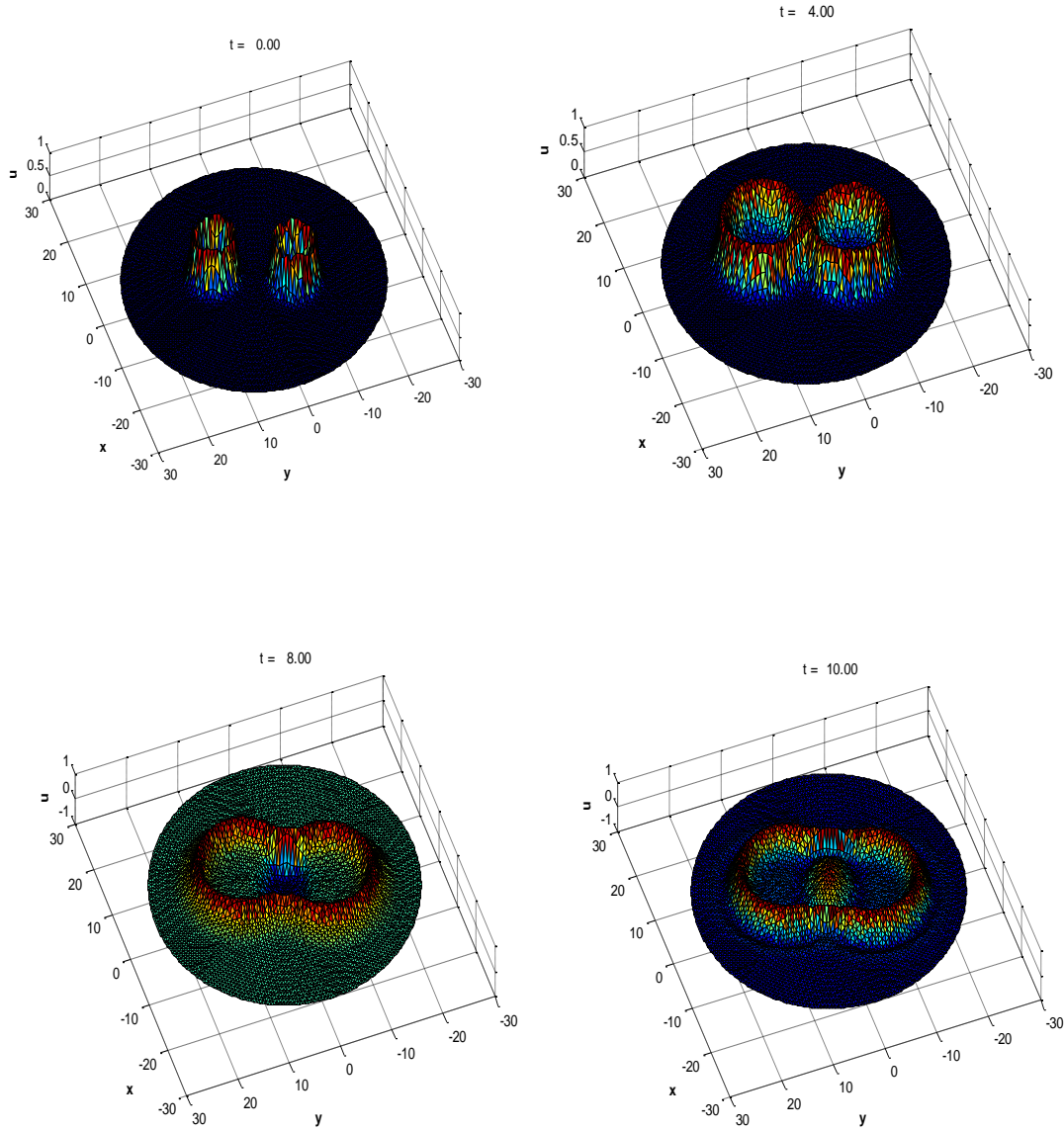
$$f_i(\xi, \eta) = 4 \tan^{-1} \exp[(4 - \sqrt{(\xi \pm 3)^2 + (\eta \pm 7)^2})/0.436], \quad (19)$$

and

$$g_i(\xi, \eta) = 4.13 \operatorname{sech}[(4 - \sqrt{(\xi \pm 3)^2 + (\eta \pm 7)^2})/0.436]. \quad (20)$$

We select  $N = 5000$  number of uniformly distributed interpolation points in the circular domain  $\Omega$  of radius  $r = 25$ . We solved the problem without using the symmetry features that was used in the earlier work [Argyris et al (1991), Sheng, Q.et

al. (2005); Dehghan, M. and Shokri, Ali. (2008)] for simulating the collision of two circular solitons. We take the interpolation points in the whole computation domain to demonstrate the robustness of local radial kernel method. This demonstrates the capability and efficiency of the present method for solving large scale problem in circular domain. The results of the present method are shown in Figure 2.



**Figure 2: Two solitons collision: the function  $\sin(s/2)$ , at  $t = 0, 4, 8$  and  $10$ , in the domain  $\Omega$  with radius  $r = 25$ ,  $N = 5000$ .**

### 5.3. Four expanding solitons collision

The collision of four expanding circular solitons are considered for the choice  $\mu(\xi, \eta) = 1$ , and with the initial solutions

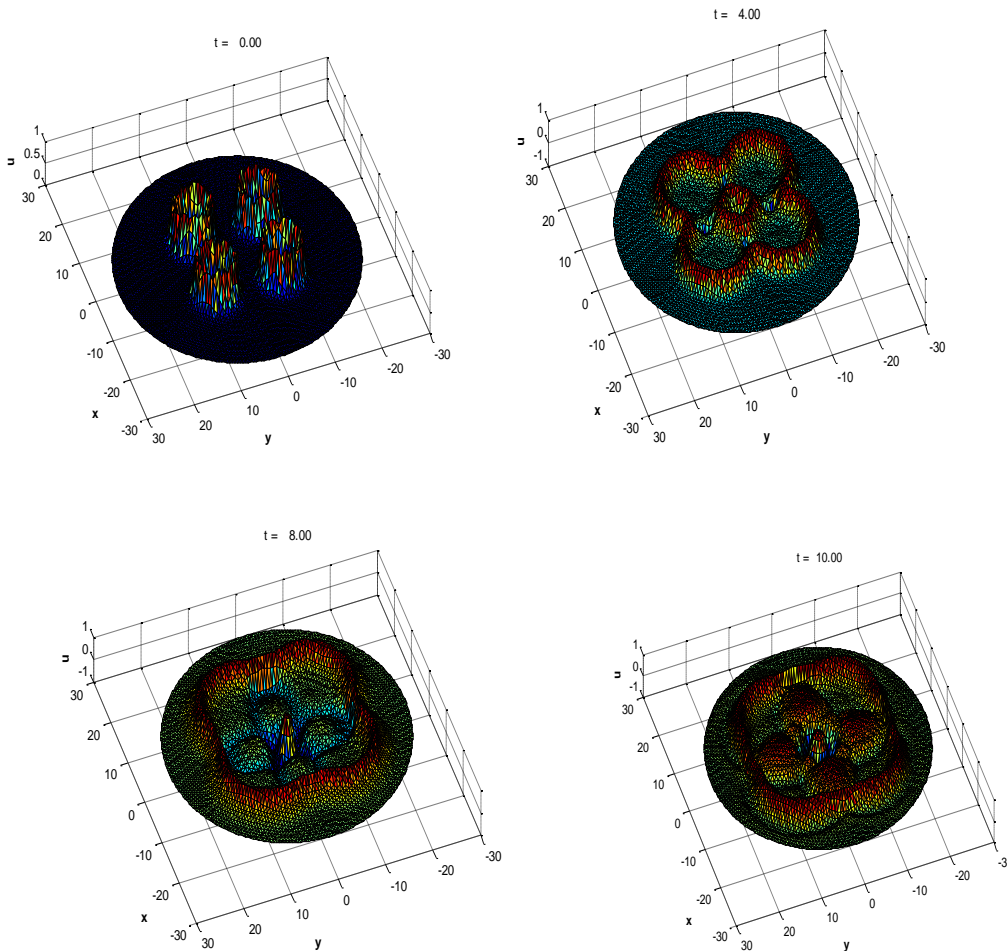
$$h_1(\xi, \eta) = \sum_{i=1}^4 f_i(\xi, \eta), h_2(\xi, \eta) = \sum_{i=1}^4 g_i(\xi, \eta) \quad (21)$$

$$f_i(\xi, \eta) = 4 \tan^{-1} \exp[(4 - \sqrt{(\xi \pm 7)^2 + (\eta \pm 7)^2})/0.436], \quad (22)$$

and

$$g_i(\xi, \eta) = 4.13/\cosh[(4 - \sqrt{(\xi \pm 7)^2 + (\eta \pm 7)^2})/0.436]. \quad (23)$$

This problem is solved in the circular domain of radius  $r = 25$  with  $N = 5000$  uniformly distributed interpolation points. Again we are not using the symmetry features used in the earlier work [Argyris et al (1991), Sheng, Q. et al. (2005); Dehghan, M. and Shokri, Ali. (2008)]. The evolution of the four expanding solitons in times are shown in Figure 3.



**Figure 3: Four expanding solitons collision: the function  $\sin(s/2)$ , at  $t = 0, 4, 8$  and  $10$ , in the domain  $\Omega$  with radius  $r = 25$ , and  $N = 5000$ .**



## 6. Conclusions

In this work we have constructed local kernel based numerical scheme for simulating the two dimensional sine-Gordon equation. As contrary to the global based kernel based methods [Dehghan, M. and Shokri, Ali. (2008)], the present local scheme performed efficiently for large data points in complex shaped domain. The present local method may be used to similar types of time-dependent partial differential equations in irregular shaped domain.

## References

- Kansa, E. (1990) Multiquadric - a scattered data approximations scheme with application to computational fluid dynamics - I. Surface approximations and partial derivative estimates, *Comput.Math. Appl* **19**, 127-415.
- Kansa, E. (1990) Multiquadrics - a scattered data approximation scheme with applications to computational fluid-dynamics-II solutions to parabolic hyperbolic and elliptic partial differential equations, *Comput. Math. Appl* **19** (8-9) 147-161
- Greengard, L. and Strain J. (1991) The fast Gauss transform, *SIAM J. Sci. Stat. Comput* **12**, 79-94.
- Cherrie, J. B., Beatson, R. K. and Newsam, G. N. (2002) Fast evaluation of radial basis functions: Methods for generalized multiquadrics in  $R^n$ , *SIAM J. Sci. Comput* **23**, 1549-1571.
- Gumerov, N. A. and Duraiswami R. (2007) Fast radial basis function interpolation via preconditioned Krylov iteration, *SIAM J. Sci. Comput* **29**, 1876-1899.
- Beatson, R. K., Light W. A., and Billings S. (2001) Fast solution of the radial basis function interpolation equations: Domain decomposition methods, *SIAM J. Sci. Comput* **22**, 1717-1740.
- Kansa, E. J. and Hon, Y. C. (2000) Circumventing the ill-conditioning problem with multiquadric radial basis functions: Applications to elliptic partial differential equations, *Comp. Math. Applic* **39**, 123-137.
- Li, J. and Hon, Y. C. (2004) Domain decomposition for radial basis meshless methods, *Numer. Methods Partial Differential Eq* **20**, 450-462.
- Schaback, R. and Wendland, H. (2000) Adaptive greedy techniques for approximate solution of large RBF systems, *Numer. Algorithms* **24**, 239-254.
- Hon, Y. C., Schaback, R. and Zhou, X. (2003) An adaptive greedy algorithm for solving large RBF collocation problems, *Numer. Algor* **32**, 13-25.
- Ling, L. and Schaback, R. (2008) Stable and convergent unsymmetric meshless collocation methods, *SIAM Journal of Numerical Analysis* **46(3)** 1097- 1115.
- Wendland, H. (2002) Fast evaluation of radial basis functions: Methods based on partition of unity. In *Approximation theory, X* (St. Louis, MO, 2001), *Innov. Appl. Math.*, pages 473-483. Vanderbilt Univ. Press, Nashville, TN,.
- Fasshauer, G. E. (1999) Solving differential equations with radial basis functions: Multilevel methods and smoothing, *Adv. Comput. Math* **11**, 139-159.
- Fasshauer, G. E. and McCourt, M. J. (2012) Stable evaluation of Gaussian RBF interpolants, *SIAM J. Sci. Comput* **34**, 737-762.
- Wendland, H. (1995) Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, *Adv. Comput. Math* **4**, 389-396.
- Floater, M. S. and Iske, A. (1996) Multistep scattered data interpolation using compactly supported radial basis functions, *J. Comput. Appl. Math* **73**, 65-78.
- Tolstykh, A. I. (2000) On using RBF-based differencing formulas for unstructured and mixed structured-unstructured grid calculations, In *Proceedings of the 16th IMACS World Congress*, volume **228**, pages 4606-4624, Lausanne.
- Chandhini G. and Sanyasiraju, Y. V. S. S. (2007) Local RBF-FD solutions for steady convection diffusion problems, *Int. J. Numer. Methods Eng* **72**, 357-378.
- Stevens, D., Power, H., Lees, M., and Morvan, H. (2009) The use of PDE centers in the local RBF Hermitean method for 3D convective-diffusion problems, *J. Comput. Phys* **228**, 4606-4624.

- Sarler, B., and Vertnik, R. (2006) Meshfree Explicit Local Radial Basis Function Collocation Method for Diffusion Problems, *Computers and Mathematics with Applications* **51**, 1269-1282.
- Sarra, Scott A. (2012) A local radial basis function method for advection-diffusion-reaction equations on complexly shaped domains, *Applied mathematics and Computation* **218**, 9853-9865.
- Chinchapatnam, P. P., Djidjeli, K. P., Nair, B. and Tan. M. (2009) A compact RBF-FD based meshless method for the incompressible Navier-Stokes equations, *Proc. IMechE, Part M: J. Eng. For Maritime Env* **223**, 275-290.
- Shan, Y., Shu, C., and Lu, Z. L. (2008) Application of local MQ-DQ method to solve 3D incompressible viscous flows with curved boundary, *Comp. Modeling in Eng. and Sci* **25**, 99-113.
- Shu, C., Ding, H., and Yeo, K. S. (2003) Local radial basis function-based differential quadrature method and its application to solve two dimensional incompressible Navier-Stokes equations, *Comput. Meth. Appl. Mech. Eng* **192**, 941-954.
- Tolstykh, A. I. and Shirobokov, D. A. (2003) On using radial basis functions in a finite difference mode with applications to elasticity problems, *Comput. Mech* **33**, 68-79.
- Wright, G. and Fornberg, B.. (2006) Scattered node compact finite difference-type formulas generated from radial basis functions, *J. Comput. Phys* **212**, 99-123.
- Wong, A. S. M., Hon, Y. C., Li, T. S., Chung, S. L., and Kansa, E. J. (1999) Multizone decomposition for simulation of time-dependent problems using the multiquadric scheme, *Comput. Math. Appl* **37**, 23-43.
- Xiao, J. R., and McCarthy, M. A., (2003) A local Heaviside weighted meshless method for two-dimensional solids using radial basis functions, *Computational Mechanics* **31**, 301-315.
- Brown, D., Ling, L., Kansa, E., and Levesley, J. (2005) On approximate cardinal preconditioning methods for solving PDEs with radial basis functions, *Eng. Anal. Boundary Elem* **29**, 343-353.
- Guo, B. Y., Pascual, P. J., Rodriguez, M. J., and Vzquez, L. (1986) Numerical solution of the sine-Gordon equation, *Appl. Math. Comput* **18**, 1-14.
- Christiansen, P. L., and Lomdahl, P. S. (1981) Numerical solution of 2+1 dimensional sine-Gordon solitons, *Physica 2D* 482-494.
- Argyris, J., Haase, M., and Heinrich, J. C. (1991) Finite element approximation to two-dimensional sine-Gordon solitons, *Comput. Methods Appl. Mech.Eng* **86**, 1-26.
- Khalik, A. Q. M., Abukhodair, B., Sheng, Q. (2000) A predictor-corrector scheme for the sine-Gordon equation, *Numer. Methods Partial Differ. Eqns* **16**, 133-146.
- Sheng, Q., Khalik, A. Q. M., and Voss, D. A. (2005) Numerical simulation of two-dimensional sine Gordon solitons via a split cosine scheme, *Math. Comput. Simulation* **68**, 355-373.
- Bratsos, A. G. (2007) A third order numerical scheme for the two-dimensional sine-Gordon equation, *Math. Comput. Simulation* **76**, 271-282.
- Trefethen, L. N. and Bau, D.(1997) *Numerical Linear Algebra*, first ed., SIAM.
- Hardy, R. L. (1971) Multiquadric equations of topography and other irregular surfaces, *Journal of Geophysical Research* **76**, 1905-1915 .
- Franke F., (1982) Scattered data interpolation: tests of some methods, *Math. Comp* **38**, 181-200.
- Carlson R. E., and Foley, T. A.(1991) The parameter  $r^2$  in multiquadric interpolation, *Comput. Math. Appl* **21**, 29-42.
- Foley T. A. (1994) Near optimal parameter selection for multiquadric interpolation, *J. Appl. Sci. Comput* **1**, 54-69 .
- Rippa S. (1999) An algorithm for selecting a good value for the parameter  $c$  in radial basis function interpolation, *Adv. Comput. Math* **11**, 193-210 .
- Trahan C. J. and Wyatt, R. E. (2003) Radial basis function interpolation in the quantum trajectory method: optimization of the multiquadric shape parameter, *J. Comput. Phys* **185**, 27-49 .
- Fasshauer G. E. and Zhang, J. G. (2007) On choosing optimal shape parameters for RBF approximation, *Numer. Algorithms* **45**, 345-368 .
- Scheuerer M. (2011) An alternative procedure for selecting a good value for the parameter  $c$  in RBF-interpolation, *Adv. Comput. Math* **34**, 105-126 .

- Schaback R. (1995) Error estimates and condition numbers for radial basis function interpolation, *Advances in Computational Mathematics* **3**, 251-264 .
- Dehghan, M. and Shokri, Ali. (2008) A numerical method for solution of the two-dimensional sine-Gordon equation using the radial basis functions, *Math. Comput. Simul.* **79**, 700-715.

## Numerical Simulation of Infrared Staking Plastic for an Automotive Part

†\*Hong- Seok Park<sup>1</sup>, Trung-Thanh Nguyen<sup>1</sup>

<sup>1</sup>School of Mechanical Engineering, University of Ulsan, 93 Daehak-ro, Nam-gu, Ulsan, South-Korea, 680-749

\*Presenting author: phosk@ulsan.ac.kr

†Corresponding author: phosk@ulsan.ac.kr

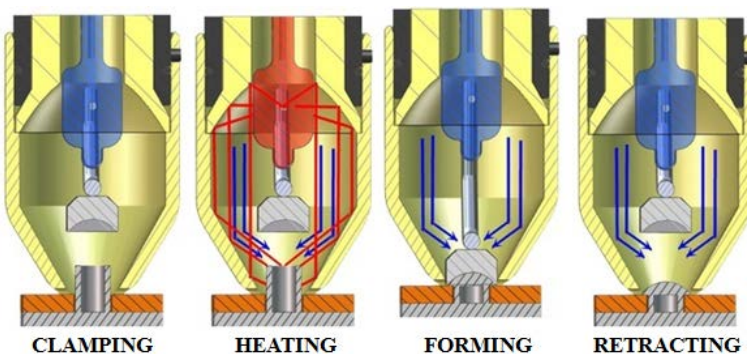
### Abstract

Staking plastic is a new technology that used infrared light as the energy sources in order to assembly the heated plastic parts. In this work, the proposed method coupling between Finite Element Method (FEM), Response Surface Method, and Genetic Algorithm (GA) is employed to simulate and optimize the staking plastic process. An automotive part, namely Door Trim with polypropylene (PP) material is implemented by using the developed technology. The effects of three key process parameters, such as heating time, cooling time, and air flow rate on the mechanical behavior of the joint has been investigated. The simulation procedure based on DEFORM-3D Multi-operation is conducted to integrate sub-process and obtain the numerical results. The optimum values, including heating time, cooling time, and air flow rate are 14 second, 14 second, and 60 (ft<sup>3</sup>/h), respectively. The optimizing results indicated that heating time is more contributed to the tensile force, following by cooling time and air flow rate. The correlation between simulation and experimental results indicates the effectiveness of the proposed method.

**Keywords:** Staking plastic, Door Trim, DEFORM-3D-based Multi-operation, Tensile strength, Genetic algorithm

### Introduction

The infrared staking process was developed by EXTOL by using halogen lighting in order to procedure the joint between different plastic part [Extolinc]. In comparison with the conventional staking technology using heat, hot air, and ultrasonic source, this process provides some advantages, such as higher energy efficiency, product quality, productivity, and better mechanical properties. The process takes place in four basic phases that can be listed as follows (Fig.1):



**Fig.1 Infrared staking process [Extolinc]**

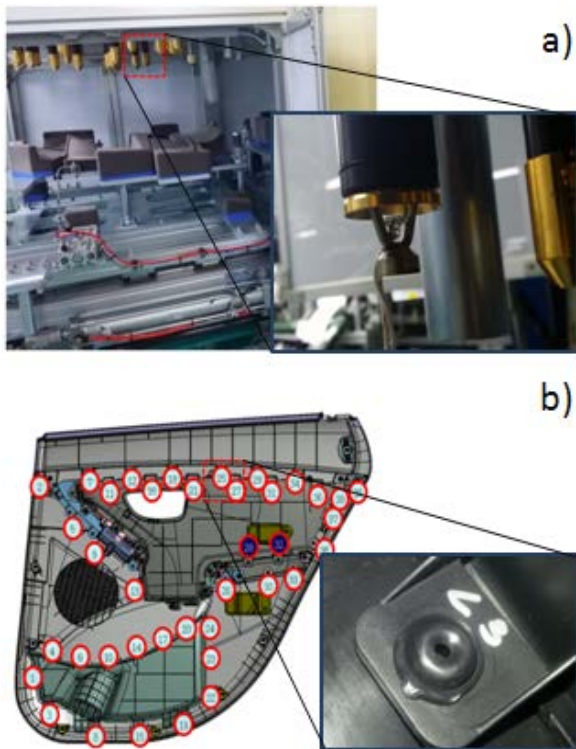
**Heating:** The energy source used in the infrared staking process is a 12-Volt, 100-Watt, technical-grade of a halogen lamp. The reflector directly transfers the infrared energy from the lamp into a stud around the punch. This column of energy travels downward until it is focused on the full perimeter of the boss, heating it from top to bottom. To make this possible, the surface geometry of both the reflector and the concentrator creates a focal area that is centered on the boss. Ideally, the

**Clamping:** The infrared staking module is positioned over the molded stud or boss to ensure perfect contact with the plastic part. The concentrator directly makes contact with the upper part, clamping it to the lower part and holding them in the proper position. As an added benefit, due to the ideal clamping at the stake points, secondary part clamps are not necessary.

stud can be heated as quickly and as evenly as possible.

**Staking & Cooling:** At the end of the heating cycle, the lamp is switched off and the low-impact air cylinder with integrated staking punch is employed which forms over the semi-molten stud. The gold plating and the perfect contact area between the punch and formed part can reduce the temperature of the molten plastic quickly. The cooling air will be flowing into the chamber in order to assist in regulating the punch temperature which prevents the plastic from sticking to it.

**Punch Retracts:** At the end of the punch time, the punch retracts. The module is retracted from the components, and the cycle is completed. The sub-assembly can be immediately handled. At this point, it is important that the plastic has re-solidified to the point where it can maintain its shape and structure.

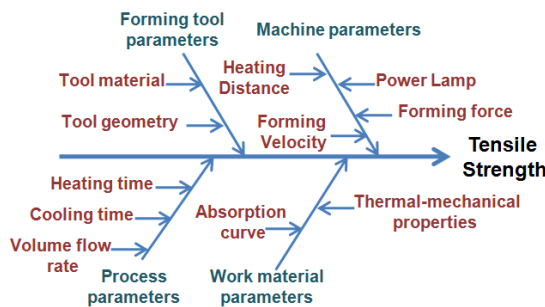


**Fig.2 Machine and completed product; a- Developed machine and dies, b-assembled door trim and completed joint**

In this project, we attempt to develop new machine, which used to assemble the different plastic components of Door Trim. The joint is generated based on PP material and an infrared staking process. The machine, dies, assembled Door Trim, and completed product are shown in Fig.2. Recently, this technology is still a lack of the publications. In addition, identification of optimal values and understanding the effects of process parameters on the mechanical behavior after process are necessary. To solve this problem, this paper proposed an approach to simulate and optimize whole staking process. The key sub-process, including heating, forming, and cooling is integrated into one simulation model. The developed finite element (FE) model considering the data history based on DEFORM-3D Multi operation [Fluhrer, J (2005)] was adopted to conduct the serial numerical experiments. The coupling RSM using Box-Behnken experimental designs [C.F. Jeff et al. (200)] and GA is used to obtain the optimal values. Finally, the numerical results are compared with physical experiment to validate the effectiveness of the proposed method.

**The proposed approach**

The effected parameters on the tensile strength are listed in Fig. 3. In this work, we only consider process parameters due to the fixed of forming tool, machine, and material properties.



**Fig.3 The effected parameters on the tensile strength**

To investigate the material behavior, the sub-process has to integrate into one simulation model, in which the elements have the connection (Fig.4). In other words, the results of preceding process are inputs for the next step. The material state is changed by every step in the production process. The behavior of the final part will depend on the results of each step. After sub-process, the material properties and the stress, and the strain state have changed. Therefore the properties of the joints have to be included in a parameterized way. Transferring data between the simulation processes includes geometric, stress, and strain history [M. Oudjene et al. (2009)], [F. Lambiase et al. (2013)].

Multi-operation in DEFORM-3D was employed to simulate process chains. In the first stage, the simulation model is developed with material model, friction, and heat transfer conditions. After that, the forming process is conducted using boundary conditions. Consequently, these results are transferred to a cooling process considering two kinds of heat transfer, including conduction and convection. Finally, the tensile testing is conducted with the results of the previous stage.

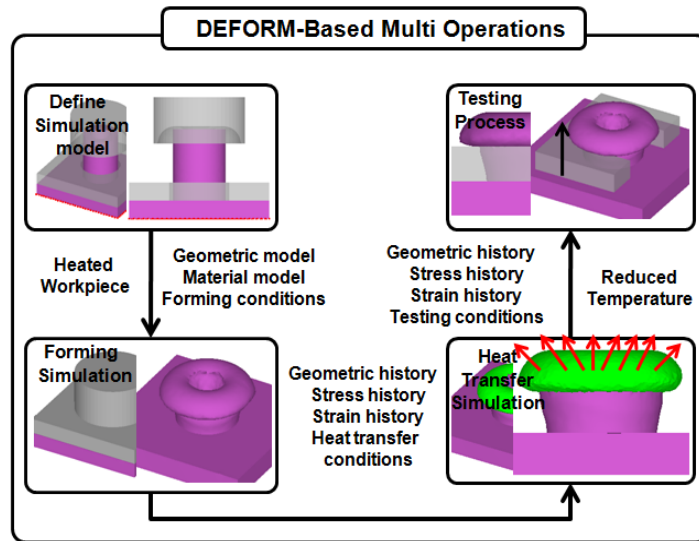


Fig.4 The approach for simulations of the whole process

**Governing equations of forming simulation**

*Material model*

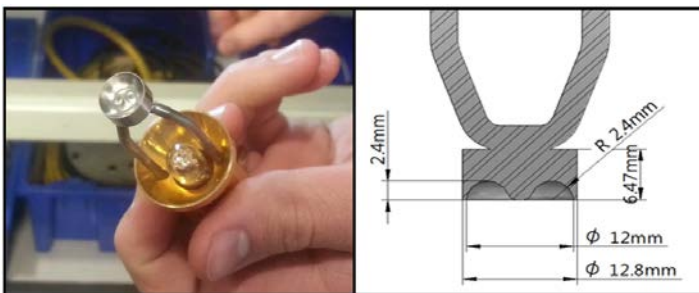


Fig.5 Geometric dimensions of die

A material model includes physical and thermal properties of the work piece and tool. Due to the fact that metal cutting and forming operations perform in high temperature and high strain rates, mechanical properties of work piece material are exactly known. Material properties of plastic-PP and die-S45C are shown in Table 1

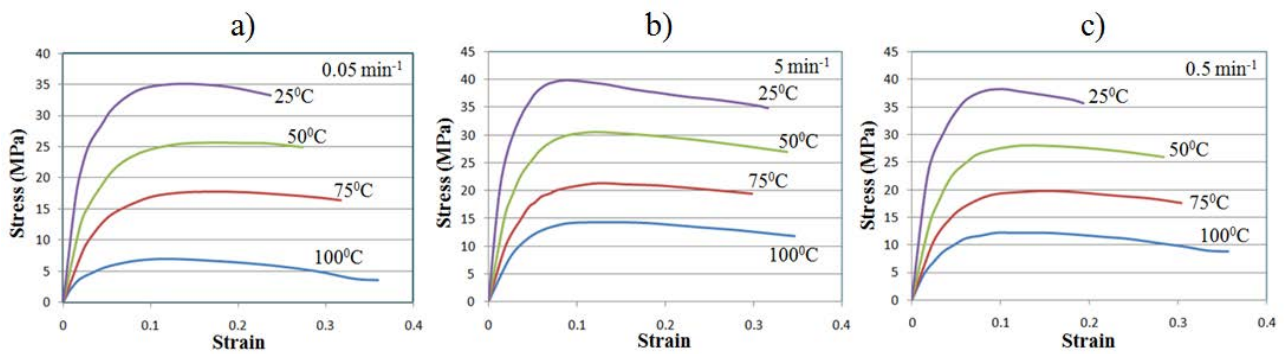
**Table 1 Material property of PP and S45C [Matweb]**

Parameters	PP	S45C
Density (g/mm <sup>3</sup> )	0.892 x10 <sup>-3</sup>	7.85 x10 <sup>-3</sup>
Young modulus (MPa)	1300	250x10 <sup>3</sup>
Poison ratio	0.45	0.29
Thermal conductivity(W/mm °C)	0.16x10 <sup>-3</sup>	49.8
Specific heat (J/g °C)	2	0.486

To save the simulation time and increase the accuracy, the tooling was modeled as perfectly rigid, while the work piece was considered as Elasto-plastic properties with isotropic hardening law and following Vonmises Stress yield criterion. FE model of die and PP work piece are consist of tetrahedral elements with four nodes. Flow stress curves of work piece materials must be used especially in the forming analysis. Flow stress curves of PP material change as a function strain, strain rate, and temperature is shown in Fig.6 at the strain 5min-1, 0.5min-1, and 0.05min-1. This law was selected due to its ability to follow the true behavior of a material [YUANXIN ZHOU et al (2002)]:

$$\bar{\sigma} = (\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T) \quad (1)$$

According to above Eq,  $\sigma$  is flow stress,  $\varepsilon$  and  $\dot{\varepsilon}$ , and T are effective plastic strain, effective strain rate, and temperature, respectively.



**Fig.6 Stress-strain curves of polypropylene; (a) at the 0.05 min-1, (b)at the 5 min-1; (c) at the 0.5 min-1**

### *Friction model*

The friction types allowed are shear and coulomb friction. Shear (sticking); constant shear friction is used mostly for bulk-forming simulations. The frictional force in the constant shear model is defined by:

$$f_s = mk \quad (2)$$

Where  $f_s$  is the frictional stress,  $k$  is the shear yield stress and  $m$  is the friction factor. Coulomb friction is used when contact occurs between two elastically deforming objects (could include an elastic-plastic object, if it is deforming elastically) or an elastic object and a rigid object, generally to model sheet-forming processes. The frictional force in the Coulomb law model is defined by:

$$f_s = \mu p \quad (3)$$

Where  $f_s$  is the frictional stress,  $p$  is the interface pressure between two bodies and  $\mu$  is the friction factor. There must be interfacial pressure between two bodies for frictional force to be present. Due to the complexity of the process, the hybrid friction model was used together both of Coulomb and shear friction models. To obtain the best simulation results, the preliminary simulations were carried out and  $\mu = 0.3$  was selected.

### *Heat transfer*

The initial temperature of the ambient was assumed as 20 °C. The plastic temperature reaches 29 °C (after 8 second), 38 °C (after 14 second) and 43 °C ( after 20 second), respectively. The heat losses to the environment from the free surface of the work piece material are determined by the heat flux:

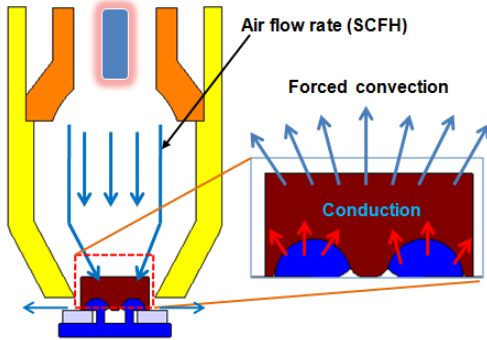
$$Q = h(T - T_0) \quad (4)$$

Where  $h = 4.5 \text{ W/m}^2\text{K}$  is the heat transfer coefficient of the work material. T and  $T_0$  are the

temperature of the work material and ambient temperature, respectively.

**Governing equations of heat transfer simulation**

Two kinds of heat transfer are parallel conducted: convection and radiation. The plastic temperature is reduced due to conductive heat exchange with the die that decreases the temperature by forced convective heat transfer (Fig.7). According to the heat transfer theory and thermal equilibrium relationship by referring to the Fourer Role, the governing equation is established and given as follows:



**Fig.7 Heat transfer in the cooling process**

$$k\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right) = \rho C \frac{\partial T}{\partial t} + Q \quad (5)$$

Where k is the thermal conductivity, ρ is density, C is the specific heat, Q is the latent heat of phase transformation, x, y, z are the coordinates, and it is the time respectively. For the thermal boundary conditions, the adiabatic condition was employed for the others. The heat flow by conduction between the polymer and die [Carlaw, H et al (1959)] can be described as bellow:

$$\dot{Q} = -\frac{kA(\theta_f - \theta_0)}{\sqrt{\pi\alpha t}} \quad (6)$$

Where, θ<sub>0</sub> is considered at the initial time and the final temperature after cooling is assumed to be at θ<sub>f</sub>, α is the thermal diffusivity of the polymer, A denotes the cross sectional area, and t is the cooling time, respectively.

**Optimization process**

Response surface methodology (RSM) is adopted to establish a relationship between process parameters and the performance of objective functions. RSM is a well-known method with higher accuracy and better ease-of use than other popular meta-models, such as radial basis function and kriging model. The second-order RSM model is suitable for modeling the moderate non-linear behavior with few design variables. Prior to the optimization process, relationships between process parameters and objective functions should be created. In this research, the experimental plans are generated using the stipulated conditions using the Box-Behnken experimental designs with 17 runs. Box-Behnken experimental method is one of the effective designs based on multi-dimensional sphere and all the design points lie on a same sphere with at least three or six runs at the center point. The optimization process was resolved based on explicit equations in regression that were obtained through the previous approximation.

The process parameters, including heating time, cooling time, and air flow rate with three levels are listed in Table 2. The sequential simulations are conducted by using boundary conditions to obtain the simulation results that can be used to generate the mathematic model of tensile force (Eq.7).

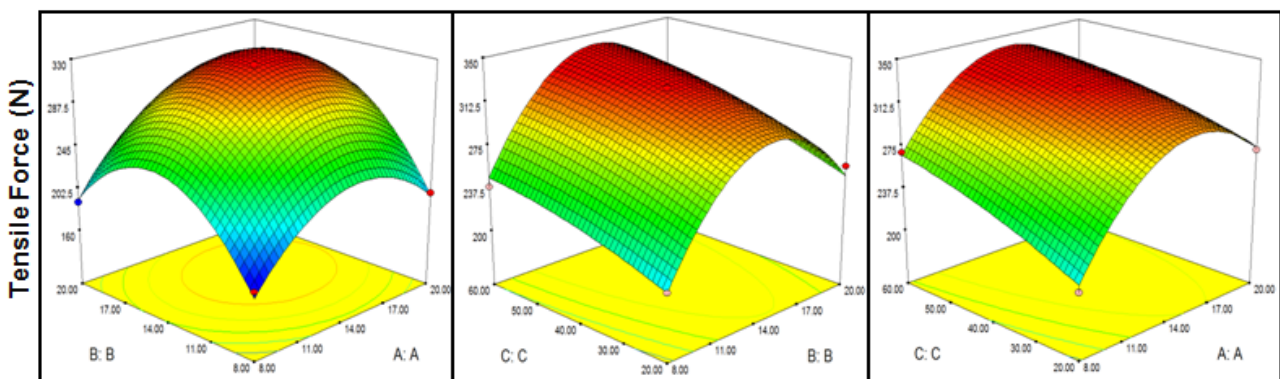
**Table 2 The level of parameters**

Code	Parameters	Value
A	Heating time (s)	8-14-20
B	Cooling time (s)	8-14-20
C	Air flow rate (ft <sup>3</sup> /h)	20-40-60

$$F=630.38+47.65*A+63.32*B+5.53*C+0.2286*A*B-0.0536*B*C-1.6*A^2-2.2*B^2-0.041*C^2 \quad (7)$$

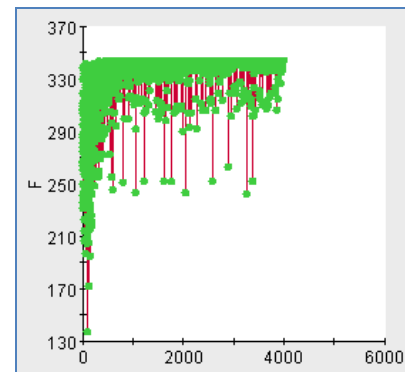


Fig.8 shows the response surfaces, which show the effect of three key process parameters at the center point in the design space of the tensile force. It can be easily seen from the picture that heating time and cooling time have a large effect on the tensile strength, while volume air flow rate has a little effect on the objective function. The main reasons can be explained as follows. Too low heating time results in poor heat transfer and poor mixing material, thus causing of weak joint. Increasing the heating time increases the heat input to the material, resulting in better fusion, consequently, joint strength increases. However, a further heating time combining with forming pressure decreases the joint strength due to material decomposed and a weak joint is formed. Increasing the cooling time results in reducing more temperature that increases the joint strength. The input energy can be loss due to the diffusion at the contact area. Thus, higher cooling time may lead to decrease the tensile strength. Air flow rate has a significant effect to the tensile force. Increasing the air flow rate leads to reduce temperature and to improve the rigidity of workpiece.



**Fig.8 The effects of process parameters on the tensile forces**

Process parameters such as heating time, cooling time, and air flow rate have significant and complex effects on objective function. These process parameters also have contradictory effects on responses. To solve the optimization problem, the evolution algorithm, namely Multi-Island GA is adopted. Specific parameters were population size, number of generations, crossover probability, crossover distribution index, and mutation distribution index, with values of 10, 20, 0.9, 20, and 100, respectively. Figure 9 describes the history of the NSGA II-based optimization process.



**Fig.9 History of GA solving process**



**Fig.10 Tensile testing at the company**

The optimal parameters, including heating time, cooling time, and air flow rate are 14, 14 second, 60, respectively. To validate the proposed method, the physical staking process was conducted in HANIL EHWA-Door Trim Manufacturer in South Korea. The error between simulation and experimental results is lower 10% that indicates the effectiveness of the proposed method.

Table 3 Optimization values

Design variables			Tensile Force (N)	
A (s)	B (s)	C (ft <sup>3</sup> /h)	Simulation	Experiment
14	14	60	348.81	362.8

## Conclusion

This paper presents a new approach to simulate and optimize infrared staking process. The FE model considering the interaction between sub-process has been developed using DEFORM-3D in order to investigate the effects of process parameters on the mechanical behavior. The maximizing tensile force can be achieved at optimal parameters in 14, 14 second, and 60 (ft<sup>3</sup>/h) of heating time, cooling time, and air flow rate, respectively. The optimal results also indicated that heating time is more contribution to tensile strength compared to the others. The proposed method can be used as the effective approach to simulate and optimize the process chains. The effects of forming tool geometry on the mechanical behavior of the joint will be analyzed in future work.

## Acknowledgement

This work was supported by the Research Fund of the University of Ulsan (2015) and Brain Korea 21 plus program.

## References

- Information <http://www.extolinc.com/infrastake>
- Fluhrer, J., DEFORMTM-3D Version 5.0 User's Manual, Scientific Forming Technologies Corporation, 2005
- C.F. Jeff, R. C. Z., Y. Lan (2000), Experiment Design & Analyses and Parameters Optimization, Wiley.
- M. Oudjene, L. Ben-Ayed, A. Delamézière, J.-L. Batoz (2009) Shape optimization of clinching tools using the response surface methodology with Moving Least-Square approximation, *Journal of Materials Processing Technology* **209**, 289-296
- F. Lambiase, A. Di Ilio (2013) Optimization of the clinching tools by means of integrated FE modeling and artificial intelligence techniques, *Procedia CIRP* **12** 163-168
- YUANXIN ZHOU, P. K. MALLICK (2002), Effects of Temperature and Strain Rate on the Tensile Behavior of Unfilled and Talc-Filled Polypropylene. Part I: Experiments, *Polymer Engineering & Science* **42**, 2461-2470
- Matweb, Material Property Data.
- Carlaw, H. and Jaeger, J., (1959) Conduction of heat in solids, *Oxford at the Clarendon Press* **2**, 17-25.

# An unsteady double diffusive natural convection in an inclined rectangular enclosure with different angles of magnetic field

Sabyasachi Mondal \*, Precious Sibanda

School of Mathematics, Statistics and Computer Science

University of KwaZulu-Natal, Pietermaritzburg

Private Bag X01 Scottsville 3209, South Africa

December 12, 2015

## Abstract

An unsteady double-diffusive natural convection flow in an inclined rectangular enclosure subject to an applied magnetic field and heat generation parameter is studied. The enclosure is heated and concentrated from one side and cooled from the adjacent side. The other two sides are adiabatic. The governing equations are solved numerically using a staggered grid finite-difference method to determine the streamline, isotherm and iso-concentration contours. We have further obtained the average Nusselt numbers and average Sherwood numbers for various values of buoyancy ratio and different angles of the magnetic field by considering three different inclination angles of the enclosure while keeping the aspect ratio fixed. The results indicate that the flow pattern, temperature and concentration fields are significantly dependent on the buoyancy ratio and the magnetic field angles. It is found that different angles of the magnetic field suppress the convection flow and its direction influences the flow patterns. This leads to the appearance of inner loop and multiple eddies.

**Keywords:** Double- diffusive natural convection; Inclined rectangular cavity; Magnetic field angles; Buoyancy ratio.

## 1 Introduction

The study of magnetic field effect and buoyancy ratio on double diffusive natural convection in a fluid-saturated cavity has received considerable attention in recent years due to its wide variety of applications in engineering and technology processes such as solar energy collection, nuclear reactor insulation, cooling of electronic devices, furnaces, drying technologies and crystal growth in liquids, etc. As the Lorentz force suppresses the convection currents by reducing the velocities when the fluid is electrically conducting and exposed to a magnetic field, an external magnetic field is used as a flow control mechanism in manufacturing industries. Rudraiah *et al.* (1995) studied the effect

---

\*Email id: sabya.mondal.2007@gmail.com

of a magnetic field on free convection inside a rectangular enclosure. They found that a circular flow was formed with a relatively weak magnetic field, the convection was suppressed and the rate of convective heat transfer decreased when the magnetic field strength was increased. Garandet *et al.* (1992) analyzed the effect of magnetic field on buoyancy driven convection in a rectangular enclosure. Sarris *et al.* (2005) examined two-dimensional unsteady simulations of MHD natural convection of a liquid-metal in a laterally and volumetrically heated square cavity. Mansour *et al.* (2010) studied the effects of an inclined magnetic field on the unsteady natural convection in an inclined cavity filled with a fluid saturated porous medium considering heat source in the solid phase. Al-Najem *et al.* (1998) showed that an increase in the Hartmann number causes reduction in the heat transfer rate from cavity sidewalls. Ece and Büyük (2006) found solutions for the velocity and temperature fields inside the rectangular enclosure and determined the effect of the strength and direction of the magnetic field, the aspect ratio and the inclination of the enclosure on the transport phenomena. Later, Ece and Büyük (2007) investigated the steady natural convection flow in an inclined square enclosure with differentially heated adjacent walls under the influence of magnetic field. Jordan (2007) studied the effects of thermal radiation and viscous dissipation on MHD unsteady free-convection flow over a semi-infinite vertical porous plate. He examined the velocity, temperature, local skin-friction and local Nusselt number for various physical parameters like the Eckert number, magnetic number and suction (or injection). The effect of heat-generation/absorption in an enclosure in the presence of magnetic field also plays an important role in convective flows. Grosan *et al.* (2009) discussed the effects of magnetic field and internal heat generation on the free convection in a rectangular cavity filled with a porous medium. Thereafter, Mahapatra *et al.* (2013a) numerically examined the effects of buoyancy ratio and thermal Rayleigh number on double diffusive natural convection in a cavity when boundaries are uniformly and non-uniformly heated and concentrated. Rahman and Sharif (2003) investigated the laminar natural convection in differentially heated inclined rectangular enclosures of various aspect ratios. Chamkha and Al-Naser (2001) considered laminar double-diffusive convective flow of a binary gas mixture in an inclined rectangular enclosure filled with a uniform porous medium. A numerical investigation of double-diffusive laminar mixed convection in an inclined cavity has been studied by Teamah *et al.* (2011). Wang *et al.* (2008) studied the natural convection heat transfer in an inclined porous cavity with time-periodic boundary conditions numerically. Teamah *et al.* (2013) studied double-diffusive convective flow in an inclined rectangular enclosure with the shortest sides being insulated and impermeable. On the other hand, studies on the combination of radiative heat transfer and heat generation with convection or conduction were intensified because the effect of the radiation heat transfer and heat generation is more important, particularly in the presence of a participating medium and/ or radiative surfaces with large emissivities [Mahapatra *et al.* (2012; 2011)]. Recent developments in hypersonic flights, missile reentry rocket combustion chambers and gas cooled nuclear reactors have focused attention of researchers on thermal radiation, heat generation and emphasize the need for inclusion of heat transfer in these processes. Moufekkik *et al.* (2013) investigated double-diffusive natural convection and radiation in an inclined cavity using lattice Boltzmann method. After that, Mahapatra *et al.* (2013b) studied mixed convection flow in an inclined enclosure under magnetic field with thermal radiation and heat generation. They had studied the radiation and heat generation effects in an inclined cavity with different magnetic field angles. Recently, Mondal and

Sibanda (2015a) studied the effects of buoyancy ratio on unsteady double-diffusive natural convection in a cavity filled with a porous medium with non-uniform boundary conditions. Again, Mondal and Sibanda (2015b) examined the effects of buoyancy ratio on an unsteady double diffusive convection in an inclined rectangular lid-driven enclosure with different magnetic field angles and non-uniform boundary conditions.

Not much attention has been given to the study of the effects of the buoyancy ratio on double diffusive natural convection flow when the magnetic field acts at different angles in an inclined rectangular enclosure with heat generation and uniform boundary conditions (i.e. the rectangular enclosure is heated and concentrated from the left vertical wall and cooled from the top wall by keeping other walls in adiabatic state). The present study deals with the effects of the buoyancy ratio on unsteady double-diffusive natural convection in an inclined rectangular enclosure with different angles of the magnetic field in the presence of a heat generation parameter. The streamline pattern, isotherms, iso-concentrations, average Nusselt and Sherwood numbers are presented graphically and in tabular forms.

## 2 Governing equations and boundary conditions

An unsteady-state flow in two-dimensional rectangular cavity of height  $H$  and length  $L$  as shown in Fig.1 is considered. The thermophysical properties of the fluid are assumed to be constant except the density variation in the buoyancy force, which is approximated according to the Boussinesq approximation. This variation, due to both temperature and concentration gradients, can be described by the following equation:

$$\rho = \rho_0[1 - \beta_T(T - T_c) - \beta_C(C - C_c)], \quad (1)$$

where  $\beta_T$  and  $\beta_C$  are the thermal and concentration expansion coefficients, respectively. The angle of inclination of the enclosure with the horizontal line in the counter-clockwise direction is denoted by  $\phi$ . The magnetic field strength  $B_0$  is applied at an angle  $\varphi$  with respect to the coordinate system. The right and the bottom walls are insulated and the fluid is isothermally heated and cooled at the left and top walls with uniform temperature of  $T_h$  and  $T_c$ , respectively. The magnetic Reynolds number is assumed to be small and the induced magnetic field due to the motion of the electrically conducting fluid is neglected Shercliff (1965). The Joule heating of the fluid and the effect of viscous dissipation are also negligible.

The governing equations under the Boussinesq approximation are written as:

$$\frac{\partial U}{\partial X} + \frac{\partial V}{\partial Y} = 0, \quad (2)$$

$$\begin{aligned} \rho_0 \frac{\partial U}{\partial t} + \rho_0 \left( U \frac{\partial U}{\partial X} + V \frac{\partial U}{\partial Y} \right) &= -\rho_0[1 - \beta_T(T - T_c) - \beta_S(C - C_c)]g \sin \phi \\ -\frac{\partial P}{\partial X} + \mu \left( \frac{\partial^2 U}{\partial X^2} + \frac{\partial^2 U}{\partial Y^2} \right) &+ \sigma B^2(V \sin \varphi \cos \varphi - U \sin^2 \varphi), \end{aligned} \quad (3)$$

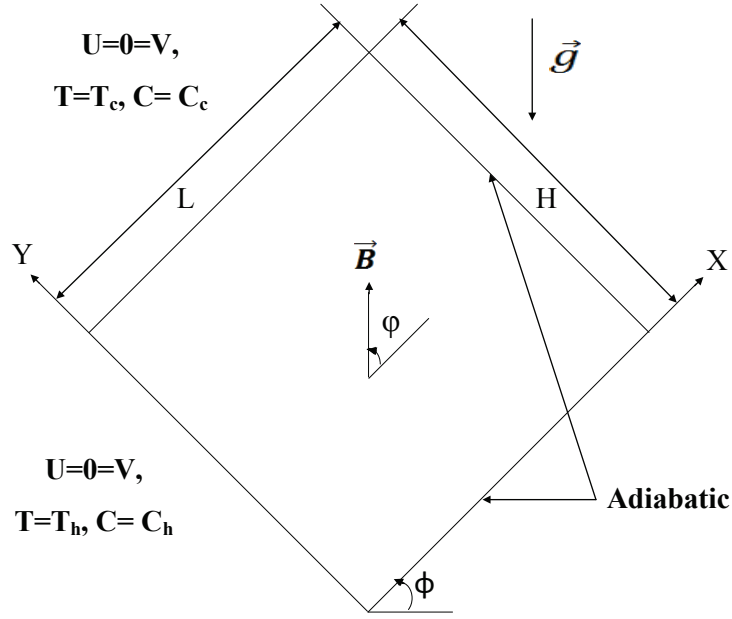


Figure 1: Geometry and the coordinate system

$$\begin{aligned} \rho_0 \frac{\partial V}{\partial t'} + \rho_0 \left( U \frac{\partial V}{\partial X} + V \frac{\partial V}{\partial Y} \right) &= -\rho_0 [1 - \beta_T (T - T_c) - \beta_S (C - C_c)] g \cos \phi \\ - \frac{\partial P}{\partial Y} + \mu \left( \frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2} \right) + \sigma B^2 (U \cos \phi \sin \phi - V \cos^2 \phi), \end{aligned} \quad (4)$$

$$\frac{\partial T}{\partial t'} + U \frac{\partial T}{\partial X} + V \frac{\partial T}{\partial Y} = \alpha \left( \frac{\partial^2 T}{\partial X^2} + \frac{\partial^2 T}{\partial Y^2} \right) + \frac{Q}{\rho_0 C_p} (T - T_c), \quad (5)$$

$$\frac{\partial C}{\partial t'} + U \frac{\partial C}{\partial X} + V \frac{\partial C}{\partial Y} = D \left( \frac{\partial^2 C}{\partial X^2} + \frac{\partial^2 C}{\partial Y^2} \right). \quad (6)$$

The associated boundary conditions are

$$U(X, 0) = U(X, H) = U(0, Y) = U(L, Y) = 0, \quad (7)$$

$$V(X, 0) = V(X, H) = V(0, Y) = V(L, Y) = 0, \quad (8)$$

$$T(0, Y) = T_h, T(X, H) = T_c, \frac{\partial T}{\partial X}(L, Y) = \frac{\partial T}{\partial Y}(X, 0) = 0, \quad (9)$$

$$C(0, Y) = C_h, C(X, H) = C_c, \frac{\partial C}{\partial X}(L, Y) = \frac{\partial C}{\partial Y}(X, 0) = 0. \quad (10)$$

Dimensionless variables used in the analysis are defined as,

$$t = \frac{\alpha t'}{LH}, x = \frac{X}{L}, y = \frac{Y}{H}, \delta = \frac{H}{L}, u = \frac{HU}{\alpha}, v = \frac{LV}{\alpha}, \quad (11)$$

$$\theta = \frac{T - T_c}{T_h - T_c}, S = \frac{C - C_c}{C_h - C_c}, p = \frac{[P + \rho_0 g(X \sin \phi + Y \cos \phi)] L^2}{\rho_0 \alpha^2}. \quad (12)$$

where,  $X$  and  $Y$  are the distances measured along the horizontal and vertical directions respectively;  $U$  and  $V$  are velocity components in the  $X$ - and  $Y$ - directions respectively;  $T$  and  $C$  denote the temperature and concentration respectively;  $\nu$ ,  $\alpha$  and  $D$  are kinematic viscosity, thermal diffusivity and mass diffusivity respectively;  $\mu$  is viscosity;  $P$  is the pressure and  $\rho$  is the density;  $T_h$  and  $T_c$  are the temperatures at the hot and cold walls respectively;  $C_h$  and  $C_c$  are the concentrations at the hot and cold walls respectively;  $C_p$  and  $Q$  are the specific heat at constant pressure and the heat generation parameter respectively.

The dimensionless governing equations are as follows:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad (13)$$

$$\begin{aligned} \frac{\partial u}{\partial t} = -\frac{\partial p}{\partial x} + Pr \left( \frac{1}{\delta} \frac{\partial^2 u}{\partial x^2} + \frac{1}{\delta^3} \frac{\partial^2 u}{\partial y^2} \right) - \frac{1}{\delta^2} \left( \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} \right) + Gr_T Pr^2 (\theta + NS) \sin \phi \\ + Ha^2 Pr (v \sin \phi \cos \phi - \frac{1}{\delta} u \sin^2 \phi), \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial y} + Pr \left( \delta \frac{\partial^2 u}{\partial x^2} + \frac{1}{\delta} \frac{\partial^2 u}{\partial y^2} \right) - \left( \frac{\partial v^2}{\partial y} + \frac{\partial uv}{\partial x} \right) + \delta Gr_T Pr^2 (\theta + NS) \cos \phi \\ + Ha^2 Pr (u \cos \phi \sin \phi - \delta v \cos^2 \phi), \end{aligned} \quad (15)$$

$$\frac{\partial \theta}{\partial t} = \left( \delta \frac{\partial^2 \theta}{\partial x^2} + \frac{1}{\delta} \frac{\partial^2 \theta}{\partial y^2} \right) - \left( \frac{\partial u \theta}{\partial x} + \frac{\partial v \theta}{\partial y} \right) + \delta He \theta. \quad (16)$$

$$\frac{\partial S}{\partial t} = \frac{1}{Le} \left( \delta \frac{\partial^2 S}{\partial x^2} + \frac{1}{\delta} \frac{\partial^2 S}{\partial y^2} \right) - \left( \frac{\partial u S}{\partial x} + \frac{\partial v S}{\partial y} \right). \quad (17)$$

Here, dimensionless parameters are

$$Gr_T = g \beta_T (T_h - T_c) L^3 / \nu^2, Gr_C = g \beta_S (C_h - C_c) L^3 / \nu^2, Ha = LB_0 \sqrt{\frac{\sigma^*}{\mu}} \quad (18)$$

$$Pr = \frac{\mu}{\rho_0 \alpha}, N = \frac{Gr_C}{Gr_T}, Ra = Gr_T Pr, He = \frac{QL^2}{k}. \quad (19)$$

The dimensionless boundary conditions are

$$u = 0, \quad v = 0 \quad \text{on} \quad x = 0, 1 \quad \text{and} \quad y = 0, 1. \quad (20)$$

$$\theta = 1 \quad \text{on} \quad x = 0 \quad \text{and} \quad \theta = 0 \quad \text{on} \quad y = 1. \quad (21)$$

$$\left( \frac{\partial \theta}{\partial x} \right) \Big|_{x=1} = 0 \quad \text{and} \quad \left( \frac{\partial \theta}{\partial y} \right) \Big|_{y=0} = 0. \quad (22)$$

$$S = 1 \quad \text{on} \quad x = 0 \quad \text{and} \quad S = 0 \quad \text{on} \quad y = 1. \quad (23)$$

$$\left( \frac{\partial S}{\partial x} \right) \Big|_{x=1} = 0 \quad \text{and} \quad \left( \frac{\partial S}{\partial y} \right) \Big|_{y=0} = 0. \quad (24)$$

Here,  $\delta$ ,  $N$ ,  $Le$ ,  $Pr$ ,  $Gr_C$ ,  $Gr_T$ ,  $\sigma^*$  and  $He$  are the aspect ratio of the enclosure, the buoyancy ratio, Lewis number, Prandtl number, solutal Grashof number, thermal Grashof number, Stefan-Boltzmann constant and the heat generation parameter respectively. We define  $Nu$  and  $\overline{Nu}_H$  by,

$$Nu = -\delta \frac{\partial \theta}{\partial x} \Big|_{x=0} \quad \text{and} \quad \overline{Nu}_H = \delta \int_0^1 Nu \, dy.$$

## 2.1 Solution procedure and numerical stability criteria

A control-volume based finite-difference discretization of the continuity, momentum, temperature and concentration equations was carried out using a staggered grid, popularly known as MAC cell method. In this approach, the velocity components  $u$  and  $v$  are evaluated at different locations in the control volume whereas the pressure, temperature and concentration are evaluated at the same location of the control volume as shown in Fig.2. We note from Fig.2 that the velocity components  $u$ ,  $v$  are stored at the mid point of the vertical and horizontal faces respectively whereas the pressure, temperature and concentration values are stored at the centre of the cells. In the MAC method we use different cells to discretize different equations. For discretizing the continuity, temperature and concentration equations, we used the cell ABCD in Fig. 2.

In the derivation of pressure-Poisson equation, the divergence term at the  $n$ -th time level, ( $D_{i,j}^n$ ) is retained and evaluated in the pressure-Poisson iteration. This is done because the discretized form of the divergence of the velocity field, i.e,  $D_{i,j}^n$  is not guaranteed to be zero initially. The solution procedure begins with the initialization of the velocity field. This is done either from the results of a previous cycle or from the prescribed initial and boundary conditions. Using this velocity field, the pressure-Poisson equation is solved using the Bi-CG-Stab method. Having determined the pressure field, the  $u$ - and the  $v$ - momentum, the temperature and concentration equations are updated to find  $u$ ,  $v$ ,  $\theta$  and  $S$  at the  $(n+1)$ th time level. Then using the values of  $u$  and  $v$  at the  $(n+1)$ th time level, the divergence of the velocity field is obtained and the limit determined. If the absolute value is less than  $0.5 \times 10^{-5}$  and a steady state is reached, the iteration process stops, otherwise the pressure-Poisson equation is solved again for pressure.

The linear stability of the fluid flow gives  $\delta t_1 \leq \text{Min} \left[ \frac{\delta x}{|u|}, \frac{\delta y}{|v|} \right]$ , (Courant, Friedrichs and Lewy condition). Here, temporal discretization is used in the explicit finite difference scheme. From Hirt's stability analysis, we have  $\delta t_2 \leq \text{Min} \left[ \frac{1}{2Pr} \cdot \frac{\delta x^2 \cdot \delta y^2}{(\delta x^2 + \delta y^2)} \right]$ . This condition roughly states that momentum cannot diffuse more than one cell width per time step. The time step is determined from  $\delta t = FCT \times [\text{Min}(\delta t_1, \delta t_2)]$ , where the factor FCT varies from 0.2 to 0.4. The upwinding parameter  $\beta$  is governed by the inequality condition  $1 \geq \beta \geq \text{Max} \left[ \left| \frac{u \delta t}{\delta x} \right|, \left| \frac{v \delta t}{\delta y} \right| \right]$ . As a rule of thumb,  $\beta$  is taken to be approximately 1.2 times larger than the result found from the above inequality condition.



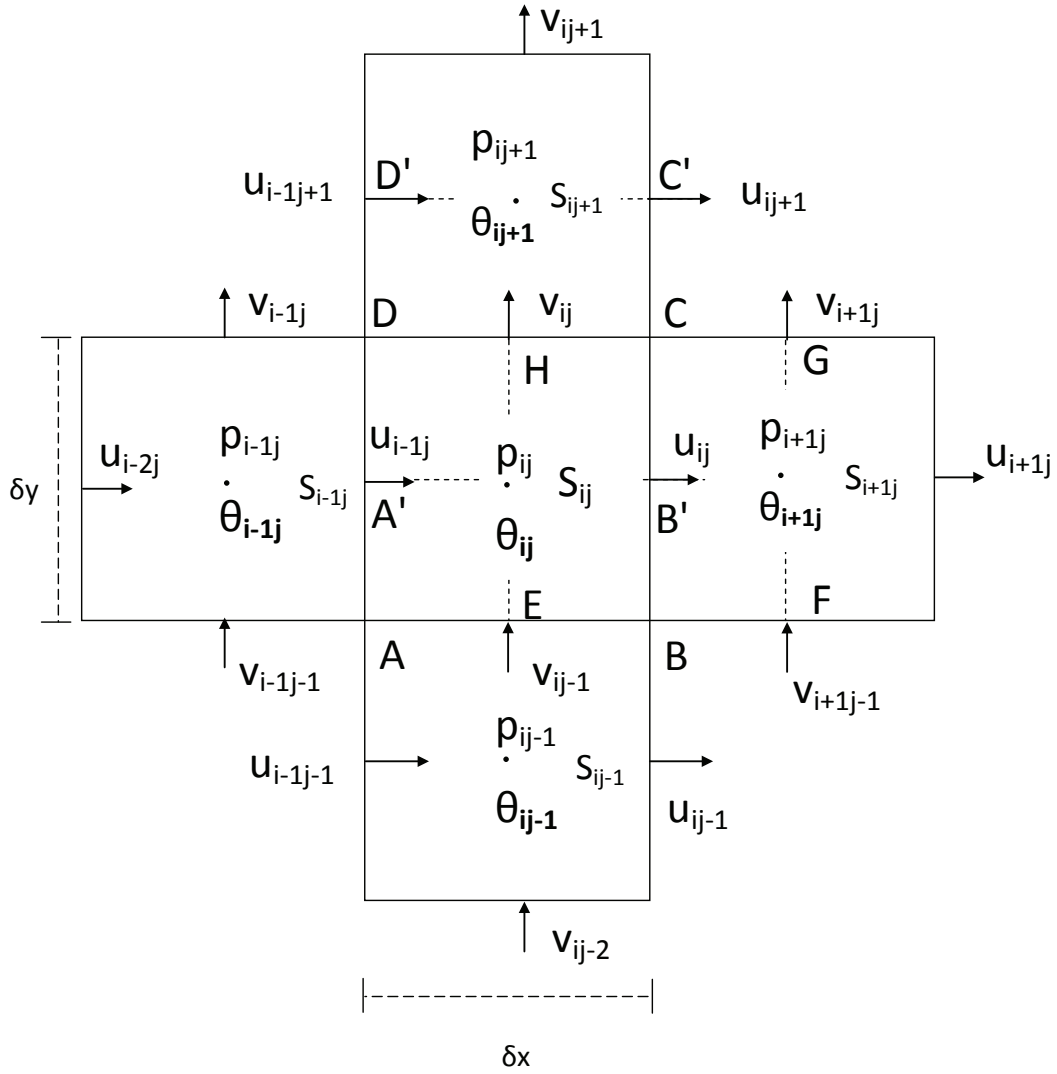


Figure 2: Control volume for  $u$ -momentum,  $v$ -momentum, temperature and concentration equations.

### 3 Results and discussions

The working fluid in this study is chosen to be air with Prandtl number  $Pr = 0.7$ , the heat generation parameter  $He = 1.0$ , Hartmann number  $Ha = 10^2$ , Lewis number  $Le = 1.0$ , Grashof number  $Gr_T =$

$10^4$  and aspect ratio  $\delta = 2$ . The enclosure inclination angle  $\phi$ , magnetic field angle  $\varphi$  in the enclosure and the buoyancy ratio are such that  $45^\circ \leq \phi \leq 135^\circ$ ,  $45^\circ \leq \varphi \leq 135^\circ$  and  $-20 \leq N \leq 20$ , respectively.

Numerical results for the streamline, isotherm and iso-concentration contours inside the inclined rectangular cavity, average Nusselt and Sherwood numbers distribution at the uniformly heated and concentrated surface of the inclined cavity for various values of the angle of magnetic field ( $\varphi$ ), buoyancy ratio ( $N$ ), inclination angle ( $\phi$ ) and dimensionless time parameter  $t$  are presented in Figures 3 - 9 and Tables 1 - 3.

In order to obtain a grid independent solution to the problem, a grid refinement study was performed and the results are presented in Table 1 for  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $N = 1.0$ ,  $He = 1.0$ ,  $\varphi = 45^\circ$ ,  $\phi = 45^\circ$  and  $N = 1$ . It is important to note that as the number of grid points are increased the value of  $|\psi_{max}|$  increases. However, when the number of grid points increases from  $80 \times 80$  to  $160 \times 160$ , no significant change is found in the value of  $|\psi_{max}|$ . Hence, all the results were computed using  $80 \times 80$  grid points.

A comparison of the average Nusselt number at the hot wall is made with Mahapatra *et al.* (2013b); Ece and Büyük (2006) for various values of  $\phi$  and  $\varphi$  in Table 2. It is noted that a very good agreement has been obtained with the previously published results.

Table 3 shows the average Nusselt and Sherwood numbers for different magnetic field angles  $\varphi$ , cavity inclination angles  $\phi$  and buoyancy ratio  $N$  when the other parameters are fixed. It is seen that when  $\varphi$  is fixed and  $N = 20$ ,  $\overline{Nu}_H$  increases with  $\phi$ . However,  $\overline{Sh}_H$  decreases between  $\phi = 45^\circ$  and  $\phi = 90^\circ$ . Beyond this interval, the opposite trend is observed. When  $\varphi$  is fixed and  $N = 1$ , both  $\overline{Nu}_H$  and  $\overline{Sh}_H$  increase with  $\phi$ . Again, when  $\varphi$  is fixed and  $N = -20$ , both  $\overline{Nu}_H$  and  $\overline{Sh}_H$  decrease from  $\phi = 45^\circ$  to  $\phi = 90^\circ$ . For any value of  $N$  both  $\overline{Nu}_H$  and  $\overline{Sh}_H$  increase between  $\phi = 45^\circ$  and  $\phi = 90^\circ$ . If  $\varphi$  and  $\phi$  are fixed then  $\overline{Nu}_H$  and  $\overline{Sh}_H$  decrease as  $N$  varies from  $N = 20$  to  $N = 1$ . However, from  $N = 1$  to  $N = -20$ , the opposite trend are observed.

The numerical results for the streamline, isotherm, iso-concentration contours are presented in Figs. 3 - 9 for uniformly heated and concentrated walls for different values of  $\varphi$ ,  $\phi$ ,  $t$  and  $N$ . The relative importance of thermal and solutal buoyancy forces is denoted by the buoyancy ratio ( $N$ ) and is defined as the ratio of the solutal buoyancy force to thermal buoyancy force. This parameter is varied through a wide range  $-20 < N < 20$ ; covering the concentration-dominated opposing flow ( $N = -20$ ), pure thermal convection dominated flow ( $N = 0$ ), and concentration-dominated aiding flow ( $N = 20$ ). When  $N$  is sufficiently small i.e, the mass buoyancy is greater than the thermal buoyancy, the buoyancy forces that drive the fluid motion are mainly due to the gradients of temperature. Negative values of  $N$  are due to the negative coefficient of concentration expansion.

When  $N = 1$ , the flow is steady; this is because in this case, the two buoyancy forces are equal to and oppose each other. When  $N > 1$ , the flow is driven by buoyancy due to solutal gradients. Clockwise and anticlockwise flows are shown using negative and positive signs of stream functions, respectively.

Fig. 3 shows the effect of  $N$  on the streamlines, isotherms and as well as on the iso-concentrations for a uniformly heated and concentrated wall when  $\varphi = 90^\circ$  and  $\phi = 45^\circ$  for  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$  and  $He = 1.0$ . When  $N = 20$  the streamlines are concentrated near the hot wall.

When  $N = 1$ , due to the cold fluid rising up from the middle portion of the bottom wall and the flow downward along the two vertical walls, we observe two symmetric rolls with clockwise and anticlockwise rotations inside the cavity. However, when  $N = -20$  the streamlines fill the whole cavity forming two eddies at the centre of the cavity in an anticlockwise direction. The value of the stream function increases due to the strong circulation of the fluid. The isotherms are dispersed in the whole cavity when  $N = 20$  but as  $N$  decreases the isotherms become concentrated near the cold wall. The stronger circulation causes the temperature contours to be concentrated near the cold wall which may result in greater heat transfer rate due to convection. The iso-concentration contours are concentrated near the edge of the hot wall and cold wall on one side and the edge of hot wall and adiabatic wall on the other. However, as  $N$  decreases from 20 to  $-20$  the iso-concentration contours on the one side of the cavity are dispersed towards the adiabatic wall due to a higher mass transfer rate.

The streamline, isotherm and iso-concentration contours for different values of  $N$  when  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 90^\circ$  and  $\phi = 90^\circ$  with uniformly heated and concentrated wall are displayed in Fig. 4. As the inclination angle increases the streamlines form a single eddy near the uniformly heated and concentration wall with anticlockwise circulation when  $N = 20$ . The center of the single eddy is slightly shifted away from the heated wall towards the adiabatic wall as  $N$  decreases from 20 to 1. It is interesting to note that for  $N = -20$ , the effect of solutal buoyancy force is in the opposite direction to the thermal buoyancy force. Therefore, the magnitude of the thermal buoyancy force is small compared to the solutal buoyancy force. Consequently, the streamlines are in a clockwise direction. As  $N$  decreases the isotherms are dispersed towards the adiabatic walls from the cold wall by dividing into two parts due to stronger heat transfer rate from the heated wall. However, as  $N$  decreases, the iso-concentrations become more concentrated towards the hot wall.

Fig. 5 shows the effect of  $N$  on the streamlines, isotherms and as well as on the iso-concentrations for a uniformly heated and concentrated wall when  $\varphi = 90^\circ$  and  $\phi = 135^\circ$  for  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$  and  $He = 1.0$ . As the buoyancy ratio decreases, the stream function values become concentrated near the hot wall due convection. When  $N = 20$  the isotherms are concentrated near the cold wall. As buoyancy ratio decreases the isotherms are dispersed in the whole cavity. The iso-concentrations are mainly concentrated towards the hot wall due to increased mass transfer rate at the heated wall.

The streamline, isotherm and iso-concentration contours for different values of  $N$  when  $P_r = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 135^\circ$  and  $\phi = 45^\circ$  with uniformly heated and concentrated walls are displayed in Fig. 6. A comparison of Figs. 3 and 6, shows that the pattern of streamlines, isotherms and iso-concentrations is similar for uniformly heated and concentrated cases except when  $N = -20$ . When  $N = -20$ , the streamlines form a single anticlockwise eddy in the centre of the cavity. Fig. 7 shows the effect of  $N$  on the streamlines, isotherms and as well as on the iso-concentrations for near a uniformly heated and concentrated wall when  $\varphi = 135^\circ$  and  $\phi = 90^\circ$  for  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $N = 1.0$  and  $He = 1.0$ . A comparison of Figs. 4 and 7, again shows that the pattern of streamlines, isotherms and iso-concentrations is almost similar for uniformly heated and concentrated cases.

Fig. 8 shows the effect of  $N$  on the streamlines, isotherms and as well as on the iso-concentrations

when  $\varphi = 135^\circ$  and  $\phi = 135^\circ$  for  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$  and  $He = 1.0$ . As the buoyancy ratio increases, the boundary layer thickness becomes thinner. The change in the flow structure for high buoyancy ratio numbers has a significant influence on the concentration field, which builds up a vertical stratification in the enclosure. The streamlines circulate clockwise in the cavity in although mainly concentrated at edge of hot and cold walls when  $N = -20$ . These streamlines circulate in an anticlockwise direction as  $N$  increases to 1 and are mainly at the edge of the hot and cold walls because the fluid motions mainly due to temperature gradients. When  $N = 20$ , the streamlines are spread in the whole cavity in an anti-clockwise rotation. In, this case the stream function value is very high compared to others in the core of the cavity. Again, when  $N = -20$ , the isotherms spread from the edge of hot and cold wall to the insulated wall due to the opposition of concentration expansion coefficient to the thermal expansion coefficient. However, isotherms become concentrated near the cold wall as the value of  $N$  increases from 1 to 20. Also, when  $N = -20$  the iso-concentration contours are concentrated near the hot wall but these are dispersed from the hot wall as  $N$  increases from  $-20$  to 1 and 1 to 20.

Fig. 9 shows the effect of the non-dimensional time parameter  $t$  on the streamlines, isotherms and iso-concentrations when  $\varphi = 90^\circ$  and  $\phi = 45^\circ$ ,  $P_r = 0.7$ ,  $Gr_T = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$  and  $N = 1.0$ . The streamlines are mainly concentrated near the uniformly heated and concentrated wall with clockwise direction when  $t = 0.279$ . We can see another eddy with anti-clockwise rotation near the corner of the cold and adiabatic wall as  $t$  increases to  $t = 0.558$ . When  $t = 0.837$ , the anti-clockwise rotating eddies become large and suppress the clockwise rotating eddies. In the steady state i.e., when  $t = 3.170$ , the cavity is divided into two parts with clockwise and anticlockwise rotating streamlines forming multiply eddies. Again, when  $t = 0.279$ , the isotherms are spread to the whole cavity from the edge of hot and cold to the adiabatic walls. But, the isotherms are dispersed form that edge to the edge of adiabatic walls forming two different branches of isotherms as  $t$  increases it's value. But, minimum changes can be found for the iso-concentration contours with the change of non-dimensional time parameter  $t$ .

## 4 Conclusion

The main objective of the current investigation was to study the effects of buoyancy ratio on an unsteady double diffusive natural convection in an inclined rectangular enclosure with different angles of magnetic field. As, the buoyancy ratio increases the boundary layer thickness becomes thinner. The change in the flow structure for high buoyancy ratio has a significant influence on the concentration field. Formation of multiple eddies of counter-clockwise rotations greatly influences the fluid flow. When  $N$  is sufficiently small i.e, the mass buoyancy is greater than the thermal buoyancy, the buoyancy forces that drive the fluid motion are mainly due to the gradients of temperature. Negative values of  $N$  represent the opposing nature of two buoyancy forces, due to the negative coefficient of concentration expansion. In this limit, the so-called heat transfer driven flows, the distribution of constituent does not influence the flow pattern and the heat transfer rate. When  $N = 1$ , the flow is steady; this is because in this case, the two buoyancy forces are equal to and oppose each other. When  $N > 1$ , the flows driven by buoyancy due to solutal gradients where the flow are mainly due to gradients of solute

concentration. The change of patterns with respect to time is described here.

**Acknowledgement** Authors are grateful to the University of KwaZulu-Natal, South Africa for financial support.

## References

- Al-Najem, N.M., Khanafer, K.M. and El-Refae, M.M. (1998). Numerical study of laminar natural convection in tilted enclosure with transverse magnetic field. *International Journal of Numerical Methods for Heat and Fluid Flow*, **8**: 651–672.
- Chamkha, A.J. and Al-Naser, H. (2001). Double-diffusive convection in an inclined porous enclosure with opposing temperature and concentration gradients. *International Journal of Thermal Sciences*, **40**: 227–244.
- Ece, M.C. and Büyük, E. (2006). Natural-convection flow under a magnetic field in an inclined rectangular enclosure heated and cooled on adjacent walls. *Fluid Dynamics Research*, **38**: 564–590.
- Ece, M.C. and Büyük, E. (2007). Natural convection flow under a magnetic field in an inclined square enclosure differentially heated on adjacent walls. *Meccanica*, **42**: 435–449.
- Garandet, J.P., Alboussiere, T. and Moreau, R. (1992). Buoyancy driven convection in a rectangular enclosure with a transverse magnetic field. *International Journal of Heat and Mass Transfer*, **35**: 741–748.
- Grosan, T., Revnic, C., Pop, I. and Ingham, D.B. (2009). Magnetic field and internal heat generation effects on the free convection in a rectangular cavity filled with a porous medium. *International Journal of Heat and Mass Transfer*, **52**: 1525–1533.
- Jordan, J.Z. (2007). Network simulation method applied to radiation and viscous dissipation effects on MHD unsteady free convection over vertical porous plate. *Applied Mathematical Modelling*, **31**: 2019–2033.
- Mahapatra, T.R., Pal, D. and Mondal, S. (2011). Natural convection in a lid-driven square cavity filled with Darcy-Forchheimer porous medium in the presence of thermal radiation. *International Journal of Nonlinear Science*, **11**: 366–379.
- Mahapatra, T.R., Pal, D. and Mondal, S. (2012). Influence of thermal radiation on non-Darcian natural convection in a square cavity filled with fluid saturated porous medium of uniform porosity. *Nonlinear Analysis: Modelling and Control*, **17**: 223–237.
- Mahapatra, T.R., Pal, D. and Mondal, S. (2013a). Effects of buoyancy ratio on double-diffusive natural convection in a lid-driven cavity. *International Journal of Heat and Mass Transfer*, **57**: 771–785.
- Mahapatra, T.R., Pal, D. and Mondal, S. (2013b). Mixed convection flow in an inclined enclosure under magnetic field with thermal radiation and heat generation. *International Communications in Heat and Mass Transfer*, **41**: 47–56.

- Mansour, M.A., Chamkha, A.J., Grecos, R.A., Abd El-Aziz, M.M. and Ahmed, S.E. (2010). MHD natural convection in an inclined cavity filled with a fluid saturated porous medium with heat source in the solid phase. *Nonlinear Analysis: Modelling and Control*, **15**: 55–70.
- Mondal, S. and Sibanda, P. (2015a). Effects of buoyancy ratio on unsteady double-diffusive natural convection in a cavity filled with porous medium with non-uniform boundary conditions. *International Journal of Heat and Mass Transfer*, **85**: 401–413.
- Mondal, S. and Sibanda, P. (2015b). Unsteady double diffusive convection in an inclined rectangular lid-driven enclosure with different magnetic field angles and non-uniform boundary conditions. *International Journal of Heat and Mass Transfer*, **90**: 900–910.
- Moufekkik, F., Moussaoui, M.A., Mezrhab, A., Bouzidi, M. and Laraqi, N. (2013). Study of double-diffusive natural convection and radiation in an inclined cavity using lattice Boltzmann method. *International Journal of Thermal Sciences*, **63**: 65–86.
- Rahman, R. and Sharif, M.A. (2003). Numerical study of laminar natural convection in inclined rectangular enclosures of various aspect ratios. *Numerical Heat Transfer Part A*, **44**: 355–373.
- Rudraiah, N., Barron, R.M., Venkatachalappa, M. and Subbaraya, C.K. (1995). Effect of a magnetic field on free convection in a rectangular enclosure. *International Journal of Engineering Science*, **33**: 1075–1084.
- Sarris, I.E., Kakarantzas, S.C., Grecos, A.P. and Vlachos, N.S. (2005). MHD natural convection in a laterally and volumetrically heated square cavity. *International Journal of Heat and Mass Transfer*, **48**: 3443–3453.
- Shercliff, J. A. (1965). *A Textbook on Magnetohydrodynamics*, Pergamon Press, Oxford.
- Teamah, M.A., Elsafty, A.F., Elfeky, M.Z. and El-Gazzar, E.Z. (2011). Numerical simulation of double-diffusive natural convective flow in an inclined rectangular enclosure in the presence of magnetic field and heat source, part A: Effect of Rayleigh number and inclination angle. *Alexandria Engineering Journal*, **50**: 269–282.
- Teamah, M.A., Sorour, M.M., El-Maghlany, W.M. and Afifi, A. (2013). Numerical simulation of double diffusive laminar mixed convection in shallow inclined cavities with moving lid. *Alexandria Engineering Journal*, **52**: 227–239.
- Wang, G., Wang, Q., Zeng, M. and Qzoe, H. (2008). Numerical study of natural convection heat transfer in an inclined porous cavity with time-periodic boundary conditions *Transport in Porous Media*, **74**: 393–409.

Table 1: Grid independence test when  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 45^\circ$ ,  $\phi = 45^\circ$  and  $N = 1$ .

Grid points	Iter	$ \psi_{\max} $
20 × 20	34368	$5.95110 \times 10^{-2}$
40 × 40	131880	$6.00418 \times 10^{-2}$
80 × 80	530782	$6.01884 \times 10^{-2}$
160 × 160	1162672	$6.01887 \times 10^{-2}$

Table 2: Comparison of average Nusselt number  $\overline{Nu}_H|_{x=0}$  in absence of concentration equation, thermal radiation parameter and heat generation parameter with same boundary conditions and same values of all other parameters.

$\phi$	$\overline{Nu}_H$								
	$\varphi = 0^\circ$			$\varphi = 45^\circ$			$\varphi = 90^\circ$		
	Mahapatra <i>et al.</i> (2013b)	Ece and Büyüük (2006)	Present Results	Mahapatra <i>et al.</i> (2013b)	Ece and Büyüük (2006)	Present Results	Mahapatra <i>et al.</i> (2013b)	Ece and Büyüük (2006)	Present Results
0°	3.5350	3.6831	3.5350	3.5354	3.6819	3.5354	3.5356	3.6813	3.5356
-45°	3.5363	3.6806	3.5363	3.5366	3.6846	3.5366	3.5363	3.6843	3.5363
45°	3.5340	3.6780	3.5340	3.5339	3.6820	3.5339	3.5341	3.6806	3.5341

Table 3: Computed values of  $\overline{Nu}_H$  and  $\overline{Sh}_H$  when  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$  and  $He = 1.0$  for various values of  $\varphi$ ,  $\phi$  and  $N$ .

$\varphi$	$\phi$	Buoyancy ratio N = 20		Buoyancy ratio N = 1		Buoyancy ratio N = -20	
		$\overline{Nu}_H$	$\overline{Sh}_H$	$\overline{Nu}_H$	$\overline{Sh}_H$	$\overline{Nu}_H$	$\overline{Sh}_H$
45°	45°	1.94699	8.93750	1.67498	7.50946	1.88540	7.85743
	90°	2.11064	7.93526	1.68330	7.5132655	1.76858	7.57010
	135°	2.42024	8.52446	1.68626	7.51558	2.44834	8.40416
90°	45°	2.65610	10.73437	1.67736	7.51126	2.19201	8.56903
	90°	2.66738	9.10950	1.70193	7.52637	1.88740	7.73783
	135°	3.00922	11.00129	1.71378	7.54384	3.00949	10.23665
135°	45°	1.931679	8.37004	1.67510	7.50954	1.86995	7.96886
	90°	2.12979	7.80844	1.68421	7.51211	1.84856	7.613725
	135°	2.46780	8.76437	1.68908	7.51589	2.42373	8.37702

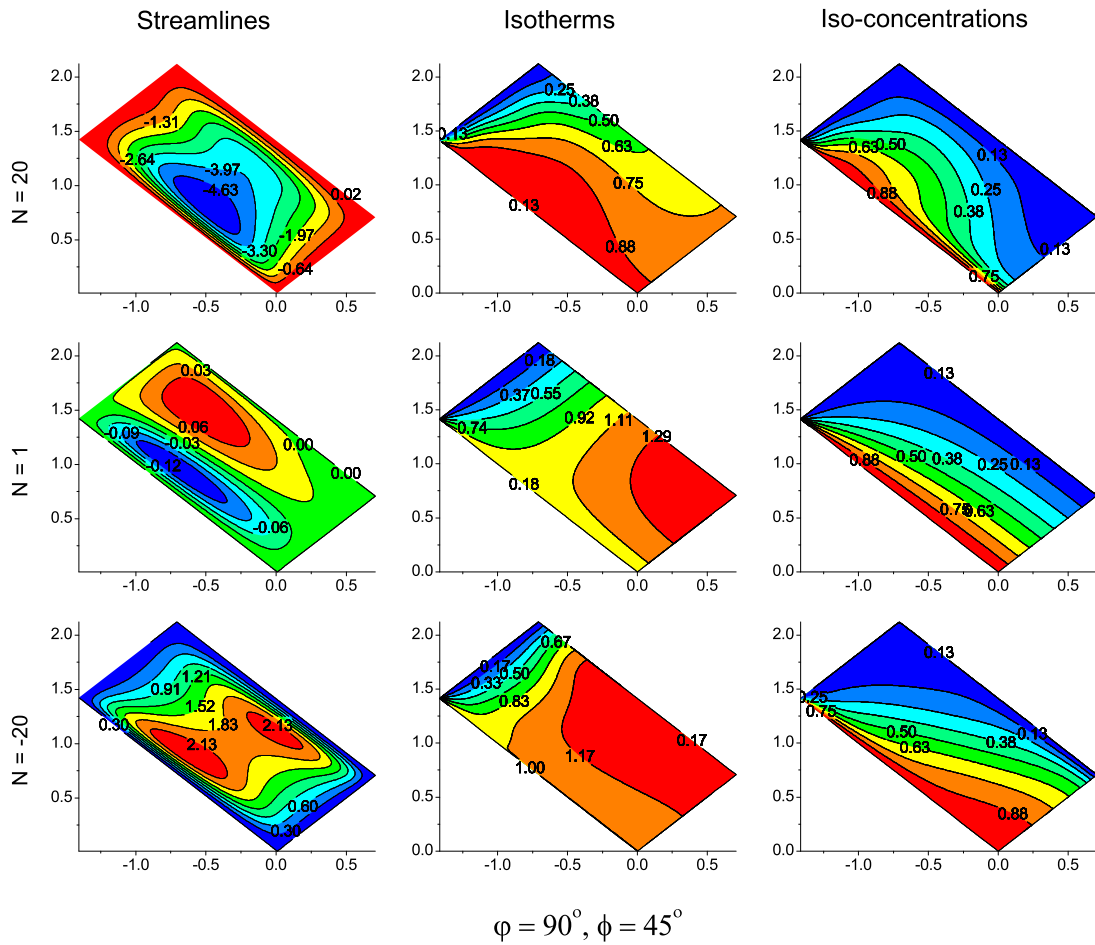


Figure 3: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 90^\circ$ ,  $\phi = 45^\circ$  for different values of  $N$ .



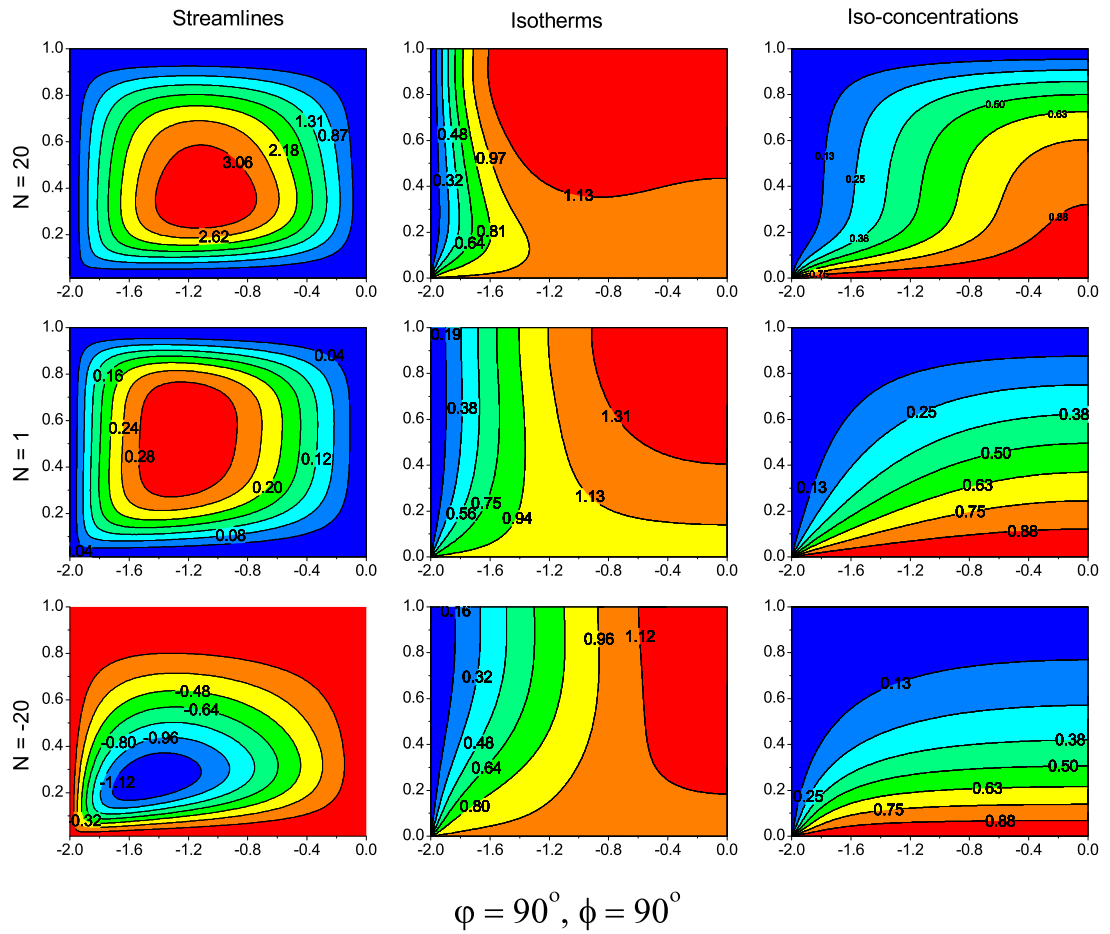


Figure 4: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 90^\circ$ ,  $\phi = 90^\circ$  for different values of  $N$ .

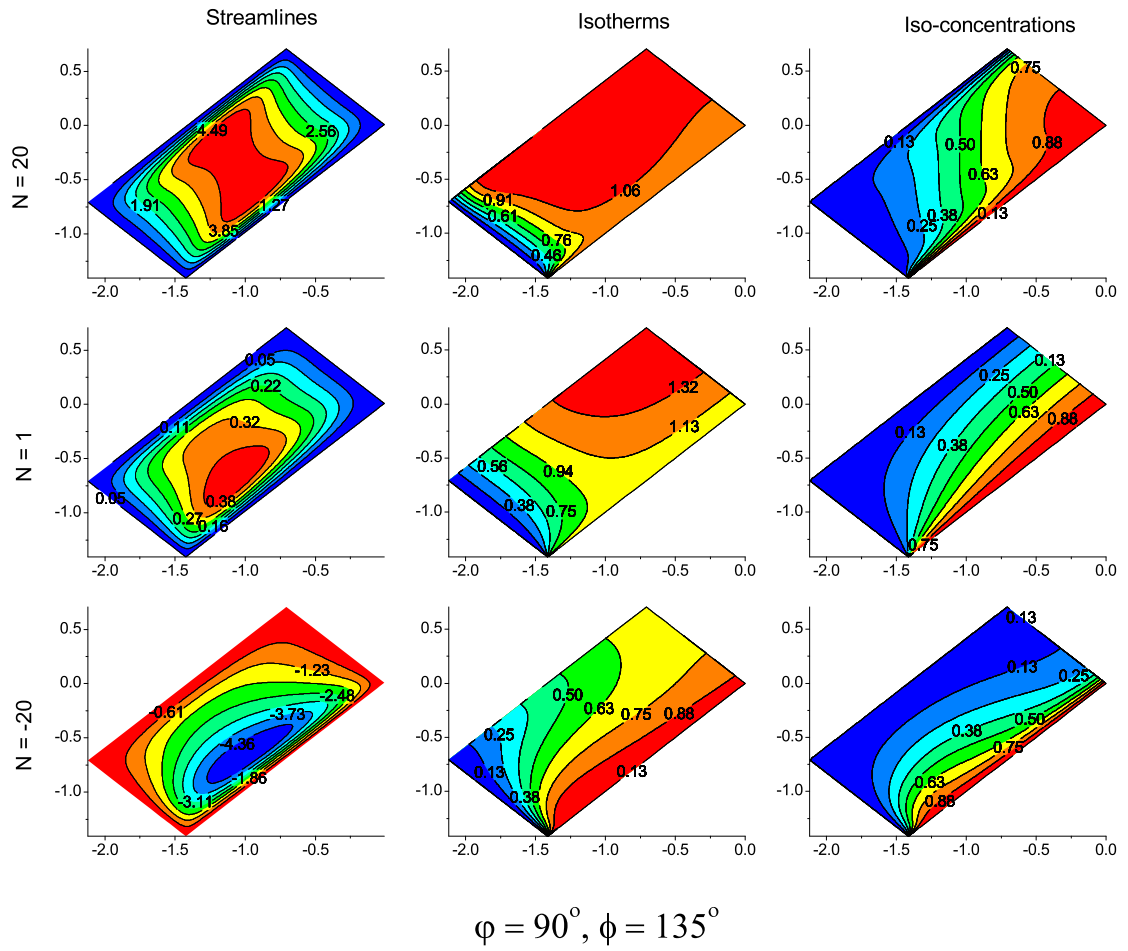


Figure 5: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 90^\circ$ ,  $\phi = 135^\circ$  for different values of  $N$ .

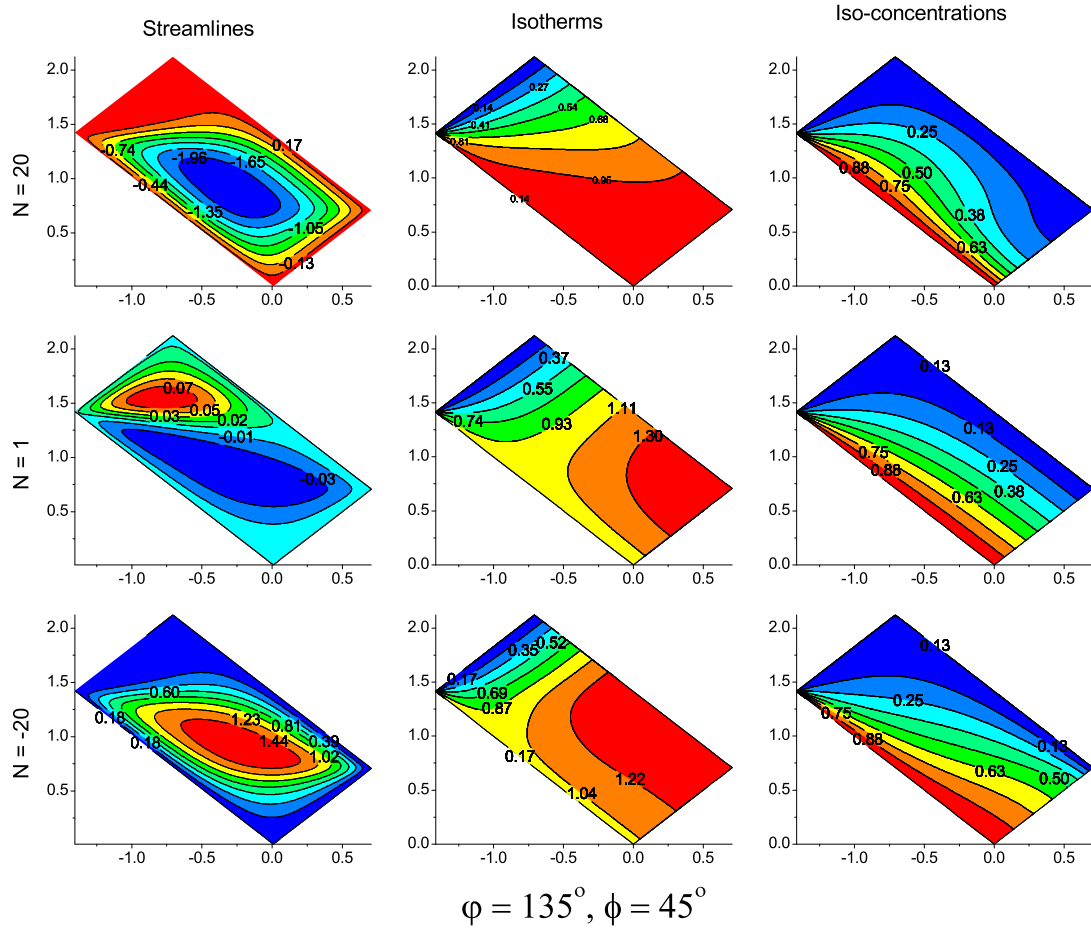


Figure 6: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 135^\circ$ ,  $\phi = 45^\circ$  for different values of  $N$ .

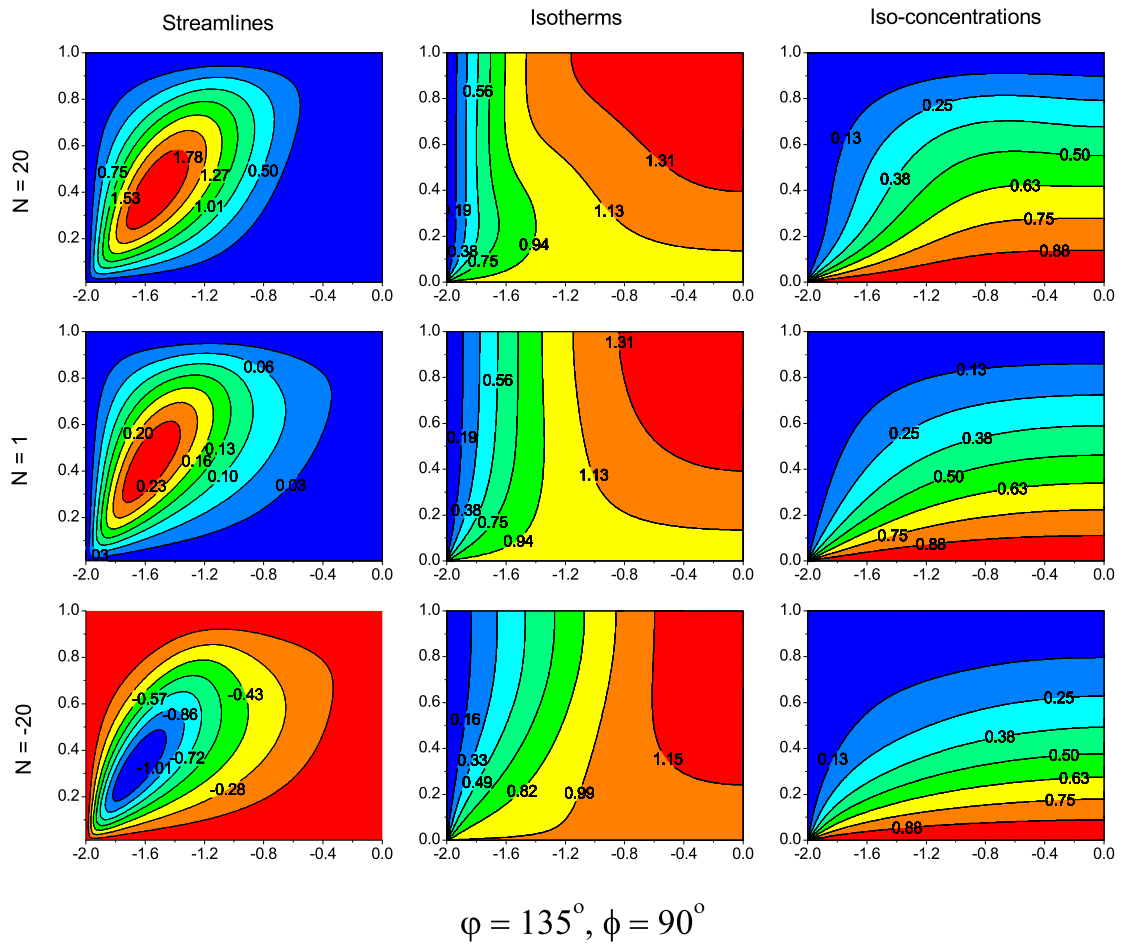


Figure 7: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 135^\circ$ ,  $\phi = 90^\circ$  for different values of  $N$ .

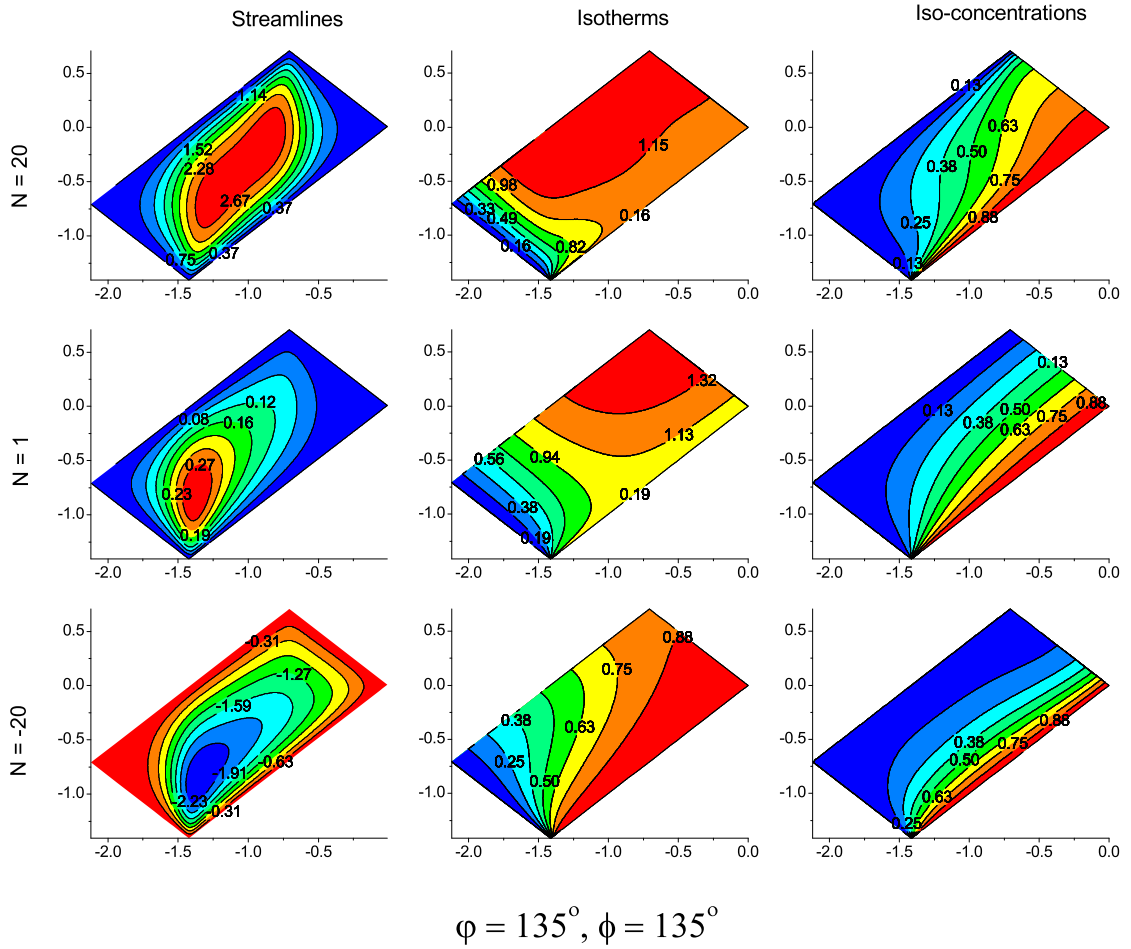


Figure 8: Contour plots for  $P_r = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 135^\circ$ ,  $\phi = 135^\circ$  for different values of  $N$ .

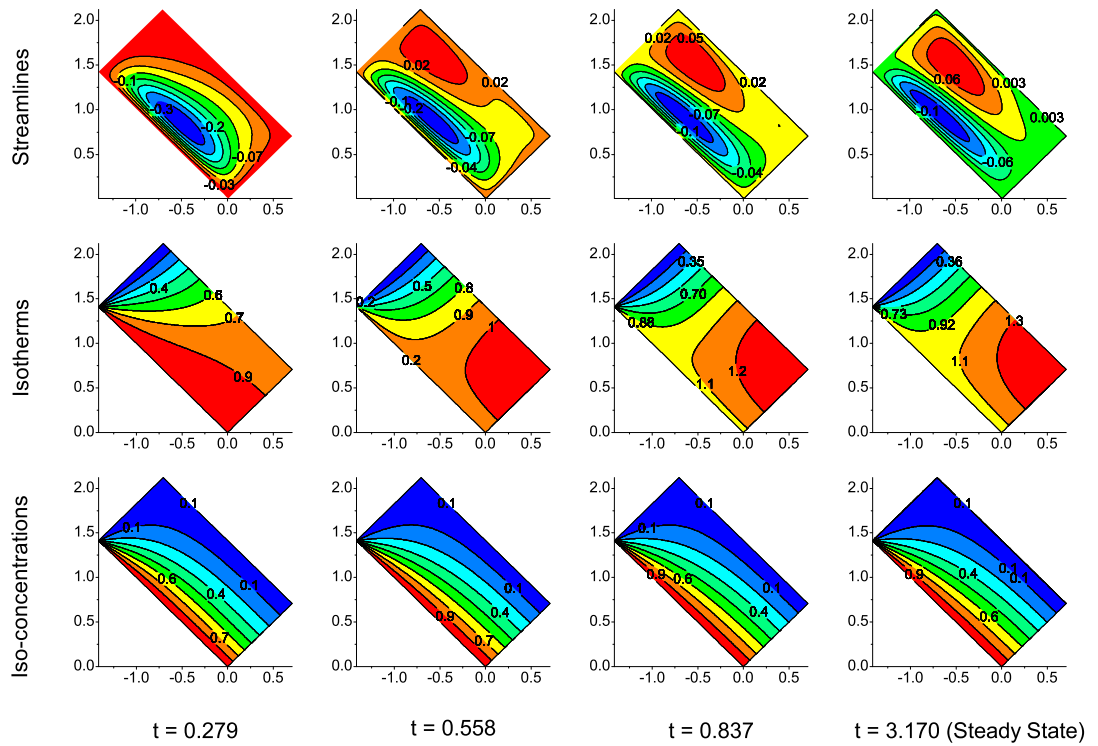


Figure 9: Contour plots for  $Pr = 0.7$ ,  $Gr = 10^4$ ,  $Ha = 10^2$ ,  $\delta = 2$ ,  $Le = 1.0$ ,  $He = 1.0$ ,  $\varphi = 90^\circ$ ,  $\phi = 45^\circ$  and  $N = 1.0$  for different values of dimensionless time parameter ( $t$ ).

## **Establishing effective criteria to link atomic and macro-scale simulations of dislocation nucleation in FCC metals**

†N. J. Burberry<sup>1</sup>, R. Das<sup>1</sup>, W.G. Ferguson<sup>2</sup>, G. Po<sup>3</sup>, N. Ghoniem<sup>3</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Auckland, Auckland, New Zealand.

<sup>2</sup>Department of Chemical and Materials Engineering, University of Auckland, Auckland, New Zealand.

<sup>3</sup>Mechanical and Aerospace Engineering Department, Department of Engineering, University of California, Los Angeles 1597, United States of America

†Presenting and corresponding author: nbur049@aucklanduni.ac.nz

### **Abstract**

Processes at fundamental length scales contribute collectively, in a statistical manner, to the macro-scale effects observed at length scales several orders of magnitude higher. To derive useful quantities pertaining to real material properties from atomic scale simulations, it is critical to evaluate the cumulative effect of multiple atomic-scale defects at the ‘meso’- and ‘micro’- scales. This study aims to develop a phenomenological model for atomic scale effects, which is a critical step towards the development of a comprehensive meso-scale simulation framework. In moderate loading conditions, dislocations in FCC metals are dictated by thermally activated processes that become energetically favourable as the stress approaches a threshold value. The nudged elastic band technique is ideal for evaluating the energetic activation parameters from atomic simulations, in order to evaluate the stress, temperature and rate dependence of a process. On this basis, a constitutive mathematical model is developed for simulations at the meso-scale with respect to the atomic activation parameters, to evaluate the critical (local) shear stress threshold. Once models are established for multiple effects, such as dislocation junction formation, cross-slip, and nucleation, the threshold temperature and stress for a transition between different effects can be evaluated. For example, the threshold temperature can be evaluated during heating, beyond which an immobilised dislocation in a junction will be activated for cross-slip and will shift into an adjacent mobile slip system. This is useful to predict the rate-limiting dislocation process at each simulation timestep, by evaluating the simulation condition-dependent criteria. Additional criteria variables for the constitutive models include properties of the dislocation, the grain boundary and the material’s chemical and elastic properties. Multi-scale modelling from a lower-scale basis is inherently limited by a reduction in the degrees of freedom required to enable large scale simulations, constrained by computational limits. To address this, we intend to use hierarchical multi-scale linking by iteratively updating the constitutive model parameters until the meso-scale method is capable of reproducing atomic scale dislocation effects. The resultant meso-scale method will be useful to study multi-dislocation interactions, which are capable of driving high-stress effects such as dislocation nucleation under low applied stresses, due to stress-concentration in dislocation pile-ups at interfaces. This study contributes to the development of a ‘fundamental basis’ to inform macro-scale models that can provide significant insights about the effect of dislocation microstructure evolution during plastic deformation.

Keywords: Multi-scale computer simulations, dislocation dynamics, FCC metals, polycrystalline plasticity, activation parameters

## 1.0 Introduction

Dislocations act as a “weak point”, defining the elastic limit and subsequent ductility of crystalline materials, such as FCC metals [Zbib and Khraishi (2005); Po, Mohamed et al. (2014)]. Due to their high mobility at low stress and temperature, once dislocations are activated they act as the primary crystal defects for mediating plastic deformation. Dislocations are atomic bond defects with a core region that is fundamentally defined by the sub-nano scale burgers vector. However, dislocations contribute to mechanical properties via multi-dislocation interactions in a cumulative statistical manner up to  $\sim 100 \mu\text{m}$  [Po, Mohamed et al. (2014)]. Hence, to fully understand the characteristics of dislocations for modelling and prediction, it is critical to utilise a multi-scale approach to evaluate the inter-atomic mechanisms and the inter-dislocation effects [Ghoniem†, Busso et al. (2003)]. Molecular dynamics (or MD) simulations are an ideal tool for developing a conceptual and constitutive modelling framework for the atomic processes which are fundamental to the macroscopic properties seen in real materials. MD is inherently limited to very small size and time simulations by computational requirements, so constitutive models need to be applied in micro-meso scale methods such as dislocation dynamics (DD) to evaluate cumulative effects [Po, Lazar et al. (2014)].

Dislocation-mediated deformation is strongly influenced by thermally activated nucleation, especially when confined within a nano-crystalline material with a grain size less than  $\sim 0.1 \mu\text{m}$  [Zhu, Li et al. (2008)]. Recent developments in computational methods have been developed to simulate the minimum energy path (MEP) for an atomic transition, which indicates the activation energy ( $E_a$ ). The nudged elastic band method (NEB) is a popular method which obtains the MEP by minimizing the potential energy of transition states interpolated between input initial and end states [Henkelman, Uberuaga et al. (2000)]. The stress dependence of the activation energy ( $E_a$ ) is known as the activation volume ( $\Omega$ ).  $\Omega$  can be evaluated by calculating the MEP between identical initial and final atomic configurations, but with various externally applied loads [McPhie, Berbenni et al. (2012)]. The fundamental activation properties ( $E_a$  and  $\Omega$ ) are time-stress and temperature independent [Voter, Montalenti et al. (2002)]. Assuming the atomic mechanism is independent of simulation scale, the activation parameters are hence ideal for linking multi-scale simulation methods. On this basis, the thermal activation parameters can be used to establish a constitutive model for predicting the threshold for nucleation, as a function of two known parameters (from stress, temperature and/or strain rate) [Zhu, Li et al. (2008)]. This study establishes a constitutive model using transition state theory to provide a critical contribution for atomic-informed meso-scale computer simulations.

## 2.0 Methods

Single crystals of pure FCC aluminium were simulated using molecular static simulations, with an embedded atom method (EAM) inter-atomic potential provided by [Mishin, Farkas et al. (1999)]. This EAM potential was chosen because it efficiently simulates the stacking fault energies and elastic properties of pure FCC Al. Single crystals are initialised without defects and with periodic boundaries in all dimensions, to represent an ‘infinite single crystal’. The length in the ‘y-axis’ was defined to be greater than 30nm, and the ‘x-’ and ‘y-’ axes were 10 lattice units each, in agreement with prior simulations of dislocation nucleation [Tschopp and McDowell (2007); Sangid, Ezaz et al. (2011); Tucker and McDowell (2011)]. Loading



was applied along the  $[1\bar{1}0]$  close-packed slip direction via constant applied uniaxial strain in the y-direction, so the single crystal was oriented as shown in Figure 1. A fully-dense, minimum energy state was obtained using the conjugate gradient minimization method [Štich, Car et al. (1989)]. The crystal was temperature-rescaled to 50K and temperature-pressure equilibrated by Nose-Hoover thermo-barostat and 0 Bar pressure [Nosé (1984; Hoover (1986)]. The crystal orientations and exact dimensions can be most easily understood schematically, by referring to Figure 1. Note that atomistic “imaging” of simulation states is performed using AtomEye visualisation tool [Li (2003)], and atoms are coloured by either: (a) potential energy or (b) centro-symmetry bond parameter ( $P_{csym}$ ) [Kelchner, Plimpton et al. (1998)].

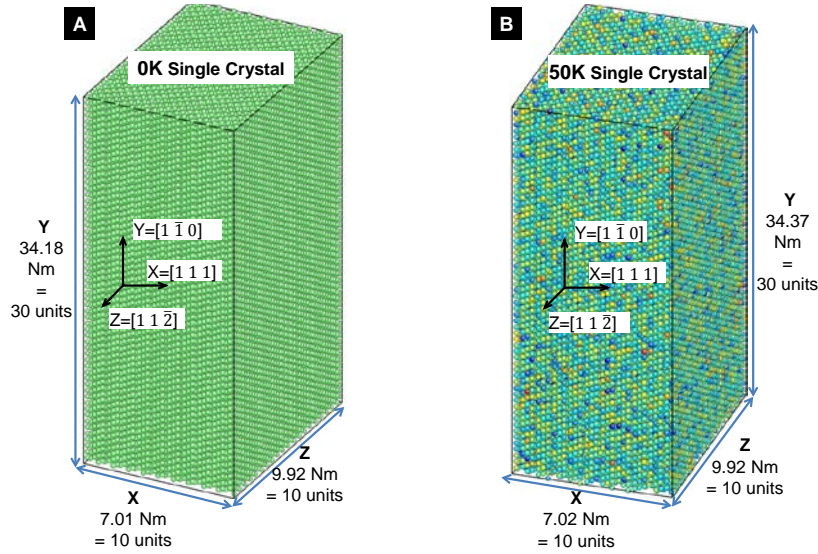


Figure 1: Dimensions and crystal orientations of pure FCC aluminium: A) at 0K; B) at 50K (Note: atoms coloured according to potential energy)

A dislocation was generated in the simulation volume at 50K, by applying a constant rate tensile strain in the  $[1\bar{1}0]$  direction at a rate of  $5.0e^8 s^{-1}$ , with outputs provided every 0.1ps for restarting and visualisation. By selectively filtering atoms according to the  $P_{csym}$ , the first interval was identified containing atoms with high  $P_{csym}$ , to indicate the initiation of crystal slip. From the associated restart file at this point, an instantaneous reverse strain was applied by rescaling the simulation y-dimension in a fashion, followed by immediate rapid quenching to 1K. The atomic velocity and kinetic energy was then reduced to 0 (i.e., 0K) and the simulation energy was minimized by the conjugate-gradient method [Štich, Car et al. (1989)], with an energy tolerance of  $1.0e^{-6}$  eV. This energy tolerance matches that used by a similar prior study, which retains a “metastable” dislocation loop at 0K [McPhie, Berbenni et al. (2012)]. By iteratively testing multiple the reverse strain magnitudes, the exact final simulation stress-strain state was identified which would establish a stable dislocation loop after minimization at 0K. This elaborate process of heating, stressing and quenching was necessary to form a single stable “loop”, as multiple dislocations will simultaneously nucleate above the yield stress at 0K. This method was used to establish the end-state atomic configuration for NEB simulations. The initial state was created by following an identical heat – stress - quench minimization protocol, however the initial strain is much lower. Specifically, the initial state was chosen so that it relaxes to 0 GPa, after the reverse strain is applied.

NEB simulations were performed using the “replica” library available within the LAMMPS code [Plimpton (1995)] to an energy tolerance of  $1e^{-6}$  eV (to match the

previously utilised tolerance). To obtain a useful result, it is critical that the simulation volumes of the initial and final states are identical. To achieve this, the initial state dimensions were re-scaled to match the final state dimensions. To study the stress-dependence of the energy threshold required for dislocation nucleation, it was necessary to repeat NEB simulations at various stresses. This was achieved by utilising an additional stage of incrementally increasing strain in the y-dimension immediately following the first stage of quench-minimization. An additional stage of minimization was also applied. Finally, NEB simulations were performed between the identically dimensioned initial defect-free state and the end state containing a dislocation loop.  $E_a$  for dislocation nucleation was evaluated for each stress state. To evaluate whether there was any relationship between the  $E_a$ , activation volume, temperature, yield strength and strain rate, dynamic simulations were performed at various temperatures from 0K – 1200K, and at strain rates between  $5.0e7 - 5.0e9$  s<sup>-1</sup>. Hence, the temperature and rate dependent stress-strain curves were obtained. For constitutive modelling, data analysis, regression modelling and mathematical validation were performed with Microsoft Excel.

### 3.0 Results and discussion

#### 3.1 Key thermal and mechanical properties for predictive modelling

It is critical to evaluate the threshold disorder temperature ( $T_{disorder}$ ) which is accurate for the inter-atomic potential used to represent the material properties in atomic simulations.  $T_{disorder}$  can be described as the threshold beyond which the yield stress for dislocation nucleation deviates pronouncedly from a linear relationship. The characteristic  $T_{disorder}$  properties of single crystal aluminium, using EAM potential from [Mishin, Farkas et al. (1999)] with the crystal orientation shown in Figure 1 is derived from a plot of yield stress as a function of temperature in Figure 2.

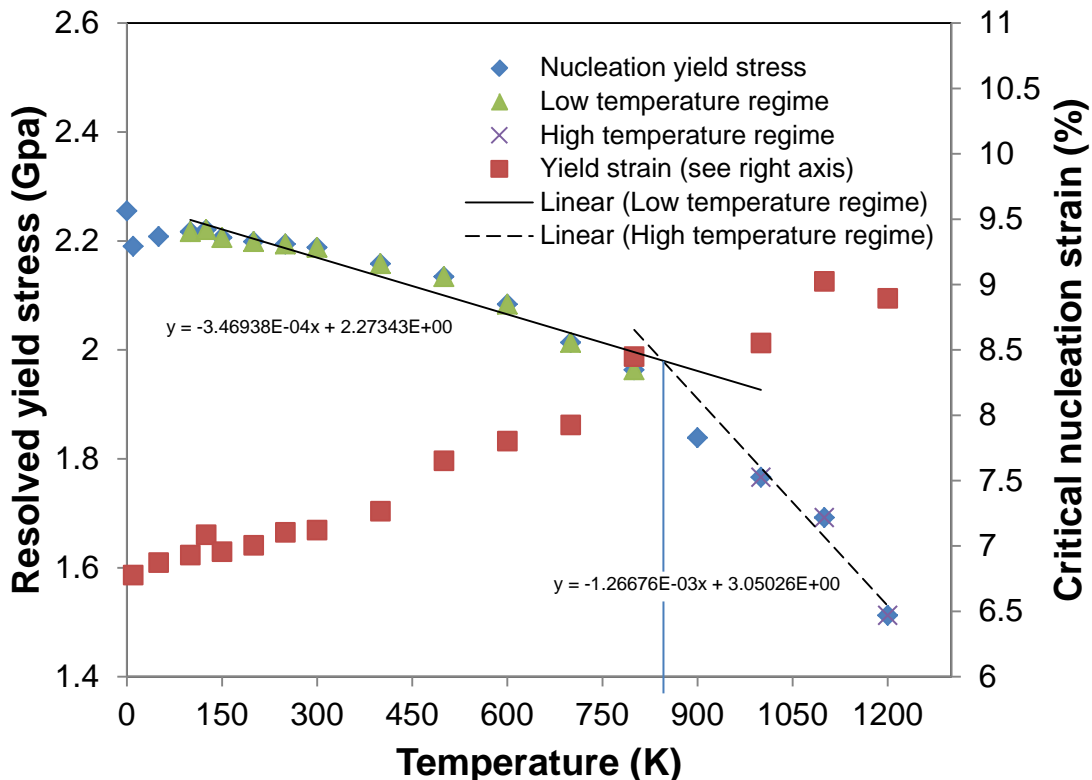


Figure 2: Temperature-dependent strength of single crystal Al in uniaxial tension

Figure 2 demonstrates that the Al atomic potential [Mishin, Farkas et al. (1999)] exhibits approximately linear temperature-strength relationships under strain rates accessible within atomic simulations. Note that the square data points representing yield strain correspond with the y-axis on the right of the figure. The results also provide a valuable analogue for the energy-based constitutive models, derived in later sections from NEB simulations.

Previous detailed analysis of the thermal decomposition temperature, identified that pure FCC aluminium will begin to destabilize at approximately 867.7 K [Nguyen, Ho et al. (1991)]. Referring to Figure 2, there are clearly two linear stress-temperature regimes. The intersection point between the extrapolated curves obtained by simple linear regression indicates the transition temperature, which lies directly between 800K and 900K. Analytical evaluation of the temperature at which the regression curves intersect, results in an exact decomposition temperature of 844.54K, which is in very close agreement with the results obtained in [Nguyen, Ho et al. (1991)]. The critical strain also appears to deviate from a linear trend above 800, with a significant reduction at 900K. Beyond this temperature, the critical strain increases sharply despite a reduction in the yield stress, confirming that the material has distinctly altered elastic properties above 800K.

### 3.2 Characteristics of homogeneous dislocation nucleation

This section will demonstrate the transitional atomic mechanism for dislocation nucleation from a homogeneous, defect free single crystal at 0K and will evaluate the critical athermal activation parameters (i.e., 0 GPa stress and 0K). These parameters are key constant parameters, which form the fundamental basis of constitutive modelling based on thermal activation parameters [Zhu, Li et al. (2008)]. This is most clearly shown by visualising the generation of defect atoms in transitional atomic states, which are identified by the centro-symmetry bond coefficient [Kelchner, Plimpton et al. (1998)]. The transition states and energy barrier are seen in Figure 3.

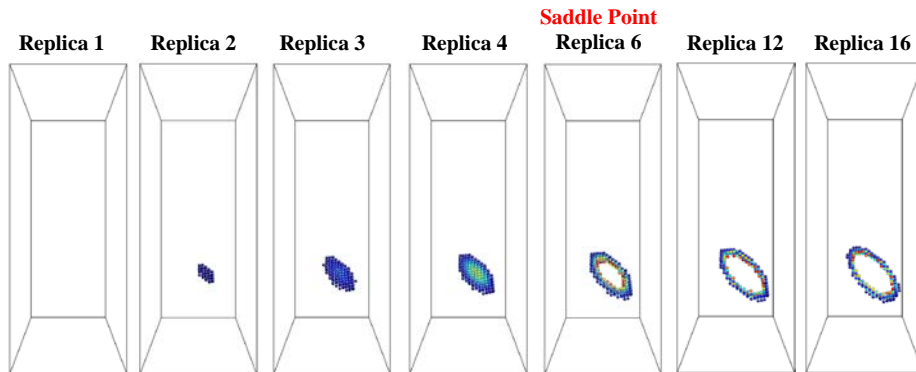


Figure 3a.: Generation of defect atoms during dislocation nucleation transition

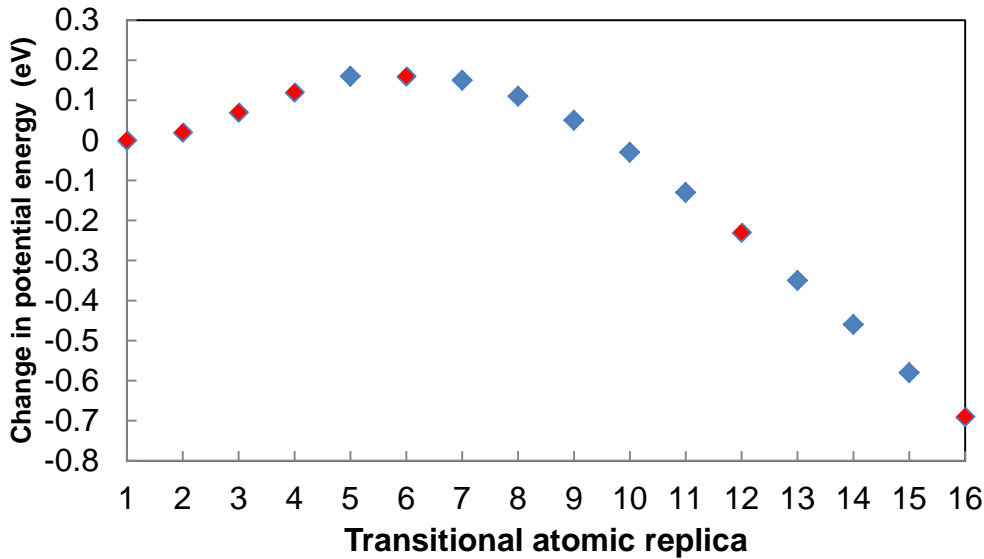


Figure 3b.: Minimum energy barrier for near-athermal dislocation nucleation

Figure 3: transition for dislocation nucleation from Nudged elastic band simulation at 0K and at 2.22 GPa (near athermal yield stress)

Figure 3 shows the results of an NEB simulation between the initial defect free state and final state containing a single full dislocation loop at approximately 2.22 GPa. The process begins with the generation of a Shockley Partial dislocation with a Burgers vector of  $[\bar{1}1\bar{2}]$ , and magnitude of  $1.672\text{\AA}$ . Close inspection shows that nucleation is initiated as fundamental atomic-scale vibration in 2 or 3 atoms, resulting in the minor bond disruption of 9 atoms in 2 adjacent  $(\bar{1}1\bar{1})$  planes. There is a substantial reduction in potential energy, which likely corresponds with the relaxation of the elastic strain energy in all non-defect atoms in the volume. Note that this also corresponds with a very minor reduction in shear stress of  $\sim 0.06\text{GPa}$ , however this change is considered negligible. The forward energy barrier approaches zero as the stress increases above 2.2 GPa, which is why nucleation can proceed without thermal input above the athermal stress-strain limit.

It is interesting to note that the dislocation loop is not at a maximum size at the saddle point. However, this is probably explained because the saddle point involves the energy jump required to nucleate the trailing partial dislocation. For this reason, the potential energy of the defected atoms is at the highest in this replica. This also indicates that the elastic strain and stacking fault energies decrease beyond this point.

### 3.3 $E_a$ for homogeneous nucleation as a function of stress

The  $E_a$  for dislocation nucleation was calculated for single crystals with a residual shear stress of between -0.04 GPa to 2.19 GPa. The stress-dependence is most clearly shown graphically according to Figure 6.

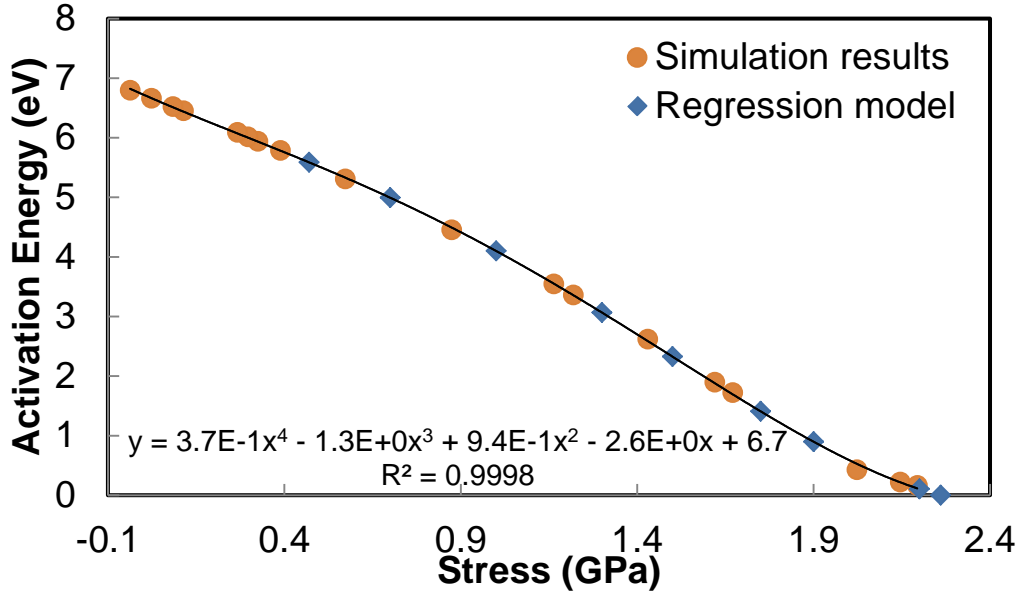


Figure 4:  $E_a$  for dislocation nucleation as a function of resolved shear stress

Figure 6 shows a nearly linear relationship exists between stress and  $E_a$  until approximately 1.4GPa. Beyond this point, the relationship is non-linear and approaches a more ‘exponential shape’. The critical stress with an  $E_a$  of 0.0eV was extrapolated from the data using a 4<sup>th</sup> order polynomial regression curve, fitted to the simulation data with an  $R^2$  value of 0.9998. Note that only 1 SF is shown in figure due to space limitations. Accordingly, the critical athermal resolved shear stress for spontaneous dislocation nucleation is exactly 2.26 GPa. This is another critical parameter that will be used for the constitutive modelling in the next section.

These results are also ideal for an explicit evaluation of the activation volume  $\Omega$ , according to the standard thermodynamic relationship shown in Equation 1:

$$\Omega = \left( \frac{\partial \Delta E_A}{\partial \sigma} \right)_{T,P} \quad (1)$$

According to Equation 1, the activation volume can be very simply evaluated from the stress dependent  $E_a$  at constant temperature and hydrostatic pressure. Temperature is, by definition of the NEB procedure, exactly 0K. The hydrostatic pressure is 0 Bar, due to the algorithms used to define the uniaxial loading and with damping to reduce fluctuations from the Hoover barostat [Hoover (1986)]. If evaluated from total data range, the  $\Omega$  is  $4.9e^{-28}$  J/Pa. For stresses greater than 2.1 GPa, the  $\Omega$  is  $2.5e^{-28}$  J/Pa and for stresses less than 2.1 GPa and greater than 0.8 GPa, the  $\Omega$  is  $5.66e^{-28}$  J/Pa. The activation volume is a critical fundamental parameter that can has been correlated with the strain rate dependence of nucleation with simulations [Deng and Sansoz (2010)] and experiments [Asaro and Suresh (2005)].

### 3.4 Constitutive model for temperature – strain rate dependent yield stress

The activation energy,  $E_a$ , is typically used to directly evaluate the thermal effect on mechanical properties, and is directly correlated with the temperature according to Equation 2:

$$E_{a,T} = \left(1 - \frac{T}{T_{disorder}}\right) E_{a,0K} \quad [\text{Zhu and Li (2010)}] \quad (2)$$

In combination with the activation volume,  $\Omega$ , the stress can be evaluated according the simplified, modified version of the relationship derived by Zhu et al. [Zhu and Li (2010)]:

$$\sigma = \left( \sigma_{athermal} - \ln \frac{k_b T N \nu_0}{E \dot{\epsilon} \Omega} \right) \quad [\text{Zhu and Li (2010)}] \quad (3)$$

where  $\sigma_{athermal}$  is the critical shear stress at 0K,  $k_b$  is the Boltzmann constant,  $N \nu_0$  is the number of transitions attempted per second,  $E$  is the Young's modulus and  $\dot{\epsilon}$  is the strain rate. Unfortunately, although this method is based on fundamental theoretical physics, it is dependent on an accurate evaluation of the exact context – dependent value of  $\Omega$  and is typically flawed by a biased “prediction” of  $N \nu_0$  [Zhu, Li et al. (2009)]. As the strain rate decreases and the temperature decreases, the stress is influenced in a multiplicative manner. Hence, the cumulative effects should be significant when comparing 0K atomic simulations at  $5.0 \times 10^8 \text{ s}^{-1}$  with experiments which are typically at strain rate  $< 1 \text{ s}^{-1}$  and at 298K. Thermally activated dislocation nucleation exhibits a physics-based correlation the yield stress and the strain rate according to:

$$\Omega = \frac{k_b T}{m \sigma} \quad [\text{Asaro and Suresh (2005)}] \quad (4)$$

where  $m$  is the strain-rate sensitivity exponent. If  $\Omega$  is assumed stress-independent (i.e., Figure 6 is linear), it is possible to evaluate the temperature dependence by rearranging Equation 4 and solving for  $m$  at  $\dot{\epsilon}_{athermal}$  and  $\sigma_{athermal}$ . The critical resolved shear stress at a given  $\dot{\epsilon}$  and  $T$  is then predicted with the well-established formula:

$$\frac{\sigma}{\sigma_{Athermal}} = \left( \frac{\dot{\epsilon}}{\dot{\epsilon}_{Athermal}} \right)^m \quad (5)$$

Using Equation 5, the yield stress is predicted as a function of temperature at a variety of strain rates between  $5.0 \times 10^8 \text{ s}^{-1}$  -  $500 \text{ s}^{-1}$ . The results are then compared with dynamic simulation results at strain rates of  $5.0 \times 10^8$ ,  $1.0 \times 10^8$ , and  $1.0 \times 10^7 \text{ s}^{-1}$ . Unfortunately, it was impossible to go to lower strain rates, due to inherent computational limitations of atomic simulations. Refer to Figure 7.

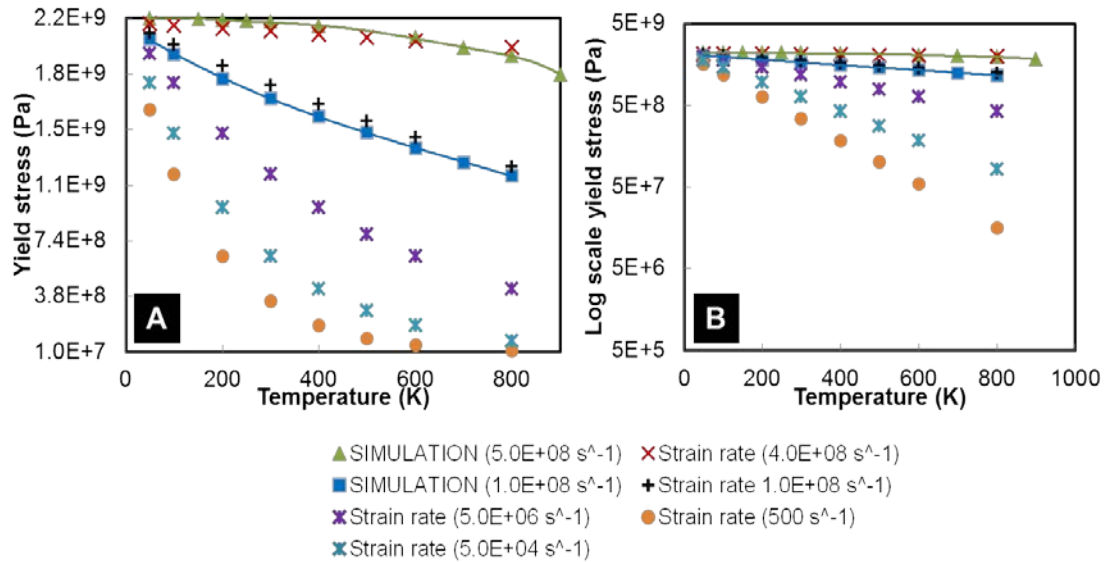


Figure 5: Comparison of yield stress-T curves at different strain rates between MD simulations and model predictions: (A) Yield stress in Pa, (B) Logarithmic scale yield stress

Figure 7 shows solid lines interpolated from the yield stress obtained with simulations at constant strain rate, corresponding with the peak stress value obtained prior to the first dislocation nucleation event. The data points without corresponding solid lines are values calculated directly using Equation 5. Figure 7B is identical to Figure 7A, however with a logarithmic scale on the y-axis. Figure 7B provides a clear demonstration that the temperature dependence does not reach an asymptote at  $T > 600\text{K}$  and  $\dot{\epsilon} > 5.0 \times 10^4 \text{ s}^{-1}$ .

The results in Figure 7 demonstrate remarkable consistency between the simulated and analytically predicted yield stress, as a function of strain rate and temperature between  $1.0 \times 10^7 \leq \dot{\epsilon} \leq 5.0 \times 10^8$  and up to 800K. This is not coincidental, as the temperature dependence is significantly strain-rate dependent, as demonstrated by the significant difference between  $\dot{\epsilon} = 5.0 \times 10^8$  and  $\dot{\epsilon} = 5.0 \times 10^7$  by 1 order of magnitude at 800K. This provides an extremely significant timescale link between atomic simulation ( $< 100\mu\text{s}$ ) and experiments ( $> 1\text{ms}$ ), and a valuable constitutive model.

### 3.5 Effectiveness of model at low strain rate and significance

This section will discuss the suitability of the model derived from an energy-based criterion to effectively predict the rate-dependence of yield stress and the significance of this for atomic simulation studies.

The primary challenge of simulation-based studies, is establishing an effective link between the idealised, theoretical model and the real-world properties. Fundamentally, energy criteria are ideal, as they are pure thermodynamic values that are time-size-temperature independent, and can also be used to explain complex “real-world” non-ideal material properties by accounting for defect energies. However, this requires a consistent mechanism when effects of strain rate and material defects are included. For example, due to time-dependent effects, “slow” processes such as mechanical creep may not be observed in atomic simulations, because they require longer timeframes ( $> 1\text{s}$ ). However, it can be assumed to be a very good approximation of homogenous dislocation nucleation processes, such as is consistently observed in nano-indentation experiments [Lorenz, Zeckzer et al. (2003; Zhao, Ma et al. (2012)].

The model obtained corresponds very well with previous studies [Asaro and Suresh (2005; Zhu, Li et al. (2008; Deng and Sansoz (2010; Zhang, Liu et al. (2013))], and there is evidence that the energy-based criterion is effective for atomic-experimental timescale linking [Zhu, Li et al. (2008)].

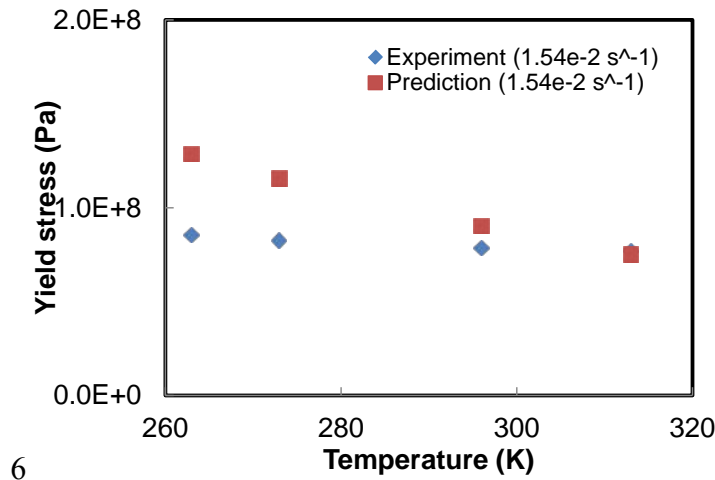


Figure 6a: Small temperature range comparison of experiment and predicted results

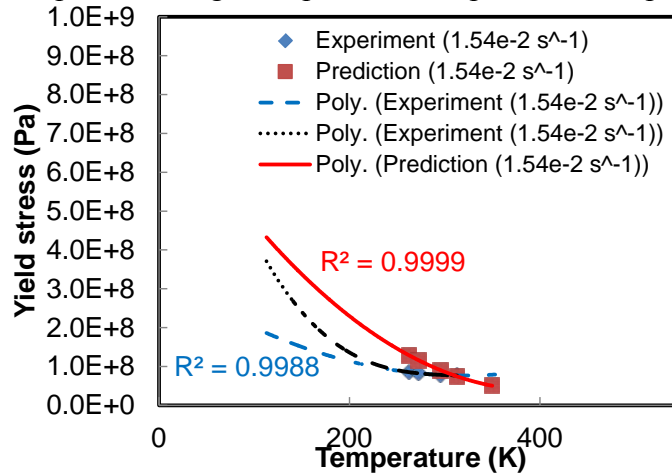


Figure 7b: Extrapolation of polynomial regression models showing results may diverge

Figure 8: Comparison of predicted and experimental flow stress at  $\dot{\epsilon} = 1.54 \times 10^{-2} \text{ s}^{-1}$

Although the accuracy of the model derived in this paper is very good for all the testable strain rates, it cannot be validated at low strain rates (i.e.,  $\dot{\epsilon} < 5.0 \times 10^6$ ) by the same atomic simulation approach, as explained in the introduction. The flow stress provides an analogue for the experimental nucleation yield stress [Deng and Sansoz (2010)]. Hence, Figure 8 validates the model by comparing the flow stress as a function of temperature from an experimental study with pure FCC Al [Rosen and Bodner (1967)] with the predicted yield stress at a strain rate of  $\dot{\epsilon} = 1.54 \times 10^{-2} \text{ s}^{-1}$ . This provides an indication of the validity of the model for low strain rate regimes by comparing with results in the literature. Note: this is limited due to the small data set available experimentally [Rosen and Bodner (1967)]. The results are very consistent around the temperature of 300K. However, the results appear to diverge slightly as the experimental temperature decreases below 0°C. To test the divergence, the results are backwards extrapolated with regression models, showing that a 2nd order model will exhibit fairly significant divergence as  $T \rightarrow 0K$ . However, the result is remarkably consistent considering that the simulation results at  $\dot{\epsilon} = 5.0 \times 10^8 \text{ s}^{-1}$  and



300K is 28 times greater than the experimental result. The result is a very promising time and size link from atomistic to experimental results.

## 4.0 Conclusions

This study demonstrates that the non-elastic mechanical properties of crystalline materials can be defined by the atomic-scale crystal defect processes, which are driven by the thermal and mechanical limits. This paper shows that the yield stress (mechanical limit) in a defect-free single crystal can be correlated with the activation energy for dislocation nucleation. The activation energy is hence used as a fundamental-basis to model the temperature-dependence (thermal limits). More significantly, this paper demonstrates that the stress-dependence of the activation energy (i.e., the activation volume) can be used to accurately predict the effect of strain rate on the temperature-dependence of mechanical strength. In other words, this paper shows that a fundamental energy criterion from atomistic simulations can be used to derive an effective constitutive model for temperature- and rate- dependent thermo mechanical properties! This is a critical insight, because it enables a timescale link between atomistic simulations at very high strain rates ( $\dot{\epsilon} > 5.0 \times 10^6 \text{ s}^{-1}$ ), and macro-scale simulations and/or experiments.

## 5.0 References

- Asaro, R. J. and S. Suresh (2005). "Mechanistic models for the activation volume and rate sensitivity in metals with nanocrystalline grains and nano-scale twins." *Acta Materialia* **53**(12): 3369-3382.
- Deng, C. and F. Sansoz (2010). "Effects of twin and surface facet on strain-rate sensitivity of gold nanowires at different temperatures." *Physical Review B* **81**(15): 155430.
- Ghoniem†, N. M., E. P. Busso, et al. (2003). "Multiscale modelling of nanomechanics and micromechanics: an overview." *Philosophical Magazine* **83**(31-34): 3475-3528.
- Henkelman, G., B. P. Uberuaga, et al. (2000). A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*. **113**: 9901-9904.
- Hoover, W. G. (1986). "Constant-pressure equations of motion." *Physical Review A* **34**(3): 2499-2500.
- Kelchner, C. L., S. J. Plimpton, et al. (1998). "Dislocation nucleation and defect structure during surface indentation." *Phys Rev B: Condens Matter and Mater Phys* **58**(17): 11085-11088.
- Li, J. (2003). "AtomEye: an efficient atomistic configuration viewer." *Modelling and Simulation in Materials Science and Engineering* **11**(2): 173.
- Lorenz, D., A. Zeckzer, et al. (2003). "Pop-in effect as homogeneous nucleation of dislocations during nanoindentation." *Physical Review B* **67**(17): 172101.
- McPhie, M. G., S. Berbenni, et al. (2012). "Activation energy for nucleation of partial dislocation from grain boundaries." *Computational Materials Science* **62**(0): 169-174.
- Mishin, Y., D. Farkas, et al. (1999). "Interatomic potentials for monoatomic metals from experimental data and ab initio calculations." *Physical Review B* **59**(5): 3393-3407.
- Nguyen, T., P. S. Ho, et al. (1991). *Molecular Dynamics Study of Thermal Disorder in a Bicrystal Model*. MRS Proceedings, Cambridge Univ Press.
- Nosé, S. (1984). "A unified formulation of the constant temperature molecular dynamics methods." *The Journal of Chemical Physics* **81**(1): 511-519.
- Plimpton, S. J. (1995). "Fast Parallel Algorithms for Short-Range Molecular Dynamics." *Journal of Computational Physics* **117**(Refer to: <http://lammps.sandia.gov>): 1-19.
- Po, G., M. Lazar, et al. (2014). "Singularity-free dislocation dynamics with strain gradient elasticity." *Journal of the Mechanics and Physics of Solids* **68**(0): 161-178.
- Po, G., M. Mohamed, et al. (2014). "Recent Progress in Discrete Dislocation Dynamics and Its Applications to Micro Plasticity." *JOM* **66**(10): 2108-2120.
- Rosen, A. and S. R. Bodner (1967). "The influence of strain rate and strain ageing on the flow stress of commercially-pure aluminium." *Journal of the Mechanics and Physics of Solids* **15**(1): 47-62.
- Sangid, M. D., T. Ezaz, et al. (2011). "Energy of slip transmission and nucleation at grain boundaries." *Acta Materialia* **59**(1): 283-296.

- Štich, I., R. Car, et al. (1989). "Conjugate gradient minimization of the energy functional: A new method for electronic structure calculation." Physical Review B **39**(8): 4997.
- Tschopp, M. A. and D. L. McDowell (2007). "Asymmetric tilt grain boundary structure and energy in copper and aluminium." Philosophical Magazine **87**(25): 3871-3892.
- Tucker, G. J. and D. L. McDowell (2011). "Non-equilibrium grain boundary structure and inelastic deformation using atomistic simulations." International Journal of Plasticity **27**(6): 841-857.
- Voter, A. F., F. Montalenti, et al. (2002). "EXTENDING THE TIME SCALE IN ATOMISTIC SIMULATION OF MATERIALS." Annual Review of Materials Research **32**(1): 321-346.
- Zbib, H. M. and T. A. Khraishi (2005). Dislocation Dynamics. Handbook of Materials Modeling. S. Yip, Springer Netherlands: 1097-1114.
- Zhang, J. Y., G. Liu, et al. (2013). "Strain rate effects on the mechanical response in multi- and single-crystalline Cu micropillars: Grain boundary effects." International Journal of Plasticity **50**(0): 1-17.
- Zhao, J., L. F. Ma, et al. (2012). "Surface dislocation nucleation by wedge indenter contacts." Materials Science and Technology **28**(9-10): 1167-1172.
- Zhu, T. and J. Li (2010). "Ultra-strength materials." Progress in Materials Science **55**(7): 710-757.
- Zhu, T., J. Li, et al. (2009). "Mechanics of ultra-strength materials." MRS bulletin **34**(3): 167.
- Zhu, T., J. Li, et al. (2008). "Temperature and Strain-Rate Dependence of Surface Dislocation Nucleation." Physical Review Letters **100**(2): 025502.

## **Magnetohydrodynamic nanofluid flow and heat transfer along a permeable stretching surface with non uniform heat generation/absorption**

Md. S. Ansari <sup>a</sup>, R. Nandkeolyar <sup>b</sup> and S. S. Motsa <sup>b</sup>

<sup>a</sup> School of Technology, Pandit Deendayal Petroleum University, Gandhinagar - 382007, India

<sup>b</sup> School of Mathematics, Statistics & Computer Science, University of KwaZulu-Natal, Pietermaritzburg-3209, South Africa

**Abstract:** This paper investigates the flow of a nanofluid past a stretching permeable sheet under the influence of transverse magnetic field. The space and temperature dependent heat source/sink effect is considered. The governing partial differential equations are transformed into ordinary differential equations with the help suitable similarity transformation. The transformed ordinary differential equations are solved using Spectral Relaxation Method. To validate the accuracy of the method, a comparison of Nusselt and Sherwood number with the data of previous work is presented, which shows an excellent agreement. The effects of magnetic field, suction, space and temperature dependent heat source/sink, Lewis number, thermophoresis and Brownian motion on the flow field, temperature and nanofluid particle concentration are studied with the help of graphs and tables.

**Keyword:** Magnetic field, Nanofluid, Non uniform heat source/sink, Spectral Relaxation Method, similar solution.

### **1. Introduction**

Nanofluids are suspensions of nanoparticles (nominally 1-100nm in size) in conventional base fluids such as water, oils or ethylene glycols. The nanoparticles used in nanofluids are typically made of metals, carbides, oxides or carbon nanotubes. Nanofluids have interesting properties that make them potentially useful in many applications in heat transfer [1] including microelectronics, fuel cells, pharmaceutical processes, and hybrid-powered engines, engine cooling, vehicle thermal management, domestic refrigerator, chiller, heat exchanger, nuclear reactor coolant, in grinding, in space technology, defense and ships, and in boiler flue gas temperature reduction. They exhibit enhanced thermal conductivity and the convective heat transfer coefficient compared to the base fluid [2]. Choi [3] is the first who used the term nanofluids to refer to the fluid with suspended nano particles. Choi et al. [4] reported that addition of small amount (less than 1% by volume) of nanoparticles to conventional heat transfer liquids increased the thermal conductivity of the fluid upto approximately two times.

Magnetohydrodynamic boundary layer flow of nanofluid and heat transfer over a linear stretching sheet has diverse applications in industrial, scientific and engineering problems such as boundary layer flow control, MHD power generators, micro MHD pumps, the cooling of nuclear reactors, plasma studies, geothermal energy extraction, cooling of large metallic plates in bathes, micro mixing of physiological samples, biological transportation and drug delivery [5-8]. Crane [8] was the first to analyze boundary layer flow of a Newtonian fluid induced by a stretching of elastic sheet moving in its own plane linearly. Several researchers [9 -13] studied analytically \ numerically the various aspects of problem of nanofluid flow over a stretching surface under different conditions. Salari et al. [14] analyzed numerically the heat transfer of nanofluid over a flat stretching sheet considering two sets of boundary conditions viz (i) a

constant and (ii) a linear streamwise variation of nanoparticle volume fraction and wall temperature. The boundary layer flow and heat transfer over a permeable stretching sheet due to a nanofluid under the influence of external magnetic field, slip boundary condition and thermal radiation was investigated by Ibrahim and Shankar [15]. Ibrahim et al. [16] studied theoretically the effect of magnetic field on stagnation point flow and heat transfer of nanofluid flow over a stretching sheet. It was reported that the heat transfer rate at the surface increases with the magnetic parameter when the free stream velocity exceeds the stretching sheet velocity. Makinde et al. [17] analyzed the combined effects of buoyancy force, convective heating, Brownian motion and thermophoresis on the stagnation point flow and heat transfer of an electrically conducting nanofluid towards a stretching sheet under the influence of magnetic field. Qasim et al. [18] investigated the problem of magnetohydrodynamic flow of ferrofluid along a stretching cylinder with velocity slip and prescribed surface heat flux. Bhattacharyya and Layek [19] considered the boundary layer flow of a nanofluid due to an exponentially permeable stretching sheet with external magnetic field.

The study of temperature dependent heat source/sink on the heat transfer characteristics is interesting owing to significant temperature differences between the surface and ambient-fluid in various engineering applications. The effect of space- and temperature dependent internal heat generation/absorption on the heat transfer characteristics of the flow have been studied by several researchers [20 – 25] under different conditions and configurations. Abel and Mahantesh [26] presented an analytical solution of the viscoelastic fluid flow and heat transfer over a stretching sheet under the influence of non uniform heat source/sink and uniform magnetic field. Ramesh et al. [27] investigated numerically the steady MHD flow of a dusty fluid near the stagnation point past a permeable stretching sheet with the effect of non-uniform heat source/sink. In view of the above referred works we consider the flow of nanofluid past a permeable stretching sheet under the influence of an external magnetic field taking into account the effect of the space and temperature dependent heat source/sink. The nonlinear governing partial differential equations for the flow, energy and nanoparticle concentration are transformed into ordinary differential equations using similarity transformation and are then solved numerically by Spectral Relaxation Method.

## 2. Mathematical Formulation

Consider two dimensional steady boundary layer flow of a viscous, incompressible, electrically conducting, heat generating/ absorbing nanofluid past a stretching sheet in a quiescent fluid. The velocity of stretching sheet is  $U_w = ax$  (where  $a > 0$  is the constant acceleration parameter). The  $x$ -axis is taken along the sheet in vertically upward direction and  $y$ -axis is taken normal to the sheet. The surface of the sheet is maintained at uniform temperature and concentration,  $T_w$  and  $C_w$ , respectively, and these values are assumed to be greater than the ambient temperature and concentration,  $T_\infty$  and  $C_\infty$ , respectively. The whole system is permeated by a uniform transverse magnetic field  $B_0$  which is applied parallel to  $y$ -axis. The induced magnetic field is neglected in comparison to applied one. It is assumed that both the fluid and nanoparticles are in thermal equilibrium state. The thermo physical properties of the nanofluid are assumed to be constant. The pressure gradient and external forces are neglected. In addition, the fluid suction is imposed at the sheet surface in the  $y$ - direction.

Under the above assumptions and usual boundary layer approximation, the magnetohydrodynamic steady nanofluid flow, heat and mass transfer with internal heat generation/absorption are governed by the following equations:

The continuity equation

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0, \quad \dots(1)$$

The momentum equation equation

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = \nu \frac{\partial^2 u}{\partial y^2} - \frac{\sigma B_0^2}{\rho_f} u, \quad \dots(2)$$

The energy equation

$$u \frac{\partial T}{\partial x} + v \frac{\partial T}{\partial y} = \alpha \frac{\partial^2 T}{\partial y^2} + \tau \left[ D_B \frac{\partial C}{\partial y} \frac{\partial T}{\partial y} + \frac{D_T}{T_\infty} \left( \frac{\partial T}{\partial y} \right)^2 \right] + \frac{q'''}{(\rho c)_f}, \quad \dots(3)$$

The nanoparticle concentration equation

$$u \frac{\partial C}{\partial x} + v \frac{\partial C}{\partial y} = D_B \frac{\partial^2 C}{\partial y^2} + \frac{D_T}{T_\infty} \frac{\partial^2 T}{\partial y^2}, \quad \dots(4)$$

where  $u$  and  $v$  are velocity components along  $x$ - and  $y$ - directions, respectively.  $\alpha$ ,  $\nu$ ,  $\rho$ ,  $c$ ,  $(\rho c)_p$ ,  $(\rho c)_f$ ,  $D_B$ ,  $D_T$  and  $\tau$  are, respectively, thermal diffusivity, kinematic viscosity, mass density, specific heat, effective heat capacity of the nanoparticle material, heat capacity of the fluid, Brownian diffusion coefficient, thermophoresis diffusion coefficient and a parameter defined by  $(\rho c)_p / (\rho c)_f$ .

The associated boundary conditions are

$$\left. \begin{array}{l} \text{at } y=0: u = U_w(x) = ax, v = -V_0; T = T_w; C = C_w \\ \text{as } y \rightarrow \infty: u \rightarrow 0; T \rightarrow T_\infty; C \rightarrow C_\infty \end{array} \right\}, \quad \dots(5)$$

In order for similarity condition to exist, the non uniform heat source/sink  $q'''$  is modeled as [28]

$$q''' = \left( \frac{kU_w(x)}{x\nu} \right) \left[ A^* (T_w - T_\infty) f' + B^* (T - T_\infty) \right], \quad \dots(6)$$

where  $k$  is thermal conductivity,  $A^*$  and  $B^*$  are parameters of space-dependent and temperature-dependent heat generation/absorption. Both  $A^*$  and  $B^*$  positive corresponds to internal heat source and negative to internal heat sink.

To transform the governing equations into a set of similarity equations, the following transformations are introduced

$$\eta = \sqrt{\frac{a}{\nu}} y, \quad \psi(x, y) = \sqrt{a\nu} x f(\eta), \quad \theta = \frac{T - T_\infty}{T_w - T_\infty}, \quad \phi = \frac{C - C_\infty}{C_w - C_\infty}. \quad \dots(7)$$

where  $\psi$  is the stream function,  $f(\eta)$  is a dimensionless stream function,  $\theta$  is dimensionless temperature,  $\phi$  is dimensionless concentration function and  $\eta$  is similarity variable. Continuity equation is satisfied if the velocity components are taken as

$$u = \frac{\partial \psi}{\partial y}, \quad v = -\frac{\partial \psi}{\partial x} \quad \dots(8)$$

Eqns. (1) – (4), after similarity transformation, are

$$f''' + ff'' - f'^2 - Mf' = 0, \quad \dots(9)$$

$$\frac{1}{Pr} [\theta'' + (A^* f' + B^* \theta)] + f\theta' + N_b \phi' \theta' + N_t \theta'^2 = 0, \quad \dots(10)$$

$$\phi'' + Le f \phi' + \frac{N_t}{N_b} \theta'' = 0, \quad \dots(11) \text{ and}$$

the associated boundary conditions become

$$f'(0) = 1, f(0) = S, \theta(0) = 1, \phi(0) = 1, f'(\infty) \rightarrow 0, \theta(\infty) \rightarrow 0, \phi(\infty) \rightarrow 0, \quad \dots(12)$$

where

$$\left. \begin{aligned} M &= \frac{\sigma B_0^2}{a\rho}, \quad Pr = \frac{\nu}{\alpha}, \quad N_b = \frac{\tau D_B (C_w - C_\infty)}{\nu} \\ N_t &= \frac{\tau D_T (T_w - T_\infty)}{\nu T_\infty}, \quad Le = \frac{\nu}{D_B}, \quad S = \frac{V_0}{\sqrt{av}}. \end{aligned} \right\} \text{governing parameters}$$

where  $M$ ,  $Pr$ ,  $N_b$ ,  $N_t$ ,  $Le$  and  $S$  are, respectively, magnetic parameter, Prandtl number, Brownian motion parameter, thermophoresis parameter, Lewis number and suction parameter.

The physical quantities of interest, the local skin friction coefficient  $Cf_x$ , the local Nusselt number  $Nu_x$  and the local Sherwood number  $Sh_x$ , are defined as:

$$Cf_x Re_x^{1/2} = f''(0), \quad \dots(13)$$

$$Nu_x Re_x^{-1/2} = -\theta'(0), \quad \dots(14)$$

$$Sh_x Re_x^{-1/2} = -\phi'(0), \quad \dots(15)$$

where  $Re_x = \frac{U_w x}{\nu}$  is the local Reynolds number.

### 3. Solution Technique

To solve the non linear boundary value problem described by eqns. (9) – (12), We adopted the spectral relaxation method (SRM) [29]. To apply SRM on eqns. (9) – (12), we set  $f'(\eta) = g(\eta)$

Then the eqns. (9) – (11) reduce to the following system of equations:

$$f' = g, \quad \dots(16)$$

$$g'' + fg' - g^2 - Mg = 0, \quad \dots(17)$$

$$\frac{1}{Pr} [\theta'' + (A^* f' + B^* \theta)] + f\theta' + N_b \phi' \theta' + N_t \theta'^2 = 0, \quad \dots(18)$$

$$\phi'' + Lef\phi' + \frac{N_t}{N_b}\theta'' = 0, \quad \dots(19)$$

And the boundary conditions (12) become

$$g(0) = 1, f(0) = S, \theta(0) = 1, \phi(0) = 1, \quad g(\infty) \rightarrow 0, \theta(\infty) \rightarrow 0, \phi(\infty) \rightarrow 0, \quad \dots(20)$$

In frame work of SRM, writing the iteration scheme as

$$f'_{r+1} = g_r, f_{r+1}(0) = S, \quad \dots(21)$$

$$g''_{r+1} + f_{r+1}g'_{r+1} - Mg_{r+1} = g_r^2, \quad g_{r+1}(0) = 1, g_{r+1}(\infty) = 0, \quad \dots(22)$$

$$\theta''_{r+1} + \text{Pr} f_{r+1}\theta'_{r+1} + B^*\theta_{r+1} = -(\text{Pr} N_b\phi'_r\theta'_r + \text{Pr} N_t\theta_r'^2 + A^*g_{r+1}), \quad \theta_{r+1}(0) = 1, \theta_{r+1}(\infty) = 0, \quad \dots(23)$$

$$\phi''_{r+1} + Lef_{r+1}\phi'_{r+1} = -\frac{N_t}{N_b}\theta''_{r+1}, \quad \phi_{r+1}(0) = 1, \phi_{r+1}(\infty) = 0. \quad \dots(24)$$

The domain is  $0 \leq \eta \leq \eta_\infty$  and  $\eta_\infty$  is the edge of the boundary layer. Using the mapping  $\xi = \frac{2\eta}{\eta_\infty} - 1$

the domain  $[0, \eta_\infty]$  is mapped into the computational domain  $[-1, 1]$ . Defining the grid points

$\xi_j = \cos\left(\frac{\pi j}{N}\right)$ , where  $N$  is the number of grid points,  $j = 0, 1, 2, \dots, N$ . . Applying the Chebyshev pseudo-

spectral method on eqns. (21) – (24), we obtain

$$A_1 F_{r+1} = B_1, \quad f_{r+1}(\xi_N) = S, \quad \dots(25)$$

$$A_2 G_{r+1} = B_2, \quad g_{r+1}(\xi_N) = 1, g_{r+1}(\xi_0) = 0, \quad \dots(26)$$

$$A_3 \Theta_{r+1} = B_3, \quad \theta_{r+1}(\xi_N) = 1, \theta_{r+1}(\xi_0) = 0, \quad \dots(27)$$

$$A_4 \Phi_{r+1} = B_4, \quad \phi_{r+1}(\xi_N) = 1, \phi_{r+1}(\xi_0) = 0, \quad \dots(28)$$

where

$$A_1 = D, \quad B_1 = g_r, \quad A_2 = D^2 + \text{diag}(f_{r+1})D - MI, \quad B_2 = g_r^2,$$

$$A_3 = D^2 + \text{diag}(\text{Pr} f_{r+1})D + B^*I, \quad B_3 = -[\text{Pr} N_b\phi'_r\theta'_r + \text{Pr} N_t\theta_r'^2 + A^*g_{r+1}],$$

$$A_4 = D^2 + \text{diag}(Lef_{r+1})D, \quad B_4 = -\frac{N_t}{N_b}\theta''_{r+1},$$

with  $I$  and  $D$  are being an  $(N+1) \times (N+1)$  identity and differentiation matrix, respectively.

$F, G, \Theta$  and  $\Phi$ , are the values of functions  $f, g, \theta$  and  $\phi$ , respectively, computed at the grid points.

The initial approximations satisfying the boundary conditions, used to start the iteration process, are

$$f_0(\eta) = S - e^{-\eta} + 1, \quad g_0(\eta) = e^{-\eta}, \quad \theta_0(\eta) = e^{-\eta} \quad \text{and} \quad \phi_0(\eta) = e^{-\eta}.$$

## 4. Results and Discussion

To demonstrate the accuracy of our numerical results, Tables 1 and 2 are the values of  $-\theta'(0)$  and  $-\phi'(0)$  for the case when  $M = 0, A^* = 0, B^* = 0$  and  $S = 0$  between our calculations and the data presented by Khan and Pop [9]. These tables show good agreement. Figs. 1 and 2 demonstrate the distribution of dimensionless velocity  $f'(\eta)$  and temperature  $\theta(\eta)$  for different values of magnetic parameter  $M$  when  $Pr = 6.7850, N_t = 0.3, N_b = 0.2, Le = 2, A^* = 0.2, B^* = 0.2$  and  $S = 0.5$ . The velocity reduces with the increase in Magnetic parameter. The presence of magnetic field in an electrically conducting fluid tends to produce a body force against the flow. This type of resistive force tends to slow down the motion of the fluid in the boundary layer region which, in turn, increases the temperature, this is exactly what we observe in Figs. 1 and 2. Consequently, the thermal boundary layer thickness becomes thicker for stronger magnetic field.

Table 1: Comparison of the results for reduced Nusselt number  $-\theta'(0)$

$N_b$	$N_t$	Present results $M = A^* = B^* = S = 0$	Khan and Pop[9]
0.1	0.1	0.9524	0.9524
	0.2	0.6932	0.6932
	0.3	0.5201	0.5201
0.2	0.1	0.5056	0.5056
	0.2	0.3654	0.3654
	0.3	0.2731	0.2731
0.3	0.1	0.2522	0.2522
	0.2	0.1816	0.1816
	0.3	0.1355	0.1355

Table 2: Comparison of the results for reduced Sherwood number  $-\phi'(0)$

$N_b$	$N_t$	Present results $M = A^* = B^* = S = 0$	Khan and Pop [9]
0.1	0.1	2.1294	2.1294
	0.2	2.2740	2.2740
	0.3	2.5286	2.5286
0.2	0.1	2.3819	2.3819
	0.2	2.5152	2.5152
	0.3	2.6555	2.6555
0.3	0.1	2.4100	2.4100
	0.2	2.5150	2.5150
	0.3	2.6088	2.6088



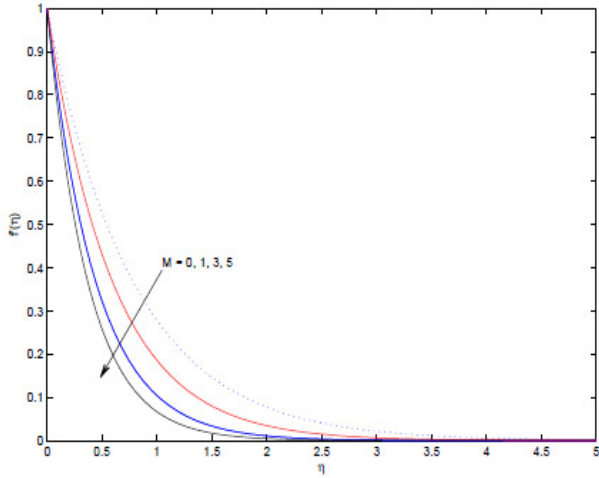


Figure 1: Velocity profiles for different  $M$ .

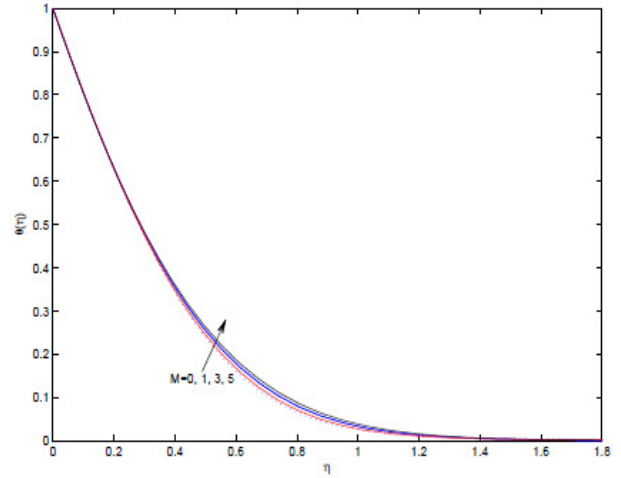


Figure 2: Temperature profiles for different  $M$ .

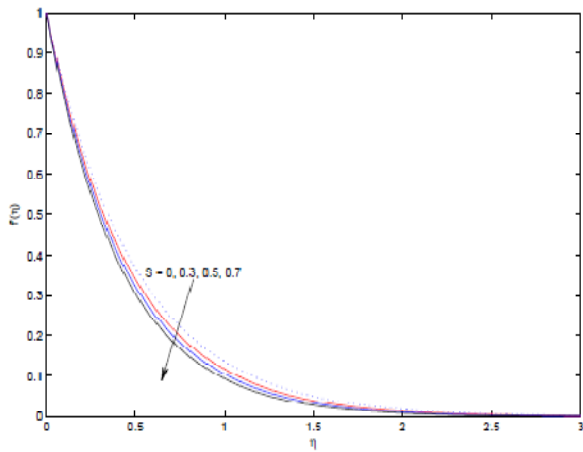


Figure 3: Variations in  $f'(\eta)$  profiles for different  $S$ .

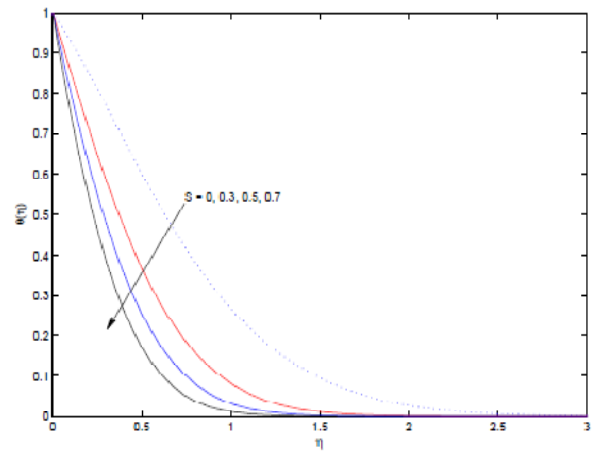


Figure 4: Variations in  $\theta(\eta)$  profiles for different  $S$ .

Figs. 3 and 4 depict the influence of suction parameter  $S$  on the velocity and temperature profiles in the boundary layer respectively. Due to imposition of wall suction, the fluid is brought closer to the sheet and it reduces momentum boundary layer thickness as well as the thermal boundary layer thicknesses. This causes reduction in the velocity and temperature profiles.

The temperature profiles for different space-dependent and temperature-dependent heat source/sink parameters are presented in Figs. 5 and 6 when  $Pr = 6.7850, N_t = 0.3, N_b = 0.2, Le = 2, A^* = 0.2, B^* = 0.2, M = 3$  and  $S = 0.5$ . On increasing values of  $A^*$  and/ or  $B^*$  produce increase in the temperature distributions of the nanofluid. This is expected since the presence of heat source  $A^* > 0$  and/ or  $B^* > 0$  in the boundary layer generates energy which causes the temperature of the fluid to increase. Heat sink  $A^* < 0$  and/ or  $B^* < 0$  has the opposite effect, namely cooling of the fluid. The thermal field exerts forces on all molecules and nano particles in the nanofluid forcing them to move, in the direction of the heat flow, i.e., from the hot side to the cold side.

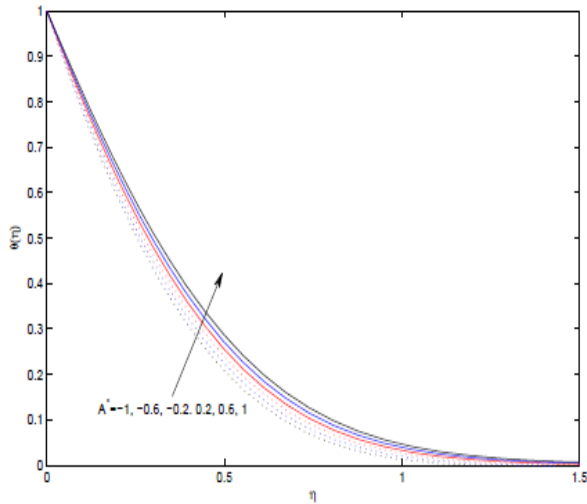


Figure 5: Temperature profiles for various values of  $A^*$ .

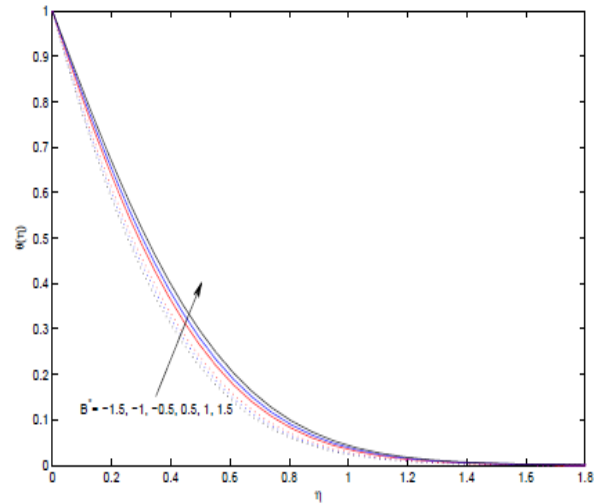


Figure 6: Temperature profiles for various values of  $B^*$ .

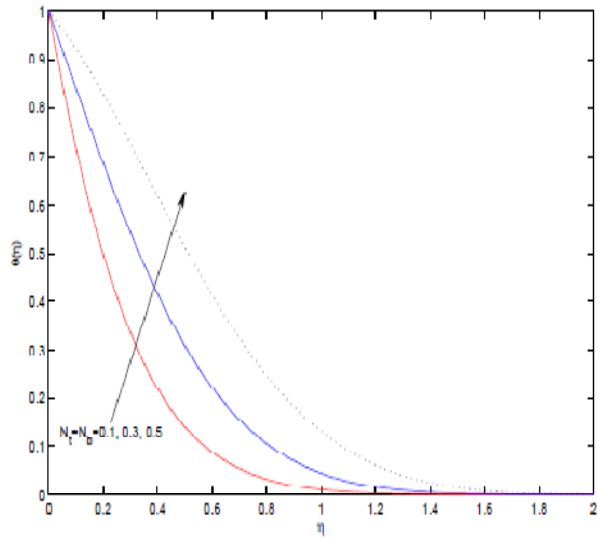


Figure 7: Effects of  $N_t$  and  $N_b$  on temperature profiles

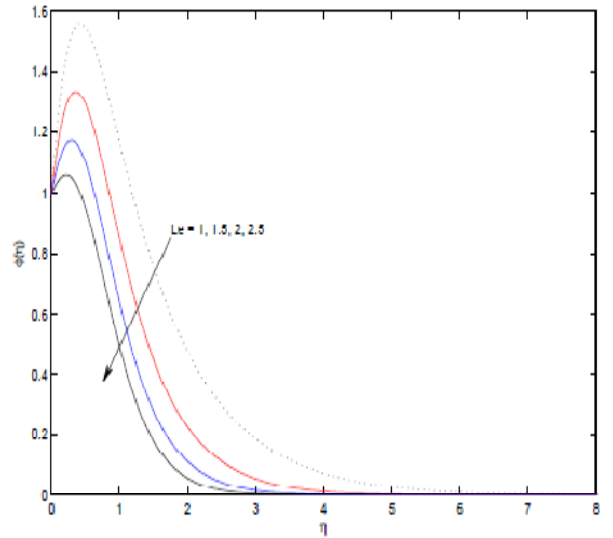


Figure 8: Effects of  $Le$  on concentration profiles

Fig. 7 presents the effects of the Bownian motion and thermophoresis effect on the temperature distribution. The figure reveals that the temperature of the fluid increases with increasing values of  $N_t$  and  $N_b$ . The Brownian motion takes place due to the presence of nanoparticles and for the increase in  $N_b$  the Brownian motion is affected and consequently the heat transfer characteristics of the fluid changes. Increase in  $N_t$  causes increment in the thermophoresis force which tends to move nanoparticles from hot to cold side and consequently increases the temperature of the fluid.

The variation of nanoparticle concentration for various values of Lewis number is shown in Fig. 8. It is observed that Lewis number significantly affects the concentration distribution. There is decrease in volume fraction on increasing Lewis number, since higher Lewis number implies a lower Brownian

diffusion coefficient  $N_b$  which result in a shorter penetration depth for the concentration boundary layer. It is also noticed that the volume fraction initially increases near the sheet and after attaining a peak it decreases away from the sheet for smaller values of  $Le$ .

**Table 3: Variations of skin friction, Nusselt number and Sherwood number for various values of  $M, S, A^*, B^*$  with  $Pr = 6.7850, Le = 2, N_t = 0.3$  and  $N_b = 0.2$ .**

$M$	$S$	$A^*$	$B^*$	$-f''(0)$	$-\theta'(0)$	$-\phi'(0)$
1	0.5	0.2	0.2	1.686141	2.000978	-0.955299
3				2.265564	2.037771	-1.173981
5				2.712214	2.056024	-1.305090
7				3.089454	2.065771	-1.394071
3	0.3	0.2	0.2	2.155617	1.427646	-0.600293
	0.5			2.265564	2.037771	-1.173981
	0.7			2.380394	2.686705	-1.805905
	0.9			2.500000	3.361406	-2.473698
3	0.5	-0.6	0.2	-	2.249771	-1.476083
		-0.2		-	2.144097	-1.325506
		0.2		-	2.037771	-1.173981
		0.6		-	1.930788	-1.021497
3	0.5	0.2	-0.6	-	2.218963	-1.434683
			-0.2	-	2.130147	-1.306829
			0.2	-	2.037771	-1.173981
			0.6	-	1.941409	-1.035540

Table 3 represents the variations in local skin friction coefficient  $-f''(0)$ , local Nusselt number  $-\theta'(0)$  and local sherwood number  $-\phi'(0)$  for various values of  $M, S, A^*$  and  $B^*$ . For stronger magnetic field the values of  $-f''(0)$  and  $-\phi'(0)$  increases. A very minor increase in  $-\theta'(0)$  is observed with the increase in  $M$ . Negative values of  $-\phi'(0)$  indicate that nanoparticle is transferred from the fluid to moving surface as discussed before.  $-f''(0)$ ,  $-\theta'(0)$  and  $-\phi'(0)$  are the increasing function of suction parameter  $S$ .  $-\theta'(0)$  and  $-\phi'(0)$  are decreasing functions of heat source that is,  $A^* > 0$  and  $B^* > 0$  while increasing with heat sink  $A^* < 0$  and  $B^* < 0$ .

**Table 4: Variations of Nusselt number and Sherwood number for various values of  $Le, N_t, N_b$  with  $Pr = 6.7850, M = 3, S = 0.5, A^* = 0.2$  and  $B^* = 0.3$ .**

$Le$	$N_t$	$N_b$	$-\theta'(0)$	$-\phi'(0)$
1	0.3	0.2	2.723933	-3.002687
1.5			2.309029	-1.962483
2			2.014073	-1.139920
2.5			1.796686	-0.457844
2	0.05	0.2	2.709139	0.904751
	0.1		2.553545	0.372000
	0.3		2.014073	-1.139920
	0.5		1.595297	-1.928164
	0.7		1.276229	-2.275771
2	0.3	0.1	2.523860	-5.265578
		0.2	2.014073	-1.139920
		0.3	1.572160	0.167136
		0.4	1.197889	0.769972
		0.5	0.888746	1.092839

The numerical values of  $-\theta'(0)$  and  $-\phi'(0)$  for different values of  $Le, N_b$  and  $N_t$  are displayed in Table 4. It is evident from table 4 that  $Le$  has decreasing influence on both  $-\theta'(0)$  and  $-\phi'(0)$ .  $-\theta'(0)$  decreases with thermophoresis and Brownian motion. As thermophoresis increases,  $-\phi'(0)$  decreases, attains a minimum value and again increases in the direction from fluid to the wall. Brownian motion decreases  $-\phi'(0)$ , approaches a minimum value and again increases in opposite direction, that is, from the wall to the fluid.

### 5. Conclusion

The steady laminar, hydromagnetic nano fluid flow, heat and mass transfer adjacent to a permeable, continuously stretching sheet with linear surface velocity in the presence of magnetic, heat generation/absorption which is a function of both space and temperature is investigated. The nonlinear governing equations for the flow are solved numerically using spectral relaxation method.

Numerical results for the velocity, temperature and concentration are presented graphically for various values of parameters. In addition, Numerical data for the local skin-friction coefficient, the local Nusselt number and the local Sherwood number are tabulated for various values of magnetic field parameter, coefficients of space-dependent and time-dependent internal heat generation/absorption, suction, Lewis number, thermophoresis and the Brownian motion parameters. It is found that: The momentum boundary layer become thinner and the thermal boundary layer thickness and nanoparticle volume fraction boundary layer thickness become thicker for stronger magnetic field. The local Nusselt number and local Sherwood number decreases as both space-dependent and temperature-dependent internal heat generation coefficients increase. The opposite impact is observed as both space-dependent and temperature-

dependent internal heat absorption coefficients increase. Wall fluid suction have increasing influence on the local skin friction coefficient, local Nusselt number and local Sherwood number.

## References

- [1] W.J Minkowycz, E. M. Sparrow, J. P. Abraham, Nanoparticle Heat Transfer and Fluid flow, CRC Press, 2012.
- [2] S. Kakac, A. Pramuanjarokij, Review of convective heat transfer enhancement with nanofluids, *Int. J. Heat Mass Tansf.* 52 (2009) 3817-3196.
- [3] S. U. S. Choi, Enhancing thermal conductivity of fluids with nanoparticle, in: *The Proceedings of the 1995 ASME International Mechanical Engineering Congress and Exposition, San Francisco, USA, ASME, FED 231/MD 66, 1995, pp.99-105, 1995.*
- [4] S. U. S. Choi, Z. G. Zhang, W. Yu, F. E. Lockwood, E. A. Grulke, Anomalously thermal conductivity enhancement in nanotube suspension, *Appl. Phys. Lett.* 79 (2001) 2252-2254.
- [5] C. Kleinstreuer, J. Li, J. Koo, Microfluidics of nano-drug delivery, *Int. J. Heat Mass Tansf* 51 (2008) 5590-5597.
- [6] L. Capretto, W. Cheng, M. Hill, X. Zhang, Micromixing within microfluidic devices, *Top Curr Chem* 304 (2011) 27-68.
- [7] M. J. Uddin, W. A. Khan, A. I. Ismail, MHD free convective boundary layer flow of a nanofluid past a vertical plate with Newtonian heating boundary condition, *PLoS ONE* 7(11) (2012) e49499.
- [8] L. J. Crane, Flow past a stretching plate, *Z. Angrew. Math. Phys.* 21 (1970) 645-647.
- [9] W. A. Khan, I. Pop, Boundary layer flow of nanofluid past a stretching sheet, *Int. J. Heat Mass Transf* 53 (2010) 2477-2483.
- [10] M. Hassani, M. M. Tabar, J. Nemati, G. Domairy, F. Noori, An analytical solution for boundary layer flow of a nanofluid past a stretching sheet, *Int. J. Thermal Sci.* 50(11), (2011) 2256-2263.
- [11] O. D. Makinde, A. Aziz, Boundary layer flow of a nanofluid past a stretching sheet with a convective boundary condition, *Int. J. Thermal Sci.* 50 (2011) 1326-1332.
- [12] M. A. A. Hamad, M. Ferdows, Similarity solutions to viscous flow and heat transfer of nanofluid over nonlinear stretching sheet, *Appl. Math. Mech.* 33(7) (2012) 923-930.
- [13] F. G. Awad, P. Sibanda, A. A. Khidir, Thermodiffusion effect on magneto nanofluid flow over stretching sheet, *Boundary Value Problem* 2013:136.
- [14] M. Salari, M. M. Tabar, A. M. Tabar, Numerical solutions of heat transfer of nanofluid over a stretching sheet subjected to variations of nanoparticle volume fraction and wall temperature, *Appl. Math. Mech.* 35 (2014) 63-72.
- [15] W. Ibrahim, B. Shankar, MHD boundary layer flow and heat transfer of nanofluid past a permeable stretching with velocity, thermal and solutal slip boundary conditions, *Computers and fluids* 75 (2013) 1-10.
- [16] W. Ibrahim, B. Shankar, M. M. Nandeppanavar, MHD stagnation point flow and heat transfer due to nanofluid towards a stretching sheet, *Int. J. Heat Mass Transf.* 56 (2013) 1-9.
- [17] O. D. Manlinda, W. A. Khan, Z. H. Khan, Buoyancy effect on MHD stagnation point flow and heat transfer of a nanofluid past a convectively heated stretching/shrinking sheet, *Int. J. Heat Mass Transf* 62 (2013) 526-533.

- [18] M. Qasim, Z. H. Khan, W. A. Khan, I. A. Shah, MHD boundary layer slip flow and heat transfer of ferrofluid along a stretching cylinder with prescribed heat flux, PLoS ONE 9(1) (2014) e83930.
- [19] K. Bhattacharyya, G. Layek, Magnetohydrodynamic boundary layer flow of a nanofluid over an exponentially stretching permeable sheet, Physics Research International 2014 (2014) 592536.
- [20] J. C. Crepean, R. Clarksean, Similarity solutions of natural convection with internal heat generation, J. Heat Transf 119(1) (1997) 183-185.
- [21] A. Postelnicu, T. Grosan, I. Pop, Free convection boundary layer over a vertical permeable flat plate in a porous medium with internal heat generation, Int. Comm. Heat Mass Transfer 27(2000) 729-738.
- [22] E. M. Abo-Eldahab, M. A. El-Aziz, Blowing/suction effect on hydromagnetic heat transfer by mixed convection from an inclined continuously stretching surface with internal heat generation/absorption, Int. J. Thermal Sci. 43 (2004) 709-719.
- [23] M. S. Abel, P. G. Siddheshwar, M. M. Nandeppanavar, Heat transfer in a viscoelastic boundary layer flow over a stretching sheet with viscous dissipation and non uniform heat source, Int. J. Heat Mass Transf. 50 (2007) 960-966.
- [24] R. Bataller, Viscoelastic fluid flow and heat transfer over a stretching sheet under the effects of a non-uniform heat source viscous dissipation and thermal radiation, Int. J. Heat Mass. Transf. 50 (2007) 3152-3162.
- [25] R. Tsai, K. Huang, J. Huang, Flow and heat transfer over an unsteady stretching surface with non-uniform heat source, Int. Comm. Heat Mass Transfer 35 (2008) 1340-1343.
- [26] M. S. Abel, M. N. Mahantesh, Heat transfer in MHD viscoelastic boundary layer flow over a stretching sheet with non uniform heat source/sink, Comm. Nonlinear. Sci. Numer. Simulate. 14 (2009) 2120-2131.
- [27] G. K. Ramesh, B. J. Gireesha, C. S. Bagewadi, MHD flow of a dusty fluid near the stagnation point over a permeable stretching sheet with non-uniform source/sink, Int. J. Heat Mass Transfer 55 (2012) 4900-4907.
- [28] L. Zheng, N. Liu, X. Zhang, Maxwell fluids unsteady mixed flow and radiation heat transfer over a stretching permeable plate with boundary slip and non-uniform heat source/sink, J. Heat Transf. 135 (2013) 031705-1
- [29] S. S. Motsa, A new spectral relaxation method for similarity variable non linear boundary layer flow systems, Chem. Eng. Comm. 201 (2014) 241-256.

# **Parametric study of FRP-strengthened reinforced concrete panels under blast loads**

**†X. Lin<sup>1</sup> and \*Y.X. Zhang<sup>2</sup>**

<sup>1</sup>School of Civil and Environmental Engineering, Nanyang Technological University, Singapore.

<sup>2</sup>School of Engineering and Information Technology, the University of New South Wales, Australian Defence Force Academy, Australia

\*Presenting author: y.zhang@adfa.edu.au

†Corresponding author: xslin@ntu.edu.sg

## **Abstract**

The use of fibre reinforced polymer (FRP) as a strengthening material for reinforced concrete structures to resist blast loads has attracted great interests in recently years. The structural responses of FRP-strengthened reinforced concrete panels under blast loads are investigated intensively in this paper by employing a finite element model recently developed by the authors. The effects of the thickness of the strengthening FRP sheet, the retrofitted surface, the standoff distance and the mass of the charge on the structural behavior of the reinforced concrete panels under blast loads are studied. The results of the parametric study are analyzed and presented in this paper.

Keywords: FRP, Reinforced concrete panel, Blast loading, Finite element model

## **Introduction**

In recent years, fibre reinforced polymers (FRPs) have been increasingly used for retrofitting reinforced concrete (RC) structural components to against blast accidents which could produce an overload pressure much greater than the design load of a structure in a very short period of time and result in severe damage to the RC structure. The FRP attachments can significantly improve the blast resistance of structures without forfeiting usable space [Nam et al. (2010)]. FRP is also considered to be one of the most suitable materials for retrofitting concrete structures under blast loads, as it can be easily installed and naturally blended to the structures [Nam et al. (2009)].

A few experimental studies on the FRP-strengthened RC panels subjected to blast loads have been reported. However, full scale experimental tests are usually costly and time consuming. By contrast, finite element analysis is a much more economical and efficient method for predicting the behaviors of RC structures and structural components especially for the investigation of the parametric effects. Several numerical studies for the analysis of FRP-strengthened RC panels under blast loads [Nam et al. (2010); Nam et al. (2009); Mosalam and Mosallam (2001)] were conducted. However, the finite element models employed in these researches were either not very well validated or didn't consider the strain rate effect on the material properties of FRP appropriately. A finite element model, which considered the strain rate effect on FRP materials, was established by Tanapornraweekit et al. [Tanapornraweekit et al. (2010)] for modelling the structural behavior of GFRP and CFRP-strengthened reinforced concrete slabs under air blasts. Although progress has been made in the numerical modelling of FRP-strengthened RC structures under blast loading, the numerical models developed are still far from ideal due to the complexity of dynamic response of reinforced concrete structures and the lack of information on the dynamic material properties of FRPs. Moreover, a comprehensive study on the structural responses of FRP-strengthened RC panels under blast loads affected by various parameters has rarely been reported.

Recently, a 3D finite element model was developed by the authors for the analysis of the structural behavior of FRP strengthened RC panels under blast loading [Lin et al. (2015)]. In this model, strain rate effects on the material models of concrete under tension and compression were considered separately, and strain rate effects on the material models of steel reinforcements and FRP laminates were also taken into account. The proposed finite element model was demonstrated to be effective and accurate for the prediction of the structural behavior of FRP-strengthened RC panels under blast loads. In this paper, the 3D finite element model is firstly introduced, and then it is employed to investigate the effects of a series of parameters on the structural behaviors of FRP-strengthened RC panels under blast loads, including the effects of the thickness of FRP sheet, the retrofitted surface, the standoff distance and the mass of the charge. The research findings are reported in this paper, which are expected to be able to provide reliable and useful references for structural design.

### A 3D Finite Element Model

The 3D nonlinear finite element model was established using the commercial software package LS-DYNA. Concrete was modelled using the Solid 164 element, which is an 8-node constant stress hexahedron brick element, and the Lagrangian formulation was applied in the analysis. Steel reinforcing bars were modelled using the Link 160 truss element, and FRP laminates were modelled using the Shell 163 element, which is a 4-noded element with both bending and membrane capabilities. A perfect bond was assumed both between concrete and steel reinforcements and between FRP laminates and concrete panel.

#### *Material Model for Concrete*

The MAT72 R3 was employed for modelling the concrete. In the study conducted by Lin et al. [Lin et al. (2014)], various material models available for the analysis of concrete structures under dynamic loading were compared, and that the CONCRETE\_DAMAGE\_REL3 (MAT72 R3) in LS-DYNA was found to be relatively simple and numerically robust. It can reproduce the key concrete behaviors which are critical to blast and impact analyses, and it is also easy to be calibrated using laboratory data [Magallanes et al. (2010)].

In addition, concrete dynamic behavior is strain rate dependent. Both tensile and compressive strengths of concrete can be increased significantly under dynamic loading. In MAT72 R3, the strain rate effect is accounted for by using a dynamic increase factor (DIF), which is the ratio of the dynamic-to-static material strength. In the developed finite element model, the values of DIF suggested by CEB-FIP Model Code 1990 [CEB-FIP (1993)] were employed for concrete in compression, while the modified formulations proposed by Malvar and Crawford [Malvar and Crawford (1998)] were used for concrete in tension. The formulas of DIF for concrete in compression and tension are given in the following equations.

For concrete in compression:

$$\text{CDIF} = \frac{f_{c,d}}{f_{c,s}} = \left( \frac{\dot{\epsilon}_c}{\dot{\epsilon}_{c0}} \right)^{1.026\alpha} \quad \text{for} \quad \dot{\epsilon}_c \leq 30 \text{ s}^{-1} \quad (1a)$$

$$\text{CDIF} = \frac{f_{c,d}}{f_{c,s}} = \gamma \left( \frac{\dot{\epsilon}_c}{\dot{\epsilon}_{c0}} \right)^{1/3} \quad \text{for} \quad \dot{\epsilon}_c > 30 \text{ s}^{-1} \quad (1b)$$

where  $f_{c,d}$  is the dynamic compressive strength of concrete,  $f_{c,s}$  the static compressive strength of concrete,  $\dot{\epsilon}_c$  the strain rate in compression, and  $\dot{\epsilon}_{c0} = 30 \times 10^{-6} \text{ s}^{-1}$  the quasi-static strain rate in compression.  $\alpha$  is the coefficient given by  $\alpha = 1/(5 + 9f_{c,s}/10)$ , and  $\gamma$  is expressed as  $\log \gamma = 6.156\alpha - 2$ .



For concrete in tension:

$$\text{TDIF} = \frac{f_{ct,d}}{f_{ct,s}} = \left( \frac{\dot{\epsilon}_{ct}}{\dot{\epsilon}_{cto}} \right)^{\delta} \quad \text{for} \quad \dot{\epsilon}_{ct} \leq 1 \text{ s}^{-1} \quad (2a)$$

$$\text{TDIF} = \frac{f_{ct,d}}{f_{ct,s}} = \beta \left( \frac{\dot{\epsilon}_{ct}}{\dot{\epsilon}_{cto}} \right)^{1/3} \quad \text{for} \quad \dot{\epsilon}_{ct} > 1 \text{ s}^{-1} \quad (2b)$$

where  $f_{ct,d}$  is the dynamic tensile strength of concrete,  $f_{ct,s}$  the static tensile strength of concrete,  $\dot{\epsilon}_{ct}$  the strain rate in tension, and  $\dot{\epsilon}_{cto} = 1 \times 10^{-6} \text{ s}^{-1}$  the quasi-static strain rate in tension.  $\delta$  is the coefficient given by  $\delta = 1/(1 + 8f_{ct,s}/10)$ , and  $\beta$  is expressed as  $\log \beta = 6\delta - 2$ .

Due to the presence of strain-softening phenomenon in concrete, non-converged or incorrect converged solutions are usually obtained from finite element analysis, and the results are not objective with regard to mesh refinement [de Borst (1987)]. Therefore, a localization limiter must be introduced in concrete material model to remedy this situation. In the MAT72 R3, a crack band model is employed for this spurious mesh sensitivity caused by the strain-softening. In the single element tests carried out by Lin et al. [Lin et al. (2014)], the softening of small elements was found to be relatively slow. With the increase in element size, the softening was accelerated to maintain constant fracture energy. The stress-strain relationships for the elements with size between 1 mm and 25 mm were found to coincide very well. Therefore, the size of finite element mesh in the present model was chosen to be 15 mm.

#### Material Model for Steel

In the proposed finite element model, steel reinforcing bars were modelled using the PLASTIC\_KINEMATIC Model (MAT3). The isotropic and kinematic hardening can be specified by varying the hardening parameter between 0 and 1. The Cowper-Symonds model [Hallquist (2006)] was used to take into account the strain rate effect under blast loading, which scaled the yield stress by a strain-rate dependent factor of  $1 + (\dot{\epsilon}/C)^{1/P}$ , where  $\dot{\epsilon}$  is the strain rate, and **C** and **P** are the strain rate parameters for Cowper-Symonds model, which were 255.4 and 7.59 respectively in the proposed model.

#### Material Model for FRPs

The material properties of FRPs are also strain rate sensitive, and their modulus and strength increase with the increasing of the loading rate. So far, very few constitutive relationships of FRPs considering dynamic loading effect have been reported in the literature, and the use of an improper material model for FRPs under high strain rate might lead to inaccuracy in modelling the structural performance of FRP-strengthened RC structures subjected to blast loading. In the proposed finite element model, FRP laminates were modelled using the PLASTICITY\_POLYMER Model (MAT89), and the strain rate effect on FRP material properties were taken into account by employing the formulas used by Tanapornraweekit et al. [Tanapornraweekit et al. (2010)] for glass fibre reinforced polymer (GFRP) and carbon fibre reinforced polymer (CFRP) laminates. The formulas for GFRP laminates are given as follows.

Elastic modulus of GFRP (unit: GPa):

$$E_d = E_s \quad \text{for} \quad \dot{\epsilon} < 0.01 \text{ s}^{-1} \quad (3a)$$

$$E_d = E_s + 1.055 \log \dot{\epsilon} + 2.11 \quad \text{for} \quad 0.01 \text{ s}^{-1} \leq \dot{\epsilon} \leq 200 \text{ s}^{-1} \quad (3b)$$

$$E_d = E_s + 13.969 \log \dot{\epsilon} - 27.606 \quad \text{for} \quad 200 \text{ s}^{-1} < \dot{\epsilon} \leq 1700 \text{ s}^{-1} \quad (3c)$$

Tensile strength of GFRP (unit: GPa):

$$f_d = f_s \quad \text{for} \quad \dot{\epsilon} < 0.01 \text{ s}^{-1} \quad (4a)$$

$$f_d = f_s + 0.2797 \log \dot{\epsilon} + 0.5594 \quad \text{for} \quad 0.01 \text{ s}^{-1} \leq \dot{\epsilon} \leq 90 \text{ s}^{-1} \quad (4b)$$

$$f_d = f_s + 0.6696 \log \dot{\epsilon} - 0.2026 \quad \text{for} \quad 90 \text{ s}^{-1} < \dot{\epsilon} \leq 1700 \text{ s}^{-1} \quad (4c)$$

Failure strain of GFRP:

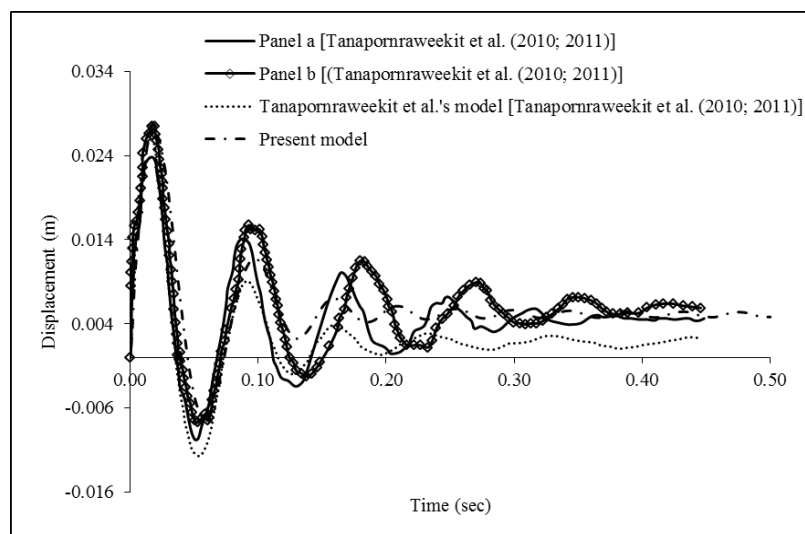
$$\epsilon_d = \epsilon_s \quad \text{for} \quad \dot{\epsilon} < 0.01 \text{ s}^{-1} \quad (5a)$$

$$\epsilon_d = \epsilon_s + 0.005 \log \dot{\epsilon} + 0.01 \quad \text{for} \quad 0.01 \text{ s}^{-1} \leq \dot{\epsilon} \leq 1700 \text{ s}^{-1} \quad (5b)$$

where  $E_d$ ,  $f_d$  and  $\epsilon_d$  are the dynamic elastic modulus, tensile strength and failure strain of GFRP respectively,  $E_s$ ,  $f_s$  and  $\epsilon_s$  the static elastic modulus, tensile strength and failure strain of GFRP respectively, and  $\dot{\epsilon}$  is the strain rate.

### Numerical Analysis

The developed finite element model has been validated by the authors [Lin et al. (2015)] by modelling two RC concrete panels strengthened with GFRP laminates. The two reinforced concrete panels (Panel b was the repeat test specimen of Panel a) had the same design and were supported on two short edges by a steel frame. They were reinforced on both top and bottom with 6 mm longitudinal steel bars spaced at 225 mm centre to centre, and 6 mm transverse steel bars at 300 mm centre to centre. In addition, 6 mm steel reinforcing bars were placed on both sides of the support zone at 65 mm centre to centre space. The yield strength, ultimate strength, ultimate strain and elastic modulus of steel reinforcing bars were 356 MPa, 412 MPa, 22.2% and 194 GPa, respectively. The compressive strength of concrete was 32 MPa. A 0.353 mm thick single layer GFRP sheet was attached to the top and bottom faces of the reinforced concrete panels. The Young's modulus and tensile strength of the GFRP were 75.6 GPa and 1331 MPa, respectively. The GFRP-strengthened RC panels were subjected to a blast load caused by a charge with an equivalent TNT mass of 0.45 kg, which was placed at 0.5 m above the centre of the concrete panel. The central displacement-time histories obtained from the developed finite element model, the tests and Tanapornraweeakit et al.'s numerical analysis [Tanapornraweeakit et al. (2010); Tanapornraweeakit et al. (2011)] are shown in Fig. 1.



**Figure 1. Central displacement-time histories of Panel a and Panel b**

It can be seen that the computed results agree very well with the test data, and they are closer to the test results than Tanapornraweeakit et al.'s prediction [Tanapornraweeakit et al. (2010)], especially

for the residual deflection. The maximum and residual deflections of the GFRP-strengthened reinforced concrete panel obtained from the present model are 27.2 mm and 5.3 mm respectively.

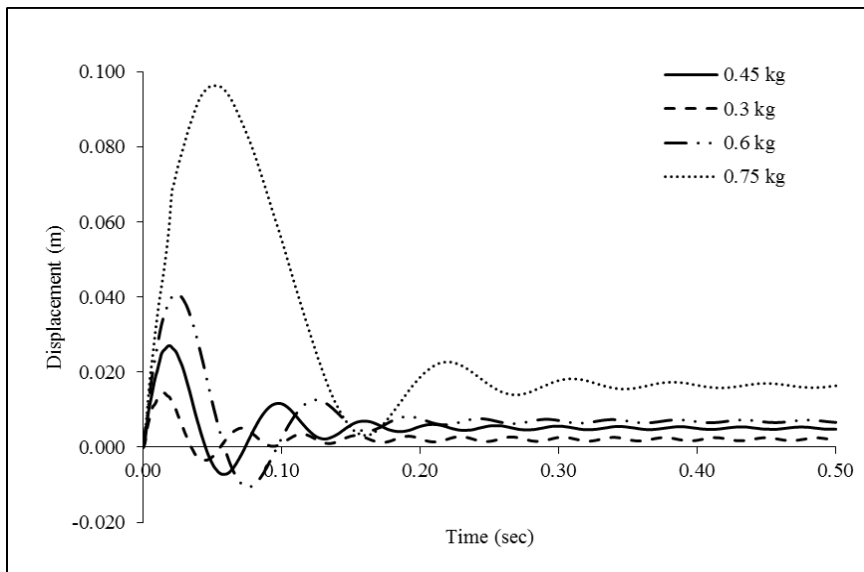
In this study, the developed finite element model is employed to further investigate the structural responses of FRP-strengthened RC panels affected by various parameters, including the effects of the thickness of the FRP sheet, the retrofitted surface, the standoff distance and the mass of the charge. The GFRP-strengthened RC panels (Panel a and Panel b) are used herein as a basic model, and the material properties of concrete, steel and GFRP sheet retain the same. In the parametric study, the TNT charge mass varies from 0.30 kg to 0.75 kg, and the standoff distance from 0.3 m to 1.0 m. In the basic model, the RC panel is strengthened by the GFRP sheets of the same thickness on both top and bottom surfaces of the panel. In order to investigate the FRP strengthening effect, the thickness of the FRP sheets on the top and bottom surfaces are changing from 0 to 1.0 mm and from 0.353 mm to 1.0 mm respectively, thus the modelled RC panels either have thicker GFRP sheet strengthened on the bottom surface or have GFRP sheet of the same thickness on both top and bottom surfaces. The details of various parameters and the predicted maximum and residual deflections of the concrete panels are listed in Tabel 1. Figs. 2 to 4 show the central displacement-time histories of GFRP-strengthened RC panels with various parameters.

As can be seen, with the increase of TNT charge mass, both the maximum and residual deflections of the GFRP-strengthened RC panels increase significantly. The maximum deflection of the panel under the blast of a charge mass of 0.30 kg TNT is 53.3%, 35.4% and 15.1% of that of the 0.45 kg, 0.60 kg and 0.75 kg TNT respectively. The residual deflection of the panel subjected to 0.30 kg TNT is 2.2 mm, which is 41.5%, 31.9% and 13.7% of that under 0.45 kg, 0.60 kg and 0.75 kg TNT blast loads respectively. For the same TNT charge mass, the maximum and residual deflections of GFRP-strengthened RC panels decrease with the increase of the standoff distance. The maximum and residual deflections for the panel with a standoff distance of 1.0 m are 9.5 mm and 2.7 mm respectively, which are only 34.9% and 50.9% respectively of those obtained for the panel with a standoff distance of 0.5 m. Both the maximum and residual deflections increase when the standoff distance reduces from 0.5 m to 0.4 m, but not by much. Whereas, when the standoff distance is further reduced from 0.4 m to 0.3 m, the maximum and residual deflections of panel suddenly jump to 126.8 mm and 48.5 mm respectively, which are about 4 times and 10 times of those for the panel at 0.4 m away from the charge.

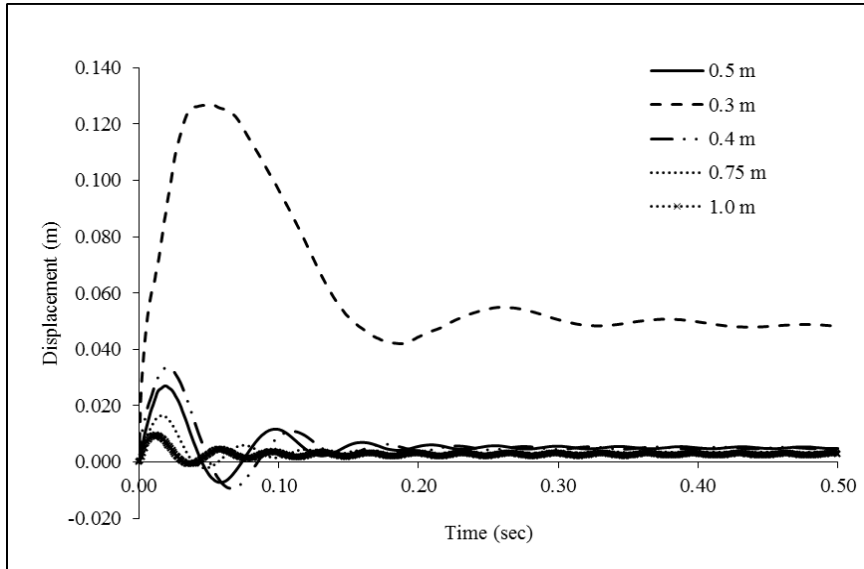
In addition, the thickness of the FRP strengthening sheet also affects the blast resistance of RC panels. In general, the maximum and residual deflections of the RC panels are reduced with the increase of the thicknesses of the FRP strengthening sheets. The maximum deflection of the RC panel with 1.0 mm GFRP sheets strengthened on both surfaces is 19.7 mm, which is 72.4% of that for the basic model with 0.353 mm GFRP sheet. The residual deflection is reduced from 5.3 mm to 1.8 mm which is only 34.0% of the deflection in the basic model. Also, the RC panels strengthened with GFRP sheets of different thicknesses on the top and bottom surfaces are modelled in this study. By comparing the basic model with the panel without GFRP sheet strengthened on the top surface, it is found that the maximum deflection of the basic model is reduced by 4.2%, whereas the residual deflection is increased from 0 to 5.3 mm. When the thickness of the GFRP sheet on the top surface is kept constant as 0.353 mm, and that on the bottom surface is increased, the computed maximum and residual deflections are always a bit less than those obtained for the panel which has the same GFRP thickness on the top and bottom surfaces. However, it should be noted that the deflection when the panel bounces back is obviously reduced by strengthening the GFRP sheets with the same thicknesses on both surfaces, as can be seen in Fig. 4. Therefore, the damage on the top surface of the RC panels could be reduced by increasing the thickness of GFRP sheets on the top surface.

**Table 1. Maximum and residual deflections affected by various parameters**

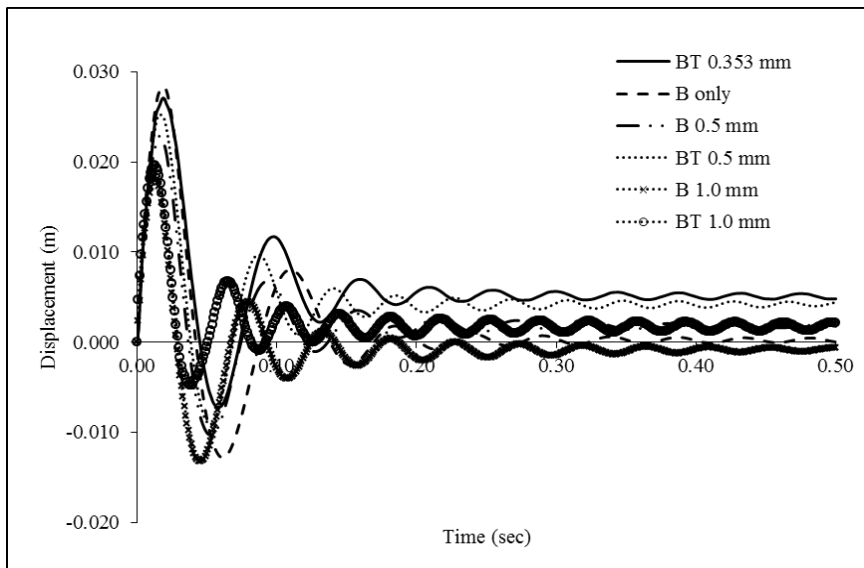
TNT Charge mass (kg)	Standoff distance (m)	GFRP thickness (mm)		Maximum deflection (mm)	Residual deflection (mm)
		Bottom (B)	Top (T)		
0.30	0.5	0.353	0.353	14.5	2.2
0.45	0.5	0.353	0.353	27.2	5.3
0.60	0.5	0.353	0.353	41.0	6.9
0.75	0.5	0.353	0.353	96.3	16.1
0.45	0.3	0.353	0.353	126.8	48.5
0.45	0.4	0.353	0.353	33.6	5.0
0.45	0.5	0.353	0.353	27.2	5.3
0.45	0.75	0.353	0.353	16.8	3.4
0.45	1.0	0.353	0.353	9.5	2.7
0.45	0.5	0.353	0	28.4	0
0.45	0.5	0.353	0.353	27.2	5.3
0.45	0.5	0.5	0.353	23.0	1.8
0.45	0.5	0.5	0.5	25.8	4.3
0.45	0.5	1.0	0.353	18.4	-0.7
0.45	0.5	1.0	1.0	19.7	1.8



**Figure 2. Displacement-time histories of GFRP-strengthened RC panels under various TNT charge masses**



**Figure 3. Displacement-time histories of GFRP-strengthened RC panels at various standoff distances**



**Figure 4. Displacement-time histories of RC panels with various strengthening laminates under blast loads**

### Conclusions

In this paper, a finite element model recently developed for modelling structural behavior of FRP-strengthened RC panels under blast loads is firstly introduced, and then employed to investigate the responses of FRP-strengthened RC panels with various parameters under blast loads. It is found that the structural behavior of FRP-strengthened RC panel is very sensitive to the charge mass and the standoff distance. Care must be taken in the structural design, especially when increasing the charge mass or reducing the standoff distance. In addition, the effect of thickness of GFRP sheet on different retrofitting surface is studied. In general, with the increase in the thickness of the GFRP sheet, both the maximum and residual deflections of RC panel are decreased. Although no improvement in reducing the maximum and residual deflections could be observed by using thicker GFRP sheet on the top surface, the deflection when the RC panel bounces back could be reduced greatly.

## References

- Nam, J.-W., Kim, H.-J., Kim, S.-B., Yi, N.-H. and Kim, J.-H. J. (2010) Numerical evaluation of the retrofit effectiveness for GFRP retrofitted concrete slab subjected to blast pressure, *Composite Structures* **92**, 1212-1222.
- Nam, J.-W., Kim, H.-J., Kim, S.-B., Kim, J.-H. J. and Byun, K. J. (2009) Analytical study of finite element models for FRP retrofitted concrete structure under blast loads, *International Journal of Damage Mechanics* **18**, 461-490.
- Mosalam, K. M. and Mosallam, A. S. (2001) Nonlinear transient analysis of reinforced concrete slabs subjected to blast loading and retrofitted with CFRP composites, *Composites Part B: Engineering* **32**, 623-636.
- Tanapornraweekit, G., Haritos, N., Mendis, P. and Ngo, T. (2010) Finite element simulation of FRP strengthened reinforced concrete slabs under two independent air blasts, *International Journal of Protective Structures* **1**, 469-488.
- Lin, X., Zhang, Y. X. and Lee, C. K. (2015) Finite element analysis of FRP-strengthened reinforced concrete panels under blast loading, *5th International Conference on Design and Analysis of Protective Structures*, 19-21 May 2015, Singapore (submitted).
- Lin, X., Zhang, Y. X. and Hazell, P. J. (2014) Modelling the response of reinforced concrete panels under blast loading, *Materials & Design* **56**, 620-628.
- Magallanes, J. M., Wu, Y., Malvar, L. J. and Crawford, J. E. (2010) Recent improvements to Release III of the K&C concrete model, *11th International LS-DYNA Users Conference*.
- Comite Euro-International du Beton (1993) *CEB-FIP model code 1990*, Lausanne: Thomas Telford Services Ltd.
- Malvar, L. J. and Crawford, J. E. (1998) Dynamic increase factors for concrete, Naval Facilities Engineering Service Center, Port Hueneme, CA.
- de Borst, R. (1987) Computation of post-bifurcation and post-failure behavior of strain-softening solids, *Computers & Structures* **25**, 211-224.
- Hallquist, J. O. (2006) *LS-DYNA theory manual*, Livermore Software Technology Corporation (LSTC), Livermore, CA.
- Tanapornraweekit, G., Haritos, N. and Mendis, P. (2011) Behavior of FRP-RC slabs under multiple independent air blasts, *Journal of Performance of Constructed Facilities* **25**, 433-440.

# Frequency Domain Elastodynamic Solutions Using Iterative Coupling of FEM and BEM

Delfim Soares Jr.<sup>1\*</sup>, Kleber A. Gonçalves<sup>2a</sup>, and José Claudio de Faria Telles<sup>2b</sup>

<sup>1</sup>Structural Engineering Department, Federal University of Juiz de Fora, CEP 36036-330, Juiz de Fora, Brazil.

<sup>2</sup>Civil Engineering Programme, COPPE - Federal University of Rio de Janeiro, CEP 21945-970, Rio de Janeiro, Brazil.

\* delfim.soares@ufjf.edu.br

<sup>a</sup> kleber.a.g@hotmail.com

<sup>b</sup> telles@coc.ufrj.br

## Abstract

The paper introduces a coupling FEM-BEM procedure for solving elastodynamic frequency domain problems. Emphasis is given to infinite domain analyses, including discrete complex heterogeneous regions, rendering a configuration in which neither the Finite Element Method (FEM) nor the Boundary Element Method (BEM), isolated, is ideally suited for the complete numerical analysis. In this case, the coupling of these methodologies is recommended, allowing for the exploration of their respective intrinsic advantages. The elastodynamic multi domain interaction is carried out here by an optimized iterative coupling procedure. This coupling technique allows for independent discretization strategies, not even needing to include matching interface nodes between methods, leading to a best of both worlds approach. In addition, optimal relaxation parameters are computed, in order to improve convergence of the iterative procedure, properly dealing with possible frequency domain ill-posed problems.

**Keywords:** Elastodynamics; frequency domain; iterative coupling; relaxation parameter

## Introduction

The numerical simulation of arbitrarily shaped continuous bodies subjected to harmonic or transient loads remains, despite much effort and progress over the last decades, a challenging area of research. In most cases, discrete techniques, such as the finite element method (FEM) and the boundary element method (BEM) have been employed and continuously further developed with respect to accuracy and efficiency. Both methodologies can be formulated in the time domain or in the frequency domain, and each approach has relative benefits and limitations. The finite element method, for instance, is well suited for inhomogeneous and anisotropic materials as well as for dealing with the nonlinear behaviour of a body. For systems with infinite extension and regions of high stress concentration, however, the use of the boundary element method is by far more advantageous.

In fact, it did not take long until some researchers started to combine the FEM and the BEM in order to profit from their respective advantages, trying to evade their disadvantages, and nowadays several works dealing with FEM-BEM coupling are available (an overview is provided by [Beskos (2003)], taking into account dynamic analyses). However, Standard coupling procedures of FEM/BEM can lead to several problems with respect to efficiency, accuracy and flexibility. First, the coupled system of equations has a banded symmetric structure only in the FEM part, while in the BEM part it is non-symmetric and fully populated. Consequently, for its solution, the optimized solvers usually used by the FEM cannot be employed anymore, which leads to rather expensive calculations with respect to computer time. Second, quite different physical properties may be

involved in the coupled model, resulting in bad-conditioned matrices when standard coupling procedures are considered. This may affect the accuracy of the methodology, providing misleading results. Third, the standard coupling methodology does not allow independent discretization for each sub-domain of the model, requiring matching nodes at common interfaces, which drastically affects the flexibility and versatility of the technique.

In order to evade these drawbacks, iterative coupling procedures have been developed. Initially, static problems were studied considering iterative coupling approaches, and linear and nonlinear behaviour have been simulated [Lin et al (1996)], [Elleithy et al (2001; 2009; 2012)], [Jahromi et al (2009)], [Boumaiza and Aour (2014)]. Later on, dynamic problems were focused, and time domain analyses were initially implemented [Soares et al (2004)], [Soares (2008; 2012)]. Recently, frequency domain iterative analyses have also been considered; but, in this case, most works are related to fluid-fluid or fluid-structure coupled models [Bendali et al (2007)], [Soares and Godinho (2012)], [Godinho and Soares (2013)]. For an overview of recent advances in the iterative analysis of coupled models considering time and frequency domain approaches, the work of [Soares and Godinho (2014)] is recommended.

Iterative coupling approaches allow BEM and FEM sub-domains to be analyzed separately, leading to smaller and better-conditioned systems of equations (different solvers, suitable for each sub-domain, may be employed). Moreover, a small number of iterations is required for the algorithm to converge and the matrices related to the smaller governing systems of equations do not need to be treated (inverted, triangularized etc.) at each iterative step, providing an efficient methodology. This coupling technique allows independent discretizations to be efficiently employed for the boundary and finite element sub-domains, without any requirement of matching nodes at the common interfaces. As a matter of fact, in the present work, constant boundary elements and linear finite elements are considered, and matching functional nodes are never provided in the common interfaces. It is important to observe, however, that frequency domain analyses usually give rise to ill-posed problems and, in these cases, the convergence of the iterative coupling algorithm can be either too slow or unachievable if no special procedure is taken into account. In order to deal with this ill-posed problem and ensure convergence of the iterative coupling algorithm, an optimal iterative procedure is adopted here, with optimal relaxation parameters being computed at each iterative step. Thus, an expression to compute optimal relaxation parameters, which is quite efficient and easy to implement, is provided and discussed, being its effectiveness illustrated at the end of the paper, where numerical examples are analyzed. In the numerical examples, soil-structure interacting models are discussed, being the results of the proposed iterative coupling formulation compared to those of the standard coupling technique. As one will observe, the proposed technique is flexible, robust and efficient, allowing a quite effective coupling of the finite element and boundary element methods for frequency domain elastodynamic analyses.

## Governing Equations

The frequency domain elastic wave equation for homogenous media is given by:

$$\rho(c_d^2 - c_s^2)u_{j,j} + \rho c_s^2 u_{i,jj} + (\omega^2 \rho - i\omega\nu)u_i + b_i(X, \omega) = 0 \quad (1)$$

where  $u_i(X, \omega)$  and  $b_i(X, \omega)$  stand for the displacement and the body force distribution components, respectively. In Eq. (1),  $c_d$  is the dilatational wave velocity and  $c_s$  is the shear wave velocity, they are given by:  $c_d^2 = (\lambda + 2\mu)/\rho$  and  $c_s^2 = \mu/\rho$ , where  $\rho$  is the mass density and  $\lambda$  and  $\mu$  are the Lamé's constants.  $\nu$  stands for viscous damping related parameters. Eq. (1) can be obtained from the combination of the following basic mechanical equations (proper to model heterogeneous media):

$$\sigma_{ij,j} + (\rho(X)\omega^2 - i\omega\nu(X))u_i + b_i(X, \omega) = 0 \quad (2a)$$



$$\sigma_{ij}(X, \omega) = \lambda(X)\delta_{ij}\varepsilon_{kk}(X, \omega) + 2\mu(X)\varepsilon_{ij}(X, \omega) \quad (2b)$$

$$\varepsilon_{ij}(X, \omega) = (1/2)(u_i(X, \omega)_{,j} + u_j(X, \omega)_{,i}) \quad (2c)$$

where  $\sigma_{ij}(X, \omega)$  and  $\varepsilon_{ij}(X, \omega)$  are, respectively, stress and strain tensor components, and  $\delta_{ij}$  is the Kronecker delta ( $\delta_{ij} = 1$ , for  $i = j$  and  $\delta_{ij} = 0$ , for  $i \neq j$ ). Eq. (2a) is the momentum equilibrium equation; Eq. (2b) represents the constitutive law of the linear elastic model and Eq. (2c) stands for kinematical relations. The boundary conditions of the elastodynamic problem are given by:

$$u_i(X, \omega) = \bar{u}_i(X, \omega) \text{ for } X \in \Gamma_1 \quad (3a)$$

$$\tau_i(X, \omega) = \sigma_{ij}(X, \omega)n_j(X) = \bar{\tau}_i(X, \omega) \text{ for } X \in \Gamma_2 \quad (3b)$$

where the prescribed values are indicated by over bars,  $\tau_i(X, \omega)$  denotes the traction vector along the boundary and  $n_j(X)$  stands for the components of the unit outward normal vector.

### Boundary Element Modelling

The BEM integral equation related to the elastodynamic model is given by:

$$c_{ij}(\xi)u_j(\xi, \omega) = \int_{\Gamma} u_{ij}^*(X; \xi, \omega)\tau_j(X, \omega) d\Gamma - \int_{\Gamma} \tau_{ij}^*(X; \xi, \omega) u_j(X) d\Gamma + \zeta_i(X; \xi, \omega) \quad (4)$$

where  $c_{ij}(\xi)$  depends on geometric aspects,  $\zeta_i(X; \xi, \omega)$  stands for possible domain integral contributions (such as body sources) and the terms  $u_{ij}^*(X; \xi, \omega)$  and  $\tau_{ij}^*(X; \xi, \omega)$  represent the fundamental displacement and traction, respectively ( $X$  is the field point and  $\xi$  is the source point). For a two-dimensional approach, the fundamental solutions can be found at [Dominguez (1993)]. By introducing spatial approximations for the variables of the model into the integral Eq. (4), the following system of equations can be obtained, once proper numerical treatment is considered [Dominguez (1993)]:

$$\mathbf{C}\mathbf{U}(\omega) = \mathbf{G}(\omega)\mathbf{T}(\omega) - \mathbf{H}(\omega)\mathbf{U}(\omega) + \mathbf{S}(\omega) \quad (5)$$

where  $\mathbf{C}$ ,  $\mathbf{G}$  and  $\mathbf{H}$  are influence matrices,  $\mathbf{S}$  is a vector related to domain integrals and  $\mathbf{U}$  and  $\mathbf{T}$  are displacement and traction vectors, respectively, at frequency  $\omega$ . After considering the boundary conditions of the problem (translating all the known variables to the right-hand-side of Eq. (5), and the unknown fields to the left-hand-side), the BEM responses for the elastic model can be computed for the given frequency  $\omega$ .

### Finite Element Modelling

The integral weak-form of the governing equations at section 2 can be written as:

$$\begin{aligned} & -\omega^2 \int_{\Omega} \rho(X)u_i(X, \omega)w_{ik}(X)d\Omega + i\omega \int_{\Omega} \nu(X)u_i(X, \omega)w_{ik}(X)d\Omega + \\ & + \int_{\Omega} \sigma_{ij}(X, \omega)w_{ik}(X)_{,j}d\Omega + \int_{\Omega} b_i(X, \omega)w_{ik}(X)d\Omega - \int_{\Gamma_2} \tau_i(X, \omega)w_{ik}(X)d\Gamma = 0 \end{aligned} \quad (6)$$

where  $w_{ik}(X)$  stands for a weight function, which is assumed to have null values in the essential boundary (i.e.,  $w_{ik}(X) = 0$  for  $X \in \Gamma_1$ ).

By introducing spatial approximations for the variables of the model into the integral Eq. (6), and by adopting these approximations to define the specified weight functions (Galerkin Method), the following system of equations can be obtained, once proper numerical treatment is considered [Bathe (1996)], [Hughes (2000)]:

$$-\omega^2 \mathbf{M}\mathbf{U}(\omega) + i\omega \mathbf{C}\mathbf{U}(\omega) + \mathbf{K}\mathbf{U}(\omega) = \mathbf{F}(\omega) \quad (7)$$

where  $\mathbf{M}$ ,  $\mathbf{C}$  and  $\mathbf{K}$  stand for the mass, damping and stiffness matrix of the model, respectively, and  $\mathbf{U}$  and  $\mathbf{F}$  stand for the nodal displacement and force vector, respectively. Matrices  $\mathbf{M}$ ,  $\mathbf{C}$  and  $\mathbf{K}$  are computed taking into account the first, second and third terms in Eq. (6), respectively, whereas vector  $\mathbf{F}$  is computed taking into account the last two terms in the l.h.s. of Eq. (6) (for the stiffness matrix computation, Eqs. (2b)-(2c) are employed to relate the stress tensor with the displacement vector). After considering the boundary conditions of the problem, the FEM responses for the elastodynamic model can be computed for the given frequency  $\omega$ , taking into account Eq. (7).

### Coupling Procedures

In order to enable the coupling between the BEM and the FEM sub-domains of the model, an iterative procedure is employed here, which performs a successive renewal of the relevant variables at the common interfaces. The proposed approach is based on the imposition of prescribed displacement at the BEM sub-domain and of prescribed nodal forces at the FEM sub-domain. Since the two sub-domains are analysed separately, the relevant systems of equations are formed independently, before the iterative process starts, and are kept constant for each frequency along the iterative process. The separate treatment of the two sub-domains allows independent discretizations to be used on both parts, without any special requirement of matching nodes along the common interfaces. Thus, the coupling algorithm can be presented for a generic case, in which the interface nodes may not match, allowing exploiting this benefit of the iterative coupling formulation.

To ensure and/or to speed up convergence, a relaxation parameter  $\lambda$  is introduced in the iterative coupling algorithm. The effectiveness of the iterative process is strongly related to the selection of this relaxation parameter, since an inappropriate selection for  $\lambda$  can significantly increase the number of iterations in the analysis or, even worse, make convergence unfeasible. At the end of the section, an optimal relaxation parameter is calculated, taking into account the coupled BEM-FEM frequency-domain formulation.

#### *Iterative coupling procedures*

Initially, in the  $k^{th}$  iterative step of the FEM-BEM coupling, the FEM sub-domain is analysed and the structure displacements at the common interfaces  ${}_f\mathbf{U}_I^{(k)}(\omega)$  (subscript  $I$  indicates the common interface, whereas  $f$  and  $b$  indicates finite and boundary element sub-domains, respectively) are computed, as described in section 4. In this case,  ${}_f\mathbf{U}_I^{(k)}(\omega)$  is evaluated taking into account prescribed nodal forces at the common interfaces  ${}_f\mathbf{F}_I^{(k)}$ , which are provided from the previous iterative step (in the first iterative step, null prescribed nodal forces are considered). Once  ${}_f\mathbf{U}_I^{(k)}(\omega)$  is computed, it is applied to evaluate the essential boundary conditions that are prescribed at the common interfaces of the BEM sub-domains. More precisely,  ${}_f\mathbf{U}_I^{(k)}(\omega)$  is used to compute BEM displacements, as indicated below:

$${}_b\mathbf{U}_I^{(k+\lambda)}(\omega) = \int_{\Gamma_I} \delta^T(X - {}_bX) {}_f\mathbf{N}(X) d\Gamma {}_f\mathbf{U}_I^{(k)}(\omega) \quad (8)$$

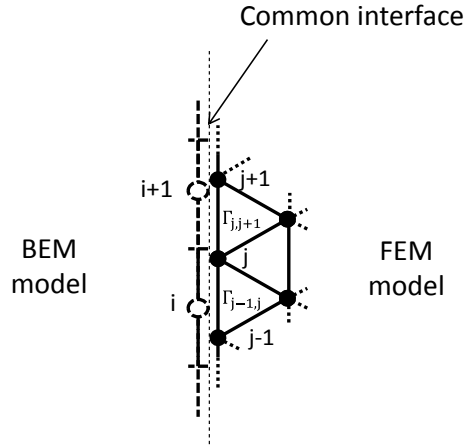
where  $\delta$  stands for a matrix representation of the Dirac's Delta function, employed here just to properly indicate the computation of the variables at the BEM nodes  ${}_bX$ , and  $N(X)$  stands for the BEM or FEM interpolation functions, according to the subscript  $b$  or  $f$ , respectively.

To better describe the proposed FEM-BEM coupling methodology, Figure 1 illustrates its application for the case of constant boundary elements and linear triangular finite elements.

As previously discussed, in this work, relaxation parameters are considered in order to ensure and/or to speed up the convergence of the iterative process. Thus, the displacements  ${}_bU_I^{(k+\lambda)}$  that are calculated by Eq. (8) are actualized as follow:

$$-\omega^2 \mathbf{M}U(w) + i\omega \mathbf{C}U(\omega) + \mathbf{K}U(\omega) = \mathbf{F}(\omega) \quad (9)$$

where  $\lambda$  stands for the relaxation parameter.



**Figure 1. Detail of a portion of the FEM-BEM interface when linear triangular finite elements and constant boundary are used. In the figure,  $j - 1$ ,  $j$  and  $j + 1$  are FEM interface nodes, while  $i$  and  $i + 1$  are BEM nodes. Displacements at BEM node  $i$  can be computed by interpolation of FEM displacements at nodes  $j - 1$  and  $j$  (Eq. (8)); FEM nodal force in  $j$  can be calculated by integration of the traction along boundaries  $\Gamma_{j-1,j}$  and  $\Gamma_{j,j+1}$ , using Eq. (10) and considering FEM linear and BEM piecewise constant shape functions along these boundaries.**

Once the BEM displacements at the common interfaces are computed, the BEM sub-domains can be analyzed, as described in section 3. As a consequence, the BEM tractions at the common interfaces are evaluated  ${}_bT_I^{(k+1)}$ , allowing the computation of the natural boundary conditions that are prescribed at the FEM sub-domains at the next iterative step. This is carried out as indicated below:

$${}_fF_I^{(k+1)}(\omega) = \int_{\Gamma_I} {}_fN^T(X) {}_bN(X) d\Gamma {}_bT_I^{(k+1)}(\omega) \quad (10)$$

Once  ${}_fF_I^{(k+1)}(\omega)$  is computed, the algorithm goes on to the next iterative step, repeating all the above described procedures, until convergence is achieved.

As it is illustrated in section 6, a proper selection for  $\lambda$  at each iterative step is extremely important for the effectiveness of the iterative coupling procedure. In order to obtain an easy to implement, efficient and effective expression for the relaxation parameter computation, in the next sub-section optimal  $\lambda$  values are deduced.

### Optimal relaxation parameter

In order to evaluate an optimal relaxation parameter, the following square error functional is minimized here:

$$f(\lambda) = \left\| {}_b\mathbf{U}_I^{(k+1)}(\lambda) - {}_b\mathbf{U}_I^{(k)}(\lambda) \right\|^2 \quad (11)$$

where  ${}_b\mathbf{U}_I$  stands for the BEM prescribed values at the common interfaces.

Taking into account the relaxation of the prescribed values for the (k+1) and (k) iterations, Eq. (12a) and Eq. (12b) may be written, based on the definition in Eq. (9):

$${}_b\mathbf{U}_I^{(k+1)} = (\lambda) {}_b\mathbf{U}_I^{(k+\lambda)} + (1 - \lambda) {}_b\mathbf{U}_I^{(k)} \quad (12a)$$

$${}_b\mathbf{U}_I^{(k)} = (\lambda) {}_b\mathbf{U}_I^{(k+\lambda-1)} + (1 - \lambda) {}_b\mathbf{U}_I^{(k-1)} \quad (12b)$$

Substituting Eqs. (12) into Eq. (11) yields:

$$\begin{aligned} f(\lambda) &= \left\| (\lambda)\mathbf{W}^{(k+\lambda)} + (1 - \lambda)\mathbf{W}^{(k)} \right\|^2 = \\ &= (\lambda^2) \left\| \mathbf{W}^{(k+\lambda)} \right\|^2 + 2\lambda(1 - \lambda)(\mathbf{W}^{(k+\lambda)}, \mathbf{W}^{(k)}) + (1 - \lambda)^2 \left\| \mathbf{W}^{(k)} \right\|^2 \end{aligned} \quad (13)$$

where the inner product definition is employed (e.g.,  $(\mathbf{W}, \mathbf{W}) = \|\mathbf{W}\|^2$ ) and new variables, as defined in Eq. (14), are considered.

$$\mathbf{W}^{(k+\lambda)} = {}_b\mathbf{U}_I^{(k+\lambda)} - {}_b\mathbf{U}_I^{(k+\lambda-1)} \quad (14)$$

To find the optimal  $\lambda$  that minimizes the functional  $f(\lambda)$ , Eq. (13) is differentiated with respect to  $\lambda$  and the result is set to zero, as described below:

$$\lambda \left\| \mathbf{W}^{(k+\lambda)} \right\|^2 + (1 - 2\lambda) (\mathbf{W}^{(k+\lambda)}, \mathbf{W}^{(k)}) + (\lambda - 1) \left\| \mathbf{W}^{(k)} \right\|^2 = 0 \quad (15)$$

Re-arranging the terms in Eq. (15), yields:

$$\lambda = \frac{(\mathbf{W}^{(k)}, \mathbf{W}^{(k)} - \mathbf{W}^{(k+\lambda)})}{\left\| \mathbf{W}^{(k)} - \mathbf{W}^{(k+\lambda)} \right\|^2} \quad (16)$$

which is an easy to implement expression that provides an optimal value for the relaxation parameter  $\lambda$ , at each iterative step. This expression requires a low computational cost, when compared to other alternatives that can be found in the literature (see, for instance, [Elleithy et al (2001)]).

Additionally, one should keep in mind that the computed relaxation parameter is a complex number, since the problem is formulated in the frequency domain. This complex number computation could be ranged (e.g., imposing  $|\lambda| \leq 1$ ), but the authors have observed that faster convergence is usually achieved in the iterative process if a non-restricted relaxation parameter selection, provided by Eq. (16), is considered. Moreover, although the authors found that the iterative process is relatively insensitive to the value of the relaxation parameter used for the first step, in all the cases discussed here, a real value of  $\lambda = 0.5$  is considered.

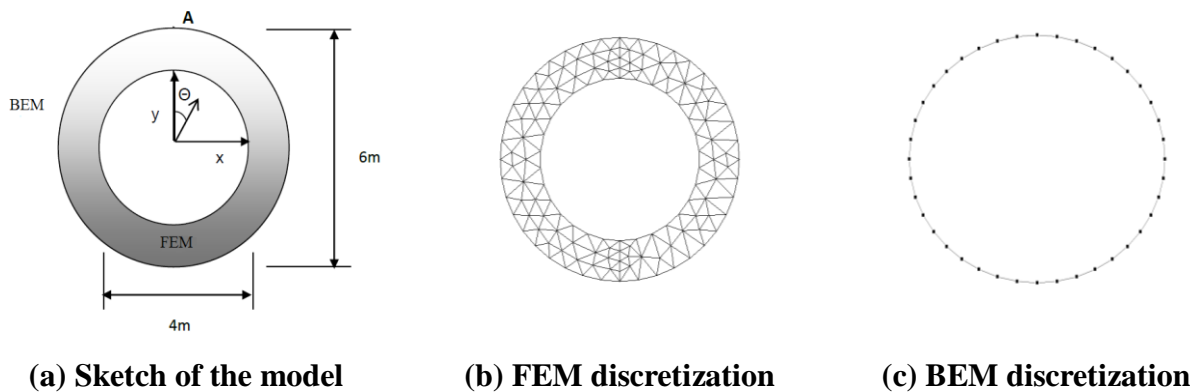
### Numerical Analysis

In order to illustrate the performance and potentialities of the discussed techniques, two application examples are considered here, corresponding to a circular ring-shaped structure involved by an

infinite soil domain. Different material properties, as well as prescribed load/displacement configurations, are considered in the analyses.

*Ring-shaped structure inside an infinite elastic domain*

Consider a circular homogeneous ring-shaped elastic inclusion, inside a homogeneous and infinite elastic environment (see Fig. 2a). The external environment has a density of  $7,85 \times 10^3 \text{ kg/m}^3$ , Young's modulus of  $20,58 \times 10^{10} \text{ N/m}^2$  and Poisson's ratio of  $0.2$  (no damping is considered). This elastic material allows dilatational and shear waves to travel at  $5397,17 \text{ m/s}$  and  $3305,08 \text{ m/s}$ , respectively. The circular inclusion has an external radius of  $3.0 \text{ m}$  and an internal radius of  $2.0 \text{ m}$  and is made of the same elastic material of the external domain.

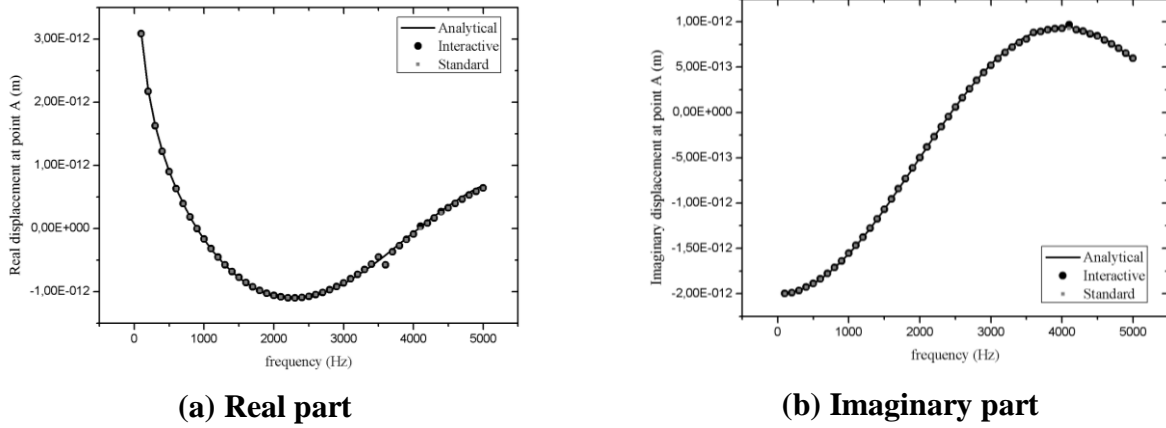


**Figure 2. Model's sketch and discretization**

The external environment is discretized by boundary elements distributed uniformly along the common interface (straight boundary elements with constant interpolation functions are adopted); the ring structure is modelled by using linear triangular finite elements. Fundamental harmonic displacements are prescribed at the internal cavity of the ring structure, which are acquired by considering a horizontal Dirac's delta force acting at the centre of the cavity. Thus, the analytical solution for the problem is known and it is provided by the model's fundamental solutions.

First, the external environment is modelled using 40 boundary elements, while a total of 210 elements (40 nodes at the interface) are considered at the finite element mesh. The corresponding FEM and BEM discretizations are illustrated in Fig. 2b and Fig. 2c, respectively.

Fig. 3 illustrates the displacements computed at point A (see Fig. 2a), taking into account the proposed iterative coupling procedure, considering a frequency range from  $100$  to  $5000 \text{ Hz}$ . Analytical answers and results computed taking into account a standard FEM-BEM direct coupling methodology are also depicted in Fig. 3, for comparison. As one can observe, the results provided by these different approaches are in good agreement. It is important to highlight that the coupled FEM-BEM results get closer to the analytical answers as the discretization of the model is refined.



**Figure 3. Vertical displacements at point A**

As a matter of fact, the convergence of the proposed technique is analyzed next, taking into account independent discretizations (and, as a consequence, no matching nodes at the common interface) for the FEM and the BEM. In order to do so, 4 discretizations for the BEM sub-domain and 4 discretizations for the FEM sub-domain are focused, as described in Table 1 (as one may observe, meshes 2 are those depicted in Fig. 2). These different discretizations are combined among each other and the errors that arise (taking into account the analytical answer of the model) are depicted in Fig. 4. Three combinations are considered here, the first one considers the FEM mesh 4 (i.e., 160 nodes on the FEM common interface) combined with all the focused BEM meshes. This combination is referred here as "FEM 160 - BEM". The second combination considers the BEM mesh 4 (i.e., 160 nodes on the BEM common interface) combined with all the focused FEM meshes. This combination is referred here as "BEM 160 - FEM". Finally, standard node-to-node combinations (i.e., considering matching geometrical nodes at the common interface) of the BEM and FEM meshes are also considered, and this combination is referred here as "node - node".

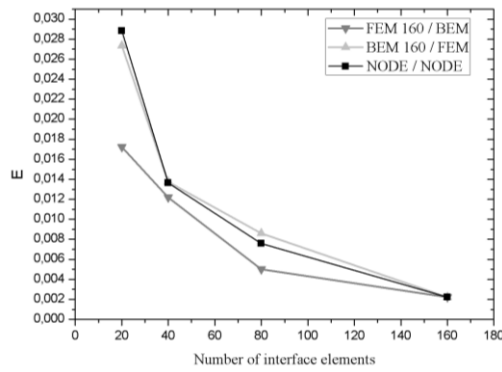
**Table 1: Discretizations for the BEM and FEM sub-domains.**

BEM straight constant elements	FEM triangular linear elements
Mesh 1: 20 elements	Mesh 1: 162 elements (20 elements at the interface)
Mesh 2: 40 elements	Mesh 2: 210 elements (40 elements at the interface)
Mesh 3: 80 elements	Mesh 3: 726 elements (80 elements at the interface)
Mesh 4: 160 elements	Mesh 4: 3436 elements (160 elements at the interface)

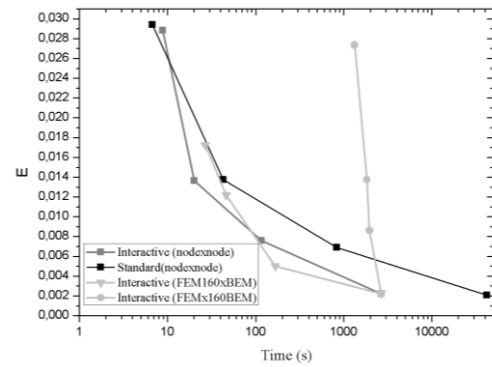
The relative errors depicted in Fig. 4 are computed as follows:

$$E = \sqrt{\frac{\sum_{i=1}^{nf} [ (|{}^iU_c|) - (|{}^iU_a|) ]^2}{\sum_{i=1}^{nf} (|{}^iU_a|)^2}} \quad (17)$$

where  ${}^iU_c$  stands for the computed numerical displacement at point A and frequency  $i$ ,  ${}^iU_a$  stands for the analytical answer at the same point and frequency, and  $nf$  is the total number of frequencies considered in the analysis.



(a) Convergence analysis (error x discretization)

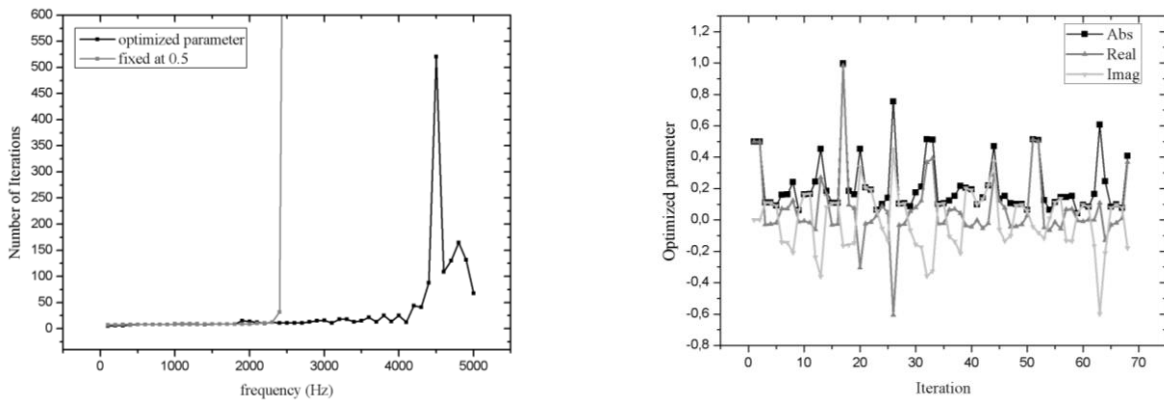


(b) Efficiency analysis (error x CPU time)

Figure 4. Error analysis

As one can observe in Fig. 4a, convergence is achieved, even considering non-matching nodes at the common interface. As it can be further observed in Fig. 4a, the "BEM 160 / FEM" and the "node / node" curves are very close, indicating that, in this case, a small amount of boundary elements are sufficient to properly discretize the model. On the other hand, better results are obtained considering the "FEM 160 / BEM" combination, which was expected, since refined FEM discretizations can better represent the prescribed boundary conditions of the model, providing more accurate analyses. In Fig. 4b, the computed errors are plotted against the CPU times of the analyses. As one can observe, considering matching nodes at the common interface, the iterative coupling procedure is usually more efficient than the standard direct coupling procedure (i.e., for a given CPU time of analysis, more accurate results can be obtained by the iterative procedure; or, for a given accuracy level, faster analyses can be provided by the iterative procedure). Moreover, as described in Fig. 4a, once proper discretizations are considered for each sub-domain of the model, even more efficient analyses may be achieved, highlighting the importance of a coupling procedure that allows flexible and independent discretizations of the involved sub-domains, taking into account non-matching nodes at the common interfaces.

In order to further analyze the performance of the iterative coupling algorithm, the evolution of the optimal relaxation parameter and the convergence of the iterative process are briefly illustrated in Fig. 5. In Fig. 5a, the total amounts of iterative steps necessary for convergence are depicted, for each frequency, considering the spatial discretizations illustrated in Fig. 2. For comparison, results are also depicted considering a constant relaxation parameter value of 0.5. As one can observe, for higher frequencies (above 2500 Hz), convergence is not achieved if  $\lambda = 0.5$  is adopted, highlighting the importance of Eq (16) for the effectiveness of the iterative coupling analysis. Moreover, for a constant value  $\lambda = 1.0$ , convergence is never achieved considering the entire adopted frequency range, further illustrating the importance of relaxation parameters in the iterative coupling technique. In Fig. 5b, the evolution of the optimally computed relaxation parameters (eq. 16) are illustrated, taking into account  $\omega = 5000$  Hz. As one can observe, its evolution is quite complex since it is based on residuals computed at consecutive iterative steps.



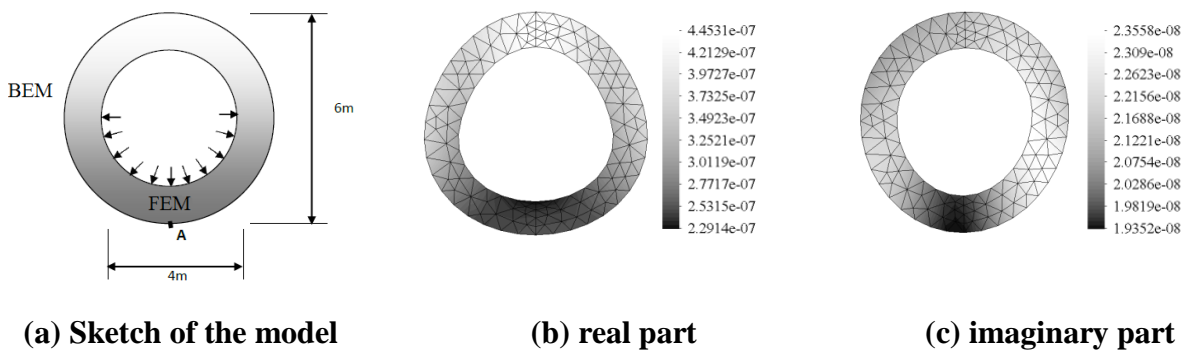
(a) Convergence of the iterative procedure

(b) Evolution of the optimal relaxation parameter

Figure 5. Convergence and optimal relaxation parameter evolution

*Ring-shaped structure inside an infinite elastic domain*

Consider, once again, a circular homogeneous ring-shaped elastic structure, inside a homogeneous and infinite soil environment. The external environment has a density of  $1900 \text{ kg/m}^3$ , Lamé constant  $\mu = 2.5 \times 10^{10} \text{ N/m}^2$  and Poisson's ratio of  $0.35$  (no damping). The tunnel structure is made of concrete and has an external radius of  $3.0 \text{ m}$  and an internal radius of  $2.0 \text{ m}$ . It has a density of  $2500 \text{ kg/m}^3$ , Young's modulus of  $2.5 \times 10^{10} \text{ N/m}^2$  and Poisson's ratio of  $0.2$  (no damping). The structure is loaded as indicated in Figure 6a, i.e., the load is applied at the bottom of the concrete ring internal cavity, with constant amplitude of  $850 \text{ kN/m}$ . The corresponding FEM and BEM discretizations are illustrated in Figure 2b and 2c, respectively. In Fig. 6b and Fig. 6c, the computed deformation of the tunnel is illustrated, considering  $\omega = 500 \text{ Hz}$ .



(a) Sketch of the model

(b) real part

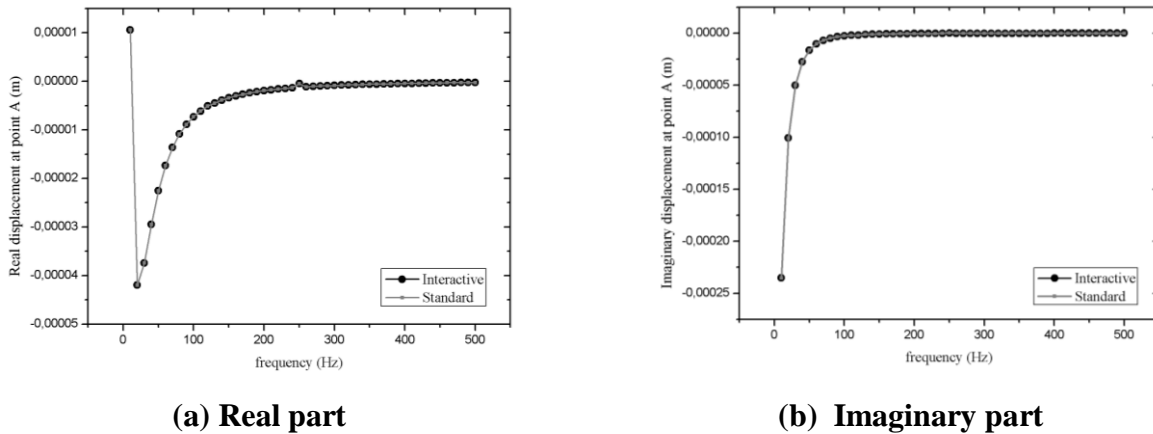
(c) imaginary part

Figure 6. model's sketch and scaled deformation of the tunnel for  $\omega=500\text{Hz}$

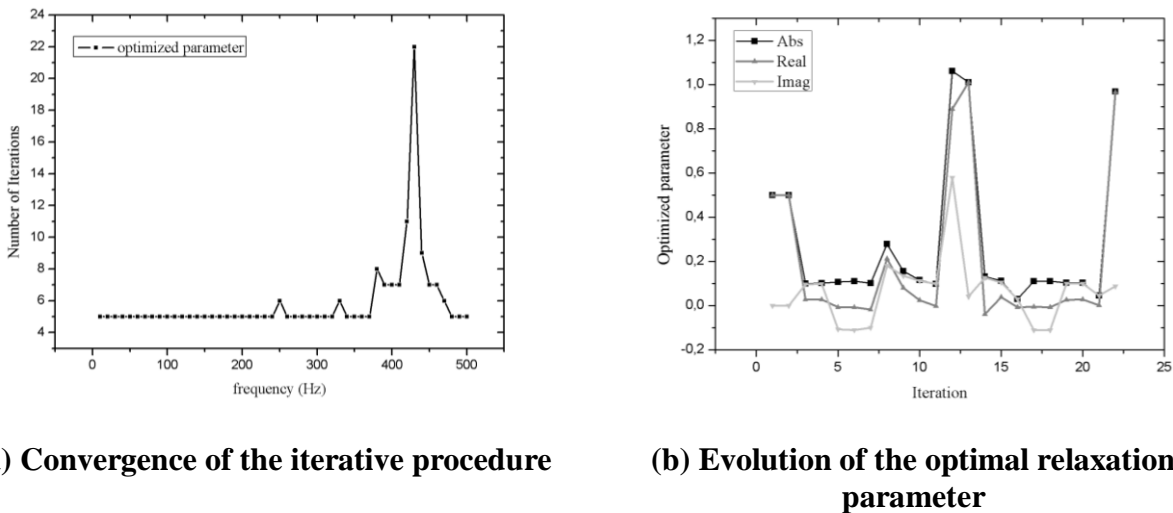
Fig. 7 illustrates the displacements computed at point A (see Fig. 6a), taking into account the proposed iterative and a standard direct FEM-BEM coupling procedure, considering a frequency range from  $10$  to  $500 \text{ Hz}$ . As one can observe, the results provided by these different approaches are once again in good agreement, indicating that the iterative solution is converging to the right solution.



In Fig. 8a, the total amounts of iterative steps necessary for convergence are depicted, taking into account the selected frequency range. As one can note, for all tested frequencies, convergence occurred with a relatively small amount of iterations, with no more than 25 iterations being necessary at any of the tested frequencies. It is important to highlight that, for the present application, for  $\lambda = 0.5$  and  $\lambda = 1.0$ , convergence is never achieved considering the entire adopted frequency range, further illustrating the importance of optimal relaxation parameters in the iterative coupling technique. In Fig. 8b, the evolution of the optimally computed relaxation parameters (Eq. 16) are illustrated, taking into account  $\omega = 430$  Hz.



**Figure 7. Vertical displacements at point A**



**Figure 8. Convergence and optimal relaxation parameter evolution**

### Conclusions

A FEM-BEM iterative coupling algorithm was discussed here to analyze elastodynamic models, taking into account frequency domain formulations. In order to deal with this ill-posed problem, optimal relaxation parameters were introduced into the iterative coupling analyses, enabling convergence at a relative low number of iterative steps. An efficient and easy to implement

expression to compute the optimal relaxation parameters was discussed and tested, providing an effective and robust iterative coupling procedure.

The use of iterative coupling approaches enables the separated analysis of different sub-domains, leading to better conditioned, smaller and easier to deal with systems of equations, as well as independent definitions of nodal points along distinct sub-domains, allowing non-matching nodes on common interfaces to be easily considered. In section 6 several results were presented, illustrating the versatility and effectiveness of the proposed procedure.

As a matter of fact, the present methodology represents an important step forward in the analyses of wave propagation in frequency domain problems considering iterative coupling procedures, which are well-known ill-posed problems, specially taking into account sub-domains governed by different physical properties and discretized by different numerical techniques.

## Acknowledgments

The financial support by CNPQ (*conselho nacional de desenvolvimento científico e tecnológico*) and FAPEMIG (*fundação de amparo à pesquisa do estado de minas gerais*) is greatly acknowledged.

## References

- Bathe KJ, (1996), *Finite Element Procedures*, Prentice Hall Inc., New Jersey.
- Bendali, A., Boubendir, Y. and Fares, M. (2007), “A FETI-like domain decomposition method for coupling finite elements and boundary elements in large-size problems of acoustic scattering”, *Computers & Structures*, **85**, 526–535.
- Beskos, D.E. (2003), “Dynamic analysis of structures and structural systems”, In: Beskos DE, Maier G (eds) *Boundary Element Advances in Solid Mechanics*, CISM International Centre for Mechanical Sciences No. 440, Springer-Verlag, Wien, New York, pp. 1–53
- Boumaiza, D. and Aour B. (2014), “On the efficiency of the iterative coupling FEM–BEM for solving the elasto-plastic problems”, *Engineering Structures*, **72**, 12–25
- Dominguez J, (1993), *Boundary elements in dynamics*, Computational Mechanics Publications, Southampton and Boston.
- Elleithy, W. (2012), “Multi-region adaptive finite element-boundary element method for elasto-plastic analysis”, *International Journal of Computer Mathematics*, **89**, 1525-1539.
- Elleithy, W.M., Al-Gahtani, H.J. and El-Gebeily, M. (2001), “Iterative coupling of BE and FE methods in elastostatics”, *Engineering Analysis with Boundary Elements*, **25**, 685–695.
- Elleithy, W. and Grzhibovskis, R. (2009), “An adaptive domain decomposition coupled finite element–boundary element method for solving problems in elasto-plasticity”, *International Journal for Numerical Methods in Engineering*, **79**, 1019–104
- Godinho, L. and Soares, D. (2013), “Frequency domain analysis of interacting acoustic-elastodynamic models taking into account optimized iterative coupling of different numerical methods”, *Engineering Analysis with Boundary Elements*, **37**, 1074–1088.
- Hughes TJR, (2000), *The Finite Element Method – Linear Static and Dynamic Finite Element Analysis*, Dover Publications, New York.
- Jahromi, H.Z., Izzuddin, B.A. and Zdravkovic, L. (2009), “A domain decomposition approach for coupled modelling of nonlinear soil-structure interaction”, *Computer Methods in Applied Mechanics and Engineering*, **198**, 2738–2749.
- Lin, C.C., Lawton, E.C., Caliendo, J.A. and Anderson, L.R. (1996), “An iterative finite element-boundary element algorithm”, *Computers & Structures*, **59**, 899–909.
- Soares, D. (2008), “An optimised FEM-BEM time-domain iterative coupling algorithm for dynamic analyses”, *Computers & Structures*, **86**, 1839-1844.

- Soares, D. (2012), “FEM-BEM iterative coupling procedures to analyze interacting wave propagation models: fluid-fluid, solid-solid and fluid-solid analyses”, *Coupled Systems Mechanics*, **1**, 19–37.
- Soares, D. and Godinho, L. (2012), “An optimized BEM-FEM iterative coupling algorithm for acoustic-elastodynamic interaction analyses in the frequency domain”, *Computers&Structures*, **106-107**, 68–80.
- Soares, D. and Godinho, L. (2014), “An overview of recent advances in the iterative analysis of coupled models for wave propagation”, *Journal of Applied Mathematics*, **2014**, Article ID 426283.
- Soares, D., von Estorff, O. and Mansur, W.J. (2004), “Iterative coupling of BEM and FEM for nonlinear dynamic analyses”, *Computational Mechanics*, **34**, 67–73.

# The self-assembly of self-interstitial-atoms caused by an electron beam

S. T. Nakagawa, K. Ohashi, and K. Sugita

Graduate School of Science, Okayama Univ. of Science, Japan.

1-1 Ridai-Cho Okayama 700-0005 Japan

## Abstract:

A low energy electron beam (EB) can let the self-interstitial atoms (SIA) cause the self-assembly (DSA), i.e.,  $\{311\}_{\text{SIA}}$  platelet in c-Si is discussed. We have studied how an SIA migrates toward metastable sites that form the platelet. We used a molecular dynamic (MD) simulation to trace all the atoms and crystallographic analysis method (PM) method to analyse crystalline defects. In MD, we skipped the EB irradiation stage that produces SIAs. Instead, before MD some Frenkel pairs (FP) or SIA coupled with a vacancy were randomly distributed in bulk. Then we monitored how contributed the surrounding target atoms to the DSA process. When the FP concentration is higher than 3 atomic percent (3 %), the number of atoms to be stayed at metastable sites looked to be saturated. However, it increased significantly at lower FP's concentration, e.g., at 1 %. At the same time, somewhat cooperative and oscillatory movement of target atoms in bulk was found in terms of the long-range-order (LRO) parameter. The oscillatory change of LRO continued for a few tens of ps, which is much slower than the so-called lattice vibration. Therefore, we suppose the cooperative motion like slow phonon might promote the migration of SIA and increase the number of atoms on  $\{311\}_{\text{SIA}}$  planes, whereas no such global wavy motion was observed in cases including much higher FPs, as if somewhat pinning effect had occurred therein.

Keywords; electron beam, MD simulation, directed self-assembly, crystallography

## 1. Introduction

The  $\{311\}$  platelet in a c-Si is known to be composed of the self-interstitial atom (SIA)s, which was found by the in-situ measurement using High Resolution Transmission Electron Microscopy by Takeda at an annealing temperature ( $T_A$ ) of 723 K [Takeda (1991)]. This is a kind of directed self-assembly (DSA). Hereafter we call

(311)<sub>SIA</sub> formation in c-Si DSA. The energy of the electron beam (EB) adopted was 300 keV, which was much higher than the threshold value of 145 keV to eject a target atom from a lattice site in c-Si. However, it was much lower than 1 MeV above which the electronic stopping is dominant [Yasuda et al. (2007)]. The very significant experimental condition to produce {311}<sub>SIA</sub> platelet was the  $T_A$ . The rate of fluence was  $\Phi = 5 \times 10^{24}$  electrons/ sec/m<sup>2</sup> and it was continued for  $t = 1080$  sec or longer. The procedure of DSA had two steps [Takeda and Kamino (1995)]; the first one was to form atomic rows along  $\langle 0\bar{1}1 \rangle$  direction, then the second one was to align those rows parallel along the direction of  $\langle 2\bar{3}\bar{3} \rangle$ , which formed (311) platelet. It was a notable experiment, because the geometry of this lattice structure was exactly the same with that predicted by the pixel mapping (PM) method; a (311)<sub>SIA</sub> plane is made of *metastable* {H} or {T} sites [Nakagawa (2007a)].

Firstly, we had reproduced the DSA at  $T_A = 723$  K using an ion beam (IB) (1keV Si ion into c-Si), making use of a molecular dynamic (MD) simulation [Nakagawa et al. (2007b)]. The two-step process of DSA that we found was exactly same with that proven by EB [Takeda and Kamino (1995)], and the areal density of the SIAs on (311)<sub>SIA</sub> agreed with that experiment due to EB irradiation. The  $\langle 110 \rangle$  alignment on (311)<sub>SIA</sub> was clearly observed. The assignment of Miller index of an atomic plane or crystallographic orientation of an atomic row was possible using PM [Nakagawa et al. (2002; 2007a; 2009)]. This calculation proved that when SIAs are produced by whatever external disturbance, they migrate into the wide space between two sets of {311} double-layers and reallocated by the help of local field surrounding the SIA.

Secondly, assuming the Frenkel pair (FP) production due to EB irradiation, we had obtained quite similar result of DSA at  $T_A = 723$  K, making use of a MD [Nakagawa (2014)]. Because the recoil cross-section of an electron against target atom is very small and it occurs so sparsely in a wide region. Then, we skipped the EB irradiation stage in MD. Instead we scattered FPs uniformly in bulk and then started MD to monitor the movement of all atoms. After heating started, SIAs migrated and formed a platelet, where  $\langle 110 \rangle$  alignment was again observed.

A recent finding was the FP's concentration dependence on DSA. Target atoms in bulk showed somewhat cooperative and oscillatory behaviour of entire atoms, when the FP's concentration was lower than 3 atomic percent (3 %). The phonon-like motion was identified by the long-range-order (LRO) parameter that defined the degree of perfectness of crystallinity. In cases of higher FP's concentration, however,

no peculiar behaviour occurred, as if somewhat pinning-effect suppressed the migration of SIAs. Present aim is to study how significantly the pinning-effect exerts on the DSA.

## 2. Method

### 2.1 Invisible but metastable sites for SIA

In the PM method, the location of an atom  $(x, y, z)$  in a cubic crystal is presented by a set of integer  $(l, m, n)$  using Gauss symbol [Nakagawa (2002)]. Namely,  $(l, m, n) = ([x/(d/4)], [y/(d/4)], [z/(d/4)])$ , where the “ $d$ ” is the lattice constant of a unit cubic of a crystal. This means to adopt marginal space for each lattice site, and we call that small cube *one pixel*. The PM method is available for 24 cubic-type crystals belonging to the space group #195-#230.

The c-Si is one of the zincblende (ZnS)-type crystals, which is made of two fcc-(1) and fcc-(2) lattices. We regard a lattice point where one atom is occupies in a perfect crystal “stable site”. These lattice points satisfy the following condition: In the fcc-(1) lattice, the address  $(l, m, n)$  of a stable site satisfies an algebraic rule of “ $l + m + n = 4k$ , and  $(l, m, n)$  are all even”. Similarly, in the fcc-(2) lattice, the corresponding rule is “ $l + m + n = 4k+3$ , and  $(l, m, n)$  are all odd”. Both lattices are visible.

**Table 1** The identification table for an atom located at  $(l, m, n)$  in a ZnS-type crystal, i.e., a compound AB. The last two planes are made of metastable {T} and {H}, both form the *invisible double-layer structure*. “Meta.” is the abbreviation of metastable. “ $(311)_A$ ” or “ $(31)_B$ ” means  $\{311\}_{SIA}$ . The “ $k$ ” and “ $J$ ” are arbitrary integers.

Site	Which site? Which crystal?		Which plane?			
	$l+m+n$	$l, m, n$	On which fcc lattice	$D = 3l+m+n$	Double-layer	Plane
(1) Stable	4k	all even	real fcc-(1)	$D = 4J$	Visible  Atoms at lattice site	$\{311\}_A$
(2) Stable	4k+3	all odd	real fcc-(2)	$D = 4J + 1$		$\{311\}_B$
{H} Meta.	4k+2	all even	virtual fcc-(1') from fcc-(1) shift by $\langle 2,0,0 \rangle$	$D = 4J + 2$	Invisible  SIAs	$\{311\}_A'$
{T} Meta.	4k+1	all odd	virtual fcc-(2') from fcc-(2) shift by $\langle 2,0,0 \rangle$	$D = 4J + 3$		$\{311\}_B'$

The PM makes it possible to identify various types of crystalline defects from micro to macroscopic scale. Another merit of the PM is to evaluate the LRO parameter [Nakagawa (2002, 2007a; 2009)], which is defined as the ratio of number of atoms allocated at stable sites to that of total atoms in the MD box. Note  $LRO = 1$  means the state of a perfect crystal while  $LRO = 0$  means amorphous state.

In Table 1, “metastable sites, {H} or {T}” are listed. They are closely related to the stable sites (1) and (2), respectively, although those lattices are invisible in a perfect crystal. The condition for {H} site is “ $l + m + n = 4k+2$ , and  $(l, m, n)$  are all even”. The similar condition for {T} site is “ $l + m + n = 4k+1$ , and  $(l, m, n)$  are all odd”. If the address of one atom does not satisfy any of those four conditions, the atom in a pixel for unstable site [Nakagawa (2002)].

Because of the translational symmetry between (1) and {H}, which are shifted by the constant vector of  $(0, 0, d/2)$  to each other, they individually form mathematically equivalent lattices with the same space group. The visible (1) and invisible {H} are like twin lattices, although the former is stable and the latter is metastable. The situation is the same for (2) and {T} [Nakagawa (2002)]. If some SIAs stay at such metastable sites, they look like a part of ZnS-type crystal and crystalline alignment as atomic row or atomic plane can be observed. Therefore, we have supposed, that invisible lattice can work as a hidden and secret chart for DSA. In this article, we examine the effect of the concentration of FPs on the DSA process, because the presence of embedded FPs should deform the potential field in bulk.

## 2.2 How to identify {311} defects;

A crystal has periodic array in a set of { $HKL$ } planes labelled by the fixed Miller index. When an atom has address  $(l, m, n)$  defined by PM, the sequential number of the ( $HKL$ ) plane on which that atom stays is known as follows [Nakagawa (2009)],

$$D = l \times H + m \times K + n \times L \quad (1).$$

When substituted the conditions for addresses of (1), (2), {T}, {H}, into Eq.(1) for a case of ( $HKL$ ) = (311); we get an interesting correspondence. That is, stable site (1) and (2) are located on the planes termed  $D = 4J$  and  $D = 4J + 1$ , respectively. These two layers can form the visible paired double-layer. On the other hand, metastable {T} and {H} sites are on the planes termed  $D = 4J + 2$  and  $4J + 3$ , respectively. These two layers can form the *invisible paired double-layer*. This alignment is repeated every four layers as is tabulated in Table 1.

Making use of Table 1, we count how many SIAs are trapped at metastable sites, {H} or {T}, on the invisible  $\{311\}_{\text{SIA}}$  planes as will be plotted later in Fig. 2. The interplanar distance ( $\Delta D$ ) of adjacent plane in a series of  $\{HKL\}$  planes is  $d/\{H^2+K^2+L^2\}^{1/2}$ . It is  $\Delta D = d/\sqrt{11}$  for the case of  $\{311\}$  planes, thus the *invisible paired double-layers* yield wide gap of  $3d/\sqrt{11}$  for an SIA between two sets of visible double-layers. A sheet of  $(311)_{\text{SIA}}$  plane has two dimensional diagonal grids made of  $\langle 0\bar{1}1 \rangle$  and  $\langle 1\bar{2}\bar{1} \rangle$  axes. The area of one oblique unit cell of  $11d^2$  is allocated to one metastable site on  $(311)_{\text{SIA}}$  plane. Since the shape of  $\{311\}$  planes inside the MD box are all polygons, because of truncation due to the surface of MD box adopted. When the MD box is composed of  $p \times p \times p$  unit cubes, the address of the lattice point at the body-diagonal corner is  $(4p, 4p, 4p)$ . Then as we look over the specific sites of  $(l, m, n) = (0, 0, 0) \rightarrow (2p, 2p, 2p) \rightarrow (4p, 4p, 4p)$ , the  $D$  number of  $(311)$  plane including those sites changes from  $D = 0 \rightarrow 10p \rightarrow 20p$ . With increasing  $D$ -value, the number of metastable sites on those planes changes from zero  $\rightarrow$  maximum  $\rightarrow$  zero, respectively. Namely the number of metastable sites can show an upheaval profile as a function of  $D$ .

### 2.3 MD Simulation model

Generally speaking, the defect density in the collision cascade caused by an EB is much sparse if compared with the case due to an IB. With a low-energy EB below 1 MeV, the dominant energy loss is due to the nuclear stopping. The nuclear stopping power of 300 keV EB for silicon material is 2 eV/10 nm [Yasuda et al. (2007)]. Thus, one electron in an EB loses 1 eV when it penetrates a MD box whose depth is 5 nm. Namely, the depth doze of EB is very low. On the other hand with respect to the mean free path, the inelastic one is 240 nm while the elastic one is 12 nm. The formation energy of one FP is 3.61 eV. These facts make it possible to assume that the FPs would be distribution uniformly.

We embed FPs into c-Si prior to MD simulation. Here MD box is composed of  $16 \times 16 \times 16$  unit cubes, thus 32768 silicon atoms are enclosed. With this box,  $D_{\text{max}}$  is 320. The range of FP's concentration is from 1 % to 10 %. We adopted the so-called NVT frame for our MD, where three quantities are kept constant; they are the number ( $N$ ) of target atoms, the volume ( $V$ ) of the MD box, and the bulk temperature ( $T$ ). After distributing FPs, we started MD simulation in the following order: (i) Thermalization to keep  $T$  constant at room temperature ( $T = 300\text{K}$ ; RT). When the deviation of  $(T - RT)/T$  is stabilized within the fluctuation less than 1.0 % in average, next step started. (ii) Heating up to  $T_A$ . At MD box walls, particle's velocity was regulated to yield the

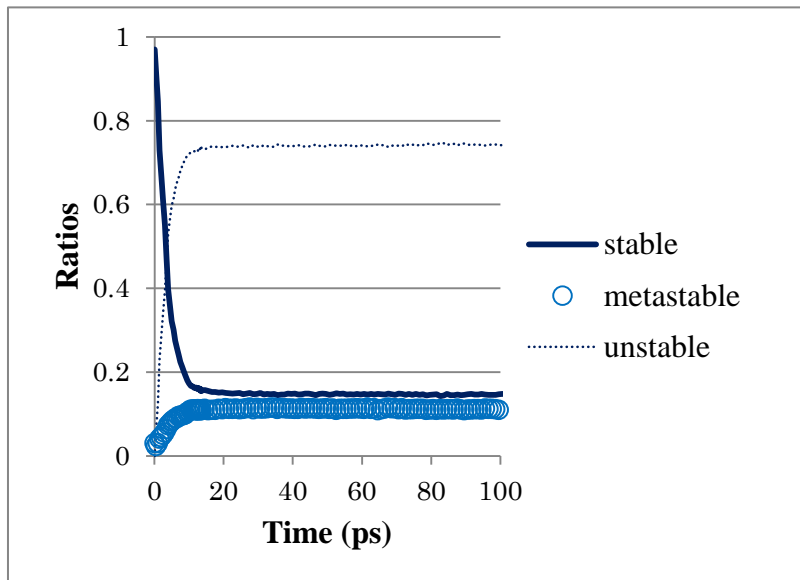


heating rate of 0.5 K/femtosecond. When the deviation  $(T - T_A)/T$  becomes less than 1.0 %, further heating was quitted. (iii) Annealing to keep  $T = T_A$  for 200 picoseconds (ps) at longest.

### 3. Result

#### 3.1 How SIA migrates towards metastable site?

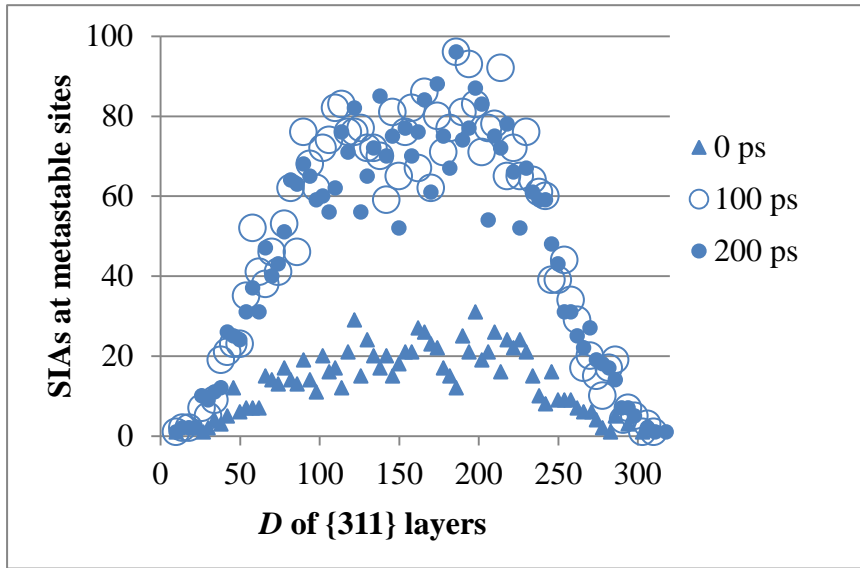
Figure 1 shows the change in the fractions of atoms staying different sites, after starting heating (time  $t = 0$ ) for a case of 3 % of FP's concentration, where  $T_A = 1000$  K. Symbols indicate the ratios of atoms located at stable (-), metastable ( $\circ$ ), unstable (--) sites to the number of total atoms.



**Figure 1** Time series of the site distribution of target atoms, since starting heating from RT to  $T_A = 1000$  K, for a case with 3 % of FP concentration. Symbols indicate the ratios of atoms located at stable (-), metastable ( $\circ$ ), unstable (--) sites.

From this figure, it is known that many atoms left the original stable site within 10 ps after heating started, which was crystal to amorphous (CA) transition. Since then, it looks no significant change in those fractions of three kinds of states after CA transition. In the case of of higher concentration, 5 % and 10 %, such situation was quite similar, there was no significant concentration dependence. In the mean time,

the number of atoms at metastable sites, {T} or {H}, increased significantly and reached at the same level with that at stable sites here.

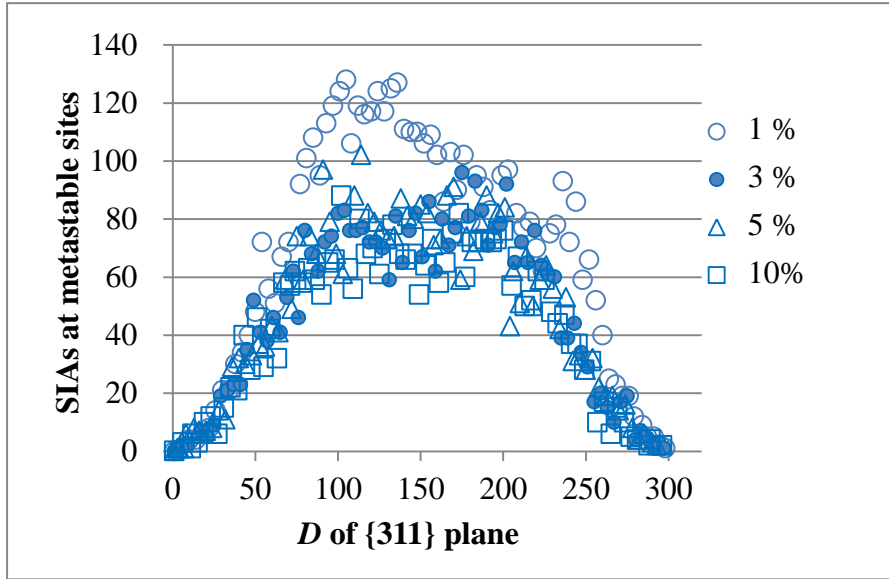


**Figure 2** The time dependence of the SIA accumulation toward metastable sites on  $(311)_{\text{SIA}}$ , where  $T = T_A = 1000\text{K}$  is kept. Each symbol indicates the sum of atoms located at metastable sites on one paired invisible double-layers.

Figure 2 shows the significant accumulation of atoms toward metastable sites due to heating up to  $T_A = 1000\text{ K}$  when FP's concentration is 3 %. The migration was promoted evidently until 100 ps. However, since then any further migration toward metastable sites looked suppressed. Even increasing FP's concentration, up to 10 %, migration looks less easy as if somewhat pinning-effect might have appeared.

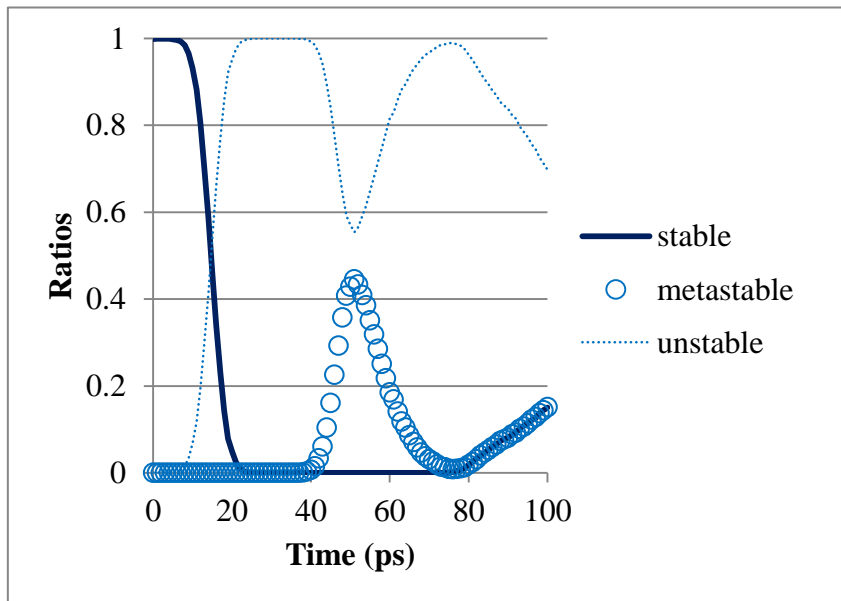
### 3.2 The pinning-effect on the DSA

Figure 3 shows the FP's concentration dependence on the number of atoms trapped at metastable sites. The number of atoms located at metastable sites did not increase with the initial FP's concentration but showed rather saturation when the initial concentration was higher than 3 %. The migration of SIAs can be assisted by atoms nearby. Therefore, we will examine why such concentration dependence happened. The movement of target atoms can be monitored in terms of the LRO parameter. Note  $LRO$  is equal to the ratio of occupied number of atoms at stable sites to that of the total atoms as was plotted in Fig. 1



**Figure 3** The FP's concentration dependence on the SIA migration toward the metastable sites within 100 ps at  $T_A = 1000$  K. Symbols for different concentrations are ( $\circ$ ; 1%), ( $\bullet$ ; 3%), ( $\Delta$ ; 5%), and ( $\square$ ; 10%).

3.3 *The cooperative motion of target atoms vs. the pinning effect:*



**Figure 4** Time series of the site distribution when the initial concentration of FP was 1% with keeping  $T_A = 1000$  K. Notations are same to Fig. 1.

If compared this Fig.4 with Fig.1, the presence of somewhat global and synergistic movement of target atoms is evident when FP concentration is low. Moreover the

trend of restoration of LRO looked, at around 50 ps and after 80 ps, although there was once disappeared by the slower CA transition if compared with Fig.1. This oscillatory and cooperative motion of atoms located at metastable sites inversely corresponds to that at unstable ones, which implies the simultaneous migration of many SIAs. This result could explain the profile of Fig. 3; more atoms were located at metastable sites when less FPs were scattered before MD. The increasing ratios of stable and metastable after 80 ps indicate the sign of further restoration by the assist of deformed potential field. In other words, the presence of much FPs may cause like pinning-effect that may suppress such a cooperative motion of target atoms as was shown in Fig.1

#### 4. Conclusion

We have studied why the SIA is apt to form the  $\{311\}$  planar defect ( $= (311)_{\text{SIA}}$  platelet) in c-Si when it was irradiated by a low-energy (sub-MeV) EB or a low energy (a few keV or less) IB. We traced all the atoms by an empirical MD simulation, and analysed the crystalline defects by the PM method.

Based on the PM, we have supposed the DSA is directed by the hidden chart for migrating SIAs. Because the measured DSA had shown exactly same geometrical lattice structure with that described by the PM method. That is, metastable sites are forming hidden fcc lattices in the wide gap between two sets of  $\{311\}$  double-layers. Therefore, when an SIA would migrate, it is apt to enter such wide space and would be trapped at metastable sites, where the hidden lattice would provide a chart for migrating SIA.

The EB irradiation first produces many FPs in c-Si. Then we embedded FPs and started MD simulation to monitor how target atoms assist SIAs to migrate and form the  $\{311\}_{\text{SIA}}$  platelet, in terms of the long-range-order parameter. We have confirmed that a global and synergistic movement of target atoms promoted the migration of SIAs, when the FPs concentration is lower than three atomic percent. Beyond the critical value of 3 %, however, pinning effect emerged, i.e., no significant cooperative wavy-motions occurred, and the defect formation is suppressed.

### Acknowledgment

This work has been supported by a grant for Scientific Research (C) (Grant No. 21510108) from JSPS (Japan Society for the Promotion of Science) and partly by a grant for Scientific Research on Innovative Areas (Grant No. 21104003) from MEXT (Ministry of Education, Culture, Sports, Science, and Technology, Japan) and partly supported.

### References

- Takeda, S. (1991), An Atomic Model of Electron-Irradiation-Induced Defects on {113} in Si, *Jpn. J. Appl. Phys.*, **30**, L639-642.
- Takeda, S.; Kamino, T. (1995), Agglomeration of self-interstitials in Si observed at 450 °C by high-resolution transmission electron microscopy, *Phys. Rev. B*, **51**, 2148-2152.
- Nakagawa, S. T. (2002), Pixel mapping analysis to characterize long-range order interactions in crystals under ion irradiation, *Phys. Rev. B*, **66**, 094103/1-7.
- Nakagawa, S. T. (2007a), Crystallographic analysis of phase-changes in cubic crystals, *J. Phys. Soc. Jpn.*, **76**, 034603/1-12.
- Nakagawa, S. T.; Whitlow H. J.; Betz, G. (2007b), Detection of planar defects caused by ion irradiation in Si using molecular dynamics, *Surf. Coating Tech.*, **201**, 8393-8397.
- Nakagawa, S. T. (2009), In *Ion beams in Nanoscience and Technology* (Chap.9), Springer
- Nakagawa, S. T. (2014), The directed self-assembly for the surface patterning by electron beam, *Rad. Effect Defect Solids*, (2014) doi: 10.1080/10420150.2014.984613.
- Yasuda, M.; Kimoto, Y.; Tada, K.; Mori, H.; Akita, S.; Nakayama, Y.; Hirai, Y. (2007), Molecular dynamics study of electron-irradiation effects in single-walled carbon nanotubes, *Phys. Rev. B*, **75**, 205406/1-5.

# Cognition of Parameters' Role on Vertical Control Device for Aerodynamic Characteristics of Aircraft Using Data Mining

Kazuhisa Chiba<sup>1</sup> \*; Taiga Omori<sup>2</sup>, Yasuto Sunada<sup>2</sup>, and Taro Imamura<sup>2</sup>

<sup>1</sup> Graduate School of Informatics and Engineering,  
The University of Electro-Communications,  
1-5-1, Chofugaoka, Chofu, Tokyo 182-8585, Japan

<sup>2</sup> Department of Aeronautics and Astronautics,  
The University of Tokyo,  
7-3-1, Hongo, Bunkyo, Tokyo 113-8656, Japan

December 3, 2015

## Abstract

The new concept to place the vertical airfoil device as control surface has been discovered so as to improve the aerodynamic performance of aircraft. The concept was predicated on not only the several devices as vortex generator and winglet but also the wing-mounted engine system of the HondaJet. Thereupon, the wind tunnel experiment has been implemented in order to investigate the influence of the vertical control device with the symmetrical airfoil shape. Furthermore, a self-organizing map as data mining has been performed for the experimental data in order to qualitatively elucidate the correlations among the aerodynamic performances as design requirements and the design parameters to place the vertical control device. Consequently, it has been revealed the design information regarding the intimate correlations. Moreover, there is the sweet spot in the design space to improve the aerodynamic performances.

Keyword: Vertical device; Control surface; Aerodynamics of aircraft;  
Data mining; Self-organizing map.

---

\*kazchiba@uec.ac.jp

## 1 Introduction

Although the surface of the main wing of aircraft, especially upper wing surface, is generally desirable to be smooth in ordinary design of an aircraft, there are several exceptions to this universal tacit knowledge[6, 9], such as small devices for flow control. Honda Aircraft Company has designed and developed a business jet aircraft named as the HondaJet[3]. Despite the fact that the devices are generally designed small on general knowledge even when devices will be on the wing surface, the HondaJet mounts its engine over the upper surface of the wing with the pylon. The design of the HondaJet astonishingly reveals that the optimum location of the nacelle and the cross section of the pylon exists to accomplish lower drag coefficient compared with the clean wing[4]. This fact indicates that the devices on the wing surface, whose size is independent on the convention of aircraft design, can uncommonly improve the aerodynamic performance of aircraft.

Thereupon, in the present study, a new basic idea regarding a vertical control device on the upper surface of the main wing will be proposed in order to improve the aerodynamic performance of aircraft due to the flow control on the wing surface. The devices are expected to be also installed on the trailing edge of the pylon in order to improve the aerodynamic performance. Therefore, the objective of the present study is to elucidate the effectiveness on the aerodynamic performance regarding the control surface which is vertically mounted on the wing. As a first step, the wind tunnel experiment is implemented in order to quantitatively reveal its effectiveness[7]. As a second step, data mining is performed by using a self-organizing map for the experimental data so that the global design information for the design space will be also efficiently revealed. Especially, the keystone of the present treatise corresponds to the second step. The objectives of the present data mining are that significant experimental conditions are efficiently addressed from  $10^3$ -order conditions. Furthermore, the obtained design knowledge will be utilized in order to generate a wind tunnel model for the next-step experiments so that a vertical control device is efficiently installed and its optimum geometry will be designed.

## 2 Problem definition

The simple symmetrical aircraft model constructed by the main and tail wings with rectangular planform and vertical control device is developed in order to utilize in the wind tunnel experiment. The specification of an

Table 1: Specification of aircraft model for wind tunnel experiment.

<i>component</i>	<i>content</i>	<i>data</i>
fuselage	length	370 [mm]
	width	44 [mm]
	height	55 [mm]
main wing	span length	404 [mm]
	chord length	80 [mm]
	airfoil	NACA2410
	aspect ratio; $AR$	5.05 [-]
	taper ratio	1.0 [-]
vertical control device	chord length	40 [mm]
	span height	40 [mm]
	airfoil	NACA0010

Table 2: Design parameters and their discretized design space.

<i>description</i>	<i>symbol</i>	<i>design space</i>
spanwise distance	$\mu$ [mm]	$10 \leq \mu \leq 170$ for every 10
deflection angle	$\delta$ [deg]	$-10 \leq \delta \leq 10$ for every 2
angle of attack of body	$\alpha$ [deg]	$-6 \leq \alpha \leq 20$ for every 2

aircraft model is shown in Table 1. The fuselage and tail wings constructed by the plane surfaces are fixed. The main wing itself is fixed, however, the vertical control device is shifted on the upper surface of the main wing[7]. Thereupon, the model geometry is defined by the following three design parameters. The first is the spanwise distance from the root of the main wing to the installed position of the vertical control device  $\mu$  [mm]. The second is the deflection angle of the vertical control device onto the upper surface of the main wing  $\delta$  [deg]. The illustrated description of these two design parameters is shown in Fig. 1. The third is the angle of attack of the body  $\alpha$  [deg]. The design space of the each design parameter is summarized in Table 2. Since the experiment cannot strictly set the values of the design parameters, the three design parameters have not continuous but discretized values.  $\mu$  is the distance between the root of the main wing (that is, body wall) and the 25% position of the mean aerodynamic chord for the vertical control device.  $\mu$  moves from 10 to 170 [mm] for every 10 [mm]. The two vertical control devices are symmetrically set on the main wing.  $\delta$  is the deflection angle of the vertical control device onto the main wing. The revolutionary center is set on the 25% mean aerodynamic chord of the vertical control device.  $\delta$  is



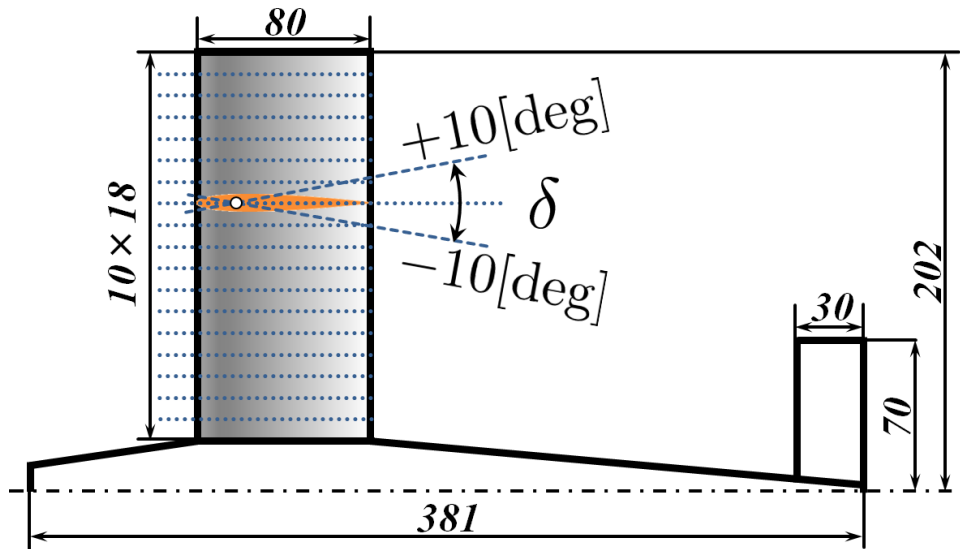


Figure 1: Bird's-eye illustration of overall geometry. The dotted lines on the main wing describe the 17 installation positions (the length from the body wall denotes  $\mu$ ) of the vertical control device colored by orange. The 25% position of the mean aerodynamic chord for the vertical control device is described by the white point in the orange color.

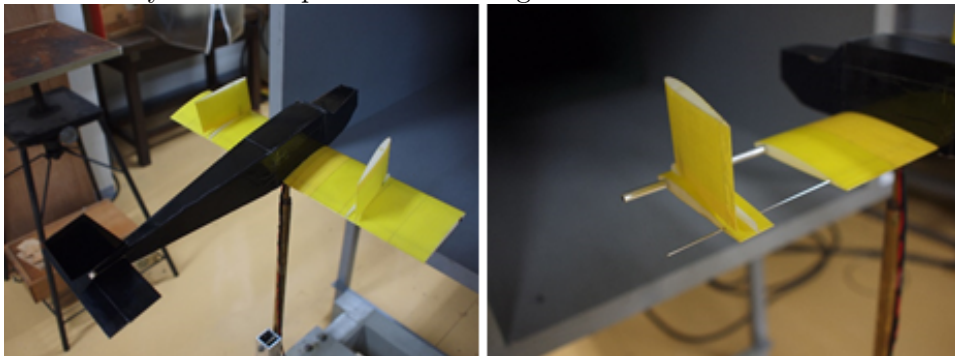


Figure 2: The wind tunnel model constructed by the separated wing blocks.

defined to be the positive value when the trailing edge of the vertical control device is installed on the outboard side shown in Fig. 1.  $\delta$  changes from  $-10$  to  $+10$  [deg] for every 2 [deg]. Note that there are no experimental data in the case of  $\mu$  of 10 [mm] and  $\delta$  of  $-10$  [deg] because the vertical control device interferes in the fuselage.  $\alpha$  changes from  $-6$  to  $+20$  [deg] for every 2 [deg].

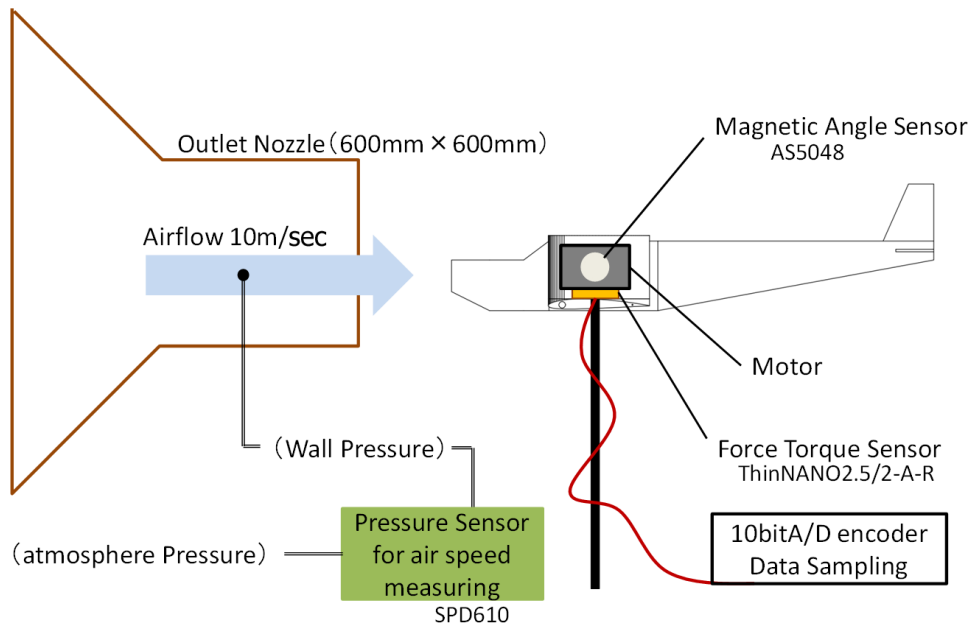


Figure 3: Schematic illustration of the system for the wind tunnel experiment.

The total number of experimental conditions is 2,604. Geometry is designed by using a computer-aided design software and it is outputted as the stereo lithography data to generate the wind tunnel model.

The wind tunnel model is made from wood. It is constructed by several elements in order to simply alter the geometry for all conditions of the wind tunnel experiments. The appearance of the wind tunnel model and the elements of the main wing are shown in Fig. 2.  $\mu$  can be moved by inserting blocks in the different order along the spar. Each wing block is made by using 3-dimensional printer. The vertical control device is also separately constructed and it is attached onto the main wing with a screw so that  $\delta$  can be simply changed. There are gaps between the leading and trailing edges of the vertical control device and the upper surface of the main wing, however, they are negligible small.

### 3 Experimental result

The experiment was performed by using the blow-down wind tunnel at the department of aeronautics and astronautics, the University of Tokyo. Its

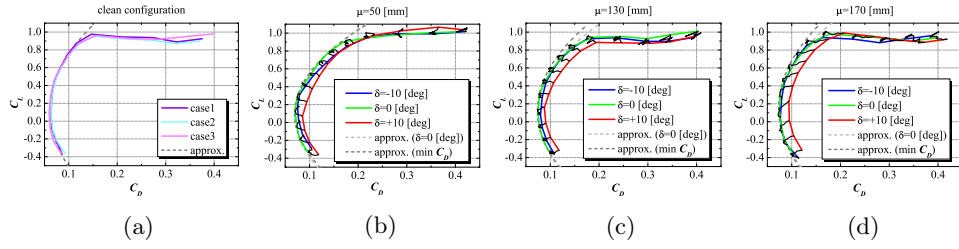


Figure 4: Polar curves. (a) clean configuration, (b) installed configuration at  $\mu = 50$  [mm], (c) installed configuration at  $\mu = 130$  [mm], and (d) installed configuration at  $\mu = 170$  [mm],

outward form has 600 [mm] height and width. The flow velocity was set to be 10 [m/sec] for all experimental conditions. The Reynolds number based on the chord length as the reference one was approximately  $5.0 \times 10^4$ . All of the experiments were carried out for 10 [sec] with the sampling frequency of 1,000 [Hz]. Therefore, all of the data regarding the aerodynamic performance obtained from the experiments are the time-averaged value of 10,000 points for 10 [sec]. Figure 3 shows the conceptual illustration of the present measurement system for the present wind tunnel experiment.  $\alpha$  was controlled by the microcomputer using a proportional-integral-derivative controller. Three aerodynamic performances of the body as a whole, the lift  $L$ , the drag  $D$ , and the pitching moment  $M_p$ , are gauged by using the wind tunnel balance. These performances are respectively transformed into the lift coefficient  $C_L$ , the drag coefficient  $C_D$ , and the pitching moment coefficient  $C_{M_p}$ , which describe the following equation divided by the dynamic pressure using the air density  $\rho$ , the velocity  $v$ , and the planform area of the main wing  $S$  as the reference one.

$$C_{\square} = \frac{\square}{\frac{1}{2}\rho v^2 \cdot S}, \quad (1)$$

where,  $\square$  denotes  $L$ ,  $D$ , and  $M_p$ .

The Oswald efficiency factor  $e$  is selected as an indicator to preliminary evaluate the aerodynamic performance of the aircraft[8]. The factor  $e$  is calculated by using the following equation.

$$e = \frac{1}{K} \cdot \frac{1}{\pi AR}, \quad (2)$$

where, the drag-due-to-lift factor  $K$  is defined as a leading coefficient of the quadratic approximation function due to  $C_L$  under the consideration of  $C_D$

as function of  $C_L$ .

$$C_D = C_{D_0} + K \cdot (C_L - C_{L_0})^2. \quad (3)$$

$C_{D_0}$  denotes  $C_D$  caused by the other drag mechanisms.  $C_{L_0}$  is physically caused by the vertical asymmetry such as a cambered wing and a finite angle of incidence. When the lift of a wing is elliptically distributed along the span,  $K$  is defined to be 1.  $AR$  denotes the aspect ratio of the main wing, whose value is summarized in Table 1.

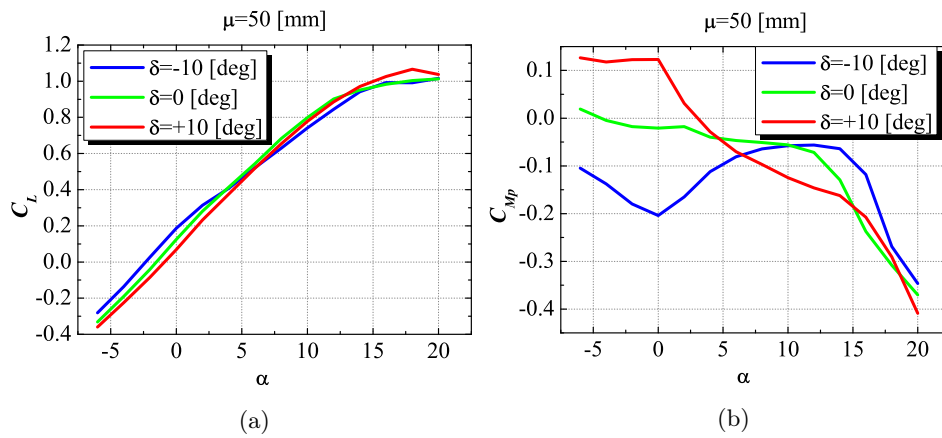
Figure 4 shows the polar curves under the several conditions. Figure 4(a) shows the repeatability of the polar curve for the clean configuration implemented three times on different days. Since the three lines precisely correspond each other, the reproducibility of the present experiment can be elucidated. When the  $K$  is calculated by using eq. (3) for the average of three data shown in Fig. 4(a), the wind tunnel model without the vertical control device found to be  $e = 0.6505$ . Note that the correlation between the dotted line and the other three lines in Fig. 4 shows the accuracy of  $K$ . Figure 4(a) shows that the curve generated by the quadratic approximation function exactly describes the polar curves by the experiment.

Figures 4(b), (c), and (d) respectively show the polar curves by changing  $\delta$  from  $-10$  to  $+10$  [deg] under the conditions of  $\mu$  of 50 [mm], 130 [mm], and 170 [mm]. The dotted curve is quadratic approximation as eq. (2) with the points of  $-4 \leq \alpha \leq 12$  [deg]. Figure 4 reveals that the shape of polar curve becomes similar to that for the clean configuration as  $\mu$  is larger. The curvature of polar curve becomes larger as  $\mu$  is smaller. Although  $C_D$  is moved to right direction due to  $C_D$  by the vertical control device, the geometry of the polar curve is similar in the case of  $\mu$  of 170 [mm]. Although  $C_D$  at  $\delta = 0$  is found to be low around low angle of attack, there are  $\delta$  that gives larger  $C_L/C_D$  than that of  $\delta = 0$ , when  $\alpha$  is higher than 6 [deg]. When optimum  $\delta$  is selected according to the angle of attack, the data is on the envelope curve and  $e$  will be improved. The results based on this procedure are summarized in Table 3. In both cases of  $\mu = 130$  and 170 [mm],  $e$  is improved. Especially, it is almost the identical as the clean configuration for the case of  $\mu = 170$  [mm].

On the other hand, in the cases of  $\mu = 50$  and 90 [mm], there is not as much improvement as cases of  $\mu = 130$  and 170 [mm]. In Fig. 4(d), the case of  $\delta = 0$  gives the best  $C_L/C_D$  except the cases of high angle of attack. In Fig. 5, there is considerably the interference between  $\delta$  and  $C_{Mp}$ , and also between  $\delta$  and  $C_L$ . When  $\delta$  is positive value,  $C_L$  tends to be lower and  $C_{Mp}$  tends to be higher. In contrast, the negative  $\delta$  oppositely affects on  $C_L$  and  $C_{Mp}$ . The reason of these effects is that the vortex generated

Table 3: Comparison of  $e$  for several experimental conditions.

$\mu$ [mm]	max $e$ [-]	
	$\delta = 0$ (fixed)	$\delta$ (variable)
clean	0.6505	
50	0.4531	0.4273
90	0.4535	0.4718
130	0.4875	0.5633
170	0.5471	0.6513

Figure 5: Comparison of the aerodynamic performance of the installed configuration at  $\mu = 50$  [mm]. (a)  $C_L$ - $\alpha$  and (b)  $C_{Mp}$ - $\alpha$ .

from the tip of the vertical control device passes in the vicinity of the tail wings, when  $\mu$  is small value such as  $\mu = 50$  [mm]. Changing the value of  $\delta$  from positive to negative reverses the rotational direction of the tip vortex by the vertical control device so that the interference for  $C_L$  and  $C_{Mp}$  is also opposite. There was little improvement on  $e$  under the condition of  $\mu = 50$  [mm] configuration because the positive effect of  $\delta \geq 0$  and negative effect of the tip vortex on the vertical control device shown in Fig. 5(a) cancelled each other.

## 4 Data-mining technique

In the present study, a self-organizing map (SOM)[5] is selected as a data-mining technique because the primary objective of data mining is the ac-

quisition of global design information in order to implement the structuring of design space. The previous study[1] indicated that SOM extracted the global design information for whole design space. The distinguishing feature of SOM is the generation of a qualitative description. The advantage of this method contains the intuitive visualization of two-dimensional colored maps of design space using bird's-eye-views. As a result, SOM reveals the tradeoffs among objective functions. Moreover, SOM addresses the effective design parameters and also reveals how a specific design parameter gives effects on objective functions and other design characteristics. One SOM is colored for one variable of objective function, design parameter, and other characteristic value so that the coloration pattern is compared with each other. Therefore, data mining using SOM might have a disadvantage to overlook important correlation in the problem with a large number of objective functions and design parameters. Since the present study has a total number of 9 at most among the design requirements, design parameters, and other variables that the influence will be observed, SOM is sufficient for the data mining manner.

In the present study, SOMs are generated by using commercial software Viscovery<sup>®</sup> SOMine 4.0 plus produced by Eudaptics, GmbH[2]. The uniqueness of the map generated by SOMine is assured due to Kohonen's Batch SOM algorithm and search of the best-matching unit for all input data and

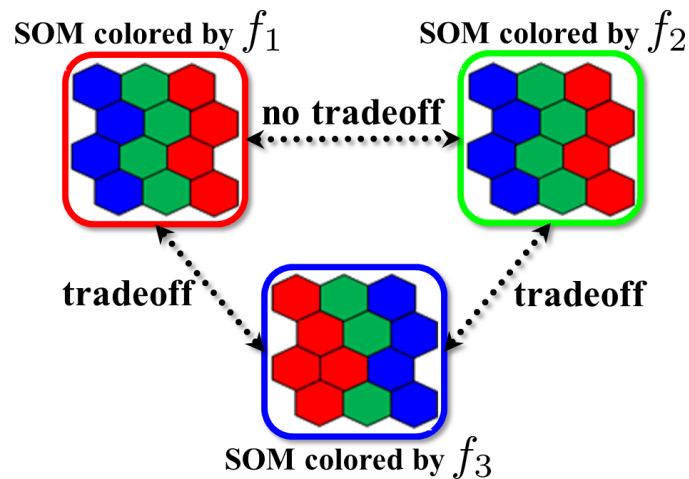


Figure 6: Comparison example of colored SOMs for minimization problem with three objective functions as  $f_1$ ,  $f_2$ , and  $f_3$ . Red describes high value and blue is low one.

the adjustment of weight vector near the best-matching unit. The decoding manner of SOM is briefly explained by using Fig. 6. This figure is assumed to be SOMs colored by three objective functions on the minimization problem of three objective functions. The generated SOM is made from hexagonal grid, which has the values of objective functions and design parameters as a vector quantity. Grids are distributed on a two-dimensional rectangular surface by the affinity of each objective-function value. Thereupon, grids with high affinity of each objective-function value clusters around a grid. There is no physical import on the vertical and horizontal lines of SOM. The comparison among SOMs to be colored by each vector quantity in each grid intuitively reveals the correlations among each vector quantity. There is similar coloration pattern between SOMs for  $f_1$  and  $f_2$  shown in Fig. 6. This comparison shows that one objective function absolutely has a low value, when another objective function has low value. Moreover, one objective function absolutely has high value, when another objective function has high value. That is, this comparison indicates that there is no tradeoff between  $f_1$  and  $f_2$ . On the other hand,  $f_3$  absolutely becomes large, when  $f_1$  becomes small, and vice versa. This comparison proves to be a severe tradeoff between  $f_1$  and  $f_3$ .

## 5 Data-mining result

The coloration pattern of SOM depends on indicator. Multiobjective optimization problems generally use objective functions as the indicator to generate SOM. However, both of the design requirements, i.e.,  $C_L$ ,  $C_D$ , and  $C_{Mp}$  and the design parameters have a major role in the present problem. Thereupon, as the first step, the SOM which the design requirements take charge of the indicator will be observed. As the second step, the SOM which the design parameters take charge of the indicator will be observed in this chapter. The especial design parameters to improve the aerodynamic performances will be specified so as to address the experimental condition and to efficiently reveal the flow mechanism.

### 5.1 Case to generate using design requirements

Figure 7 shows the SOM generated by the values of the three design requirements. As this SOM learning is implicated based on the values of the design requirements as the indicator for the similarity on the neural network, the SOMs colored by the design requirements have absolutely gradation shown in Fig. 7(a). The SOM colored by design requirement can generally indicate

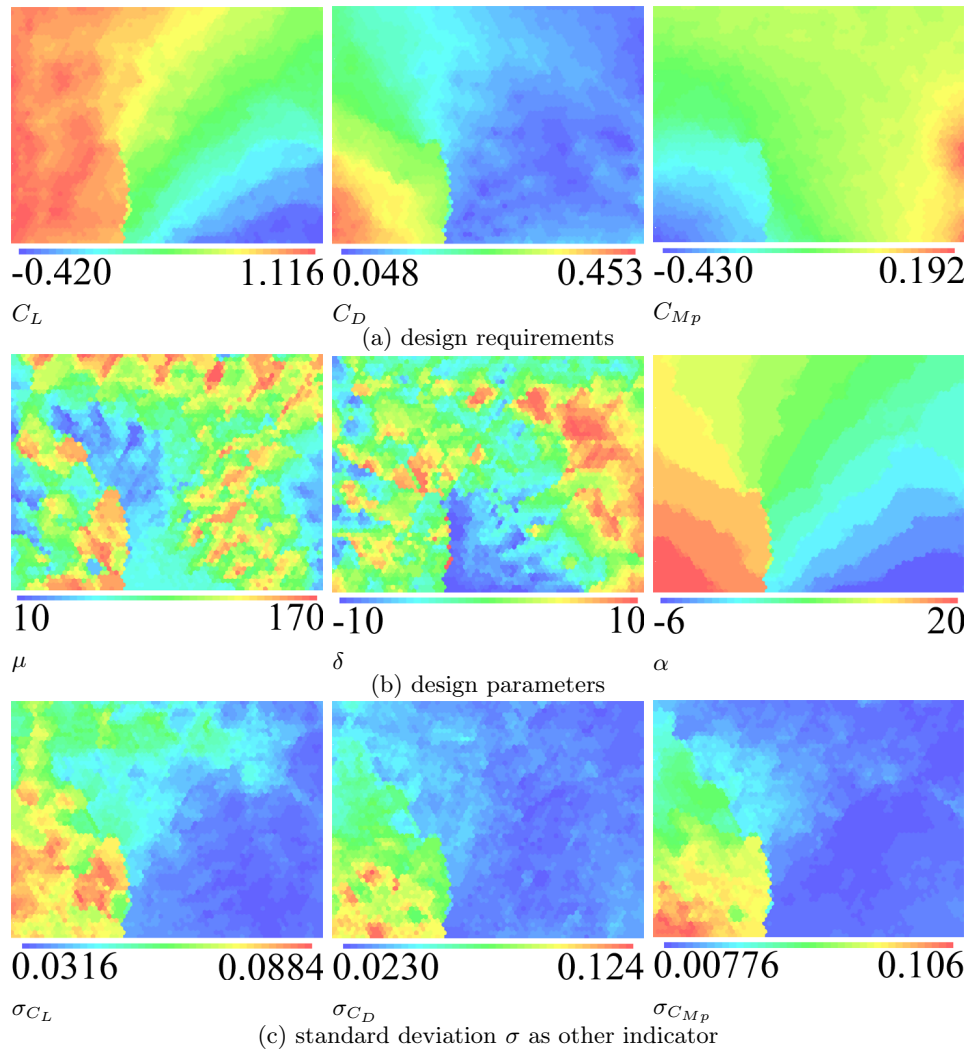


Figure 7: SOM generated by design-requirements values.

not only tradeoff information but also optimum and pessimum direction on SOM due to the gradation. In addition, the directions of the influence of design parameters for design requirements can be observed by comparison between the SOMs colored by the design requirements and those by the design parameters.

The SOMs colored by  $C_L$  and  $C_D$  in Fig. 7(a) reveal that there is a tradeoff between them. However, coloration patterns of  $C_L$  and  $C_D$  for



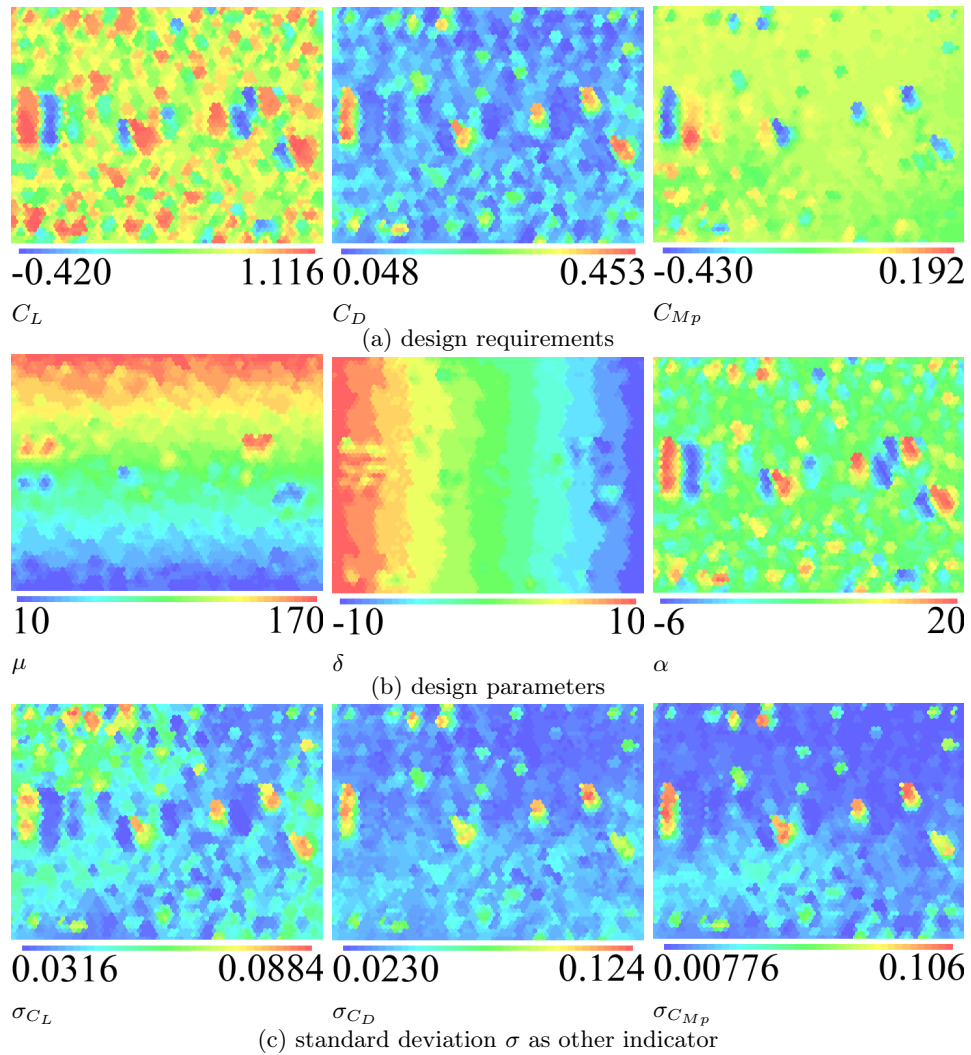


Figure 8: SOM generated by design-parameters values.

both the maximum and minimum directions are different. The compromise design region can be relatively found out on the SOM. The SOM colored by  $C_{Mp}$  in Fig. 7(a) reveals that the SOM's region to be the low value of  $C_{Mp}$  corresponds to that to be the high value of  $C_D$ . On the other hand, although the SOM's region to be the high value of  $C_{Mp}$  exists the bottom right on the SOM, the coloration pattern of it is unique. Note that  $C_{Mp}$  should be generally zero for the trim of the aircraft. The trim is practically gained by

controlling the elevators. Since the elevators of the present body are fixed, the present  $C_{Mp}$  cannot indicate the optimum and pessimum directions. Correlations between  $C_{Mp}$  and the other two aerodynamic characteristics as  $C_L$  and  $C_D$  are merely observed.

The SOMs colored by the three design parameters as  $\mu$  [mm],  $\delta$  [deg], and  $\alpha$  [deg] are shown in Fig. 7(b). The SOM colored by  $\mu$  reveals that  $\mu$  does not have direct influence on the three design requirements. Although there is a possibility that the combination between  $\mu$  and  $\delta$  gives the effects on the design requirements, Figs. 7(a) and (b) does not indicate it. The SOM colored by  $\delta$  reveals that the low value of  $\delta$  gives an effect on the low value of  $C_D$ . The high value of  $\delta$  does not directly give effects on the three design requirements. The SOM colored by  $\alpha$  reveals that the high value of  $\alpha$  directly affects on the high value of  $C_D$  and also the low value of  $\alpha$  directly gives an effect on the low value of  $C_L$ . Since  $\alpha$  generally has the effects on the aerodynamic performance, these results make sense. Since the coloration pattern shown in Figs. 7(a) and (b) depends on  $\alpha$ ,  $\alpha$  should be omitted so that the influences of  $\mu$  and  $\delta$  are observed.

Figure 7(c) shows the SOMs colored by the standard deviation  $\sigma$  for the three design requirements as  $C_L$ ,  $C_D$ , and  $C_{Mp}$ . The present  $\sigma$  is defined as the standard deviation for the data of 10,000 points for 10 [sec] in an experimental condition. These figures reveal that these have similar coloration pattern, and  $\sigma$  has high value when  $\alpha$  becomes high. This fact suggests that  $\sigma$  increases after the stall. The SOM generated by the three design parameters as  $\mu$ ,  $\delta$ , and  $\alpha$  is prepared in Fig. 8 in order to directly observe the influence of them on the three design requirements. The coloration patterns of  $C_L$  and  $C_D$  reveal that there is no regularity for those of  $\mu$  and  $\delta$ . That is, the coloration patterns of the design requirements indicate that the design requirements strictly depend on  $\alpha$ . Thereupon, the influence of  $\alpha$  on the three design requirements should be erased in order to directly observe the influence of  $\mu$  and  $\delta$ .

## 5.2 Case to generate using two design parameters as $\mu$ and $\delta$

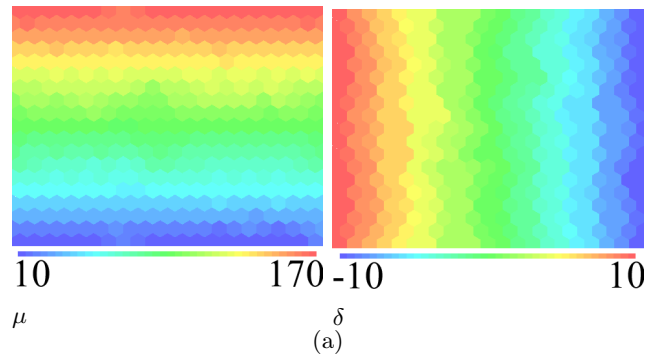
The SOM generated by  $\mu$  and  $\delta$  is shown in Fig. 9. Figure 9(a) shows the SOMs colored by  $\mu$  and  $\delta$  themselves, which are the straightforward coloration patterns. The coloration pattern for  $\mu$  is from upper to bottom and the upper region has high value of  $\mu$  and the bottom region has low value of  $\mu$ . On the other hand, the coloration pattern for  $\delta$  is from left to right. The left region has high value of  $\delta$  and the right region has low value

of  $\delta$ . Figures 9(b) to (o) show the SOMs colored by  $C_L$ ,  $C_D$ , and  $C_{M_p}$  for each  $\alpha$  from  $-6$  [deg] to  $20$  [deg] with  $2$  [deg] interval. The influence of the combination between  $\mu$  and  $\delta$  on each design requirement will be observed step by step. Note that the results of the latest calibration experiment of the wind tunnel balance show to ensure the sufficient accuracy of  $C_D$  for the narrow range of  $C_D$  in Fig. 9. Therefore, discussion which Fig. 9 is employed can be implemented because Fig. 9 has the significant difference of the design requirements.

### 5.2.1 Effectiveness on $C_L$

In the first place, influence on  $C_L$  will be observed. The effectiveness of the design parameters on  $C_L$  is roughly clustered for three  $\alpha$  regions as  $\alpha \leq 0$ ,  $2 \leq \alpha \leq 12$ , and  $\alpha \geq 14$  [deg].

In the case of  $\alpha \leq 0$  [deg], specific combinations of  $\mu$  and  $\delta$  give effects on  $C_L$ . The combinations of  $\mu \geq 140$  [mm] and  $\delta \geq 8$  [deg], and the  $\mu \leq 90$  [mm] and  $\delta \leq 0$  [deg] give the effect on increasing  $C_L$ . Effectiveness on  $C_L$  is stronger as  $\alpha$  is greater in the case of the former combination. On the other hand, the combinations of  $\mu \geq 150$  [mm] and  $\delta \leq -8$  [deg], and  $\mu \leq 40$  [mm] and  $\delta \geq 6$  [deg] give the adverse effect on decreasing  $C_L$ . The magnitude of the latter adverse effectiveness is stronger than that of the former one. The adverse effectiveness on  $C_L$  is weaker as  $\alpha$  increases in the former case. That is, the effectiveness on the increase of  $C_L$  in the case of high  $\mu$  is stronger as  $\alpha$  increases. Since the separation near the tip of the main wing is restrained when the vertical control device is in the vicinity of there,  $C_L$  increases. In addition, the main wing generates the positive  $C_L$  at greater than  $\alpha_{C_{L_0}}$ . The clean configuration does not have this effectiveness. On the other hand, the latter adverse effectiveness is independent on  $\alpha$ . When the vertical control device with  $+\delta$  installs in the vicinity of the fuselage, the fuselage and the



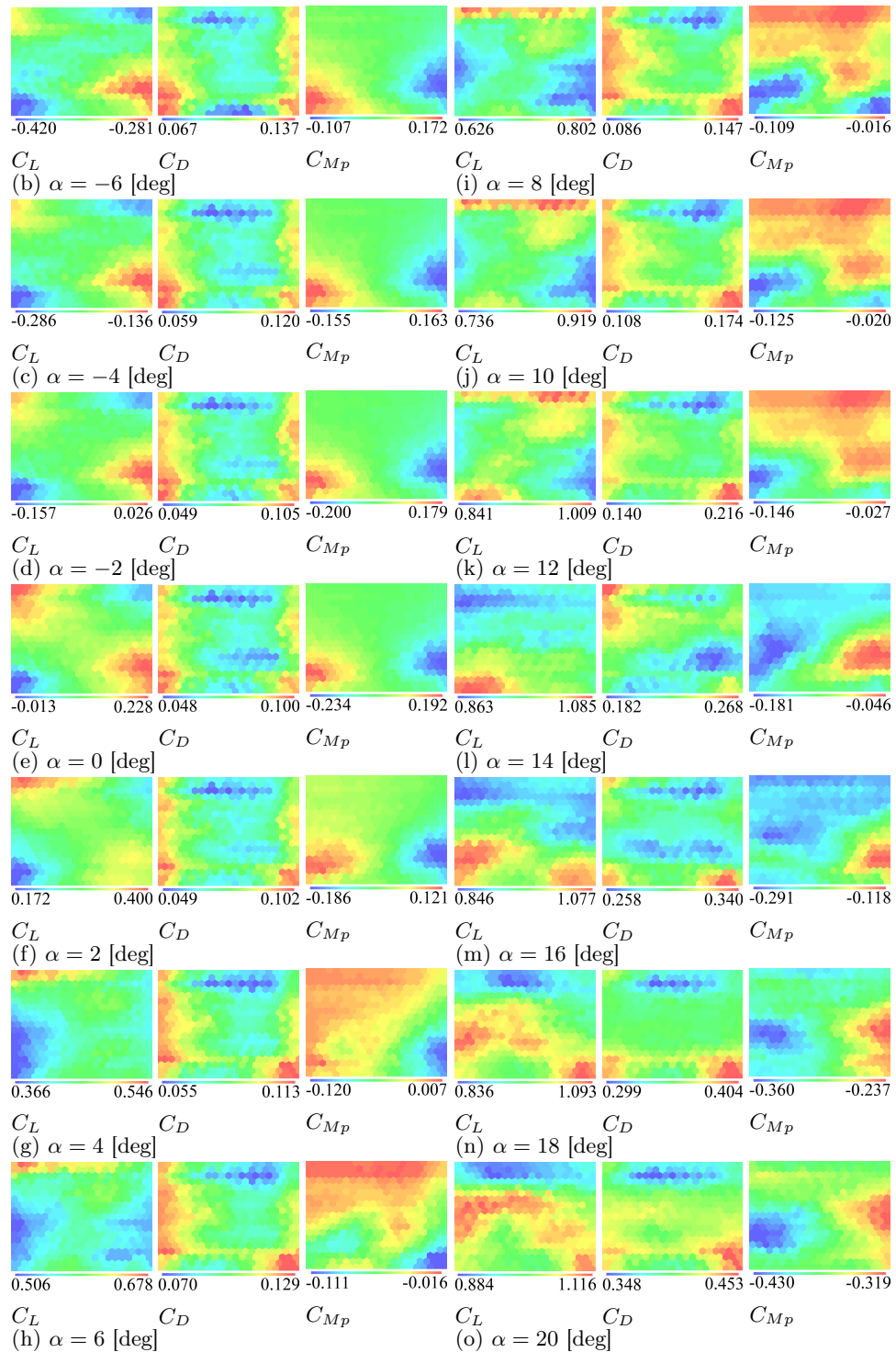


Figure 9: SOMs, (a) colored by each value of the two design parameters as  $\mu$  and  $\delta$ , (b) to (o) colored by the design requirements at each  $\alpha$ .

device generate quasi-throat flow. Since it is difficult to flow on the main-wing region where the fuselage and the device sandwich, this region is not functioning as a wing. Therefore,  $C_L$  is reduced as much.

In the case of  $2 \leq \alpha \leq 12$  [deg], the combinations between  $\mu \geq 160$  [mm] and  $\delta \geq 0$  [deg] at  $2 \leq \alpha \leq 6$  [deg], and between  $\mu \geq 160$  [mm] and all  $\delta$  at  $\alpha \geq 8$  [deg] give the effect on increasing  $C_L$ . When the vertical control device installs near the wing tip, the device functions as winglet. Therefore,  $C_L$  increases under the condition. The effectiveness under the condition that the vertical control device is near the tip disappears in the case of over  $\alpha$  of 14 [deg] because of the stall. The combination between  $\mu \leq 100$  [mm] and  $\delta \geq 6$  [deg] affects on decreasing  $C_L$ . This effectiveness is weaker as  $\alpha$  increases. The combination between  $\mu \leq 90$  [mm] and  $\delta \leq -2$  [deg] also affects on decreasing  $C_L$ . This effectiveness is stronger as  $\alpha$  increases. When the vertical control device installs around the middle of the wing, the device discourages the wing function. Since the wetted area of the vertical control device for the uniform flow is especially larger as  $|\delta|$  becomes larger, the adverse effectiveness on  $C_L$  is strong.

In the case of  $\alpha \geq 14$  [deg],  $\mu \geq 140$  [mm] affects on decreasing  $C_L$ . Especially,  $\delta \geq 0$  [deg] at  $\alpha$  of 14 [deg],  $\delta \geq 0$  [deg] and  $\delta \leq -6$  [deg] at  $\alpha$  of 16 [deg], and  $\delta$  of roughly 0 [deg] at  $\alpha$  of 18 and 20 [deg] have this effectiveness. Since the vertical control device with  $+\delta$  in the vicinity of the wing tip amplifies the tip stall,  $C_L$  sharply decreases. On the other hand, the combination between  $\mu \leq 40$  [mm] and  $\delta \geq 4$  [deg] gives the effect on increasing  $C_L$ . The upper limit of  $\mu$  to increase  $C_L$  grows as  $\alpha$  increases. In addition, the combinations between  $\mu \leq 70$  [mm] and  $\delta \geq 0$  [deg] at  $\alpha$  of 16 [deg], between  $30 \leq \mu \leq 110$  [mm] and  $\delta \geq -2$  [deg] at  $\alpha$  of 18 [deg], and between  $50 \leq \mu \leq 130$  [mm] and  $\delta \geq -4$  [deg] also give the effect on increasing  $C_L$ . Since the vertical control device at the middle of the wing exists the inside of separation due to the stall, the device reduces the pressure of its wake. As a result,  $C_L$  increases. The combination between  $\mu \leq 50$  [mm] and  $\delta \leq -2$  [deg] at  $\alpha \geq 16$  [deg] gives the effect on increasing  $C_L$ . Since the vertical control device with  $-\delta$  maintains the wing tip vortex,  $C_L$  increases.  $C_L$  is easily increased by  $\mu$  and  $\delta$  in the case of high  $\alpha$ .

### 5.2.2 Effectiveness on $C_D$

In the second place, influence on  $C_D$  will be observed. The effectiveness of the design parameters on  $C_D$  is clustered for three  $\alpha$  region as  $\alpha \leq 12$  [deg],  $\alpha$  of 14 [deg], and  $\alpha \geq 16$  [deg]. However, since  $\mu$  primarily has the effectiveness on  $C_D$ , the results will be summarized by using  $\mu$ .

$\mu$  of 150 [mm] always gives the effect on decreasing  $C_D$ . The effectiveness is not dependent on  $\alpha$ . The combination between  $\mu$  of 150 [mm] and  $-6 \leq \delta \leq 4$  [deg] especially gives more powerful effect on decreasing  $C_D$ . The magnitude of this effectiveness is similar among the cases at  $\alpha \leq 4$  [deg] and  $\alpha \geq 16$  [deg]. The magnitude of this effectiveness of  $0 < \delta \leq 4$  [deg] is stronger at  $\alpha \leq 2$  [deg]. In contrast, that of  $-6 \leq \delta \leq 0$  [deg] is stronger at  $4 \leq \alpha \leq 12$  [deg]. The magnitude of this effectiveness at  $\alpha$  of 14 [deg] is the weakest due to the existence of another combination between  $\mu$  and  $\delta$  to reduce  $C_D$  more. The separation in the vicinity of the tip of the main wing will be restrained in the case at 150 [mm]. The flow visualization of the three-dimensional space should be additionally performed in order to reveal the physical mechanism that  $\mu$  of 150 [mm] has the effectiveness on reducing  $C_D$ .

$\mu \leq 20$  and 70 [mm] also give the effect on decreasing  $C_D$ . The effectiveness does not depend on  $\alpha$ . The separation which occurs due to the interference with the fuselage will be restrained in the case at  $\mu \leq 20$  [mm] position. On the other hand, the wake of the vertical control device interferes in the tip of the horizontal tail wing in the case at  $\mu$  of 70 [mm] position. Both of these cases should not have a large  $|\delta|$  because of the larger wetted area of the vertical control device for the uniform flow.

In contrast,  $\mu$  of 40 [mm] affects on increasing  $C_D$ . The influence does not depend on  $\alpha$ . Since the wake of the vertical control device interferes the horizontal tail wing, the  $C_D$  of it increases. The flow visualization of the wake of the device should be implemented. In addition,  $C_D$  of each component should be elucidated by using computational fluid dynamics analysis.

$\delta \geq 8$  [deg] and  $\delta \leq -8$  [deg] affects on increasing  $C_D$  although  $\mu$  of 70 and 150 [mm] restricts the influence because the wetted area of the vertical control device for the uniform flow becomes large. Thereupon, a large number of  $|\delta|$  such as  $\delta \geq 8$  [deg] and  $\delta \leq -8$  [deg] should not be set in order to reduce  $C_D$ .

The case of  $\alpha$  of 14 [deg] has unique effectiveness on  $C_D$ . The combination between  $\mu$  around 60 [mm] and  $\delta$  of  $-4$  [deg] gives the effect on decreasing  $C_D$ . Since the wake of the vertical control device interferes the tip of the horizontal tail wing, the flow around the horizontal tail wing will be changed. On the other hand, the combination between  $\mu$  of 10 [mm] and  $\delta \geq 8$  [deg] and  $\delta \leq -6$  [deg] affects on increasing  $C_D$  in the case of  $\alpha \geq 16$  [deg]. The wing tip vortex is broke down because the vertical control device interferes it. The circumstantial physical mechanism to give the influence on  $C_D$  should be elucidated by using the flow visualization.

### 5.2.3 Effectiveness on $C_{Mp}$

In the third place, influence on  $C_{Mp}$  will be observed. The effectiveness of the design parameters on  $C_{Mp}$  is clustered for three  $\alpha$  regions as  $\alpha \leq 2$ ,  $4 \leq \alpha \leq 12$ , and  $\alpha \geq 14$  [deg], whose clustering is similar to that for  $C_L$ . The influence on  $C_{Mp}$  is easily understood because it depends on  $\alpha$ .

In the case of  $\alpha \leq 2$  [deg], the combination between  $\mu$  around 50 [mm] and  $\delta \geq 4$  [deg] affects on increasing  $C_{Mp}$ . On the other hand, the combination between  $\mu$  around 60 [mm] and  $\delta \leq -8$  [deg] affects on decreasing  $C_{Mp}$ . This change of  $C_{Mp}$  is explained by the function on the main-wing region where the fuselage and the vertical control device sandwich, that is similar mechanism of  $C_L$ .

In the case of  $4 \leq \alpha \leq 12$  [deg], the effectiveness is clustered by using  $\mu$ . The case of  $\mu \geq 150$  [mm] affects on increasing  $C_{Mp}$ . It is independent of  $\delta$ . Since this area on SOM has large  $C_L$  and small  $C_D$ ,  $C_{Mp}$  naturally increases. The combination between  $\mu$  around 50 [mm] and  $\delta \geq 4$  [deg] affects on decreasing  $C_{Mp}$ . The result is occurred by the similar mechanism in the above case of  $\alpha \leq 2$  [deg]. The combination between  $\mu \leq 90$  [mm] and  $\delta \leq -4$  [deg] also affects on decreasing  $C_{Mp}$ . Since the wake of the vertical control device interferes the tip of the horizontal tail wing, the tip vortex of the horizontal tail wing is induced. Therefore, the total  $C_{Mp}$  is reduced.

In the case of  $\alpha \geq 14$  [deg], the combination between  $50 \leq \mu \leq 80$  [mm] and  $\delta \leq -4$  [deg] affects on increasing  $C_{Mp}$ . This is caused by the interference of the wake of the device with the tip of the horizontal tail wing. On the other hand, the combination between  $40 \leq \mu \leq 70$  [mm] and  $\delta \geq 4$  [deg] affects on decreasing  $C_{Mp}$ . The result is occurred by the similar mechanism in the above case of  $\alpha \leq 2$  [deg]. Moreover,  $\mu \geq 140$  [mm] also affects on decreasing  $C_{Mp}$  except for the case of  $\alpha$  of 20 [deg]. This does not depend on  $\delta$ . The result is induced by decreasing  $C_L$ .

$C_{Mp}$  directly depends on  $C_L$ ,  $C_D$ , and  $\alpha$ . In addition, the trim of the aircraft is practically gained to control elevators. Thereupon, it is considerable that the design knowledge regarding  $C_L$  and  $C_D$  is primary and the design knowledge regarding  $C_{Mp}$  is secondary.

## 6 Conclusions

The new concept to place the vertical airfoil device as control surface has arrived in so as to improve the aerodynamic performance. The wind tunnel experiment has been implemented in order to investigate the influence of the vertical control device with symmetrical airfoil shape. Moreover, data min-

ing has been performed by using a self-organizing map for the experimental data in order to qualitatively reveal the correlations among the aerodynamic performances and the design parameters to place the vertical control device. Consequently, it has been revealed the correlations among them. Furthermore, there is a sweet spot, where is at  $\mu$  around 150 [mm] and  $-4 \leq \delta \leq 4$  [deg], in the present design space. In addition, the especial design parameters to improve the aerodynamic performance have been specified by using the data mining so that the detailed flow condition is observed. The three-dimensional geometry of vertical control device in the sweet spot will be optimized as the subsequent design phase based on the extracted design knowledge.

## Acknowledgment

The present study was supported by Japan Society for the Promotion of Science through a Grant-in-Aid for Challenging Exploratory Research 26630440.

## References

- [1] K. Chiba and S. Obayashi. Knowledge discovery in aerodynamic design space for flyback-boosted wing using data mining. *Journal of Spacecraft and Rockets*, 45(5):975–987, 2008.
- [2] G. Deboeck and T. Kohonen. *Visual Explorations in Finance with Self-Organizing Maps*. London, Springer Finance, 1998.
- [3] M. Fujino. Design and development of HondaJet. *Journal of Aircraft*, 42(3):755–764, 2005.
- [4] M. Fujino and Y. Kawamura. Wave-drag characteristics of an over-the-wing nacelle business-jet configuration. *Journal of Aircraft*, 40(6):1177–1184, 2003.
- [5] T. Kohonen. *Self-Organizing Maps*. Springer, Berlin, Heidelberg, 1995.
- [6] J. C. Lin. Review of research on low-profile vortex generators to control boundary-layer separation. *Progress in Aerospace Sciences*, 38(4):389–420, 2002.



- [7] T. Omori, Y. Sunada, and T. Imamura. Experimental and numerical research on aerodynamic characteristics of rectangular fin mounted vertically over the wing. AIAA Paper 2015-0774, 2015, 2015.
- [8] O. Samoylovitch and D. Strelets. Determination of the Oswald efficiency factor at the aeroplane design preliminary stage. *Aircraft Design*, 3(3):167–174, 2000.
- [9] D. A. Solfelt and R. C. Maple. CFD analysis of a T-38 wing fence. AIAA Paper 2008-0331, 2005.

## **Evidence-theory-based analysis for structural-acoustic field with epistemic uncertainties**

**†Jian Liu<sup>1</sup>, \*Longxiang Xie<sup>2</sup>, Xianfeng Man<sup>3</sup>, and Yongchang Guo<sup>4</sup>**

<sup>1</sup>State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,  
Hunan University, Changsha, Hunan 410082, China.

<sup>2</sup>State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,  
Hunan University, Changsha, Hunan 410082, China.

<sup>3</sup>State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,  
Hunan University, Changsha, Hunan 410082, China.

<sup>4</sup>State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body,  
Hunan University, Changsha, Hunan 410082, China.

\*Presenting author: d\_x20140920@126.com

†Corresponding author: JianL2004@126.com

### **Abstract**

Evidence theory has a strong capacity to deal with epistemic uncertainty, in view of the overestimation in interval analysis, the responses of structural-acoustic problem with epistemic uncertainty could be untreated. In this paper, a numerical method is proposed for structural-acoustic system response analysis under epistemic uncertainties based on evidence theory. To improve the calculation accuracy and reduce the computational cost, the interval analysis technique and radial point interpolation method are adopted to obtain the approximate frequency response characteristics for each focal element, and the corresponding formulations of structural-acoustic system for interval response analysis are deduced. Numerical examples are introduced to illustrate the efficiency of the proposed method.

**Key words:** Structural-acoustic system response analysis; Evidence theory; Radial point interpolation method; Interval analysis; Finite element method; Epistemic uncertainty

### **Introduction**

In the last two decades, with the increasing of people's interest in the performance of NVH (noise, vibration and harshness), researches on the structural-acoustic field have been experienced a rapid development in engineering [1-3]. In most engineering cases, the structural-acoustic problems have been analysed by Probabilistic methods, in which the probability distribution, the boundary conditions and the external loads are defined unambiguously. However, due to the effects of manufacturing/assembling errors, original algorithm defect, imprecise environment factors and external excitations, uncertainties associated with geometric tolerances, material properties

and boundary conditions are unavoidable [4,5]. Generally, uncertainty can be divided into epistemic and aleatory categories based on the source of uncertainty. Epistemic uncertainty is related to the incomplete knowledge or imprecise information in any activity, which can be reduced by collecting more knowledge or experimental data. Aleatory uncertainty, on the other hand, derives from inherent variation in a physical system or environment, which is always regarded as random variables in probability theory [6]. Numerous mathematical theories or methods are developed to deal with the structural-acoustic problems under epistemic uncertainties, including possibility theory, D-S evidence theory, Bayesian theory, interval analysis, p-box method, Monte-Carlo method, spectral stochastic method, etc [7-9].

Among the approaches above, evidence theory seems to be more capable or more flexible to define epistemic uncertainty in the practical engineering problems. According to the D-S theory, it defines BPAs (basic probability assignment) to each focal element, which can provide corresponding formulations as possibility theory. Besides, the basic axioms in evidence theory can also deal with hybrid uncertainties in which aleatory and epistemic uncertainties combined in a very natural way. Thus, evidence theory has been widely used in artificial intelligence related fields and has been extended to conduct engineering structures and mechanical systems design, and reliability analysis, recently. The benefits and drawbacks of evidence theory in reliability analysis were summed by Oberkampf and Helton through a simple algebraic function [10]. An evidence-theory-based reliability analysis method was developed by Jiang et al., in which the concept of focal element was proposed firstly [11,12]. H. R. Bae proposed an efficient method based on evidence theory for reliability analysis using a multi-point approximation [13,14]. Helton et al. combined evidence theory with sampling-based sensitivity analysis when determining the epistemic uncertainty in model inputs [15]. A non-probability convex model was created by Elishakoff et al. to handle uncertain problems without sufficient information [16]. Qiu et al. proposed an interval perturbation method for narrow parameter intervals due to the unpredictable effect of neglecting the higher order terms of Taylor series or Neumann series [17]. An exploration of evidence theory has been conducted by J. C. Helton by using three uncertain quantification methods to address the challenge problems at model predictions [18]. An evidence-theory-based interval method was proposed by Rao et al. to analyse uncertain structural systems [19]. The application of fuzzy set theory in finite element method had developed the fuzzy finite element method (FFEM) for non-deterministic models [20–23]. Bae et al. applied an efficient method under a multi-point approximation to process evidence-theory-based reliability analysis [24,25]. The evidence theory and Bayesian theory were used for decision-making problems to compare the effectiveness of uncertainty quantification [26].

The response characteristics of structural-acoustic system is one of the hot points in noise prediction, which is important for NVH performance in engineering design and manufacturing [27]. From the works above, some inspiring progresses have been made for the response analysis of structure-acoustic coupling system with epistemic

uncertainties and evidence-theory-based reliability analysis. However, from an overall perspective, research on the hybrid uncertain analysis and response characteristics of complex system are still at the very beginning. Moreover, some crucial issues have not yet been solved [28]. Traditional numerical methods for the structural–acoustic problems are possibility theory or FEM (Finite Element method) in which the parameters are always regarded as random variables and the probability distributions are defined unambiguously. This assumption would ignore the influence of uncertainty and correlation in complex system [29].

In this paper, an evidence-theory-based radial point interpolation method (DSRPIM) is proposed for structure-acoustic coupling system under epistemic uncertainties, which can acquire the frequency response characteristics of complex system. The remainder of this paper is organized as follows. In chapter 2, the fundamentals of evidence theory are introduced. The equilibrium equation for structure-acoustic coupling system is deduced in chapter 3. In chapter 4, DS-RPIM is proposed to predict the frequency response characteristics of structural-acoustic problems. Two numerical examples are investigated in chapter 5. In chapter 6, some conclusions are given.

## 1. Evidence theory

### 1.1. Fundamentals of evidence theory

Evidence theory, also called as DS (Dempster-Shafer) theory, was firstly introduced by Dempster through studying statistical problems in 1976. And further developed by Shafer who defined probability to make it more suitable for general cases [30]. Compared with probability theory, evidence theory uses a prior probability distribution to get a posterior evidence interval, which quantifies the belief and plausibility of each proposition to handle the uncertainty in system response.

As probability theory, evidence theory firstly defines FD (a frame of discernment)  $\Theta$ , which contains a set of mutually exclusive propositions.  $2^\Theta$  is a non-blank finite set that always denotes the power set of  $\Theta$ , which means all possible various propositions. For example, if the frame of discernment  $\Theta$  includes three mutually exclusive elementary propositions  $X_1$ ,  $X_2$  and  $X_3$ , the power set of  $\Theta$  can be illustrated as follows

$$2^\Theta = \{\emptyset, \{X_1\}, \{X_2\}, \{X_3\}, \{X_1, X_2\}, \{X_1, X_3\}, \{X_2, X_3\}, \{X_1, X_2, X_3\}\} \quad (1)$$

In evidence theory, the probability is assigned not only to a single matter but also to any subset of possible propositions.  $m: 2^\Theta \rightarrow [0, 1]$ , called as the BPAF (basic probability assignment function) of  $\Theta$ , defines the elementary belief of each proposition, which should satisfy the following three theorems

Theorem 1:  $m(A) \geq 0$  for any  $A \in 2^\Theta$

Theorem 2:  $m(\emptyset) = 0$

Theorem 3:  $\sum_{A \subset \Theta} m(A) = 1$

where  $m(A)$  represents the corresponding BPAs of  $A$ . And every set  $A$  satisfying  $m(A) > 0$  be defined as a focal element.

It is hard to construct a precise PDF (probability density function) for proposition  $A$  because of the insufficient information or knowledge. Thus, it seems more reasonable to provide a confidence interval instead of a deterministic value to depict the total degree of belief in a proposition. In general, evidence theory uses the belief and plausibility to quantify the lower and upper bounds of an interval  $[Bel(A), Pl(A)]$ , which is defined as

$$Bel(A) = \sum_{B \subseteq A} m(B) \quad (\forall A \subset \Theta) \quad (2)$$

$$Pl(A) = \sum_{A \cap B \neq \phi} m(B) \quad (3)$$

where  $Bel: 2^\Theta \rightarrow [0, 1]$  is called as belief which is obtained by adding the evidence of propositions in  $A$ . Meanwhile,  $Pl: 2^\Theta \rightarrow [0, 1]$  is the summation of BPAs that belong to the propositions of  $A$  totally or partially, which is defined as the Plausibility function of  $\Theta$ .

### 1.2. Characteristic function with interval variables based on DS theory

Considering a general function with  $q$ -dimensional independent variables

$$Y = f(X) \quad X_i \in X, \quad i = 1, 2, \dots, q \quad (4)$$

Similar to the probability theory, the uncertain parameters are generally seen as relatively independent and the joint frame of discernment  $S$  is defined as

$$S = X_1 \times X_2 \times \dots \times X_q = \{s_k = [x_1, x_2, \dots, x_q], \quad x_j \in X_j, \quad j = 1, 2, \dots, q\} \quad (5)$$

where  $s_k$  and  $x_j$  represent the focal element of joint FD and the focal element of the  $j$ th evidence variable, respectively. The joint BPAs can be expressed as

$$m_s(s_k) = \left\{ \begin{array}{l} \prod_{j=1}^q m(x_j) \\ 0, \text{ otherwise} \end{array} \right\} \quad (6)$$

In probability theory, the mean value  $E(X)$  and the evidence variable  $X_i^I$  are relatively independent. However, the evidence variable  $X_i^I$  is an interval rather than a deterministic value. Thus,  $E(X)$  and  $X_i^I$  are related rather than independent. Based on the concepts mentioned above, considering the overestimation phenomenon in interval analysis, the characteristic function of evidence variables are provided below [31].

### 1.2.1. The relevant expectance $E(X)$

Through the analysis above, the relevant expectance  $E(X_i)$  is expressed as

$$E(X_i) = \sum_{\substack{j=1 \\ j \neq i}}^n X_j^l / n \quad (7)$$

where  $n$  is the amount of evidence variables.  $X_j^l$  is the others except  $X_i^l$ .

### 1.2.2. The relevant variance $D(X)$

Similar to the expectance  $E(X)$ , the overestimation characteristics is also existed in the variance  $D(X)$  calculation. To eliminate the phenomenon above, expanding the  $E(X)$ , the variance formula is defined as

$$D(X) = \sum_{i=1}^n (X_i^l - (X_1^l m(X_1^l) + X_2^l m(X_2^l) + \dots + X_{i-1}^l m(X_{i-1}^l) + X_i^l m(X_i^l) + X_{i+1}^l m(X_{i+1}^l) + \dots + X_n^l m(X_n^l)))^2 m(X_i^l) \quad (8)$$

where  $X_i^l$  is the  $i$ th evidence variable and  $m(X_i^l)$  is the corresponding BPAs.

Obviously,  $D(X)$  changes with the change of  $X_i$  in the interval  $[\text{Bel}(A), \text{Pl}(A)]$ . Thus, the relevant variance  $D'(X)$  can be defined as

$$D'(X) = \sum_{i=1}^n \left( \frac{n^2 - (1 + \beta)n}{n(n - \beta)} X_i^l - \left( \sum_{\substack{j=1 \\ j \neq i}}^n X_j^l / n \right) \right)^2 m(X_i^l) \quad (9)$$

where  $\beta$  is the interval correction coefficient and its range is from 0.01 to 0.30. The coefficient factor  $\partial$  is introduced to the relevant variance, which is expressed as

$$\partial = \frac{n^2 - (1 + \beta)n}{n(n - \beta)} \quad (10)$$

So, the relevant variance formula is rewritten as

$$D'(X) = \sum_{i=1}^n (\partial X_i^l - E(X_i))^2 m(X_i^l) \quad (11)$$

### 1.2.3. The relevant covariance $\text{Cov}'(X_1, X_2)$

Similarly, the co-relevant expectance  $E'(X_i)$  is introduced for covariance  $\text{Cov}(X_1, X_2)$ , which is defined as

$$E'(X_i) = \sum_{\substack{j=1 \\ j \neq i}}^n X_j^l m(X_j^l) \quad (12)$$

By introducing the coefficient factor  $\delta$  to covariance, the relevant covariance  $Cov'(X_1, X_2)$  can be expressed as

$$Cov'(X_1, X_2) = \sum_{i=1}^n \sum_{j=1}^m (\delta_1 X_{1i}^l - E'(X_{1i}^l)) (\delta_2 X_{2j}^l - E'(X_{2j}^l)) m(X_{1i}^l X_{2j}^l) \quad (13)$$

where  $\delta_1 = \frac{n}{n - \varepsilon_1} - m(X_{1i}^l)$ ,  $\delta_2 = \frac{m}{m - \varepsilon_2} - m(X_{2j}^l)$ ,  $\varepsilon_1$  and  $\varepsilon_2$  are the interval

combined coefficients whose range is from 0.01 to 0.25.

## 2. FEM/RPIM for structural-acoustic coupling system

In this paper, the coupled FEM/RPIM method is proposed to solve the structural-acoustic field problem, in which the FEM/RPIM model is used to simulate the plate structure and the acoustic medium. Due to the  $c_0$  continuity characteristic of fluid element, the Reissner-Mindlin plate is elected to the plate structure, in which the normals to the mid-plane of the plate remain straight during the deformation[32]. And the acoustic medium satisfies the linear constitutive equations which is assumed to be inviscid and incompressible. On the interface of the plate and the acoustic medium, only the acoustic medium exerts the normal loads on the plate and the normal displacement of the plate is just coupled with the acoustic medium[33].

### 2.1. FEM/RPIM model of the plate structure

In the frequency domain, without considering structural damping, the steady-state dynamic equation Galerkin weak form of the plate structure can be defined as

$$\begin{aligned} & \int_{\Omega} \delta \kappa^T D_b \kappa d\Omega + \int_{\Omega} \delta \gamma^T D_s \gamma d\Omega + \int_{\Omega} \delta \mu^T \rho t \omega^2 \ddot{\mu} d\Omega \\ & + \int_{\partial\Omega} \delta \mu^T t_s dS - \int_{\Omega} \delta \mu^T b_s d\Omega = 0 \end{aligned} \quad (14)$$

where  $\mu$  is the displacement,  $\ddot{\mu}$  is the acceleration,  $\rho$  is the material density,  $t$  is the thickness of plate element,  $t_s$  is the surface loading plate structure and  $b_s$  is volume force, respectively.

$\gamma$  and  $\kappa$  are the plate shear strain and bending strain, respectively, which can be expressed as:

$$\gamma = \left[ \frac{\partial w}{\partial x} - \theta_x \quad \frac{\partial w}{\partial y} - \theta_y \right]^T \quad (15)$$

$$\kappa = \left[ -\frac{\partial \theta_x}{\partial x} \quad -\frac{\partial \theta_y}{\partial y} \quad -\left( \frac{\partial \theta_x}{\partial y} = \frac{\partial \theta_y}{\partial x} \right) \right]^T \quad (16)$$

$D_s$  and  $D_b$  are the transverse shear stiffness constitutive matrix and the bending plate stiffness constitutive matrix, respectively, which are written as:

$$D_s = \frac{Et\nu}{2(1+\nu)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (17)$$

$$D_b = \frac{Et^3}{12(1-\nu^2)} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{1-\nu}{2} \end{bmatrix} \quad (18)$$

where  $E$  is the Young's modulus,  $\nu$  is the Poisson's ratio and  $\nu=5/6$  is the shear correction factor, respectively.

From Eqs.(15-18), we can get that  $\begin{bmatrix} \kappa \\ \nu \end{bmatrix} = \begin{bmatrix} B_b \\ B_s \end{bmatrix} \mu$ , the steady-state dynamic equation of the plate structure is defined as

$$Ku - M \ddot{u} = F_f + F_b \quad (19)$$

where  $K$  denotes the plate stiffness matrix which is given as

$$K = K_b + K_s = \int_{\Omega} (B_b)^T D_b B_b d\Omega + \int_{\Omega} (B_s)^T D_s B_s d\Omega \quad (20)$$

$K_b$  denotes the the bending stiffness matrix,  $K_s$  denotes the shear stiffness matrix,  $M$  denotes the plate element mass matrix,  $M$  is defined as

$$M = \int_{\Omega} \rho Q^T \text{diag} \left[ \frac{t^3}{12} \quad \frac{t^3}{12} \quad t \right] Q d\Omega \quad (21)$$

$F_f$  and  $F_b$  are the surface load matrix and a volume force array, which are expressed as

$$F_f = \int_{\partial\Omega} Q^T t_s dS \quad (22)$$



$$F_b = \int_{\Omega} Q^T b_s d\Omega \quad (23)$$

## 2.2. FEM/RPIM model for the acoustic medium

In the engineering application, the fluid is generally regarded as compressible and inviscid which is seen to undergo small translational movement[34]. Considering an acoustic field problem with domain  $\Omega_f$  and boundary  $\Gamma_b$ , the speed of sound  $c$  and the field acoustic pressure  $p$  are provided, the acoustic wave equation is defined as

$$\Delta p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0, \text{ in } \Omega_f \quad (24)$$

where  $\Delta$  is the Laplace operator,  $p$  is the field acoustic pressure,  $c$  and  $t$  are the speed of sound traveling in the fluid medium and its time, respectively.

The boundary condition of acoustic field is written as

$$\nabla p \cdot n = 0, \quad \text{on } F_b \quad (25)$$

where  $n$  denotes the boundary surface normal to the acoustic fluid domain.

On the interface between the plate structure and the acoustic medium, the momentum balance requires that

$$\nabla p \cdot n = -\rho \ddot{u}_f \text{ on } \Omega_f \quad (26)$$

where  $\rho$  is the density of acoustic medium,  $\ddot{u}_f$  is the normal acceleration component of acoustic fluid on the interface and  $\Omega_f$  is the interface between the plate structure and acoustic fluid.

If the acoustic pressure  $p$  is regarded as a time harmonic variable, the Eq.(24) can be re-written as

$$\nabla^2 p + k^2 p = 0 \quad (27)$$

where  $k = \omega / c$  represents the wavenumber,  $\omega$  is the angular frequency,  $c$  denotes the sound speed.

The smoothed Galerkin weak form for acoustic problems can be expressed as

$$-\int_{\Omega} \nabla \Psi \cdot \nabla \Psi P d\Omega + \frac{1}{c^2} \int_{\Omega} \Psi \cdot \Psi P d\Omega - \rho \int_{\partial\Omega_{sf}} \Psi P \cdot \ddot{u}_f d\Gamma - \int_{\Omega} \Psi \cdot \Psi P \cdot \frac{\partial q_f}{\partial t} d\Omega = 0 \quad (28)$$

where  $q_f$  is the additional load of unit volume and  $\Psi$  expresses the shape function matrix of FE-RPIM.

For numerical computation, the acoustic wave equation should be discretized by using the Radial Point Interpolation method[35]. This leads to the discretized equation of node sound pressure  $\mathbf{p}$  which is re-written as

$$\mathbf{p} = \sum_{i=1}^m N_{f_i} p_i = N_f \mathbf{P} \quad (29)$$

where  $\mathbf{p}$  is the vector of nodal pressure,  $m$  expresses the number of nodal variables per element, and  $N_f$  denotes the FE-RPIM shape function of fluid domain.

By substituting Eq.(29) into Eq.(28), the matrix form equation of acoustic domain can be obtained as

$$\mathbf{K}_f \mathbf{p} + \mathbf{M}_f \ddot{\mathbf{p}} = \mathbf{F}_f \quad (30)$$

where  $\mathbf{K}_f$  is the acoustic stiffness matrix and it can be expressed as

$$\mathbf{K}_f = \int_{\Omega} \mathbf{B}_f^T \mathbf{B}_f d\Omega \quad (31)$$

$\mathbf{B}_f$  denotes the smoothed gradient matrix that is defined as

$$\mathbf{B}_f = \begin{bmatrix} \Psi_{1,x} & \Psi_{2,x} & \cdots & \Psi_{M,x} \\ \Psi_{1,y} & \Psi_{2,y} & \cdots & \Psi_{M,y} \\ \Psi_{1,z} & \Psi_{2,z} & \cdots & \Psi_{M,z} \end{bmatrix} \quad (32)$$

$\mathbf{M}_f$  is the acoustic mass matrix and it is written as

$$\mathbf{M}_f = \frac{1}{c^2} \int_{\Omega} \Psi^T \Psi d\Omega \quad (33)$$

$\mathbf{p}$  denotes the nodal pressure of the acoustic domain, which can be expressed as

$$\mathbf{p} = \{p_1, p_2, \dots, p_n\}^T \quad (34)$$

$F_s$  and  $F_f$  are the vectors of nodal acoustic forces that are given as

$$F_s = \rho \int_{\partial\Omega_{sf}} \Psi^T \cdot \Psi d\Gamma \quad (35)$$

$$F_f = \int_{\Omega} \Psi^T \frac{\partial q_f}{\partial t} d\Omega \quad (36)$$

### 2.3. Coupled FEM/RPIM for structural-acoustic problem

Considering that the structural domain  $\Omega_s$  coupled with fluid domain  $\Omega_f$  on the interface  $\Omega_{sf}$ , the boundary conditions of structural-acoustic coupling system are denoted by  $\Gamma_b$ ,  $\Gamma_u$  and  $\Gamma_t$  which are illustrated in Fig.1. In this section, the coupled FEM/RPIM equation is proposed for structural-acoustic problem.

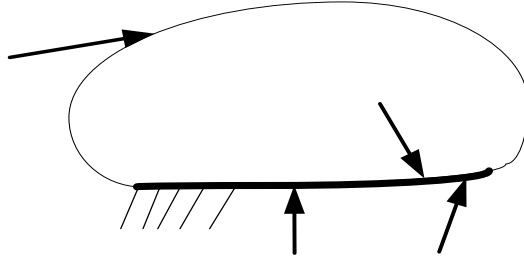


Figure 1. Schematic illustrating of the structural-acoustic system

The fluid particle and the structure move in the normal direction of the interface are written as

$$u_s \mathbf{n}_s = u_f \mathbf{n}_f \quad (37)$$

where  $\mathbf{n}$  is the normal vector,  $u_s$  is the displacement of structure on the interface and  $u_f$  is the displacement of fluid contacting the structure.

On the interface, based on the continuity and equilibrium conditions, we obtain that

$\mathbf{n} = \mathbf{n}_f = -\mathbf{n}_s$  [36]. The fluid force loading on the structure  $F_s$  can be expressed as

$$F_s = \int_{\Omega_{sf}} N_s^T \mathbf{n}_s \sigma_s d\Gamma_f = \int_{\Omega_{sf}} N_s^T \mathbf{n}_f p d\Gamma_f \\ = \left( \int_{\Omega_{sf}} N_s^T \mathbf{n}_f N_f d\Gamma_f \right) \mathbf{p} \quad (38)$$

The structural force loading on the fluid  $F_f$  is also expressed as

$$\begin{aligned}
 F_f &= -\rho \int_{\Omega_f} N_f^T \ddot{u}_f d\Gamma = -\rho \int_{\Omega_f} N_f^T \ddot{u}_s d\Gamma \\
 &= -\rho \left( \int_{\Omega_f} N_f^T n_f N_s d\Gamma \right) \ddot{u}_s
 \end{aligned} \tag{39}$$

The spatial coupling matrix  $\mathbf{H}$  can be defined as

$$\mathbf{H} = \int_{\Omega_f} N_s n_f N_f dS \tag{40}$$

By substituting the Eq.(40) into Eq.(38) and Eq.(39), the equations are rewritten as

$$\mathbf{F}_s = \mathbf{H}\mathbf{p} \quad \mathbf{M}_f \ddot{\mathbf{p}} = -\rho \mathbf{H}^T \ddot{\mathbf{u}}_s \tag{41}$$

Thus, the governing equation for coupled structure-acoustic system is expressed as

$$\begin{bmatrix} \mathbf{M} & 0 \\ \rho \mathbf{H}^T & \mathbf{M}_f \end{bmatrix} \begin{Bmatrix} \ddot{\mathbf{u}}_s \\ \ddot{\mathbf{p}} \end{Bmatrix} + \begin{bmatrix} \mathbf{K} & -\mathbf{H} \\ 0 & \mathbf{K}_f \end{bmatrix} \begin{Bmatrix} \mathbf{u}_s \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_s \\ \mathbf{F}_f \end{Bmatrix} \tag{42}$$

Assuming that the displacement and pressure are all time-harmonic[37], Eq.(42) can be rewritten as

$$\begin{bmatrix} \mathbf{K} - \omega^2 \mathbf{M} & -\mathbf{H} \\ \rho \omega^2 \mathbf{H}^T & \mathbf{K}_f - \omega^2 \mathbf{M}_f \end{bmatrix} \begin{Bmatrix} \mathbf{u}_s \\ \mathbf{p} \end{Bmatrix} = \begin{Bmatrix} \mathbf{F}_s \\ \mathbf{F}_f \end{Bmatrix} \tag{43}$$

To simplify the process of analyzing the FE/RPI equation of the structural-acoustic system[38,39], we rewrite Eq. (43) into the following form

$$\mathbf{Z}\mathbf{U} = \mathbf{F} \tag{44}$$

where  $\mathbf{Z}$  is the structural-acoustic dynamic stiffness matrix,  $\mathbf{U}$  is the response vector and  $\mathbf{F}$  is the external excitation vector which can be expressed as

$$\mathbf{Z} = \begin{bmatrix} \mathbf{K} - \omega^2 \mathbf{M} & -\mathbf{H} \\ \rho \omega^2 \mathbf{H}^T & \mathbf{K}_f - \omega^2 \mathbf{M}_f \end{bmatrix}, \quad \mathbf{U} = [\mathbf{u}_s \quad \mathbf{p}]^T, \quad \mathbf{F} = [\mathbf{F}_s \quad \mathbf{F}_f] \tag{45}$$

### 3. DS-FE/RPIM for epistemic uncertainty structural-acoustic problem

Discretizing the structural-acoustic coupling system, the discretization form of the structural-acoustic dynamic stiffness matrix  $\mathbf{Z}$  and the external excitation vector  $\mathbf{F}$  can be rewritten as

$$\mathbf{Z} = \begin{bmatrix} \sum_{i=1}^N K_i - \omega^2 \sum_{i=1}^N M_i & -\sum_{i=1}^N H_{fi} \\ \rho \omega^2 \left[ \sum_{i=1}^N H_{fi} \right]^T & \sum_{i=1}^{NA} K_{fi} - \omega^2 \sum_{i=1}^{NA} M_{fi} \end{bmatrix} \tag{46}$$

$$F = \left\{ \sum_{i=1}^N F_{si} \sum_{i=1}^{NA} F_{fi} \right\} \quad (47)$$

where  $N$  denotes the number of plate elements and  $NA$  denotes the number of acoustic field elements, respectively.

According to D-S evidence theory, the FPD of the dynamic stiffness matrix and the external excitation vector in evidence focal element are expressed as

$$\frac{\partial Z(X_{ik}^m)}{\partial X_{ik}} = \begin{bmatrix} \sum_{i=1}^N \frac{\partial K_i}{\partial X_{ik}} - \omega^2 \sum_{i=1}^N \frac{\partial M_i}{\partial X_{ik}} & - \sum_{i=1}^N \frac{\partial H_{fi}}{\partial X_{ik}} \\ \rho \omega^2 \left[ \sum_{i=1}^N \frac{\partial H_{fi}}{\partial X_{ik}} \right]^T & \sum_{i=1}^{NA} \frac{\partial K_{fi}}{\partial X_{ik}} - \omega^2 \sum_{i=1}^{NA} \frac{\partial M_{fi}}{\partial X_{ik}} \end{bmatrix} \quad (48)$$

$$F = \left\{ \sum_{i=1}^N \frac{\partial F_{si}}{\partial X_{ik}} \sum_{i=1}^{NA} \frac{\partial F_{fi}}{\partial X_{ik}} \right\}^T \quad (49)$$

where  $X_{ik}^m$  is the interval variable which denotes the  $k$ th focal element of the  $i$ th evidence variable.

Combined with the interval perturbation theory, ignoring the higher order perturbation[40], the approximate formula of node pressure response is defined as

$$p_{ik}^m = (Z_{ik}^m)^{-1} F_{ik}^m \quad (50)$$

$$\begin{aligned} \Delta p_{ik}^I &= (Z_{ik}^m)^{-1} (\Delta F_{ik}^I - \Delta Z_{ik}^I p_{ik}^m) \\ &= (Z_{ik}^m)^{-1} \left\{ \Delta X_{ik} \left( \frac{\partial F(X_{ik}^m)}{\partial X_{ik}} - \frac{\partial Z(X_{ik}^m)}{\partial X_{ik}} p_{ik}^m \right) \Delta e^I \right\} \end{aligned} \quad (51)$$

where  $p_{ik}^m$  is the node pressure and  $\Delta e^I = [-1, 1]$ .

According to Eq.(51), the estimated value of  $\Delta p_{ik}^I$  interval radius is expressed as

$$\Delta p_{ik} = \left| (Z_{ik}^m)^{-1} \Delta X_{ik} \frac{\partial F(X_{ik}^m)}{\partial X_{ik}} \right| + \left| (Z_{ik}^m)^{-1} \Delta X_{ik} \frac{\partial Z(X_{ik}^m)}{\partial X_{ik}} p_{ik}^m \right| \quad (52)$$

Based on the value range of evidence vector, under the effects of the evidence variable  $X_{ik}$ , the upper and lower bounds of the node pressure response value  $p_{ik}^m$  can be write as

$$p_{ik}^U = p_{ik}^m + \Delta p_{ik} \quad (53)$$

$$p_{ik}^L = p_{ik}^m - \Delta p_{ik} \quad (54)$$

By substituting Eq.(53) and Eq.(54) to Eq.(7), the expectance interval of the steady-state sound pressure response can be expressed as

$$E(p)^U = \sum_{i=1}^I \sum_{\substack{k=1 \\ k \neq j}}^n P_{ik}^U / n \quad (55)$$

$$E(p)^L = \sum_{i=1}^I \sum_{\substack{k=1 \\ k \neq j}}^n P_{ik}^L / n \quad (56)$$

where  $I$  is the number of the evidence variables and  $n$  denotes the number of focal elements, respectively.

By substituting Eqs.(53-56) to Eq.(9), the deviation interval of the sound pressure response is expressed as

$$D(p)^U = \sum_{i=1}^I \left( \sum_{j=1}^n \left( \frac{n^2 - (1 + \beta)n}{n(n - \beta)} p_{ij}^U - E(p)^U \right)^2 m_{ij} \right) \quad (57)$$

$$D(p)^L = \sum_{i=1}^I \left( \sum_{j=1}^n \left( \frac{n^2 - (1 + \beta)n}{n(n - \beta)} p_{ij}^L - E(p)^L \right)^2 m_{ij} \right) \quad (58)$$

#### 4. Numerical example

In this section, a 3D structural-acoustic problem is provided to verify the approach mentioned above. A square flexible plate model coupled with the acoustic field of dimensions  $500 \times 500 \times 500$ mm is depicted in Fig.2. The plate structure is discretized by 144 four-node quadrilateral elements and the acoustic field is discretized by 1152 eight-node hexahedron elements. The acoustic field is surrounded by five rigid walls and a flexible plate. The plate is excited by a unit normal harmonic point force at the middle point and the boundary conditions for it are:  $w = 0$ , and  $\theta_x$  and  $\theta_y$  are free at the edges.

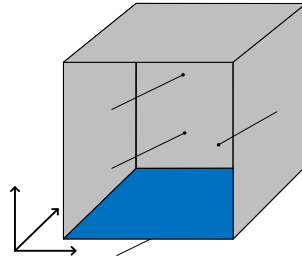


Figure 2. A cubic structural-acoustic coupling model

The density  $\rho_s$  and the Poisson's ratio  $\nu$  of the plate are  $2.5 \times 10^3 \text{ kg/m}^3$  and 0.37, respectively. The sound speed of the air in the acoustic field  $c$  is 346m/s. The Young's modulus and the thickness of the plate, the density of the air in the acoustic field are considered to be the independent uncertain parameters which are treated as evidence variables. To compare with the probability method, assumed that the evidence variables are the truncated normal distribution:  $\mu(E)=21 \times 10^4 \text{ Mpa}$ ,  $\sigma(E)=0.84 \times 10^4 \text{ Mpa}$ ,  $\mu(\rho_f)=1.30 \text{ kg/m}^3$ ,  $\sigma(\rho_f)=0.03 \text{ kg/m}^3$  and  $\mu(t)=1.25 \text{ mm}$ ,  $\sigma(t)=0.05 \text{ mm}$ . The BPA of uncertain parameters with 4, 8 and 16 focal elements are given in Table 1. Simulations of the cubic structural-acoustic coupling model are carried out by MATLAB R2009a on a 3.30 GHz Xeon(R) CPU E3 1230 v3.

Table 1. The BPA for uncertain parameters with 4, 8 and 16 focal elements

focal elements	E ( $10^4$ MPa)		$\rho_f$ ( $\text{kg/m}^3$ )		t (mm)	
	focal element	BPA	focal element	BPA	focal element	BPA
4	[18.50, 19.75]	6.69	[1.21, 1.26]	6.55	[1.10, 1.18]	6.55
	[19.75, 21.00]	43.16	[1.26, 1.30]	43.30	[1.18, 1.25]	43.30
	[21.00, 22.25]	43.16	[1.30, 1.34]	43.30	[1.25, 1.33]	43.30
	[22.25, 23.50]	6.69	[1.34, 1.39]	6.55	[1.33, 1.40]	6.55
8	[18.50, 19.13]	1.13	[1.21, 1.23]	1.08	[1.10, 1.14]	1.09
	[19.13, 19.75]	5.56	[1.23, 1.26]	5.45	[1.14, 1.18]	5.46
	[19.75, 20.38]	16.00	[1.26, 1.28]	15.98	[1.18, 1.21]	15.97
	[20.38, 21.00]	27.16	[1.28, 1.30]	27.34	[1.21, 1.25]	27.33
	[21.00, 21.63]	27.16	[1.30, 1.32]	27.34	[1.25, 1.29]	27.33
	[21.63, 22.25]	16.00	[1.32, 1.35]	15.98	[1.29, 1.33]	15.97
16	[22.25, 22.88]	5.56	[1.35, 1.37]	5.45	[1.33, 1.36]	5.46
	[22.88, 23.50]	1.13	[1.37, 1.39]	1.08	[1.36, 1.40]	1.09
	[18.50, 18.81]	0.31	[1.21, 1.22]	0.30	[1.10, 1.12]	0.29
	[18.81, 19.13]	0.82	[1.22, 1.23]	0.79	[1.12, 1.14]	0.78
	[19.13, 19.44]	1.86	[1.23, 1.24]	1.82	[1.14, 1.16]	1.82
	[19.44, 19.75]	3.69	[1.24, 1.26]	3.64	[1.16, 1.18]	3.64
	[19.75, 20.06]	6.38	[1.26, 1.27]	6.35	[1.18, 1.19]	6.35

[20.06, 20.38]	9.62	[1.27, 1.28]	9.63	[1.19, 1.21]	9.63
[20.38, 20.69]	12.65	[1.28, 1.29]	12.71	[1.21, 1.23]	12.72
[20.69, 21.00]	14.52	[1.29, 1.30]	14.61	[1.23, 1.25]	14.62
[21.00, 21.31]	14.52	[1.30, 1.31]	14.61	[1.25, 1.27]	14.62
[21.31, 21.63]	12.65	[1.31, 1.32]	12.71	[1.27, 1.29]	12.72
[21.63, 21.94]	9.62	[1.32, 1.33]	9.63	[1.29, 1.31]	9.63
[21.94, 22.25]	6.38	[1.33, 1.35]	6.35	[1.31, 1.33]	6.35
[22.25, 22.56]	3.69	[1.35, 1.36]	3.64	[1.33, 1.34]	3.64
[22.56, 22.88]	1.86	[1.36, 1.37]	1.82	[1.34, 1.36]	1.82
[22.88, 23.19]	0.82	[1.37, 1.38]	0.79	[1.36, 1.38]	0.78
[23.19, 23.50]	0.31	[1.38, 1.39]	0.30	[1.38, 1.40]	0.29

The relevant expectance and standard deviation of the sound pressure response at the points with the distances of 50mm, 100mm, 150mm, 200mm, 250mm, 300mm, 350mm, 400mm, 450mm and 500mm are calculated. In Fig. 3, the results of frequency 100 Hz are depicted. The lower and upper bounds of the relevant expectance and standard deviation of the sound pressure response at the Point 1 with the distance of 400mm in the frequency range of 20 to 200 Hz are plotted in Fig. 4. The results obtained by the Monte Carlo method with 100000 samples are used as the reference. From Fig. 3 and Fig. 4, when the uncertain parameters are treated as evidence variables, the relevant expectance and standard deviation of the sound pressure response are intervals. Besides, the lower and upper bounds of the relevant expectance and standard deviation contain the reference. With the number of focal elements increasing, the width of the expectance and standard deviation will be decreased. Because of each evidence variable follows the truncated normal distribution in which the BPA of focal element is the cumulative probability distribution in the corresponding interval. With the amount of information increasing, the evidence uncertainty could be reducible. Thus, the analysis results will more approach to the probability computational results with more BPAs in a certain interval range. In the numerical example, the precision and effectiveness of the proposed approach for structural-acoustic fields with epistemic uncertainty is validated by comparing the analysis results with evidence variables to the probability computational results.

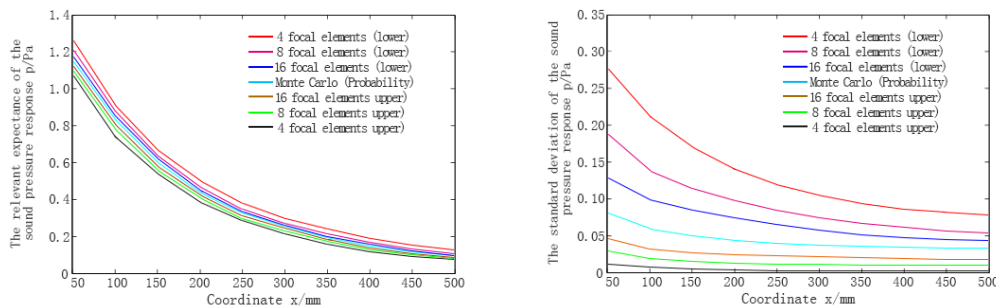


Figure 3. Bounds of the relevant expectance and standard deviation of the sound pressure response with 4, 8 and 16 focal elements (100Hz)



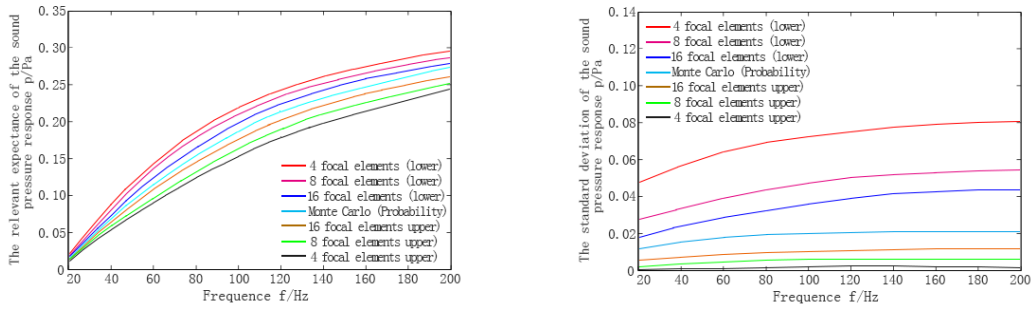


Figure 4. Bounds of the relevant expectance and standard deviation of the sound pressure response at the Point 1 under 4, 8 and 16 focal elements (20 - 200 Hz)

Assuming that  $\mathbf{x}$  is an evidence variable and  $\mathbf{X}$  denotes the sound pressure response, the belief  $\text{Bel}(\mathbf{X} \leq \mathbf{x})$  and the plausibility  $\text{Pl}(\mathbf{X} \leq \mathbf{x})$  of the sound pressure response at the Point 2 for the frequency 100 Hz are depicted in Fig. 5. The probability density function (PDF) of probability computational results obtained by the Monte Carlo method with 100000 samples are also regarded as the reference. From Fig.5, the PDFs are surrounded by the Bel and the Pls. Furthermore, with the number of focal elements increasing, the width between Bel and Pl will be decreased which further indicates the precision and effectiveness of the proposed method.

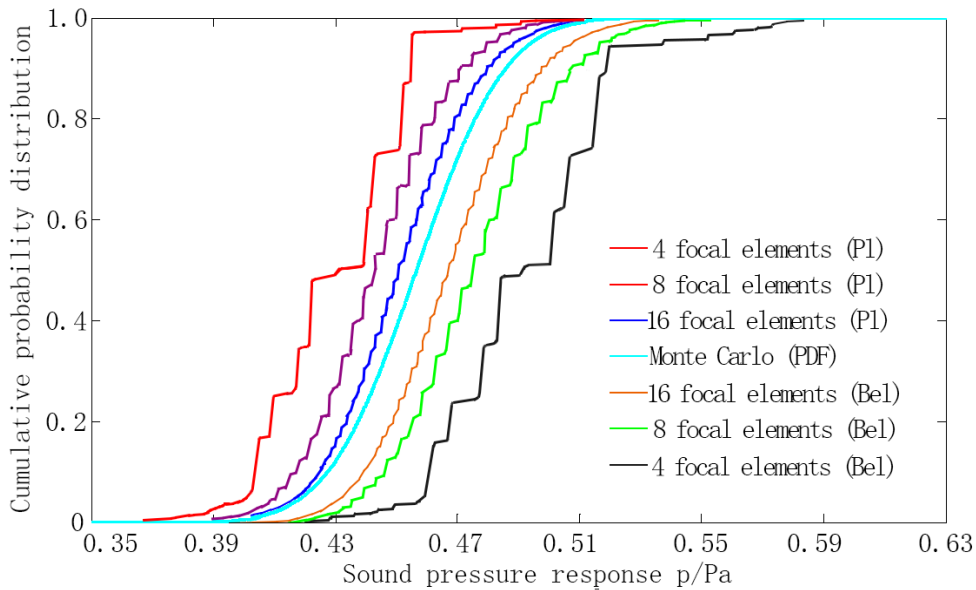


Figure 5. Cumulative probability distribution of the sound pressure response at the Point 2 (100Hz)

## Conclusions

In this paper, an evidence-theory-based approach is proposed for structural-acoustic problem with epistemic uncertainty. The evidence theory is used to handle the epistemic uncertainty in which there is no enough information or sufficient knowledge to construct the precise probability distribution for uncertain parameters. The numerical example of a plate structure-acoustic coupling system is investigated. The conclusions are as follows:

(1) The overestimation phenomenon, which derives from the correlation between parameters, is widely existent in the analysis of complex systems. The proposed method is suggested to overcome the characteristic of overestimation. The results of the numerical example shows that the proposed approach is much more efficient than the original method as the focal elements increases. Therefore, we can control the form, size and quantity of focal element to improve the analytical accuracy in practical applications.

(2) The relevant expectance, standard deviation and probability density distribution of sound pressure response are intervals not deterministic values. As the amount of information and knowledge increasing, the epistemic uncertainty could be eliminated. In other words, the bandwidths of the relevant expectance, standard deviation and probability density distribution of sound pressure response will be narrower which means the analysis results will more approach to the probability computational results.

It should be noted that this paper is focused on the epistemic uncertainty. In practical engineering problems, epistemic uncertainty and aleatory uncertainty may exist simultaneously. Thus, in further research, on the one hand, the hybrid evidence variables and random variables involved in structural-acoustic field will be investigated. On the other hand, the proposal method could be widely applied in engineering fields, such as dynamic thermal field analysis, thermal-coupling field analysis, heat-pressure field analysis and so on.

### **Acknowledgments**

The paper is supported by the Independent Research Project of State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body in Hunan University (Grant No. 71375004 and Grant No. 51375002) and the Hunan Provincial Innovation Foundation for Postgraduate (Grant No. CX2014B147). The authors would also like to thank reviewers for their valuable suggestions.

### **REFERENCE**

- [1] Du, J. B., Olhoff, N. (2010) Topological design of vibrating structures with respect to optimum sound pressure characteristics in a surrounding acoustic medium, *Struct. Multi. Optim.* 42 (1), 43–54.
- [2] He, Z. C., Liu, G. R., Zhong, Z. H., Zhang, G. Y., Cheng, A. G. (2011) A coupled ES-FEM/BEM method for fluid–structure interaction problems, *Eng. Anal. Bound Elem.* 35, 140–7.
- [3] Chen, N., Yu, D., Xia, B. (2015) Evidence-theory-based analysis for the prediction of exterior acoustic field with epistemic uncertainties, *Eng. Anal. Bound Elem.* 50, 402-411.
- [4] Stefanou, G. (2009) The stochastic finite element method: past, present and future, *Comput. Method Appl. M.* 198 (9–12), 1031–1051.
- [5] Hoffman, F. O., Hammonds, J.S. (1994) Propagation of uncertainty in risk assessment: the need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability, *Risk Anal.* 14 (5), 707–12.
- [6] Zhan, K., Luo, Y. (2009) Non-probabilistic reliability-based topology optimization of

- geometrically nonlinear structures using convex models, *Comput. Methods Appl. Mech. Eng.* 198 (41), 3228–38.
- [7] Adhikari, S. (2011) Doubly spectral stochastic finite-element method for linear structural dynamics, *J. Aerospace Eng.* 24 (3), 264–276.
- [8] Feng, Y. T., Li, C. F., Owen, D. R. J. (2010) A directed Monte Carlo solution of linear stochastic algebraic system of equations, *Finite Elem. Anal. Des.* 46 (6), 462–473.
- [9] Hua, X. G., Ni, Y. Q., Chen, Z. Q., Ko, J. M. (2007) An improved perturbation method for stochastic finite element model updating, *Int. J. Numer. Meth. Eng.* 73 (13), 1845–1864.
- [10] Oberkampf, W. L., Helton, J. C. Investigation of evidence theory for engineering applications, AIAA2002-1569, Non-Deterministic Approaches Forum, Denver, CO; 2002.
- [11] Jiang, C., Zhang, Z., Han, X., Liu, J. (2013) A novel evidence-theory-based reliability analysis method for structures with epistemic uncertainty, *Comput. Struct.* 129, 1–12.
- [12] Bai, Y. C., Jiang, C., Han, X., Hu, D. A. (2013) Evidence-theory-Based structural static and dynamic response analysis under epistemic uncertainties, *Finite Elem. Anal. Des.* 68, 52–62.
- [13] Bae, H. R., Grandhi, R. V., Canfield, R. A. (2006) An approximation approach for uncertainty quantification using evidence theory, *Reliab. Eng. Syst. Saf.* 86, 215–225.
- [14] Bae, H. R., Grandhi, R. V., Canfield, R. A. (2004) Epistemic uncertainty quantification techniques including evidence theory for large-scale structures, *Comput. Struct.* 82, 1101–1112.
- [15] Helton, J. C., Johnson, J. D., Oberkampf, W. L., Sallaberry, C. J. (2006) Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty, *Reliab. Eng. Syst. Saf.* 91, 1414–1434.
- [16] Haim, Y. Ben, Elishakoff, I. *Convex Models of Uncertainty in Applied Mechanics*, Elsevier Science Publishers, Amsterdam, 1990.
- [17] Qiu, Z. P., Xia, Y. Y., Yang, J. L. (2007) The static displacement and the stress analysis of structures with bounded uncertainties using the vertex solution theorem, *Comput. Methods Appl. Mech. Eng.* 196 (49–52), 4965–84.
- [18] Helton, J. C., Johnson, J. D., Oberkampf, W. L. (2004) An exploration of alternative approaches to the representation of uncertainty in model predictions, *Reliab. Eng. Syst. Saf.* 85, 39–71.
- [19] Rao, S. S., Berke, L. (1997) *Analysis of Uncertain Structural Systems Using Interval Analysis*, AIAA 35 (4), 727–735.
- [20] Massa, F., Tison, T., Lallemand, B. (2006) A fuzzy procedure for the static design of imprecise structures, *Comput. Methods Appl. Mech. Eng.* 195 (9–12), 925–941.
- [21] Gersem, H. De, Moens, D., Desmet, W., Vandepitte, D. (2007) Interval and fuzzy dynamic analysis of finite element models with superelements, *Comput. Struct.* 85 (5–6), 304–319.
- [22] Balu, A.S., Rao, B.N. (2012) High dimensional model representation based formulations for fuzzy finite element analysis of structures, *Finite Elem. Anal. Des.* 50, 217–230.
- [23] Hanss, M., Turrin, S. (2010) A fuzzy-based approach to comprehensive modeling and analysis of systems with epistemic uncertainties, *Struct. Saf.* 32 (6), 433–441.
- [24] Bae, H. R., Grandhi, R. V., Canfield, R. A. (2006) An approximation approach for uncertainty quantification using evidence theory, *Reliab. Eng. Syst. Saf.* 86, 215–225.
- [25] Bae, H. R., Grandhi, R. V., Canfield, R. A. (2004) Epistemic uncertainty quantification techniques including evidence theory for large-scale structures, *Comput. Struct.* 82, 1101–1112.
- [26] Soundappan, P., Nikolaidis, E., Haftka, R. T., Grandhi, R., Canfield, R. (2004) Comparison of evidence theory and Bayesian theory for uncertainty modeling. *Reliab. Eng. Syst. Saf.* 85 (1), 295–311.

- [27] Xia, B., Yu, D., Liu, J. (2013) Interval and subinterval perturbation methods for a structural-acoustic system with interval parameters, *J. Fluid. Struct.* 38, 146-163.
- [28] Heaney, K. D., Cox, H. (2006) A tactical approach to environmental uncertainty and sensitivity, *IEEE J. Oceanic Eng.* 31, 356–367.
- [29] Xia, B., Yu, D., Liu, J. (2013) Hybrid uncertain analysis for structural–acoustic problem with random and interval parameters, *J. Sound Vib.* 332, 2701-2720.
- [30] Dehghan, M., Hashemi, B. (2006) Iterative solution of fuzzy linear systems, *Appl. Math. Comput.* 175, 645–74.
- [31] Elishakoff, I., Thakkar, K. (2014) Overcoming Overestimation Characteristic to Classical Interval Analysis, *AIAA Journal* 52 (9), 2093-2097.
- [32] Xia, B., Yu, D. (2014) An interval random perturbation method for structural-acoustic system with hybrid uncertain parameters, *Int. J. Numer. Meth. Engng.* 97, 181-206.
- [33] Xia, B., Yu, D., Liu, J. (2013) Probabilistic Interval Perturbation Methods for Hybrid Uncertain Acoustic Field Prediction, *J. Vib. Acoust.* 135, 1-12.
- [34] Everstine, G. C., Henderson, F. M. (1990) Coupled finite element/boundary element approach for fluid structure interaction, *J. Acoust. Soc. Am.* 87, 1938-1947.
- [35] Wenterodt, C., Estorff, O. Von (2009) Dispersion analysis of the meshfree radial point interpolation method for the Helmholtz equation, *Int. J. Numer. Meth. Engng.* 77, 1670-1689.
- [36] Li, K., Huang, Q. B., Wang, J. L., Lin, L. G. (2011) An improved localized radial basis function meshless method for computational aeroacoustics, *Eng. Anal. Bound Elem.* 35, 47-55.
- [37] Zhang, Z., Jiang, C., Han, X., Hu, Dean, Yu, S. (2014) A response surface approach for structural reliability analysis using evidence theory, *Adv. Eng. Softw.* 69, 37-45.
- [38] He, Z. C., Liu, G. R., Zhong, Z. H., Cui, X. Y., Zhang, G. Y., Cheng, A. G. (2010) A coupled edge-/face-based smoothed finite element method for structural-acoustic problems, *Appl. Acoust.* 71, 955-964.
- [39] Wang, C., Qiu, Z., Wang, X., Wu, D. (2014) Interval finite element analysis and reliability-based optimization of coupled structural-acoustic system with uncertain parameters, *Finite Elem. Anal. Des.* 91, 108-114.
- [40] Xia, B., Yu, D., Han, X., Jiang, C. (2014) Unified response probability distribution analysis of two hybrid uncertain acoustic fields, *Comput. Methods Appl. Mech. Engrg.* 276, 20-34.

## Towards wideband mechanical metamaterials: Comparing nonlinear oscillator mechanisms

†\*A. Banerjee<sup>1,2</sup>, E.P. Calius<sup>2</sup>, and R. Das<sup>1</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Auckland, Auckland, New Zealand

<sup>2</sup>Callaghan Innovation, Auckland, New Zealand

\*Presenting author: aban991@aucklanduni.ac.nz

†Corresponding author: aban991@aucklanduni.ac.nz

### Abstract

Metamaterial is a designed material, having some exotic phenomena in resonating frequency range, such as negative properties. This is often achieved through resonant electromagnetic, acoustic or mechanical structures inside the metamaterial. Mechanical metamaterials have a comprehensive range of applications in sound, vibration and seismic engineering. However, the effectiveness of metamaterials is limited to a relatively narrow frequency band as they are generally based on linear resonance mechanisms. These linear metamaterials do not perform well under the broadband excitation spectra that are common in real life applications. Towards the first step to widen the bandwidth of the metamaterial, different classes of nonlinear oscillations, namely Duffing type monostable and bistable, and piecewise linear, are studied in non-dimensional way and compared with each other, to identify the best one according to resonating bandwidth increment. A straightforward time history based iterative methodology FRSFTI is developed to get the frequency amplitude plot of a nonlinear system without employing any approximate perturbation method. The frequency-amplitude plot from this method shows a good agreement with the conventional perturbation method at the resonating frequency range; moreover, this method enables to compute the response away from the resonating range. From the analysis it can be concluded that the bandwidth increment of bistable Duffing type oscillator is largest compared to others.

**Keywords:** Nonlinear metamaterial, Mechanical metamaterial, Nonlinear oscillation, Bandwidth comparison, Multi-stability, Steady state response, Non-dimensional analysis

### Introduction

Metamaterials are generalized composites that can exhibit unconventional behaviors and responses that are not commonly encountered in natural materials [Banerjee, B. (2011)], such as negative properties. This is often achieved through resonant electromagnetic [Lei, Z. (2008); Willis, J. R. (2011)], acoustic [Lee, S. H. et al. (2009); Huang, H. H. and Sun, C. T. (2012); Pai, P. F. et al. (2014); Sun, H. et al. (2014)] or mechanical structures inside the metamaterial. Mechanical metamaterials extensively used in the field of sound, vibration and seismic engineering. However, as they are generally based on linear resonance mechanisms, their effectiveness tends to be limited to a relatively narrow frequency band. These linear metamaterials do not perform well under the broadband excitation spectra that are common in real life applications.

Nonlinearity has a potential to widen the bandwidth of oscillator-based metamaterials by exploiting features, such as sub- and super-harmonic resonances, period multiplication, and chaotic response. Nonlinear metamaterials have already been studied in the context of electromagnetic wave propagation [Lapine, M. et al. (2014)], but to date the applicability of nonlinear metamaterials in other fields has received little attention. On the other hand, nonlinear oscillation and its effect on bandwidth have been well studied in the context of energy harvesting. Vibration of a ferromagnetic beam [Holmes, P. (1979); Moon, F. and Holmes, P. J. (1979); Erturk, A. and Inman, D. J. (2011)] or a beam with tip magnet [Stanton, S. C. et al. (2009); Zhou, S. et al. (2014)] in the presence of magnetic field can show monostable and bistable Duffing type oscillation depending on the position of magnets. Similarly, bistable and monostable oscillation can be found in the case of transverse [Sneller, A. J. et al. (2011); Cottone, F. et al. (2013); Andò, B. et al. (2014)] and axial [Masana, R. and Daqaq, M. F. (2011); Cottone, F. et al. (2012); Masana, R. and Daqaq, M. F. (2012)] vibration

of post and pre-buckled beams, respectively. Transverse vibration of Euler spring systems [Winterflood, J. et al. (2002); Zhang, G. et al. (2013)] or inclined springs systems can also result in bistable oscillations. Comparative studies between bistable and monostable harvesters show that the output power and the bandwidth exhibited a greater increase during bistable chaotic response [Masana, R. and Daqaq, M. F. (2011); Daqaq, M. F. et al. (2014)]. Ferrari et al [Stanton, S. C. et al. (2010); Ferrari, M. et al. (2011)] have shown that as the slope of the inner wall of the potential well becomes steeper, the system performs like a monostable system, and the response of the system is reduced.

As a first step towards the development of a nonlinear mechanical metamaterial, the behaviors of three types of nonlinear oscillators are examined in non-dimensional form. The oscillators under consideration are characterized by the shape of their potential energy well, namely cubic monostable, cubic bistable, and piecewise linear, as shown in Figure 1.

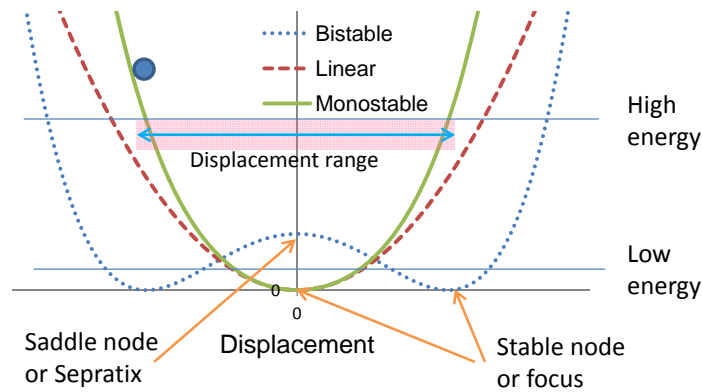


Figure 1 Potential well of typical bistable, monostable and linear system

Figure 1 shows that in a monostable system the potential well has only one stable point; whereas, the bistable system has two stable points and one saddle node. Displacement range is the peak to peak deflection of the system, as shown in Figure 1. It can be seen that the displacement range of a monostable resonator is less than that of linear and bistable oscillators, although due to the presence of sub- and super-harmonic resonances the monostable resonator can also generate a wider resonant bandwidth. On the other hand, the displacement range of a bistable oscillator is higher compared to other oscillators at high potential energy, but for lower energy vibration (below the separatrix), the displacement range is considerably reduced. The main objective of this paper is to investigate the effect of the steepness of the inner wall of the potential well and the relative distance between the stable nodes on the bandwidth for bistable systems and also compare that with monostable and piecewise type systems. The amplitude of the non-dimensional velocity in response to monochromatic excitation is determined and used to infer the bandwidth over which its response is greater than that of an equivalent linear oscillator. Thus, in this paper a comparative study is carried out to determine the bandwidth of monostable, bistable, piecewise linear system with that of a linear system.

## Methodology

### Linear oscillator

Equation of motion of linear oscillation can be written as

$$\ddot{u} + \omega^2 u = -\ddot{u}_g \sin \bar{\omega} t \quad (1)$$

where  $u$  and  $\ddot{u}$  are the displacement and acceleration of the system,  $\ddot{u}_g$  is the base excitation,  $\omega$  and  $\bar{\omega}$  are the natural frequency and excitation frequency of the system, respectively.

To non-dimensionalize Eq.(1), we set  $u = Ax(\tau)$  and  $t = \tau/\omega$ . So, the modified Eq.(1) is

$$A\omega^2 \frac{d^2x}{d\tau^2} + A\omega^2 x = -\ddot{u}_g \sin\left(\frac{\bar{\omega}}{\omega} \tau\right) \quad (2)$$

$$\ddot{x} + x = -\sin(\eta\tau)$$

where  $A = \frac{\ddot{u}_g}{\omega^2}$  and  $\eta = \frac{\bar{\omega}}{\omega}$

Now, introducing a nonlinear stiffness term  $\kappa$  in Eq.(2) the nonlinear oscillation equation can be formulated.

#### *Duffing type monostable cubic nonlinear system*

Monostable Duffing oscillation is a common example of nonlinear oscillation. Vibration of a ferromagnetic beam under the influence of magnets [Ferrari, M. et al. (2010); Kang-Qi, F. et al. (2014)], axial and transverse vibration of a pre-buckled beam [Min, G.-B. and Eisley, J. G. (1972)] result in cubic monostable Duffing type nonlinearity. By introducing a nonlinear stiffness term  $\kappa$  in Eq.(2), the equation of motion of a typical monostable Duffing oscillator can be written as:

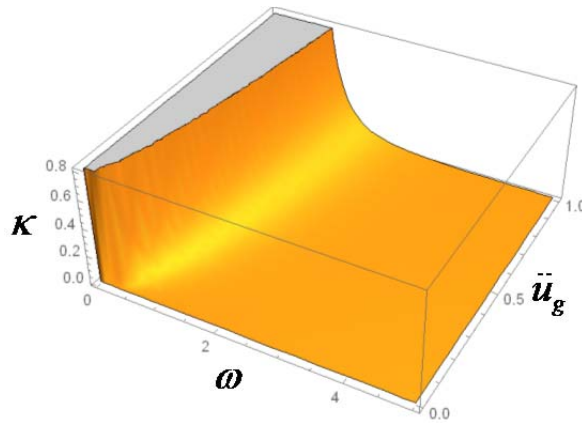
$$\ddot{u} + \omega^2 u + r\omega^2 u^3 = -\ddot{u}_g \sin \bar{\omega} t \quad (3)$$

where  $r$  is ratio of the nonlinear spring constant to the linear spring constant. To non-dimensionalize Eq.(3),  $u = Ax(\tau)$  and  $t = \tau/\omega$  substitutions are considered. So, the modified form of Eq.(3) is:

$$A\omega^2 \frac{d^2x}{d\tau^2} + A\omega^2 x + r\omega^2 A^3 x^3 = -\ddot{u}_g \sin\left(\frac{\bar{\omega}}{\omega} \tau\right) \quad (4)$$

$$\ddot{x} + x + \kappa x^3 = -\sin(\eta\tau)$$

where  $\kappa = rA^2 = r \frac{\ddot{u}_g^2}{\omega^4}$  shows that the non-dimensional nonlinear spring constant  $\kappa$  is directly proportional to the ratio of nonlinear to linear spring constant and the square of the amplitude of the acceleration input, and inversely proportional to the fourth power of the natural frequency of the system.



*Figure 2 Plot of non-dimensional nonlinear spring constant ( $\kappa$ ) with natural frequency ( $\omega$ ) and excitation amplitude ( $\ddot{u}_g$ ) for a constant value of ratio of spring coefficient ( $r$ )*

From Figure 2, it can be seen that the non-dimensional spring coefficient increases parabolically with the amplitude of vibration ( $\ddot{u}_g$ ), and is much higher at low natural frequencies.

The restoring non-dimensional force (F) and potential energy (U) can be expressed as:

$$F(x) = x + \kappa x^3$$

$$U(x) = \int_0^x F(x) dx = \frac{1}{2}x^2 + \frac{\kappa}{4}x^4 \quad (5)$$

### Duffing type bistable cubic nonlinear system

Bistable Duffing oscillation is another common form of nonlinear oscillation. Vibration of ferromagnetic beam under the influence of magnets [Stanton, S. C., McGehee, C. C. et al. (2010); Ferrari, M., Baù, M. et al. (2011)], axial and transverse vibration of post-buckled beam [Virgin, L. and Davis, R. (2003); Ibrahim, R. A. (2008); Comesca, B. et al. (2013)], vibration of Euler spring system [Winterflood, J., Barber, T. A. et al. (2002); Ibrahim, R. A. (2008); Huang, X. et al. (2014)], result in cubic bistable Duffing type nonlinearity. Unlike monostable oscillation, bistable Duffing equation has two stable nodes and one saddle node. The non-dimensional bistable Duffing equation can be written as:

$$\ddot{x} - x + \kappa x^3 = -\sin(\eta\tau) \quad (6)$$

So the restoring force ( $F$ ) and the potential energy ( $U$ ) can be expressed as:

$$F(x) = -x + \kappa x^3$$

$$U(x) = \int F(x) dx = -\frac{1}{2}x^2 + \frac{\kappa}{4}x^4 + c \quad (7)$$

Now at the stable nodes the restoring force must be 0. So, the location of the stable node can be calculated as  $x_{eq} = \pm\sqrt{\frac{1}{\kappa}}$ . If the minimum value of potential well is assumed to be zero, then the value of  $c$  is:

$$U\left(\pm\sqrt{\frac{1}{\kappa}}\right) = -\frac{1}{2}\frac{1}{\kappa} + \frac{\kappa}{4}\frac{1}{\kappa^2} + c = 0 \rightarrow c = \frac{1}{4\kappa} \quad (8)$$

The different shapes of the potential well are plotted for various values of  $\kappa$ .

### Piecewise linear system

In the piecewise linear system, the restoring force varies with the distance linearly up to a certain range. Thereafter due to impact with stopper, the slope of the restoring force curve changes significantly, but still remains linear. As the restoring force curve consists of different straight lines, this type of system is commonly known as a piecewise linear system. Vibro-impacting devices [Soliman, M. S. M. et al. (2008); Vandewater, L. and Moss, S. (2013)], where stoppers are placed at some distance from the stable point of the beam, are an example of this type of oscillator. The impact between the stopper and the beam creates the nonlinear term. The equation of motion of piecewise linear system can be written as:

$$\ddot{u} + \omega^2 u = -\ddot{u}_g \sin(\eta\tau); u + \tilde{g} \geq 0 \quad (9)$$

The Eq.(9) can be written in non-dimensional form by considering,  $u = Ax(\tau)$ ,  $t = \tau/\omega$  and

$$\tilde{g} = g \frac{\ddot{u}_g}{\omega^2}.$$

$$\ddot{x} + x = -\sin(\eta\tau); x + g \geq 0 \quad (10)$$

The restoring force ( $F$ ) and the potential energy ( $U$ ) can be expressed as:



$$\begin{aligned}
 F(x) &= \begin{cases} x & x+g \geq 0 \\ \infty & x+g < 0 \end{cases} \\
 U(x) &= \begin{cases} \frac{x^2}{2} & x+g \geq 0 \\ \infty & x+g < 0 \end{cases}
 \end{aligned} \tag{11}$$

### Identification of parameter

To compare the bandwidth of these different systems, their frequency response must be computed for a range of oscillator parameters. In linear systems, the amplitude corresponding to a specific excitation frequency of motion is conventionally solved because this is sufficient to completely describe the vibration. On the other hand, as there may be no linear correlation among all the parameters in nonlinear systems, displacement amplitude does not fully describe the oscillator dynamics. In this paper, the non-dimensional velocity amplitude  $\max(\dot{x})$  is used to describe the response, because it can represent not only the maximum velocity, but also the maximum momentum and the maximum kinetic energy of the system, which have greater significance in the context of metamaterials. The amplitude of the velocity of monochromatic excitation of frequency ( $\omega$ ) is  $\left(\frac{\ddot{u}_g}{\omega}\right)$ , which results the input momentum of the system ( $P_s$ ) is  $\frac{M\ddot{u}_g}{\omega}$ . The non-dimensional momentum ( $P$ ) of the system can be expressed as the ratio of the momentum ( $\tilde{P} = M\dot{u}$ ) and the input momentum corresponding to system ( $P_s$ ), in Eq.(12).

$$P = \frac{\tilde{P}}{P_s} = \frac{M\dot{u}\omega}{M\ddot{u}_g} = \frac{A\omega\dot{x}\omega}{\ddot{u}_g} = \dot{x} \tag{12}$$

### Solution

The frequency content of a signal is generally a more important parameter than its time history, because the corresponding response to a particular frequency can be easily calculated by convolution. In linear systems Fourier transformation of the equation of motion produces its frequency response directly; whereas in nonlinear systems Fourier transformation is not applicable [Cameron, T. M. and Griffin, J. H. (1989)]. In order to determine the frequency response of a nonlinear system, several methods have been developed mainly based on perturbation techniques. The harmonic balance (HB) method [Nayfeh, A. H. and Mook, D. T. (2008)] is the most popular and widely used technique to determine the frequency response of a nonlinear system [Beléndez, A. et al. (2007); Beléndez, A. et al. (2009); Cochelin, B. and Vergez, C. (2009); Wang, X. et al. (2012); García-Saldaña, J. D. and Gasull, A. (2013); Karkar, S. et al. (2013)]. The HB method is adequate where only single harmonic description is sufficient, as it is in the case of weakly nonlinear systems. Among the techniques that can deal with strong nonlinearity, alternating frequency/time (AFT) method [Cameron, T. M. and Griffin, J. H. (1989)], homotopy method [Liao, S. (2004); Vyasrayani, C. P. et al. (2012)] and the max-min method [He, J.-H. (2008); Azami, R. et al. (2009); Ibsen, L. B. et al. (2010)] have gained popularity.

In this paper a simple time domain based iterative method, an in house program (FRSFTI), is developed to compute the frequency response of a particular nonlinear system. To apply the method, a frequency domain of interest needs to be first defined. The program FRSFTI enables the user to control the number of steps and discretize the frequency domain into finite number of input excitation frequencies. Sinusoidal excitation of each frequency is applied to a specific system to calculate its response. To reduce the frequency leakage and the aliasing, a sine wave of fifty cycles is considered in the analysis [Lynch, S. (2011)]. The ODE45 solver of MATLAB version 8.4

[Mathworks, Natick, MA] is used to solve the equation of motion for a particular monochromatic excitation. An event identifier is amalgamated with the ODE in response to the case of a piece-wise linear system, to locate the occurrence of impact and restart integration with new initial values. Then, maximum of velocity ( $\max(\dot{x})$ ) is plotted for the corresponding frequency to obtain the frequency response because according to Eq.(12) it can represent the non-dimensional momentum which is the most important parameter in metamaterial. As the method is based on time domain solutions, there are certain advantages of this method over existing methods, such as

- No stabilization check is required and it results in a single output for a particular frequency; whereas the HB method results in a sixth order correlation between the amplitude and frequency. [Nayfeh, A. H. and Mook, D. T. (2008)].
- It can solve any type of nonlinearity or discontinuity; whereas perturbation methods are approximate and sometimes cannot deal with high level of nonlinearity or discontinuity [Nayfeh, A. H. and Mook, D. T. (2008)].
- As every frequency excitation is applied to the system, the method can work in every frequency range; whereas, most of the perturbation methods can only work near the resonating frequency or some specific sub or super harmonics.
- The method is programmable, so rigorous mathematical calculations are not needed.

To compare the bandwidth increment of three different classes of nonlinear oscillators, in each case the initial condition is assumed to be the stable node to ensure the initial potential energy is zero. That is why, initial condition is assumed to be  $(0,0)$  for the linear, monostable and piecewise linear system; whereas for bistable condition the initial condition is  $(\pm\sqrt{1/\kappa}, 0)$ . Maximum non-dimensional velocity  $\max(\dot{x})$  is considered as the comparing parameter. Figure 3 shows the flow chart of the full process.

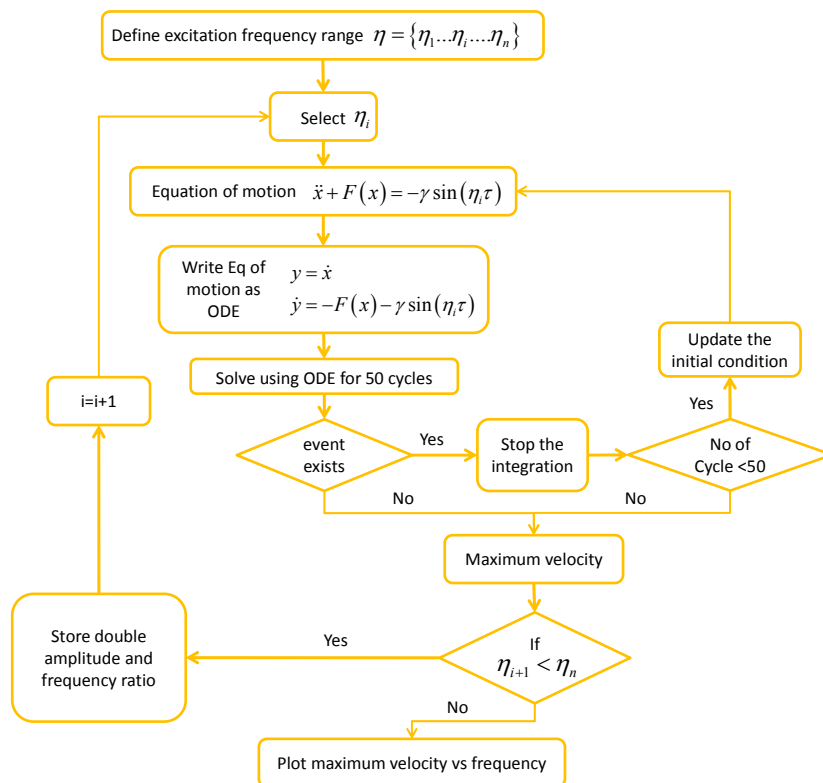


Figure 3 Steps for frequency-amplitude plot

**Results and Discussions**

*Monostable cubic nonlinear system*

The potential well and the restoring force profile of the monostable system for varying nonlinearity is plotted in Figure 4, based on Eq.(5).

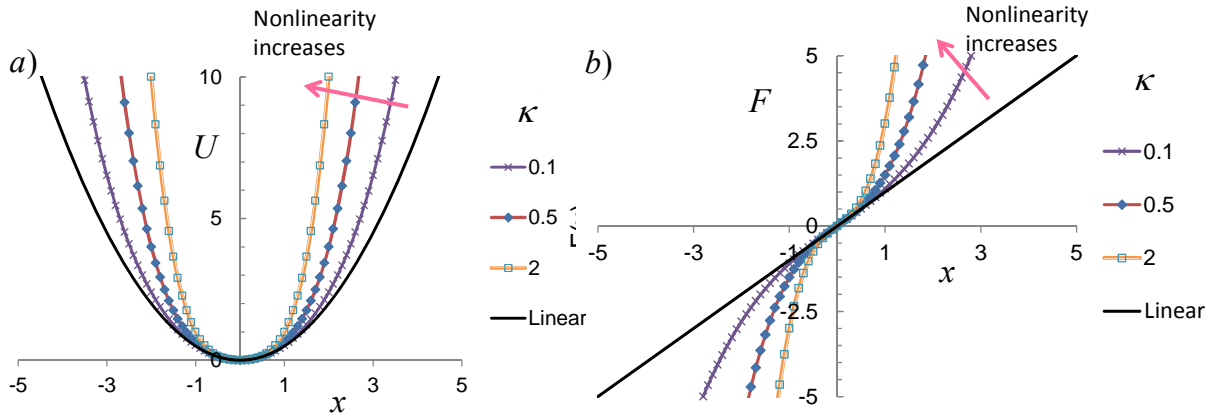


Figure 4 a) Potential energy well and b) restoring force of a linear and a monostable cubic nonlinear system for three level of nonlinearity

Figure 4 shows the variation of potential energy and the restoring force of a linear and a monostable cubic nonlinear system for three different levels of nonlinearity. The potential well becomes steeper and the restoring force shows hardening behavior as the nonlinearity increases. The resulting frequency response is shown in Figure 5.

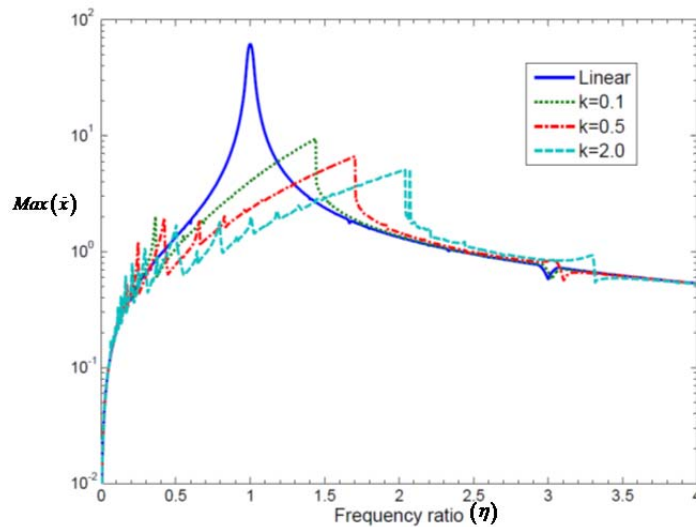


Figure 5 Frequency response of the maximum non-dimensional velocity ( $\dot{x}$ ) that relates to the amplitude of non-dimensional momentum of linear and monostable systems of varying nonlinearity

From Figure 5 it can be concluded that the resonance peak shifts to a higher side and its amplitude reduces as nonlinearity increases. Due to the presence of sub and super harmonic resonances, amplification is observed in the low frequency range. The frequency response of monostable systems is generally less than the equivalent linear system.

*Bistable cubic nonlinear system*

The potential well and the restoring force profile of bistable system for varying nonlinearity is plotted in Figure 6, based on Eq.(7).

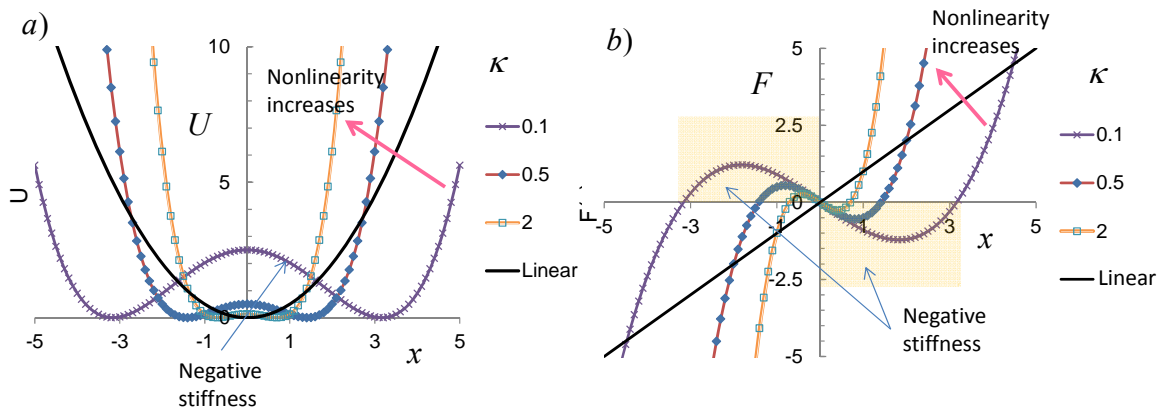


Figure 6 a) Potential energy well and b) restoring force of a linear and a monostable cubic nonlinear system for three levels of nonlinearity

Figure 6 shows the variation of potential energy and the restoring force of a linear and a bistable cubic nonlinear system for three different levels of nonlinearity. It can be noted that the stable nodes approach closer with increasing level of nonlinearity. Negative stiffness resulted at the region between two stable nodes as shown in Figure 6. The distance between the outer walls of potential wells decreased with increment of nonlinearity, hence the motion confined between two walls. Simultaneously, the energy required to overcome the separatrix barrier decreases with nonlinearity, which enables the motion to overcome the separatrix barrier easily and results a bistable motion. Therefore, an optimum level of nonlinearity should exist where the maximum response can be obtained in the case of bistable systems.

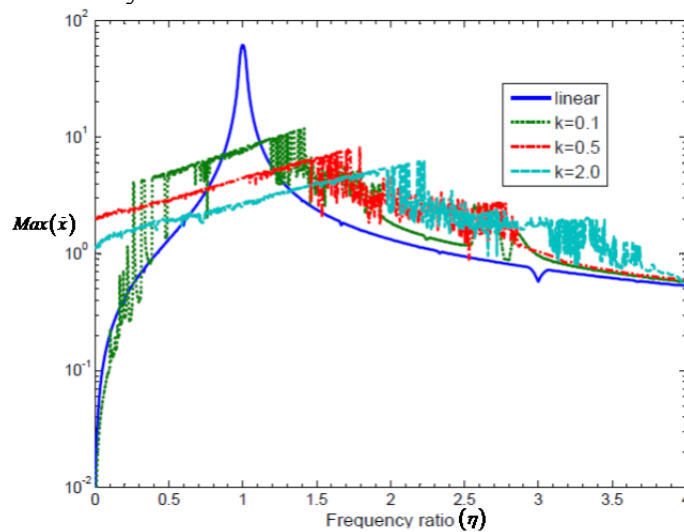


Figure 7 Frequency response of the maximum non-dimensional velocity ( $\dot{x}$ ) that relates to the amplitude of non-dimensional momentum of linear and bistable systems of varying nonlinearity

Figure 7 shows that the nonlinear system's response (maximum non-dimensional velocity) is above the equivalent linear response over a considerable range of frequencies. The increased response is particularly marked at low frequencies, which is of significant practical interest, and upon examination this frequency range is found to be greatest for  $\kappa = 0.5$  among the three nonlinear systems. Bistable systems with high nonlinearity show much higher response in the low frequency range because the energy required to cross the separatrix barrier is less.

*Piecewise linear system*

The potential well and the restoring force profile of piecewise linear system for varying gap distance is plotted in Figure 8, based on Eq.(11).

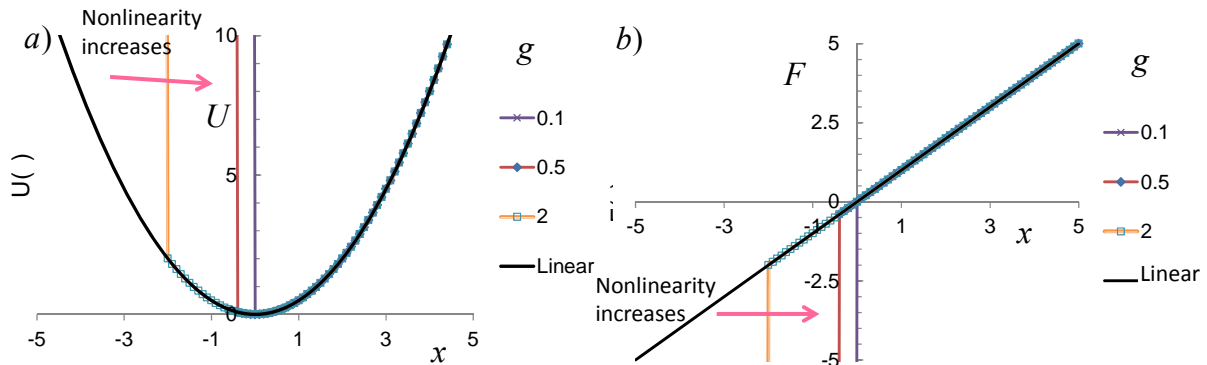


Figure 8 a) Potential energy well and b) restoring force of a linear and a piecewise linear system for three level of nonlinearity

Figure 8 shows that the potential well and the restoring force of a piecewise linear system, such as a vibro-impacting device, are the same as those for a linear system until the point of impact, whereupon a large impacting force is suddenly applied.

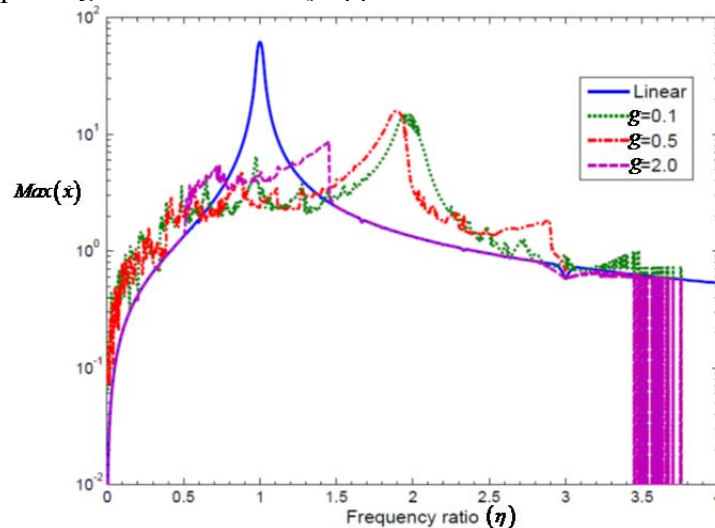


Figure 9 Frequency response of the maximum non-dimensional velocity ( $\dot{x}$ ) which represents the amplitude of non-dimensional momentum of linear and bistable systems of varying nonlinearity

Figure 9 shows that the nonlinear system’s response (maximum non-dimensional velocity) is above the equivalent linear response over a considerable range of frequencies. Figure 9 shows a remarkable increase in the frequency response at low frequency range and at a frequency ratio ( $\eta$ ) around 2. Due to the impact, the motion path reduces which shifts the frequency towards  $\eta = 2$ . When the displacement of the system is more than the gap distance then impact happens, otherwise for non-impacting case the frequency response curve follows the linear response curve. That is why the systems having large gap yields very less bandwidth and almost follows the linear path.

*Bandwidth*

To compare the bandwidth of these three nonlinear systems considered with that of an equivalent linear system, the frequency range is divided into two main categories: the low range when the frequency ratio is below the linear resonant frequency ( $\eta < 1$ ), and the high range when the

frequency ratio is above the linear resonant frequency ( $\eta > 1$ ). The low and high ranges are further subdivided into two parts, the nonlinearity dominated range (NLD) and the linearity dominated range (LD). In the NLD range, the maximum momentum of the nonlinear system is greater than that of the linear one; whereas, in the LD range the linear response is higher than the nonlinear one. A bandwidth index (*BWI*) for each system at different ranges, such as low, high and overall, is calculated based on the ratio:

$$BWI = \frac{NLD}{LD} \quad (13)$$

The bandwidth comparison of all the proposed systems is given in Table 1.

*Table 1 Summary of nonlinearity-dominated frequency range (non-dimensional) and the associated bandwidth index (BWI) for three different types of nonlinear systems*

Type of nonlinearity	Level of nonlinearity	Frequency band increment					
		Low range		High range		Overall	
		NLD ( $\eta$ )	BWI	NLD ( $\eta$ )	BWI	NLD ( $\eta$ )	BWI
Monostable	$\kappa = 0.1$	0.28	0.39	2.02	2.06	2.30	1.31
	$\kappa = 0.5$	0.27	0.37	1.86	1.63	2.13	1.14
	$\kappa = 2.0$	0.20	0.25	1.88	1.67	2.08	1.08
Bistable	$\kappa = 0.1$	0.57	1.34	2.86	19.83	3.43	5.99
	$\kappa = 0.5$	0.76	3.10	2.78	12.89	3.54	7.70
	$\kappa = 2.0$	0.63	1.69	2.67	8.15	3.30	4.71
Piecewise	$g = 2.0$	0.69	2.25	1.07	0.55	1.76	0.79
	$g = 0.5$	0.62	1.66	1.64	1.21	2.27	1.31
	$g = 0.1$	0.65	1.87	1.94	1.84	2.60	1.85

### Remarks

From the above discussion the following remarks can be made:

- The relative magnitude of the nonlinear term is proportional to the square of amplitude of excitation and the ratio of nonlinear to linear stiffness, and inversely proportional to the forth power of natural frequency of the system. The strength of the nonlinearity increases rapidly as the frequency of excitation decreases.
- In a monostable system, the potential well becomes stiffer with increasing nonlinearity. The nonlinearity dominated bandwidth in the low frequency range is lower than that of the linear system and much lower than that of the other nonlinear systems, such as bistable and piecewise linear.
- In a bistable system, the energy required to overcome the separatix decreases and the stable nodes approach closer as nonlinearity increases. This lowers the threshold to achieve the bistable response that produces maximum displacement and velocity in the low frequency excitation. On the other hand, the potential well becomes narrower and steeper, which reduces the velocity peak. For this reason, the nonlinearity dominated bandwidth increases up to an optimum value and thereafter it decreases.
- A piecewise linear system having large gap size results least nonlinear dominated bandwidth, but the bandwidth increases significantly as the gap reduces.

### Conclusion

In this paper, a comparative study of three types of nonlinear systems, namely, cubic monostable and bistable, and piecewise linear, is carried out using non-dimensional variables. The equations of motion for these systems were non-dimensionalized and solved numerically for monochromatic

excitation over a range of frequencies. To examine the effect of nonlinearity on the response bandwidth, the non-dimensional velocity response of these systems is compared to that of a linear system. A new parameter, bandwidth index (BWI), is proposed and used to quantify the relative performance of these nonlinear systems. Of the three systems studied, the bistable system with intermediate nonlinearity has the largest nonlinear dominated bandwidth (NLD) and BWI, both for low range and overall.

## References

- Andò, B., Baglio, S., Bulsara, A. R. and Marletta, V. (2014). A bistable buckled beam based approach for vibrational energy harvesting *Sensors and Actuators A: Physical* **211**(0): 153-161.
- Azami, R., Ganji, D. D., Babazadeh, H., Dvavodi, A. G. and Ganji, S. S. (2009). He's Max–Min Method for the relativistic oscillator and high order Duffing equation *International Journal of Modern Physics B* **23**(32): 5915-5927.
- Banerjee, B. (2011). *An Introduction to Metamaterials and Waves in Composites*, Auckland, Taylor & Francis.
- Beléndez, A., Gimeno, E., Alvarez, M. L. and Méndez, D. I. (2009). Nonlinear oscillator with discontinuity by generalized harmonic balance method *Computers & Mathematics with Applications* **58**(11–12): 2117-2123.
- Beléndez, A., Hernández, A., Beléndez, T., Álvarez, M. L., Gallego, S., Ortuño, M. and Neipp, C. (2007). Application of the harmonic balance method to a nonlinear oscillator typified by a mass attached to a stretched wire *Journal of Sound and Vibration* **302**(4–5): 1018-1029.
- Cameron, T. M. and Griffin, J. H. (1989). An Alternating Frequency/Time Domain Method for Calculating the Steady-State Response of Nonlinear Dynamic Systems *Journal of Applied Mechanics* **56**(1): 149-154.
- Camescasse, B., Fernandes, A. and Pouget, J. (2013). Bistable buckled beam: Elastica modeling and analysis of static actuation *International Journal of Solids and Structures* **50**(19): 2881-2893.
- Cochelin, B. and Vergez, C. (2009). A high order purely frequency-based harmonic balance formulation for continuation of periodic solutions *Journal of Sound and Vibration* **324**(1–2): 243-262.
- Cottone, F., Basset, P., Vocca, H., Gammaitoni, L. and Bourouina, T. (2013). Bistable electromagnetic generator based on buckled beams for vibration energy harvesting *Journal of Intelligent Material Systems and Structures*: 1045389X13508330.
- Cottone, F., Gammaitoni, L., Vocca, H., Ferrari, M. and Ferrari, V. (2012). Piezoelectric buckled beams for random vibration energy harvesting *Smart materials and structures* **21**(3): 035021.
- Daqaq, M. F., Masana, R., Erturk, A. and Dane Quinn, D. (2014). On the Role of Nonlinearities in Vibratory Energy Harvesting: A Critical Review and Discussion *Applied Mechanics Reviews* **66**(4): 040801-040801.
- Erturk, A. and Inman, D. J. (2011). Broadband piezoelectric power generation on high-energy orbits of the bistable Duffing oscillator with electromechanical coupling *Journal of Sound and Vibration* **330**(10): 2339-2353.
- Ferrari, M., Baù, M., Guizzetti, M. and Ferrari, V. (2011). A single-magnet nonlinear piezoelectric converter for enhanced energy harvesting from random vibrations *Sensors and Actuators A: Physical* **172**(1): 287-292.
- Ferrari, M., Ferrari, V., Guizzetti, M., Andò, B., Baglio, S. and Trigona, C. (2010). Improved energy harvesting from wideband vibrations by nonlinear piezoelectric converters *Sensors and Actuators A: Physical* **162**(2): 425-431.
- García-Saldaña, J. D. and Gasull, A. (2013). A theoretical basis for the Harmonic Balance Method *Journal of Differential Equations* **254**(1): 67-80.
- He, J.-H. (2008). Max-min approach to nonlinear oscillators *International Journal of Nonlinear Sciences and Numerical Simulation* **9**(2): 207-210.
- Holmes, P. (1979). A nonlinear oscillator with a strange attractor *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* **292**(1394): 419-448.
- Huang, H. H. and Sun, C. T. (2012). Anomalous wave propagation in a one-dimensional acoustic metamaterial having simultaneously negative mass density and Young's modulus *The Journal of the Acoustical Society of America* **132**: 2887.
- Huang, X., Liu, X., Sun, J., Zhang, Z. and Hua, H. (2014). Vibration isolation characteristics of a nonlinear isolator using Euler buckled beam as negative stiffness corrector: A theoretical and experimental study *Journal of Sound and Vibration* **333**(4): 1132-1148.
- Ibrahim, R. A. (2008). Recent advances in nonlinear passive vibration isolators *Journal of Sound and Vibration* **314**(3–5): 371-452.
- Ibsen, L. B., Barari, A. and Kimiaefar, A. (2010). Analysis of highly nonlinear oscillation systems using He's max-min method and comparison with homotopy analysis and energy balance methods *Sadhana* **35**(4): 433-448.
- Kang-Qi, F., Chun-Hui, X., Wei-Dong, W. and Yang, F. (2014). Broadband energy harvesting via magnetic coupling between two movable magnets *Chinese Physics B* **23**(8): 084501.

- Karkar, S., Cochelin, B. and Vergez, C. (2013). A high-order, purely frequency based harmonic balance formulation for continuation of periodic solutions: The case of non-polynomial nonlinearities *Journal of Sound and Vibration* **332**(4): 968-977.
- Lapine, M., Shadrivov, I. V. and Kivshar, Y. S. (2014). Colloquium: Nonlinear metamaterials *Reviews of Modern Physics* **86**(3): 1093.
- Lee, S. H., Park, C. M., Seo, Y. M., Wang, Z. G. and Kim, C. K. (2009). Acoustic metamaterial with negative density *Physics Letters A* **373**(48): 4464-4469.
- Lei, Z. (2008). Effective 1-D material properties of coplanar-waveguide-based EBG- and meta-materials, *Metamaterials, 2008 International Workshop on*.
- Liao, S. (2004). On the homotopy analysis method for nonlinear problems *Applied Mathematics and Computation* **147**(2): 499-513.
- Lynch, S. (2011). *MATLAB Programming for Engineers, Applications of Chaos and Nonlinear Dynamics in Engineering-Vol. 1*, Springer: 3-35.
- Masana, R. and Daqaq, M. F. (2011). Electromechanical modeling and nonlinear analysis of axially loaded energy harvesters *Journal of Vibration and Acoustics* **133**(1): 011007.
- Masana, R. and Daqaq, M. F. (2011). Relative performance of a vibratory energy harvester in mono- and bi-stable potentials *Journal of Sound and Vibration* **330**(24): 6036-6052.
- Masana, R. and Daqaq, M. F. (2012). Energy harvesting in the super-harmonic frequency region of a twin-well oscillator *Journal of Applied Physics* **111**(4): -.
- Min, G.-B. and Easley, J. G. (1972). Nonlinear Vibration of Buckled Beams *Journal of Manufacturing Science and Engineering* **94**(2): 637-645.
- Moon, F. and Holmes, P. J. (1979). A magnetoelastic strange attractor *Journal of Sound and Vibration* **65**(2): 275-296.
- Nayfeh, A. H. and Mook, D. T. (2008). *Nonlinear oscillations*, John Wiley & Sons.
- Pai, P. F., Peng, H. and Jiang, S. (2014). Acoustic metamaterial beams based on multi-frequency vibration absorbers *International Journal of Mechanical Sciences* **79**: 195-205.
- Sneller, A. J., Cette, P. and Mann, B. P. (2011). Experimental investigation of a post-buckled piezoelectric beam with an attached central mass used to harvest energy *Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering* **225**(4): 497-509.
- Soliman, M. S. M., El-Saadany, E. M., Mansour, E. F. and Abdel-Rahman, R. R. (2008). A wideband vibration-based energy harvester *Journal of Micromechanics and Microengineering* **18**(11).
- Stanton, S. C., McGehee, C. C. and Mann, B. P. (2009). Reversible hysteresis for broadband magnetopiezoelectric energy harvesting *Applied Physics Letters* **95**(17): 174103.
- Stanton, S. C., McGehee, C. C. and Mann, B. P. (2010). Nonlinear dynamics for broadband energy harvesting: Investigation of a bistable piezoelectric inertial generator *Physica D: Nonlinear Phenomena* **239**(10): 640-653.
- Sun, H., Yan, F., Gu, H. and Li, Y. (2014). Acoustic metamaterial with negative parameter, *Proceedings of SPIE - The International Society for Optical Engineering*.
- Vandewater, L. and Moss, S. (2013). Probability-of-existence of vibro-impact regimes in a nonlinear vibration energy harvester *Smart Materials and Structures* **22**(9): 094025.
- Virgin, L. and Davis, R. (2003). Vibration isolation using buckled struts *Journal of Sound and Vibration* **260**(5): 965-973.
- Vyasrayani, C. P., Uchida, T. and McPhee, J. (2012). Single-shooting homotopy method for parameter identification in dynamical systems *Physical Review E* **85**(3): 036201.
- Wang, X., Li, C., Huang, T. and Duan, S. (2012). Predicting chaos in memristive oscillator via harmonic balance method *Chaos: An Interdisciplinary Journal of Nonlinear Science* **22**(4): 043119.
- Willis, J. R. (2011). Effective constitutive relations for waves in composites and metamaterials *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* **467**(2131): 1865-1879.
- Winterflood, J., Barber, T. A. and Blair, D. G. (2002). Mathematical analysis of an Euler spring vibration isolator *Physics Letters A* **300**(2-3): 131-139.
- Zhang, G., Zhu, X., Xu, R. and Li, T. (2013). Vibration isolation characteristics of euler strut spring used in low frequency vibration isolation system. **248**: 475-480.
- Zhou, S., Cao, J., Inman, D. J., Lin, J., Liu, S. and Wang, Z. (2014). Broadband tristable energy harvester: Modeling and experiment verification *Applied Energy* **133**(0): 33-39.



## Prediction of Human Elbow Joint Torque Based on Improved BP Neural Network

†Gao yongsheng<sup>1\*</sup>, Sun xiaoying<sup>2</sup>, Fan jizhuang<sup>3</sup>, Zhu yanhe<sup>4</sup>, and Liu gangfeng<sup>5</sup>

1,3,4,5 Robotics Institute, Harbin Institute of Technology, Harbin, China.

2 Department of Civil Engineering, Harbin Institute of Technology, Harbin, China.

\*Presenting author: gaoy@hit.edu.cn

†Corresponding author: gaoy@hit.edu.cn

### Abstract

Pathological tremor brings too much inconvenience to patients in life and work. For better tremor suppression, a suitable biomechanical model must be established. Based on the Hill skeleton-muscle model, quantitative relations between EMG and static torque of elbow joint can be identified with improved neural network. The weights of improved neural network are adjusted according to the need, and muscle activation grade is confirmed. Through this method, a biomechanical model is established. Using OpenSim software we can simulate the drive of skeleton model by EMG signals and the validity of the model is tested by experiment.

Keywords: skeleton-muscle model, EMG(Electromyographic signal) , joint torque, improved neural network(NN)

### Introduction

Pathological tremor is common in middle and old ages and gives patients too much inconvenience in life and work. Now there are no effective methods in medical field. FES(Functional electrical stimulation) method is good for physiological control of human body and is in deeper research now. Biomechanical model is helpful to solve questions such as the actions of a series of relevant muscles, control mode of FES and effectiveness. All these provide basis for tremor suppression.

Biomechanical model is the hotspot now and many institutions are in deep work. Modeling methods mainly include model or non-model methods. Typical model methods are Hill model with phenomena presentation and Huxley model with physiological presentation. Pennestri E.et al established virtual skeleton muscle model of upper limb which is fit for the movement description of skeletons simply. Muscle activation grade is also calculated. Zhang D.G studied on the effectiveness of tremor suppression by wearable exo-skeleton and FES. They presented physical model aimed to muscle electrical stimulation first and creatively led in muscle damper character, electrical stimulation, and activation when muscles are in contraction.

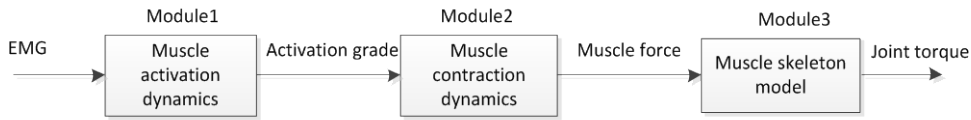
Non-model method mainly establish non-linear mapping by artificial neural network or polynomial fitting. There are many ways to establish the relations of muscle activation and joint dynamics using NN. Jer-JunnLuh estimated relation model between EMG and static moment of elbow joint using 3-layer feedforward adaptive NN. Sepulveda F. got the relations of force and EMG signals by extracting EMG eigenvalue using time domain. Freriks extracted EMG eigenvalue using RMS and meso-position MF to gain the amplitude and spectral characteristics of bicipital muscle of arm.

\*Corresponding author: Gao yongsheng. E-mail: gaoy@hit.edu.cn.

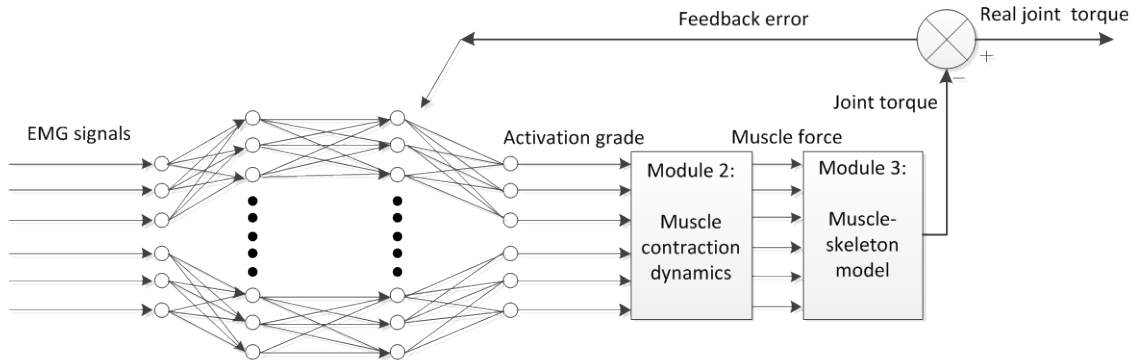
Muscle activation grade is the conceptual expression of muscle stimulation and skeleton-muscle model is driven by it. In this paper using non-linear system identification ability of NN to calculate muscle activation grade, proper activation grade is got to reflex muscle excitement status. Using activation grade as input, with the comparison of the realistic moment and calculated moment generated in joints, the activation grade is verified and the valid model is then obtained.

## Materials

Control action of EMG and joint moment is expressed in figure 1, which includes multiple complex transmission. Usually NN algorithm is used to identify the non-linear relations. Thinking about Hill's research in the expression of module 2 and 3 in mathematical formula, only muscle activation dynamics model(module 1) is not clear in math, so we think the module 1 as a blackbox and identify it using NN. For the activation grade is not measured directly and joint moment  $M$  can be measured, EMG and moment error training sets which are measured can train to setup NN (Fig.2).



**Fig.1 Control relationship of EMG and joint torque**



**Fig.2 activation dynamics of NN with the error training of joint torque**

When choosing BP NN to predict joint moment, the prediction error of joint moment can be used to prediction error of muscle activation grade(Fig.1). It is that when the prediction error of joint moment is  $\Delta M$ , the muscle I is  $\Delta a_i$

$$\Delta M = \frac{\partial M}{\partial a_1} \Delta a_1 + \frac{\partial M}{\partial a_2} \Delta a_2 + \dots + \frac{\partial M}{\partial a_s} \Delta a_s = \sum_{i=1}^s \frac{\partial M}{\partial a_i} \Delta a_i \quad (1)$$

In which  $a_i$  represents the activation grade of muscle I,  $s$  represents the number of muscles which participates contribution of torque,  $M$  represents output joint moment.  $\Delta M$  represents variable quantity of joint moment after the change of activation grade of muscle.

In description of Hill muscle model, joint moment is the function of  $a_i$ (the activation grade of muscle) and  $\theta$  (joint angle). The function can be expressed in formula 2

$$M = M(a_1, a_2, \dots, a_s, \theta) \quad (2)$$

Joint moment  $M$  has the partial derivative  $\frac{\partial M}{\partial a_i}$  of the activation grade of muscle, which can be expressed in the main part of the first order of Taylor expansion as formula 3:

$$\begin{aligned} \delta M &= M(a_1, a_2, \dots, a_s, \theta) - M(a_1, a_2, \dots, (a_i + \delta a_i), \dots, a_s, \theta) \\ \frac{\partial M}{\partial a_i} &\approx \frac{\delta M}{\delta a_i} \end{aligned} \quad (3)$$

In which  $\delta a_i$  is a very small variable at point  $a_i$ ,  $\delta M$  is a variable of joint moment  $M$  which is derived when a  $\delta a_i$  increased in  $a_i$  while others parameters is fixed. Though joint moment prediction error  $\Delta M$  and  $\partial M / \partial a_i$  can be got during the course of prediction, it is not enough to equate  $\Delta a_1, \Delta a_2, \dots, \Delta a_s$  of the prediction error from equation 1. So we need to gain prediction error further by network weights management.

According to the method of counter propagation algorithm, when a network layer has no direct error signals, the signals can be derived from the next network layer. It is similar to the course of using NN to identify EMG and the activation grade of muscle. So the output layer of NN can be seen as a hidden layer and the muscle-skeleton model as an output layer. The weights of the hidden layer are adjusted as:

$$\Delta w_{ji} = -\eta \frac{\partial \xi(n)}{\partial w_{ji}} = \eta \delta_j(n) y_i(n) \quad (4)$$

$$\delta_j(n) = -\frac{\partial \xi(n)}{\partial y_j(n)} \frac{\partial y_i(n)}{\partial v_j(n)} = -\frac{\partial \xi(n)}{\partial y_i(n)} \psi'_j(v_j(n)) \quad (5)$$

The key point of weights adjustment is to calculate the  $\partial \xi(n) / \partial y_i(n)$

The NN signals' output is joint moment after muscle-skeleton model. The error of mean square of the sample is defined as:

$$\xi(n) = \frac{1}{2} (M_{measured} - M_{output})^2; \quad y_i(n) = a_j(n) \quad (6)$$

In which  $M_{measured}$  is the measured joint moment,  $M_{output}$  is the calculated joint moment using activation grade of muscle in NN.  $a_j(n)$  is the activation grade of muscle  $j$ , which is also the output of the output layer of NN. So the calculation formula for the weights of output layer is as follows:

$$\frac{\partial \xi(n)}{\partial y_j(n)} = -(M_{measure} - M_{output}) \frac{\partial M_{output}}{\partial a_j} \quad (7)$$

$$\delta_j(n) = (M_{measure} - M_{output}) \frac{\partial M_{output}}{\partial a_j} \phi'_j(v_j(n)) \quad (8)$$

$$\Delta w_{ji} = \eta \delta_j(n) y_i(n) = \eta (M_{measure} - M_{output}) \frac{\partial M_{output}}{\partial a_j} \phi'_j(v_j(n)) y_i(n) \quad (9)$$

Up to now, the weight adjustment of the NN output is clear. Using traditional formula we can calculate the former layer weights.

## Experiment

### Experimental platform

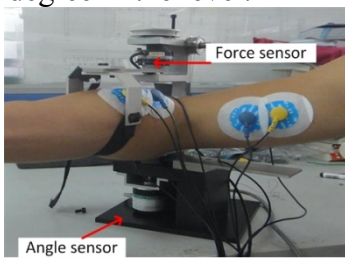
A single freedom experimental platform is designed. During the course of the test, motion of the joints of shoulder and upper arm of the testers is restrained. In the meantime the motion of the wrist is limited in the horizontal plant to avoid the effect of gravity torque. Device structure is as figure 3 which includes rotated stick, fixed stick and a set of pulley. The pulley is to transmit vertical load to horizontal load. Joint angle is measured by subjacent fixed angle encoder and joint torque is

measured by 6-D force sensor which is installed between the rotated stick and fixed stick. The force sensor is SI-80-4 made by ATI Corp.. Data acquisition program of force sensor is compiled with VC++ and acquisition frequency is 256Hz.

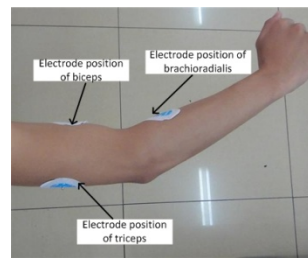
Using multiple channel EMG recorder and one-time Ag/AgCl electrode, main muscles including biceps, brachioradialis and triceps are measured from the muscle group of elbow joint. Electrodes are pasted on the tester's arm, and the position of fossa cubitalia and olecranon is the reference to paste electrode (fig. 4). The use of electrodes refers to Freriks's presentation .

*Experimental procedure*

4 male testers in average age of 25 are selected and they are all healthy with no nervous system or motion disorder. In the rotation course, rotation axis of elbow joint is coaxial with rotation axis of experiment table. So the measured angle of angle sensor is the direct position of elbow angle. Shoulder joint is fixed in 90 degree's position and motion range of elbow's joint is limited during 0 and 130 degree in the level.



**Fig.3 Single freedom platform**



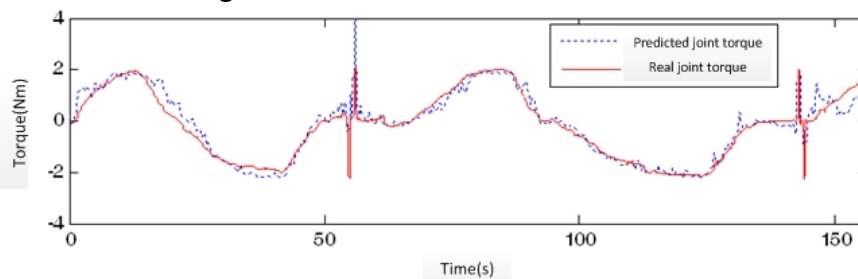
**Fig.4 Electrode position**

Main steps are as follows:

1. Each tester will finish isometric contraction of elbow joint in 120° ,90° ,60° separately, and EMG signals and joint moments are measured.
2. Each tester will finish stretching and curving motions during 3 cycles separately in each degree.
3. Recording EMG signal values in the max isometric contraction to normalize processing of EMG.
4. Choosing No.1 tester to simulate tremor for NN to predict tremor moment.

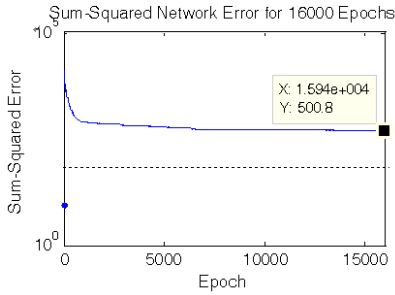
**Results and discussion**

Testers finished experiments in three joint degrees. Using established 0 output network as the original training network, first NN's training is gained by joint angle 120° . With measured EMG signals and corresponding joint moment, after 16000 times training, joint moment after the output of muscle-skeleton model is in figure 5.

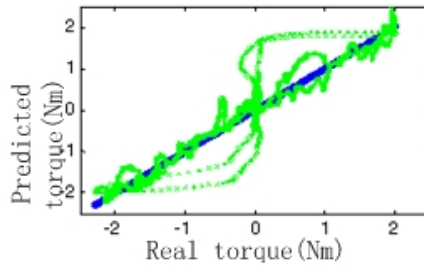


**Fig.5 Result of No.1 tester's NN training in joint angle 120°**

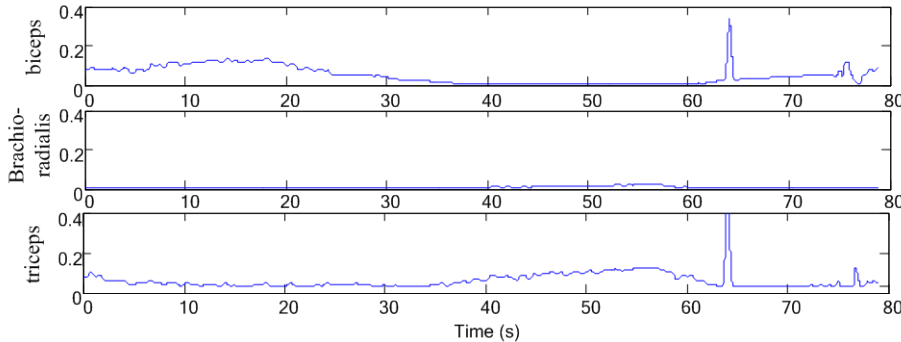
During the training of NN, the change of all samples' error sum of squares of output moment is in figure 6. The figure reflects the good study performance of NN. Figure 7 is the correspondence between real moment and prediction value. The muscle activation grade of NN calculation is in figure 8.



**Fig.6 Error during the torque training**



**Fig.7 Relationship between real and predict value of torque**



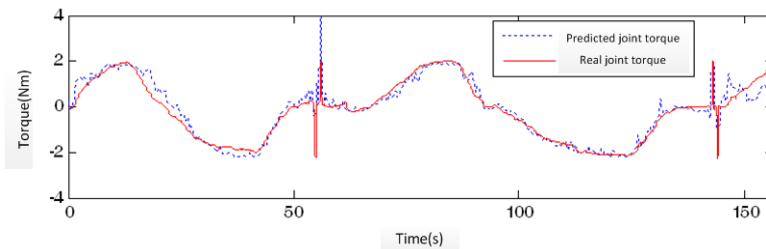
**Fig.8 Activation grade of NN's calculation result**

After the training network of  $120^\circ$ , the network is trained using experimental data of  $90^\circ$  and  $60^\circ$  separately. The average error after training is in table 1.

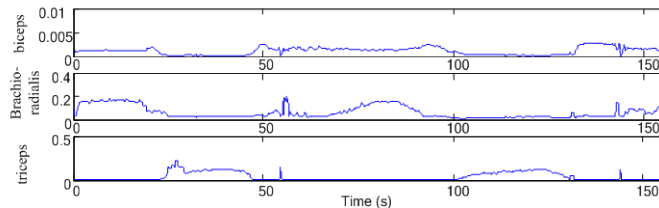
**Table 1 Average error of joint torque training**

Joint angle( $^\circ$ )	60	90	120
Average error(Nm)	0.052	0.065	0.058

We use the trained network to predict joint moment and the prediction data is the RMS values of EMG when No.1 tester did twice stretch and bend motions. The prediction result of joint moment is in figure 9(a). The average error of prediction value and real value is 0.102. The activation grade of NN output is in figure 9(b), in which the activation grade of bicipital muscle of arm is smaller than brachioradialis'. That means that the NN predicts the power of elbow bending from the action of brachioradialis. It results from the multiple solutions of NN.



(a) Prediction result of elbow joint torque

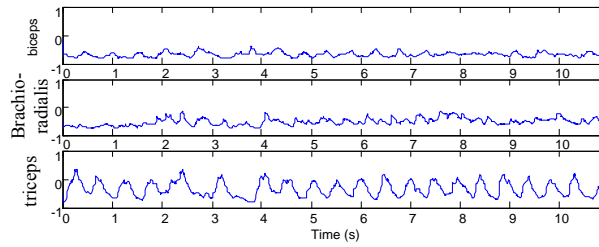


(b) Prediction result of the muscle activation grade

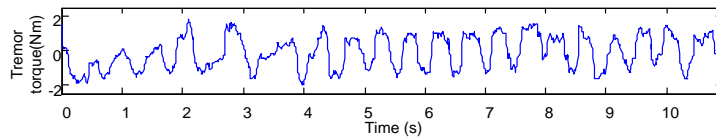
**Fig.9 NN identification for EMG and elbow joint torque**

The results above illustrate that the prediction accuracy of the NN meets requirements and then we can use the NN to predict the joint moment of tremor. A set of EMG signals of elbow tremor is

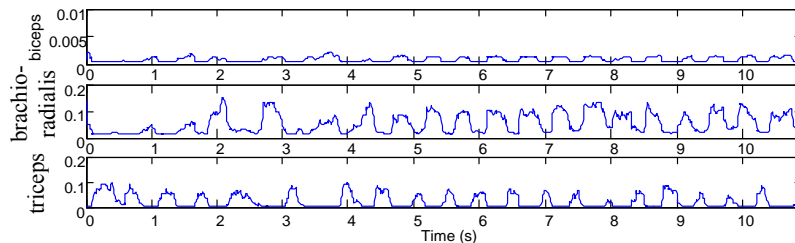
measured as figure 10 and the predicted joint moment through NN model calculation is in figure 11. The activation grade of NN calculation is in figure 12.



**Fig.10 Normalized RMS value of EMG signal**

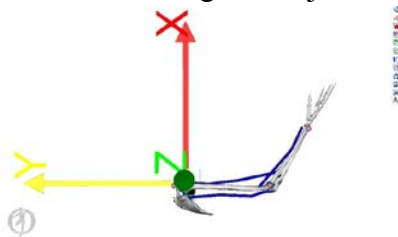


**Fig.11 Prediction value of joint torque of tremor**

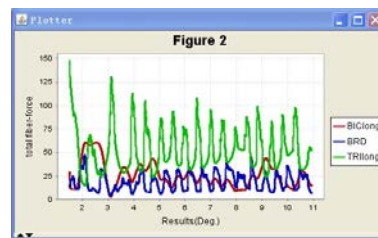


**Fig.12 Activation grade of tremor motion**

The activation grade of NN calculation controls the muscle-skeleton model in OpenSim software environment (fig. 13). The variable curve of each muscle force is in figure 14 during the course of calculating elbow joint tremor.



**Fig.13 Model in OpenSim software**



**Fig.14 Calculated muscle forces**

## Conclusions

An improved neural network is developed to solve the problem of a blackbox of muscle activation dynamics. With the proper weight adjustment in hidden layer, muscle activation grade can be identified. And then using Hill skeleton-muscle model parameters we can get the relationship between the input of EMG signals and the output of joint torque. The experiment verified the valid model.

## Acknowledgments

This study is supported by the National Science Council of the Republic of China (Project No. 60975067)

## References

- Kotovskiy, J., Rosen, M. J. (1998) A wearable tremor-suppression orthosis, *Journal of rehabilitation research and development* **35**, 373-387.
- Hallett, M. (1998) Overview of Human Tremor Physiology, *Movement Disorders* **13**, 43-48.
- Juan, G., Jaime, I., Dideriksen, J. A. (2012) Multimodal Human-Robot Interface to Drive a Neuroprosthesis for Tremor Management, *IEEE Transactions on Systems Man and Cybernetics* **42**, 1159-1168.
- Lanari, B., Antonio, P., Azevedo-Coste, C., Poignet, P. (2011) On the Use of FES to Attenuate Tremor by Modulating Joint Impedance, *50TH IEEE Conference on Decision and Control and European Control Conference (Cdc-Ecc)*, 6498-6503.
- Pennestri, E., Stefanelli, R., Valentini, P. (2007) Virtual musculo-skeletal model for the biomechanical analysis of the upper limb, *Journal of biomechanics* **40**, 1350-1361.
- Zhang, D., Poignet, P., Bo, A. P., et al (2009) Exploring peripheral mechanism of tremor on neuromusculoskeletal model: A general simulation study, *IEEE Transactions on Biomedical Engineering* **56**, 2359-2369
- Hunt, K. J., Fang, J., Saengsuwan, J., et al (2012) On the efficiency of FES cycling: A framework and systematic review, *Technology and Health Care* **20**, 395-422
- Tadej Bajd, Marko Munih (2010) Basic functional electrical stimulation (FES) of extremities: An engineer's view, *Technology and Health Care* **18**, 361-369
- Luh, J. J., Chang, G. C., Cheng, C. K. (1999) Isokinetic elbow joint torques estimation from surface EMG and joint kinematic data: using an artificial neural network model, *Journal of Electromyography and Kinesiology* **9**, 173-183.
- Sepulveda, F., Wells, D. M., Vaughan, C. L. (1993) A neural network representation of electromyography and joint dynamics in human gait, *Journal of biomechanics* **26**, 101-109.
- Freriks, B., Merletti, R., Stegeman, D., et al (1999) European recommendations for surface electromyography, *The Netherlands: Roessingh Research and Development*, 1-8.

# Solution of computational acoustics and wave propagation problems using a high order, high resolution coupled compact difference scheme

Jitenjaya Pradhan, †Y. G. Bhumkar and \*S. D. Dhandole

School of Mechanical Sciences, Indian Institute of Technology Bhubaneswar, Odisha, India, 751013.

\* Presenting author: satish@iitbbs.ac.in

†Corresponding author: bhumkar@iitbbs.ac.in

## Abstract

Here, we report a high order, high resolution coupled compact difference scheme for solving computational acoustics problems. Proposed coupled compact difference scheme displays significant spectral resolution while estimating spatial derivatives and has a physical dispersion relation preserving (DRP) ability over a wide range of wave number when a fourth order four stage Runge-Kutta scheme is used for time integration. Proposed scheme simultaneously computes the first, the second and the fourth derivative in a coupled manner at all the grid points in the domain. We have purposefully evaluated the fourth derivative term using coupled compact difference scheme to add numerical diffusion for the attenuation of unphysical spurious waves in the computed solution.

**Keywords:** Computational Acoustics, Coupled Compact Difference Scheme, DRP Property, High resolution scheme

## Introduction

Propagation of an acoustic wave over a small distance inside a homogeneous medium displays non-dissipative, non-dispersive characteristics [Kinsler et al. (1999)]. Simulation of acoustic wave propagation problem involves computation of acoustic wave field either directly from the linearized compressible flow equations [Tam & Webb (1993)] or by solving hyperbolic partial differential equation for wave propagation [Sengupta (2013)]. The numerical scheme used for solving acoustic problems must have a significant spectral resolution to effectively resolve all the scales present in the acoustic field. Compact schemes provide higher spectral resolution as compared to the explicit difference schemes for same stencil size [Lele (1992), Fung et al. (1995), Chu & Fan (1998), Sengupta et al. (2003), Zhou et al. (2007), Bhumkar et al. (2014)] and are preferred for obtaining highly accurate solutions. As propagation of an acoustic wave displays non-dispersive, non-dissipative and isotropic nature, numerical schemes used for simulating computational acoustic problems must be neutrally stable and preserve the physical dispersion relation numerically [Tam & Webb (1993), Sengupta (2013)]. For an adopted numerical scheme, it is not only important to resolve all physical spatial and temporal scales but also display neutrally stable, DRP nature [Sengupta (2013)]. This has prompted researchers to search for a high resolution, dispersion relation preserving schemes which are useful for computing wave propagation problems [Chiu & Sheu (2009), Tam & Webb (1993), Hu et al. (1996)].

Here, we are proposing a new high order, high resolution coupled compact difference scheme to compute the spatial derivative terms while a fourth order four stage Runge-Kutta scheme has been used for time integration. The derived coupled compact difference scheme evaluates the first derivative, the second derivative and the fourth derivative simultaneously at all the grid points in the domain.



## Methodology

*Stencil for the coupled compact difference scheme:*

Almost all discrete difference computations involve implicit filtering and corresponding solution components in high wavenumber range show spurious nature which are often responsible for numerical instabilities [Sengupta (2013)]. Unphysical amplification of high wavenumber components can be controlled by either using upwind scheme or by using explicit filters [Yu et al. (2015), Visbal & Gaitonde (2002)]. A general stencil for the upwind scheme is given as [Sengupta (2013)],

$$\frac{\partial f}{\partial x} = \left( \frac{\partial f}{\partial x} \right)_{CD} + \alpha \left( \frac{\partial^{2n} f}{\partial x^{2n}} \right) \quad (1)$$

In upwind schemes, one adds explicit numerical diffusion through the even order derivative term in Eq. 1. For adding controlled amount of numerical diffusion, diffusion coefficient  $\alpha$  has been used. In Eq. 1,  $n$  is an integer and the subscript  $CD$  shows derivative has been obtained using a central difference scheme. Diffusion coefficient  $\alpha$  can be either positive or negative based on the direction of propagation of information at a particular point in the domain.

A high accuracy, spectrally optimized upwind CCD scheme has been proposed in [Chiu et al. (2009)], to evaluate the first and second derivative terms together. In the present work, we propose following coupled compact difference scheme with a central stencil to evaluate the first, the second and the fourth derivative terms together. Information associated with the fourth derivative term has been used here to attenuate unphysical spurious waves.

Consider a domain discretized using equi-spaced grid points with a grid spacing  $h$ . The coupled compact difference scheme for simultaneous evaluation of the first, the second and the fourth derivative term is given as,

$$\frac{87}{176} (u'_{j+1} + u'_{j-1}) + u'_j + \left( \frac{-13}{176} \right) h (u''_{j+1} - u''_{j-1}) + \frac{1}{528} h^3 (u^{IV}_{j+1} - u^{IV}_{j-1}) = \frac{1}{h} \left( \frac{47}{48} (u_{j+1} - u_{j-1}) + \frac{1}{132} (u_{j+2} - u_{j-2}) \right) \quad (2)$$

$$\frac{1805}{1224h} (u'_{j+1} - u'_{j-1}) + \left( \frac{-79}{408} \right) (u''_{j+1} + u''_{j-1}) + u''_j + \frac{1}{408} h^2 (u^{IV}_{j+1} + u^{IV}_{j-1}) = \frac{1}{h^2} \left( \frac{1624}{459} (u_{j+1} + u_{j-1}) + \left( \frac{-241}{34} \right) u_j + \frac{11}{1836} (u_{j+2} + u_{j-2}) \right) \quad (3)$$

$$\left( \frac{-13945}{408h^3} \right) (u'_{j+1} - u'_{j-1}) + \frac{755}{136h^2} (u''_{j+1} + u''_{j-1}) + \left( \frac{-13}{136} \right) (u^{IV}_{j+1} + u^{IV}_{j-1}) + u^{IV}_j = \frac{1}{h^4} \left( \frac{-8600}{153} (u_{j+1} + u_{j-1}) + \frac{1920}{17} u_j + \frac{-40}{153} (u_{j+2} + u_{j-2}) \right) \quad (4)$$

In order to obtain the first, the second and the fourth derivative terms, equations (2)-(4) are solved in an iterative and coupled manner. For iterative approach, we propose to evaluate the various derivative terms using explicit central difference schemes as an initial guess. This will reduce the computational cost required for iterative approach as the iterations are performed with reasonable initial guess which is more close to final solution as compared to some random initial guess. Equations (2)-(4) are solved in an iterative manner till the maximum residue becomes less than the prescribed tolerance value which is chosen as  $10^{-6}$  in present study. Here, a simple traditional Gauss-Seidel iterative algorithm has been used.

One can use equations (2)-(4) at all the grid points for the periodic problem. However for the non-periodic problems, different stencils for the boundary and the near boundary nodes are required. In

this regard, we propose to use following stencils for the respective derivative terms at the inlet boundary [Sengupta (2013)]. The second and the fourth derivative terms are usually not required at the boundary nodes where one usually prescribe a Dirichlet boundary condition and hence are not evaluated at the boundary nodes while the fourth derivative term is not evaluated at the second and second last node. Same stencils can be used for the boundary and the near boundary points on the other end by reverting the stencils and adding a minus sign [Sengupta (2013)]. Thus, we have used explicit stencils for the boundary and the near boundary nodes as,

$$u^I_1 = (-1.5u_1+2u_2-0.5u_3)/h; u^{II}_1 = 0; u^{IV}_1=0 \quad (5)$$

$$u^I_2 = (u_3-u_1)/(2h); u^{II}_2 = (u_1-2u_2+u_3)/h^2; u^{IV}_2=0 \quad (6)$$

$$u^I_3 = (-u_5+u_1+8(u_4-u_2))/(12h); u^{II}_3 = (u_1-2u_2+u_3)/h^2; u^{IV}_3=(u_1+u_5-4(u_2+u_4)+6u_3)/h^4 \quad (7)$$

Finite difference schemes depend on information available at the nearby points to estimate derivative values. Taylor series approximation is used to derive stencil for difference scheme in such a way that the lower order derivative terms are matched accurately while the higher order derivatives terms are truncated. Thus the numerically estimated derivative and the exact derivative values differ due to the truncation error. This is also known as implicit filtering associated with the difference schemes [Sengupta (2013)]. Different finite difference schemes can have same order of accuracy however different spectral resolution while evaluating derivative terms. In such case, the scheme having higher spectral resolution will produce more accurate results as compared to low resolution schemes. Thus, it is imperative to evaluate performance of numerical scheme based on its spectral resolution and not on its order of accuracy [Sengupta (2013)]. Here, we have estimated spectral resolution of the proposed scheme using the full domain matrix global spectral analysis (GSA) technique in [Sengupta (2013)]. Details about this technique are not provided here to avoid repetition.

Following the work in [Sengupta (2013)], if one denotes  $K$  and  $K_{eq}$  as the exact and the numerically obtained wavenumber in a difference computation then the discretization effectiveness for the first, the second and the fourth derivative can be obtained as shown in Figure 1(a)-(c), respectively. Figure 1(a) compares the effectiveness of spectral resolution ( $K_{eq}/K$ ) for the proposed coupled compact difference scheme and a 12<sup>th</sup> order compact difference scheme. Figure shows proposed scheme has even better spectral resolution as compared to the 12<sup>th</sup> order compact difference scheme. Discretization effectiveness for the second and the fourth derivative terms show near spectral resolutions. Thus the proposed coupled compact difference scheme has significantly improved spectral resolution as compared to the existing difference schemes.

Apart from estimation of effectiveness for spatial derivative terms, one needs to estimate combined effects of spatial and temporal discretization terms for solving unsteady problems. Consider a 1D wave equation which also serves as a model equation for computational acoustics problems as,

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \quad (8)$$

Simulation of the computational acoustics problems involves estimation of both space and time derivative terms. For such unsteady problems, one needs to first estimate important numerical properties such as the numerical amplification factor  $|G|$  and the numerical group velocity  $V_{gN}/c$ . Variation of these important numerical properties with respect to non-dimensional wavenumber  $Kh$  and CFL number  $Nc$  has been discussed in [Sengupta (2013)] for the solution of 1D wave equation using different discretization schemes. Using the same methodology, we obtained the contours of the

numerical amplification factor  $|G|$  and the numerical group velocity  $V_{gN}/c$  for the solution of 1D wave equation when coupled compact difference scheme has been used for spatial discretization while RK4 scheme is used for time integration.

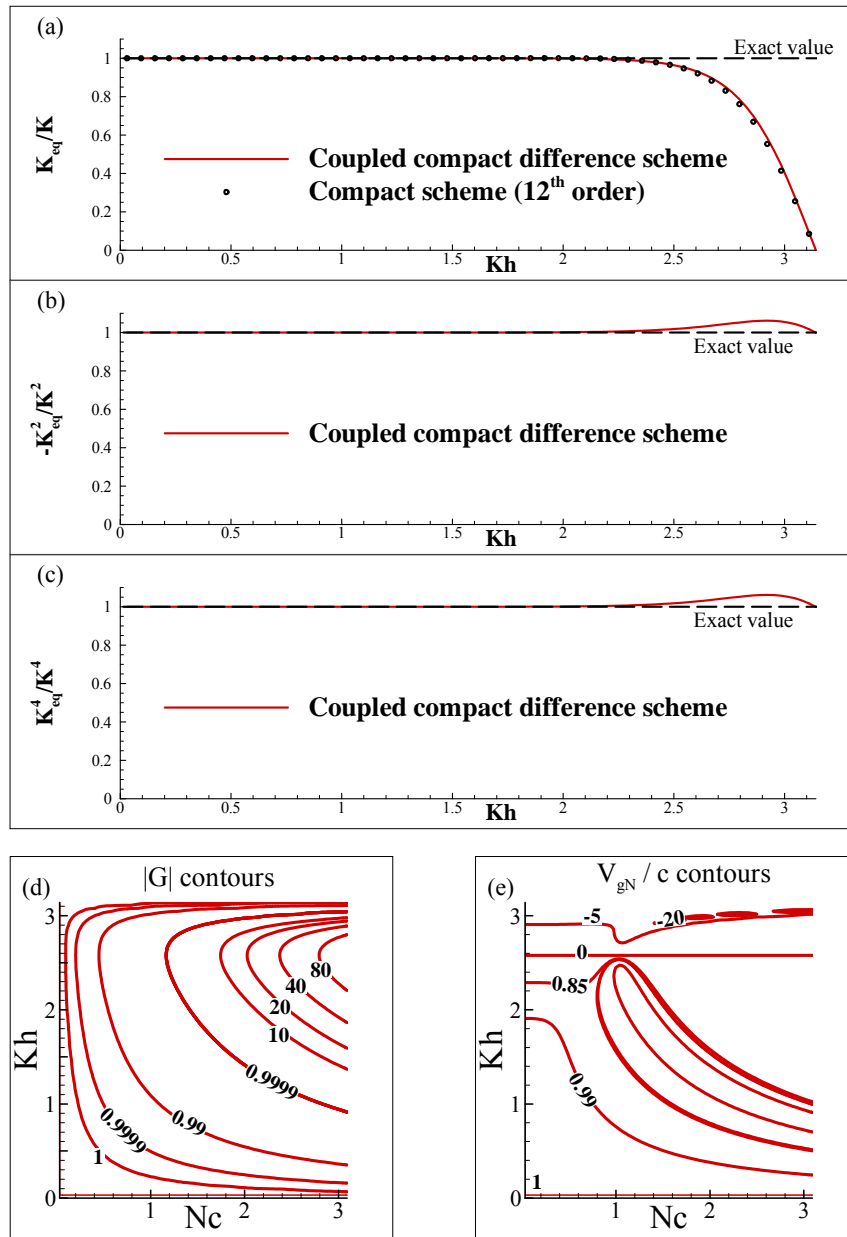


Figure 1: (a) Comparison of the effectiveness of spectral resolution ( $K_{eq}/K$ ) for different spatial discretization schemes to evaluate the first derivative term at the central node; Discretization effectiveness of the second and the fourth derivative terms is shown in (b) and (c) respectively; Contours for the variation of the numerical amplification factor  $|G|$  and the numerical group velocity  $V_{gN}/c$  are shown in (d) and (e), respectively when the coupled compact difference scheme is used with the fourth order RK4 scheme for time integration to solve Eq. (8).

Figure 1 (d) and (e) shows numerical properties corresponding to the central node. Numerical amplification factor contours in Fig. 1(d) show that for a small CFL number  $Nc$ , one observes a

neutrally stable region across a complete wavenumber range. Non-dimensionalized numerical group velocity contours in Fig. 1(e) show that for a small CFL number, physical dispersion relation has been preserved accurately up to non-dimensional wavenumber  $Kh = 1.9$ . Thus present scheme has significant DRP ability. One observes presence of negative group velocity above  $Kh=2.6$ . This region has been identified as a q-wave region [Sengupta (2013)] in which solution components not only travel with wrong velocity but also in wrong direction. Such waves are often responsible for numerical instabilities.

Spurious waves are often triggered due to presence of sharp discontinuities, irregularly spaced grid points, discontinuities in the initial and boundary conditions [Sengupta (2013)]. One needs to attenuate these spurious waves by addition of numerical diffusion to avoid numerical instabilities. One can add numerical diffusion using the information associated with the fourth derivative evaluated at each grid point as shown in Eq. (1). Amount of added numerical diffusion directly depends on diffusion coefficient  $\alpha$ . Figure 2 (a)-(b) show variation of numerical amplification factor and numerical group velocity contours for the solution of 1D wave equation when the indicated diffusion coefficient is used to obtain the upwind coupled compact difference scheme. Figure 2 shows that with increase in  $\alpha$ , scheme displays more and more stable behavior for a small CFL number across a complete wavenumber range  $Kh$ . Thus one can construct upwind coupled compact difference scheme to damp out unphysical, spurious components from the solution.

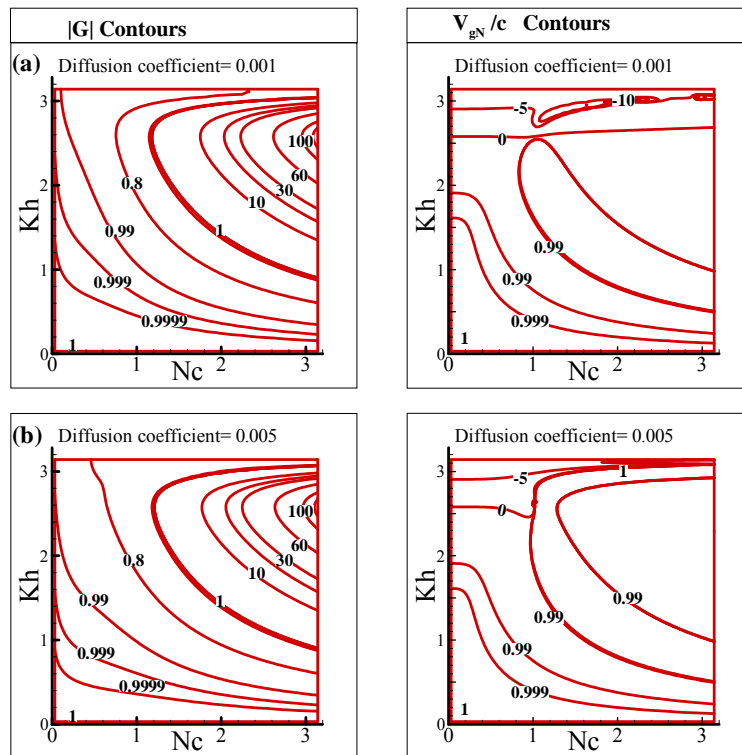


Figure 2: Comparison of the numerical amplification factor  $|G|$  contours and the numerical group velocity  $V_{gN}/c$  contours for the central node is shown for different diffusion coefficients. Contours are obtained for the solution of Eq. (8) when the coupled compact difference scheme has been used with the fourth order RK4 scheme for time integration following work in [Sengupta (2013)]. Note that with addition of numerical diffusion, stability is achieved ( $|G| < 1$ ) for a small CFL number.

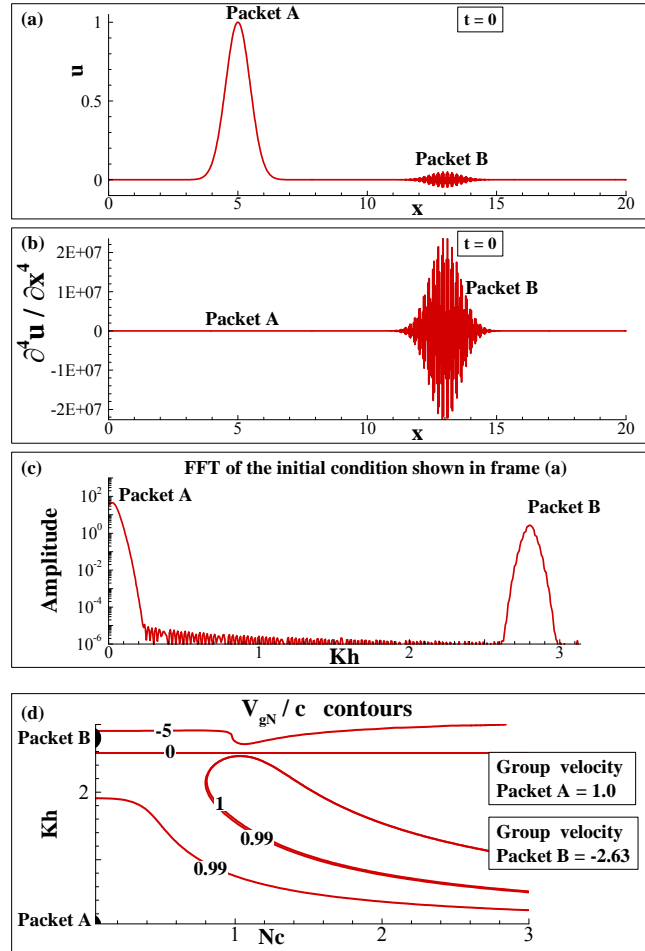


Figure 3: Initial condition of the wave packet along with its fourth derivative and FFT of the initial condition have been shown in (a), (b) and (c), respectively. Numerical group velocity contours for the solution of Eq. (8) have been shown in (d).

## Results and Discussion

In this section, we use the coupled compact difference scheme to solve the model wave equation problems as well as for solving computational acoustics problems.

### 1. Solution of 1-D wave equation for wave propagation problem.

We have chosen this problem to verify the efficacy as well as advantages of the proposed scheme while solving unsteady problems. Here, we have obtained solution of 1D wave equation Eq. (8) subjected to the initial condition as shown in Fig. 3(a). We have purposefully designed the initial condition as a combination of two different wave packets, packet A and packet B. Figure 3(b) shows the fourth derivative of the initial condition and it indicates large values corresponding to packet B due to rapid variation of amplitude associated with packet B. The FFT of the initial condition has been shown in Fig. 1(c). Central wavenumber of packet A is very small while that for packet B is close to the Nyquist limit. In figure 1(d), we have shown the numerical group velocity contours for the solution of Eq. (8) when the spatial discretization has been obtained using coupled compact

difference scheme while RK4 scheme has been applied for time integration. We have also marked the central wavenumbers corresponding to packet A and packet B.

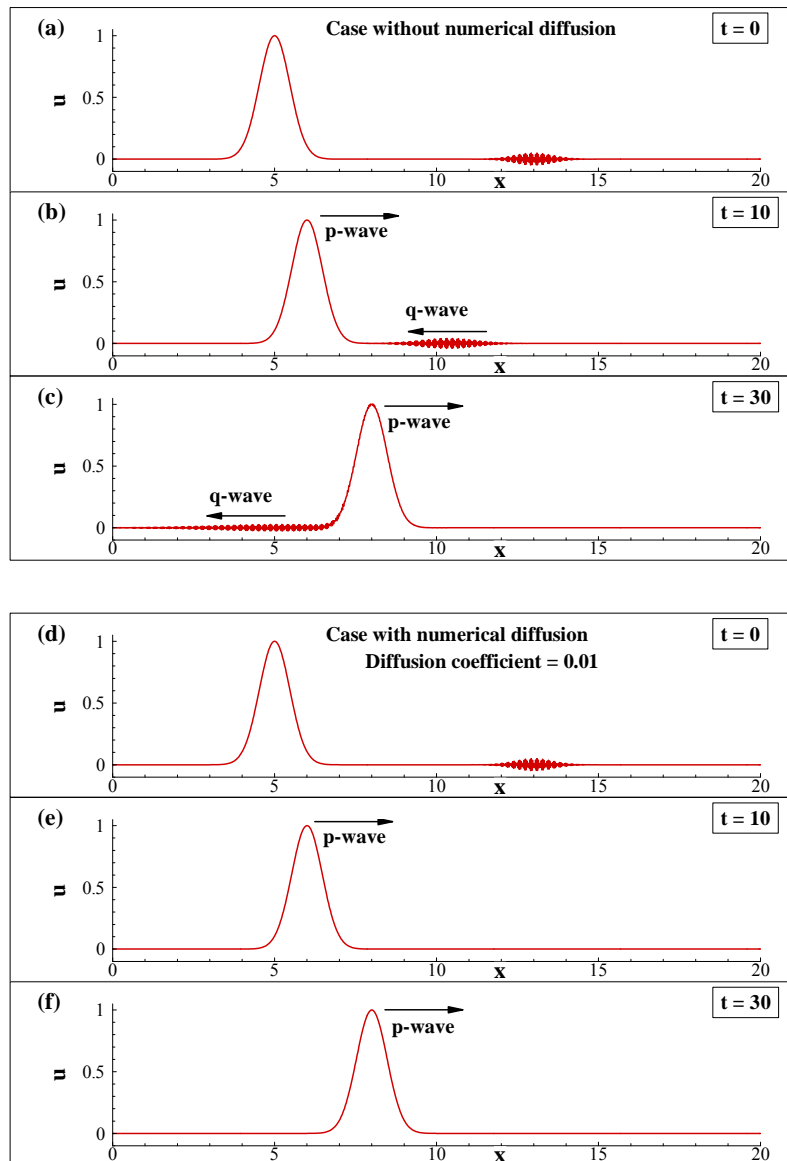


Figure 4: Solutions of 1D wave equation Eq. (8) without and with addition of numerical diffusion have been shown in (a)-(c) and (d)-(f), respectively. Note, in frames (d)-(f) addition of numerical diffusion attenuates spurious packet B which is present in frames (a)-(c) when no diffusion has been added.

We have considered a domain  $0 \leq x \leq 20$  with 101 equi-spaced grid points. Phase speed  $c$  is kept as 0.10. Computations are performed using the coupled compact difference scheme for the spatial discretization and RK4 scheme for the time integration by keeping CFL number as 0.01. Corresponding  $V_{gN}/c$  contours show group velocity for the packet A as 1.0 while that for packet B as -2.63. Figures 4(a)–(c) show the solution of Eq. (8) subjected to the initial condition in Fig. 3(a). Due to positive group velocity packet A propagates towards right hand side while packet B displays spurious nature and propagates in completely opposite direction towards left. In order to attenuate

such spurious waves and prevent numerical instabilities, one can add numerical diffusion as shown in Eq. (1). Figures 4(d)-4(f) show propagation of wave packet when numerical diffusion has been added with a diffusion coefficient as 0.01. Due to addition of numerical diffusion, spurious packet B gets attenuated completely while packet A retain itself and travel towards correct direction with correct velocity. This shows the advantage of coupled contact difference scheme.

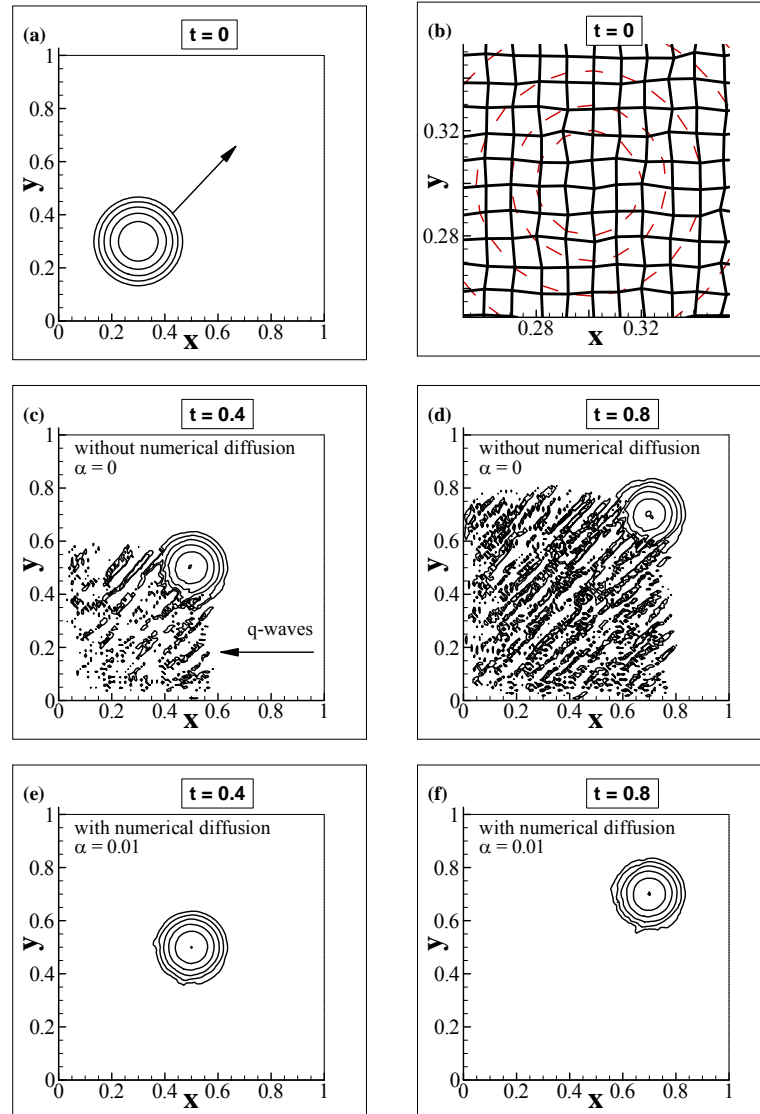


Figure 5: Initial condition of a 2D wave packet and zoomed view of a grid are shown in (a) and (b), respectively. Solutions of 2D wave equation Eq. (9), without and with addition of numerical diffusion are shown in (c)-(d) and (e)-(f), respectively. Note, in frames (e)-(f) addition of numerical diffusion attenuates spurious packets present in frames (c)-(d) when no diffusion has been added.

### 2. Propagation of a wave packet on a discontinuous grid

Next, we consider a propagation of a wave packet inside a 2D domain following the 2D wave equation given as,

$$\frac{\partial u}{\partial t} + c_x \frac{\partial u}{\partial x} + c_y \frac{\partial u}{\partial y} = 0 \quad (9)$$

In Eq. (9),  $c_x$  and  $c_y$  are phase speeds in x- and y-direction. In order to check the performance and applicability of the present scheme, we have purposefully considered wave propagation on a discontinuous grid as the discontinuous distribution of grid points triggers spurious high wavenumber oscillations. Figures 5 (a) and (b) show the initial condition of a 2D wave packet and zoomed view of a grid, respectively. We have obtained solutions of 2D wave equation following Eq. (9) using coupled compact difference scheme for the spatial discretization terms and RK4 scheme for time integration. We have constructed a domain  $0 < x, y < 1$ , with 101 points in either direction. We have purposefully assigned a random distribution to grid point spacing so as to test the efficacy of the coupled compact difference scheme. Figures 5(c) and (d) show propagation of 2D wave packet at the indicated instants. Due to discontinuous distribution of grid points one observes large amount of spurious q-waves in the domain. However, when a fourth order numerical diffusion has been added to the solution, spurious waves are attenuated. This shows the advantage of proposed coupled compact difference scheme while working on a discontinuous grid.

### 3. Propagation of the acoustic and the entropic disturbances.

Next, we solve the computational acoustic wave propagation problem which consists of simultaneous propagation of acoustic and entropic disturbances. These disturbances propagate following the linearized compressible Navier-Stokes equations given by [Tam et al. (1995)],

$$\frac{\partial U}{\partial t} + \frac{\partial E}{\partial x} + \frac{\partial F}{\partial y} = 0 \quad (10)$$

Where,

$$U = \begin{bmatrix} \rho \\ u \\ v \\ p \end{bmatrix}, \quad E = \begin{bmatrix} M_x \rho + u \\ M_x u + p \\ M_x v \\ M_x p + u \end{bmatrix}, \quad F = \begin{bmatrix} M_y \rho + u \\ M_y u \\ M_y v + p \\ M_y p + v \end{bmatrix} \quad \text{with } M_x = 0.5, M_y = 0$$

Initial condition for this problem is given as,

$$\begin{aligned} p &= \exp\left[-(\ln 2)\left(\frac{x^2 + y^2}{9}\right)\right]; \quad \rho = \exp\left[-(\ln 2)\left(\frac{x^2 + y^2}{9}\right)\right] + 0.1 \exp\left[-(\ln 2)\left(\frac{(x-67)^2 + y^2}{25}\right)\right] \\ u &= 0.04y \exp\left[-(\ln 2)\left(\frac{(x-67)^2 + y^2}{25}\right)\right] \\ v &= -0.04(x-67) \exp\left[-(\ln 2)\left(\frac{(x-67)^2 + y^2}{25}\right)\right] \end{aligned} \quad (11)$$



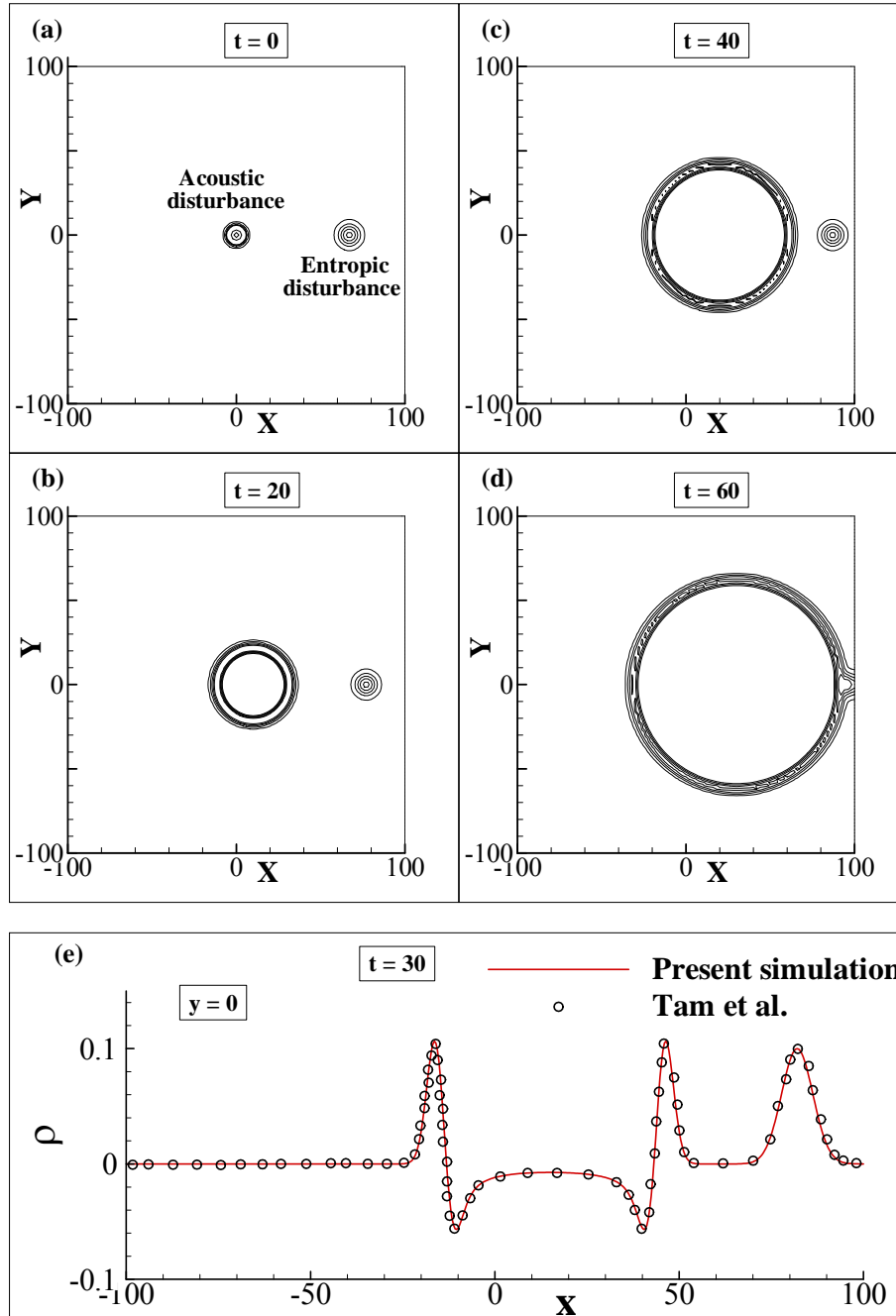


Figure 6: Initial condition and propagation of acoustic and the entropic disturbances following Eq. (10) are shown in frames (a)-(d). Comparison of density variation on the line  $y=0$  obtained from present simulation with that of [Tam et al. (1995)] is shown in (e).

This case consists of an acoustic pulse generated by a Gaussian pressure distribution at the center of the computational domain as shown in the initial condition in Fig 6. The mean flow Mach number is 0.5. We have constructed the domain using 501 X 501 grid points. Downstream of the pressure pulse, at  $x = 0.67$  an entropy pulse has also been superimposed. Acoustic pulse travels faster than entropy disturbances in the downstream direction as observed in Figs. 6(b) to 6(d) which show development and propagation of acoustic as well as entropic disturbances with time. We have compared the density

variation on the line  $y=0$  obtained from present simulation with that of [Tam et al. (1995)] in Fig. 6(e). Comparison shows a good match and justifies use coupled compact difference scheme for obtaining high accuracy solutions of computational acoustics problems.

## Conclusions

Here, we have proposed a new coupled compact difference scheme to solve computational acoustics problems. Proposed scheme has significant resolution and physical dispersion relation preserving ability. In addition, one can add controlled amount of numerical diffusion to attenuate spurious waves in the solution. Solution of model wave propagation problems and computational acoustic problem highlights the advantages of the proposed coupled compact difference scheme.

## References

- Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (1999). Fundamentals of acoustics. Fundamentals of Acoustics, 4<sup>th</sup> Edition, by Lawrence E. Kinsler, Austin R. Frey, Alan B. Coppens, James V. Sanders, pp. 560. ISBN 0-471-84789-5. Wiley-VCH, December 1999.
- Tam, C. K., & Webb, J. C. (1993). Dispersion-relation-preserving finite difference schemes for computational acoustics. *Journal of computational physics*, 107(2), 262-281.
- Sengupta, T. K. (2013). High accuracy computing methods: fluid flows and wave phenomena. Cambridge University Press.
- Chiu, P. H., & Sheu, T. W. (2009). On the development of a dispersion-relation-preserving dual-compact upwind scheme for convection–diffusion equation. *Journal of Computational Physics*, 228(10), 3640-3655.
- Hu, F. Q., Hussaini, M. Y., & Manthey, J. L. (1996). Low-dissipation and low-dispersion Runge–Kutta schemes for computational acoustics. *Journal of Computational Physics*, 124(1), 177-191.
- Lele, S. K. (1992). Compact finite difference schemes with spectral-like resolution. *Journal of computational physics*, 103(1), 16-42.
- Fung, K. Y., Man, R. S., & Davis, S. (1995, May). A compact solution to computational acoustics. In *ICASE/LARC workshop on benchmark problems in computational aeroacoustics (CAA)* (Vol. 1, pp. 59-72).
- Chu, P. C., & Fan, C. (2000). A three-point sixth-order staggered combined compact difference scheme. *Mathematical and computer modelling*, 32(3), 323-340.
- Sengupta, T. K., Ganeriwal, G., & De, S. (2003). Analysis of central and upwind compact schemes. *Journal of Computational Physics*, 192(2), 677-694.
- Zhou, Q., Yao, Z., He, F., & Shen, M. Y. (2007). A new family of high-order compact upwind difference schemes with good spectral resolution. *Journal of Computational Physics*, 227(2), 1306-1339.
- Bhumkar, Y. G., Sheu, T. W., & Sengupta, T. K. (2014). A dispersion relation preserving optimized upwind compact difference scheme for high accuracy flow simulations. *Journal of Computational Physics*, 278, 378-399.
- Yu, C. H., Bhumkar, Y. G., & Sheu, T. W. (2015). Dispersion Relation Preserving Combined Compact Difference Schemes for Flow Problems. *Journal of Scientific Computing*, 62(2), 482-516.
- Visbal, M. R., & Gaitonde, D. V. (2002). On the use of higher-order finite-difference schemes on curvilinear and deforming meshes. *Journal of Computational Physics*, 181(1), 155-185.
- Tam, C. K.W., Shen H., Kurbatskii K.A., Auriault L., Dong Z. and Webb, J.C. (1995). Solutions of the benchmark problems by the dispersion-relation-preserving scheme. In *ICASE/LARC workshop on benchmark problems in computational aeroacoustics (CAA)* (Vol. 1, pp. 149-172).

## **New model and method simulating effective protection of space vehicles from high-velocity debris**

**<sup>\*,†</sup>A.V. Gerasimov<sup>1</sup>, and Yu. F. Khristenko<sup>1</sup>**

<sup>1</sup>Research Institute of Applied Mathematics and Mechanics, Tomsk State University, Russia)

\*Presenting author: [ger@mail.tomsknet.ru](mailto:ger@mail.tomsknet.ru)

†Corresponding author: [ger@mail.tomsknet.ru](mailto:ger@mail.tomsknet.ru)

### **Abstract**

Protection of space vehicles from natural and man-made debris is a significant practical problem. Thin barriers, placed in front of the main body of the object, crush high-velocity particles into fragments and thereby reduce probable penetration. The replacement of the solid plate by a grid of high-strength material lightens the protective structure, with clear benefit for space vehicle applications.

**Keywords:** Space Vehicles, Debris, Grid, High - velocity interaction, Lagrangian method, Probabilistic approach, Fragmentation

### **Introduction**

The process of high-velocity interaction between grid barriers and compact elements (aluminum balls) is considered in Lagrangian 3-D formulation. To solve this problem it is necessary to have a reliable and sufficiently universal method to enable adequate reproduction of the processes occurring in solids under high-velocity collision. The natural heterogeneity of the structures of barrier and projectile materials affects the distribution of physical-mechanical characteristics (PMC) of the material, and is one of the most important factors determining the fracture behavior of real materials. One can account for it in the equations of deformable solid mechanics using a random distribution of the initial deviations of the strength properties from a nominal value (simulating the initial heterogeneity of the material). The relations of deformable solid mechanics, used in major recent works on dynamic fracture of structures and materials, ignore this factor. It can distort a real pattern of impact and explosive fractures of the bodies under consideration. The latter is particularly evident in the solution of axisymmetric problems, where the properties at all points on the circumferential coordinate of a calculated element are initially equal due to the use of standard equations of continuum mechanics in the numerical simulation. However in practice there are many problems where fragmentation is mainly a probabilistic process, for example the explosive fracture of axisymmetric shells, where the fragmentation pattern is unknown beforehand, or penetration and fracture of thin barriers by a projectile along surface normal, etc. The introduction of a random distribution of the initial deviations of the strength properties from a nominal value in the PMC of the body leads to the fact that, in these cases, the process of fracture becomes probabilistic in nature, and more consistent with the experimental data.

## Basic relations and solution method

To describe the processes of deformation and fracture of solids we used a model of a compressible and perfectly elastoplastic body. Basic relations describing the medium motion are based on the laws of conservations of mass, momentum and energy [Stanjukovich (1975); Wilkins M.L. (1964)], and are closed by Prandtl - Reuss relations under the von Mises flow condition. The equation of state was taken in the form of Tate and Mie - Gruneisen [Stanjukovich (1975)]. Plastic deformations, pressure and temperature are known to affect yield strength and shear modulus, therefore the model was supplemented by the relations recommended in [Steinberg et al. (1980)]. Reaching a limiting value for equivalent plastic deformation was used as a fracture criterion at intensive shear deformations [Stanjukovich (1975); Kreyenhagen et al. (1970)].

To calculate elastoplastic flows we used a procedure implemented on tetrahedral cells and based on the combined application of Wilkins method [Wilkins M.L.(1964; (1985)] intended to calculate the internal points of the body and Johnson method [Johnson et al. (1979); Johnson (1981)] to calculate contact interactions. The number of computational cells was about 800 000. To solve the problem, the authors used their own 3-D program [Gerasimov (2007)].

The initial heterogeneities of the structure were simulated by imposing a distribution of ultimate equivalent plastic strain in the cells of the calculated domain by means of a modified random number generator issuing a random variable obeying the selected distribution law. The probability densities of the random variables used were in the form of a normal Gaussian distribution with arithmetical mean being equal to the tabulated value and variable dispersion. The ideology and methodology of a probabilistic approach to the problem of the fracture of solids is completely described in [Gerasimov (2007)].

## The experimental results

High - velocity collisions between projectiles and grid barriers were also experimentally investigated using two-stage light-gas installations modified to throw particles 0.5 mm in diameter at velocities (2.5 - 5) km/s [Gerasimov (2007)]. Projectile - aluminum particles (2 mm in diameter) and a barrier - two steel grids (wire diameter 0.32 mm and 0.2 mm, respectively). The first grid: steel, density = 7.9 g/cm<sup>3</sup>, the shear modulus = 86 GPa, yield point = 0.28 GPA, cell size 0.5 × 0.5 mm, wire diameter = 0.32 mm. The second grid: steel, density = 7.9 g/cm<sup>3</sup>, the shear modulus = 86 GPa, yield point = 0.28 GPA, the cell size 0.356 × 0.356 mm, wire diameter = 0.2 mm. The tank: duralumin, density = 2.64 g/cm<sup>3</sup>, shear modulus = 26.9 GPa, yield point = 0.12 GPA, thickness = 1.9 mm.

Figure 1 shows the investigated assembly, consisting of a tank element of a space vehicle and two grids, which were placed at some distance from each other and the tank. The numbers in Figure 1 indicate the results of separate experiments. The Table 1 presents a brief description of two typical experiments on the interactions of aluminum projectiles, a grid barrier and the element of the protected tank, where V =

initial velocity of the projectile. As is seen from the results of the experiments there were no penetrations or damage to the tank wall, thus the grid barrier has successfully fulfilled its function.

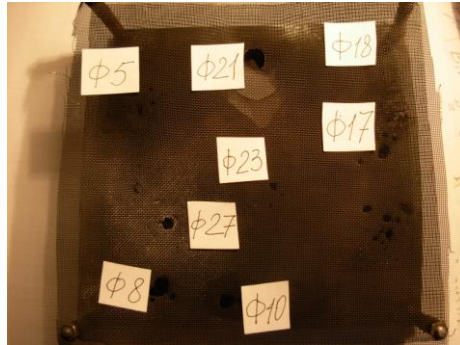


Figure 1. The face of assembly "grid - tank element"

Table 1. Experimental data

Test number	Projectile	Tested barrier and protection	V, km / sec	Experiment results
Φ 5	Duralumin ball 2 mm in diameter	"tank" + 0.32 mm grid + 0.2 mm grid	2.0	Through penetration both grids. The depth of craters from the fragments is 0.3-0.5 mm
Φ 23	Duralumin ball 2mm in diameter	"tank" + 0.32 mm grid + 0.2 mm grid	3.0	Through penetration the both grids. The depth of craters from the fragments is 0.27-0.45 mm

**Numerical results**

The results of numerical calculations of the projectile interaction with the first grid are presented below. Figure 2 presents a top view of a spherical particle and of the first grid element, as well as a two-dimensional section of the configuration.

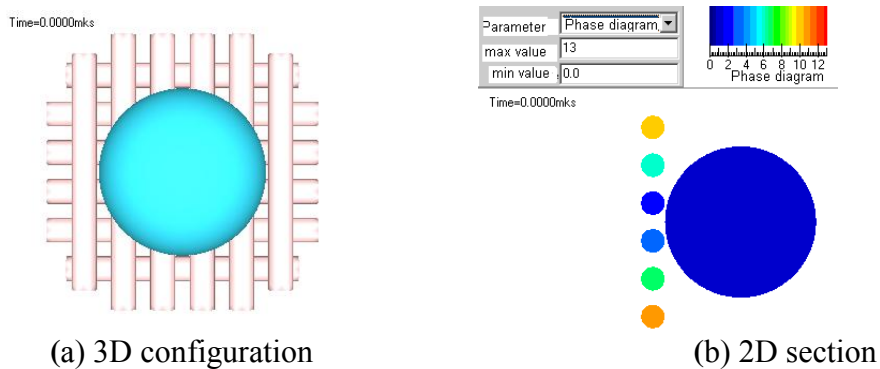


Figure 2. The initial configuration of the grid and a spherical particle

The calculations proved the experimental data on the protective properties of grid barriers. Then numerical experiments were carried out for the velocities of the particle collision with the first grid barrier  $V = 2 \text{ km/s}$ .

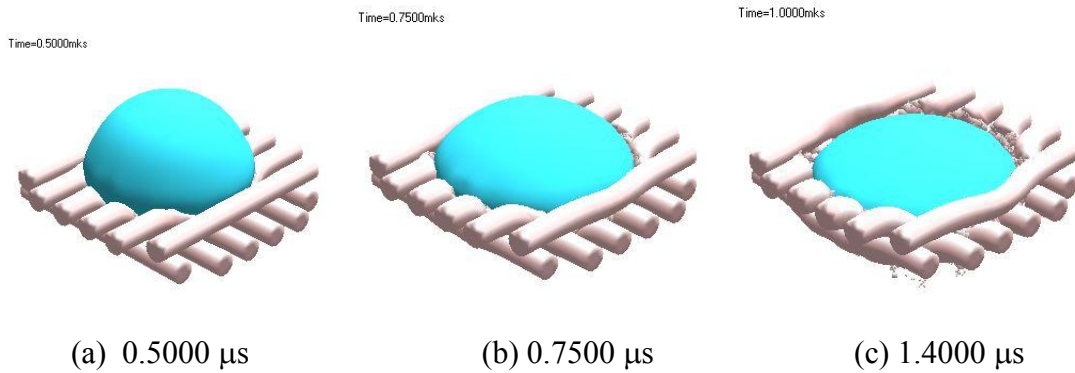


Figure 3. The configurations of the grid and the spherical particle in 3D image for  $V=2 \text{ km/s}$

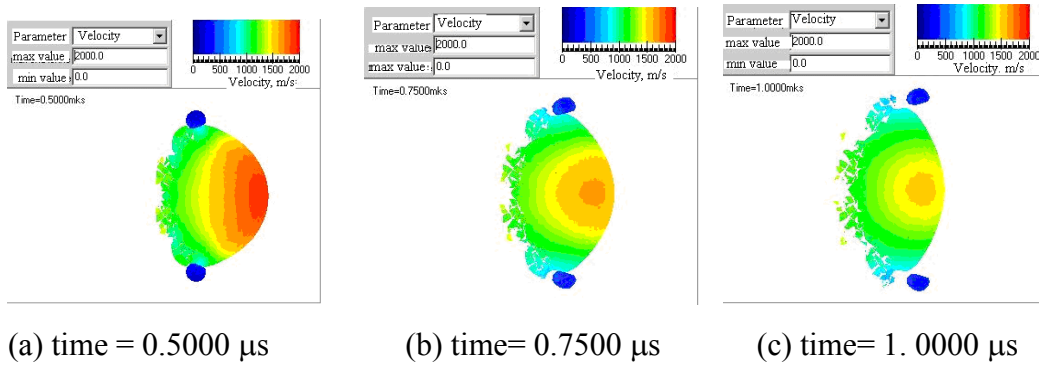


Figure 4. 2D Section of the grid and the spherical particle. Velocity distribution.  $V=2 \text{ km/s}$

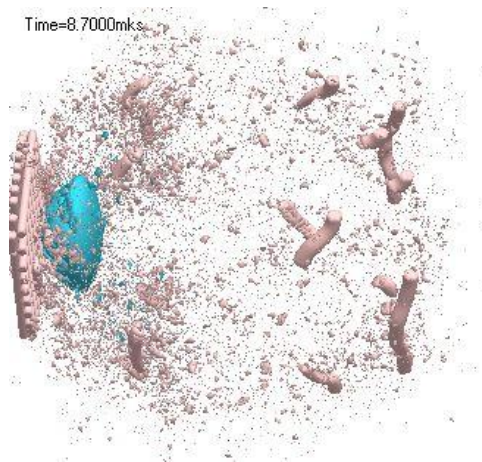


Figure 5. Contact of cloud fragments with the second barrier (grid)  $t=8.7000\mu\text{s}$   $V=2 \text{ km/s}$

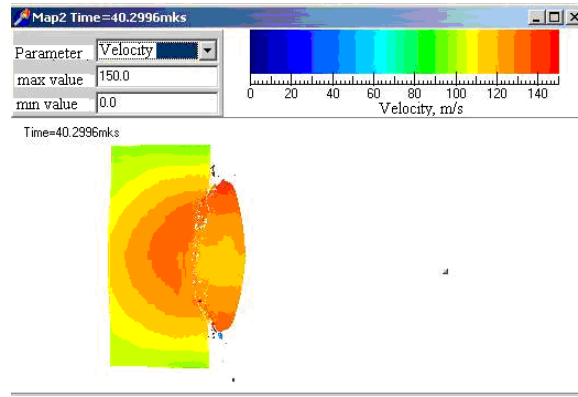


Figure 6. Particle fragment impact on the main body: distribution of current velocity, time = 40.2996  $\mu$ s,  $V=2$  km/s

Figure 6 demonstrates that the particle fragment did not penetrate the main body of the space vehicle and made only a small dent on the surface. The same result was obtained for the collision velocity of  $V = 3$  km/s.

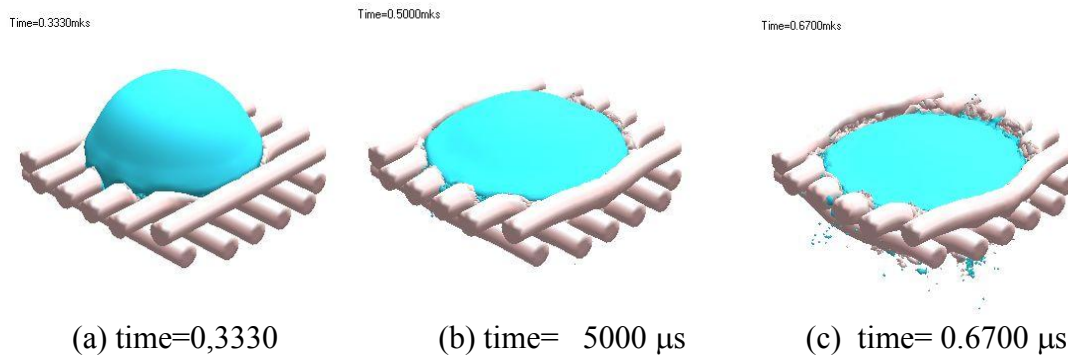


Figure 7. The configuration of the grid and the spherical particle in 3D image for  $V=3$  km/s

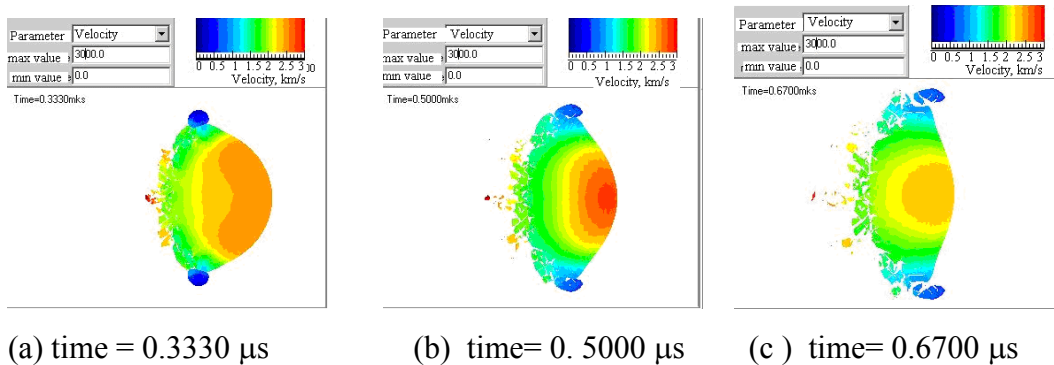


Figure 8. 2D Section of the grid and the spherical particle. Velocity distribution.  $V=3$  km/s

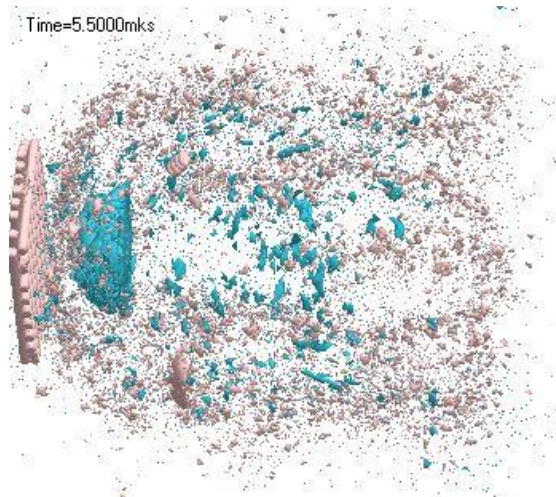


Figure 9. Contact of cloud fragments with the second barrier (grid)  
 $t=5.5000\mu\text{s}$   $V=3\text{ km/s}$

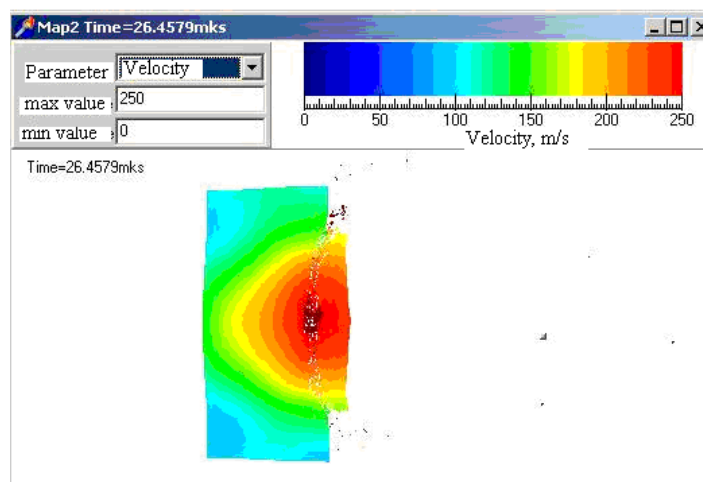


Fig. 10. Particle fragment impact on the main body: distribution of current velocity,  
 $\text{time} = 26.4579\ \mu\text{s}$ ,  $V=3\text{ km/s}$

As seen in Figures 3-10, the fragmentation of the incident particle intensifies with increasing collision velocity, as evidenced by the intense formation of material jets penetrating through the meshes (Fig. 7c). In figures number 6 and number 10 show a portion of the ball passing through the two grids, which does not penetrate the tank shell. The results for the initial velocity  $V = 2\text{ km/s}$  and  $V = 3\text{ km/s}$ , respectively. Dimensions calculated crater close to the size of the experimental craters (Table 1).

### Conclusion

This paper presents numerical research into the interaction of high-velocity particles and barriers – grids which are used to protect space vehicles — and demonstrates the efficiency of grid shielding structures.

The data obtained proved the possibility of the proposed approach and Lagrangian numerical techniques were used to reproduce completely, from the physical point of view, in three-dimensional formulation the interactions of the spaced grid barriers and



protected space vehicle elements with high- velocity particles from fractured structures and vehicles, as well as fragments of cosmic bodies, and to inform the selection of the most effective protection systems.

In the course of numerical experiments the protection consisting of the two grids proved to be more effective as compared with an equivalent by mass solid barrier. The experimental results proved the adequacy of the numerical method and the simulations provide some qualitative support for the proposed protection mechanism.

## References

- Stanjukovich, K.P. (Ed) (1975) *Physics of Explosion*. Nauka, Moscow [in Russian].
- Wilkins, M.L. (1964) Calculation of elastic-plastic flow. *Methods of Computational Physics*, Academic Press, New-York.
- Steinberg, D.J., Cochran, S.G., Guinan, M.W. (1980) A constitutive model for metals applicable at high – strain rate, *Journal of Applied Physics* **51**, 1496-1504.
- Kreyenhagen, K.N., Wagner, M.H., Piechocki, J. J., Bjork, R. L. (1970) Ballistic limit determination in impact on multimaterial laminated targets, *AIAA Journal* **8**, 2147– 2151.
- Johnson, G.R., Colby, D.D., Vavrick, D.J. (1979) Tree-dimensional computer code for dynamic response of solids to intense impulsive loads, *International Journal of Numerical Methods Engineering* **14**, 1865-1871.
- Johnson, G.R. (1981) Dynamic analysis of explosive-metal interaction in three dimensions, *Transactions of ASME. Journal of Applied Mechanics* **48**, 30-34.
- Gerasimov, A.V. (Ed) (2007) *Theoretical and experimental researches of high - velocity body interactions*, Tomsk State University Press, Tomsk [in Russian].

# Development of a Discontinuous Galerkin Method for Supersonic Flow Simulations on Hybrid Mesh

\* SU Peng-hui, ZHANG Liang

China Academy of Aerospace Aerodynamics, Beijing, 100074, China

\*Presenting author: drsubest@163.com

## Abstract

A two dimensional discontinuous Galerkin finite element method for supersonic flow field simulation on hybrid meshes is proposed. The governing equations are Euler equations, and the 3rd order explicit Runge–Kutta method is used for temporal discretization. The Hermit WENO limiter is introduced to increase the stability of this scheme when it is applied to supersonic flow fields. Two dimensional hybrid unstructured meshes are used in spatial discretizations, which contain both triangle elements and quadrangle elements. This method is validated with supersonic test problems, the results show that this method can solve supersonic flow fields using hybrid unstructured meshes.

**Keywords:** Supersonic, Euler Equations, Discontinuous Galerkin, Hybrid Mesh

## Introduction

Supersonic flow field simulation plays an important role in flight aerodynamic predictions and space craft designs. There are many computational fluid dynamic methods for solving supersonic flow problems, these methods can be put into three categories: finite difference method (FDM), finite volume method (FVM) and finite element method (FEM). The FDM is suitable for constructing high order numerical schemes, and widely used in academic researching, but this method is mainly developed under Cartesian grids, it is difficult to extend this method to unstructured or hybrid meshes, which are common when dealing with real world complex geometries. The FVM, on the other hand, has no limitation on mesh types or geometry complexities, but it is difficult for FVM to achieve a scheme higher than second order on an unstructured mesh, mainly due to the difficulties to implement a compact reconstruction stencil for high order FVM.

The discontinuous Galerkin method (DGM) is a special kind of FEM, this method was first proposed by Reed and Hill [Reed and Hill (1973)], in the 1990s, Cockburn and Shu proposed the Runge-Kutta Discontinuous Galerkin method [Cockburn and Shu (1998)]. After that, the DGM is widely used in many areas, such as aerodynamics, hydrodynamics, wave propagations and computational acoustics. The DGM has both the advantages of FDM and FVM, it is suitable for constructing high order numerical schemes by using high order basis functions, the computational mesh and element shape has no limitations. The stencil in DGM is compact with any order of basis functions, that means to get the solutions of unknown variables in one element, only the unknown variables in its neighbor elements are needed. All these characteristics of the DGM make it a promising method for solving real world engineering flow problems.

The present authors have developed a two dimensional discontinuous Galerkin method for compressible Euler equations on unstructured and hybrid meshes. In order to suppress the non-physical oscillation, a Hermit WENO limiter [Hong et al. (2007)] is introduced. The numerical tests show that this scheme provides an attractive way for solving supersonic flow problems with complex geometries.

### Governing equations

The governing equations are two dimensional inviscid Euler equations, which can be expressed in the form as:

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} = 0 \quad (1)$$

The variable  $U$  is conservative state vector,  $F$  and  $G$  are the inviscid flux vectors:

$$U = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{bmatrix}; F = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ (\rho E + p)u \end{bmatrix}; G = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ (\rho E + p)v \end{bmatrix} \quad (2)$$

The equation of state for perfect gas is used:

$$p = \rho RT \quad (3)$$

These form the complete set of equations, ready to be solved with proper numerical methods.

### Discontinuous Galerkin method

#### *Spatial discretization*

Assuming that the computational domain is divided into a set of non-overlapping elements  $K_j$ , the governing equations are solved in a weak form, we introduce test functions  $v$ , multiply test functions with the governing equations, then integrate over element  $K_j$ , we have:

$$\int_{K_j} \frac{\partial U}{\partial t} v dV + \int_{K_j} \left( \frac{\partial F}{\partial x} + \frac{\partial G}{\partial y} \right) v dV = 0 \quad (4)$$

After integrating by parts, we obtain the final form:

$$\int_{K_j} \frac{\partial U}{\partial t} v dV + \oint_{\partial K_j} v \hat{f} \cdot n d\Omega - \int_{K_j} \nabla v \cdot f dV = 0 \quad (5)$$

Where the  $\hat{f}$  is numerical flux between the two adjacent elements, any numerical flux calculation schemes can be used, here the Van Leer scheme [Toro (2009)] is chosen to calculate  $\hat{f}$ .

The approximate solution is defined in each element as a polynomial:

$$u_h = \sum_{i=1}^N u_i \phi_i(x, t) \quad (6)$$

Where  $\phi_i(x,t)$  is the basis function,  $u_i$  is the solution coefficient, the order of discontinuous Galerkin scheme is defined as the maximum order of the basis functions.

#### *Time discretization*

Replace the solution vectors and test functions in Eq.(5) by their approximation polynomials, a system of ordinary differential equations is obtained:

$$M \frac{dU}{dt} = R(U) \quad (7)$$

$R(U)$  is the residual vector, this ODE system can be solved step forward in time using explicit Runge-Kutta scheme, a third order scheme is used:

$$\begin{aligned} U^{(1)} &= U^n + \Delta t M^{-1} R(U^n), \\ U^{(2)} &= \frac{3}{4} U^n + \frac{1}{4} [U^{(1)} + \Delta t M^{-1} R(U^{(1)})] \\ U^{(3)} &= \frac{1}{3} U^n + \frac{2}{3} [U^{(2)} + \Delta t M^{-1} R(U^{(2)})] \\ U^{n+1} &= U^{(3)} \end{aligned} \quad (8)$$

Where  $U^n = U(t)$  and  $U^{n+1} = U(t + \Delta t)$ .

#### *Slope limiter*

When there are strong discontinuities in the flow field, the discontinuous Galerkin solving procedure may fail due to the severe oscillations near strong discontinuity regions, these oscillations will cause non-physical solutions such as negative pressure or negative density. When shock waves exist in flow fields, certain amount of numerical dissipation is crucial for the successful solving. The DGM with piecewise constant basis functions could offer enough dissipations by itself, but when the order of basis functions is equal or greater than unity, some additional dissipations is often needed, limiters are most commonly used tools to do this.

There are many kinds of limiters proposed by researchers, the Hermit WENO limiter proposed by Luo and Shu is becoming popular among them. The Hermit WENO limiter is based on the idea of Hermit polynomial reconstruction and WENO reconstruction. The major advantage of this limiter is the compactness of its stencils, this makes it suitable for hybrid unstructured mesh, detailed implementations of Hermit WENO limiter on unstructured mesh can be found in [Hong et al. (2007; 2010)].

### **Numerical tests**

The following numerical tests are obtained by the discontinuous Galerkin method using piecewise linear basis functions in two dimensional space.

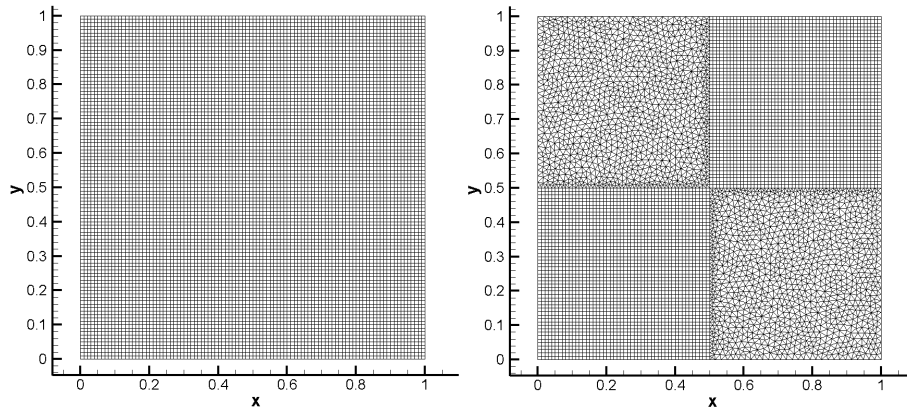
#### *Riemann problem*

The two dimensional Riemann problem is designed to test the performance of numerical methods when there are shock waves and contact discontinuities in supersonic flow fields. The computational domain is  $[0,1] \times [0,1]$ , the domain is

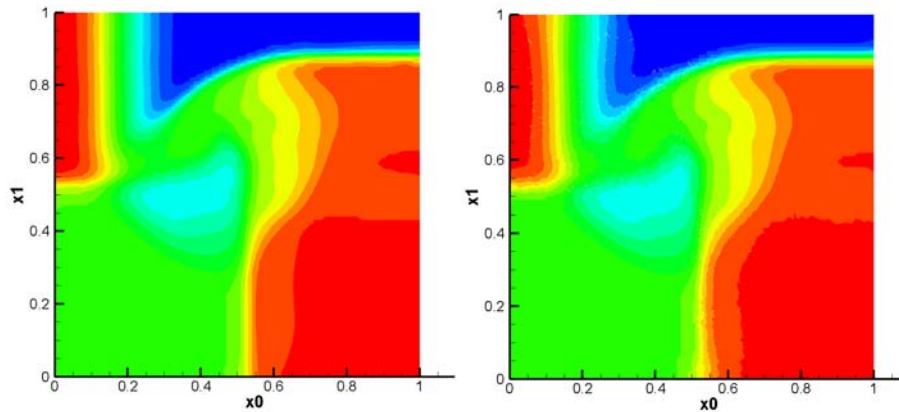
divided into four parts, they are (a):  $[0,0.5] \times [0,0.5]$ , (b):  $[0.5,1] \times [0,0.5]$ , (c):  $[0,0.5] \times [0.5,1]$  and (d):  $[0.5,1] \times [0.5,1]$ , each part has its own initial condition, the initial conditions are:

$$(a): \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 0.8 \\ 0.1 \\ 0.1 \\ 1 \end{pmatrix}; (b): \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 1 \\ 0.1 \\ 0.8276 \\ 1 \end{pmatrix}, (c): \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 1.0222 \\ -0.6179 \\ 0.1 \\ 1 \end{pmatrix}, (d): \begin{pmatrix} \rho \\ u \\ v \\ p \end{pmatrix} = \begin{pmatrix} 0.5313 \\ 0.1 \\ 0.1 \\ 0.4 \end{pmatrix}$$

Two kinds of Meshes are used in this numerical test case, Fig.1 shows the elements distributions of these meshes, the first mesh is a Cartesian grid which contains only uniform distributed quadrangle elements, the second mesh is a hybrid mesh with two blocks of Cartesian grids and two blocks of unstructured grids, this mesh contains both quadrangle and triangle elements. The solutions of the flow field are calculated until the time  $t=0.23$ .



**Figure 1. Uniform cartesian grid (left) and hybrid mesh (right) used in the numerical simulation**



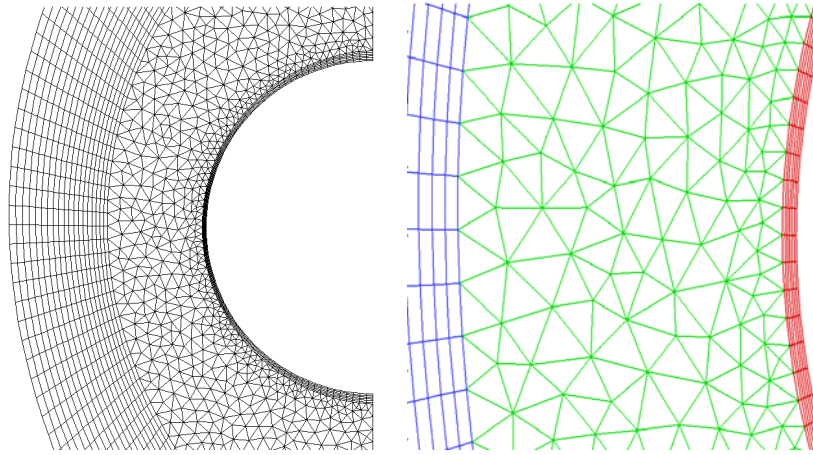
**Figure 2. Comparison of density distributions at  $t=0.23$  obtained by cartesian grid (left) and hybrid mesh (right)**

Fig. 2 shows the flow field density distribution when  $t=0.23$ , the Cartesian grid and hybrid mesh give similar density profiles, which means that the discontinuous Galerkin method has the capable of solving Euler equations on unstructured hybrid meshes, this characteristic of DGM gives a lot flexibilities in modeling complex flow geometries.

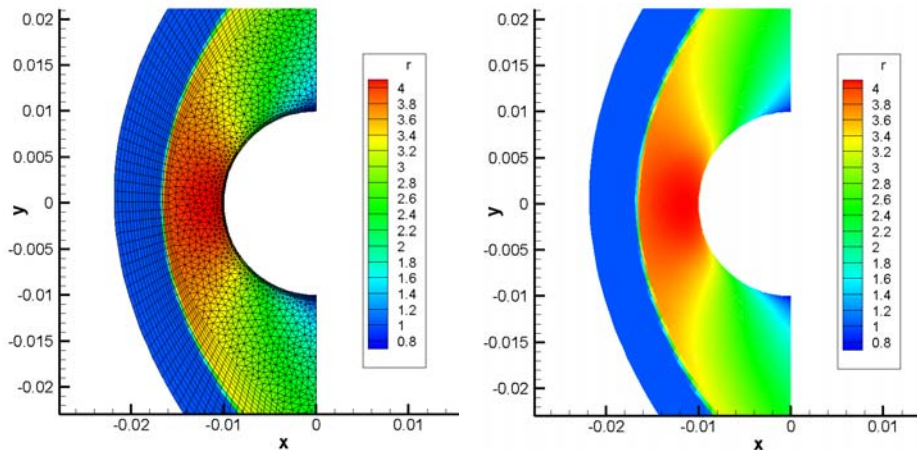
*Supersonic cylinder*

Supersonic cylinder flow is a common test case in supersonic flow field simulations, the flow field contains a strong bow shock wave in front of the cylinder, the radius of

cylinder is chosen as 0.01, inflow Mach number  $M=3$ , and the non-dimensional inflow parameters are chosen as :  $\rho=1$ ,  $u=1$ ,  $v=0$ ,  $p=1/(\gamma M^2)$ , Fig.3 shows the computational mesh, which contains both quadrangle elements and triangle elements. There is a quadrangle element block near the cylinder surface, and a quadrangle element block designed to capture the bow shock wave, between these two quadrangle element blocks are triangle elements.



**Figure 3. Computational mesh (left) and its local details (right)**



**Figure 4. Solution of density distribution compared with mesh (left) and the density contour (right)**

Fig. 4 shows the solution of density distribution using this hybrid mesh, the shock wave is distributed in the quadrangle element region and has a sharp resolution with these quadrangle elements, The density profile in the flow field between shock wave and cylinder gets a smooth distribution with triangle elements. The result gives a demonstration of the advantage to use discontinuous Galerkin method combined with hybrid meshes in supersonic flow simulations, the parallel distributed quadrangle elements are suitable for capturing shock waves, mean while for smooth flow regions, the usage of triangle elements can offer more geometrical flexibilities.

## Conclusions

A two dimensional discontinuous Galerkin finite element method for supersonic flow field simulations is proposed. The governing equations are Euler equations, and the computational mesh is unstructured hybrid mesh. Numerical tests show that the discontinuous Galerkin method provides an effective way of solving engineering supersonic problems on hybrid meshes.

**References**

- 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.
- Cockburn, B. and Shu, C. W. (1998) The Runge–Kutta discontinuous Galerkin method for conservation laws V: Multidimensional system, *Journal of Computational Physics* **141**, 199–224.
- Hong, L., Joseph, D. B. and Rainald, L. (2007) A Hermite WENO-based limiter for discontinuous Galerkin method on unstructured grids, *Journal of Computational Physics* **225**, 686–713.
- Toro, E. F. (2009) *Riemann Solvers and Numerical Methods for Fluid Dynamics –A Practical Introduction*, Springer, New York
- Hong, L., Luqing, L., Robert, N., Vincent, A. M. and Nam, D. (2010) A reconstructed discontinuous Galerkin method for the compressible Navier–Stokes equations on arbitrary grids, *Journal of Computational Physics* **229**, 6961–6978

# Mode decomposition method for non-classically damped structures using acceleration responses

J.-S. Hwang<sup>1</sup>, \*S.-H. Shin<sup>2</sup>, and †H. Kim<sup>2</sup>

<sup>1</sup>Department of Architectural Engineering, Chonnam National University, Gwangju, Korea.

<sup>2</sup>Department of Architectural Engineering, Kyungpook National University, Daegu, Korea

\*Presenting author: ssh10004ok@knu.ac.kr

†Corresponding author: hjk@knu.ac.kr

## Abstract

For structures with non-classical damping or closely distributed modes, it is not easy to apply the traditional modal analysis method because the damping matrix is not diagonalized by the modal matrix obtained from the mass and stiffness matrices. In this paper, a new mode decomposition method for structures with non-classical damping ratio and structures with very closely distributed modes is proposed. This method defines the generalized modes in state space, and uses the differential state variables estimated from measured acceleration responses to decompose modal responses. A Kalman filtering is utilized to calculate the linear transformation matrix of governing modes, and the linear transformation matrix is updated in the optimization process to maximize the performance index cooperated with a power spectral density of a target mode. For the verification of the proposed method, a numerical simulation is performed using a single degree of freedom (SDOF) system coupled with a tuned mass damper (TMD) which represents a non-classically damped system with closely distributed modes. The results from the simulations show that the proposed method estimates the modal responses more precisely than conventional mode decomposition methods such as the independent component analysis (ICA) method and the proper orthogonal decomposition (POD) method.

**Keywords:** Mode decomposition, Non-classical damping, Closely distributed mode, Linear transformation matrix, Differential state variable, Averaged power spectrum

## Introduction

The response of a linear multi-degree-of-freedom (MDOF) structure is often estimated using a few governing mode responses after transforming the system into single-degree-of-freedom (SDOF) systems in the modal space. The transformation into the modal space in the modal analysis requires the modal matrix that is composed of mode shapes, and thereby it is necessary to obtain the mode shapes primarily.

The mode shapes or modal matrix is generally obtained from the eigenvalue analysis using mass and stiffness matrices of the finite element analysis model. The mass and stiffness matrices of the actual structure, however, differ from those of the analysis model yielding the discrepancy in the dynamic characteristics and mode shapes. Further, it is not possible to separate modes using the mode shapes obtained from the mass and stiffness matrices if a structure has non-classical damping that is not proportional to mass and stiffness matrices or the structure has closely distributed modes.

In order to estimate the actual mode shapes for accurate modal separation, the mode decomposition method using measured structural responses has been studied. In special, the mode separation methods using the measured responses from appropriately numbered sensors have been developed because the behavior of a building structure subjected to wind load is governed by a few lower mode responses.

The proper orthogonal decomposition (POD) method is one of the mode decomposition methods using the linear transformation of measured responses [Feeny (2002); Han and Feeny (2003)]. The POD method, also called as the principal component analysis (PCA) method, Karhunen-Loeve method, or the singular value decomposition (SVD) method [Gramaand and Subramanian (2014); Khalil and Sarkar (2014)], transforms the higher order model into the lower order model with



orthogonal basis minimizing the loss of higher order model information. The independent component analysis (ICA) method is another mode decomposition method using the linear transformation of measured responses based on the assumption that modes are independent each other [Roberts and Everson (2001); Kerschen and Poncelet (2007)]. It is also possible to perform the mode decomposition using the output-only method such as the stochastic subspace identification (SSI) and frequency domain decomposition (FDD) methods, which estimates the modal characteristics of a structure using the measured responses [Van Overschee and De Moor (1996); Brincker et al. (2001); Ku et al. (2007)].

These mode decomposition methods are applicable to structures with classical damping, which is proportional to mass and stiffness matrices. Further, they yield reliable results when buildings have very small damping ratio and thereby possess the characteristics of structures with classical damping. The ICA method, which is mostly close to the method proposed in this paper, assumes that modes are separated enough to be independent each other. However, the damping matrix is not proportional to mass and stiffness matrices of actual structures and it is not appropriate to assume that the closely distributed modes are independent each other. Therefore, there exists a limit when the previous mode decomposition methods are applied to structures with non-classical damping or with closely distributed modes.

In this paper, the new mode decomposition method using only measured responses is presented for structures with non-classical damping or with closely distributed modes. This mode decomposition method applies the linear transformation to measured response for calculating the modal responses similarly to the ICA method. The linear transformation matrix differs from that of the ICA method such that it is obtained by optimizing the objective function. The objective function is given to maximize the energy at the certain mode and to minimize the difference between averaged modal response spectrum and the linear transformation matrix assuming that each mode possesses unique pole with one natural frequency and one damping ratio.

For the verification of the proposed mode decomposition method, the numerical simulation of a two DOF system with a tuned mass damper (TMD) that is a representative system with non-classical damping and very closely distributed modes. It is assumed in the numerical simulation that the external load has wide spectral range like wind loads and the only responses of the main structure and TMD are measurable. The mode shapes and modal responses obtained from the measured responses are compared to the analytical ones to verify the proposed mode decomposition method.

## Mode decomposition in state-space domain

### *Estimation of unmeasured state variables*

The mode decomposition of a structure with non-classical damping is not possible because the damping matrix cannot be diagonalized using the mass and stiffness matrix. This requires having linear combination of state-space variables to construct modal responses.

The mode shapes of an MDOF system whose governing equation is given in Eq. (1) are defined as the linear combination as in Eq. (2). The mode separation is possible only when the damping matrix,  $C$ , is diagonalized by the mode shape matrix,  $\Phi$ , in Eq. (3), which is obtained from the eigenvalue analysis of mass matrix,  $M$ , and stiffness matrix,  $K$ .

$$M\ddot{x} + C\dot{x} + Kx = Ef \quad (1)$$

$$x = \Phi\eta \quad (2)$$

$$\ddot{\eta} + \Phi^T C \Phi \dot{\eta} + \Omega\eta = \Phi^T Ef \quad (3)$$

where  $f$  is the external force,  $E$  is the force location matrix,  $\Omega$  is the diagonal matrix with entries of squared natural frequencies, and  $x$  and  $\eta$  are the response vectors in time domain and modal space, respectively.

If the structure has non-classical damping, the term  $\Phi^T C \Phi \dot{\eta}$  in Eq. (3) is not a diagonal matrix, and thereby the mode decomposition is not attainable. Consequently, it is required to expand the modal responses into the state-space domain for the mode decomposition. Eq. (1) is transformed into Eq.

(4) in state-space domain, and the state variable,  $z$ , can be transformed into modal space using the newly defined modal responses in state-space domain,  $p$ , as in Eqs. (5) and (6).

$$\dot{z} = Az + Bf \quad (4)$$

$$z = \Psi p \quad (5)$$

$$\dot{p} = \Psi^{-1} A \Psi p + \Psi^{-1} B f \quad (6)$$

where

$$A = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} \quad (7)$$

$$B = \begin{bmatrix} 0 \\ -M^{-1}E \end{bmatrix} \quad (8)$$

and the mode shape in state-space domain,  $\Psi$ , satisfies

$$\hat{A} = \Psi^{-1} A \Psi = \begin{bmatrix} 0 & I \\ -\Omega & -Z \end{bmatrix} \quad (9)$$

where  $Z = \text{diag}(2\xi_i\omega_i)$ ,  $\text{diag}()$  is the diagonalization function, and  $\xi_i$  and  $\omega_i$  are the damping ratio and natural frequency of the  $i$ -th mode, respectively.

All of state variable  $z$  of displacement and velocity or differential variable  $\dot{z}$  of velocity and acceleration are necessary in order to obtain the modal responses in state-space domain of Eq. (5). However, it is not practical to measure every state and it is often to measure acceleration responses in practice. Therefore, it is assumed in this paper that the number of sensors is equal to the number of governing modes and velocity and displacement responses are obtainable from the measured acceleration using the Kalman filter.

Given that the order of Kalman filter is twice the number of sensors, the initial estimate of the system matrix  $A$  is given as

$$A_0 = S_{21} S_{dv}^{-1} \quad (10)$$

where  $A_0$  is the initial estimate of the system matrix  $A$ ,  $S_{21} = E \begin{bmatrix} \dot{z} & z^T \end{bmatrix}$  and  $S_{dv} = E \begin{bmatrix} z & z^T \end{bmatrix}$ . Multiplying Eq. (4) by  $z^T$  and averaging yields Eq. (10). The external force term is ignored since it is not known or measurable.

Since the velocity and displacement are required in Eq. (10), the following simple integrating filter is introduced.

$$\begin{pmatrix} \dot{q} \\ \ddot{q} \end{pmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} q \\ \dot{q} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ddot{x}_m \quad (11)$$

where  $\ddot{x}_m$  is the measured acceleration, and  $q$  and  $\dot{q}$  are displacement and velocity integrated from the measured acceleration, respectively.

If the measured acceleration in Eq. (11) is biased, the integrated displacement and velocity have considerable amount of errors and often diverge during integration. In order to avoid the divergence and to minimize the errors, the control variable,  $u$ , is added to Eq. (11) as in Eq. (12) where the control gain,  $G$ , of size  $1 \times 2$  is decided to minimize the squared displacement and squared control variable in Eq. (13).

$$\begin{pmatrix} \dot{q} \\ \ddot{q} \end{pmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} q \\ \dot{q} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \ddot{x}_m + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \quad (12.a)$$

$$u = -G \begin{pmatrix} q \\ \dot{q} \end{pmatrix} \quad (12.b)$$

$$J = \int_0^T \frac{1}{2} (Q_1 q^2 + R u^2) dt \quad (13)$$

where  $Q_1$  and  $R$  are weights. Note that the control variable  $u$  is equal to the difference between the actual measured acceleration and the estimated one.

The control gain,  $G$ , can also be obtained by modifying weights  $Q_1$  and  $R$  such that the difference between the actual measured acceleration and the estimated one is in a certain range. The more detailed process for control gain calculation including the Kalman filter method is omitted here since it has been widely introduced in many references [Hwang et al. (2011)].

#### *Objective function for mode decomposition*

Once the state variables are estimated from the measured acceleration responses using the Kalman filter, the relationship between the state variables and the modal responses can be defined using Eq. (5). Because the purpose is the mode decomposition using the measured responses, Eq. (5) is rewritten as

$$p = W^T z \quad (14)$$

where  $W^T = \Psi^{-1}$ . Since it is assumed that the number of the sensors,  $n$ , is equal to the number of governing modes, the transformation matrix,  $W$ , is a square matrix of  $2n \times 2n$  and its inverse matrix exists.

It can be noted that each column of the transformation matrix  $W$  is the combination of linear transformation coefficients that separate certain modes from the measured state variables. Since the measured acceleration and its integral value, velocity, are mostly used, Eq. (14) can be rewritten using differentiated state variables as

$$\dot{p} = W^T \dot{z} \quad (15)$$

where the entries of the transformation matrix  $W$  are constant and are not affected by differentiation.

From Eq. (15), it can be noted that the number of differentiated state variable,  $\dot{z}$ , is  $2n$  and the number of corresponding generalized modes is also  $2n$ . The first  $n$  modes obtained from Eq. (15) have relationship with the rest of modes defined as in Eq. (6). If the effect of external force is negligible in Eq. (6), the relationship becomes velocity to acceleration. That is, the relationship between  $i$ -th mode and  $(i+n)$ -th mode is velocity to acceleration, if  $i \leq n$  and the effect of external force is negligible.

The  $i$ -th mode can be presented using the  $i$ -th column of  $W$  of Eq. (15) as

$$\dot{p}_i = (W_i)^T \dot{z} \quad (16)$$

In order for the  $i$ -th mode obtained from Eq. (16) to be decomposed into a true vibration mode with single pole that consists of natural frequency,  $\omega_i$ , and damping ratio,  $\xi_i$ , the modal power spectrum obtained from Eq. (16) needs to have only one peak near the natural frequency when there is no special poles in the external force. That is, the effect of other modes should not be appeared showing no peaks near other modes.

In this paper, the following necessary conditions are defined for true mode decomposition described above. These conditions also are the preconditions to define the objective function for mode decomposition. Note that these conditions are not necessary and sufficient conditions for mode decomposition and that other necessary conditions based on other idea can also be adopted.

Necessary condition 1: The total energy of decomposed modes is always constant. This condition is satisfied by setting the integral value of modal response spectrum, which is equal to the variance value, to be '1'.

Necessary condition 2: The modal energy is max near its natural frequency. The corresponding natural frequency can be obtained from the system matrix  $\mathbf{A}_0$  in Eq. (10).

Necessary condition 3: The effect by neighboring modes is minimized. This condition can be satisfied by minimizing the differences between the modal power spectrum and averaged power spectrum at neighboring modal frequencies.

The objective function satisfying the above necessary conditions 1 and 2 can be defined as

$$J_{12} = \int_{\omega_i - \Delta\omega}^{\omega_i + \Delta\omega} S_{ii}(\omega) d\omega + \lambda \left( \int_0^\infty S_{ii}(\omega) d\omega - 1 \right) \quad (17)$$

where  $\lambda$  is a Lagrange multiplier for constraining the necessary condition 1,  $\Delta\omega$  is the infinitesimal change of frequency, and  $S_{ii}(\omega)$  is the power spectrum of the decomposed mode.  $S_{ii}(\omega)$  is one-sided spectrum given as

$$S_{ii}(\omega) = W_i^T S_{va}(\omega) W_i \quad (18.a)$$

$$S_{va}(\omega) = z(\omega) \bar{z}(\omega) \quad (18.b)$$

where  $z(\omega)$  is the Fourier transformation of differential state variable,  $\dot{z}(t)$  and  $\bar{z}(\omega)$  is the complex conjugate of  $z(\omega)$ . Substituting Eq. (18) into Eq. (17) simplifies the objective function of Eq. (17) as

$$J_{12} = W_i^T S_{peak} W_i + \lambda (W_i^T S_{var} W_i - 1) \quad (19.a)$$

$$S_{peak} = \int_{\omega_i - \Delta\omega}^{\omega_i + \Delta\omega} S_{va}(\omega) d\omega \quad (19.b)$$

$$S_{var} = \int_0^\infty S_{va}(\omega) d\omega \quad (19.c)$$

$S_{peak}$  and  $S_{var}$  are readily obtainable from the differentiated state variables directly. Consequently, the transformation matrix,  $W_i$ , for the  $i$ -th mode that satisfies the condition 1 and 2 can be derived by differentiating  $J_{12}$  of Eq. (19.a) with respect to  $W_i$  and setting the resulting value to be '0'. The result of differentiation is given as

$$(S_{peak} + \lambda S_{var}) W_i = 0 \quad (20.a)$$

$$S_{peak} W_i = -\lambda S_{var} W_i \quad (20.b)$$

It can be noted from Eq. (20) that the value of  $(-\lambda)$  is the eigenvalue of two matrices,  $S_{peak}$  and  $S_{var}$ , while  $W_i$  is the corresponding eigenvector. This means that the largest eigenvalue becomes the maximum value of objective function and the corresponding eigenvector  $W_i$  becomes the linear transformation matrix.

If modes are separated enough to affect each other marginally, it is possible to perform the mode decomposition accurately using the transformation matrix obtained from Eq. (20). When modes are closely distributed, the reciprocal effect between modes becomes significant. In that case, the objective function that satisfies the necessary conditions 1 and 2 only cannot yield the accurate mode decomposition. In order to minimize the effect of neighboring modes, the following objective function that satisfies the necessary conditions 3 as well as 1 and 2 is defined.

$$J_{123} = \frac{\int_{\omega_i - \Delta\omega}^{\omega_i + \Delta\omega} S_{ii}(\omega) d\omega}{\int_{\omega_k - \Delta\omega}^{\omega_k + \Delta\omega} \left( \log \left( \frac{S_{ii}(\omega)}{S_H(\omega)} \right) \right)^2 d\omega} + \lambda \left( \int_0^\infty S_{ii}(\omega) d\omega - 1 \right) \quad (21)$$

where  $S_H$  is the averaged power spectrum given as

$$S_H(\omega) = S_o |H(s)|^2 \quad (22.a)$$

$$H(s) = \frac{s}{s^2 + 2\xi_i \omega_i + \omega_i^2} \quad (22.b)$$

$$S_o = \frac{\int_{\omega_i - \Delta\omega}^{\omega_i + \Delta\omega} S_{ii}(\omega) d\omega}{\int_{\omega_i - \Delta\omega}^{\omega_i + \Delta\omega} |H(s)|^2 d\omega} \quad (22.c)$$

where  $s$  is the Laplace variable,  $H(s)$  is the transfer function of velocity response from the external force of a SDOF system, and  $S_o$  is the constant that represents the ratio of the  $i$ -th modal power spectrum to the transfer function near the  $i$ -th mode frequency.

The difference between Eqs. (17) and (21) is that the logarithmic ratio of the  $i$ -th modal spectrum to the averaged spectrum near the frequency of neighboring mode,  $\omega_k$ , is included in the denominator. Minimizing the ratio in the denominator maximizes the objective function, while the logarithmic ratio accentuates the difference between two spectrums. The objective function in Eq. (21) can be simplified using  $W_i$  as

$$J_{123} = \frac{W_i^T S_{peak} W_i}{\int_{\omega_k - \Delta\omega}^{\omega_k + \Delta\omega} (\log(W_i^T q_{ii}(\omega) W_i))^2 d\omega} + \lambda (W_i^T S_{var} W_i - 1) \quad (23)$$

where

$$q_{ii}(\omega) = \frac{S_{ii}(\omega)}{S_H(\omega)} \quad (24)$$

The natural frequency,  $\omega_i$ , in Eq. (22.b) can be directly obtained from the system matrix of Eq. (10) while the damping ratio,  $\xi_i$  obtained from the system matrix has large error. Therefore, the damping ratio needs to be selected such that  $S_o$  of Eq. (22.c) satisfies the following relationship derived from the necessary condition 1.

$$\int_0^\infty S_o(\xi_i) |H(s)|^2 d\omega - 1 = \int_0^\infty S_{ii}(\omega) d\omega - 1 = 0 \quad (25)$$

Once the values of all variables in Eq. (23) are calculated, the transformation matrix,  $W_i$ , can be obtained by differentiating the objective function with respect to  $W_i$ . However, the closed-form similar to one in Eq. (20) cannot be derived due to the nonlinearity. Therefore, the sensitivity of objective function is utilized in the optimization process to obtain the transformation matrix,  $W_i$ .

## Validation of the proposed method

### *Example structure and its modal characteristics*

A numerical simulation using an example structure with non-classical damping and very closely distributed modes is carried out to verify the proposed mode decomposition method. The example structure is a two-DOF system with a TMD which is a representative system with non-classical damping and very closely distributed modes. The dynamic characteristics of the structure and external load are summarized in Table 1. It is assumed that the low-pass filtered white noise is applied to the main structure only.

In Tables 2 and 3, the mass and stiffness matrices along with the corresponding mode shapes are presented in time and state-space domains, respectively. These mode shapes will be compared to ones obtained using the proposed mode decomposition method. It can be noted that the damping matrix is not diagonalized by the mode shape obtained from the eigenvalue analysis of mass and stiffness matrices in time domain from Table 2, while modes are apparently separated in state-space domain from Table 3.

**Table 1. Dynamic characteristics of the example structure**

Description		Value	Remark
Structure	Main structure mass ( $m_s$ )	100 kg	Natural frequency of the main structure ( $f_0$ ) = 0.2 Hz
	TMD mass ( $m_t$ )	1 kg	Mass ratio of TMD mass to main structure mass = 0.01
	Main structure stiffness ( $K_s$ )	157.9137 N/m	
	TMD stiffness ( $K_t$ )	1.5635 N/m	
	Main structure damping	2.5133 N·s/m	Main structure damping ratio ( $\xi_s$ ) = 0.01
	TMD damping	0.1250 N·s/m	TMD damping ratio ( $\xi_t$ ) = 0.05
External load	Filter	Low pass filter	Zero to 4 Hz
	Sampling time	0.005 s.	Sampling frequency = 200 Hz
	Duration	3600 s.	

**Table 2. Analytically obtained modal properties in time domain**

Matrix	Symbol	Value
Mass matrix	$M$	$\begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix}$
Damping matrix	$C$	$\begin{bmatrix} 2.6383 & -0.1250 \\ -0.1250 & 0.1250 \end{bmatrix}$
Stiffness matrix	$K$	$\begin{bmatrix} 159.4772 & -1.5635 \\ -1.5635 & 1.5635 \end{bmatrix}$
Natural frequencies	$\omega_1$ $\omega_2$	1 <sup>st</sup> mode: 1.1925 (0.1898 Hz) 2 <sup>nd</sup> mode: 1.3177 (0.2097 Hz)
Mode shapes	$\Phi$	$\begin{bmatrix} -0.0671 & -0.0741 \\ -0.7415 & 0.6710 \end{bmatrix}$
Transformed damping matrix	$\Phi^T C \Phi$	$\begin{bmatrix} 0.0682 & -0.0503 \\ -0.0503 & 0.0832 \end{bmatrix}$ Damping ratio of diagonal terms = (0.0286 0.0316)

It can be noted that even though the transformed damping matrix in Table 2 is not a diagonal matrix, the diagonal entries, (0.0682 0.0832), are very close to those of damping part of mode transformed system matrix,  $A$ , in Table 3. It can be also noticed that the natural frequencies in time domain, 0.1898 Hz and 0.2097 Hz, and those in state-space domain, 0.1910 Hz and 0.2095 Hz, are very close each other, while the difference between the first and second modes is only 0.02 Hz indicating the very closely distributed modes.

In Table 3, the modal matrix,  $\Psi$ , in state-space domain is shown in the ascending order of natural frequencies, and its inverse matrix, i.e. the linear transformation matrix,  $W$ , is also provided. Considering that the first two rows of load participation matrix,  $B$ , are zeros, it can be noted that the first and second columns of modal responses in state-space domain are integral values of the third and fourth columns, respectively, indicating the displacement-velocity and velocity-acceleration relationships.

**Table 3. Analytically obtained modal properties in state-space domain**

Matrix	Symbol	Value									
System matrix	$A$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1.5948 & 0.0156 & -0.0264 & 0.0013 \\ 1.5635 & -1.5635 & 0.1250 & -0.1250 \end{bmatrix}$									
Load participation matrices	$B^T$	$[0 \ 0 \ 0.01 \ 0]$									
Modal matrix	$\Psi$	$\begin{bmatrix} -0.6367 & -0.6543 & -0.2392 & 0.2131 \\ -7.4864 & 6.6730 & -0.2046 & 0.1823 \\ 0.3431 & -0.3670 & -0.6204 & -0.6721 \\ 0.2934 & -0.3139 & -7.4725 & 6.6578 \end{bmatrix}$									
Linear transformation matrix	$W$	$\begin{bmatrix} -0.7306 & -0.8195 & 0.0396 & 0.0380 \\ -0.0704 & 0.0707 & -0.0396 & -0.0380 \\ 0 & 0 & -0.7274 & -0.8164 \\ 0.0253 & 0.0243 & -0.0736 & 0.0676 \end{bmatrix}$									
Mode transformed system matrix	$\hat{A}$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1.4340 & 0 & -0.0681 & 0 \\ 0 & -1.7217 & 0 & -0.0833 \end{bmatrix}$									
Eigenvalues	$\omega_i$ and $\xi_i$	<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th>Natural frequency</th> <th>Damping ratio</th> </tr> </thead> <tbody> <tr> <td>1<sup>st</sup> mode</td> <td>1.20 (0.1910 Hz)</td> <td>2.85e-02</td> </tr> <tr> <td>2<sup>nd</sup> mode</td> <td>1.31 (0.2085 Hz)</td> <td>3.17e-02</td> </tr> </tbody> </table>		Natural frequency	Damping ratio	1 <sup>st</sup> mode	1.20 (0.1910 Hz)	2.85e-02	2 <sup>nd</sup> mode	1.31 (0.2085 Hz)	3.17e-02
	Natural frequency	Damping ratio									
1 <sup>st</sup> mode	1.20 (0.1910 Hz)	2.85e-02									
2 <sup>nd</sup> mode	1.31 (0.2085 Hz)	3.17e-02									

#### *Characteristics of mode decomposition depending on the objective function*

A numerical analysis of the coupled main structure-TMD is performed to obtain the acceleration responses. The external load presented in Table 1 is applied in the numerical analysis.

The displacement and velocity responses are obtained using the integral filter given in Eq. (12), and the initial estimate of the system matrix is calculated using Eq. (10). Table 4 presents the covariance matrices used for calculating the initial estimate of the system matrix along with the modal characteristics.

Table 4 indicates that the natural frequencies obtained from the initial estimate of the system matrix present insignificant error compared to the exact natural frequencies given in Table 2. The damping ratios are, however, negative values indicating significant error. The modal matrix and linear transformation matrices also differ from the exact ones while the correlation of sign between matrices is very large.

The linear transformation matrices obtained using the proposed mode decomposition method are compared to the exact one in Table 5. First, the first mode linear transformation matrix that maximizes the objective function  $J_{12}$  in Eq. (19) is obtained using Eq. (20). The vector with norm value of '1' is also presented in Table 5 for easier comparison. It can be seen that the linear transformation matrix obtained from the initial estimate of the system matrix is closer to the exact one than one obtained using the objective function  $J_{12}$ . The values in the first three rows show very close results to exact ones while the value of the last row is about 2.5 times to that of exact one.

**Table 4. Covariance matrices and modal characteristics of initial estimate of system matrix**

Matrix	Symbol	Value									
Covariance matrix of state variable	$S_{dv}$	$\begin{bmatrix} 17.51 & 21.86 & 0.01 & 105.39 \\ 21.86 & 866.02 & -105.30 & 0.18 \\ 0.01 & -105.30 & 27.15 & 20.09 \\ 105.39 & 0.18 & 20.09 & 1333.47 \end{bmatrix}$									
Cross covariance matrix	$S_{21}$	$\begin{bmatrix} 0.01 & -105.30 & 27.15 & 20.09 \\ 105.39 & 0.18 & 20.09 & 1333.47 \\ -27.17 & -20.17 & 0.01 & -165.63 \\ -19.98 & -1333.03 & 165.54 & 0.27 \end{bmatrix}$									
Initial estimate of system matrix	$A_o$	$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1.5883 & 0.0321 & 0.1260 & -0.0006 \\ 1.5635 & -1.5635 & 0.1250 & -0.1250 \end{bmatrix}$									
Modal matrix	$\Psi$	$\begin{bmatrix} -0.8265 & -0.7677 & -0.6655 & 0.5842 \\ -8.2899 & 7.2777 & -0.6621 & 0.5812 \\ 0.9368 & -1.0098 & -0.8268 & -0.7675 \\ 0.9320 & -1.0046 & -8.2903 & 7.2780 \end{bmatrix}$									
Linear transformation matrix	$W$	$\begin{bmatrix} -0.5944 & -0.6757 & 0.0815 & 0.0757 \\ -0.0555 & 0.0728 & -0.0815 & -0.0757 \\ 0 & 0 & -0.5879 & -0.6697 \\ 0.0521 & 0.0484 & -0.0620 & 0.0668 \end{bmatrix}$									
Eigenvalues	$\omega_i$ and $\xi_i$	<table style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th></th> <th>Natural frequency</th> <th>Damping ratio</th> </tr> </thead> <tbody> <tr> <td>1<sup>st</sup> mode</td> <td>1.19 (0.1894 Hz)</td> <td>-2.08e-04</td> </tr> <tr> <td>2<sup>nd</sup> mode</td> <td>1.31 (0.2085 Hz)</td> <td>-1.74e-04</td> </tr> </tbody> </table>		Natural frequency	Damping ratio	1 <sup>st</sup> mode	1.19 (0.1894 Hz)	-2.08e-04	2 <sup>nd</sup> mode	1.31 (0.2085 Hz)	-1.74e-04
	Natural frequency	Damping ratio									
1 <sup>st</sup> mode	1.19 (0.1894 Hz)	-2.08e-04									
2 <sup>nd</sup> mode	1.31 (0.2085 Hz)	-1.74e-04									

**Table 5. Comparison of the first mode linear transformation matrix: values in parenthesis are normalized vectors**

Method	Symbol	Value
Exact solution	$W_1^T$	$\begin{bmatrix} -0.1705 & -0.0164 & 0 & 0.0059 \\ -0.9948 & -0.0959 & 0 & 0.0345 \end{bmatrix}$
Initial estimate of system matrix	$W_1^T$	$\begin{bmatrix} -0.1844 & -0.0172 & 0 & 0.0162 \\ -0.9919 & -0.0926 & 0 & 0.0870 \end{bmatrix}$
Objective function $J_{12}$ in Eq. (19)	$W_1^T$	$\begin{bmatrix} -0.1405 & -0.0212 & 0.0001 & 0.0130 \\ -0.9847 & -0.1485 & 0.0006 & 0.0914 \end{bmatrix}$
Objective function $J_{123}$ in Eq. (23)	$W_1^T$	$\begin{bmatrix} -0.1735 & -0.0157 & -0.0001 & 0.0055 \\ -0.9945 & -0.0900 & -0.0005 & 0.0316 \end{bmatrix}$
ICA	$W_1^T$	$\begin{bmatrix} -0.0295 & -0.0094 & -0.0337 & 0.0010 \\ -0.6442 & -0.2054 & -0.7365 & 0.0212 \end{bmatrix}$



The first mode linear transformation matrix that maximizes the objective function  $J_{123}$  in Eq. (23) is also presented in Table 5. It can be seen that the values of the first three rows are almost identical to exact ones while the value of the last row has error of about 20%.

The optimization process for maximizing the objective function  $J_{123}$  is presented in Fig. 1. The initial values used in the optimization iteration are the values that maximize the objective function  $J_{12}$ . It can be seen from Fig.1 that the value of the objective function increases gradually as the iteration number increases, and it converges to a certain value as the iteration number is about 200. Among the various optimization methods, the simple gradient method is used in this paper. The linear transformation matrix is updated at the  $i$ -th iteration as

$$W_{i+1} = W_i + 0.001\delta W \quad (26)$$

where

$$\delta W = \frac{\partial J_{123}}{\partial W_1} \quad (27)$$

Figure 1 also presents the iteration results of the denominator and numerator of Eq. (23) along with the error between the estimated first mode linear transformation matrix and exact one. The error is calculated as

$$\varepsilon = \sum_{j=1}^4 \left( \left| \frac{(W_1)_j}{\|W_1\|} \right| - \left| \frac{(W_{exact})_j}{\|W_{exact}\|} \right| \right) \quad (28)$$

where  $W_{exact}$  is the exact first mode linear transformation matrix presented in Table 5. It can be seen that the error approaches to zero as the iteration numbers increases.

Figure 2 shows the damping ratio estimation process for the averaged power spectrum of Eq. (22) used in the optimization of the objective function  $J_{123}$ . It can be noticed that the area of the power spectral function becomes almost same to that of the averaged spectrum near the damping ratio of 0.026. Using this damping ratio and the first mode frequency, the objective function  $J_{123}$  is optimized.

The first modal spectrums obtained from the different mode decomposition method are compared to the exact one in Figure 3. It can be seen that the modal spectrums decomposed using the initial estimate of the system matrix and the objective function  $J_{12}$  are distorted considerably near the second mode frequency. In special, the decomposed modal spectrum obtained using the objective function  $J_{12}$  is continuously smaller above the second mode frequency. On the contrary, the decomposed modal spectrum obtained using the objective function  $J_{123}$  matches the exact one closely.

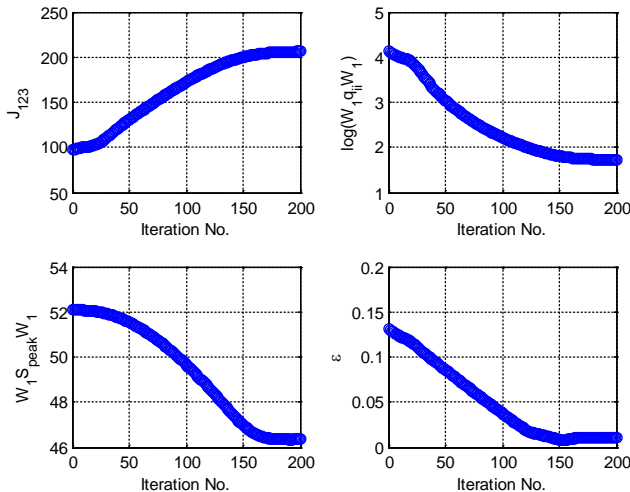


Figure 1. Iteration result

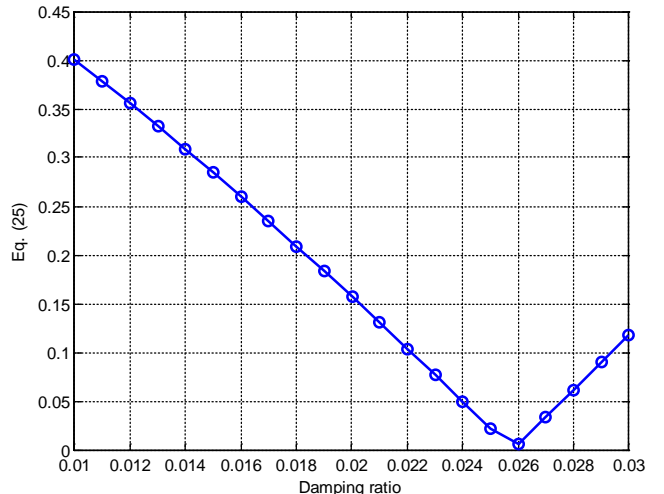
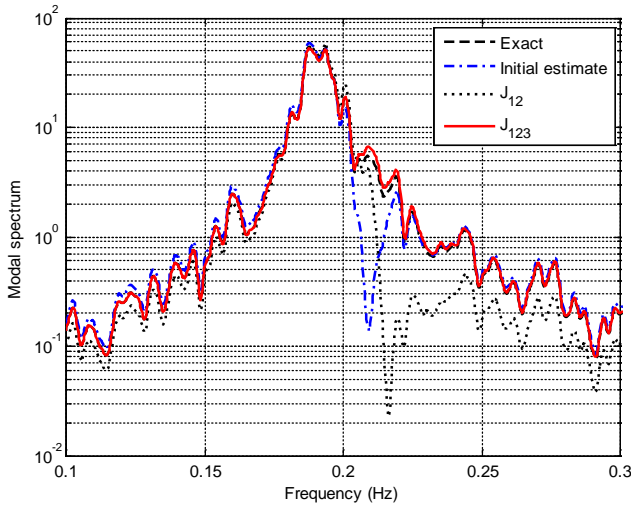
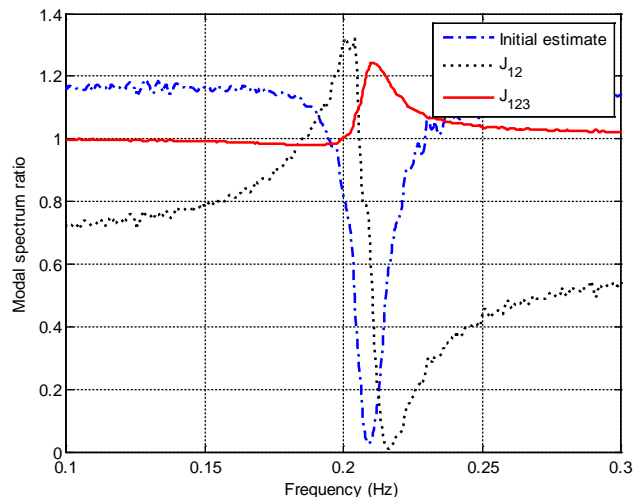
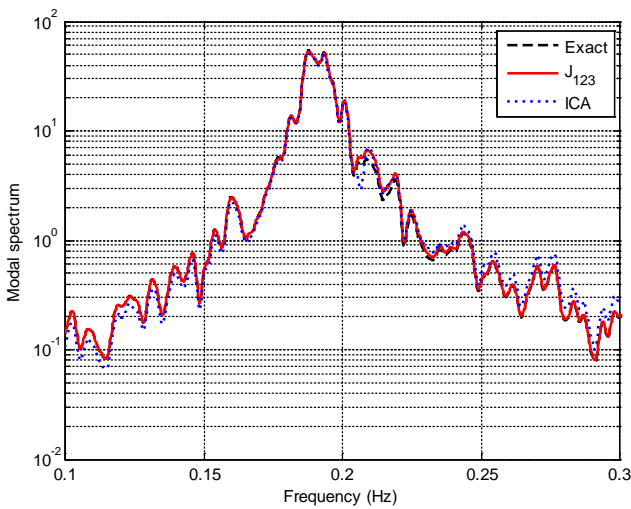
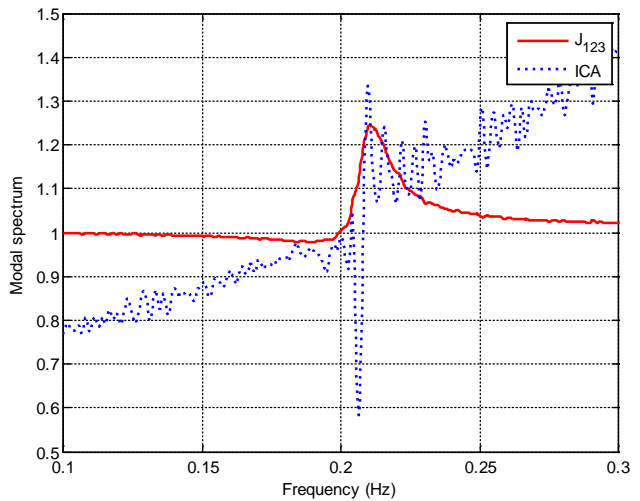


Figure 2. Damping ratio estimation


**Figure 3. Modal spectrum comparison**

**Figure 4. Modal spectrum ratio to exact one**

**Figure 5. Modal spectrum comparison**

**Figure 6. Modal spectrum ratio to exact one**

In order to compare the decomposed modal spectrum more closely, the ratios of decomposed modal spectrums to exact one are presented in Figure 4. It can be noticed the more distinguished error in the decomposed modal spectrum obtained using the initial estimate of the system matrix and the objective function  $J_{12}$ . The decomposed modal spectrum obtained using the objective function  $J_{123}$  shows the ratio near one meaning almost identical result except near the second mode frequency. Therefore, it can be concluded that the objective function defined in this paper yields the decomposed mode with minimum effect from the neighboring modes even when the structure has very closely distributed modes.

For the comparison of the proposed method to the previous mode decomposition methods, the decomposed first modal spectrum using the ICA method is compared in Figure 5. The modal spectrum ratios to exact one are also compared in Figure 6. The corresponding linear transformation matrix for the first mode is presented in Table 5.

It can be noticed that the modal spectrum ratio obtained using the ICA method is close to unity only near the first mode frequency. However, the ratio abruptly decreases near the second mode frequency and increases continuously above that frequency. This is because the ICA method matches the spectral area in average sense trying to maximize the modal independency from the neighboring modes. This feature of the ICA method leads the decrease or increase of the ratio where the modal frequencies do not exist. The other decomposition methods such as the POD and

PCA methods are also examined, but their results are provided here because their decomposition resolutions are far less than the ICA method.

In addition to the first mode decomposition, the second to fourth mode decompositions are also performed and their results are compared to exact ones. The results show that the mode decomposition using the objective function  $J_{123}$  also yields very close modal spectrums to exact ones for higher modes. Since the results are almost identical to that of the first mode, they are not presented here.

## Conclusions

The new mode decomposition method is proposed and validated numerically. The proposed method can improve the decomposition resolution for structures with non-classical damping and closely distributed modes whose mode decomposition is difficult due to non-diagonalization of damping matrix and strong correlation between neighboring modes.

The proposed method defines a generalized mode in state-space domain and performed the mode decomposition using the state variable estimated from the measured responses. The numerical simulation using a SDOF-TMD system indicates that the objective function using the averaged spectrum with single pole yields the best mode decomposition results. Further, it is shown that the proposed method yields the decomposed mode with minimum effect from the neighboring modes even when the structure has very closely distributed modes compared to results to those of the ICA and POD methods.

## References

- Li, Y., Liu, G. R., Luan, M. T., Dai, K. Y., Zhong, Z. H., Li, G. Y. and Han, X. (2007) Contact analysis for solids based on linearly conforming radial point interpolation method, *Computational Mechanics* **39**, 537–554.
- Grama, S. N. and Subramanian, S. J. (2014) Computation of Full-field Strains Using Principal Component Analysis, *Experimental Mechanics* **54**(6), 913-933.
- Khalil, M. and Sarkar, A. (2014) Independent Component Analysis to enhance performances of Karhunen–Loeve expansions for non-Gaussian stochastic processes: Application to uncertain systems, *Journal of Sound and Vibration* **333**(21), 5600–5613.
- Hwang, J.-S., Kareem, A., and Kim, H. (2011) Wind load identification using wind tunnel test data by inverse analysis, *Journal of Wind Engineering and Industrial Aerodynamics* (99), 18-26.
- Feeny, B. (2002) On Proper Orthogonal Coordinates as Indicators of Modal Activity, *Journal of Sound and Vibration* (255), 805-817.
- Han, S. and Feeny, B. (2003) Application of Proper Orthogonal Decomposition to Structural Vibration Analysis, *Mechanical System and Signal Processing* (17), 989-1001.
- Roberts S. and Everson.R. (2001) Independent Component Analysis: Principles and Practice, Cambridge Univ.
- Kerschen G, Poncelet F., and Golinval J.-C. (2007) Physical interpretation of independent component analysis in structural dynamics, *Mechanical Systems and Signal Processing* (21), 1561-1575.
- Van Overschee, P. and De Moor, B. (1996) Subspace identification for Linear Systems, Kluwer Academic Publishers.
- Brincker, R., Zhang, L., and Andersen, P. (2001) Modal identification of output-only systems using frequency domain decomposition, *Smart Materials and Structures* (10), 441-445.
- Ku, C.J., Cermak, J.E., and Chou, L.S. (2007) Random decrement based method for modal parameter identification of a dynamic system using acceleration responses, *Journal of Wind Engineering and Industrial Aerodynamics* (95), 389-410.

# An Effective Three Dimensional MMALE Method for Compressible Fluid

## Dynamics

† Shudao Zhang, Zupeng Jia, Jun Liu, Jun Xiong, and Haibing Zhou

Institute of Applied Physics and Computational Mathematics, Beijing 100094, China.

\*Presenting and corresponding author: zhang\_shudao@iapcm.ac.cn

### Abstract

An effective second-order three-dimensional unstructured multimaterial arbitrary Lagrangian–Eulerian (MMALE) method was presented for compressible fluid dynamics, which uses Moment of Fluid (MOF) method to reconstruct material interface for immiscible fluids. It is of the explicit time-marching Lagrange plus remap type. Comparing with traditional ALE method, MMALE method permits multimaterials in a single cell, thus has the extra advantage of accurately modeling problems involving severe mesh distortion as well as interface fragmentations and coalitions induced by strong shearing deformation. Because the stencil used in the staggered compatible discretization involves only the nearest neighbouring cells and the MOF algorithm does not need information from the neighboring cells, the MMALE method in this paper is suitable for parallel computation while keeps second-order accurate. Several numerical tests on three-dimensional structured and unstructured meshes have proved the accuracy and robustness of the present method.

**Keywords:** MMALE method, MOF method, compatible Lagrangian method, multimaterial fluid dynamics

### Introduction

There are generally two kinds of numerical methods for the computation of fluid Dynamics according to the movement of the mesh during computation. The first one is Lagrangian method, in which the mesh is moving with material velocity. It has the advantage of capturing the material interface precisely and explicitly which is very important for moving boundary problems where material interfaces are of great concern, but with the limitation of severe grid distortion due to strong shear deformation which always stops the calculation. The second one is Eulerian method in which the mesh is fixed avoiding the problem of grid distortions but at the expense of precise material interface construction. In order to overcome the drawbacks of the two methods above, an arbitrary Lagrangian Eulerian (ALE) method was introduced to combine the advantages of both the Lagrangian and Eulerian approaches [Donea et al. (1982)]. In the ALE methodology, the mesh may be moved in some arbitrarily specified way to improve the resolution and enhance the robustness of the simulation. Most ALE algorithms consist of three phases, a Lagrangian phase in which the physical variables and mesh are updated, a rezoning phase for defining a new mesh with better quality, and a remapping phase wherein the physical variables are conservatively interpolated from the old Lagrangian mesh onto the new rezoned one. ALE algorithms have much more flexibilities to deal with multi-material problems such as strong fluid-structure interaction and inertial confinement fusion (ICF) problems.

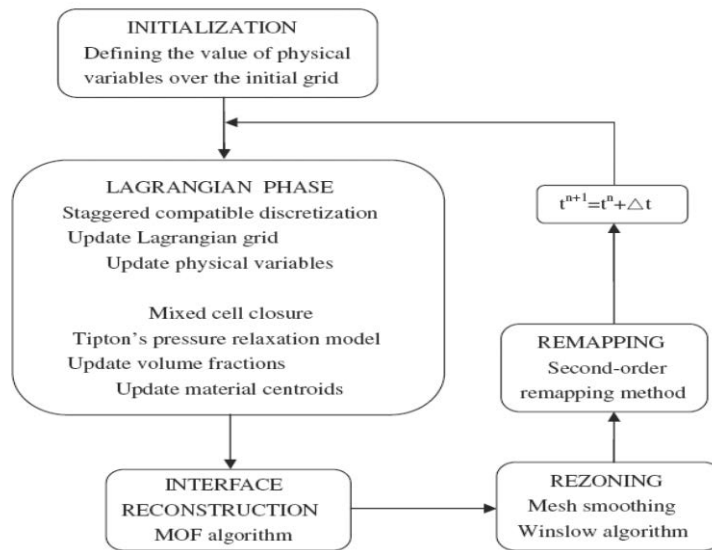
For the traditional ALE method, only one material is allowed to be contained in each mesh cell, so the material interface must be described explicitly by cell edges. When the mesh and the interface deform severely, it is very difficult to generate a new mesh with good quality. In some cases such as interface fragmentations and coalitions emerge due to strong shearing deformation, it is even impossible to perform rezoning successfully and the traditional ALE method often fails. Thus, a new approach called multimaterial ALE method (MMALE) was developed for these problems [Peery et al. (2000)]. The MMALE method is based on a flexible strategy. It allows for multiple materials in a single cell and therefore affords additional flexibility over the traditional ALE method. In other words, MMALE methods permit the interface to cut through cell edges and pass across the cells, and no material-interface-fitted mesh is required, thus the difficulty of mesh adjustment in the rezoning phase is decreased. With these flexibilities, the MMALE method can

accurately model problems involving severe mesh distortion as well as interface fragmentations and coalitions induced by strong shearing deformation.

In this paper an effective second-order three dimensional unstructured grid MMALE method was developed for simulating multi-material compressible fluid flows involving strong shearing deformation [Jia et al. (2013)]. It combined the staggered compatible Lagrangian method and momentum of fluid (MOF) algorithm for interface reconstruction, which has the advantages of second-order accuracy and compact stencils. Numerical results of several test problems including Rayleigh-Taylor instability have shown the accuracy and robustness of the method.

## Numerical Method

The flowchart of our MMALE method is displayed in Fig. 1. In the initialization stage, the distribution of all the physical variables over the initial mesh is defined. The volume fractions and positions of material centroids are initialized using the method which is an extension of [Aulisa et al. (2007)].



**Figure 1. Flowchart of the multi-material ALE algorithm**

During the Lagrangian phase, the gas dynamics equations of the Lagrangian form are solved using the staggered compatible discretization, the velocity, density, internal energy, pressure and the Lagrangian mesh are updated. For mixed cells, Tipton's pressure relaxation model is used to define a thermodynamic closure. The material centroids are updated using a method which can be seen as an extension from 2D to 3D of the method presented in [Kucharik et al. (2010)]. The interface reconstruction is performed using MOF algorithm [Ahn et al. (2007)]. In the rezoning phase, the quality of the Lagrangian mesh is improved by means of mesh smoothing using Winslow's algorithm [Winslow (1997)]. Finally, in the remapping phase, all the physical variables are conservatively interpolated from the Lagrangian mesh onto the new rezoned mesh using a cell intersection based second-order remapping method.

In the Lagrangian phase, it is assumed that the computational frame is following the material motion. For pure cells which contain just one material, staggered compatible discretization for Lagrangian gas dynamics is used to update the velocity and position of the node for the movement of the mesh [Caramana et al. (1998)]. For mixed cells containing more than one material, a thermodynamic closure model is needed to define how the volume fractions and the thermodynamic states of the individual materials evolve during the Lagrangian step.

There are several closure models in the literature. The first one is the mean strain rate model [Benson (1992); (1997)] in which each material in the mixed cell takes the mean strain rate of the cell. Actually, it simply assumes that the volume fraction of each material remains unchanged during the Lagrangian step, which may produce nonphysical results in some cases. The second one is the pressure equilibration model [Benson et al. (2004)]. It imposes instantaneous equilibration of the pressure at the cell level. The equilibration problem is nonlinear and sophisticated iteration

schemes are necessary for a robust implementation. The third one is the pressure relaxation model [Tipton (1989); Kamm and Shashkov(2010)]. In this model, a relaxation mechanism like viscosity is introduced to make the pressure within a mixed cell move toward pressure equilibration. The fourth one is the bulk modulus weighting model [Miller et al. (2007)]. In this model, when the mixed cell is in compression, the bulk modulus weighting algorithm is used; but the volume fraction keeps unchanged when the mixed cell is expanding. The fifth one is the contact mixture model, which was developed to permit slip and separation by solving the jump conditions for the stress and the strain rate across each interface [Benson (1997)]. The sixth one is a one-dimensional model called the sub-cell dynamics model [Barlow (2001)]. In this model, one first estimates the velocity normal to the interface between materials using the acoustic Riemann solver and then approximates the change of volume fraction for each material. In general, the first and the second closure models are very simple and only thermodynamic state dependent while the third and the fourth models are path or process dependent. The fifth and sixth models are more complex and more realistic which are wave structures dependent.

Among these models, the pressure relaxation model is used in this paper because it is more efficient and more effective for three-dimensional MMALE computation.

When the MOF interface reconstruction algorithm is coupled to a Lagrangian hydrodynamics scheme in MMALE methods, it is required that the positions of material centroids in mixed cells be updated during the Lagrangian phase. Following the idea of the constant parametric coordinate method presented in [Kucharik et al. (2010) ], it is assumed that the parametric coordinates of the material centroids keep unchanged during the Lagrangian phase. It is proved in [Kucharik et al. (2010) ] that this method gives a second-order approximation provided that the time step is small enough. Here we extend the method from 2D to 3D, and present a new approach to compute the parametric coordinates in a hexahedron [Jia et al. (2013)]. In this approach, a good approximation for the initial value of the parametric coordinates is given at first, and then Newton iteration is used to obtain accurate value of it. The convergence of this algorithm is quite fast.

Under the assumption that the materials of the fluid are immiscible, MOF algorithm is used to reconstruct the interface in mixed cells. MOF algorithm is second-order accurate. This method uses information not only about volume fraction but also about position of the centroid for each material. Also it provides automatic ordering of the materials in the process of interface reconstruction. In the case of three-dimensional unstructured meshes, the reconstructed interface is a plane which is chosen to match exactly the volume fraction and to provide the best possible approximation to the centroid positions of the materials. For more details about the numerical implementation refer to [Ahn and Shashkov (2007); Anbarlooei and Mazaheri(2009); Dyadechko and Shashkov (2008)].

In the remapping phase, the physical variables are interpolated from the Lagrangian mesh onto the rezoned mesh. In this paper, by simplifying and improving the method in [Goncharov and Yanilkin (2004)], we develop a second-order accurate remapping method on three-dimensional unstructured mesh [Jia et al. (2013)]. It is a cell-intersection-based method which calculates the volume and centroid of the intersection polyhedron between the old and new cells. It is suitable for remapping between two meshes with different topology.

## Numerical Examples

### *Example 1 3D Periodic Vortex Problem*

The three dimensional periodic vortex problem is constructed following the idea of manufactured analytical solution [Salari and Knupp (2000)]. It is the extension of the two-dimensional periodic vortex problem in [Shu (1998); Jia et al. (2013)]. The numerical results proved that MMALE method reaches second-order with mesh refinement.

### *Example 2 3D Noh Problem*

The Noh problem [Noh (1987)] has been used extensively to validate Lagrangian and ALE schemes in the regime of strong shock waves. A perfect gas with  $\gamma = 5/3$  is given an initial unit inward radial velocity. The initial thermodynamic state is given by  $\rho = 1$  and  $p = 0$ . A spherical shock wave is generated at the origin and moves with constant speed  $1/3$ . At time  $t = 0.6$ , the shock wave has radial coordinate  $0.2$ . The density behind shock is  $\rho = 64$ . The initial domain is  $[0, 1] \times [0, 1] \times [0, 1]$  decomposed with a  $48 \times 48 \times 48$  orthogonal mesh. At the initial time, in the vicinity of the spherical  $x^2 + y^2 + z^2 = 1/4$ , we place a layer of mixed cells. We note that in these mixed cells the two materials

are indeed perfect gases with the same polytropic index  $\gamma$ , and that we treat them as mixed cells to compare the numerical solutions obtained by the MMALE method with the analytical solutions. To run this test we do not need MMALE strategy, traditional pure Lagrangian schemes usually performs well. However, we run this test with MMALE method just for the sake of validation.

The mesh and the interface at  $t = 0.6$  obtained by Lagrangian computation (left) and by MMALE computation (right) are shown in Fig. 2. It can be seen that the final meshes obtained by Lagrangian computation and by MMALE computation both have good quality, the position of interface for both methods are almost the same. The density distributions at  $t = 0.6$  are shown in Fig. 3. The peak densities obtained by Lagrangian computation and by MMALE computation reach the value 52.6785 and 60.0365, respectively. It is obvious that MMALE result is better.

More numerical results can be found in reference [Jia et al. (2013)].

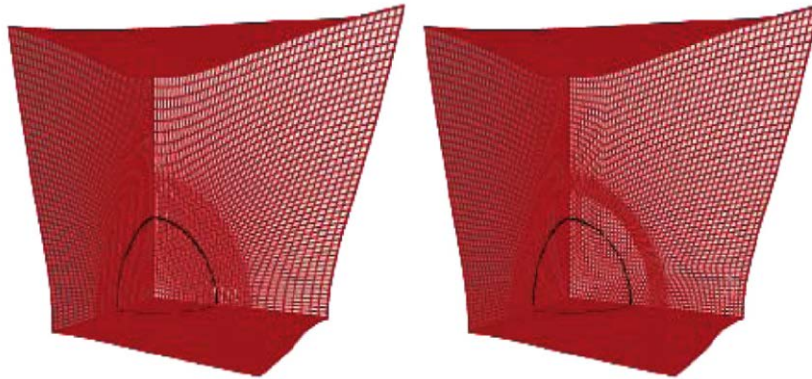


Figure 2. mesh and the interface of Noh problem (Left ,Lagrangian ; Right MMALE)

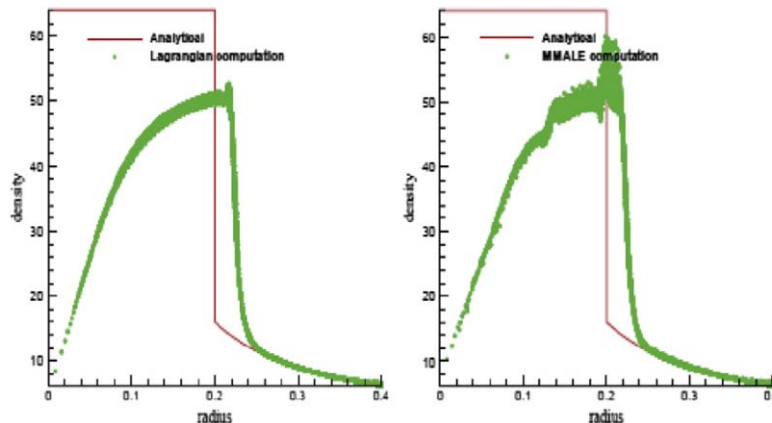


Figure 3. Density distribution at  $t=0.6$  of Noh problem (Left ,Lagrangian ; Right MMALE)

### Concluding Remark

A second order three dimensional unstructured MMALE method is presented in this paper, which uses MOF interface reconstruction for simulating multi-material compressible fluid flows involving strong shearing deformation. Numerical test proved the method is of second-order accuracy for continuous solutions. The application for 3D Noh Problem has shown the effectiveness and robust of the MMALE method.

### Acknowledgements

This research was sponsored by the National Nature Science Foundation of China (11072040, 11472060), and Projects (2012A0202010, 2012A0201011, 2011B0201042, 2009A 0202014) supported by CAEP.

## References

- J. Donea, S. Giuliani, J.P. Halleux, *Comput. Methods Appl. Mech. Eng* 33, 689–723(1982)
- J.S. Peery, D.E. Carroll, *Comput. Methods Appl. Mech. Eng* 187, 591–619(2000)
- Zupeng Jia, Jun Liu, Shudao Zhang, *J. Comput. Phys.* 236, 513-562(2013)
- E. Aulisa, S. Manservigi, R. Scardovelli, S. Zaleski, *J. Comput. Phys.*, 225, 2301–2319(2007)
- M. Kucharik, R.V. Garimella, S.P. Schofield, M. Shashkov, *J. Comput. Phys.* 229,2432–2452(2010)
- E.J. Caramana, D.E. Burton, M.J. Shashkov, P.P. Whalen, *J. Comput. Phys.* 146, 227–262(1998)
- D.J. Benson, *Comput. Methods Appl. Mech. Eng.* 99:,235-394(1992)
- D.J. Benson, *Comput. Methods Appl. Mech. Eng.* 140,59-86(1997)
- D.J. Benson, S. Okazawa, *Comput. Methods Appl. Mech. Eng.* , 193,4277–4298(2004)
- R.E. Tipton, CALE mixed zone pressure relaxation model, personal notes, 1989, unpublished.
- J.R. Kamm, M.J. Shashkov, *Commun. Comput. Phys.* 7,927–976(2010)
- D.S. Miller, G.B. Zimmerman, An Algorithm For Time Evolving Volume Fractions in Mixed Zones in Lagrangian Hydrodynamics Calculations, Numerical Methods For Multi-Material Fluid Flows, Prague, Czech Republic Sept 10–14, 2007. Available at: <<http://www.troja.fjfi.cvut.cz/multimat07/presentations/friday/Miller.pdf>>.
- A. Barlow, A new Lagrangian scheme for multi-material cells, in: Proceedings of European Congress on Computational Methods in Applied Sciences and Engineering, ECCOMAS Computational Fluid Dynamics Conference 2001, Swansea, Wales, U.K., 4–7 September 2001, 235–294.
- M. Kucharik, R.V. Garimella, S.P. Schofield, M. Shashkov, *J. Comput. Phys.*, 229,2432–2452( 2010)
- H.T. Ahn, M. Shashkov, *J. Comput. Phys.* ,226,2096–2132, 2007
- H.R. Anbarlooei, K. Mazaheri, *Comput. Methods Appl. Mech. Eng.*, 198,3782–3794(2009)
- V. Dyadechko, M. Shashkov, *J. Comput. Phys.*,227,5361–5384( 2008 )
- E.A. Goncharov, Y.V. Yanilkin, New method for computations of thermodynamical states of the materials in the mixed cells, VANT, *Math. Model. Phys. Process.* 3 (2004) 16–30 (in Russian).
- K. Salari, P.Knupp, Code Verification by the Method of Manufactured Solutions, Sandia National Laboratories, SAND 2000-1444, June, 2000
- C.-W. Shu, Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws, in: B. Cockburn, C. Johnson, C.-W. Shu, E. Tadmor (Eds.), *Advanced Numerical Approximation of Non-linear Hyperbolic Equations*, in: A. Quarteroni (Ed.), *Lecture Notes in Mathematics*, vol. 1697, Springer, Berlin, 1998, 325–432.
- W.F. Noh, *J. Comput. Phys.* 72 , 78–120(1987)
- A.M. Winslow, *J. Comput. Phys.* 135 , 128–138(1997)



# Modelling interaction of incompressible fluids and deformable particles with the Material Point Method

Rachel Gelet<sup>1,3\*</sup>, Giang Nguyen<sup>2,3</sup>, Pierre Rognon<sup>3</sup>

<sup>1</sup> GeM laboratory, University of Nantes, 58 rue Michel Ange, BP 420, 44606, Saint-Nazaire Cedex, France.

<sup>2</sup> School of Civil, Environmental and Mining Engineering, The University of Adelaide, Adelaide, SA 5005, Australia.

<sup>3</sup> School of Civil Engineering, The University of Sydney, Sydney, NSW, 2006 Australia.

December 4, 2015

## Abstract

The long-range interaction between one incompressible fluid surrounding solid objects is quite common and includes suspensions, sedimentation, fluid motion around obstacles, and erosion. Concerning the active research field of fluid-solid interactions, the challenging point nowadays is to describe the fluid dynamics in the pore space of soils or concrete samples and to establish a full coupling between the fluid and the movable deformable solid phase. This paper describes an extension of the material-point method (MPM) to modelling the interactions of incompressible fluids and multi-body deformable particles, which are discretized by a collection of unconnected, Lagrangian, material points. Primary variables, such as displacement, velocity, pressure and acceleration, and material variables, such as mass, stress and strain are associated with these points. To solve the equations of motion, data mapped from the material points are used to update variables on a background Eulerian mesh. The mesh solution is then mapped back to material points. This standard particle-like method treats all materials in a uniform way, thus avoiding complicated mesh construction and automatically possessing a no-slip contact algorithm at no additional cost.

---

\*rachel.gelet@univ-nantes.fr

In this study, the solid phase is treated as elastic, but general inelastic descriptions can also be later included to explore the interaction with the fluid phase. On the other hand, problems of incompressibility introduce numerical difficulties which need to be treated. Hence the enhanced strain method is adapted to the MPM analysis and specified to the study of long-range hydrodynamic interactions between incompressible fluid and solid deformable objects. Numerical examples including a fluid flow around an obstacle, the collapse of a water column and a sedimentation test are used to illustrate the proposed approach and its potential. The results of the MPM are compared with those obtained with classical FEM, XFEM and a modified immersed boundary method. In addition, the MPM results also compare well with existing experimental measurements of the collapsing water problem.

Keywords: Material Point Method (MPM); Fluid-Solid Interaction (FSI); Incompressible fluid; Enhanced Strain Method

## 1 Introduction

The long-range interaction between one incompressible fluid surrounding solid objects is quite common and includes suspensions, sedimentation, fluid motion around obstacles, and internal erosion. For example, the modelling of internal erosion at the scale of pore constrictions requires a complete description of both the grains/particles and the fluid within the pores [12]. While continuous methods are used at the scale of civil engineering structures, such as embankment dams [4] or sandstones reservoirs [19], discrete and micro-mechanical methods are being developed to describe the fluid dynamics in the pore space of soils and to establish a full coupling between the fluid and the movable *deformable* solid phase.

Concerning the active research field of fluid-solid interactions (FSI), a number of numerical methods are been developed [10]: Eulerian-Eulerian, Lagrangian-Lagrangian and Eulerian-Lagrangian methods.

Lagrangian-Lagrangian methods are based on the idea that calculation points (i.e. nodes or particles) of both phases are fixed to the domain being modelled throughout the analysis. This leads to the disappearance of convective terms in the governing equations of the model. The codes are then conceptually simpler and faster in this aspect, in contrast to Eulerian-Eulerian approaches. Since nodes/particles are placed and remain on material surfaces, interface tracking should be trivial. Those methods can be further subclassified into grid-based methods and meshless/particles methods. Un-

like grid-based methods [1, 15], meshless or particles methods approximate partial differential equations only based on a set of points without the need for an additional mesh (no nodal connectivity is introduced) [5]. The advantages of those methods, absence of a mesh, continuity of the shape functions and convergence, are damped by a few difficulties such as the application of essential boundary conditions, the computational effort and a certain sensitivity of the solutions to the inhomogeneous repartition of the particles.

The material point method (MPM) is a particle method that is based on the approximation of the weak form of partial differential equations [23, 24]. Aside from the advantages from the Lagrangian frame of reference and from the particle approach in terms of interface tracking, the key characteristic of the MPM is to use a background finite element mesh to solve the governing equations. Since this mesh does not carry any information, a regular grid can be utilized throughout the simulations. Thereby, the MPM combines the good features of both the finite element methods and the purely particle based methods. Finally, the MPM enjoys a single-valued velocity field which allows a natural treatment of no-slip contact and hence straightforward simulations of multi-phase materials.

Regarding geomechanical problems, different numerical methods have been applied to FSI problems with some success [2, 17, 12]. However, several drawbacks remain and there exists a need for a method combining the possibility of accurate representation of large deformation and displacement, together with an easy representation of solid-fluid interfaces and particles of random size and shape. At the moment, the use of the MPM is in the development of the research activities of our group [20, 8, 16] and pertains to a wider project targeting the development of grain contact laws in a fluid environment for the DEM. In this context, this work is a first step to establish a numerical framework and tools for the investigation of fluid-solid interactions, towards the ultimate goals of exploring immersed grains and developing corresponding DEM contact laws in the next step.

This paper describes an extension of the MPM to model the interactions of incompressible fluids and multi-body deformable particles. In the second section of this article, the governing equations are presented and the framework of the MPM is briefly addressed. The third section is devoted to the weak form of the governing equations, in which an enhanced strain element is used to represent the incompressible fluid phase. To highlight the potentials of the proposed method, we present in a final section three numerical examples: a fluid flow around an obstacle, the collapse of a water column and a sedimentation test.

## 2 Governing equations

In this section, the fundamental equations governing the problem are introduced: i.e. balance equations, and constitutive laws for the solid and liquid phases. Prior to the weak form, the MPM framework is presented which implies the spatial discretisation of each phases with material points endowed with point masses.

Motion of a continuum is governed by conservation of momentum and mass. Let's introduce the following global notations:  $\rho$  is the mass density,  $\mathbf{a}$  is the acceleration,  $\mathbf{v}$  is the velocity,  $\boldsymbol{\sigma}$  is a symmetric Cauchy stress tensor, and  $\mathbf{b}$  is the specific body force. For the whole domain (both the fluid and the solid phases), the global balance of momentum in the Lagrangian frame of reference and the global balance of mass read,

$$\begin{aligned} \operatorname{div} \boldsymbol{\sigma} + \rho \mathbf{b} &= \rho \mathbf{a}, \\ \dot{\rho} &= -\rho \operatorname{div} \mathbf{v}, \end{aligned} \tag{1}$$

For solid and fluid material points, the general form of the constitutive equation relates the stress rate, or the stress, to the strain rate *via* a tangent modulus,

$$\begin{aligned} \dot{\boldsymbol{\sigma}}_s &= \mathbf{T}_s : \dot{\boldsymbol{\epsilon}}, \\ \boldsymbol{\sigma}_f &= \lambda_f \operatorname{tr}(\dot{\boldsymbol{\epsilon}}) \mathbf{I} + 2\mu_f \dot{\boldsymbol{\epsilon}} = \mathbf{T}_f : \dot{\boldsymbol{\epsilon}}, \end{aligned} \tag{2}$$

with  $\lambda_f = 2\nu_f\mu_f/(1 - 2\nu_f)$  the Lamé parameter,  $\mu_f$  the dynamic viscosity and  $\nu_f$  the Poisson's ratio. Clearly, as  $\nu_f \rightarrow \frac{1}{2}$ , the Lamé parameter approaches infinity, so that nearly incompressible cases are characterised by  $\mu_f \ll \lambda_f$ . For simplicity, constitutive equations are presented in terms of the small deformation theory, and the strain rate tensor is related to the velocity through,

$$\dot{\boldsymbol{\epsilon}} = \frac{1}{2} [\nabla \mathbf{v} + (\nabla \mathbf{v})^T] = \nabla^* \mathbf{v}. \tag{3}$$

More general responses for the solid have already been implemented in the MPM [26, 16, 13] and are left for future work.

This work pertains to the classical MPM framework [23, 24], so that the governing equations are solved in a Lagrangian frame on a finite element mesh and the global mass density can then be written as a sum of point masses  $M^p$  by use of the Dirac delta function,

$$\rho(\mathbf{x}) = \sum_{p=1}^{n^p} M^p \delta(\mathbf{x} - \mathbf{X}^p). \quad (4)$$

The superscript  $p$  indicates a material point which is endowed with a fixed mass  $M^p$ , a position  $\mathbf{X}^p$ , a stress  $\boldsymbol{\sigma}(\mathbf{X}^p)$  and specific material parameters. Specific to our FSI problem, elements can be mixed, i.e. composed of both solid and fluid phases. However, material points are either solid or fluid. Hence, the finite collection of material points  $n^p$  gathers both the fluid material points  $n_f^p$  and the solid ones  $n_s^p$ , i.e.  $n^p = n_f^p + n_s^p$ . As a consequence, the grid forces accumulate the internal forces from both phases,

$$\mathbf{f}_i \propto \sum_{p=1}^{n_f^p} M^p \operatorname{div} \bar{\boldsymbol{\sigma}}_f(\mathbf{X}^p) + \sum_{p=1}^{n_s^p} M^p \operatorname{div} \bar{\boldsymbol{\sigma}}_s(\mathbf{X}^p) = \sum_{p=1}^{n^p} M^p \operatorname{div} \bar{\boldsymbol{\sigma}}(\mathbf{X}^p). \quad (5)$$

### 3 Method of solution

Problems of incompressibility are well known to introduce numerical difficulties, such as mesh locking, in finite element (FE). The numerous solutions developed in the realm of fluid mechanics are summarized in the comprehensive reviews of [7, 11]. A review of those methods is beyond the scope of this paper. Rather, a simple approach is tested here within the MPM and prove quite effective for our purpose (See section 4). This approach assumes a nearly incompressible fluid and uses the assumed strain method [21]. The weak form of the problem, specified to FSI problems and the MPM, is summarized below.

First, let's assume strain and stress discontinuity across the grid elements [25] holding fluid particles. The key point is to use a three-field variational formulation for the fluid domain  $\Omega_f$  and a standard formulation for the solid domain  $\Omega_s$ . Regarding the fluid phase we introduce the following enhanced forms of the strain rate field  $\dot{\boldsymbol{\epsilon}}$  and the strain rate variation field  $\boldsymbol{\gamma}$ ,

$$\begin{aligned} \dot{\boldsymbol{\epsilon}} &= \nabla^* \mathbf{v} & + & \tilde{\boldsymbol{\epsilon}} \\ \boldsymbol{\gamma} &= \nabla^*(\delta \mathbf{v}) & + & \tilde{\boldsymbol{\gamma}} \end{aligned} \quad (6)$$

in which  $\nabla^* \mathbf{v}$  is the symmetric gradient of the velocity field  $\mathbf{v}$  and  $\delta \mathbf{v}$  is the velocity variation field. Within the three-field formulation, two stresses are introduced:  $\boldsymbol{\sigma}$  the actual stress tensor and  $\boldsymbol{\sigma}^\epsilon$  the stress tensor which

satisfies the constitutive equation (2)<sub>2</sub>. In addition,  $\delta\boldsymbol{\sigma}$  is the stress variation field. Next, we consider the three standard variational equations:

$$\begin{aligned} \int_{\Omega} \delta\mathbf{v} \cdot (\operatorname{div} \boldsymbol{\sigma} + \rho\mathbf{b} - \rho\mathbf{a}) \, d\Omega &= 0, \\ \int_{\Omega_f} \delta\boldsymbol{\sigma} \cdot (\nabla^*\mathbf{v} - \dot{\boldsymbol{\epsilon}}) \, d\Omega_f &= 0, \\ \int_{\Omega_f} \tilde{\boldsymbol{\gamma}} \cdot (-\boldsymbol{\sigma} + \boldsymbol{\sigma}^{\boldsymbol{\epsilon}}) \, d\Omega_f &= 0. \end{aligned} \quad (7)$$

A modified form of this latter formulation is obtained by performing a series of steps: integrate by part the divergence term and use the divergence theorem, split the obtained stress term in two parts related to  $\Omega_s$  and  $\Omega_f$ , substitute eq. (6)<sub>1</sub> into (7)<sub>2</sub> and (6)<sub>2</sub> into (7)<sub>3</sub>, and acknowledge that the standard strain rate variation  $\nabla^*(\delta\mathbf{v})$  and the enhanced strain rate variation  $\tilde{\boldsymbol{\gamma}}$  are independent [21]. Finally, by following the second idea of the assumed strain method [21], the explicit presence of the stress term (within the fluid domain) is eliminated from the modified three-field variational formulation by choosing the space of the stress field  $L_2$ -orthogonal to the space of enhanced strains. As a result, the following modified three-field variational problem writes,

$$\begin{aligned} \int_{\Omega_s} \nabla^*(\delta\mathbf{v}) : \boldsymbol{\sigma} \, d\Omega_s + \int_{\Omega_f} \nabla^*(\delta\mathbf{v}) : \boldsymbol{\sigma}^{\boldsymbol{\epsilon}} \, d\Omega_f + \int_{\Omega} \delta\mathbf{v} \cdot \rho\mathbf{a} \, d\Omega = \\ \int_{\Omega} \delta\mathbf{v} \cdot \rho\mathbf{b} \, d\Omega + \int_{\Gamma_s^t} \delta\mathbf{v} \cdot \mathbf{t} \, d\Gamma_s^t. \\ \int_{\Omega_f} \tilde{\boldsymbol{\gamma}} \cdot \boldsymbol{\sigma}^{\boldsymbol{\epsilon}} \, d\Omega_f = 0. \end{aligned} \quad (8)$$

in which  $\boldsymbol{\sigma}_s \cdot \mathbf{n} = \mathbf{t}$  on  $\partial\Omega_s^t$  and  $\mathbf{n}$  is the unit vector outward normal to the boundary. Specific to this FSI framework, no distinction is introduced between the actual stress tensor of the solid phase and the stress tensor which satisfies the constitutive equation (2)<sub>1</sub>.

The following development follows the standard MPM approach [23]. The substitutions of eq. (4) into eq. (8) convert integrals to sums of quantities evaluated at material points,

$$\begin{aligned}
 & \sum_{p=1}^{n_s^p} \frac{M^p}{\rho^p} \nabla^*(\delta \mathbf{v})|_{\mathbf{x}=\mathbf{X}^p} : \boldsymbol{\sigma}(\mathbf{X}^p) + \sum_{p=1}^{n_f^p} \frac{M^p}{\rho^p} \nabla^*(\delta \mathbf{v})|_{\mathbf{x}=\mathbf{X}^p} : \boldsymbol{\sigma}^\epsilon(\mathbf{X}^p) \\
 & + \sum_{p=1}^{n^p} M^p \delta \mathbf{v}(\mathbf{X}^p) \cdot \mathbf{a}(\mathbf{X}^p) = \sum_{p=1}^{n^p} M^p \delta \mathbf{v}(\mathbf{X}^p) \cdot \mathbf{b}(\mathbf{X}^p) + \int_{\Gamma_s^t} \delta \mathbf{v} \cdot \mathbf{t} \, d\Gamma_s^t \quad (9) \\
 & \sum_{p=1}^{n_f^p} \frac{M^p}{\rho_f} \tilde{\boldsymbol{\gamma}}(\mathbf{X}^p) \cdot \boldsymbol{\sigma}^\epsilon(\mathbf{X}^p) = 0.
 \end{aligned}$$

The Galerkin method is adopted for the spatial discretization of variables and test functions. The spatial discretization uses an enhanced strain element. Each element is endowed with four displacement nodes and five enhanced strain nodes (Q1E5) [21]. A grid of isoparametric quadrilateral elements is used to define standard nodal basis functions,  $\mathbf{N}_u(\mathbf{x})$ , with  $n_u^n$  being the total number of displacement nodes,

$$\begin{aligned}
 \delta \mathbf{u}(\mathbf{x}) &= \sum_{i=1}^{n_u^n} \delta \mathbf{u}_i^e N_u^i(\mathbf{x}) = \mathbf{N}_u(\mathbf{x}) \delta \mathbf{u}^e, \quad \delta \mathbf{v}(\mathbf{x}) = \mathbf{N}_u(\mathbf{x}) \delta \mathbf{v}^e, \\
 \mathbf{u}(\mathbf{x}) &= \mathbf{N}_u(\mathbf{x}) \mathbf{u}^e, \quad \mathbf{v}(\mathbf{x}) = \mathbf{N}_u(\mathbf{x}) \mathbf{v}^e, \quad \mathbf{a}(\mathbf{x}) = \mathbf{N}_u(\mathbf{x}) \mathbf{a}^e \quad (10) \\
 \tilde{\boldsymbol{\gamma}}(\mathbf{x}) &= \mathbf{G}(\mathbf{x}) \tilde{\boldsymbol{\gamma}}^e, \quad \tilde{\boldsymbol{\epsilon}}(\mathbf{x}) = \mathbf{G}(\mathbf{x}) \tilde{\boldsymbol{\epsilon}}^e,
 \end{aligned}$$

while  $\mathbf{G}(\mathbf{x})$  is the enhanced strain rate interpolation matrix of size  $5 \times 4$ , for a 2D plain strain problem.  $\delta \mathbf{u}_i^e$ ,  $\delta \mathbf{v}^e$ ,  $\mathbf{u}_i^e$ ,  $\mathbf{v}_i^e$ ,  $\mathbf{a}_i^e$ ,  $\tilde{\boldsymbol{\gamma}}^e$  and  $\tilde{\boldsymbol{\epsilon}}^e$  denote the nodal vectors of the approximated functions.

For an arbitrary field  $\delta \mathbf{v}^e$  and by use of definitions (2)<sub>2</sub> and (6)<sub>1</sub>, the fluid contribution to the internal force vector may be expressed as,

$$\begin{aligned}
 \sum_{p=1}^{n_f^p} V^p \mathbf{B}^T(\mathbf{X}^p) \boldsymbol{\sigma}_f^\epsilon(\mathbf{X}^p) &= \sum_{p=1}^{n_f^p} V^p \mathbf{B}^T(\mathbf{X}^p) \mathbf{T}_f (\nabla^* \mathbf{v}(\mathbf{X}^p) + \tilde{\boldsymbol{\epsilon}}(\mathbf{X}^p)) \\
 &= \sum_{p=1}^{n_f^p} V^p \mathbf{B}^T(\mathbf{X}^p) \mathbf{T}_f (\mathbf{B}(\mathbf{X}^p) \mathbf{v}^e + \mathbf{G}(\mathbf{X}^p) \tilde{\boldsymbol{\epsilon}}^e) \quad (11)
 \end{aligned}$$

in which  $\mathbf{B}(\mathbf{X}^p) = \nabla \mathbf{N}_u(\mathbf{x})|_{\mathbf{x}=\mathbf{X}^p}$  is the strain displacement matrix evaluated at  $\mathbf{X}^p$ . Finally, the matrix form of the system is obtained for the arbitrary components  $\delta \mathbf{v}^e$  and  $\tilde{\boldsymbol{\gamma}}^e$ ,

$$\begin{aligned} \mathbf{K}^e(\mathbf{X}^p) \mathbf{u}^e + \mathbf{D}^e(\mathbf{X}^p) \mathbf{v}^e + \mathbf{M}^e(\mathbf{X}^p) \mathbf{a}^e + (\boldsymbol{\Gamma}^e(\mathbf{X}^p))^T \tilde{\boldsymbol{\epsilon}}^e &= \mathbf{F}_{\mathbf{v}}^e(\mathbf{X}^p) \\ &+ \boldsymbol{\Theta}^e(\mathbf{x}), \quad (12) \\ \boldsymbol{\Gamma}^e(\mathbf{X}^p) \mathbf{v}^e &+ \mathbf{H}^e(\mathbf{X}^p) \tilde{\boldsymbol{\epsilon}}^e = 0, \end{aligned}$$

in which we have introduced the discrete stiffness matrix  $\mathbf{K}^e(\mathbf{X}^p)$ , the discrete diffusion matrix  $\mathbf{D}^e(\mathbf{X}^p)$ , the discrete mass matrix  $\mathbf{M}^e(\mathbf{X}^p)$ , the discrete coupled matrix  $\boldsymbol{\Gamma}^e(\mathbf{X}^p)$ , the discrete body force vector  $\mathbf{F}_{\mathbf{v}}^e(\mathbf{X}^p)$ , the discrete traction vector  $\boldsymbol{\Theta}^e(\mathbf{x})$  and the discrete enhanced strain rate matrix  $\mathbf{H}^e(\mathbf{X}^p)$ . The definitions of those matrices are given in Appendix A.

The system of equations (12), which is a compact system of  $(n_{dim} \times n_u^n + n_{\boldsymbol{\epsilon}}^n)$  equations and unknowns, is solved fully implicitly by use of a standard Newmark scheme and a Newton Raphson procedure. The obtained formulation is further condensed out so that the enhanced strain unknowns disappear. It is proposed that the obtained formulation be simulated with the MPM. The idea is straightforward in that the algorithm is setup as any other type of MPM simulation. More details on the complete algorithm can be found in the following seminal references [23, 24].

## 4 Results and discussions

The potentials of the proposed method are highlighted with three numerical simulations, in two-dimensions. First a fluid flow around a cylindrical obstacle demonstrates the performance of the proposed algorithm for incompressible fluids. Next, the method is tested for problems involving free surfaces. The classic results of a collapsing water column are compared with experimental and XFEM numerical data. Finally, the accuracy of the method is tested on a sedimentation test. It was found that the results obtained with the MPM are more accurate than those of an immersed boundary method.

### 4.1 Fluid flow around a cylindrical obstacle

To test the properties of the model for an incompressible fluid, several calculations are performed by using a fixed elastic but very stiff obstacle. Fluid flow is scrutinized past a unit-radius ( $r$ ) cylindrical obstacle located at the



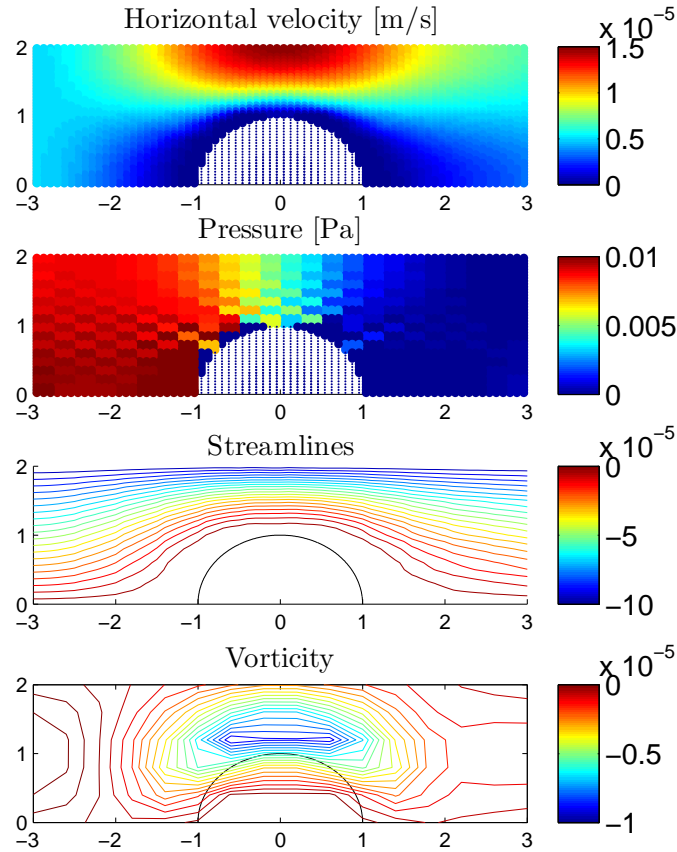


Figure 1: The particles defining the incompressible fluid and the obstacle along an axis of symmetry are shown for a  $24 \times 16$  grid with 9 MPs per Q1E5 element. The horizontal velocity and the pressure contours reflect the good capabilities of the method, while the streamlines and the vorticity are characteristic of ‘sticky’ Stokes flows.

center of the following domain  $-3 \leq x \leq 3$  and  $-2 \leq y \leq 2$  [7]. The symmetry of the setup allows us to restrict ourselves to the upper half of the domain. The boundary conditions are as follows:  $v_x = v_0 f(t/t_{\text{endload}})$  and  $v_y = 0$  at the inlet ( $x = -3$ ) with  $f(x) = x^3 * (10 - 15 * x + 6 * x^2)$  for  $x < 1$ ,  $f(x) = 1$  for  $x \geq 1$  and  $t_{\text{endload}}$  being the acceleration time constant; traction free boundary conditions at the outlet ( $x = 3$ ) and symmetry elsewhere ( $y = 0$  and  $y = 2$ ) with  $v_y = 0$ . Initial velocities assume a no-flow configuration. The time step in the Newton Raphson scheme is chosen as

$\Delta t = 0.001$  s. The characteristic lengths defining the Reynolds number are  $v_0 = 5 \times 10^{-6}$  m/s,  $\nu_f = 0.1$  m<sup>2</sup>/s and  $r = 1$  m. The flow regime is hence close to a Stokes flow.

The calculated response is examined to test linear stability (Figure 1). The particles defining the incompressible fluid and the obstacle are attached to a  $24 \times 16$  uniform grid on which the system (12) solved. Both the horizontal velocity and the pressure contours are reasonable, the latter being symmetric about  $x = 0$  and close to zero at the exit [7]. Note that the pressure solution in mixed elements (in the vicinity of the obstacle) is inaccurate highlighting the limit of the method. The streamlines in Figure 1 show the ‘large’ displacement thickness characteristic of ‘sticky’ Stokes flows; while the vorticity has diffused in a nearly-symmetric shape.

Clearly, the classical FEM would lead to results of greater quality in the vicinity of the obstacle to the price of adapting the mesh to this obstacle. A similar result, would be obtained with the MPM and an unstructured mesh that respects the boundary of the obstacle. Yet the aim here is to show that even with an unstructured coarse mesh consisting of  $24 \times 16$  elements, the MPM can provide sufficiently accurate results.

## 4.2 A collapsing water column

This test case considers a domain of size  $0.584$  m  $\times$   $0.45$  m containing a water column on the left-hand-side of size  $a \times b = 0.146$  m  $\times$   $0.292$  m [6, 14]. While the water properties are  $\rho_f = 1000$  kg/m<sup>3</sup> and  $\mu_f = 1 \times 10^{-3}$  Pa.s, the rest of the domain is filled with an incompressible fluid endowed with the following properties  $\rho = 1$  kg/m<sup>3</sup> and  $\mu = 1 \times 10^{-5}$  Pa.s, representing air. No surface tension is considered. The loading is merely restricted to a volumetric gravitation force  $g = -9.81$  m/s<sup>2</sup>. Free-slip boundary conditions are assumed along the boundaries of the domain. Two meshes are used, consisting of  $12 \times 9$  and  $36 \times 27$  elements. Each mesh is endowed with 36 MPs per cell and the simulation time is restricted to 0.3 s with  $\Delta t = 3 \times 10^{-4}$  s.

The water column positions and pressure contour at some selected points in time are compared in Figure 2. The evolutions of the dimensionless water column width and height with dimensionless times are compared with experimental data [14] and the intrinsic XFEM solution [6] in Figure 3. Dimensionless displacements and times are defined as follows,

$$w^* = \frac{w}{a}, \quad h^* = \frac{h}{b}, \quad \tau_w = t\sqrt{\frac{2g}{a}}, \quad \tau_h = t\sqrt{\frac{2g}{b}} \quad (13)$$

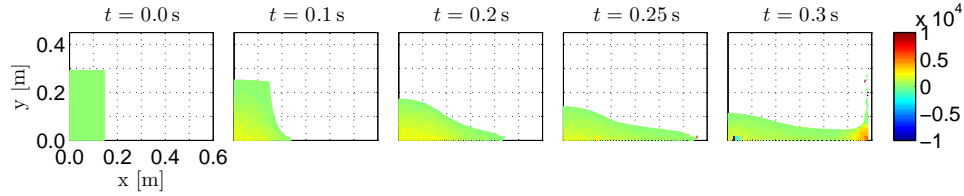


Figure 2: Water column position and pressure contour at 5 selected times for the Q1E5 element. The  $36 \times 27$  elements mesh uses 36 MPs per cell.

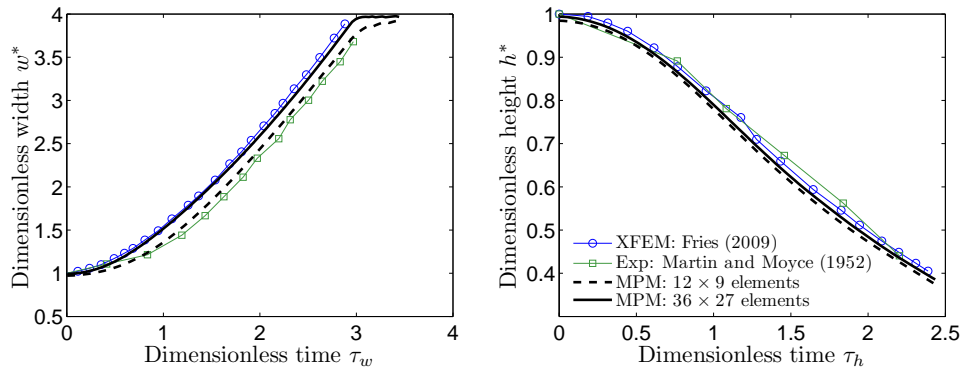


Figure 3: Dimensionless (left) width and (right) height over time, see eq. (13), for the collapsing water column test case. Each MPM mesh uses 36 MPs per cell. The MPM solution agrees well with other results in spite of the coarse nature of the meshes.

in which  $w$  and  $h$  correspond to the intersection points of the water column interface with the bottom and left walls of the domain, respectively, so that at  $t = 0$ ,  $w^* = h^* = 1$ . An excellent agreement of the MPM solution with other results is found where classical FEM results are known to be unsatisfactory [6].

### 4.3 Sedimentation test

This last test case mimics the sedimentation of a deformable cylinder in an incompressible fluid. We calculate the motion of the circular cylinder of radius  $a$  in between two parallel walls, of width  $2L = 2.0$  m. The motion of the cylinder, directed perpendicular to its axis with its axis positioned midway between the walls, is solely induced by gravity  $g = -9.81$  m/s<sup>2</sup>. The cylinder properties are  $\rho_s = 2700$  kg/m<sup>3</sup>,  $E_s = 36$  GPa and  $\nu_s = 0.25$  and

the incompressible fluid properties are  $\rho_f = 1000 \text{ kg/m}^3$  and  $\mu_f = 1000 \text{ Pa}\cdot\text{s}$ . The boundary conditions assume zero vertical and horizontal velocities on the vertical boundaries and zero vertical velocities only on the horizontal boundaries. The time step in the Newton Raphson scheme is chosen as  $\Delta t = 0.001 \text{ s}$ .

For a cylinder settling along the axis of an infinite channel, Faxen [3] presents a close form solution for small  $a/L$  that relates the ratio  $a/L$  to the steady state vertical velocity  $v_y^\infty$  (reported in Happel and Brenner [9], p. 345). This theory is compared to MPM simulations and to the work of Sulsky and Brackbill [22] for cylinders of various radius in Table 1.

The results are in agreement with the approximated theory [3] and the error remains small for  $a/L < 0.5$ . By using a modified immersed boundary technique, Sulsky and Brackbill [22] report an error decreasing in magnitude with increased resolution of the cylinder ( $a/\Delta x$ ) and increasing in magnitude up to 10% error with  $a/L = 0.4$ . In comparison, the error on the MPM solution seems less dependent of the mesh resolution and is most probably due to the influence of mixed elements.

$a/L$	$a/\Delta x$	$v_y^\infty$ [m/s] theory	$v_y^\infty$ [m/s] MPM	error% MPM	error% Sulsky and Brackbill [22]
0.1	5	0.0585	0.0591	-0.96	1.8
0.2	10	0.1267	0.1263	0.38	2.6
0.3	15	0.1611	0.1594	1.05	-0.8
0.4	20	0.1549	0.1557	-0.52	-10.5

Table 1: Velocity comparison of a settling cylinder between two rigid walls computed on a  $50 \times 100$  grid with 9 MPs per cell. In comparison with the results of Sulsky and Brackbill [22], the MPM response is less dependent on the mesh resolution ( $a/\Delta x$ ) and remains accurate within  $\approx 1\%$  for  $a/L \leq 0.4$ .

## 5 Conclusion

The MPM is applied to the interaction of incompressible fluids and deformable particles. This method uses Lagrangian material points and an

Eulerian grid or mesh to define the computational domain. The material points move through the Eulerian grid on which the balance equations are solved. This paper presents the modifications necessary to simulate the interactions between incompressible fluids and solid materials by using the assumed strain method. The spacial discretisation uses an enhanced strain element (Q1E5) [21].

Several two-dimensional test problems are presented to demonstrate the methodology. A fluid flow test past an obstacle is used to test the incompressible fluid model, for Stokes flows. The method performs well despite the fact that the mesh does not fit the obstacle. Next, a collapsing water column test is presented to evaluate the accuracy of the method. The MPM response compares very well with both experimental [14] and XFEM results [15]. Finally, a sedimentation test is performed to validate the fluid-solid interactions. Again the proposed method performs well with respect to the analytical solution [3] and better than other numerical methods such as the modified immersed boundary formulation [22].

The enhanced strain element Q1E5 used throughout this paper matches well the MPM procedure since it is compatible with the use of a lumped mass matrix during the initialization step. Yet this element does not satisfy of the inf-sup test so that one possible extension of the method is to test other elements such as the combined mixed displacement/pressure enhanced finite element [18]. Also the proposed element does not satisfy strictly the incompressibility condition  $\text{div } \mathbf{v} = 0$  and the results presented herein will be compared in a near future with that obtained with a bi-linear mixed element Q1P0 [7].

## Acknowledgments

The authors would like to thank the Australian Research Council for the Discovery Projects funding scheme (project DP110102645).

## A Discrete matrices used in eq. (12)

The discrete coupled matrix  $\mathbf{\Gamma}^e(\mathbf{X}^p)$ , the discrete diffusion matrix  $\mathbf{D}^e(\mathbf{X}^p)$ , and the discrete enhanced strain rate matrix  $\mathbf{H}^e(\mathbf{X}^p)$  write,

$$\begin{aligned}
 \mathbf{\Gamma}^e(\mathbf{X}^p) &= \sum_{p=1}^{n_f^p} V^p \mathbf{G}^T(\mathbf{X}^p) \mathbf{T}_f \mathbf{B}(\mathbf{X}^p), \\
 \mathbf{D}^e(\mathbf{X}^p) &= \sum_{p=1}^{n_f^p} V^p \mathbf{B}^T(\mathbf{X}^p) \mathbf{T}_f \mathbf{B}(\mathbf{X}^p), \\
 \mathbf{H}^e(\mathbf{X}^p) &= \sum_{p=1}^{n_f^p} V^p \mathbf{G}^T(\mathbf{X}^p) \mathbf{T}_f \mathbf{G}(\mathbf{X}^p).
 \end{aligned} \tag{14}$$

Also we have introduced the discrete mass matrix  $\mathbf{M}^e(\mathbf{X}^p)$ , the discrete body force vector  $\mathbf{F}_v^e(\mathbf{X}^p)$ , the discrete traction vector  $\mathbf{\Theta}^e(\mathbf{x})$  and the discrete stiffness matrix  $\mathbf{K}^e(\mathbf{X}^p)$  as,

$$\begin{aligned}
 \mathbf{M}^e(\mathbf{X}^p) &= \sum_{p=1}^{n^p} M^p \mathbf{N}_u^T(\mathbf{X}^p) \mathbf{N}_u(\mathbf{X}^p), \\
 \mathbf{F}_v^e(\mathbf{X}^p) &= \sum_{p=1}^{n^p} M^p \mathbf{N}_u^T(\mathbf{X}^p) \mathbf{b}(\mathbf{X}^p), \\
 \mathbf{\Theta}^e(\mathbf{x}) &= \int_{\Gamma_s^t} \mathbf{N}_u^T(\mathbf{x}) \mathbf{t} \, d\Gamma_s^t, \\
 \mathbf{K}^e(\mathbf{X}^p) &= \sum_{p=1}^{n_s^p} V^p \mathbf{B}^T(\mathbf{X}^p) \mathbf{T}_s \mathbf{B}(\mathbf{X}^p)
 \end{aligned} \tag{15}$$

It is worth noting that the number of integration points used to obtain the various discrete matrices is adapted depending on the nature of each term, i.e. the diffusion, coupled and enhanced strain rate matrices are summed over the fluid material points, the stiffness matrix is summed over the solid material points, and the mass matrix is evaluated using all material points of the element.

## References

- [1] T. Belytschko, B. Liu, W.K. and Moran, and K. Elkhodary. *Nonlinear finite elements for continua and structures*. John Wiley & Sons, 2013.
- [2] H.H. Bui, K. Sako, and R. Fukagawa. Numerical simulation of soil–water interaction using smoothed particle hydrodynamics (sph) method. *Journal of Terramechanics*, 44(5):339–346, 2007.

- [3] H. Faxen. Hydrodynamical resistance formula. *Neuvième Congrès des mathématiciens scandinaves à Helsingfors du 23 au 26 août 1938*, 1(9):165, 1939.
- [4] R. Fell and J.J. Fry. *Internal erosion of dams and their foundations*. Taylor & Francis, Oxon, Royaume Uni, 2007.
- [5] T.-P. Fries and H.-G. Matthies. Classification and overview of meshfree methods. Technical report, Institute of Scientific Computing Technical University Braunschweig Brunswick, Germany, 2004.
- [6] T.P. Fries. The intrinsic xfm for two-fluid flows. *International Journal for Numerical Methods in Fluids*, 60(4):437–471, 2009.
- [7] P.M. Gresho and R.L. Sani. *Incompressible flow and the finite element method. Volume 1: Advection-diffusion and isothermal laminar flow*. John Wiley and Sons, Inc., New York, NY (United States), 1998.
- [8] I. Guiamatsia and G.D. Nguyen. Crack modelling using the material point method and a strong discontinuity approach. In *Key Engineering Materials*, volume 525, pages 513–516. Trans Tech Publ, 2013.
- [9] H. Happel, J. and Brenner. *Low Reynolds number hydrodynamics: with special applications to particulate media*, volume 1. Springer Science & Business Media, 1983.
- [10] S. Ii, K. Sugiyama, S. Takeuchi, S. Takagi, and Y. Matsumoto. An implicit full eulerian method for the fluid-structure interaction problem. *International Journal for Numerical Methods in Fluids*, 65(1-3):150–165, 2011.
- [11] H.P. Langtangen, K.-A. Mardal, and R. Winther. Numerical methods for incompressible viscous flow. *Advances in Water Resources*, 25(8):1125–1146, 2002.
- [12] F. Lominé, L. Scholtès, L. Sibille, and P. Poullain. Modeling of fluid–solid interaction in granular media with coupled lattice boltzmann/discrete element methods: application to piping erosion. *International Journal for Numerical and Analytical Methods in Geomechanics*, 37(6):577–596, 2013.
- [13] E Love and DL Sulsky. An energy-consistent material-point method for dynamic finite deformation plasticity. *International Journal for Numerical Methods in Engineering*, 65(10):1608–1638, 2006.

- [14] J.C. Martin and W.J. Moyce. An experimental study of the collapse of liquid columns on a rigid horizontal plane. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 244(882):312–324, 1952.
- [15] J.E. Mindel. *Interface tracking and solid-fluid coupling techniques with coastal engineering applications*. PhD thesis, Imperial College London, 2008.
- [16] G.D. Nguyen. An enriched constitutive model for fracture propagation analysis using the material point method. *Applied Mechanics and Materials*, 553:731–736, 2014.
- [17] E. Oñate, S.R. Idelsohn, M.A. Celigueta, and R. Rossi. Advances in the particle finite element method for the analysis of fluid–multibody interaction and bed erosion in free surface flows. *Computer Methods in Applied Mechanics and Engineering*, 197(19):1777–1800, 2008.
- [18] D. Pantuso and K.-J. Bathe. A four-node quadrilateral mixed-interpolated element for solids and fluids. *Mathematical models and methods in applied sciences*, 5(08):1113–1128, 1995.
- [19] E. Papamichos and I. Vardoulakis. Sand erosion with a porosity diffusion law. *Computers and Geotechnics*, 32(1):47–58, 2005.
- [20] P. Rognon and C. Gay. Soft dynamics simulation. 1. normal approach of two deformable particles in a viscous fluid and optimal-approach strategy. *The European Physical Journal E: Soft Matter and Biological Physics*, 27(3):253–260, 2008.
- [21] J.C. Simo and M.S. Rifai. A class of mixed assumed strain methods and the method of incompatible modes. *International Journal for Numerical Methods in Engineering*, 29(8):1595–1638, 1990.
- [22] D. Sulsky and J.U. Brackbill. A numerical method for suspension flow. *Journal of Computational Physics*, 96(2):339–368, 1991.
- [23] D. Sulsky, Z. Chen, and H.L. Schreyer. A particle method for history-dependent materials. *Computer Methods in Applied Mechanics and Engineering*, 118(1):179–196, 1994.
- [24] D. Sulsky, S.-J. Zhou, and H.L. Schreyer. Application of a particle-in-cell method to solid mechanics. *Computer physics communications*, 87(1):236–252, 1995.



- [25] K. Washizu. *Variational methods in elasticity and plasticity*. Pergamon Press, New York, 1982.
- [26] A.R. York, D. Sulsky, and H.L. Schreyer. Fluid–membrane interaction based on the material point method. *International Journal for Numerical Methods in Engineering*, 48(6):901–924, 2000.

## Hydroelastic Analysis in Frequency Domain and Time Domain

\*Frank Lin, Michael Johnson, Zhenhong Wang, and Nigel White

Lloyd's Register Group Ltd

\*Corresponding author: frank.lin@lr.org

### Abstract

Hydroelasticity has been included in ship seakeeping assessment for more than three decades, and it has finally become an essential tool in the marine industry for design of some ship types. In the 35 years of evolution, the hydroelasticity methods applied in the marine and offshore energy industries have grown from two-dimensional to three-dimensional and now feature linear analysis models in the frequency domain and nonlinear models in the time domain. In this paper, we present the three-dimensional hydroelasticity theory model in the frequency domain and time domain, show the difference in the approach, and discuss their applications in wave-structure interaction.

**Keywords:** Hydroelasticity, Springing, Frequency domain, Time domain, Boundary element method, Linear, Nonlinear.

### Introduction

In the design process for floating structures, like ships and offshore structures, hydrodynamic analysis of wave-structure interaction is the first important step. The methods of rigid-body based seakeeping analysis have been applied successfully in this type of work for many decades, but suffer failures on some of latest mega-ships, like a container ship over 350 meters in length. It has been found that the predicted fatigue life of a large container ship based on a rigid-body approach is significantly longer than when the effects of elastic body responses are taken into account. The elastic-body based analysis method explicitly allows for the interaction of water waves and elastic structures. An ultimate hydroelasticity solution comes from a CFD approach, but this is too expensive to be applied for routine work; for example, the number of required regular wave cases will typically be 3,000 to 5,000 in a design process, and coupled with a few hundred combinations of ship speed, wave headings and sea states, leads to hundreds of thousands of hours of real time simulation. The boundary element hydroelasticity model remains the only tool practical for routine work. In this paper, a general approach for 3D hydroelasticity is presented. Differences between the rigid-body approach and the hydroelasticity approach are discussed. We also look into the theoretical details of the frequency domain hydroelasticity model orientated for conditions of low and moderate sea state, and that of the time domain hydroelasticity model orientated for high sea state conditions.

### Methodology of Hydroelasticity

Accurate prediction of hydrodynamic structural load is key to a successful strength assessment for a structure operated in waves. The hydrodynamic pressure is determined by the location and velocity of the wetted surface of the structure. A rigid-body approach will be accurate enough if the elastic deformation of the structure's wetted surface is small compared to that induced by rigid-body motion. Elastic deformation needs to be considered in the boundary condition of the boundary value problem of flow solution for ships or structures with less stiffness, such as a container ship

longer than 350 meters. The Hydroelasticity Method has been developed for the interaction of waves and elastic structures. Due to the interaction of the flow and the structure's motion and deformation, the hydrodynamic problem and structural dynamic problem is coupled together and needs to be solved simultaneously. Direct finite element structural analysis can be combined with a flow solver like RANS or BEM in time domain to form a robust nonlinear tool for hydroelastic assessment, but it will be too expensive to be practical for routine design assessment. FEM based modal analysis is usually used for the structural assessment portion, and its solution, eigen-values and eigen vectors are used with a boundary element method for hydrodynamic analysis, and this forms the so-called hydroelasticity method for seakeeping analysis. The first 2D frequency domain hydroelastic method was proposed by Bishop and Price in 1979. In their method, a ship was represented by a Timoshenko beam and discretized to a number of 2D beam elements for structural analysis, and the wave flow solution around the ship was determined by strip theory. This hydroelasticity method has continued evolving, now supporting a fully 3D structural FEM model with 3D BEM model for the hydrodynamic solution in both frequency domain and time domain.

In a finite element model of a structure, stress in an element can be estimated by the displacement of the node points of the element. A vector of model node displacements,  $\mathbf{U}$ , can be determined from the model elastic motion equation

$$[\mathbf{M}]\{\ddot{\mathbf{U}}\} + [\mathbf{B}]\{\dot{\mathbf{U}}\} + [\mathbf{K}]\{\mathbf{U}\} = \{\mathbf{P}\} + \{\mathbf{F}\} + \{\mathbf{G}\}, \quad (1)$$

where  $[\mathbf{M}]$ ,  $[\mathbf{B}]$  and  $[\mathbf{K}]$  are the matrix of model mass, structural damping and stiffness;  $\{\mathbf{P}\}$  is the vector of external surface force;  $\{\mathbf{F}\}$  is the vector of external concentrated force, and  $\{\mathbf{G}\}$  the vector of external mass force. Dot represents the gradient w.r.t time.

Introducing the homogenous solution of the node displacement vector,  $\{\mathbf{U}\} = \{\mathbf{D}\}e^{-i\omega t}$  and ignoring structural damping and all external forcing terms from equation (1), solution of equation

$$(-\omega^2[\mathbf{M}] + [\mathbf{K}])\{\mathbf{D}\} = \{\mathbf{0}\}, \quad (2)$$

gives the eigen value  $\omega_r$  and eigen vector  $\{\mathbf{D}\}_r$ , that define the dry eigenmodes. The number of eigenmodes of a FEA model will be the same as the number of degrees of freedom, being six times of the number of node points. Displacement at point  $(x, y, z)$  can be expressed by those dry eigenmodes in terms of summary

$$\{\mathbf{U}(x, y, z; t)\} = [\mathbf{D}(x, y, z)]\{\mathbf{q}(t)\} = \sum_r \mathbf{D}_r(x, y, z)q_r(t) \quad (3)$$

where  $q_r(t)$  is the amplitude of mode  $r$ , the so-called general coordinate, and

$$\{\mathbf{D}\}_r^j = (\mathbf{u}^r, \bar{\mathcal{G}}^r)_j^T = (u_r, v_r, w_r, \alpha_r, \beta_r, \chi_r)_j^T \quad (4)$$

is the displacement of point  $j$  induced by mode  $r$  with unit modal amplitude.

Multiplying  $[\mathbf{D}]^T$  on each term of equation (1), and right multiplying  $[\mathbf{D}]$  on the matrix of model mass, damping and stiffness, we have an equation to determine the modal amplitude

$$[\mathbf{m}]\{\ddot{\mathbf{q}}\} + [\mathbf{b}]\{\dot{\mathbf{q}}\} + [\mathbf{k}]\{\mathbf{q}\} = \{\mathbf{p}\} + \{\mathbf{f}\} + \{\mathbf{g}\}, \quad (5)$$

One of the advantages of using the dry eigenmode approach is the modal orthogonality. For any elastic dry modes  $r$  and  $s$ , using Kronecker delta, we have

$$\{\mathbf{D}\}_r^T [\mathbf{M}] \{\mathbf{D}\}_s = \delta_{rs} m_{rs} \quad \text{and} \quad \{\mathbf{D}\}_r^T [\mathbf{K}] \{\mathbf{D}\}_s = \delta_{rs} k_{rs} \quad (6)$$

where  $m_{rr}$  and  $k_{rr}$  are the modal mass and modal stiffness.

We can solve the modal amplitude by applying a location and velocity given by equation (3) on the wetted surface in a hydrodynamic analysis model and expressing and estimating the three forcing terms in the hydrodynamic model.

In a boundary element hydrodynamic model, linearized boundary surface condition for unsteady velocity potential  $\Phi^U$  can be given by the surface displacement  $(\mathbf{u}, \vec{g})$  and the steady flow velocity  $\mathbf{W}$

$$\frac{\partial \Phi^U}{\partial n} = \left[ \frac{\partial \mathbf{u}}{\partial t} + \vec{g} \times \mathbf{W} - (\mathbf{u} \cdot \nabla) \mathbf{W} \right] \cdot \mathbf{n} + O(|\mathbf{u}|^2). \quad (7)$$

Here  $\mathbf{n}$  is the surface normal vector, and surface displacement  $(\mathbf{u}, \vec{g})$  can be estimated from the shape functions, or eigen vectors, of the model.

Another fact worth noting in a hydroelastic model is that the eigenmode with nonzero displacement on the wetted surface will receive hydrodynamic pressure force, and we call these modes the “wetable modes”. All other modes, “non-wetable modes”, have no external force from hydrodynamic pressure. The wettable modes will be coupled to each other through hydrodynamic pressure force, which means the motion of the  $i$ -th wettable mode will induce a surface forcing term on  $j$ -th wettable mode. On the other hand, non-wetable modes are uncoupled. In a hydroelastic model, we only need to consider those wettable modes, usually only the first few wettable modes in practice.

Another difference between a rigid structure and an elastic structure is on the location of the center of gravity, COG. The COG of a rigid structure is a point fixed with the structure when it oscillates in waves. On the other hand, the COG is changing due to elastic deformation and not fixed with the elastic structure when it oscillates in waves. This difference leads to a much complicated equation for the rigid body motion mode of the elastic case.

Let's introduce two Cartesian coordinate systems: 1) the body-fixed frame, **HMF**,  $\tilde{o} - \tilde{x}\tilde{y}\tilde{z}$  with  $\tilde{x}$ -axis pointing to the bow,  $\tilde{o} - \tilde{x}\tilde{y}$  coordinate plane lying on the undisturbed water surface when the ship has no oscillations, and  $\tilde{z}$ -axis positive upward; 2) the moving reference frame, **HRF**,  $o - xyz$ , which is an inertial frame moving at the constant ship speed  $U$  and which is identical with the body-fixed frame if the ship has no oscillations. The coordinates of the body-fixed frame origin,  $\tilde{o}$ , in the reference frame **HMF**, namely  $\vec{\eta} = (\eta_1, \eta_2, \eta_3)$  define the translational motion of the ship, so called **Surge**, **Sway** and **Heave**. Three Euler angles  $(\eta_4, \eta_5, \eta_6)$  between the body-fixed frame  $\tilde{o} - \tilde{x}\tilde{y}\tilde{z}$  and the reference frame  $o - xyz$  define the rotational motion of the ship, also referred to as **Roll**, **Pitch** and **Yaw**. Supposing  $\tilde{o} - \tilde{x}\tilde{y}\tilde{z}$  rotates from the position of  $o - xyz$  with the angle  $\eta_6$  about

$\tilde{z}$ -axis first, then the angle  $\eta_3$  about  $\tilde{y}$ -axis, and finally the angle  $\eta_1$  about  $\tilde{x}$ -axis, we will have the relation between  $\tilde{o}-\tilde{x}\tilde{y}\tilde{z}$  and  $o-xyz$  as follow

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix} + \mathbf{R} \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} \quad (8)$$

where  $\mathbf{R}$  is the mapping matrix defined by the three Euler angles.

By setting the origin on the gravitational center of the structure with zero elastic deformation, the equations for rigid body motion modes, translational and rotational, can be given by

$$\begin{aligned} M \frac{d^2 \bar{\eta}}{dt^2} + \mathbf{R} \sum_s \frac{d^2 \eta_s}{dt^2} \mathbf{M}_s + \mathbf{R} \left\{ \frac{d\mathbf{w}}{dt} \times \left( \sum_s \eta_s \mathbf{M}_s \right) \right. \\ \left. + \mathbf{w} \times \left[ \mathbf{w} \times \left( \sum_s \eta_s \mathbf{M}_s \right) \right] + 2\mathbf{w} \times \left( \sum_s \frac{d\eta_s}{dt} \mathbf{M}_s \right) \right\} = \mathbf{F} \end{aligned} \quad (9)$$

$$\begin{aligned} M \left\{ \left( \sum_s \eta_s \tilde{\mathbf{r}}_s^g \right) \times \frac{d^2}{dt^2} + \left( \sum_s \frac{d\eta_s}{dt} \tilde{\mathbf{r}}_s^g \right) \times \frac{d}{dt} + \mathbf{w} \times \left[ \left( \sum_s \eta_s \tilde{\mathbf{r}}_s^g \right) \times \frac{d}{dt} \right] \right\} [\mathbf{R}^T(\bar{\eta})] + \\ \mathbf{J} \left( \frac{d\mathbf{w}}{dt} \right) + \mathbf{w} \times \mathbf{J}(\mathbf{w}) + \sum_s \frac{d^2 \eta_s}{dt^2} [\mathbf{J}_s + \mathbf{I}_s(\bar{\mathcal{G}}^s)] = \mathbf{M} \end{aligned} \quad (10)$$

where  $M$  and  $\mathbf{J}$  are the total mass and moment of mass inertia of the structure;  $\eta_s$  the amplitude of the  $s$ -th elastic mode;  $\mathbf{w}$  the vector of rotational velocity;  $\tilde{\mathbf{r}}_s^g$ ,  $\mathbf{M}_s$ , and  $\mathbf{J}_s$  are the modal mass center, modal mass vector and modal mass moment of the  $s$ -th elastic mode defined by

$$\tilde{\mathbf{r}}_s^g = \frac{1}{M} \mathbf{M}_s; \quad \mathbf{M}_s = \int \begin{Bmatrix} \tilde{u}^s \\ \tilde{v}^s \\ \tilde{w}^s \end{Bmatrix} dm; \quad \mathbf{J}_s = \int_M \tilde{\mathbf{r}} \times \tilde{\mathbf{u}}^s dm. \quad (11)$$

$\mathbf{I}_s(\bar{\mathcal{G}}^s)$  is the modal moment of mass inertia and represents the effect due to the elastic rotational deformation of the  $s$ -th elastic mode

$$\mathbf{I}_s(\bar{\mathcal{G}}^s) = \begin{pmatrix} \int_M \alpha_s (\tilde{y}^2 + \tilde{z}^2) dm & - \int_M \beta_s \tilde{x} \tilde{y} dm & - \int_M \chi_s \tilde{x} \tilde{z} dm \\ - \int_M \alpha_s \tilde{y} \tilde{x} dm & \int_M \beta_s (\tilde{z}^2 + \tilde{x}^2) dm & - \int_M \chi_s \tilde{y} \tilde{z} dm \\ - \int_M \alpha_s \tilde{z} \tilde{x} dm & - \int_M \beta_s \tilde{z} \tilde{y} dm & \int_M \chi_s (\tilde{x}^2 + \tilde{y}^2) dm \end{pmatrix} \quad (12)$$

### Boundary Element Hydroelasticity Method in the Frequency Domain

In a linearized frequency domain model, the external disturbance, the wave, is assumed “small” and responses induced by this small disturbance follow the time function  $e^{-i\omega_e t}$ . Where the encounter frequency  $\omega_e$  is a function of incident wave frequency  $\omega$ , ship speed  $U$ , and wave heading  $\chi$

$$\omega_e = \omega - Uk \cos \chi, \quad (13)$$

where wave number  $k = \omega^2 / g$ ,  $g$  is the gravitational acceleration, for deep water, and  $\chi = 0^\circ$  represents the following sea, the moving ship and propagating wave have the same direction, and  $\chi = 180^\circ$  represents the head sea condition.

Unsteady flow velocity potential is defined by

$$\Phi^U(\mathbf{r}; t) = \text{Re} \left\{ a e^{-i\omega_e t} \left[ -i \frac{g}{\omega} (\phi_I(\mathbf{r}) + \phi_D(\mathbf{r})) - i\omega_e \sum_{j=1}^L \frac{\xi_j}{a} \phi_j(\mathbf{r}) \right] \right\}, \quad (14)$$

where  $L$  is the number of system freedoms.  $L$  is 6 for the rigid-body model and 6 plus the number of involved elastic eigenmodes of the Hydroelasticity model.  $\eta_j(t) = \text{Re}(\xi_j e^{-i\omega_e t})$  is the displacement of mode  $j$  at time  $t$ , and  $\xi_j$  is the complex mode amplitude containing information for amplitude and phase.  $a$  is the amplitude of the incident wave and its space velocity potential for deep water is given by

$$\phi_I(x, y, z) = e^{kz} \exp\{ik[(\cos \chi)x + (\sin \chi)y]\}. \quad (15)$$

The first responsibility of the hydroelastic model is determination of the diffraction potential  $\phi_D(\mathbf{r})$  and radiation potential  $\phi_j(\mathbf{r})$  for each system freedom. Both diffraction and radiation potential satisfy the Laplace equation  $\nabla^2 \phi(\mathbf{r}) = 0$  and linearized free surface condition

$$\left[ g \frac{\partial}{\partial n} + \left( \frac{\partial}{\partial t} + \mathbf{W} \cdot \nabla \right)^2 \right] \phi = 0, \quad \text{on calm water surface } z = 0. \quad (16)$$

Additionally, the radiation condition requires the diffraction and radiation wave due to the existence of the ship propagating outward. The velocity potential in the fluid domain and on the boundary surface can be estimated by an integration of a singularity distribution on the wetted hull surface  $S_h$

$$\left\{ \frac{1}{\nabla_p} \right\} \phi(\mathbf{r}_p) = \iint_{S_h} \sigma(\mathbf{r}_q) \left\{ \frac{1}{\nabla_p} \right\} G(\mathbf{r}_p, \mathbf{r}_q) ds_q \quad (17)$$

$G(\mathbf{r}_p, \mathbf{r}_q)$  is the Green's function that satisfies the Laplace equation, free surface condition and radiation condition. The strength of singularity  $\sigma$  can be solved from the boundary integral equation

$$\frac{\partial \phi(\mathbf{r}_p)}{\partial n_p} = \alpha(\mathbf{r}_p) \sigma(\mathbf{r}_p) + \iint_{S_h} \sigma(\mathbf{r}_q) \frac{\partial G(\mathbf{r}_p, \mathbf{r}_q)}{\partial n_p} ds_q. \quad (18)$$

And  $\alpha(\mathbf{r}_p)$  is the interior solid angle of field point  $\mathbf{r}_p$  on the wetted hull surface  $S_h$ . The required surface condition for the diffraction problem is

$$\frac{\partial \phi_D}{\partial n} = -\frac{\partial \phi_I}{\partial n} \quad (19)$$

and for the radiation problem of  $j$ -th motion/elastic mode

$$\frac{\partial \phi_j}{\partial n} = n_j + \frac{1}{i\omega_e} m_j \quad (20)$$

Modal normal component  $n_j$ , the so-called n-term, and m-term can be estimated by

$$\left. \begin{aligned} n_j &= \mathbf{u}^j \cdot \mathbf{n} \\ m_j &= \left[ \bar{\mathbf{g}}^j \times \mathbf{W} - (\mathbf{u}^j \cdot \nabla) \mathbf{W} \right] \cdot \mathbf{n} \end{aligned} \right\} \quad (21)$$

Hydrodynamic pressure on wetted hull surface comes from Benoulli's equation using velocity potential and its gradient. Its linearized form is of

$$p(\mathbf{r};t) - p_0 = -\rho \left( \frac{\partial \Phi^U}{\partial t} + \mathbf{W} \cdot \nabla \Phi^U \right) \quad (22)$$

The hydrodynamic force on mode  $i$  can be computed by integration

$$p_i = -\iint_{S_h} n_i [p(\mathbf{r};t) - p_0] ds \quad (23)$$

After an order analysis of the perturbation expansion of this theoretical approach, this surface modal force can be expressed in terms of modal amplitude

$$p_i = \text{Re} \left\{ e^{-i\omega_e t} \left[ E_i - (i\omega_e)^2 \sum_{j=1}^L \xi_j \left( A_{ij} + \frac{1}{i\omega_e} B_{ij} - \frac{1}{\omega_e^2} C_{ij} \right) \right] \right\} \quad (24)$$

and the modal wave exciting force  $E_i$ , modal wave making added-mass and damping coefficient  $A_{ij}$  and  $B_{ij}$  is computed as follows

$$E_i = \rho g a \iint_{S_h} n_i \left[ \frac{\omega_e}{\omega} (\phi_I + \phi_D) + \frac{1}{i\omega} \mathbf{W} \cdot \nabla (\phi_I + \phi_D) \right] ds \quad (25)$$

$$A_{ij} + \frac{B_{ij}}{i\omega_e} = \rho \iint_{S_h} n_i \left[ \phi_j + \frac{1}{i\omega_e} \mathbf{W} \cdot \nabla \phi_j \right] ds \quad (26)$$

The modal restoring coefficient  $C_{ij}$  has a similar, but more lengthy, integration formula.

The modal amplitude for a model without concentrated force can then be solved from

$$\sum_{j=1}^L \left[ (i\omega_e)^2 (m_{ij} + A_{ij}) + i\omega_e (b_{ij} + B_{ij} + B_{ij}^V) + (k_{ij} + C_{ij}) \right] \xi_j = E_i, \quad (27)$$

including the linearized equation of rigid body motion

$$\begin{aligned} M \frac{d^2 \bar{\eta}}{dt^2} + \sum_s \frac{d^2 \eta_s}{dt^2} \mathbf{M}_s &= \mathbf{F} \\ \mathbf{J} \left( \frac{d\mathbf{w}}{dt} \right) + \sum_s \frac{d^2 \eta_s}{dt^2} \left[ \mathbf{J}_s + \mathbf{I}_s (\bar{g}^s) \right] &= \mathbf{M} \end{aligned} \quad (28)$$

These formulae are expressed in a hydrodynamic reference coordinate system. This system is on the calm water surface and moving at the constant speed  $U$  toward the ship moving direction. Therefore it is an inertial coordinate system and the dynamic mass modal force is nil, as the existing mass force is gravity and it is a constant in time in this coordinate system. In this deformation/motion equation,  $B_{ij}^V$  is the coefficient of viscous flow induced damping which is important to those modes having small wave making damping, such as roll motion mode. It is possible to involve nonlinear viscous flow damping in the analysis.  $b_{ij}$  is the coefficient of structural modal damping, which is still a challenge for structural engineers to estimate reliably. So far, this coefficient is mainly determined from model tests and/or sea trials.

In most of the available tools, the structural analysis is performed in a ship fixed coordinate system that leads to nonzero mass modal force with two components, one being induced by the inertial acceleration of the rigid body ship motion, and the other coming from the dynamic gravitational acceleration in this ship fixed coordinate system. This mass modal force will excite every non-wettable mode and therefore those modes may need to be involved in structural assessment.

### Boundary Element Hydroelasticity Method in Time Domain

As described in the previous section, the frequency domain hydroelastic BEM model combines linear structural FEA, linear hydrodynamic BEM with linear or nonlinear deformation/motion equations. It is basically a linear Hydroelasticity approach. This frequency domain approach works well for a ship or offshore structure in low and moderate sea conditions, but it becomes unreliable for large wave cases and a nonlinear model is required. For ship or offshore structures in large waves, the dominant nonlinear factors in hydrodynamic problems are the vertical shape change of the hull surface, i.e. the hull flare, and high wave itself. A cheaper and more efficient time domain hydroelastic BEM model could be considered instead of the ultimate CFD model. The time domain BEM model simulates responses in waves by using a retardation function, the response function to an impulse disturbance and carries the memory effects of disturbances in the past. The retardation function in this time domain BEM model can be determined either by time domain Green's function or by the Fourier transformation of hydrodynamic results obtained from frequency domain analysis. Compared to the first approach, the second approach usually takes less computing time and requires much less computer memory and it is the method we will use in this paper. The important difference between the frequency domain hydroelastic BEM model and the time domain model is that the small wave restriction is removed in the time domain model, and it leads to large responses



like structural motion/deformation, internal load and so on. The nonlinear terms need to be involved in the computation of motion/deformation equations and internal loads.

One of the main differences from the linearized frequency domain model is that the rigid-body motion mode can have large amplitude, and equation (9) and (10) are adopted. Also in the time domain model, the modal surface force, i.e. the hydrodynamic force, will be determined by

$$F_i = F_i^{rad} + F_i^{dif} + F_i^{FK} + F_i^{rst} + F_i^{others}, \quad i = 1, 2, \dots, L \quad (29)$$

where subscript  $i$  stands for deformation/motion mode;  $F_i^{rad}$  is the radiation force on mode  $i$ ;  $F_i^{dif}$  is the diffraction force;  $F_i^{FK}$  is the so-called Froude-Krylov force due to incident waves;  $F_i^{rst}$  is the restoring force due to hydrostatic pressure; and  $F_i^{others}$  represents the force due to all other external effects like mooring, viscous damping, tank sloshing, maneuvering, etc. In contrast with the frequency domain model, the maneuvering force on rigid-body motion mode for a ship with forward speed is a mandatory factor in the time domain simulation. Similarly the mooring force or position stationary force for a floating offshore platform is also important.

The radiation force on mode  $i$  that is induced by the motion/deformation of the structure is determined by the convolution

$$F_i^{rad}(t) = -\sum_{j=1}^L \left[ A_{ij}^{\infty} \ddot{\eta}_j(t) + B_{ij}^{\infty} \dot{\eta}_j(t) + C_{ij}^{\infty} \eta_j(t) + \int_0^t \kappa_{ij}(t-\tau) \dot{\eta}_j(\tau) d\tau \right], \quad i = 1, 2, \dots, L. \quad (30)$$

$\eta_j(t)$  is the displacement of mode  $j$  at time  $t$ ,  $A_{ij}^{\infty}$ ,  $B_{ij}^{\infty}$  and  $C_{ij}^{\infty}$  is the coefficient of wave making added-mass, damping and restoring force at infinite encounter frequency.  $\kappa_{ij}(t)$  is the retardation function and can be estimated from

$$\kappa_{ij}(\tau) = \frac{2}{\pi} \int_0^{\infty} [A_{ij}^{\infty} - A_{ij}(\omega_e)] \frac{\sin(\omega_e \tau)}{\omega_e} d\omega_e \quad \text{or} \quad \kappa_{ij}(\tau) = \frac{2}{\pi} \int_0^{\infty} [B_{ij}(\omega_e) - B_{ij}^{\infty}] \cos(\omega_e \tau) d\omega_e \quad (31)$$

The diffraction modal force can be expressed in a similar way

$$F_i^{dif} = \int_{-\infty}^{\infty} \kappa_i^D(t-\tau) \zeta_0(\tau) d\tau, \quad (32)$$

and

$$\kappa_i^D(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [\text{Re}\{E_i^D(\omega_e)\} \cos(\omega_e t) + \text{Im}\{E_i^D(\omega_e)\} \sin(\omega_e t)] d\omega_e \quad (33)$$

Accuracy of the retardation function for radiation and diffraction modal force will directly affect the analysis results and development of a reliable algorithm for the infinite integration is one of the challenges in this time domain model.

In time domain hydroelastic BEM model, nonlinear rigid body motion equations, (9) and (10), are applied. The nonlinear terms in modal force computation needs to be consistently involved as well.

## Examples of Application

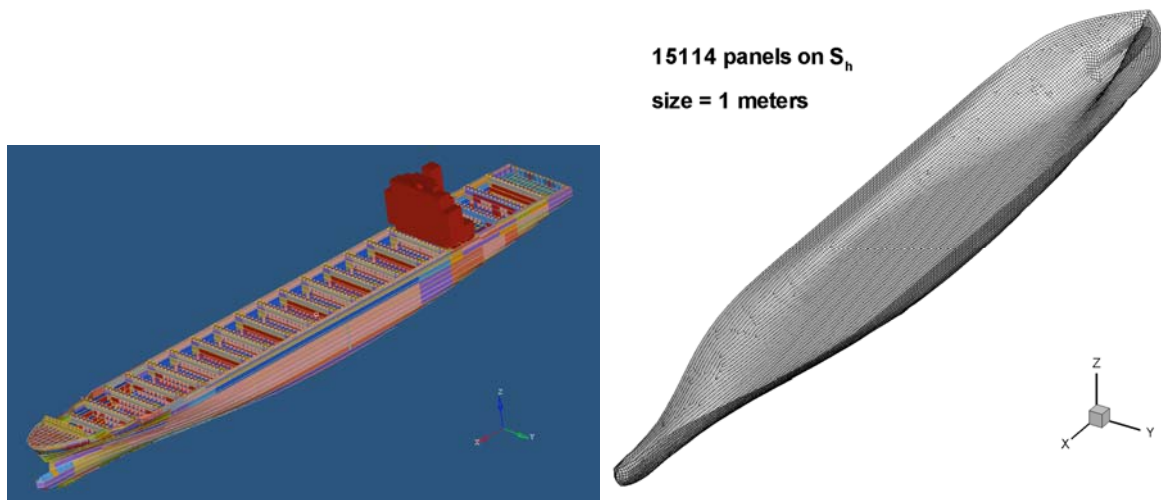
Two hydroelasticity codes have been developed in Lloyd's Register, HydroE-FD for the frequency domain model and HydroE-TD for the time domain model. The results of frequency domain hydroelastic analysis are compared against the results of Lloyd's Register's rigid body frequency domain code WAVELOAD-FD. The structural analysis was performed by using Lloyd's Register's FEA code, Trident.

### Example 1 - Linear frequency domain hydroelasticity

A container ship was selected as an analysis example. Particulars of this vessel, including ship dimensions, draft at fore and aft perpendicular, displacement, COG and radii of gyrations, are shown in Table 1. The full ship global FEA model and hydroelastic panel model are shown in Figure 1. A one meter sized panel model was selected to capture the responses in the high wave frequency range. Figure 2 and 3 plot the shape of first 10 wettable elastic dry eigenmodes and their  $n$ -term. The  $n$ -term of rigid motion mode roll and pitch are also presented in Figure 3. It can be observed that modes 9, 21 and 28 are the first three vertical bending modes, all the other modes are horizontal modes representing bending, torque or their combinations.

**Table 1: Particulars of container ship model**

Lbp ( $m$ )	325.0
B ( $m$ )	43.8
T_fp ( $m$ )	11.075
T_ap ( $m$ )	11.405
Displacement ( $m^3$ )	94428
Wetted hull area ( $m^2$ )	14910
LOG ( $m$ )	154.5
VOG ( $m$ )	15.342
Kxx ( $m$ )	16.088
Kyy ( $m$ )	78.235
Kzz ( $m$ )	78.343
Kxz ( $m$ )	6.112



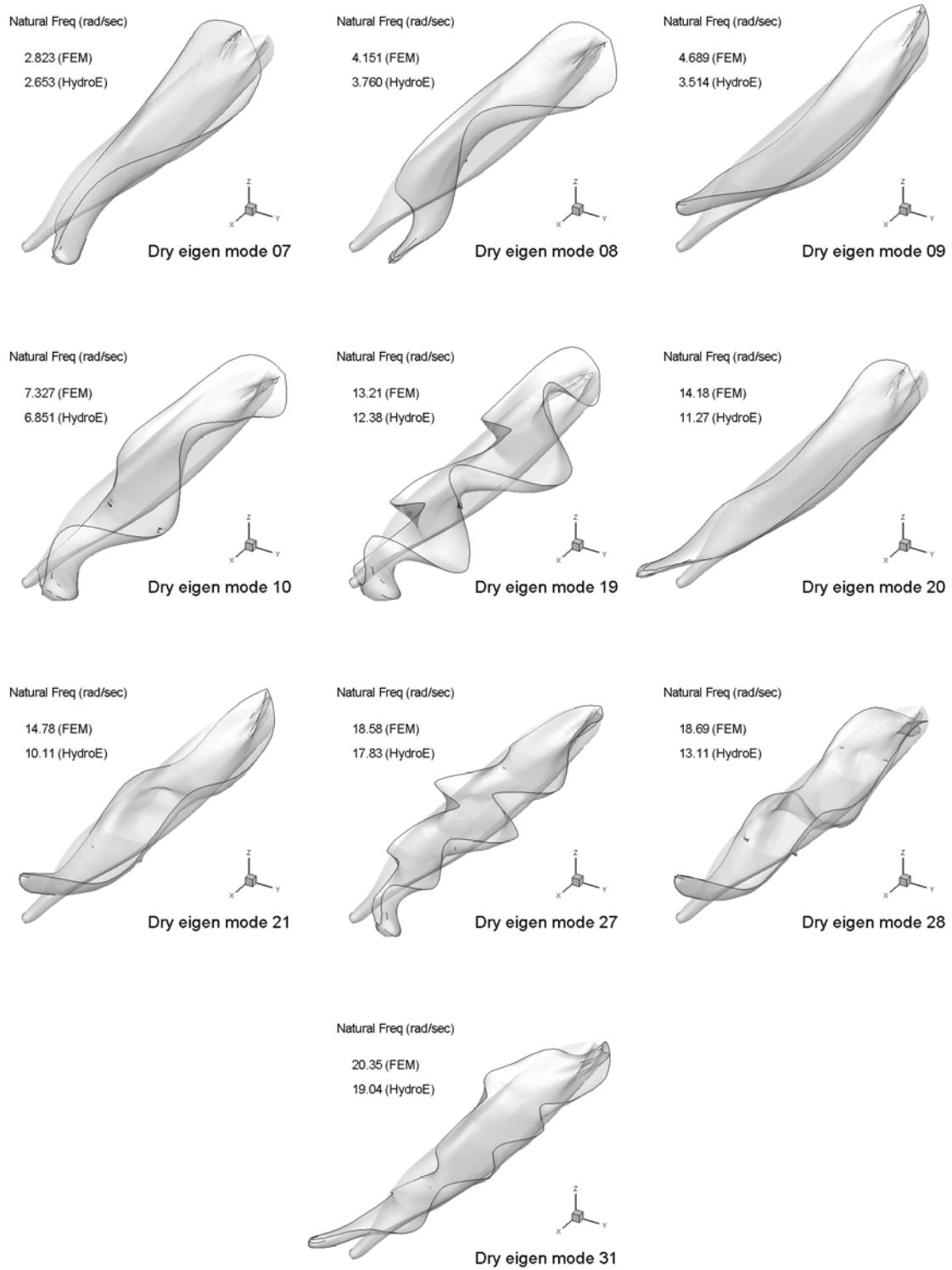
**Figure 1: FEA model (left) and hydroelastic panel model (right) of the container vessel.**

Figure 4 shows the first key result, being the comparison of the vessel deformation due to the hydrostatic pressure and gravitational force when the vessel is floating in calm water. The upper plot is the result by a direct static 3D FEM analysis and the lower one is the result by HydroE-FD using the modal pressure force due to the hydrostatic pressure and modal mass force due the gravitation. Static modal amplitudes from HydroE-FD are given in Table 2. The main contribution to the static deformation was from the three vertical bending modes as the model is very close to a symmetric case. The ship is in a static hogging state and maximum static deformation by direct FEA using Trident was 324 mm while that from HydroE-FD was 323 mm, so they correlate very well.

**Table 2: Static mode amplitude by Hydro-FD**

Mode	07	08	09	10	19	20	21	27	28	31
$\xi_j^S$	7.975	11.30	-101.0	0.2505	-0.1286	1.077	-39.87	-0.689	25.46	-0.03120

The natural frequency of each elastic eigenmode is one of most important results in a structural assessment and the results of the “dry natural frequency” and “wave making natural frequency” are listed in Table 3. The “dry natural frequency” represents the natural frequency when the structure oscillates in air or “in vacuum” and these are the eigen values calculated by the whole ship FEA. The “wave making natural frequency” is the natural frequency when the structure oscillates in water and generates the so-called radiation waves. The restoring force and encounter frequency dependent wave making added-mass are considered together with the modal mass and stiffness. From the results, we can see that the wave making effect always decreases the natural frequency and it can even change the sequence of some eigenmodes. In this example, the first vertical bending mode is the third elastic eigenmode (09) in the original dry eigenmode list, but it jumps up to the second elastic mode in the wave making list due to the significant increase of mass due to the wave making added-mass associated with this modal shape. For ocean waves, the typical wave period is on the order of 10 seconds and in general waves of this period do not directly excite resonant oscillation for those eigenmodes with a wave making frequency higher than 5 *rad/sec*. Waves with higher frequency (> 2 *rad/sec*) can excite eigenmodes of natural frequency larger than 5 *rad/sec*. But those shorter waves have smaller amplitude and are usually ignored in many analyses.



**Figure 2: Wetted surface of the first 10 elastic mode of the container ship model.**

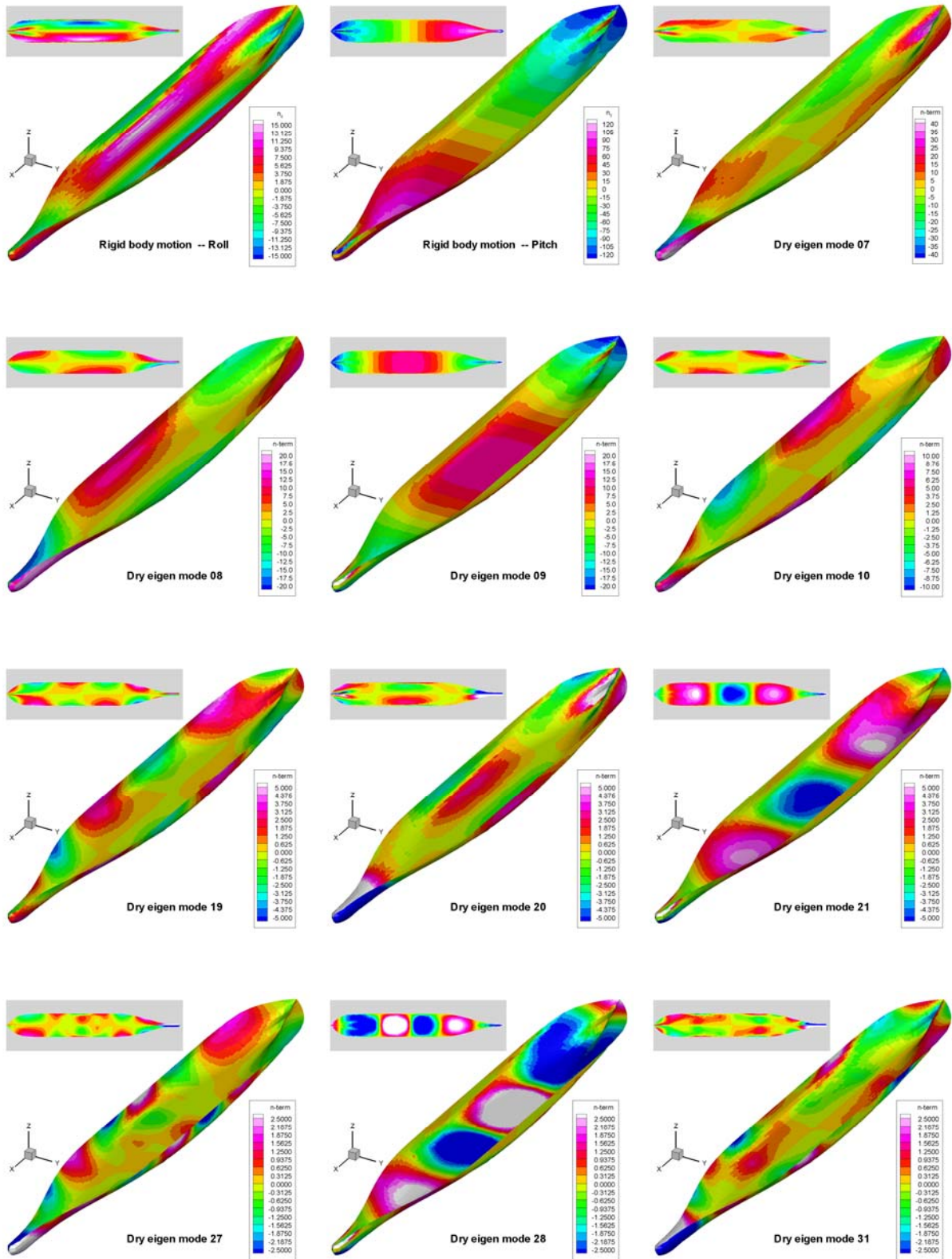
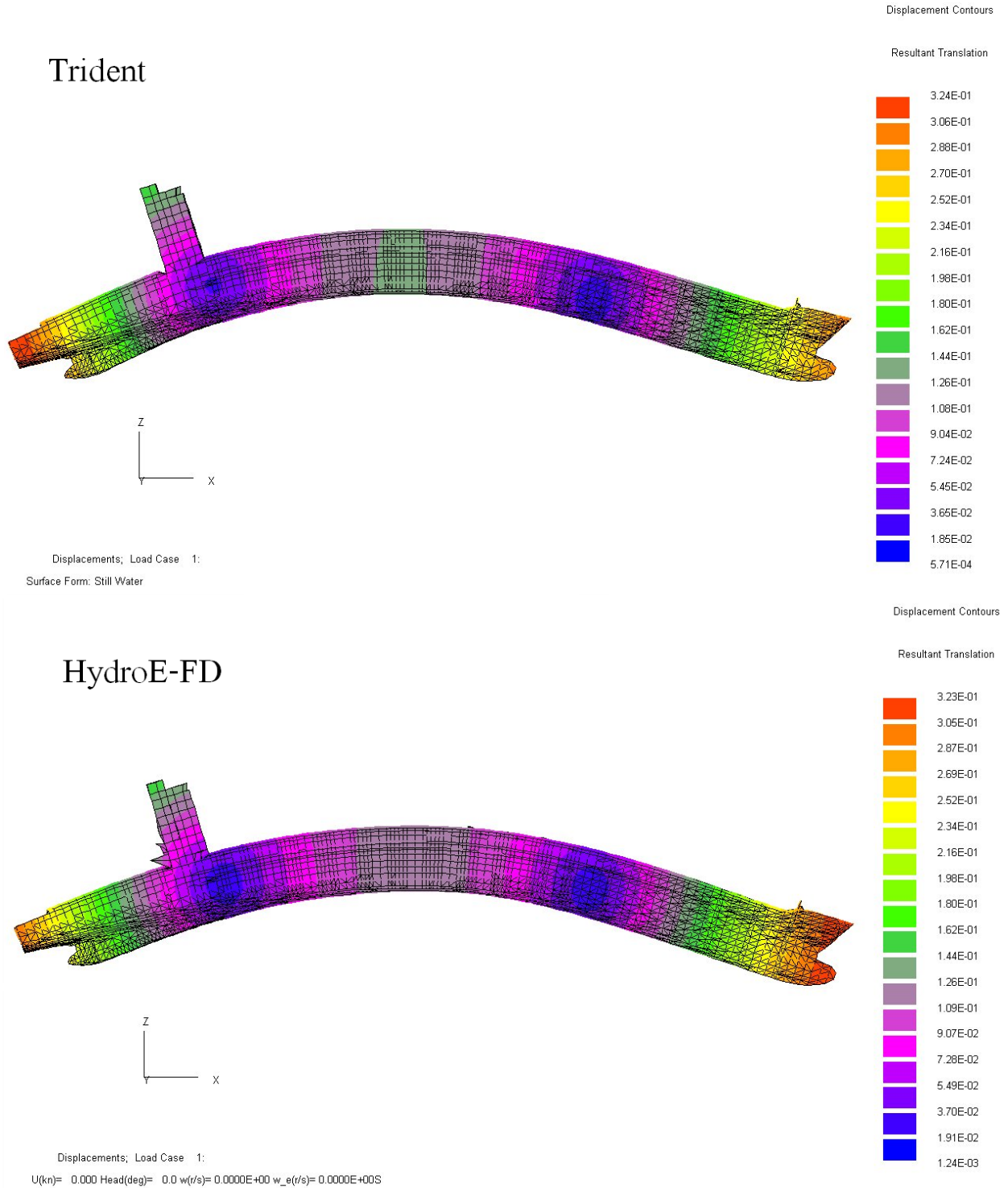


Figure 3: The n-term of roll, pitch, and the first 10 elastic dry eigenmodes.



**Figure 4: Static displacement of the model by Trident (upper) and HydroE-FD (lower).**

**Table 3: Natural frequency of the container ship**

Eigen Mode	Dry natural frequency from FEA ( <i>rad/sec</i> )	Wave making natural frequency from HydroE-FD ( <i>rad/sec</i> )
07	2.823	2.655
08	4.151	3.754
09	4.689	<a href="#">3.501</a>
10	7.327	6.851
19	13.21	12.38
20	14.18	<a href="#">11.27</a>
21	14.78	<a href="#">10.11</a>
27	18.58	17.83
28	18.69	<a href="#">13.11</a>
31	20.35	19.04

**Table 4: Rayleigh coefficients of structural modal damping for S8100R model**

Mode	07	08	09	10	19	20	21	27	28	31
%	<b>1.27</b>	<b>3.29</b>	<b>2.28</b>	<b>1.15</b>	<b>0.67</b>	<b>14.3</b>	<b>1.13</b>	<b>0.0</b>	<b>2.82</b>	<b>0.0</b>

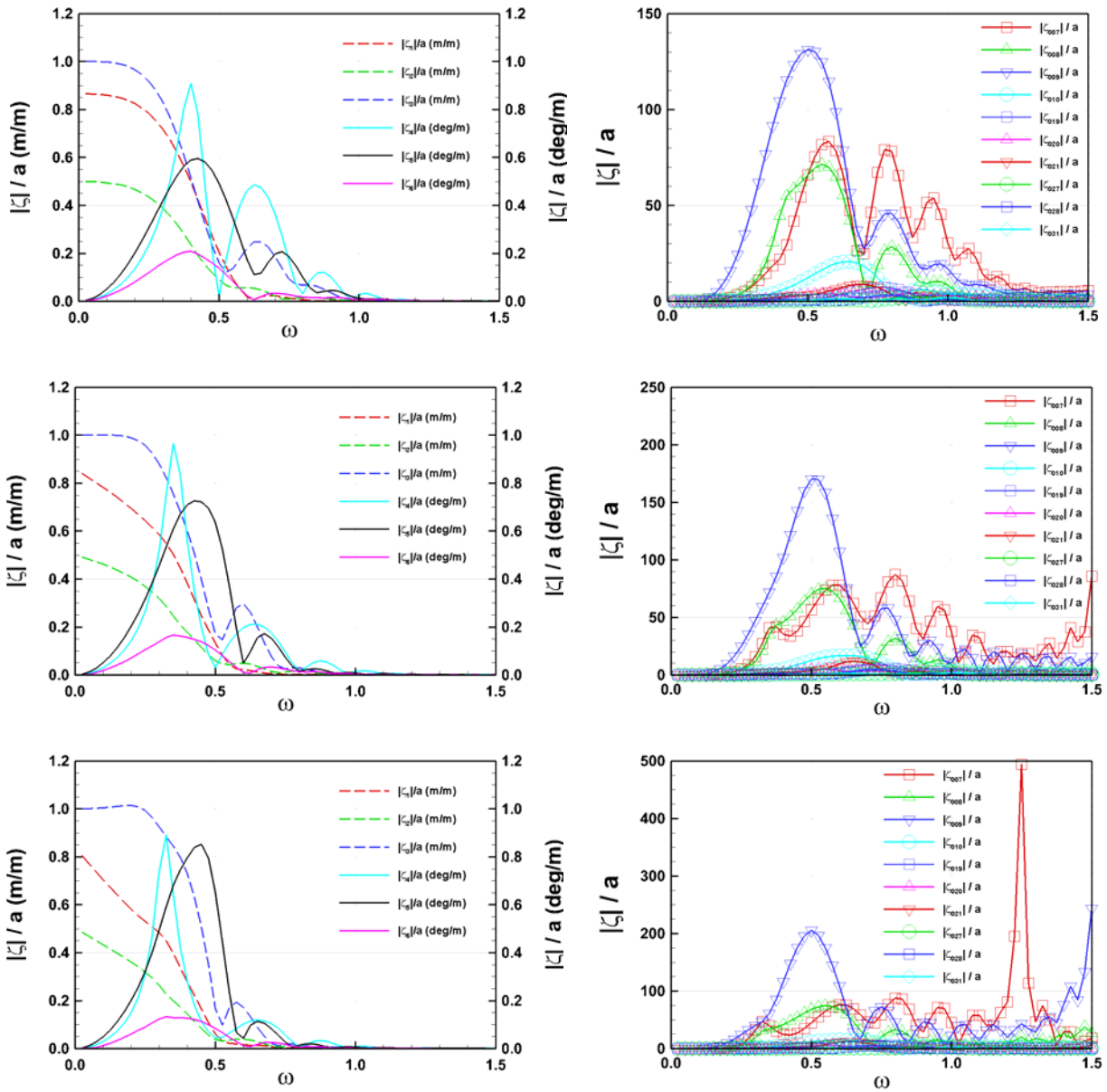
Structural modal damping applied in this example is listed in Table 4.

RAO curves (Response Amplitude Operator curves which represent the response amplitude induced by a wave of 1 meters amplitude) of rigid-body motion (modes 1 to 6) and elastic deformation (modes 7 to N) over a range of incident wave frequency 0.0 to 1.2 rad/sec and 150 degree heading at three ship speeds are plotted in Figure 5. The modal amplitude of the elastic eigenmodes over a wave frequency range 0 to 5 rad/sec at the same ship speed and heading are shown in Figure 6, and the resonant responses of modes 7, 8 and 9 can be found.

Distribution of hydrodynamic pressure RAO is shown in Figure 7 in a resonant condition of the 3<sup>rd</sup> elastic eigenmode, the first vertical bending mode. Compared to the rigid-body analysis results, hydroelastic analysis received much higher pressure in this case. Note the n-term pattern shown in Figure 3, from this we can find that the pattern of the pressure distribution is similar to that of the n-term of the resonant eigenmode, and the radiation pressure is the dominant component in this case.

The modal resonant is determined by two factors, frequency and strength of external excitation. For the case of a ship, an elastic structural eigenmode can be excited when the incident wave has an encounter frequency close to its modal natural frequency. The scale of this resonant eigen response is determined strongly by the pattern of external exciting pressure, including pressure of incident waves and diffraction waves as well as the radiation waves of other modes. The third elastic eigenmode, the 1<sup>st</sup> vertical bending mode, has a resonant amplitude around 6.5 for the zero speed case at a wave frequency of 3.5 rad/sec where the incident wave length is very short, around 5 meters; see Figure 6, top graph. When the wave length increases to 28 meters in 20 the knots case, a

wave frequency of 1.5 rad/sec, the amplitude of this eigenmode jumps up to a level of 50 times larger than that of zero ship speed.

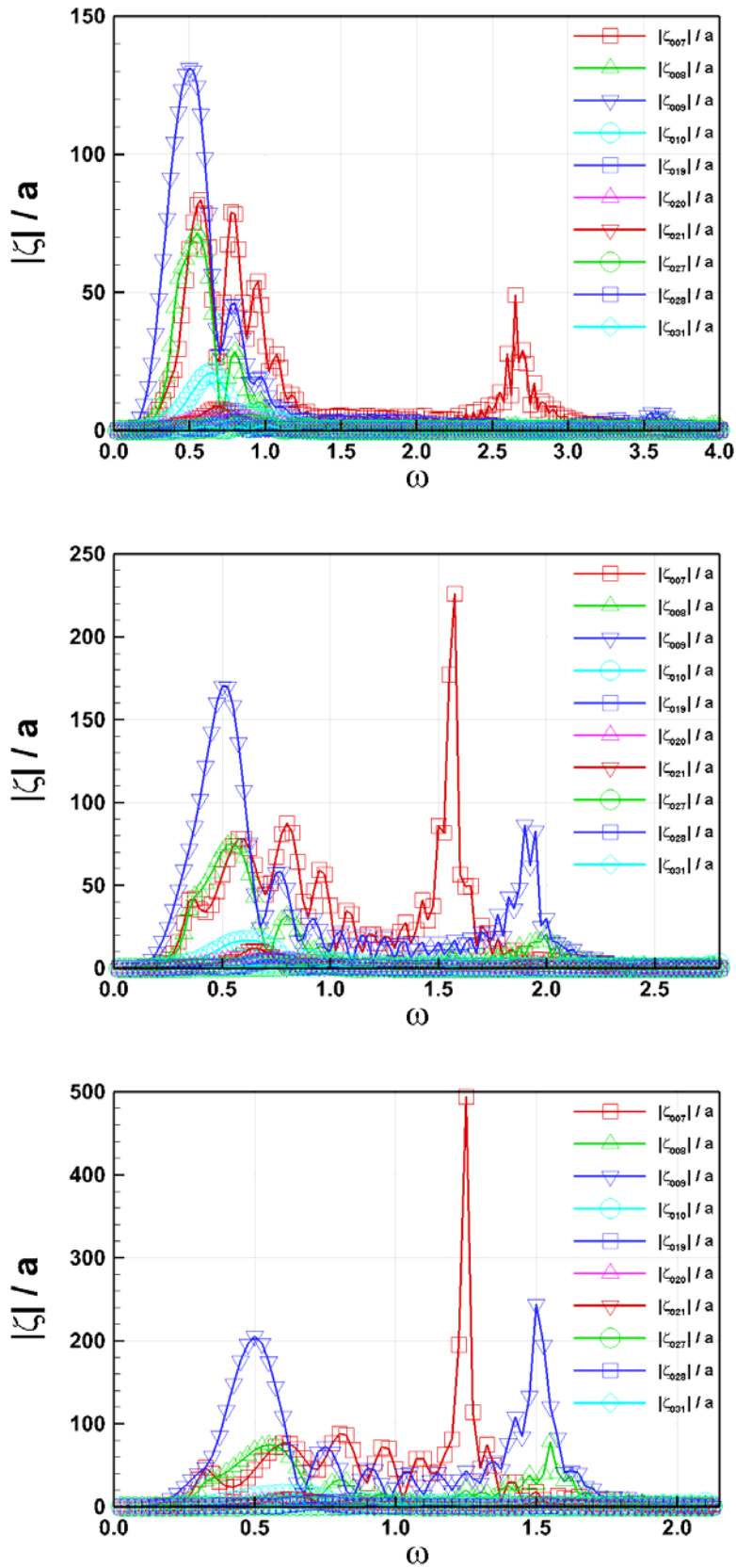


**Figure 5: RAO of rigid-body motion and elastic deformation in waves of  $\chi = 150^\circ$**

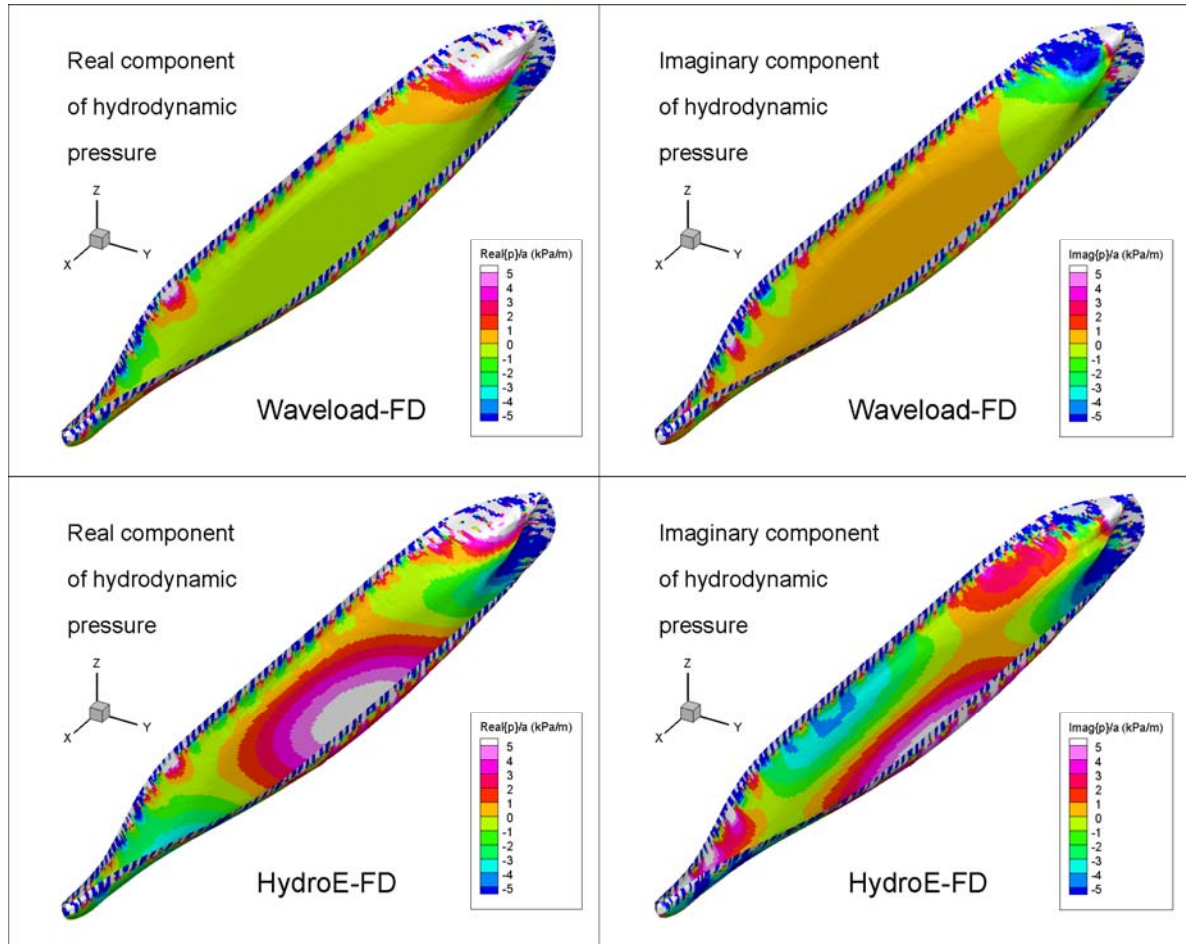
**1<sup>st</sup> row: U= 0 knots; 2<sup>nd</sup> row: U= 10 knots, 3<sup>rd</sup> row: U= 20 knots**

**1<sup>st</sup> column: rigid body motion modes; 2<sup>nd</sup> column: elastic deformation modes.**





**Figure 6: RAO of elastic deformation in waves of  $\chi = 150^\circ$**   
**1<sup>st</sup> column: U=0 knots; 2<sup>nd</sup> column: U=10 knots; 3<sup>rd</sup> column: U=20 knots**



**Figure 7: RAO of hydrodynamic pressure of rigid body model (upper) and hydroelasticity model (lower) for condition  $U = 12$  knots,  $\chi = 150^\circ$  and  $\omega = 1.875$  rad/s.**

The internal load values, shear forces, bending moments and torsional moments, are obtained by summation the hydrodynamic pressure and ship motion inertial loads. To check the effects of elastic deformation, internal load results of tension force, horizontal shear force and vertical shear force on a section at 25% of ship length from aft perpendicular are shown on the left column of Figure 8. The torque, vertical bending moment and horizontal bending moment on the mid-ship section are shown in the right column of Figure 8. Blue lines show the results of the frequency domain hydroelasticity model and the results from a rigid-body hydrodynamic analysis are shown as red diamond marks. From this figure, we can see that the internal loads calculated by a hydroelastic analysis method will be close to the results from a rigid-body analysis method when the ship is in waves with encounter frequencies well away from its structural natural frequencies. When the ship is in waves with encounter frequencies close to the structural natural frequencies, then significant feedback can be expected from the internal loads and those resonant internal loads will induce structural vibration, so-called springing, which ultimately may lead to early onset of damage to the ship structure. For certain types of ship or floating structures, the fatigue life on some structural components based on hydroelasticity assessment can be more than 50% shorter than that by a rigid-body based analysis. A hydroelastic assessment will definitely be required for such cases.

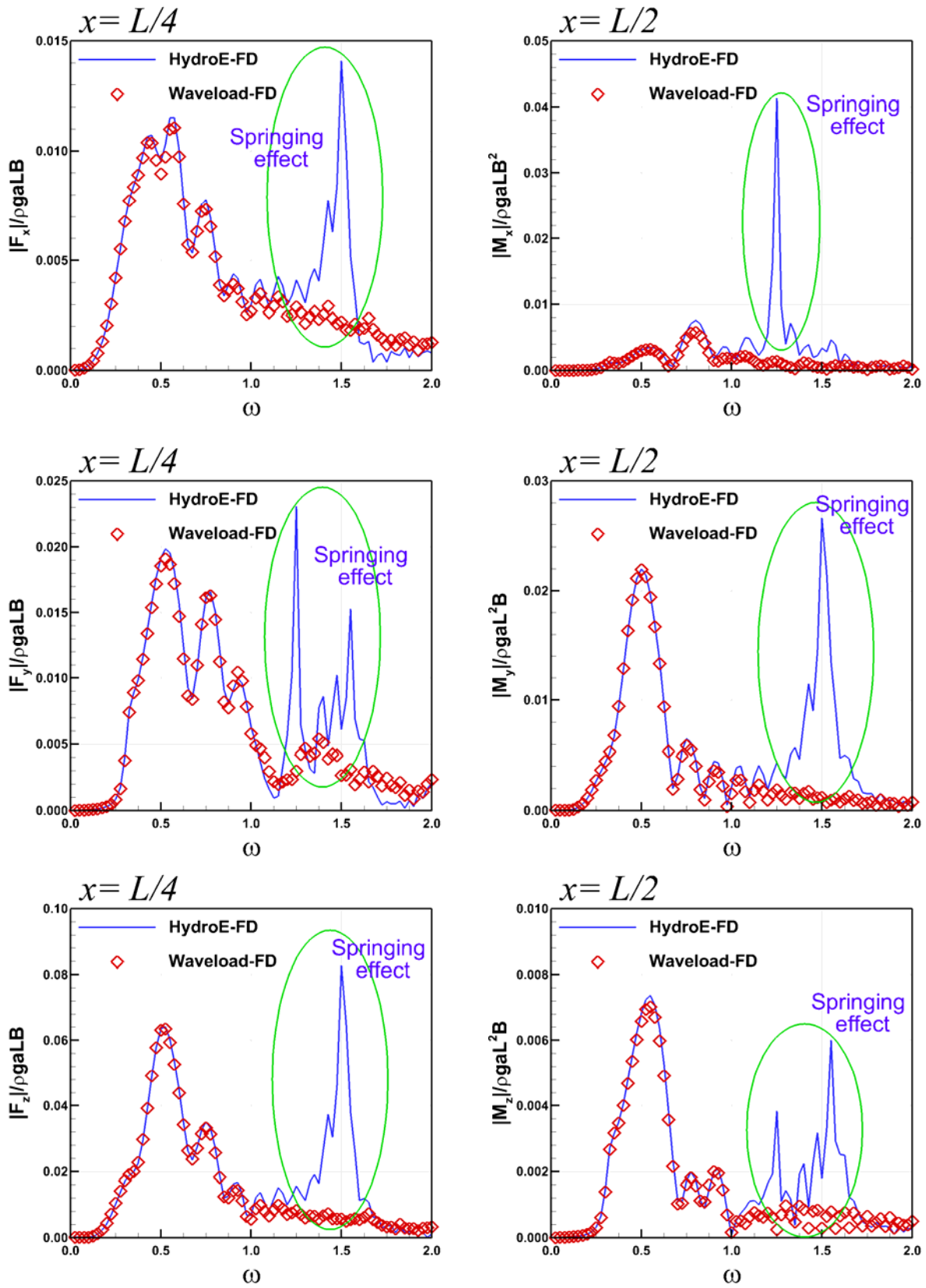
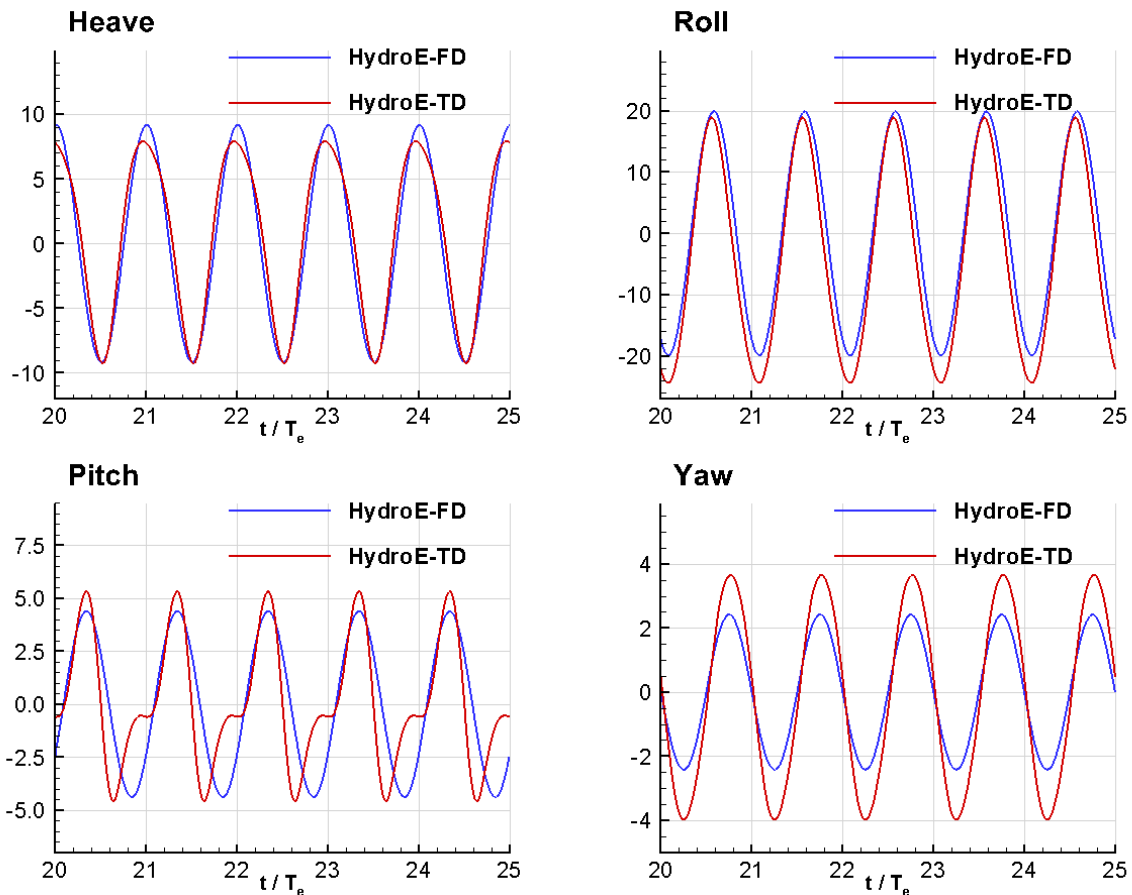


Figure 8: Non-dimensional RAO of internal load in condition of  $U=20$  knots and  $\chi = 150^\circ$ .

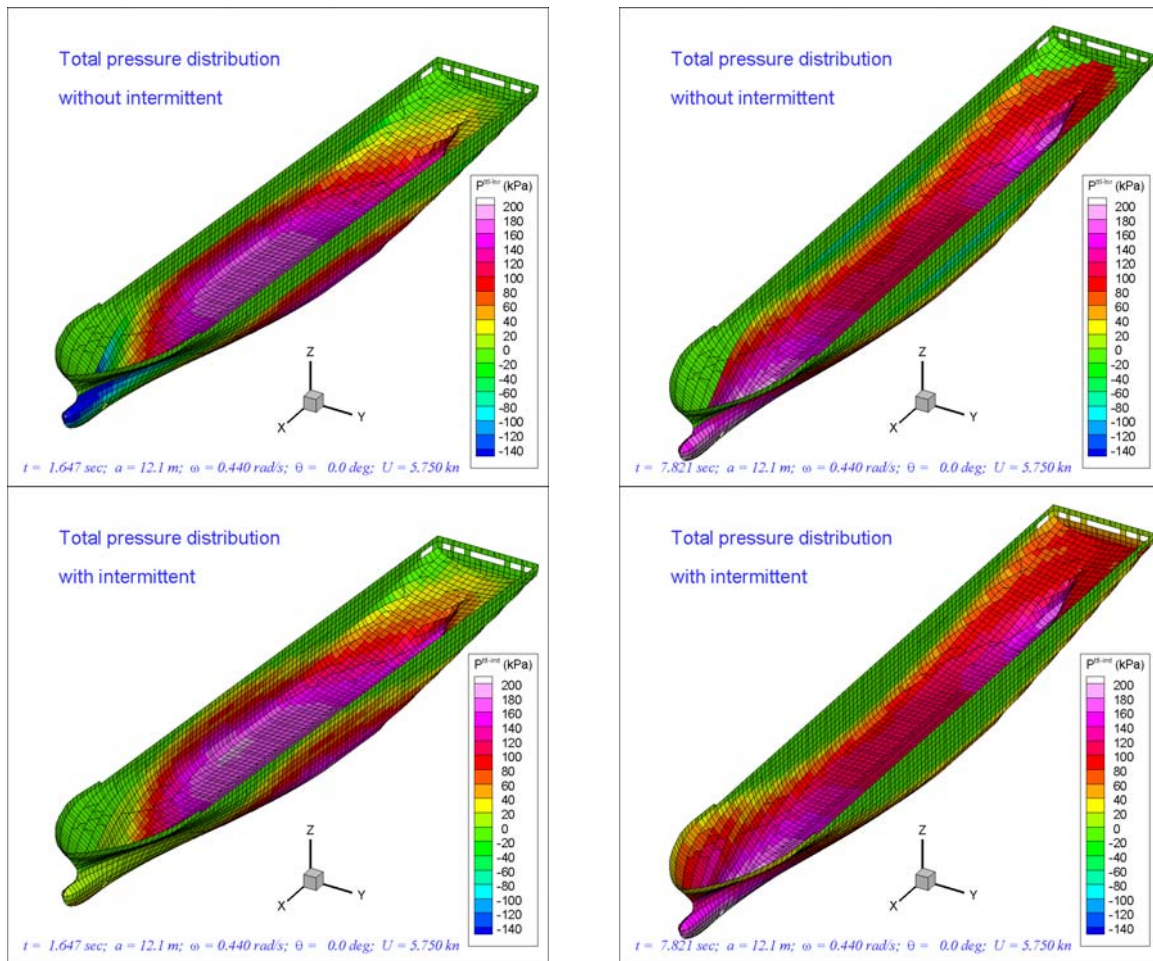
### Example 2 - Nonlinear time domain hydroelasticity

From analyses of frequency domain hydroelastic results, one can see the reasons for considering hydroelasticity and what impact a hydroelasticity analysis can have on an engineering assessment. We should bear in mind that the frequency domain hydroelasticity is based on a “linear response” condition and works well for cases of waves with small wave steepness. A time domain hydroelasticity model needs to be considered for large wave cases.

We present some results of time domain hydroelasticity analysis for the same ship model below. Time trace of four rigid-body motion modes is shown in Figure 10, for a case in waves of 120 degree heading, 0.375 rad/sec of frequency, 20 meter wave height and ship speed of 10 knots. We can see the time domain model predicts larger motion and nonlinear characteristics for pitch in this extremely high wave condition. The differences in results from the frequency domain analysis and time domain analysis are caused mainly by the so-called geometry nonlinear problem, i.e. the effect of “flare bow” and “flat stern” for the present model above the mean waterline and the lack of buoyancy at both ends below the mean waterline. The frequency domain model and time domain will have similar responses if the wettable surface of the structure has a purely vertical shape, but this is not the case for most ships. Another reason for this difference is the large nonlinear rotation motions. For example, the roll motion amplitude of the present case goes up to 20 degrees, and this roll motion will affect the yaw motion. The effects of roll on yaw in the nonlinear rotation term  $\mathbf{w} \times \mathbf{J}(\mathbf{w})$  and mapping matrix  $\mathbf{R}$  are ignored in a frequency domain model therefore coupling between rotation motions are weaker there.



**Figure 9:** Amplitude of some rigid-body motion modes in condition  $U=10$  knots,  $\chi = 120^\circ$ ,  $\omega = 0.375$  rad/sec and wave height = 20 meters.

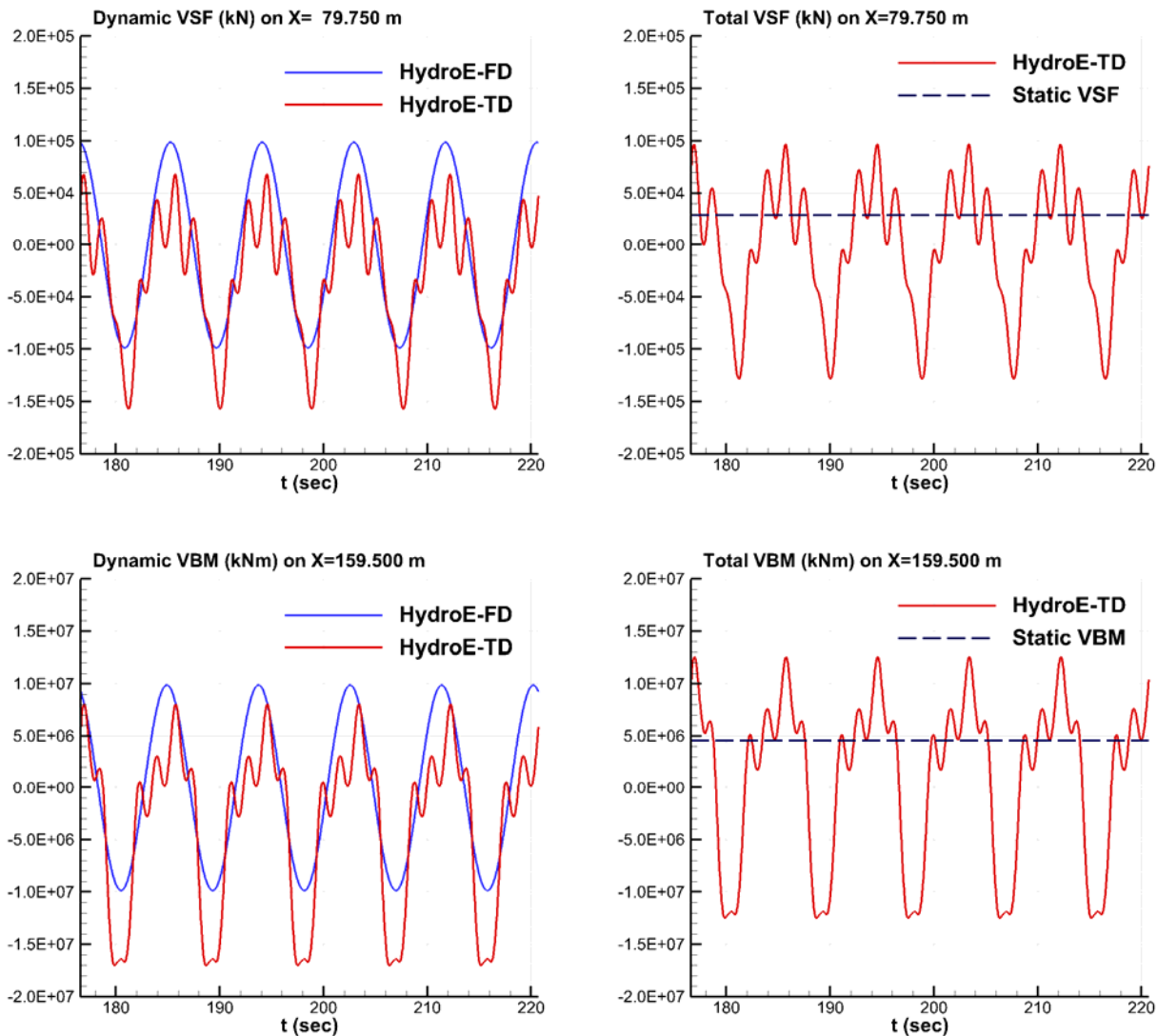


**Figure 10: Pressure distribution by frequency domain model (top) and time domain model (bottom) in condition  $U=5.75$  knots,  $\chi = 0^\circ$ ,  $\omega = 0.44$  rad/sec and wave height = 24.2 meters, in a hogging state (left) and a sagging state (right)**

The most obvious difference between the frequency domain and the time domain approaches is the pressure distribution, as illustrated in Figure 10. The hydrodynamic pressure only acts on the mean wetted surface of the ship in the frequency domain mode and changes in a pure sinusoidal style. From the top row of Figure 10, labelled "Total pressure distribution without intermittent," showing the linear frequency domain results with the addition of the hydrostatic pressure, we can see that there is no pressure on the mean dry hull surface and the total pressure on the bow can have a negative value in a frequency domain model. On the other hand, the total combined hydrostatic and hydrodynamic pressure calculated by a time domain model has no such problem and the pressure distribution looks much closer to reality, as shown in the bottom row of Figure 10 labelled "Total pressure distribution with intermittent".

A problem in the frequency domain model is the "symmetric vertical load". The magnitude of the dynamic vertical bending moment and shear force will be the same for both the hogging and sagging conditions, but this is clearly not the case due to "intermittent" pressure effects which can become significant for higher waves. For a conventional ship, the dynamic vertical load in a sagging wave condition is normally always larger than in a hogging wave condition. In a time domain model we simulate the instantaneous wetted surface and solve the pressure on this surface, and as a consequence the vertical loads become much closer to reality. The dynamic and total

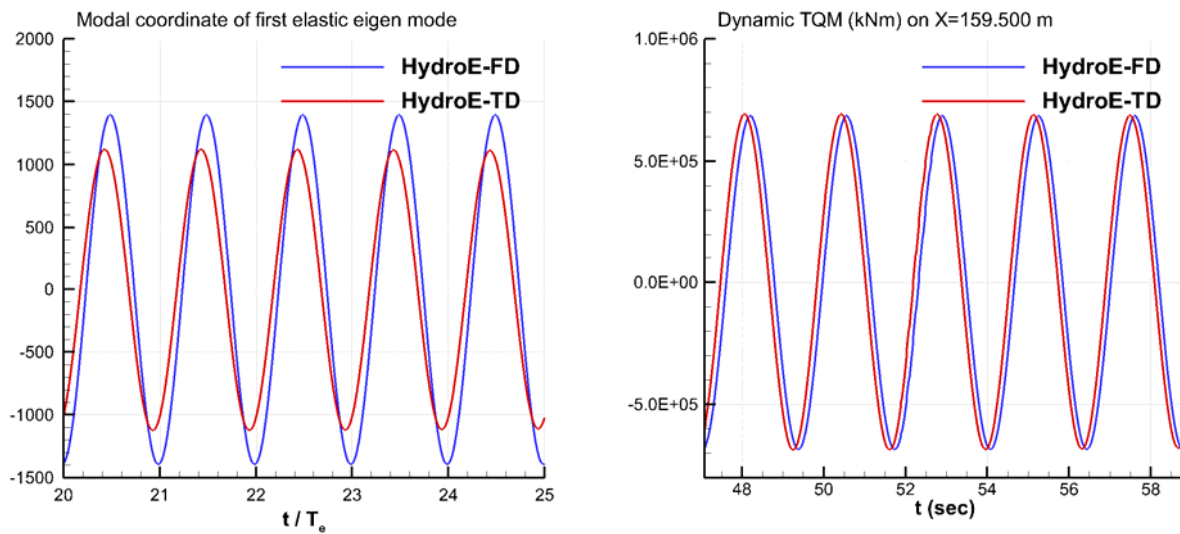
vertical shear force on a section at  $\frac{1}{4}$  of ship length are presented in the top row of Figure 11, and those for the vertical bending moment on a mid-ship section are presented in the bottom row. This shows that for this example, the dynamic vertical bending moment in the sagging wave case is about twice of that of the hogging wave case. The vertical shear force has similar tendency. It is worth noting that the results of this extreme wave case are for demonstration of the difference of the frequency and time domain models, and we would not expect the ship to be operated at 20 knots in a 20 meters wave environment. The asymmetric ratio of vertical load in sagging/hogging condition will decrease as the wave height becomes smaller.



**Figure 11: Vertical shear force and bending moment in condition  $U=20$  knots,  $\chi = 180^\circ$ ,  $\omega = 0.475$  rad/sec and wave height = 20 meters**

The last result presented is the time trace of the first elastic eigenmode, which is an almost pure torque mode for the present ship model, shown in the left of Figure 12, and the time trace of torque load on the mid-ship section. The amplitude of the first elastic mode predicted by the time domain model is smaller than that predicted by the frequency domain, but the torque load predicted by both

methods are almost the same due to the effect of the intermittent and nonlinear pressure correction in the time domain approach.



**Figure 12: Time trace of first elastic eigenmode amplitude and dynamic torque at mid-ship in condition  $U=20$  knots,  $\chi = 150^\circ$ ,  $\omega = 1.250$  rad/sec (resonant frequency of the 1<sup>st</sup> (torque only) mode) and wave height = 2.82 meters**

## Conclusions

The theory of hydroelasticity is outlined for both the linear frequency domain approach and non-linear time domain approach. Results of a container ship of 320 meters in length are used to demonstrate the improvement for hydrodynamic analysis going from rigid-body to elastic body and from frequency domain to time domain. Those results also prove the importance of considering the application of hydroelastic analysis for assessment of ships or structures where structural vibration plays a dominant role.

The authors would like to gratefully acknowledge the support of many colleagues in Lloyd's Register for their contributions to this work. The authors also wish acknowledge support from the Lloyd's Register Global Technology Centre in Southampton, UK, and the Lloyd's Register Applied Technology Group in Halifax, Canada, to publish this work.

The views expressed in this work are those of the authors alone and do not necessarily represent the policy of Lloyd's Register or any of its affiliates or subsidiaries.

## References

- Bishop, R.E.D. and Price, W.G. (1979): Hydroelasticity of Ships, Cambridge University Press.
- Wu, Y. (1984): Hydroelasticity of Floating Bodies, Ph.D. Thesis, Brunel Univ., UK
- Wang, D. (1996): Three-Dimensional Hydroelastic Analysis of Ships in Time Domain, Ph.D. Thesis, China Ship Scientific Research Centre, China

## **A comparison of two computational methods (RANS and LES) in analysing of ventilation flow through a room fitted with a two-sided windcatcher for free flow condition**

†A.Niktash<sup>1</sup> and B.P. Huynh<sup>1</sup>

<sup>1</sup> School of Electrical, Mechanical and Mechatronic Systems,  
Faculty of Engineering & IT, University of Technology, Sydney, Australia

†Corresponding author: amirreza.niktash@uts.edu.au

### **Abstract**

A windcatcher is a structure for providing natural ventilation using wind power; it is usually fitted on the roof of a building to exhaust the inside stale air to the outside and supplies the outside fresh air into the building interior space working by pressure difference between outside and inside of the building. In this paper, the behaviour of free wind flow through a three-dimensional room fitted with a centred position two-canal bottom shape windcatcher model is investigated numerically, using a commercial computational fluid dynamics (CFD) software package and LES (Large Eddy Simulation) CFD method. The results have been compared with the obtained results for the same model but using RANS (Reynolds Averaged Navier-Stokes) CFD method. The model with its surrounded space has been considered in both method. It is found that the achieved results for the model from LES method are in good agreement with RANS method's results for the same model.

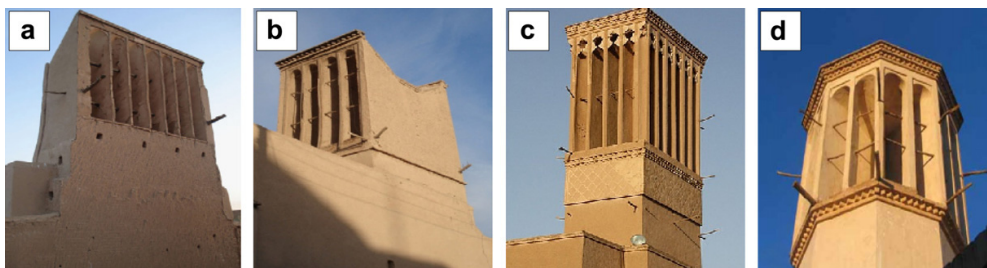
**Keywords:** CFD, Human comfort, K-ε, LES, RANS, Simulation, Ventilation, Windcatcher

### **Introduction**

Windcatcher is one of the green features for providing natural ventilation using wind power which has been employed over centuries in the hot arid parts of Iran and the other Persian Gulf countries to provide natural ventilation, passive cooling and thermal comfort [Fathy and Walter(1986); Karakastanis et al. (1986); Montazeri et al. (2010)].

The low cost of windcatcher system in comparison with mechanical ventilation system, being noiseless, durability, requiring no fossil energy, supplying clean air and using sustainable energy of wind power have led to use of the windcatcher as a passive and environmental friendly system.

The cross sections of all windcatchers which have circular or square shapes are divided internally into various segments to get one-sided, two-sided, three-sided, four-sided, hexahedral, and octahedral windcatchers to make the windcatcher system safer to periodic wind changes [Liu and Mak (2007)]. Some traditional windcatchers with different number of opening have been shown in Figure 1.



**Figure 1. Traditional wind towers with different number of openings (a) one-sided, (b) two-sided, (c) four-sided, (d) octahedral [Hughes et al. (2012)]**

The experimental studies of windcatcher systems for all different cases are obviously costly or even impossible. Employing computational fluid dynamics (CFD) for the assessment of windcatcher

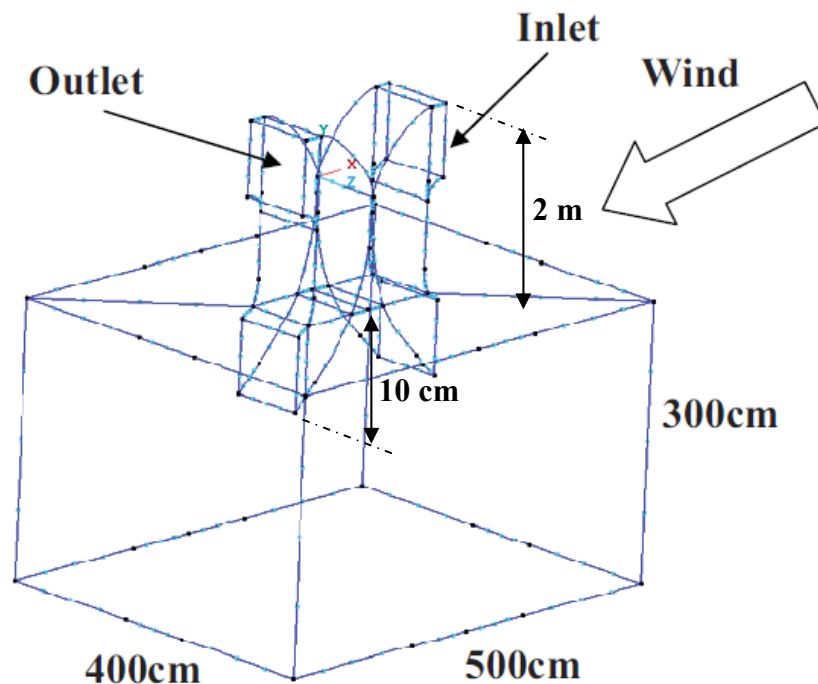


systems' performance is a new, useful and reliable tool with reasonable accuracy [Karakastanis et al. (1986) ; Elmualim (2006)].

LES and RANS are two applicable and popular CFD techniques for simulations. In this paper, a two-sided windcatcher is modeled by using LES CFD method which is more accurate but requiring more computational efforts and the results will be analysed and compared with the achieved results from previous studies has been done by the same authors for the same model by using RANS K- $\epsilon$  CFD technique Niktash and Huynh (2014a; 2014b)].

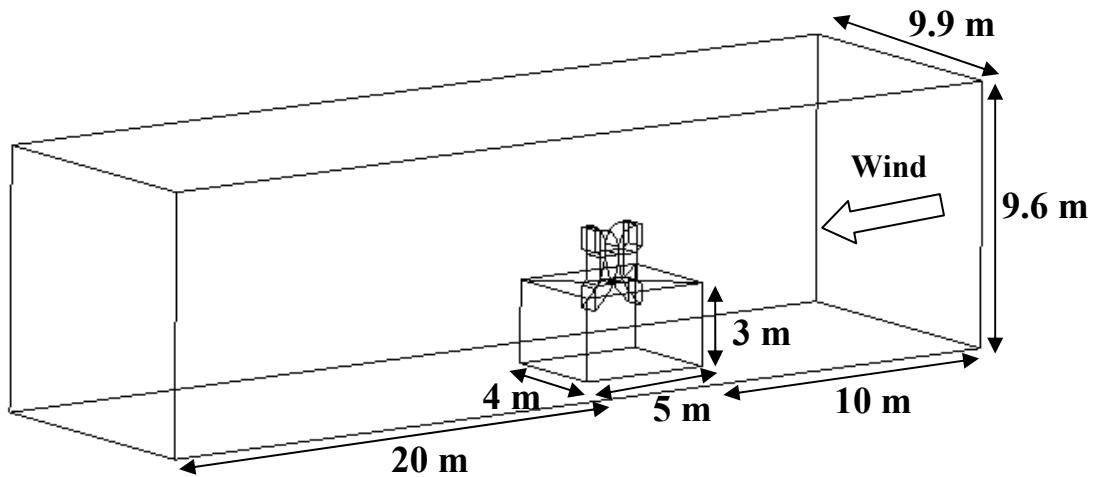
### Modelling and computation

A three dimensional room with the size of  $5 \times 4 \times 3 \text{ m}^3$  has been fitted with a two-sided centred position windcatcher including two-canal bottom shape and 10cm bottom length with square shape is simulated using CFD-ACE+, a CFD software package from the ESI group; it is assumed that the height of windcatcher's canal type inlet/outlet duct is 2 m and the inlet/outlet area is  $80 \times 80 \text{ cm}^2$ . Figure 2 shows the simulated room and its fitted windcatcher.



**Figure 2. 3D modelled room fitted with a two-sided windcatcher**

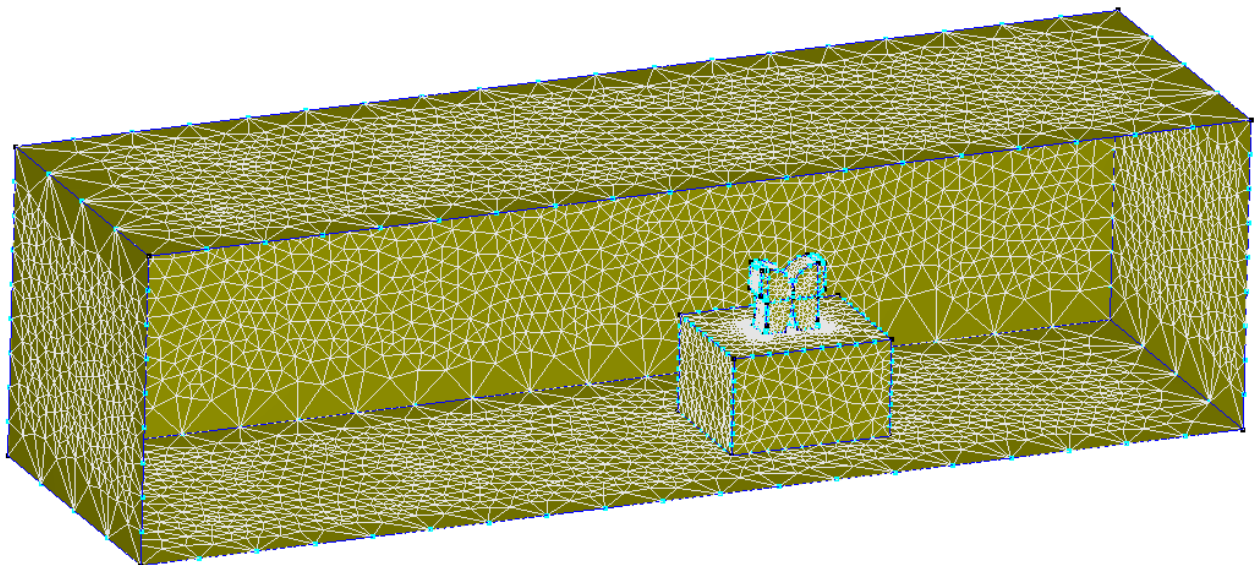
Figure 3 shows the complete system for simulation including the model with the windcatcher surrounded by the large space. The size of surrounded space has been chosen based on minimising the effect of forced flow and providing the conditions for free wind flow very similar to the real conditions.



**Figure 3. A complete system for simulation including the model and its large surrounded space**

It is assumed that the prevailing wind direction is perpendicular to the windcatcher's inlet/outlet surface and it blows with the velocity of 3 m/s from the right side of the surrounded space towards the modelled room and windcatcher and part of it enters the room via windcatcher's inlet and after circulation inside the room, it will exhaust through windcatcher's outlet to the outdoor and finally it will exit from left side of the surrounded space. All air properties are assumed to be constant and corresponding to air temperature as 300 K and air standard pressure at sea level as 101.3 kPa. To provide similar conditions with the previous studies, it has not been considered any door or window or any home accessory inside the room. Also the ventilation process is assumed to be an isothermal process to minimise the influence of thermal changes on the ventilation's quality.

To reach the sufficient accuracy in CFD simulation and due to complexity in the geometry of windcatcher's model, unstructured triangle meshes have been used throughout the whole system including room, windcatcher and the surrounded space; due to save on computational efforts and reduce accumulated errors, mesh distribution is less dense in the expected near stagnant flow regions like the room corners while dense mesh is applied for 3D model of the living area which is far from the corners of the room and windcatcher to observe the ventilation quality (Figures 4).

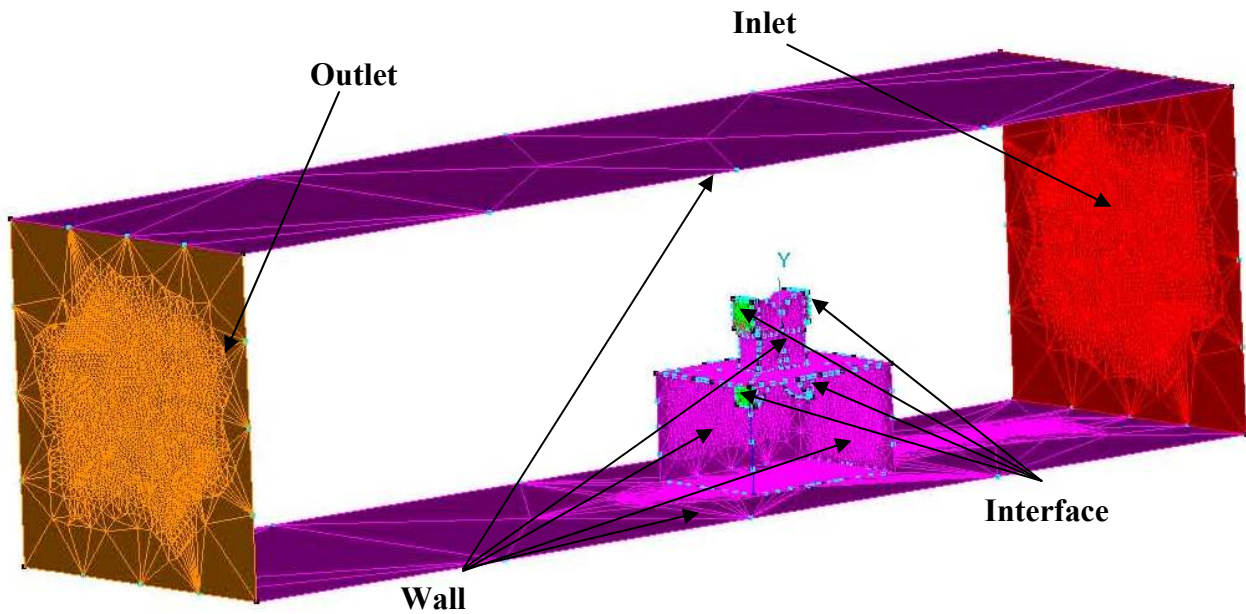


**Figure 4. Cross section of unstructured triangle meshes in a 3D model with surrounded area**

The grid-independence study has been done for different grid numbers in the model including surrounding space to make sure that the grid pattern used is adequate.

Consequently, the total number of grids in the complete system is about 135,000 and the maximum and the minimum grid areas are about  $1.4 \times 10^{-3} m^2$  and  $1.4 \times 10^{-7} m^2$ , respectively.

The right side of the surrounded space is defined as the inlet, the opposite side is considered as outlet. In the boundary condition part, four faces are defined as interface which let the flow pass through them; these faces are including the model's inlet/outlet and the inlet and the outlet of windcatcher's duct to the room. Figure 5 shows the defined faces of model. Some of the walls have been hidden in the figure to provide a clear view of other faces.



**Figure 5. Defined boundary conditions for different faces of the model**

Large eddy simulation (LES) is essentially a simulation that directly solves the large scale motion but approximates the small scale motion. A filtering process is used to separate out the large-scale components of the flow field from the small-scale ones. Within the finite volume method, it is rather sensible and natural to define the filter width as an average of the grid volume. The flow eddies larger than the filter width are large eddies while eddies smaller than the filter width are small eddies which require modelling. When filtering is performed on the incompressible Navier-Stocks equations, a set of equations very similar to the RANS equations in previous parts are obtained. Similar to RANS, there are additional terms where a modelling approximation must be introduced. In the context of LES, these terms are the subgrid scale turbulent stresses which require subgrid scale (SGS) models to close the set of equations [Tu et al. (2008)].

The most popular subgrid scale model for engineering applications is the one proposed by Smagorinsky [Smagorinsky (1963)] where in the eddy viscosity is proportional to the square of the grid spacing and the local strain rate. Since it is prescribed through the eddy viscosity assumption, it therefore shares many similarities to the formulation of the Reynolds stresses as obtained through the RANS approach [Tu et al. (2008)]. For the unsolved subgrid scale turbulent stresses, these are modeled accordingly as:

$$\tau_{ij} - \frac{\delta_{ij}}{3} \tau_{kk} = -2g_T^{SGS} \bar{S}_{ij} \quad (1)$$

$$\bar{S}_{ij} = \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \quad (2)$$

Where  $g_T^{SGS}$  is the subgrid scale kinematic viscosity and  $\bar{S}_{ij}$  is the strain rate of the large scale or resolved field.

The form of the subgrid scale eddy viscosity  $\mu_T^{SGS}$  (noting that  $g_T^{SGS} = \mu_T^{SGS} / \rho$ ) can be derived by dimensional arguments and is given by:

$$\mu_T^{SGS} = C_s^2 \rho \Delta^2 |\bar{S}_{ij}| \quad (3)$$

$$|\bar{S}_{ij}| = \sqrt{2\bar{S}_{ij}\bar{S}_{ij}} \quad (4)$$

Where  $\Delta$  is denoted by the grid filter width and the model constant  $C_s$  varies between 0.0065 and 0.3 depending on the particular fluid flow problem [Tu et al. (2008)]. In this work, Smagorinsky SGS model is applied and the averaged value is assumed for  $C_s$  which is 0.15325.

It is estimated that for the model and surrounded space totally it takes about 150 seconds to transfer flow from the inlet to the outlet of surrounded space by considering the full circulation inside the room's space. So the procedure would have 300 steps of 0.5 second which totally is 150 seconds. The convergence criteria has been defined as 0.0001 for higher accuracy.

## Results and discussion

Figure 7 shows velocity magnitude at level 1.2 m above the floor along the room's central plane.

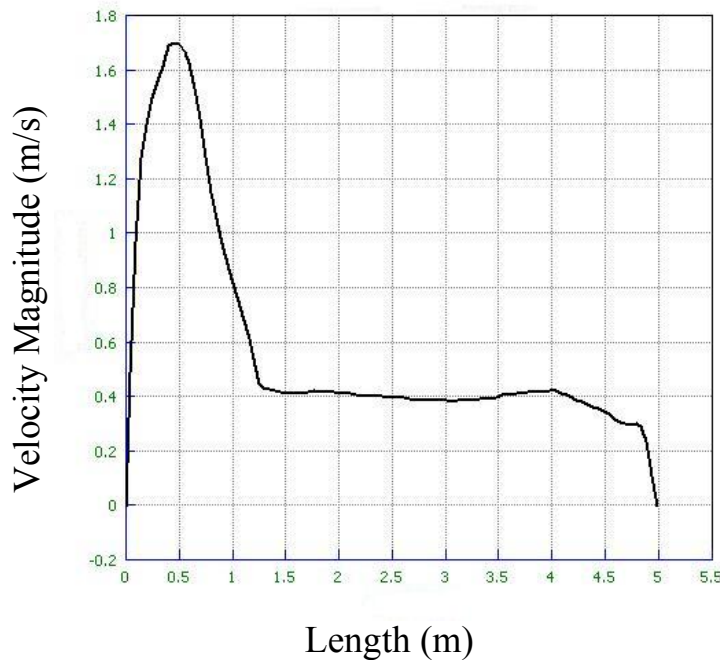
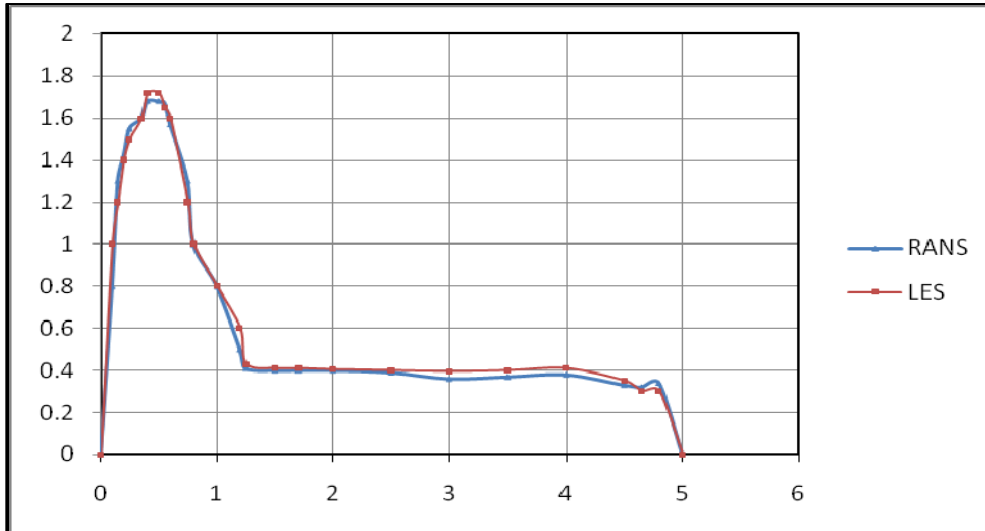


Figure 6, Velocity Magnitude at level 1.2 m above the floor for the model using LES method

It is seen that the velocity magnitude at this level is approximately stable in the range of 0.36~0.41 m/s at the distance of 1.4 m to 4.4 m from the right wall which is about 3 m of the total room's length (60% of the total room's length).

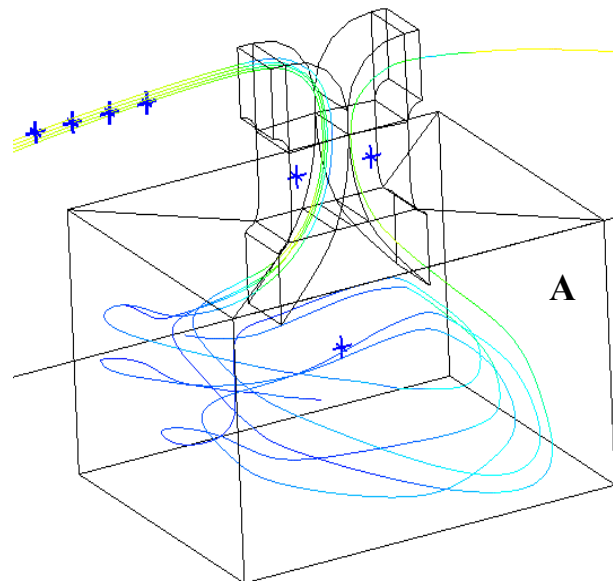
Figure 7 compares the velocity magnitudes profiles at level 1.2 m above the floor in the middle of the room for the model which has been achieved via RANS K- $\epsilon$  and LES methods.



**Figure 7. Comparing the calculated velocities in LES and K- $\epsilon$  methods for the model**

It is seen from the profiles, there is minor difference in calculated velocity via these two CFD techniques. This small difference is due to the accuracy of LES method while the trends of velocity changes are very similar to each other. Moreover, LES method confirms the acceptable range of velocity for human comfort (less than 0.8 m/s for indoor air speed in hot climates based on ASHRAE 2004) across the most regions of the room's length which already has been calculated and shown by using RANS method. The average difference of the two methods' results is about 5.45 %.

The flow traces for this model achieved by LES method is shown in Figure 8.

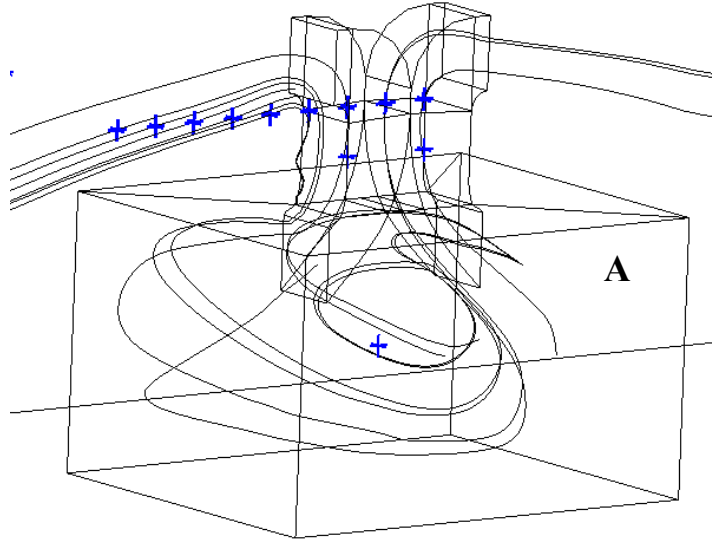


**Figure 8. Some traces of flow path for the model using LES method**

As it is seen from the above flow path, there is full ventilation in the living area and the uniform flow distribution in the room while there is only a small stagnation region in top corner of the room at the right side (indicated as “A” in Figure 8). The average flowrate for this model is  $1.075 \text{ m}^3/\text{s}$

during the defined time period.

Based on the previous studies, the flow traces for this model by using RANS K- $\epsilon$  method is shown in Figure 9.



**Figure 9. Some traces of flow path for the model by using RANS K- $\epsilon$  method [Niktash and Huynh (2014b)]**

As it is seen, the flowtraces are very similar to the achieved one from LES method. It verifies the RANS K- $\epsilon$  method's results with minor difference of % 4.1 (the flowrate was computed  $1.119 \text{ m}^3/\text{s}$  by RANS k- $\epsilon$  method in previous studies [Niktash and Huynh (2014a;2014b)]).

## Conclusions

The achieved results by LES method are so close to the RANS method results for this model and it confirms that two-sided centred position windcatcher with two-canal bottom shape, 10cm bottom length with square shape, canal type inlet/outlet, 2 m inlet/outlet ducts height, and perpendicular inlet/outlet's surface and the wind direction provides full circulation for most part of the room and large region of stable velocity magnitude in the acceptable range of indoor air velocity for human comfort.

RANS K- $\epsilon$  method is a useful computational method and good alternative for more accurate but time consuming LES method where very high accuracy is not required.

## References

- Elmualim, A.A. (2006) Dynamic modelling of a windcatcher/tower turret for natural ventilation, *Building Service Engineering Research Technology*, **27**, 165-182.
- Fathy, H. (1986) *Natural energy and vernacular architecture: principles and examples with reference to hot arid climates*, The university of Chicago press, USA.
- Hughes, B.R., Calautit, J.K. and Ghani, S.A. (2012), The development of commercial wind towers for natural ventilation: A review, *Applied energy*, **92**, 606-627.
- Karakatsanis, C., Bahadori, M.N. and Vickery, B.J. (1986) Evaluation of pressure coefficients and estimation of air flow rate in buildings employing wind tower, *Solar Energy* **37**, 363-374.
- Liu, L. and Mak, C.M. (2007) The assessment of the performance of a windcatcher system using computational fluid dynamics, *Building and Environment* **42**, 1135-1141.
- Montazeri, H., Montazeri, F., Azizian, R. and Mostafavi, S. (2010) Two-sided windcatcher performance evaluation using experimental, numerical and analytical modeling, *Renewable Energy* **35**, 1424-1435.
- Niktash A.R., Huynh B.P. (2014a) *Numerical Simulation and Analysis of the Two-sided Windcatcher Inlet/Outlet Effect in Ventilation Flow Through a Three Dimensional Room*, ASME 2014 Power Conference, USA.
- Niktash A.R., Huynh B.P. (2014b), *CFD Simulation and Analysis of a Two-sided Windcatcher's Inlet/Outlet Geometric Shape Effect in Ventilation Flow Through a Three Dimensional Room*, 19th Australasian Fluid Mechanics Conference, Australia.
- Smagorinsky, J. (1963) General circulation experiments with the primitive equations: I. The basic experiment, *Monthly weather review* **91**, 99-164.
- Tu, J. , Yeoh, G.H. and Liu, C. (2008) *Computational Fluid Dynamics - A Practical Approach*, Butterworth-Heinemann, MA, USA.

## Mass-redistributed method in the evaluation of eigenfrequency of solid systems

Eric Li<sup>1,2\*</sup>, ZC He<sup>1</sup>, Xu Xu<sup>2</sup>

<sup>1</sup>State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha, 410082 P. R. China

<sup>2</sup>College of Mathematics, Jilin University, Changchun 130012, China

### Abstract

Due to overly-stiff effect, standard Finite Element Method (FEM) using triangular and tetrahedral elements gives upper bound solution of natural frequencies. In order to improve the simulation results using low-order elements, one approach is to soften the overly-stiff of stiffness matrix to simulate the exact system, and the other is to match the mass matrix to the overly-stiff system. In this paper, the mass-redistributed method is further extended to analyse the eigenfrequency of solid systems. The mass-redistributed method is to modify the mass matrix of the discrete systems by shifting the integration points away from the Gaussian locations, while ensuring the mass conservation. Numerical examples including 2D and 3D problems have verified that Gaussian integration point in the mass matrix has a significant effect in the evaluation of eigenfrequency.

*key words: Mass-redistributed Method; Finite Element Method; Eigenfrequency*

### 1. Introduction

The overly-stiff finite element method predicts upper bound solution of eigenfrequency [1]. In order to soften the overly-stiff effect, smoothed finite element method (SFEM) is developed [2]. With different types of smoothing domain, edge-based smoothed finite element method (ES-FEM) [3-9], node-based smoothed finite element method (NS-FEM) [10], alpha finite element method ( $\alpha$ FEM) [11-13] and hybrid smoothed finite element method (HS-FEM) [14-17] have been formulated and applied to heat transfer, biomechanics, acoustic problems. The SFEM with right softened effect is able to achieve a close to exact stiffness, which provides more accurate solution of eigenfrequencies compared with finite element method.

Alternatively, the balance of discrete model between mass and stiffness can be achieved with modification of mass matrix. For example, a weighted average of the consistent and lumped mass matrices has been proposed by Marfurt in the computation of the mass matrix for acoustic and elastic wave propagation problems [18]. In addition, a new modified integration rules in the calculation of the mass and stiffness for acoustic problems for quadrilateral mesh is developed by Murthy [19]. He et al have further proposed mass-redistributed method to reduce the dispersion error in acoustic problems [20]. The re-distribution of mass matrix of the discrete systems will be modified by shifting the integration points away from the usual Gaussian locations [20]. In this paper, we further extend mass-redistributed method in the analysis of eigenfrequency.

The paper is organized as follows: Section 2 briefly describes the formulation of mass-redistributed method in the computation of eigenfrequency. Numerical examples including 2D and 3D are presented in Section 3 to investigate the effect of Gaussian integration point on eigenfrequency using mass-redistributed method. Finally the conclusions from the numerical results are made in Section 4.

---

<sup>1\*</sup>Corresponding author.

Email address: ericsg2012@gmail.com



## 2. Materials and Methods

### 2.1. 2D Mass-redistributed method

The standard FEM behaves stiffer than the exact system, which causes larger eigenvalues compared to exact one.

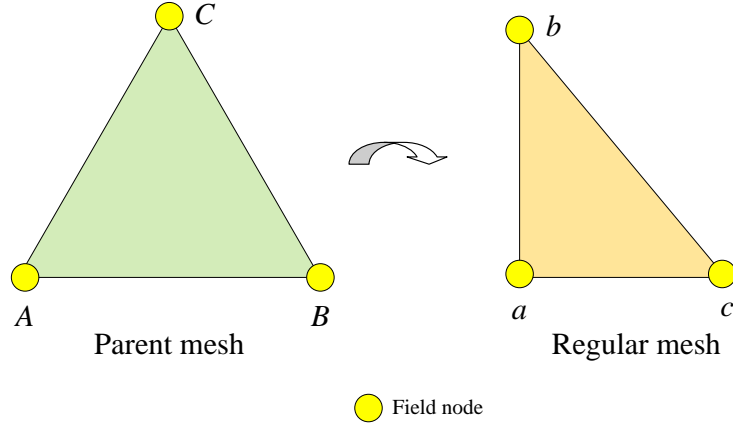


Figure 1: Transformation of triangular mesh

In the formulation of 2D mass-redistributed method, the first step is to transform parent mesh to regular mesh as shown in Fig. 1 [20]. Hence, the integration of stiffness and mass matrix is transformed into the natural coordinate system as follows:

$$\mathbf{K}^e = \int_{-1}^{+1} \int_{-1}^{+1} (\nabla \mathbf{N})^T (\nabla \mathbf{N}) \det(J) d\xi d\eta, \quad \mathbf{M}^e = \int_{-1}^{+1} \int_{-1}^{+1} \mathbf{N}^T \mathbf{N} \det(J) d\xi d\eta, \quad (1)$$

At the same time, the linearly shape function  $N_i(x)$  can be written in the area coordinates form as shown in Fig. 2.

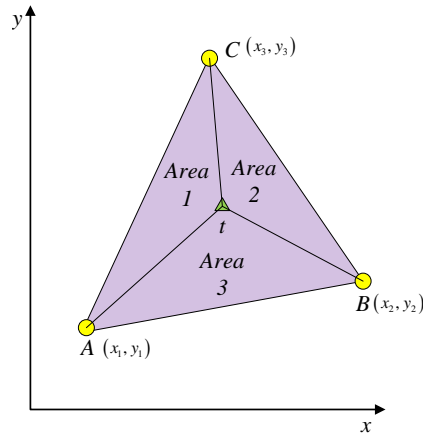


Figure 2: Area coordinates for a linear triangular element

With connection  $t$  to three vertices 1, 2, 3, three sub-areas  $A_1, A_2,$  and  $A_3$  corresponding to the triangles  $tAC, tBC$  and  $tBA$ , respectively can be determined for any given point  $t$  within the element. Using area coordinates, the shape functions for each node are expressed as follows:

$$N_1 = \frac{A_1}{A} = 1 - \xi - \eta ; \quad N_2 = \frac{A_2}{A} = \xi ; \quad N_3 = \frac{A_3}{A} = \eta \quad (2)$$

With Gauss quadrature rule, Eq. (1) can be evaluated:

$$\int_{-1}^{+1} \int_{-1}^{+1} \phi d\xi d\eta = \sum_{i=1}^{ng} w_i \phi(\xi_i, \eta_i) \quad (3)$$

where  $ng$  is the number of Gauss points,  $w_i$  are the weights and  $\xi, \eta$  are the local coordinates of Gauss points. Then the generalized stiffness and mass matrix can be obtained by the integration rule.

$$\begin{aligned} \mathbf{K}^e &= w_i (\nabla \mathbf{N}(\xi_i, \eta_i))^T (\nabla \mathbf{N}(\xi_i, \eta_i)) \det(J) \\ \mathbf{M}^e &= \sum_{i=1}^3 w_i^k (\mathbf{N}(\xi_i^k, \eta_i^k))^T (\mathbf{N}(\xi_i^k, \eta_i^k)) \det(J) \end{aligned} \quad (4)$$

From Eq. (4), we can notice that the gradient of shape function using triangular mesh is constant within the elements and thus the integration of stiffness matrix in Eq.(4) equals to the constant multiplied by the Jacobi.

The consistent mass matrix is evaluated by three Gauss points with  $w_i = 1/3$ .

$$\text{point 1: } \xi_i = 2/3, \eta_i = 1/6 \quad \text{point 2: } \xi_i = 1/6, \eta_i = 2/3 \quad \text{point 3: } \xi_i = 1/6, \eta_i = 1/6 \quad (5)$$

On the other hand, the lumped mass matrix is formed with the following points:

$$\text{point 1: } \xi_i = 1, \eta_i = 0 \quad \text{point 2: } \xi_i = 1/6, \eta_i = 2/3 \quad \text{point 3: } \xi_i = 1/6, \eta_i = 1/6 \quad (6)$$

The generalized mass matrix is evaluated by introducing a variable integration point  $t$ :

$$\begin{aligned} \xi_i &= t, \quad \eta_i = \frac{1-t}{2} \quad (\text{Point 1}) \\ \xi_i &= \frac{1-t}{2}, \quad \eta_i = t \quad (\text{Point 2}) \\ \xi_i &= \frac{1-t}{2}, \quad \eta_i = \frac{1-t}{2} \quad (\text{Point 3}) \end{aligned} \quad (7)$$

where  $t \in [0, 1]$ , and weight  $w_i = 1/3$  so that  $\sum_{i=1}^3 w_i = 1$ .

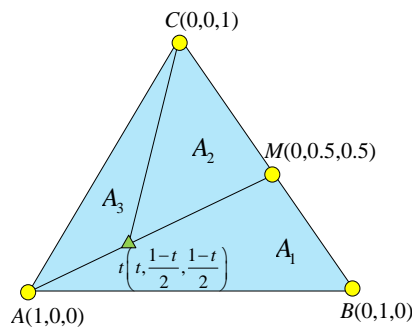


Figure 3: Modified integration point for the mass matrix of linear triangular element

It is noticed that the integration point will move from a triangular vertex to midpoint of the edge opposite to the vertex, as shown in Fig. 3 as  $t$  varies from 0 to 1.

Using these three sampling points, the mass matrix can be expressed in terms of parameter

$t$  as:

$$\mathbf{M}^e = \frac{1}{3} A^e \begin{bmatrix} a_{11} & 0 & a_{13} & 0 & a_{15} & 0 \\ 0 & a_{22} & 0 & a_{24} & 0 & a_{26} \\ a_{31} & 0 & a_{33} & 0 & a_{35} & 0 \\ 0 & a_{42} & 0 & a_{44} & 0 & a_{46} \\ a_{51} & 0 & a_{53} & 0 & a_{55} & 0 \\ 0 & a_{62} & 0 & a_{64} & 0 & a_{66} \end{bmatrix} \quad (8)$$

where

$$a_{11} = a_{22} = a_{33} = a_{44} = a_{55} = a_{66} = t^2 + \frac{(1-t)^2}{2}$$

$$a_{13} = a_{15} = a_{24} = a_{26} = a_{31} = a_{35} = a_{42} = a_{46} = a_{51} = a_{53} = a_{62} = a_{64} = t(1-t) + \frac{(1-t)^2}{4}$$

The summation of all matrix elements of  $\mathbf{M}^e$  equals to the area of the element regardless of value of  $t$ :

$$\sum_{i=1}^{36} \frac{1}{6} A^e \mathbf{M}_i^e = \underset{\text{total area of the element}}{A^e} \quad (9)$$

The eigenvalues are determined by the stiffness and mass matrix:

$$\omega^2 I = \frac{\boldsymbol{\phi}^T \mathbf{K} \boldsymbol{\phi}}{\boldsymbol{\phi}^T \mathbf{M} \boldsymbol{\phi}} = \begin{bmatrix} \frac{\lambda_{k1}}{\lambda_{m1}} & & & & & \\ & \frac{\lambda_{k2}}{\lambda_{m2}} & & & & \\ & & \ddots & & & \\ & & & & \frac{\lambda_{kn}}{\lambda_{mn}} & \\ & & & & & \lambda_{mn} \end{bmatrix} \quad (10)$$

where  $\lambda_{k1} \sim \lambda_{kn}$  are the eigenvalues of stiffness matrix  $\mathbf{K}$ ,  $\lambda_{m1} \sim \lambda_{mn}$  are the eigenvalues of mass matrix.

## 2.2. 3D Mass-redistributed method

In the formulation of 3D mass-distributed method using linear tetrahedral element, the mapping from a parent tetrahedral mesh to the regular tetrahedral mesh is based on volume coordinates. Thereby, the integration of stiffness and mass matrix is transformed into the natural coordinate system as follows:

$$\mathbf{K}^e = \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} (\nabla \mathbf{N})^T (\nabla \mathbf{N}) \det(J) d\xi d\eta d\zeta, \quad (11)$$

$$\mathbf{M}^e = \int_{-1}^{+1} \int_{-1}^{+1} \int_{-1}^{+1} \mathbf{N}^T \mathbf{N} \det(J) d\xi d\eta d\zeta,$$

Similarly, the linearly shape functions is expressed in the volume coordinates form as shown in Fig. 4.

Four sub-volumes  $V_1, V_2, V_3$  and  $V_4$  corresponding to the tetrahedron  $tBCD, tACD, tABD$  and  $tABC$ , respectively can be determined for any given point  $t$  within the element. Hence, the

shape function is formulated as follows:

$$\begin{aligned}
 N_1 &= \frac{V_{tBCD}}{V_{ABCD}} = 1 - \xi - \eta - \zeta ; & N_2 &= \frac{V_{tACD}}{V_{ABCD}} = \xi ; \\
 N_3 &= \frac{V_{tABD}}{V_{ABCD}} = \eta ; & N_4 &= \frac{V_{tABC}}{V_{ABCD}} = \zeta
 \end{aligned} \tag{12}$$

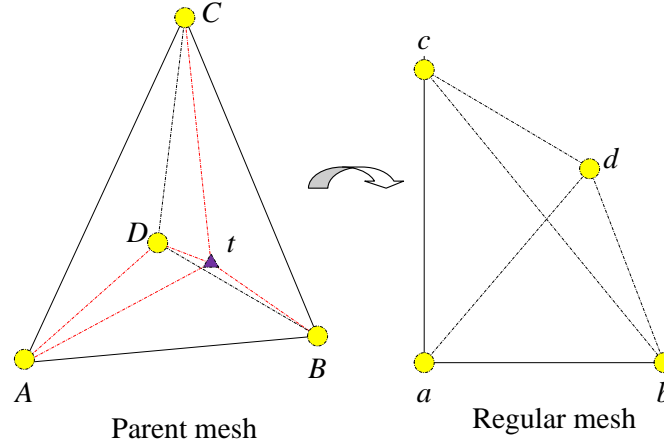


Figure 4: Volume coordinate using tetrahedral mesh

Similar to 2D case, the Gaussian integration point  $t$  controls the formulation of 3D mass matrix. For example, the consistence mass matrix is evaluated by using following four Gauss points with weight  $w_i = 1/4$ :

$$\text{point 1: } \xi_i = 0.5854102, \eta_i = 0.1381966, \zeta_i = 0.1381966,$$

$$\text{point 2: } \xi_i = 0.1381966, \eta_i = 0.5854102, \zeta_i = 0.1381966$$

$$\text{point 3: } \xi_i = 0.1381966, \eta_i = 0.1381966, \zeta_i = 0.5854102$$

$$\text{point 4: } \xi_i = 0.1381966, \eta_i = 0.1381966, \zeta_i = 0.1381966$$

Using the following four Gauss points with weight  $w_i = 1/4$ , the lumped mass matrix is obtained:

$$\text{point 1: } \xi_i = 1, \eta_i = 0, \zeta_i = 0, \quad \text{point 2: } \xi_i = 0, \eta_i = 1, \zeta_i = 0$$

$$\text{point 3: } \xi_i = 0, \eta_i = 0, \zeta_i = 1, \quad \text{point 4: } \xi_i = 0, \eta_i = 0, \zeta_i = 0$$

Here, the general formulation of mass matrix with a variable Gaussian integration point  $t$  is written as follows:

$$\text{point 1: } \xi_i = t, \eta_i = \frac{1-t}{3}, \zeta_i = \frac{1-t}{3}, \tag{13}$$

$$\text{point 2: } \xi_i = \frac{1-t}{3}, \eta_i = t, \zeta_i = \frac{1-t}{3}$$

$$\text{point 3: } \xi_i = \frac{1-t}{3}, \eta_i = \frac{1-t}{3}, \zeta_i = t$$

$$\text{point 4: } \xi_i = \frac{1-t}{3}, \eta_i = \frac{1-t}{3}, \zeta_i = \frac{1-t}{3}$$

In Eq. (11), weight  $w_i = 1/4$  so that  $\sum_{i=1}^4 w_i = 1$ , and  $t \in [0,1]$ . Using these four sampling points, the mass matrix can be formulated as:

$$\mathbf{M}^e = \frac{1}{4} V^e \begin{pmatrix} b_{11} & 0 & 0 & b_{14} & 0 & 0 & b_{17} & 0 & 0 & b_{110} & 0 & 0 \\ 0 & b_{22} & 0 & 0 & b_{25} & 0 & 0 & b_{28} & 0 & 0 & b_{211} & 0 \\ 0 & 0 & b_{33} & 0 & 0 & b_{36} & 0 & 0 & b_{39} & 0 & 0 & b_{312} \\ b_{41} & 0 & 0 & b_{44} & 0 & 0 & b_{47} & 0 & 0 & b_{410} & 0 & 0 \\ 0 & b_{52} & 0 & 0 & b_{55} & 0 & 0 & b_{58} & 0 & 0 & b_{511} & 0 \\ 0 & 0 & b_{63} & 0 & 0 & b_{66} & 0 & 0 & b_{69} & 0 & 0 & b_{612} \\ b_{71} & 0 & 0 & b_{74} & 0 & 0 & b_{77} & 0 & 0 & b_{710} & 0 & 0 \\ 0 & b_{82} & 0 & 0 & b_{85} & 0 & 0 & b_{88} & 0 & 0 & b_{811} & 0 \\ 0 & 0 & b_{93} & 0 & 0 & b_{96} & 0 & 0 & b_{99} & 0 & 0 & b_{912} \\ b_{101} & 0 & 0 & b_{104} & 0 & 0 & b_{107} & 0 & 0 & b_{1010} & 0 & 0 \\ 0 & b_{112} & 0 & 0 & b_{115} & 0 & 0 & b_{118} & 0 & 0 & b_{1111} & 0 \\ 0 & 0 & b_{123} & 0 & 0 & b_{126} & 0 & 0 & b_{129} & 0 & 0 & b_{1212} \end{pmatrix} \quad (14)$$

where

$$b_{11} = b_{22} = b_{33} = b_{44} = b_{55} = b_{66} = b_{77} = b_{88} = b_{99} = b_{10} = b_{11} = b_{12} = t^2 + \frac{(1-t)^2}{3}$$

The remaining parameters in Eq. (14) equal to  $2t \frac{(1-t)}{3} + \frac{2(1-t)^2}{9}$

The summation of all matrix elements of  $\mathbf{M}^e$  equals to the volume of the element regardless of value of  $t$ :

$$\begin{aligned} & \sum_{i=1}^{144} \frac{1}{12} V^e \left[ 12 \left( t^2 + \frac{(1-t)^2}{3} \right) + 36 \left( 2t \frac{(1-t)}{3} + \frac{2(1-t)^2}{9} \right) \right] \\ & = V^e \\ & \text{total mass of the element} \end{aligned} \quad (15)$$

### 3. Numerical examples

In this section, some examples will be analysed to demonstrate the properties of the present method. The triangular elements in 2D and tetrahedral elements in 3D are used.

### 4.1 2D examples

The first 2D cantilever beam with length  $l=1\text{m}$  and width  $d=0.2\text{m}$  is studied, which is subjected to fixed boundary condition at left hand side as outlined in Fig. 5

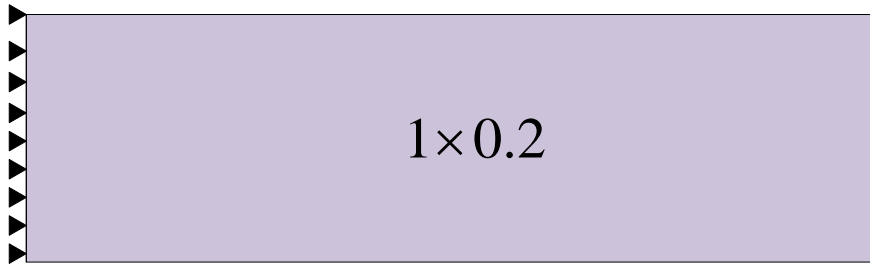


Figure 5: 2D cantilever beam

#### 4.1.1 Effect of Gaussian integration point

As shown in Fig. 6, the effect of Gaussian integration point on the computation of eigenfrequency using FEM is presented. In order to show the applicability of the proposed method, mode shape 1 and 4 are studied. It is easily observed that the Gaussian integration point controls the value of eigenfrequency.

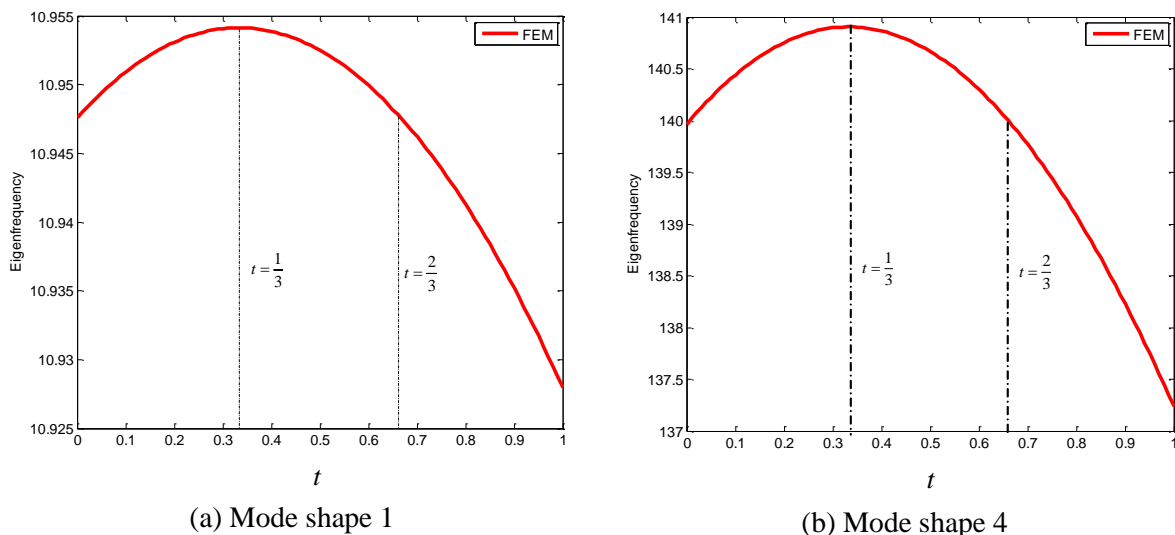
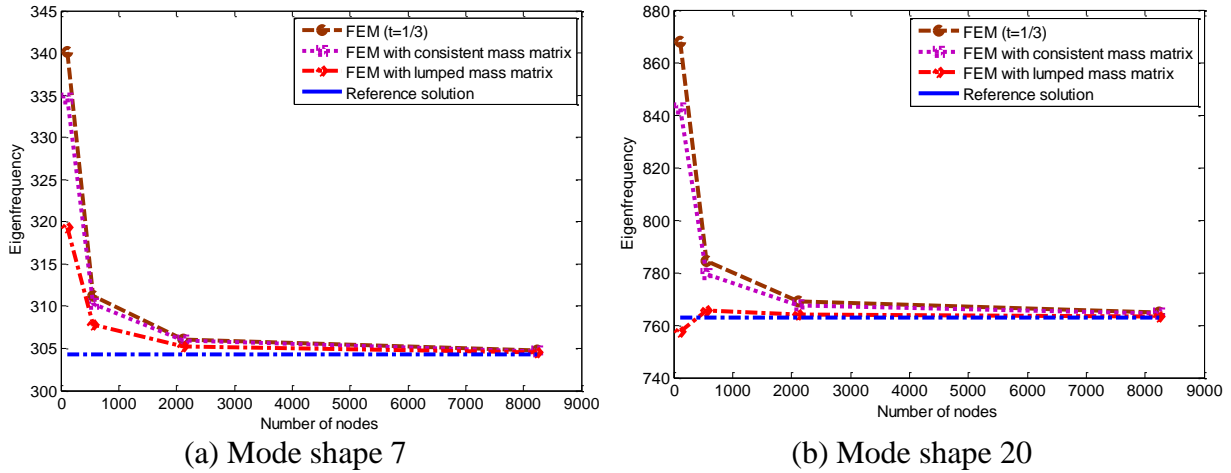


Figure 6: Effect of Gaussian integration point on eigenfrequency using FEM

As shown in Fig. 6, it is seen that  $t=1/3$  provides the maximum eigenfrequency. As  $t=1$  is corresponding to lumped mass matrix, the minimum eigenfrequency is obtained. While the eigenfrequency from  $t=2/3$  which gives consistent mass matrix is less than value for  $t=1/3$ , but greater than the one using lumped mass matrix.

The convergence rate of eigenfrequency for mode shape 7 and 20 using different Gaussian integration point is presented in Fig. 7. It is clearly indicated that FEM with lumped mass matrix gives the most accurate results compared with consistent mass matrix. In addition, it is noticed that Gaussian integration  $t=1/3$  always gives the maximum eigenfrequency.



(a) Mode shape 7 (b) Mode shape 20  
 Figure 7: Convergence rate for different Gaussian integration point using FEM

### 4.2 3D examples

In this section, the mass-redistributed method is further extended to 3D problem.

#### 4.2.1 Effect of Gaussian integration point

As shown in Fig. 8, a three dimensional cantilever beam with dimension  $1 \times 0.2 \times 0.2$  subjected to fixed boundary condition is studied. The eigenfrequency is computed to analyze the performance of mass-redistributed method.

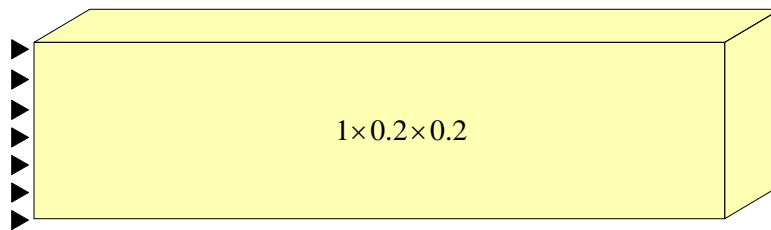
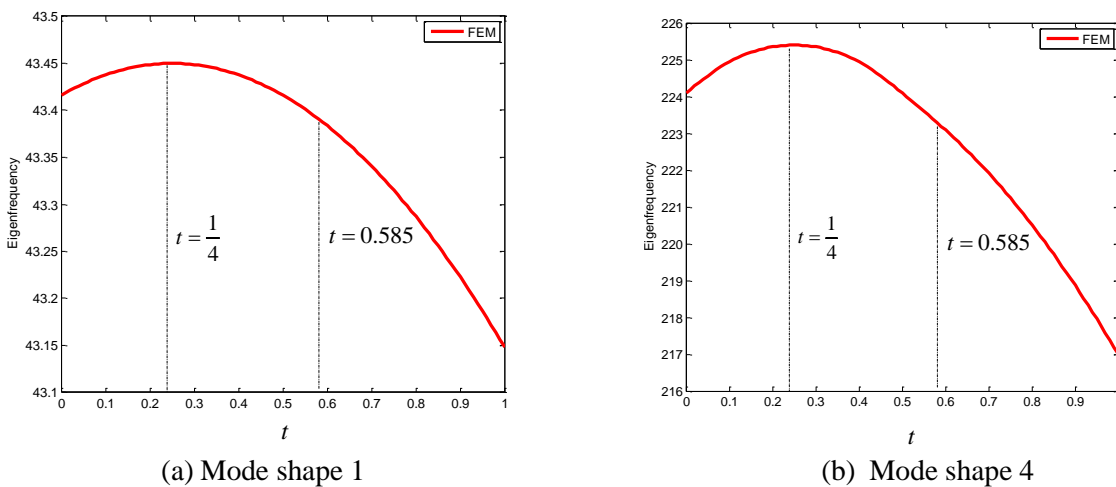


Figure 8: 3D cantilever beam

As outlined in Fig. 9, it is seen that the consistent mass matrix is corresponding to  $t=0.585$ . The minimum eigenfrequency is obtained as  $t$  equals to 1 ( $t=1$  formulates the lumped mass matrix). It is seen that  $t=1/4$  always provides the largest eigenfrequency as presented in Fig.9.



(a) Mode shape 1 (b) Mode shape 4  
 Figure 9: Effect of Gaussian integration point on mode shape using FEM

To further investigate the effect of mass-redistributed method in the evaluation of eigenfrequency, the convergence rate for mode shape 5 and 20 using different Gaussian integration point is shown in Fig. 10. Again, it is found that the lumped mass matrix always predicts much better solution compared with consistent mass matrix. The Gaussian integration point  $t=1/4$  gives the largest solution of eigenfrequency. The solution from consistent mass matrix is always between the Gaussian integration point  $t=1$  (corresponding to lumped mass matrix) and  $t=1/4$ .

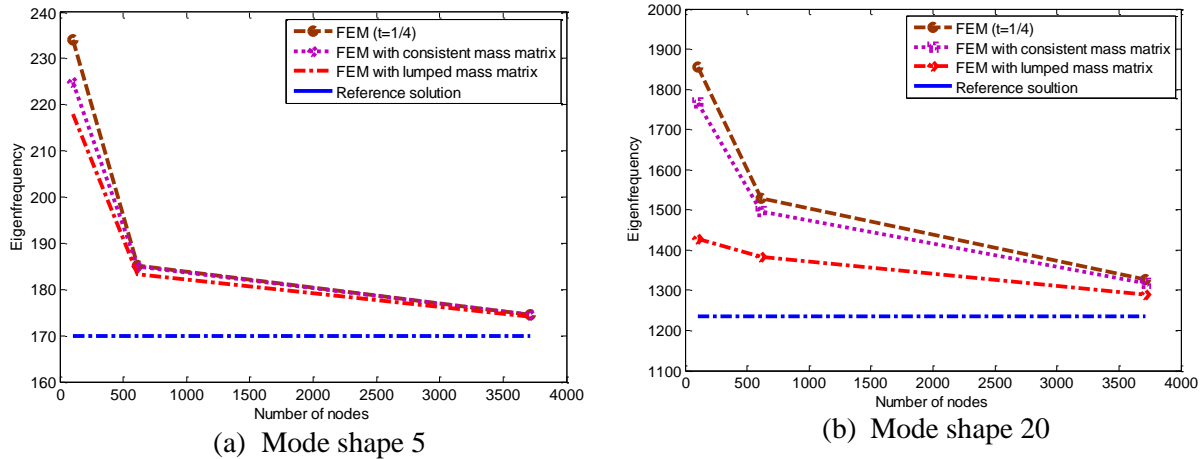


Figure 10: Convergence rate for different Gaussian integration point using FEM

## 4. Conclusion

In this paper, a mass-redistributed method is further developed for solving eigenfrequency by modification of Gaussian integration point in the computation of mass matrix. In the present mass-redistributed method, the alternation of Gaussian integration point always ensures the mass conservation. The triangular and tetrahedral elements are focused in this study. Both theoretical and numerical results have demonstrated that the Gaussian integration point in the mass matrix has a significant effect on the prediction of eigenfrequency, and the following conclusions can be summarized:

- With adjustment of Gaussian integration point, the eigenfrequency of systems can be modified.
- The  $t=1/3$  gives the maximum eigenfrequency in the 2D model; while  $t=1/4$  results in the maximum eigenfrequency of 3D model
- Due to the correct balance between the stiffness and mass matrix, the lumped mass matrix has the best solution in the computation of eigenfrequency.

## Acknowledgements

The project is supported by the Science Fund of State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body Nos. 31315002 and 51375001. The authors wish to thank the National Natural Science Foundation of China (Grant No. 11202074) and Research Project of State Key Laboratory of Mechanical System and Vibration MSV201403 for the support.

## Reference

- GR Liu. Meshfree methods: Moving Beyond the Finite Element Method, 2nd Edition, CRC Press:



- Boca Raton, USA, 2009.
2. GR Liu, TT Nguyen. Smoothed Finite Element Methods. Boca Raton: CRC Press; 2010.
  3. GR Liu, T Nguyen-Thoi, KY Lam. An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analyses in solids. *Journal of Sound and Vibration*, 320: 1100-1130, 2009.
  4. Eric Li, G. R. Liu, Vincent Tan. Simulation of Hyperthermia Treatment Using the Edge-Based Smoothed Finite-Element Method. *Numerical Heat Transfer, Part A: Applications*. 2010; 57: 11, 822 -847
  5. Z. C. He, G. Y. Li, Z. H. Zhong, A. G. Cheng, G. Y. Zhang, Eric Li, G. R. Liu. An ES-FEM for accurate analysis of 3D mid-frequency acoustics using tetrahedron mesh. *Computers and Structures*. 2012; 106-107: 125-134.
  6. Z.C. He, A.G. Cheng, Z.H. Zhong, G.Y. Zhang, G.Y. Li, Eric Li. An improved eigenfrequencies prediction for three-dimensional problems using face-based smoothed finite element method. *Journal of ActaMechanicaSolidaSinica*. 2013; 26:140-150.
  7. Z. C. He, G. Y. Li, Z. H. Zhong, A. G. Cheng, G. Y. Zhang, G. R. Liu, Eric Li, Z. Zhou. An edge-based smoothed tetrahedron finite element method (ES-T-FEM) for 3D static and dynamic problems. *Computational mechanics*. 2013; 2:221 -236.
  8. Eric Li, Z.C. He, Xu Xu. A novel edge-based smoothed tetrahedron finite element method (ES-T-FEM) for thermomechanical problems. *International Journal of Heat and Mass Transfer* 66 (2013) 723–732.
  9. Z. C. He, G. Y. Li, Eric Li, Z. H. Zhong, G. R. Liu. Mid-frequency acoustics analysis using edge-based smoothed tetrahedron radial point interpolation method (ES-T-RPIM). *Computational methods* DOI: 10.1142/S021987621350103X.
  10. GR Liu, TNguyen-Thoi, H Nguyen-Xuan, KY Lam. A node-based smoothed finite element method (NS-FEM) for upper bound solution to solid mechanics problems. *Computers and Structures*, 87: 14-26, 2009.
  11. GR Liu, T Nguyen-Thoi, KY Lam. A novel Alpha Finite Element Method ( $\alpha$ FEM) for exact solution to mechanics problems using triangular and tetrahedral elements. *Computer Methods in Applied Mechanics and Engineering*, 197: 3883-3897, 2008.
  12. Eric Li, G. R. Liu, Vincent Tan, and Z. C. He. Modeling and simulation of bioheat transfer in the human eye using the 3D alpha finite element method ( $\alpha$ FEM). *Intentional Journal for Numerical Methods in Biomedical Engineering*. 2010; 26:955–976
  13. Eric Li, G.R. Liu, Vincent Tan, Z.C. He. An efficient algorithm for phase change problem in tumor treatment using  $\alpha$ FEM. *International Journal of Thermal Sciences*. 2010; 49: 10, 1954-1967.
  14. Eric Li, ZP Zhang, ZC He, Xu Xu, GR Liu, Q Li. Smoothed finite element method with exact solutions in heat transfer problems. *International Journal of Heat and Mass Transfer* 78(2014) 1219-1231.
  15. Eric Li, ZC He, L Chen, Xu Xu, GR Liu. An ultra-accurate hybrid smoothed finite element method for Piezoelectric problem. *Engineering Analysis with boundary element* 50 (2015) 188–197.
  16. Eric Li, Xu Xu, Z. C. He, G.R. Liu. Hybrid smoothed finite element method for acoustic problems. *Computer Methods in Applied Mechanics and Engineering* 283 (2015) 664–688.
  17. Eric Li, Zhongpu Zhang, CC Chang, GR Liu, Q Li. Numerical homogenization for incompressible materials using selective smoothed finite element method. *Composites Structures*123 (2015) 216–232.
  18. Marfurt KJ, Accuracy of finite difference and finite element modeling of the scalar and elastic wave equation, *Geophysics*.1984; 49: 533–549.
  19. Murthy N, Guddati; Bin Yue. Modified integration rules for reducing dispersion error in finite element methods. *Computer Methods in Applied Mechanics and Engineering* 2004;193(3-5):275-287.
  20. ZC He, GY Li, GR Liu, AG Cheng, Eric Li. Numerical investigation of ES-FEM with various Mass re-distribution for acoustic problems. 89 (2015) 222–233

# Viscoelastic Response of Composite Laminated Shells based on Efficient Higher-Order Theory

\*S-N Nguyen, J. Lee and †M. Cho

Department of Mechanical and Aerospace Engineering, Seoul National University, Republic of Korea.

\*Presenting author: nguyensyngoc@snu.ac.kr

†Corresponding author: mcho@snu.ac.kr

## Abstract

The Laplace transformation has been applied for analyzing viscoelastic behaviors of composite laminated shells. The constitutive equation in integral can be reduced to linear stress-strain relationship in Laplace domain. Therefore, the formulation such as displacement field, can be expressed similarly with elastic one. In order to describe accurately the viscoelastic behaviors of laminated shells, the efficient higher-order shell theory (EHOST) has employed. The general tensor formulation is used for accurate description and arbitrary shell. The time-dependent results of strain, stress distributions through the thickness of the shell are obtained by converting back to real time domain for creep and relaxation process.

**Keywords:** Laplace transform, viscoelastic, composite laminated shell, efficient higher-order shell theory

## Introduction

Due to the time-dependent effects of matrix, viscoelastic behavior is one of important factors in the analysis of composite structures. These responses significantly depend on the applied loading and environmental conditions such as temperature or moisture. Nonetheless, the Boltzmann's superposition principle in integral form of viscoelastic analysis is much more complicate than Hook's law of elastic one [Yi and Hilton (1995); Venkat et al. (2012)]. Consequence, the computation costs extremely expensive for thick laminates. It needs to find a method which satisfies both efficiency and accuracy.

Some researcher have employed the Laplace or Fourier transforms to avoid above limitation [Lin and Hwang (1989); Chen (1995)]. The converted constitutive equation of viscoelastic materials in Laplace domain is very similar to the Hook's laws of elastic one. However, all of formulation procedures are done in Laplace domain. Finally, the results in real time domain are obtained by inversion Laplace transforms [Hassanzadeh and Pooladi-Darvish (2007); Dubner and Abate (1968)]. Hence, these methods are effective for long-term problems of viscoelastic analysis.

On the other hand, plates and shells more and more are popular structures, especially in aerospace engineering such as wings and fuselages. Because of potential application, various plate and shell theories have been developed to improve the accuracy as well as efficiency [Reddy (2004); Pagano (1969); Cho and Parmerter (1993; 1994); Kim and Cho (2003); Nguyen et al. (2015)]. To express adequately viscoelastic behaviors and the advance of present method, EHOST is good choice which can predict accurately and efficiently strain stress distribution through the thickness of composite laminated shell [Kim and Cho (2003)].

The present study aims to develop EHOST for viscoelastic composite laminates. Besides, the Laplace transform is employed to reduce the mortal formulation and save an expanse computation amount. Thus, this study provides a powerful tool for investigating the time-dependent response of composite laminated shell with the advantage of accuracy and efficiency.

## Mathematical Formulation

### 1. Constitutive equation for viscoelastic material and Laplace transform

The constitutive equation for a linear viscoelastic materials can be expressed, as follows:

$$\sigma_i(t) = \int_0^t Q_{ij}(t-t') \frac{\partial \varepsilon_j(t')}{\partial t'} dt \quad (1)$$

where  $t$  denotes time,  $t'$  is a dummy variable for integration,  $Q_{ijkl}(t)$  is relaxation modulus which can be well represented by a series of decaying exponentials, as Prony series.

By applying Laplace transform, the above complex equation can be reduced as linear relationship in Laplace domain as follows:

$$\sigma_i^*(s) = s Q_{ij}^*(s) \varepsilon_j^*(s) \quad (2)$$

where  $( )^*$  denotes parameters in the Laplace domain. Therefore, the procedures of viscoelastic composite laminated shell analysis in Laplace domain is the same to elastic counterpart in time domain.

Finally, the viscoelastic responses in real time domain can be obtained by using inverse Laplace techniques. In present study, the Fourier series algorithm is employed to convert the strain, stress values back into the real time domain as follows:

$$f(t) = \frac{e^{at}}{t} \left\{ \frac{1}{2} F(a) + \text{Re} \sum_{k=1}^{\infty} F\left(a + j \frac{k\pi}{t}\right) (-1)^k \right\} \quad (3)$$

where  $F(s)$  is a function in the Laplace domain that need to converted,  $f(t)$  is the converted function in real time domain, and  $\mathbf{a}$  and  $\mathbf{n}$  are the parameters that must be optimized for accuracy.

### 2. Efficient higher-order shell theory for viscoelastic composite laminates

The composited shells are considered, consisting of  $n$  orthotropic layers with uniform thickness  $h$  in a curvilinear orthogonal coordinate system  $\mathbf{x}^\alpha$  as shown in Fig. 1.  $(\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3=0)$  represents curvature of the mid-surface which is chosen as the reference surface. The position vector  $\mathbf{R}$  can be expressed by position vector  $\mathbf{r}$  pointing to reference surface and unit vector  $\mathbf{a}^3$ , which is perpendicular to the surface at the point  $\mathbf{x}^\alpha$ , as follows:

$$\bar{\mathbf{R}}(x^\alpha, x^3) = \bar{\mathbf{r}}(x^\alpha) + x^3 \bar{\mathbf{a}}_3(x^\alpha) \quad (4)$$

where  $x^3$  is the distance from the reference surface to the material point. The covariant base vector is defined as:

$$\bar{\mathbf{g}}_i = \frac{\partial \bar{\mathbf{R}}}{\partial x^i} = \frac{\partial \bar{\mathbf{r}}}{\partial x^i} + \frac{\partial}{\partial x^i} (z \bar{\mathbf{a}}_3) \quad (5)$$

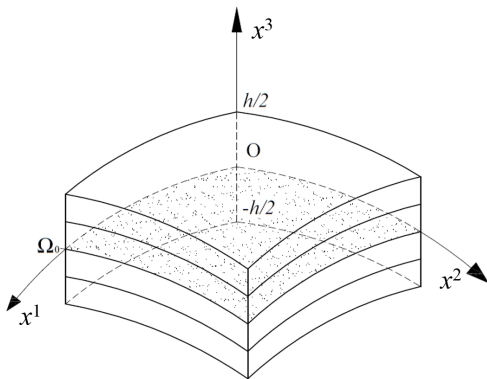


Fig. 1. Geometry and coordinates of the rectangular laminated shell.

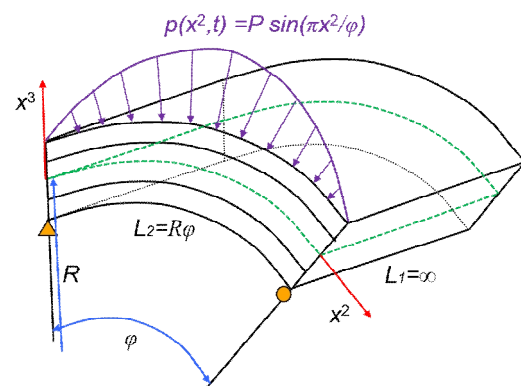


Fig. 2. Geometry and loading of cylindrical shell

where  $\mathbf{g}^3 = \mathbf{a}^3$  and  $\mathbf{a}_i = \mathbf{r}_{,i}$  are vector tangent to the surface coordinate curves. Hence, the displacement vector of the shell can be expressed as follows:

$$\vec{V}(x^i) = V_\alpha \bar{\mathbf{g}}^\alpha + V_3 \bar{\mathbf{g}}^3 = U_\alpha \bar{\mathbf{a}}^\alpha + U_3 \bar{\mathbf{a}}^3 \quad (6)$$

$$V_\alpha = \mu_\alpha^\beta U_\beta, \quad V_3 = U_3$$

where  $\mathbf{a}_i, \mathbf{g}_i$  are covariant base vectors and  $\mathbf{a}^i, \mathbf{g}^i$  are contravariant base vectors;  $\mu_\alpha^\beta$  denote the shifter tensor which is expressed as follows:

$$\bar{\mathbf{g}}_\alpha = \mu_\alpha^\beta \bar{\mathbf{a}}_\beta = (\delta_\alpha^\beta - x^3 b_\alpha^\beta) \bar{\mathbf{a}}_\beta \quad (7)$$

The metric tensor  $\mathbf{a}_{\alpha\beta}$  and surface tensor  $\mathbf{b}_{\alpha\beta}$  are defined from first and second fundamental forms respectively, as follows:

$$ds^2 = d\vec{r} \cdot d\vec{r} = a_{\alpha\beta} dx^\alpha dx^\beta \quad (8)$$

$$d\vec{r} \cdot d\bar{\mathbf{a}}_3 = -b_{\alpha\beta} dx^\alpha dx^\beta = -b_\beta^\alpha dx^\beta dx_\alpha$$

The in-plane displacement fields of laminated shells are assumed as the one of the original EHOST, which superimposes linear zigzag displacements, with different slope in each layer, as follows:

$$\begin{cases} U_\alpha(x^i, t) = u_\alpha(x^\alpha, t) + x^3 \psi_\alpha(x^\alpha, t) + (x^3)^2 \xi_\alpha(x^\alpha, t) + (x^3)^3 \phi_\alpha(x^\alpha, t) + \sum_{k=1}^{n-1} S_\alpha^{(k)}(x^\alpha, t) (x^3 - x^3_{(k)}) H(x^3 - x^3_{(k)}) \\ U_3(x^i, t) = w(x^\alpha, t) \end{cases} \quad (10)$$

where  $\mathbf{u}_\alpha$  and  $\mathbf{w}$  are displacements defined at the mid-surface, and  $\mathbf{H}(x^3 - x^3_{(k)})$  is the Heaviside unit step function. The requirement that the transverse stress should be vanished at the top and bottom surface, and continuous through the thickness of the shells are employed to reduce the number of unknown primary variables. Due to the advantage of Laplace transform, this work can be executed conveniently in Laplace domain, and the displacements can be obtained as follows:

$$\begin{cases} U_\alpha^*(x^i, s) = \mu_\alpha^\beta \mathbf{u}_\beta^*(x^\alpha, s) - x^3 \mathbf{u}_{3||\alpha}^*(x^\alpha, s) + f_\alpha^{\beta*}(x^3) \phi_\beta^*(x^\alpha, s) \\ U_3^*(x^i, s) = \mathbf{w}^*(x^\alpha, s) \end{cases} \quad (11)$$

where

$$f_\alpha^{\beta*}(x^3) = \nu_\alpha^{\beta*} x^3 + \kappa_\alpha^{\beta*} (x^3)^2 + \delta_\alpha^\beta (x^3)^3 + \sum_{k=1}^{N-1} (\varpi^k)_\alpha^{\beta*} (x^3 - x^3_{(k)}) H(x^3 - x^3_{(k)}) \quad (12)$$

where  $\nu_\alpha^{\beta*}, \kappa_\alpha^{\beta*}$  and  $(\varpi^k)_\alpha^{\beta*}$  denote the matrices in Laplace domain which depend on the material properties and the thickness of each ply.

Then, the converted virtual work principle () in Laplace domain is employed and integrated by part to obtain the equilibrium (), as follows:

$$\int_V (\sigma^{\alpha\beta} \delta e_{\alpha\beta} + 2\sigma^{\alpha 3} \delta e_{\alpha 3}) - \int_\Omega p \delta u^3 d\Omega = 0 \quad (13)$$

$$\delta u_\omega : (N^{1*})^\omega - (N^{2*})^{\omega\gamma} - (M^{1*})^\omega + (M^{2*})^{\omega\nu} = 0$$

$$\delta \phi_\omega : (R^{1*})^\omega - (R^{2*})^{\gamma\omega} + (V^{*})^{\omega 3} = 0 \quad (14)$$

$$\delta u_3 : N^{3*} + (M^{3*})^\omega - (M^{4*})^{\omega\nu} = p^*$$

where the stress resultants will be shown in final version.

### 3. Cylindrical bending for cylindrical shell

For cylindrical bending problem for cylindrical shallow shell as shown in Fig. 2, the coordinates are specified as  $(x^1, x^2, x^3) = (x, \theta, z)$ , the curvature teams will be reduced as follows:

$$\begin{aligned} \mu_\alpha^\beta &= \delta_\alpha^\beta, & g_{\alpha\beta} &= a_{\alpha\beta} = \delta_{\alpha\beta}, & \mu &= 1, \\ \bar{\Gamma}_{\alpha\beta}^v &= \mathbf{0}, & b_{\alpha\beta} &= b_\beta^\alpha = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{1}{R} \end{bmatrix} \end{aligned} \quad (15)$$

The sinusoidal loading is applied on the upper surface of laminated shell, as follows:

$$p_z(x, \theta, t) = P_z(t) \sin\left(\frac{\pi\theta}{\varphi}\right) \quad (16)$$

The simply supported boundary conditions can be satisfied by the following forms of displacements:

$$u_2^0 = U(t) \cos\left(\frac{\pi\theta}{\varphi}\right), \quad \phi_2 = \Phi(t) \cos\left(\frac{\pi\theta}{\varphi}\right), \quad w = W(t) \sin\left(\frac{\pi\theta}{\varphi}\right) \quad (17)$$

where  $\varphi$  is chosen as  $\pi/3$ .

Substituting the form of displacement in Eq. (17) into Eq. (1) with the help of resultants-displacements relations, we can obtain the algebraic relations between the primary variables and the external force in the Laplace domain as follows:

$$\begin{bmatrix} K_{11}^* & K_{12}^* & K_{13}^* \\ K_{21}^* & K_{22}^* & K_{23}^* \\ K_{31}^* & K_{32}^* & K_{33}^* \end{bmatrix} \begin{Bmatrix} U^* \\ \Phi^* \\ W^* \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ P^* \end{Bmatrix} \quad (18)$$

where  $K^*$  is the stiffness matrix in Laplace domain. The detail formulation of stiffness matrix will be given in the presentation. After solving the above equation, the primary variables as well as the deflection, in-plane displacements, stress distributions in real time domain can be obtained by applying the numerical inverse Laplace transform  $\mathcal{L}^{-1}$ .

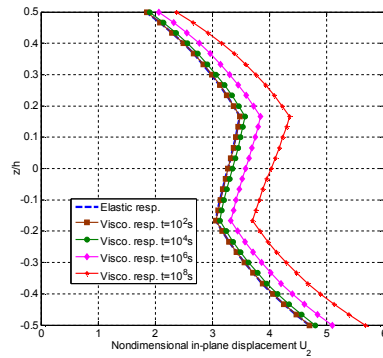
## Results and discussion

To compare with previous studies, the laminated shell (90/0/90), assumed to be made of Graphite/Epoxy GY70/339 at reference condition ( $T=75^\circ\text{C}$ ,  $M=0.1\%$ ), is chosen as numerical example with elastic properties are shown as follows:

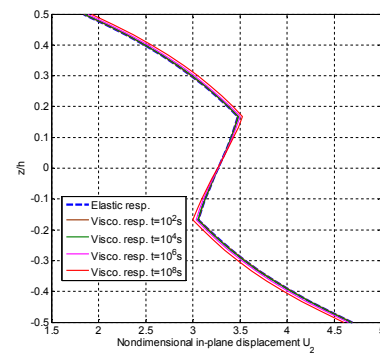
$$\begin{aligned} E_L &= 2.89 \times 10^5 \text{ MPa} & E_T &= 6.063 \times 10^3 \text{ MPa} \\ G_{LT} &= 4.134 \times 10^3 \text{ MPa} & G_{TT} &= 2.067 \times 10^3 \text{ MPa} \\ \nu_{LT} &= \nu_{TT} = 0.31 \end{aligned} \quad (19)$$

where  $L$  denotes a fiber direction and  $T$  denotes a perpendicular direction to the fiber. The time-dependent function  $\varphi(t)$  for GY70/339 determined by mastering curve Crossman's experimental data is omitted because of limited space [14]. In present study, for the static creep process, the applied force is kept as constant  $p(t)=1.0$ . For the relaxation process, after the force  $p|_{t=0}=1.0$  is applied, the deflection will be kept as constant as the time go by. In this version, the in-plane displacement distribution  $U_2$  is represented; the other stress-strain distributions through the thickness of the shell will be shown in the presentation.

The Fig. 3 shows the distribution of in-plane displacement  $U_2$  through the thickness for creep process. The viscoelastic solution at initial time is the same to elastic solution. Then, it changes as the time goes by, but still have good agreement with elastic one. Due to the decay of the stiffness of composite matrix, the amplitude of  $U_2$  increases respect to time. The increment at upper and lower



**Fig. 3. The in-plane displacement  $U_2$  distribution through the thickness of creep process**



**Fig. 4. The in-plane displacement  $U_2$  distribution through the thickness of relaxation process**

is smaller because of dominant elastic fiber. These responses make the structures in risks with larger deformation.

The Fig. 4 shows the distribution of in-plane displacement  $U_2$  through the thickness for relaxation process. The viscoelastic solution at initial time is the same as elastic solution. Then, similar with 1-D viscoelastic response of relaxation, the amplitude of  $U_2$  changes insignificantly. The amplitude lightly decreases at top and bottom of shell, and lightly increases at interface surface.

## Conclusion

The viscoelastic response of laminated composite plate of EHOST has been analyzed with the help of Laplace transform without any integral transformation or any time step scheme. The accuracy and efficiency of the analysis are retained since the viscoelastic EHOST formulation was made in the linear elastic regime in the Laplace transformed domain. The numerical results for composite laminated shell, which are converted by employing inversion Laplace transform, adequately show the change of time-dependent mechanical behaviors for creep process as well as relaxation process.

## Acknowledgement

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2012R1A3A2048841).

## References

- Chen, T. M. (1995) The hybrid Laplace transform/ finite element method applied to the quasi-static and dynamic analysis of viscoelastic Timoshenko beams, *International Journal for Numerical Methods Engineering* **38**, 509–522.
- Cho, M. and Parmerter, R. R. (1993) Efficient higher order composite plate theory for general Lamination Configurations, *AIAA Journal* **31**(7), 1299–1306.
- Cho, M. and Parmerter, R. R. (1994) Finite element for composite plate bending based on efficient higher order theory, *AIAA Journal* **32**(11), 2241–2248.
- Crossman, F. W., Mauri, R. E. and Warren, W. J. (1978) Moisture altered viscoelastic response of graphite/ epoxy composite, *Advanced Composite Materials Environmental Effects*, ASTM STP 658, 205–220.
- Dubner, H. and Abate, J. (1968) Numerical inversion of Laplace transforms by relating them to the finite Fourier cosine transform, *Journal of the Association for Computing Machinery* **15**(1), 115–123.
- Hilton, T. H. and Yi, S. (1993) Anisotropic viscoelastic finite element analysis of mechanically and hydrothermally loaded composite, *Composite Engineering* **3**(2), 123–135.
- Hassanzadeh, H. and Pooladi-Darvish, M. (2007) Comparison of different numerical Laplace inversion methods for engineering applications, *Applied Mathematics and Computation* **189**, 1966–1981.
- Kim, J. S. and Cho, M. (2003) Efficient higher-order shell theory for laminated composites with multiple delaminations, *AIAA Journal* **41**, 941–950.
- Lin, K. J. and Hwang, I. H. (1989) Thermo-viscoelastic analysis of composite materials, *Journal of Composite Materials* **23**, 554–569.

- Narayanan, G. V. and Beskos, D. E. (1982) Numerical operational methods for time-dependent linear problems, *International Journal for Numerical Methods in Engineering* **18**, 1829–1854.
- Nguyen, S-N, Lee, J. and Cho, M. (2015) Efficient higher-order zigzag theory for viscoelastic laminated composite plates, *International Journal of Solid and Structure* (accepted).
- Pagano, N. J. (1969) Exact solutions for composite laminates in cylindrical bending, *Journal of Composite Materials* **3**, 398–411.
- Reddy, J. N. (2004) *Mechanics of Laminated Composite Plates and Shells Theory and Analysis 2<sup>nd</sup>*, CRC Press LLC.
- Venkat V., Annie R. and Reddy, J. N. (2012) Nonlinear viscoelastic analysis of orthotropic beams using a general third-order theory, *Composite Structures* **94**, 3759–3768.
- Yi, S. and Hilton, H. H. (1995) Hygrothermal effects on viscoelastic responses of laminated composites, *Composites Engineering* **5**(2), 183–193.

## **A better way for managing all of the physical sciences under a single unified theory of analytical integration.**

**Mike Mikalajunas**

CIME, 38 Neuville, Montreal, Canada J7V 8L1

[michelmikalajunas@bellnet.ca](mailto:michelmikalajunas@bellnet.ca)

[jpnelson\\_mfc@yahoo.ca](mailto:jpnelson_mfc@yahoo.ca)

### **Abstract**

If a new system of computational logic would be entirely based on the application of a true unified *computational-based* analytical theory of integration then what better way of validating such a system of mathematical logic then through the complete development of a unified theory of physics. The outcome of having successfully arrived at such a monumental theory in physics would represent a much greater expansion of our knowledge in terms of engineering science. This would be the direct consequence of having analytically resolved under one unified theory of analytical integration the vast majority of PDEs some of which would prove very similar in appearance to those encountered in theoretical physics.

**Keywords:** Polynomial Transform, ODEs, PDEs, Analytical Integration, Engineering Science Theoretical & Experimental Physics, CERN.

### ***0. Introduction***

The new development for the physical science begins with the introduction of what appears to be a *computational-based* unified theory of *analytical* integration. This would eventually lead to the formation of some type of a new physical ideology by which a unified theory of physics could eventually be constructed over time.

Even *Albert Einstein* who has always claimed that God never plays with dices would have conceded if he were alive today, that somewhere out there in the vast realm of mathematics there has to exist some form of an algebraic system capable of addressing all of physics under one unified theory of "*computation*".

And I would like to quote from the ending of his book *The meaning of Relativity* [Einstein (1974)]:

*" This does not seem to be in accordance with continuum theory, and must lead to an attempt to find a purely algebraic theory for the description of reality. But nobody knows how to obtain the basis of such a theory "*.

With a unified theory of integration now possibly well within our grasp, this would represent a very important step towards becoming much less dependent on experimental scientific and engineering methods of analysis. As a result of this, we would be adopting a far more theoretical approach towards every aspect of the physical sciences on a much wider universal scale than ever thought possible under existing traditional methods of analysis.

*" The greatest problem in having to rely on traditional methods of mathematical analysis is mainly due to a severe lack of universality as a direct consequence of not having uncovered a unified theory of analytical integration in the past. "*

Throughout the remaining of this article, I will first and foremost attempt to briefly summarize in more *layman's term* an entirely new mathematical ideology. Next, I will proceed to demonstrate how such a very powerful new ideology in mathematics can be mutated into a whole new branch of physics that I would like to introduce everyone as an ***idealistic physics***.



As a direct consequence of what appears to be a unified theory of analytical integration, I would like to address the importance of arriving at some unified theory of physics for a much greater expansion of our knowledge in terms of engineering science.

***1. First and foremost, the new mathematical ideology***

What gives the new ideology its own and very unique flavor in mathematics is that all analysis pertaining to anti-differentiation whether in the form of a DE or an integral is always performed at the differential level. The main reason for this is to insure that the concept of continuity be preserved throughout the entire anti-differentiation process. Because the laws of algebra apply equally well to finite quantities as they do to differential quantities without regards to any limiting process near zero, there is never a risk of violating any known mathematical principles. The type of continuity I am referring to can ideally be described by a DE or a system of DEs.

Under the new proposed mathematical ideology when it comes to solving for any type of DE or system of DEs, rather than working with complete mathematical equations, we instead only become interested in working with complete *differential* form representations.

This is where the new mathematical ideology now begins to deviate from the old traditional thinking of Calculus.

If we were to have complete access to every imaginable type of mathematical equations just from the computed values alone that would originate from the application of some very unique mathematical ideology then this would certainly represent a very significant discovery in mathematics. This would no doubt represent an extremely valuable tool for completely eliminating our most fundamental problem of not being able to select the most suitable type of mathematical equation for handling *all* aspects of the physical and biological sciences under one "*unified theory of computation*". Over time, this would inevitably lead towards the development of some form of a unified theory of physics by which some type of a "*theory of everything*" would be constructed from.

The very first place we might want to look for the possible existence of such a potentially formidable mathematical theory of integration is in the following very simple integral equation:

$$t = \int \frac{dy}{ay^2 + by + c} \tag{1.001}$$

Everyone would certainly agree that only because "y(t)" was initially presented in its complete differential form, this has provided us with the capability of defining a wide range of mathematical expressions just by varying the numerical values present inside this integral. From this very simple observation in Calculus, we can immediately deduce that differential forms could at least potentially represent a very powerful link between *numerical computation* and *complete mathematical expressions*.

So our primary objective now is to determine what possible variations in terms of differential form representations can we expect for including "*all*" types of mathematical equations regardless of the degree of complexity involved. Such mathematical equations would be constructed from the use of algebraic and elementary basis functions that would involve the presence of composite functions with no limit whatsoever as to each of their degree of composition. Furthermore, there would be no restriction whatsoever on the number of dependent and independent variables involved and finally, the entire mathematical equation may be expressible not only in explicit form but also in implicit form as well.

Such an ideal universal differential expansion form can only be described mathematically in terms of *two* fundamental parts that would involve the use of *multivariate polynomials* as well as complete *differentials* of *multivariate polynomials*.

For a general system of "k" number of implicitly defined *multivariate* mathematical equations in the form of " $f_k(z_m, x_n) = 0$ " that consist of "m" number of dependent variables and "n" number of independent variables this may be described as follow:

(1). Primary Expansion:

$$F_i(W_j) = 0 = \sum_r a_{i,r} \left( \prod_j^p W_j^{E_{i,s}} \right) \quad (1 \leq i \leq k) \quad (1.002)$$

where " $W_j$ " are auxiliary variables, " $p$ " is the total number of such auxiliary variables each of which are raised to some floating point value and " $r$ " is the total number of terms present in each of the "k" number of implicitly defined multivariate polynomial equations.

(2). Secondary Differential Expansion:

$$dz_i = dW_i \quad (1 \leq i \leq m) \quad (1.003)$$

$$dx_i = dW_{m+i} \quad (1 \leq i \leq n) \quad (1.004)$$

$$\begin{aligned} \sum_{t=1}^m N_{i(m+n+1)-m-n-1+t} dz_t + \sum_{t=1}^n N_{i(m+n+1)-n-1+t} dx_t &= \\ = N_{i(m+n+1)} dW_j \quad [1 \leq i \leq p - m - n] \quad [m + n + 1 \leq j \leq p] & \quad (1.005) \end{aligned}$$

$$\begin{aligned} N_c(W_j) &= \sum_{t=(c-1)r+1}^{cr} b_{c,t} \left( \prod_j^p W_j^{E'_{c,s}} \right) \quad (1.006) \\ & \quad [1 \leq c \leq i(m+n+1)] \quad [1 \leq i \leq p - m - n] \end{aligned}$$

There is at present no other possible differential form capable of representing *all* mathematical equations with such a high degree of universality then the one suggested above.

In complete expanded form we would write this as follow:

(1). Primary Expansion:

$$\begin{aligned} F_1 = 0 &= a_{1,1} W_1^{m_{11}} W_2^{m_{12}} \dots W_p^{m_{1p}} + a_{1,2} W_1^{m_{1,p+1}} W_2^{m_{1,p+2}} \dots W_p^{m_{1,2p}} + \dots + \\ & \quad + \dots + a_{1,r} W_1^{m_{1,p(r-1)+1}} W_2^{m_{1,p(r-1)+2}} \dots W_p^{m_{1,rp}} \quad (1.007) \end{aligned}$$

$$F_2 = 0 = a_{2,1}W_1^{m_{2,1}}W_2^{m_{2,2}} \dots W_p^{m_{2,p}} + a_{2,2}W_1^{m_{2,p+1}}W_2^{m_{2,p+2}} \dots W_p^{m_{2,2p}} + \dots +$$

$$+ \dots + a_{2,r}W_1^{m_{2,p(r-1)+1}}W_2^{m_{2,p(r-1)+2}} \dots W_p^{m_{2,rp}} \quad (1.008)$$

. . .  
 . . .  
 . . .

$$F_k = 0 = a_{k,1}W_1^{m_{k,1}}W_2^{m_{k,2}} \dots W_p^{m_{k,p}} + a_{k,2}W_1^{m_{k,p+1}}W_2^{m_{k,p+2}} \dots W_p^{m_{k,2p}} + \dots +$$

$$+ \dots + a_{k,r}W_1^{m_{k,p(r-1)+1}}W_2^{m_{k,p(r-1)+2}} \dots W_p^{m_{k,rp}} \quad (1.009)$$

(2). Secondary Differential Expansion:

$$dz_i = dW_i \quad (1 \leq i \leq m) \quad (1.010)$$

$$dx_i = dW_{m+i} \quad (1 \leq i \leq n) \quad (1.011)$$

$$[N_1dz_1 + N_2dz_2 + \dots + N_mdz_m] + [N_{m+1}dx_1 + N_{m+2}dx_2 + \dots +$$

$$+ \dots + N_{m+n}dx_n] = N_{m+n+1}dW_{m+n+1} \quad (1.012)$$

$$[N_{m+n+2}dz_1 + N_{m+n+3}dz_2 + \dots + N_{2m+n+1}dz_m] + [N_{2m+n+2}dx_1 +$$

$$+ N_{2m+n+3}dx_2 + \dots + N_{2(m+n+1)-1}dx_n] = N_{2(m+n+1)}dW_{m+n+2} \quad (1.013)$$

. . .  
 . . .  
 . . .  
 . . .

$$\begin{aligned}
 & \left[ N_{(p-1)(m+n+1)+1}dz_1 + N_{(p-1)(m+n+1)+2}dz_2 + \dots + N_{(p-1)(m+n+1)+m}dz_m \right] + \\
 & + \left[ N_{(p-1)(m+n+1)+m+1}dx_1 + N_{(p-1)(m+n+1)+m+2}dx_2 + \dots + N_{p(m+n+1)-1}dx_n \right] = \\
 & = N_{p(m+n+1)}dW_p \quad (1.014)
 \end{aligned}$$

Assuming a system of implicitly defined mathematical equations consisting of 3 dependent variables and 5 independent variables with a total number of 12 auxiliary variables.

We will determine the correct index value for each of the multivariate polynomials present inside the *Secondary Differential Expansion* that would be responsible for defining the complete expression for the 10<sup>th</sup> auxiliary variable:

Starting with equation (1.005):

$$\begin{aligned}
 & \sum_{t=1}^m N_{i(m+n+1)-m-n-1+t}dz_t + \sum_{t=1}^n N_{i(m+n+1)-n-1+t}dx_t = \\
 & = N_{i(m+n+1)}dW_j \quad [1 \leq i \leq p - m - n] \quad [m + n + 1 \leq j \leq p] \quad (1.015)
 \end{aligned}$$

The results are for "p=12", "m=3", "n=5" we have "i = 12 - 10 = 2" and "j = 10":

$$\sum_{t=1}^3 N_{2(3+5+1)-3-5-1+t}dz_t + \sum_{t=1}^5 N_{2(3+5+1)-5-1+t}dx_t = N_{2(m+n+1)}dW_{10} \quad (1.016)$$

$$\begin{aligned}
 & [N_{10}dz_1 + N_{11}dz_2 + N_{12}dz_3] + [N_{13}dx_1 + N_{14}dx_2 + N_{15}dx_3 + \\
 & + N_{16}dx_4 + N_{17}dx_5] = N_{18}dW_{10} \quad (1.017)
 \end{aligned}$$

Before proceeding any further, a few simple mathematical definitions need to be in order.

The first one, the actual process of transforming a complete mathematical equation in terms of the above universal differential form representation is referred to as taking its *Multivariate Polynomial Transform*.

Next, the complete *reverse* process of going from a differential form representation back to its original complete mathematical equation would be referred to as taking the *inverse* of a *Multivariate Polynomial Transform*. This would require following a very unique integration process to be des-

cribed later for determining the complete analytical expression corresponding to each auxiliary variable " $W_j$ ". They each in turn would be substituting back into the *Primary Expansion* for arriving at the complete original expression that we started with being in the form of " $f_k(z_m, x_n) = 0$ ".

As we are dealing mainly with multivariate polynomials and complete differentials of multivariate polynomials, new types of coefficients are being introduced along the way. During the process of inverting from a differential form back to the original complete mathematical equation, some of these coefficients would be entirely responsible for defining the basis functions by which the complete mathematical equation was originally constructed from. These particular types of coefficients are present only in the *Secondary Differential Expansion* of a *Multivariate Polynomial Transform*. The remaining types of coefficients will be described in more detail later on.

**Example (1.1).** Let us consider the simplest *two* dimensional case which would correspond to the case for " $k = m = n = 1$ " and by replacing the dependent variable " $z$ " with " $y$ ", we arrive at the following corresponding general *Multivariate Polynomial Transform* for " $y(x)$ ":

(1). Primary Expansion:

$$F(W_j) = 0 \quad [1 \leq j \leq p] \quad (1.018)$$

(2). Secondary Differential Expansion:

$$dx = dW_1 \quad (1.019)$$

$$dy = dW_2 \quad (1.020)$$

$$N_{3i-2}dx + N_{3i-1}dy = N_{3i}dW_j \quad [1 \leq i \leq p - 2] [3 \leq j \leq p] \quad (1.021)$$

For this general univariate two dimensional case, the *Secondary Differential Expansion* may be written in the following more general format upon replacing each auxiliary variable on the left hand side with the dependent and independent variables:

$$M(x, y)dx + N(x, y)dy = P(W_j)dW_j \quad (1.022)$$

The left hand side of this equation appears in exactly the same format by which all first order ODEs are written prior to applying Euler's method for specifically targeting those that are considered as exact differentials.

This test is well known in Calculus and is defined by:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} \quad (1.023)$$

When this condition is met than Euler's general integral formula can then be applied and the result is a vastly simplified integration process.

The formula has two equivalent form.

The first:

$$\int Mdx + \int \left( N - \frac{\partial}{\partial y} \int Mdx \right) dy = C \quad (1.024)$$

and the second one:

$$\int Ndy + \int \left( M - \frac{\partial}{\partial x} \int Ndy \right) dx = C \quad (1.025)$$

**Example (1.2).** We will apply the concept of an exact differential for demonstrating in detail the exact process involved for inverting the following *Multivariate Polynomial Transform* corresponding to a univariate implicitly defined equation in two dimension.

(1). Primary Expansion:

$$F(W_1, W_2, W_3, W_4) = 0 = W_4 + 2W_2 \quad (1.026)$$

(2). Secondary Differential Expansion:

$$dx + 0 \cdot dy = dW_1 \quad (1.027)$$

$$0 \cdot dx + dy = dW_2 \quad (1.028)$$

$$-2W_1 dx + 0 \cdot dy = W_3 dW_3 \quad (1.029)$$

$$2W_1 dx - W_3 dy = W_3(W_2 + W_3) dW_4 \quad (1.030)$$

The first step is to naturally begin by integrating in ascending order of complexity each first order ODE that is present in the *Secondary Differential Expansion* for the expression of each auxiliary variable.

We begin first by defining " $W_1(x) = x$ " and " $W_2(y) = y$ ".

For " $W_3(x)$ ", we integrate equation (1.029) by parts to arrive at:

$$W_3(x) = \pm \sqrt{C_3 - 2x^2} \quad (1.031)$$

For " $W_4(x, y)$ ", the corresponding first order ODE to integrate is obtained by substituting the expression for " $W_1(x)$ " and " $W_2(y)$ " into (1.030) to afterwards rearrange the resultant equation in the form of:

$$\frac{2x dx}{W_3(y + W_3)} - \frac{dy}{y + W_3} = dW_4 \quad (1.032)$$

Let:

$$M(x, y) = \frac{2x}{W_3(y + W_3)} \quad (1.033)$$

so that since " $W_3 = W_3(x)$ ":

$$\frac{\partial M}{\partial y} = \frac{-2x}{W_3(y + W_3)^2} \quad (1.034)$$

Next, define:

$$N(x, y) = \frac{-1}{(y + W_3)} \quad (1.035)$$

so that:

$$\frac{\partial N}{\partial x} = \frac{1}{(y + W_3)^2} \frac{dW_3}{dx} \quad (1.036)$$

From equation (1.029):

$$\frac{dW_3}{dx} = \frac{-2W_1}{W_3} = \frac{-2x}{W_3} \quad (1.037)$$

Substituting this equation into equation (1.036), we get:

$$\frac{\partial N}{\partial x} = \frac{-2x}{W_3(y + W_3)^2} \quad (1.038)$$

Since:

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x} = \frac{-2x}{W_3(y + W_3)^2} \quad (1.039)$$

it follows that equation (1.030) is an exact differential whose solution may be obtained using any one of Euler's integral formula mentioned earlier in equation (1.024) and (1.025).

The following general form will be used:

$$\int Ndy + \int \left( M - \frac{\partial}{\partial x} \int Ndy \right) dx = c \quad (1.040)$$

For the first integral:

$$\int Ndy = \int \frac{-dy}{y + W_3} = -\ln(y + W_3) \quad (1.041)$$

For the second integral:

$$\frac{\partial}{\partial x} \int N dy = - \frac{\partial}{\partial x} \ln(y + W_3) = - \frac{1}{y + W_3} \frac{dW_3}{dx} \quad (1.042)$$

From the differential that defined the third auxiliary variable as given by equation (1.029), we can write:

$$\frac{dW_3}{dx} = \frac{-2x}{W_3} \quad (1.043)$$

Thus equation (1.042) may be rewritten as follow:

$$\frac{\partial}{\partial x} \int N dy = - \frac{1}{y + W_3} \frac{dW_3}{dx} = \frac{2x}{W_3(y + W_3)} = \mathbf{M}(x, y) \quad (1.044)$$

so that :

$$\int \left( M - \frac{\partial}{\partial x} \int N dy \right) dx = \int (M - \mathbf{M}) dx = 0 \quad (1.045)$$

Euler's integral formula may now be rewritten in the following final form:

$$\int N dy + \int \left( M - \frac{\partial}{\partial x} \int N dy \right) dx = \int N dy = \int \frac{-dy}{y + W_3} = -\ln(y + W_3) \quad (1.046)$$

The complete exact solution of the differential form that define " $W_4$ " is obtained by integrating equation (1.030) using the above integral solution.

The results are:

$$-\ln(y + W_3) = W_4 + K \quad (1.047)$$

Substituting the expression for " $W_3(x)$ " as defined by equation (1.031) into the above equation, we obtain:

$$-\ln(y \pm \sqrt{C_3 - 2x^2}) = W_4 + K \quad (1.048)$$

Solving for " $W_4$ " :

$$W_4(x, y) = C_4 - \ln(y \pm \sqrt{C_3 - 2x^2}) \quad (1.049)$$

The complete inverse *Multivariate Polynomial Transform* of the given implicitly defined equation is obtained by substituting the expression for " $W_1(x)$ ", " $W_2(y)$ ", " $W_3(x)$ " and " $W_4(x, y)$ " into the *Primary Expansion* defined by equation (1.026).



The results are:

$$f(x, y) = 0 = C_4 - \ln\left(y \pm \sqrt{C_3 - 2x^2}\right) + 2y \quad (1.050)$$

where the constants of integration defined by " $C_3$ " and " $C_4$ " are each determined from the initial condition of " $f(x, y) = 0$ ".

**I**n *higher dimension* than two, the basic principle behind the main test for exactness is still applicable but requires some very minor modifications in order to account for the multivariate nature of the corresponding general differential form representation.

In view of equation (1.002) through (1.006), an example of a *single* first order multivariate ODE that can be present inside a *Secondary Differential Expansion* may be expressed in the following general form:

$$\begin{aligned} (M_1 dz_1 + M_2 dz_2 + \dots + M_m dz_m) + (M_{m+1} dx_1 + M_{m+2} dx_2 + \dots + \\ + \dots + M_{m+n} dx_n) = M_{m+n+1} dW_j \end{aligned} \quad (1.051)$$

where as by eliminating each auxiliary variable in terms of the dependent and independent variables on the left hand side of this equation, we can also define:

$$M_i = M_i(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n) \quad (1 \leq i \leq m+n) \quad (1.052)$$

The *right hand* side of this equation can be expressed only in terms of the auxiliary variable " $W_j$ " so that:

$$M_i = M_i(W_j) \quad (i = m+n+1) \quad (1.053)$$

The auxiliary variable " $W_j$ " is actually a "*multivariate composite function*" and is to be determined assuming of course that an exact expression for each of the auxiliary variables " $W_1, W_2, \dots, W_{j-1}$ " have all been previously obtained in ascending order of complexity.

Equation (1.051) may be rewritten as:

$$dH_1 = dH_2 \quad (1.054)$$

where:

$$\begin{aligned} dH_1 = (M_1 dz_1 + M_2 dz_2 + \dots + M_m dz_m) + (M_{m+1} dx_1 + M_{m+2} dx_2 + \dots + \\ + \dots + M_{m+n} dx_n) \end{aligned} \quad (1.055)$$

and where:

$$dH_2 = M_{m+n+1} W_j \quad (1.056)$$

If each side of equation (1.054) is an exact differential then from the chain rule:

$$dH_1 = \sum_{k=1}^m \frac{\partial H_1}{\partial z_k} dz_k + \sum_{k=1}^n \frac{\partial H_1}{\partial x_k} dx_k \quad (1.057)$$

and since " $H_2 = H_2(W_j)$ " :

$$dH_2 = \frac{\partial H_2}{\partial W_j} dW_j \quad (1.058)$$

It follows from equation (1.055) and (1.057) that :

$$M_1 = \frac{\partial H_1}{\partial z_1} \quad (1.059)$$

$$M_2 = \frac{\partial H_1}{\partial z_2} \quad (1.060)$$

⋮  
⋮  
⋮

$$M_m = \frac{\partial H_1}{\partial z_m} \quad (1.061)$$

$$M_{m+1} = \frac{\partial H_1}{\partial x_1} \quad (1.062)$$

$$M_{m+2} = \frac{\partial H_1}{\partial x_2} \quad (1.063)$$

⋮  
⋮  
⋮

$$M_{m+n} = \frac{\partial H_1}{\partial x_n} \quad (1.064)$$

It also follows from equation (1.056) and (1.058) that:

$$M_{m+n+1} = \frac{\partial H_2}{\partial W_j} \tag{1.065}$$

From multivariate calculus, the condition that both sides of equation (1.054) each define an exact differential is of course when:

$$\frac{\partial M_1}{\partial z_2} = \frac{\partial M_2}{\partial z_1}, \quad \frac{\partial M_1}{\partial z_3} = \frac{\partial M_3}{\partial z_1}, \quad \frac{\partial M_1}{\partial z_4} = \frac{\partial M_4}{\partial z_1}, \quad \dots, \quad \frac{\partial M_1}{\partial z_m} = \frac{\partial M_m}{\partial z_1} \tag{1.066}$$

$$\frac{\partial M_1}{\partial x_1} = \frac{\partial M_{m+1}}{\partial z_1}, \quad \frac{\partial M_1}{\partial x_2} = \frac{\partial M_{m+2}}{\partial z_1}, \quad \dots, \quad \frac{\partial M_1}{\partial x_n} = \frac{\partial M_{m+n}}{\partial z_1} \tag{1.067}$$

$$\frac{\partial M_2}{\partial z_3} = \frac{\partial M_3}{\partial z_2}, \quad \frac{\partial M_2}{\partial z_4} = \frac{\partial M_4}{\partial z_2}, \quad \frac{\partial M_2}{\partial z_5} = \frac{\partial M_5}{\partial z_2}, \quad \dots, \quad \frac{\partial M_2}{\partial z_m} = \frac{\partial M_m}{\partial z_2} \tag{1.068}$$

$$\frac{\partial M_2}{\partial x_1} = \frac{\partial M_{m+1}}{\partial z_2}, \quad \frac{\partial M_2}{\partial x_2} = \frac{\partial M_{m+2}}{\partial z_2}, \quad \dots, \quad \frac{\partial M_2}{\partial x_n} = \frac{\partial M_{m+n}}{\partial z_2} \tag{1.069}$$

$$\begin{matrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

$$\frac{\partial M_m}{\partial x_n} = \frac{\partial M_{m+n}}{\partial z_m} \tag{1.070}$$

If each of the above conditions are met then the solution for "H<sub>1</sub>" and "H<sub>2</sub>" is obtained as follow:

For "H<sub>1</sub>", we integrate equation (1.059) :

$$H_1 = \int M_1(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n) \partial z_1 \tag{1.071}$$

where in this case, z<sub>i</sub> and x<sub>j</sub> for 1 < i ≤ m, 1 ≤ j ≤ n and i ≠ 1 are all treated as constants when evaluating this indefinite integral.

We can also use as another alternative:

$$H_1 = \int M_k(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n) \partial z_k \quad (1.072)$$

where in this case,  $z_i$  and  $x_j$  for  $2 \leq i \leq m$ ,  $1 \leq j \leq n$  and  $i \neq k$  are all treated as constants when evaluating this indefinite integral.

Other alternatives for the same expression of " $H_1$ " can also be obtained from:

$$H_1 = \int M_{m+k}(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n) \partial x_k \quad (1.073)$$

where in this case,  $z_i$  and  $x_j$  for  $1 \leq i \leq m$ ,  $1 \leq j \leq n$  and  $j \neq k$  are all treated as constants when evaluating this indefinite integral.

As for the expression of " $H_2$ " defined by equation (1.056), it can be determined using the following integral :

$$H_2 = H_2(W_j) = \int M_{m+n+1}(W_j) dW_j \quad (1.074)$$

because " $W_j = W_j(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n)$ " is a multivariate composite function.

The complete exact solution of the first order multivariate ODE defined by equation (1.054) that would be present inside a *Secondary Differential Expansion* is:

$$H_1(z_1, z_2, \dots, z_m, x_1, x_2, \dots, x_n) - H_2(W_j) = 0 \quad (1.075)$$

from which " $W_j$ " can be obtained explicitly whenever possible.

Once the complete expression for each auxiliary variable is obtained, they can afterwards be substituted along with each of their initial condition(s) into the *Primary Expansion* for arriving at the required system of implicitly defined equations in the form of " $f_k(z_i, x_j) = 0$ " for  $1 \leq i \leq m$  and  $1 \leq j \leq n$ .

The initial condition(s) that belong to each auxiliary variable all take part in satisfying the initial condition(s) of a system implicitly defined equations that can be used for completely representing the exact or approximate analytical solution of a *system of PDEs*.

**F**or inverting a *Multivariate Polynomial Transform* defined in much higher dimension follows the same type of logic as for the simple two dimensional case. The following example illustrates this in greater detail.

**Example (1.3).** Assuming the following *Secondary Differential Expansion* as a part of a large *Multivariate Polynomial Transform* that would correspond to some large system of implicitly defined equations involving several dependent variables and one single independent variable. Furthermore, for the sake of simplicity, let us assume that every first order ODE present in the

*Secondary Differential Expansion* would satisfy the condition for exactness everywhere and thus readily integrable using the method described earlier.

*Secondary Differential Expansion:*

$$dx + 0 \cdot dy_1 + 0 \cdot dy_2 + 0 \cdot dy_3 = dW_1 \quad (1.076)$$

$$0 \cdot dx + dy_1 + 0 \cdot dy_2 + 0 \cdot dy_3 = dW_2 \quad (1.077)$$

$$0 \cdot dx + 0 \cdot dy_1 + dy_2 + 0 \cdot dy_3 = dW_3 \quad (1.078)$$

$$0 \cdot dx + 0 \cdot dy_1 + 0 \cdot dy_2 + dy_3 = dW_4 \quad (1.079)$$

$$W_1^2 dx + W_2^2 dy_1 + W_3^2 dy_2 + W_4^2 dy_3 = W_5^2 dW_5 \quad (1.080)$$

$$W_1 dx + W_2 dy_1 + W_3 dy_2 + 0 \cdot dy_3 = W_6 dW_6 \quad (1.081)$$

$$\begin{aligned} W_1 W_6^{-1} dx + (W_2 W_6^{-1} + 2W_3) dy_1 + (W_3 W_6^{-1} + 2W_2) dy_2 + 0 \cdot dy_3 = \\ = \frac{dW_7}{1 + W_7^2} \end{aligned} \quad (1.082)$$

subjected to " $y_1(x_0)$ ", " $y_2(x_0)$ " and " $y_3(x_0)$ "

We will now determine its complete inverse where it is assumed that for the sake of simplicity each first order ODE present in the above *Secondary Differential Expansion* have already been factored out in order to filter out any unnecessary multivariate polynomials. These do not contribute in any manner on the overall integration process as they would tend to naturally cancel each other out by computation.

The first step is to naturally begin by integrating in ascending order of complexity each first order ODE present in the *Secondary Differential Expansion* for an expression of each auxiliary variable.

The results are:

$$\text{For "W}_1\text{"}, W_1(x) = x \quad (1.083)$$

$$\text{For "W}_2\text{"}, W_2(y_1) = y_1 \quad (1.084)$$

$$\text{For "W}_3\text{"}, W_3(y_2) = y_2 \quad (1.085)$$

$$\text{For "W}_4\text{"}, W_4(y_3) = y_3 \quad (1.086)$$

For " $W_5(x, y_1, y_2, y_3)$ ", this is more involved.

Equation (1.080) can be rewritten as:

$$M_1 dx + M_2 dy_1 + M_3 dy_2 + M_4 dy_3 = M_5 dW_5 \quad (1.087)$$

where:

$$M_1 = W_1^2 = x^2 \quad (1.088)$$

$$M_2 = W_2^2 = y_1^2 \quad (1.089)$$

$$M_3 = W_3^2 = y_2^2 \quad (1.090)$$

$$M_4 = W_4^2 = y_3^2 \quad (1.091)$$

$$M_5 = W_5^2 \quad (1.092)$$

Since " $M_1 = M_1(x)$ ", " $M_2 = M_2(y_1)$ ", " $M_3 = M_3(y_2)$ ", and " $M_4 = M_4(y_3)$ ", our test for exactness using equation (1.066) through (1.070) reveals that:

$$\frac{\partial M_2}{\partial x} = \frac{\partial M_1}{\partial y_1} = 0 \quad (1.093)$$

$$\frac{\partial M_3}{\partial x} = \frac{\partial M_1}{\partial y_2} = 0 \quad (1.094)$$

$$\frac{\partial M_4}{\partial x} = \frac{\partial M_1}{\partial y_3} = 0 \quad (1.095)$$

$$\frac{\partial M_3}{\partial y_1} = \frac{\partial M_2}{\partial y_2} = 0 \quad (1.096)$$

$$\frac{\partial M_4}{\partial y_1} = \frac{\partial M_2}{\partial y_3} = 0 \quad (1.097)$$

$$\frac{\partial M_4}{\partial y_2} = \frac{\partial M_3}{\partial y_3} = 0 \quad (1.098)$$

so that equation (1.087) is an exact differential equation with solution:

$$\int (x^2 dx + y_1^2 dy_1 + y_2^2 dy_2 + y_3^2 dy_3) = \int W_5^2 dW_5 \quad (1.099)$$

or:

$$W_5 = \sqrt[3]{x^3 + y_1^3 + y_2^3 + y_3^3 + c_5} \quad (1.100)$$

For " $W_6(x, y_1, y_2)$ ", equation (1.081) can be rewritten as:

$$M_1 dx + M_2 dy_1 + M_3 dy_2 = M_4 dW_6 \quad (1.101)$$

where:

$$M_1 = W_1 = x \quad (1.102)$$

$$M_2 = W_2 = y_1 \quad (1.103)$$

$$M_3 = W_3 = y_2 \quad (1.104)$$

$$M_4 = W_6 \quad (1.105)$$

Since " $M_1 = M_1(x)$ ", " $M_2 = M_2(y_1)$ " and " $M_3 = M_3(y_2)$ ", our test for exactness using equation (1.066) through (1.070) reveals that:

$$\frac{\partial M_2}{\partial x} = \frac{\partial M_1}{\partial y_1} = 0 \quad (1.106)$$

$$\frac{\partial M_3}{\partial x} = \frac{\partial M_1}{\partial y_2} = 0 \quad (1.107)$$

$$\frac{\partial M_3}{\partial y_1} = \frac{\partial M_2}{\partial y_2} = 0 \quad (1.108)$$

so that equation (1.101) is an exact differential equation with solution:

$$\int (x dx + y_1 dy_1 + y_2 dy_2) = \int W_6 dW_6 \quad (1.109)$$

or:

$$W_6 = \sqrt{x^2 + y_1^2 + y_2^2 + c_6} \quad (1.110)$$

For " $W_7(x, y_1, y_2)$ ", equation (1.082) can be rewritten as:

$$dH_1 = dH_2 \quad (1.111)$$

where:

$$dH_1 = M_1 dx + M_2 dy_1 + M_3 dy_2 \quad (1.112)$$

$$M_1 = W_1 W_6^{-1} \quad (1.113)$$

$$M_2 = W_2 W_6^{-1} + 2W_3 \quad (1.114)$$

$$M_3 = W_3 W_6^{-1} + 2W_2 \quad (1.115)$$

and:

$$dH_2 = M_4(W_7) dW_7 \quad (1.116)$$

$$M_4 = \frac{1}{1 + W_7^2} \quad (1.117)$$

It follows that:

$$\frac{\partial M_1}{\partial y_1} = \frac{\partial W_1}{\partial y_1} W_6^{-1} + W_1(-W_6^{-2}) \frac{\partial W_6}{\partial y_1} = 0 - W_1 W_6^{-2} (W_2 W_6^{-1}) \quad (1.118)$$

$$= -W_1 W_2 W_6^{-3} \quad (1.119)$$

$$\frac{\partial M_1}{\partial y_2} = \frac{\partial W_1}{\partial y_2} W_6^{-1} + W_1(-W_6^{-2}) \frac{\partial W_6}{\partial y_2} = 0 - W_1 W_6^{-2} (W_3 W_6^{-1}) \quad (1.120)$$

$$= -W_1 W_3 W_6^{-3} \quad (1.121)$$

$$\frac{\partial M_2}{\partial x} = \frac{\partial W_2}{\partial x} W_6^{-1} + W_2(-W_6^{-2}) \frac{\partial W_6}{\partial x} + 2 \frac{\partial W_3}{\partial x} = 0 - W_2 W_6^{-2} (W_1 W_6^{-1}) + 0 \quad (1.122)$$

$$= -W_1 W_2 W_6^{-3} \quad (1.123)$$

$$\frac{\partial M_2}{\partial y_2} = \frac{\partial W_2}{\partial y_2} W_6^{-1} + W_2(-W_6^{-2}) \frac{\partial W_6}{\partial y_2} + 2 \frac{\partial W_3}{\partial y_2} = 0 - W_2 W_6^{-2} (W_3 W_6^{-1}) + 2 \quad (1.124)$$

$$= -W_2 W_3 W_6^{-3} + 2 \quad (1.125)$$

$$\frac{\partial M_3}{\partial x} = \frac{\partial W_3}{\partial x} W_6^{-1} + W_3(-W_6^{-2}) \frac{\partial W_6}{\partial x} + 2 \frac{\partial W_2}{\partial x} = 0 - W_3 W_6^{-2} (W_1 W_6^{-1}) + 0 \quad (1.126)$$

$$= -W_1 W_3 W_6^{-3} \quad (1.127)$$

$$\frac{\partial M_3}{\partial y_1} = \frac{\partial W_3}{\partial y_1} W_6^{-1} + W_3(-W_6^{-2}) \frac{\partial W_6}{\partial y_1} + 2 \frac{\partial W_2}{\partial y_1} = 0 - W_3 W_6^{-2} (W_2 W_6^{-1}) + 2 \quad (1.128)$$

$$= -W_2 W_3 W_6^{-3} + 2 \quad (1.129)$$

Our test for exactness using equation (1.066) and (1.070) reveals that:

$$\frac{\partial M_2}{\partial x} = \frac{\partial M_1}{\partial y_1} = -W_1 W_2 W_6^{-3} \quad (1.130)$$

$$\frac{\partial M_3}{\partial x} = \frac{\partial M_1}{\partial y_2} = -W_1 W_3 W_6^{-3} \quad (1.131)$$

$$\frac{\partial M_3}{\partial y_1} = \frac{\partial M_2}{\partial y_2} = -W_2 W_3 W_6^{-3} + 2 \quad (1.132)$$

Furthermore:



$$M_4 = M_4(W_7) \quad (1.133)$$

so that equation (1.112) is an exact differential equation with solution:

$$\int dH_1 = \int M_1 \partial x = \int \frac{W_1}{W_6} \partial x \quad (1.134)$$

$$= \int \frac{x}{\sqrt{x^2 + y_1^2 + y_2^2 + C_6}} \partial x \quad (1.135)$$

Solving for " $H_1$ ":

$$H_1 = \sqrt{x^2 + y_1^2 + y_2^2 + C_6} + f_1(y_1, y_2) \quad (1.136)$$

$$= W_6 + f_1(y_1, y_2) \quad (1.137)$$

We can also define as a second alternative for " $H_1$ " the following integral equation:

$$\int dH_1 = \int M_2 \partial y_1 = \int \left( \frac{W_2}{W_6} + 2W_3 \right) \partial y_1 \quad (1.138)$$

$$= \int \left( \frac{y_1}{\sqrt{x^2 + y_1^2 + y_2^2 + C_6}} + 2y_2 \right) \partial y_1 \quad (1.139)$$

so that:

$$H_1 = \sqrt{x^2 + y_1^2 + y_2^2 + C_6} + 2y_1y_2 + f_2(x, y_2) \quad (1.140)$$

$$= W_6 + 2W_2W_3 + f_2(x, y_2) \quad (1.141)$$

A third alternative for " $H_1$ " can be derived from:

$$\int dH_1 = \int M_3 \partial y_2 = \int \left( \frac{W_3}{W_6} + 2W_2 \right) \partial y_2 \quad (1.142)$$

$$= \int \left( \frac{y_2}{\sqrt{x^2 + y_1^2 + y_2^2 + C_6}} + 2y_1 \right) \partial y_2 \quad (1.143)$$

so that:

$$H_1 = \sqrt{x^2 + y_1^2 + y_2^2 + C_6} + 2y_1y_2 + f_3(x, y_2) \quad (1.144)$$

$$= W_6 + 2W_2W_3 + f_3(x, y_1) \quad (1.145)$$

From equation (1.140) and (1.144) we arrive at the conclusion that:

$$f_2(x, y_2) = f_3(x, y_1) \quad (1.146)$$

The only condition for this equation to be satisfied is of course when:

$$f_2 = f_3 = F(x) \quad (1.147)$$

because " $y_1 \neq y_2$ ".

Substituting equation (1.147) into equation (1.145), we obtain:

$$H_1 = W_6 + 2W_2W_3 + F(x) \quad (1.148)$$

Since " $f_1(y_1, y_2)$ " in equation (1.137) is not a function of " $x$ " then it is safe to assume in equation (1.147) that:

$$F(x) = 0 \quad (1.149)$$

Substituting the expression for " $W_2$ ", " $W_3$ ", " $W_6$ " and " $F(x)$ " into equation (1.148), the expression for " $H_1$ " can now be completely defined as:

$$H_1 = W_6 + 2y_1y_2 + 0 \quad (1.150)$$

$$= \sqrt{x^2 + y_1^2 + y_2^2 + C_6} + 2y_1y_2 \quad (1.151)$$

The expression for " $H_2$ " can be determined by integrating equation (1.116) using equation (1.117):

$$H_2 = \int M_4 dW_7 = \int \frac{dW_7}{1 + W_7^2} = \tan^{-1}(W_7) + K \quad (1.152)$$

Since from equation (1.111) " $H_1 = H_2$ " we thus arrive at the following complete expression for " $W_7$ ":

$$\sqrt{x^2 + y_1^2 + y_2^2 + C_6} + 2y_1y_2 = \tan^{-1}(W_7) + K \quad (1.153)$$

or:

$$W_7 = \tan\left(\sqrt{x^2 + y_1^2 + y_2^2 + c_6} + 2y_1y_2 + c_7\right) \quad (1.154)$$

The complete inverse of the *Multivariate Polynomial Transform* whose *Secondary Differential Expansion* is defined by equation (1.076) through (1.082) is obtained by substituting each of the expression for the auxiliary variables " $W_1(x)$ ", " $W_2(y_1)$ ", " $W_3(y_2)$ ", " $W_4(y_3)$ ", " $W_5(x, y_1, y_2, y_3)$ ", " $W_6(x, y_1, y_2)$ " and " $W_7(x, y_1, y_2)$ " into a *Primary Expansion* that could be described in the following general form:

$$F_k(W_j) = 0 \quad (1 \leq k \leq 3) \quad (1 \leq j \leq 7) \quad (1.155)$$

## 2. Complete analytical theory of integration under one universal system of computational logic

The universal representation of all mathematical equations presented in the differential expansion form described by equation (1.002) through (1.006) should really be referred to as a general *initially assumed Multivariate Polynomial Transform* (IAMPT) when it comes to solving for DEs and systems of DEs. The only difference between traditional methods of series expansion and the one presented here, is that ours can succeed in arriving at complete *exact* analytical solution to any type of DEs and systems of DEs. All other known traditional methods of series solutions are incapacitated right from the beginning for arriving at *exact* analytical solutions since they were originally meant only to be utilized as part of some functional approximation theory. This being the direct consequence for all tradition methods of series solutions for not having originated from the application of some form of a unified theory of integration.

For those functional expressions that are present inside a DE or a system of DEs, they somehow would have to be totally accounted for in our initially assumed *Multivariate Polynomial Transform*. This is made possible only if we append at the end of our initially assumed expansion the *Multivariate Polynomial Transform* of each functional expression by introducing additional new supplemental auxiliary variables. Each of these additional auxiliary variables in turn are most likely to reappear in the final analytical solution of the DE or system of DEs. This would thus providing us with a real sense of measure in the manner by which such individual functional expressions can succeed in influencing the complete behavior of a physical system.

For including these types of DEs and systems of DEs, our general initially assumed *Multivariate Polynomial Transform* would have to be modified accordingly as follow:

### (1). Primary Expansion:

$$F_i(W_j) = 0 = \sum_r a_{i,r} \left( \prod_j^{p+q} W_j^{E_{i,s}} \right) \quad (1 \leq i \leq k) \quad (2.01)$$

where " $W_j$ " are auxiliary variables, " $q$ " is the total number of auxiliary variables required for defining the *Multivariate Polynomial Transform* of each functional expression that is present in a DE or a system of DEs. The total number of auxiliary variables now grows from " $p$ " to " $p + q$ " when functional expressions are present in these types of DEs. Each of the " $p$ " number of auxiliary variables are always assumed raised to some floating point value and finally, " $r$ " is the total number of terms present in each of the " $k$ " number of implicitly defined multivariate polynomial equations.

### (2). Secondary Differential Expansion:

$$dz_i = dW_i \quad (1 \leq i \leq m) \quad (2.02)$$

$$dx_i = dW_{m+i} \quad (1 \leq i \leq n) \quad (2.03)$$

$$\begin{aligned} \sum_{t=1}^m N_{i(m+n+1)-m-n-1+t} dz_t + \sum_{t=1}^n N_{i(m+n+1)-n-1+t} dx_t &= \\ = N_{i(m+n+1)} dW_j & \quad [1 \leq i \leq p - m - n] \quad [m + n + 1 \leq j \leq p] \end{aligned} \quad (2.04)$$

$$\begin{aligned} N_c(W_j) &= \sum_{t=(c-1)r+1}^{cr} b_{c,t} \left( \prod_j^{p+q} W_j^{E'_{c,s}} \right) \\ & \quad [1 \leq c \leq i(m+n+1)] \quad [1 \leq i \leq p - m - n] \end{aligned} \quad (2.05)$$

$$\begin{aligned} \sum_{t=1}^m T_{i(m+n+1)-m-n-1+t} dz_t + \sum_{t=1}^n T_{i(m+n+1)-n-1+t} dx_t &= \\ = T_{i(m+n+1)} dW_j & \quad [1 \leq i \leq q] \quad [p \leq j \leq p + q] \end{aligned} \quad (2.06)$$

where " $T_g(W_j)$ " are the special multivariate polynomials that would be reserved exclusively for only representing those functional expressions that would be present inside a DE or system of DEs.

Just as we can represent any mathematical equation in universal differential form, we can also express and type of DE and system of DEs also in complete universal differential form.

The *Primary Expansion* representation for the following general system of DEs:

$$g_k \left( z_i, x_j, \frac{\partial}{\partial x_u} \left( \frac{\partial z_r}{\partial x_v} \right) \right) = 0 \quad (2.07)$$

can be defined as follow:

$$G_k \left( W_t, \frac{P_{ruv}}{Q_{ruv}} \right) = 0 \quad [1 \leq t \leq m + n + q] \quad (2.08)$$

where:

$$\frac{P_{ruv}}{Q_{ruv}} = \frac{\partial}{\partial x_u} \left( \frac{\partial z_r}{\partial x_v} \right) \quad (2.09)$$

As for the *Secondary Differential Expansion* representation, it becomes exactly identical to the one present inside the initially assumed *Multivariate Polynomial Transform* that would have been selected for solving the general system of DEs.

**Example (2.1).** The following system of second order ODEs is used to describe the motion of a dumbbell of length "L" in space consisting of masses " $m_1$ " and " $m_2$ " both rigidly attached at its extremities and free to rotate under the influence of gravity:

$$(m_1 + m_2) \frac{d^2 x_1}{dt^2} - m_2 L \frac{d^2 \theta}{dt^2} \sin(\theta) - m_2 L \left( \frac{d\theta}{dt} \right)^2 \cos(\theta) = 0 \quad (2.10)$$

$$(m_1 + m_2) \frac{d^2 y_1}{dt^2} - m_2 L \frac{d^2 \theta}{dt^2} \cos(\theta) - m_2 L \left( \frac{d\theta}{dt} \right)^2 \sin(\theta) = -(m_1 + m_2)g \quad (2.11)$$

$$L \frac{d^2 \theta}{dt^2} - \frac{d^2 x_1}{dt^2} \sin(\theta) + \frac{d^2 y_1}{dt^2} \cos(\theta) = -g \cos(\theta) \quad (2.12)$$

For this system of equations, " $x_i$ " and " $y_i$ " represent the horizontal and vertical linear displacements of mass " $m_i$ " respectively and " $\theta$ " is the angle of rotation of the dumbbell with respect to the X-axis. We will assume for the sake of simplicity that the mass of the rod is negligible compared to mass " $m_1$ " and " $m_2$ ".

The complete *Multivariate Polynomial Transform* of the system of second order ODEs will now be determine.

For the sake of simplicity, we will need to express the *Sine* and *Cosine* function as a rational combination of the *Tangent* function by selecting:

$$h_1 = \tan(\theta/2) \quad (2.13)$$

so that:

$$\sin \theta = \frac{2 \tan(\theta/2)}{1 + \tan^2(\theta/2)} = \frac{2h_1}{1 + h_1^2} \quad (2.14)$$

and

$$\cos \theta = \frac{1 - \tan^2(\theta/2)}{1 + \tan^2(\theta/2)} = \frac{1 - h_1^2}{1 + h_1^2} \quad (2.15)$$

We can arbitrarily select each auxiliary variable as:

$$W_1 = x_1 \quad (2.16)$$

$$W_2 = y_1 \quad (2.17)$$

$$W_3 = \theta \quad (2.18)$$

$$W_4 = \tan\left(\frac{\theta}{2}\right) \quad (2.19)$$

The *Multivariate Polynomial Transform* of the single external input " $h_1$ " as defined by equation (2.13) is:

(1). Primary Expansion:

$$H_1 = W_4 \quad (2.20)$$

(2). Secondary Differential Expansion:

$$0 \cdot dx_1 + 0 \cdot dy_1 + (1 + W_4^2)d\theta = 2 dW_4 \quad (2.21)$$

Using our standard notation in equation (2.08) and (2.09), we can now define the complete *Multivariate Polynomial Transform* of this entire system of second order ODEs starting with the *Primary Expansion* as:

(1). Primary Expansion:

$$G_1 = (m_1 + m_2) \left( \frac{P_{21}}{Q_{21}} \right) - m_2 L \left( \frac{P_{23}}{Q_{23}} \right) \left( \frac{2W_4}{1 + W_4^2} \right) - m_2 L \left( \frac{P_{13}}{Q_{13}} \right)^2 \left( \frac{1 - W_4^2}{1 + W_4^2} \right) = 0 \quad (2.22)$$

$$G_2 = (m_1 + m_2) \left( \frac{P_{22}}{Q_{22}} \right) - m_2 L \left( \frac{P_{23}}{Q_{23}} \right) \left( \frac{1 - W_4^2}{1 + W_4^2} \right) - m_2 L \left( \frac{P_{13}}{Q_{13}} \right)^2 \left( \frac{2W_4}{1 + W_4^2} \right) + (m_1 + m_2)g = 0 \quad (2.23)$$

$$G_3 = L \left( \frac{P_{23}}{Q_{23}} \right) - \left( \frac{P_{21}}{Q_{21}} \right) \left( \frac{2W_4}{1 + W_4^2} \right) + \left( \frac{P_{22}}{Q_{22}} \right) \left( \frac{1 - W_4^2}{1 + W_4^2} \right) + g \left( \frac{1 - W_4^2}{1 + W_4^2} \right) = 0 \quad (2.24)$$

Where:

$$\frac{P_{n1}}{Q_{n1}} = \frac{d^n x_1}{dt^n} \quad (2.25)$$

$$\frac{P_{n2}}{Q_{n2}} = \frac{d^n y_1}{dt^n} \quad (2.26)$$

$$\frac{P_{n3}}{Q_{n3}} = \frac{d^n \theta}{dt^n} \quad (2.27)$$

The complete *Secondary Differential Expansion* of this system of second order ODEs is the combination of the *Secondary Differential Expansion* of "h<sub>1</sub>" as defined by equation (2.21) and

the same one present inside the initially assumed *Multivariate Polynomial Transform* that would have been selected for solving this particular system of second order ODEs.

**Example (2.2).** For the following system of *second order* PDEs,

$$z_2 \frac{\partial z_1}{\partial x_1} + \sin(2x_2) \frac{\partial}{\partial x_1} \left( \frac{\partial z_2}{\partial x_2} \right) + x_1 x_2 = 0 \quad (2.28)$$

$$\left( \frac{\partial^2 z_1}{\partial x_2^2} \right) \left( \frac{\partial^2 z_2}{\partial x_1^2} \right) + \left( \frac{\partial z_1}{\partial x_1} \right)^2 + \left( \frac{\partial z_2}{\partial x_2} \right)^2 + x_1^2 + x_2^2 = 6z_1 e^{x_1} \quad (2.29)$$

we can define each external input as:

$$h_1 = z_1 \quad (2.30)$$

$$h_2 = z_2 \quad (2.31)$$

$$h_3 = x_1 \quad (2.32)$$

$$h_4 = x_2 \quad (2.33)$$

$$h_5 = \sin(2x_2) \quad (2.34)$$

$$h_6 = 6e^{x_1} \quad (2.35)$$

We can also select each auxiliary variable as:

$$W_1 = z_1 \quad (2.36)$$

$$W_2 = z_2 \quad (2.37)$$

$$W_3 = x_1 \quad (2.38)$$

$$W_4 = x_2 \quad (2.39)$$

$$W_5 = \tan(x_2) \quad (2.40)$$

$$W_6 = e^{x_1} \quad (2.41)$$

The *Multivariate Polynomial Transform* of the first external input " $h_1$ " is:

(1). Primary Expansion:

$$H_1 = W_1 \quad (2.42)$$

(2). Secondary Differential Expansion:

$$dz_1 + 0 \cdot dz_2 + 0 \cdot dx_1 + 0 \cdot dx_2 = dW_1 \quad (2.43)$$

The *Multivariate Polynomial Transform* of the second external input " $h_2$ " is:

(1). Primary Expansion:

$$H_2 = W_2 \quad (2.44)$$

(2). Secondary Differential Expansion:

$$0 \cdot dz_1 + dz_2 + 0 \cdot dx_1 + 0 \cdot dx_2 = dW_2 \quad (2.45)$$

The *Multivariate Polynomial Transform* of the third external input " $h_3$ " is:

(1). Primary Expansion:

$$H_3 = W_3 \quad (2.46)$$

(2). Secondary Differential Expansion:

$$0 \cdot dz_1 + 0 \cdot dz_2 + dx_1 + 0 \cdot dx_2 = dW_3 \quad (2.47)$$

The *Multivariate Polynomial Transform* of the fourth external input " $h_4$ " is:

(1). Primary Expansion:

$$H_4 = W_4 \quad (2.48)$$

(2). Secondary Differential Expansion:

$$0 \cdot dz_1 + 0 \cdot dz_2 + 0 \cdot dx_1 + dx_2 = dW_4 \quad (2.49)$$

The *Multivariate Polynomial Transform* of the fifth external input " $h_5$ " is:

(1). Primary Expansion:

$$H_5 = \frac{2W_5}{1 + W_5^2} \quad (2.50)$$

(2). Secondary Differential Expansion:

$$0 \cdot dz_1 + 0 \cdot dz_2 + 0 \cdot dx_1 + (1 + W_5^2)dx_2 = dW_5 \quad (2.51)$$



The *Multivariate Polynomial Transform* of the sixth external input " $h_6$ " is:

(1). Primary Expansion:

$$H_6 = 6W_6 \quad (2.52)$$

(2). Secondary Differential Expansion:

$$0 \cdot dz_1 + 0 \cdot dz_2 + W_6 dx_1 + 0 \cdot dx_2 = dW_6 \quad (2.53)$$

Using the notation defined in equation (2.08) and (2.09), the complete *Multivariate Polynomial Transform* of the entire system of second order PDEs may now be completely defined as:

(1). Primary Expansion:

$$G_1 = H_2 \left( \frac{P_{110}}{Q_{110}} \right) + H_5 \left( \frac{P_{212}}{Q_{212}} \right) + H_3 H_4 = 0 \quad (2.54)$$

$$G_2 = \left( \frac{P_{122}}{Q_{122}} \right) \left( \frac{P_{211}}{Q_{211}} \right) + \left( \frac{P_{110}}{Q_{110}} \right)^2 + \left( \frac{P_{201}}{Q_{201}} \right)^2 + H_3^2 + H_4^2 - H_1 H_6 = 0 \quad (2.55)$$

The complete *Secondary Differential Expansion* of this system of second order PDEs is the combination of the *Secondary Differential Expansion* of " $h_1$ " through " $h_6$ " and the same one present inside an initially assumed *Multivariate Polynomial Transform* that would have been selected for solving this particular system of second order PDEs.

**B**y substituting an initially assumed *Multivariate Polynomial Transform* into any type of DE or system of DEs would always result into defining a complete system of *Nonlinear Simultaneous Equations* to solve for. Each exact numerical solution set obtained will always define a complete exact analytical solution of the DE or system of DEs by *inverting* the corresponding initially assumed *Multivariate Polynomial Transform*. This is provided of course that each of the first order ODEs present inside the *Secondary Differential Expansion* have all been determined as being exact differentials and therefore always completely integrable.

Some of the *unknown coefficients* present inside an initially assumed *Multivariate Polynomial Transform* would be reserved exclusively for defining all the basis function that are to be present inside the *analytical* solution of a DE or a system of DEs. Others would be mainly responsible for assuring that the boundary conditions of the DE or system of DEs would be completely satisfied.

As a consequence of the fundamental laws of algebra, a completely differentiable mathematical equation as well as its many equivalent differential form representation in terms of a *Multivariate Polynomial Transform* can always appear in various disguise form. That is, any mathematical equation as well as its equivalent differential form representation can always have many alternative equivalent representations. However, to an observer each may appear quite distinct from one another and yet are completely identical with each other purely from a computational point of view.

Such a unique mathematical property about equations in general would guarantee that there will always be an infinite number of *numerical solution* sets of the *Nonlinear Simultaneous Equations* corresponding to a DE or a system of DEs. As a result of this, we acquire the ability of being able

to select among an infinite number of *numerical solution* sets obtained only those that would *translate* into defining much simpler *Secondary Differential Expansion* to integrate. This would have the effect of significantly facilitating the entire integration process involved in the *Secondary Differential Expansion* when attempting to invert an initially assumed *Multivariate Polynomial Transform* for acquiring an exact analytical solution to a DE or a system of DEs.

No analytical method of integration has ever been devised in the history of Calculus that could offer us with this much flexibility for selecting out of an *infinite* number of integrals only those that are considered more friendly to evaluate than others while in the process of attempting to solve for a DE or a system of DEs. Other well known traditional methods of analytical integration have shown weaknesses in that area mainly as a result of some major integrability issues due to a very restricted number of integrals that could be resolved in the end while leaving behind a vast majority of them as completely unsolved.

When the *Nonlinear Simultaneous Equations* cannot be solved in terms of an *exact numerical solution set*, this in turn would indicate that the *exact analytical solution* of the DE or system of DEs in question cannot be resolved as some exact combination of algebraic and elementary basis functions whether explicitly or implicitly defined. It is then always possible to establish some form of a measure on the degree of accuracy that a particular *numerical solution set* can satisfy a system of *Nonlinear Simultaneous Equations* by using various well known methods of optimization techniques. This in turn would provide us with some real measure of accuracy on how well the resultant *analytical solution* obtained can satisfy the DE or system of DEs. Of course only when an *exact numerical solution set* of the *Nonlinear Simultaneous Equations* has been found then this would automatically indicate that the DE or system of DEs in question can be completely resolved in terms of an *exact analytical solution*. All of this is provided of course that each first order ODE present in the *Secondary Differential Expansion* are determined to be exact and thus always completely integrable.

The simplicity in appearance for the analytical solution of a particular DE or a system of DEs is very crucial towards a complete understanding of a physical system so that only those appearing in its simplest form would be of greatest interest to the physical science. If we were to apply this very general principle directly into the world of physics under the new proposed unified theory of integration, then Albert Einstein's assertion that "*God does not play with dices*" could certainly be put to the real test with potential major historical implications !

### **3. A universal method of proof for the quadratic equation and the superposition theorem**

As a direct consequence of having established a unified theory of integration, a universal method of proof can be devised for proving a variety of classical theorems that were once proven under old traditional methods of pure mathematical logic. Only those theorems that can be formulated through some form of a DE or a system of DEs would be included.

The simple quadratic formula would fall into such category of theorems since it can always be reformulated computationally using a method that is based entirely on the use of successive partial differentiation. In this case, the unique computational method of proof for the quadratic equation begins by first computing the various partial derivatives of an initially assumed *Multivariate Polynomial Transform* that has been selected solely on the basis of representing only the class of *multivariate mathematical equations* that are defined in *explicit* form only.

This would correspond to the case for " $k = m = 1$ " in equation (1.002) through (1.006) such that instead of assuming a *Primary Expansion* in the form of " $f(z_i, x_j) = 0$ ", we would instead assume an *explicit version* in the form of " $z = z(x_j)$ " as being a ratio of two general multivariate polynomials. Note that since an explicitly defined equation is just a special case of an implicitly defined equation, we could have selected the original implicit form representation in the *Primary Expansion* and still arrive at an explicitly defined analytical solution in the end.

(1). Primary Expansion:

$$z(W_j) = \frac{P(W_j)}{Q(W_j)} \quad (1 \leq j \leq p) \quad (3.01)$$

where "P" and "Q" are each multivariate polynomials each consisting of a total number of "p" auxiliary variables each of which are raised to some floating point value.

(2). Secondary Differential Expansion:

$$dz = dW_1 \quad (3.02)$$

$$dx_i = dW_{i+1} \quad (1 \leq i \leq n) \quad (3.03)$$

$$\begin{aligned} \sum_{t=1}^1 N_{i(n+2)-n-2+t} dz + \sum_{t=1}^n N_{i(n+2)-n-1+t} dx_t &= \\ = N_{i(n+2)} dW_j & \quad [1 \leq i \leq p-1-n] \quad [n+2 \leq j \leq p] \end{aligned} \quad (3.04)$$

$$\begin{aligned} N_c(W_j) &= \sum_{t=(c-1)r+1}^{cr} b_{c,t} \left( \prod_j^p W_j^{E'_{c,s}} \right) \\ & \quad [1 \leq c \leq i(n+2)] \quad [1 \leq i \leq p-1-n] \end{aligned} \quad (3.05)$$

The computed values for the various partial derivatives of " $Z = Z(W_j)$ " would then be equated with the various partial derivatives that are calculated based entirely on a very unique change of variables involving the coefficients and the root of the quadratic equation.

This unique change of variable would include the root of the quadratic equation " $r = r(A,B,C)$ " that would be regarded as the *dependent* variable while the coefficients A, B and C would be defined as the *independent* variables. This would correspond to " $m = 1$ " and " $n = 3$ " in the above differential expansion form representation. We would setup our complete system of *Nonlinear Simultaneous Equations* to solve for by simply equating the various partial derivatives of " $r(A,B,C)$ " with respect to each of the coefficient A, B and C with the various partial derivatives of our initially assumed *Multivariate Polynomial Transform* that was setup to only represent all multivariate mathematical equations defined in explicit form only. We can also apply the same logic for determining the root formulas corresponding to *higher degree* polynomials.

By restricting our initially assumed *Multivariate Polynomial Transform* to represent all mathematical equations in explicit form, this will guarantee the presence of *exact* numerical solution sets corresponding to the *Nonlinear Simultaneous Equations* to solve for. Each of these exact numerical solution sets obtained would lead towards the formation of many complete snapshots of the actual general formula such that by some very special algebraic manipulations, will enable confirmation of its very unique existence. The type of algebraic manipulation involved that is to be conducted will be referred to in the following section as being a special type of mathematical interpolation.

Another and far more interesting example mainly for the physical sciences is arriving at the famous *superposition theorem* by once again beginning with an initially assuming *Multivariate Polynomial Transform*. This time we would be selecting our differential expansion strictly in terms of representing all *univariate* mathematical equations defined in *explicit* form only.

This would correspond to the case for " $k = m = n = 1$ " in equation (1.002) through (1.006) such that instead of assuming a *Primary Expansion* in the form of " $f(z_i, x_j) = 0$ ", we would instead assume the explicit version of " $y = y(x)$ " as being a ratio of two general multivariate polynomials:

(1). Primary Expansion:

$$y(W_j) = \frac{P(W_j)}{Q(W_j)} \quad (1 \leq j \leq p) \quad (3.06)$$

where "P" and "Q" are each multivariate polynomials each consisting of a total number of "p" auxiliary variables each of which are raised to some floating point value.

(2). Secondary Differential Expansion:

$$dx = dW_1 \quad (3.07)$$

$$dy = dW_2 \quad (3.08)$$

$$N_{3i-2}dx + N_{3i-1}dy = N_{3i}dW_j \quad [1 \leq i \leq p-2] [3 \leq j \leq p] \quad (3.09)$$

$$N_c(W_j) = \sum_{t=(c-1)r+1}^{cr} b_{c,t} \left( \prod_j^p W_j^{E'_{c,s}} \right) \quad (3.10)$$

$$[1 \leq c \leq 3i] [1 \leq i \leq p-2]$$

We would define the *Nonlinear Simultaneous Equations* to solve for by substituting the above generally assumed *Multivariate Polynomial Transform* into the following general class of *second order* ODEs.

$$h_1(x) \frac{d^2y}{dx^2} + h_2(x) \frac{dy}{dx} + h_3(x)y = h_4(x) \quad (3.11)$$

Next, we would be performing a very complete and detailed analysis on all computational results obtained by solving for the corresponding *Nonlinear Simultaneous Equations*.

The generalized form of this second order ODE would have been selected purely on the basis of its reoccurrence in describing various types of linear mechanical and electrical models.

As in the case of the quadratic equation, by restricting our initially assumed *Multivariate Polynomial Transform* to represent all mathematical equations in *explicit form* only, the presence of exact numerical solution sets corresponding to the *Nonlinear Simultaneous Equations* to solve for will confirm the unique explicit nature of the superposition theorem. It is with some very special type of algebraic manipulations to be discussed in the next section that we will succeed in identifying a number of *subclasses* of ODEs by which the general explicitly defined analytical solution obtained would be applicable to. By method of conjecture this would eventually lead us directly towards a purely *computational* proof of the famous *superposition theorem* thereby completely bypassing all forms of non-computationally based mathematical methods of analysis !

#### ***4. A new form of mathematical interpolation as a means of establishing a main pathway by which a unified theory of physics may be obtained***

Only from the relentless application of the new unified theory of integration on a very large scale over a substantial class of DEs and systems of DEs can we expect to begin slowly unravelling many potentially new and yet undiscovered theorems similar to the *superposition theorem*. It is only from the long term cumulative effect of gathering a large collection of such universal theorems that can only lead towards the development of some unified theory of physics. This would be the result of having meticulously consolidate each of the most fundamental theorems ever discovered into one gigantic universal theory of physics.

All existing experimentally based methods of physics could never succeed in achieving such a monumental objective for the physical sciences. That is because during the process of gathering the physical data there would be a severe loss of continuity that only mathematical equations are capable of maintaining throughout.

There are of course more advanced examples that can be selected other than the ones involving the computational proof of the quadratic equation and the superposition theorem especially from someone with a remarkable understanding of mathematics and the physical sciences. But no matter what example in whatever subject matter anyone decides to choose from, the bottom line is that *by following a very unique brand of mathematical ideology such as the one being proposed in this article, the new unified theory of integration will always computationally arrive at the same mathematical equations that all traditional methods have succeeded in arriving at in the past*. By doing so, this would undoubtedly provide just the ideal fertile testing ground for any real future software development related to the unified theory of integration.

It is expected that we would be following an extremely long computational trajectory for achieving in some cases the same exact results as with traditional methods of analysis. However, it should be very obvious to everyone of the enormous potential benefits involved especially on a long term basis.

***"Our unique computational approach will always certainly succeed in solving those "other" problems by which classical methods of analysis have completely failed as a result of not having provided an adequate solution to certain key DEs or systems of DEs"***.

So in order to take full advantage of what the unified theory of integration can offer to everyone, it must be implemented in a very methodological manner. That is, each DE and system of DEs that is being solved for must absolutely undergo a very thorough examination in terms of determining the best analytical solution that can be extracted from the relentless numerical application of the initially assumed *Multivariate Polynomial Transform* described in equation (1.002) through (1.006). All boundary conditions related to the DE or system of DEs must also become included as part of this gigantic computational process.

We would therefore need to construct some form of a very unique presentation by which a very special type of mathematical database would have to be created for storing all empirical results obtained. This would then be entirely converted in the form of pure mathematical equations. Beyond this computational stage, much further scrutiny would then be necessary for potentially recognizing certain key fundamental theorems that over time would eventually contribute towards the complete development of some unified theory of physics.

The exact nature of such a presentation that would be applicable for solving *all types* of DEs and systems of DEs under the new unified theory of integration can be described through the following general mathematical template.

$g\left(x, y, \frac{dy}{dx}\right) = 0$		
<i>Initial Condition</i>	<i>Coefficient Values</i>	<i>Exact analytical solution obtained using the Multivariate Polynomial Transform method</i>
$x_0, y_0$	$a_0, b_0, c_0, \dots$	$U_1(x, y) = 0$
$x_0, y_0$	$a_1, b_0, c_0, \dots$	$U_2(x, y) = 0$
$x_1, y_1$	$a_0, b_2, c_2, \dots$	$U_3(x, y) = 0$
$x_1, y_1$	$a_3, b_0, c_0, \dots$	$U_4(x, y) = 0$
$x_2, y_2$	$a_4, b_3, c_2, \dots$	$U_5(x, y) = 0$
$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$
$\dots$	$\dots$	$\dots$

Table 4.1

This tailored designed template was produced to accommodate only first order ODEs. However due to the universality nature of the fundamental logic behind introducing such a new type of template in mathematics, it can easily be modified to accommodate other far more complex types of DEs and systems of DEs.

In the following example, we have included a very simple *live* demonstration by which the proposed unified theory of integration would succeed in resolving a randomly selected "general" first order ODE uniquely in terms of a complete "general" analytical solution.

Only by following this example very closely would it become very apparent that our unique mathematical template has succeeded in developing a more *generalized* approach for arriving at *general analytical solutions to any type of DEs and systems of DEs*. This would certainly go a long way towards uncovering the many well hidden potential mathematical theorems that lay very deep beneath many unresolved DEs and systems of DEs.

***"It is only by being in complete possession of a very large collection of powerful mathematical theorems that can succeed in carving a whole new pathway by which a unified theory of physics can eventually be uncovered. "***

**Example (4.1).** Starting with the following *general* first order ODE,

$$x \frac{dy}{dx} + ay + bx^n y^2 = 0 \tag{4.01}$$

we can begin by constructing the following table:

$x \frac{dy}{dx} + ay + bx^n y^2 = 0$		
<i>Initial Condition</i>	<i>Coefficient Values</i>	<i>Exact analytical solution obtained using the Multivariate Polynomial Transform method</i>
$x_0 = 1$ $y_0 = 1$	$a = 1.0$ $b = 1.0$ $n = -1.0$	$(-3x + x^{-1})y + 2 = 0$
$x_0 = 1$ $y_0 = 2$	$a = 1.2$ $b = -1.0$ $n = 2.0$	$(1.4x^{1.2} - x^2)y - 0.80 = 0$
$x_0 = 1$ $y_0 = -1$	$a = 1.2$ $b = 1.5$ $n = -2.0$	$(1.7x^{1.2} + 1.5^{-2})y + 3.2 = 0$
$x_0 = 1$ $y_0 = 2$	$a = 2.0$ $b = -1.0$ $n = 2.0$	$x^2 y(0.5 - \ln(x)) - 1 = 0$
$x_0 = 1$ $y_0 = -2$	$a = 1.5$ $b = 2.0$ $n = 3.0$	$(-2.75x^{1.5} + 2x^3)y - 1.5 = 0$
$x_0 = 1$ $y_0 = 1$	$a = 1.0$ $b = 1.0$ $n = 1.0$	$xy(1 + \ln(x)) - 1.0 = 0$
$x_0 = 1$ $y_0 = -1$	$a = -1.0$ $b = 1.5$ $n = -1.0$	$x^{-1}y(-1 + 1.5 \ln(x)) - 1.0 = 0$

Table 4.2

The evidence gathered purely by observation from this table suggests by *conjecture* that:

$$f_1(x, y) = 0 = (Ax^B + Cx^D)y + E \quad (4.02)$$

and:

$$f_2(x, y) = 0 = x^A y(B + C \ln(x)) + D \quad (4.03)$$

both appear to be perfect candidates for the general exact solution of the ODE where the coefficients "A", "B", "C", "D" and "E" are to be expressed in terms of the coefficients "a", "b", "n" and the initial condition of the ODE.

For the first expression defined by " $f_1(x, y) = 0$ ", we substitute this equation into the ODE and equate like terms to zero.

The first derivative of equation (4.02) is defined as:

$$\frac{dy}{dx} = -\frac{\partial f_1}{\partial x} / \frac{\partial f_1}{\partial y} = \frac{-y(ABx^{B-1} + CDx^{D-1})}{Ax^B + Cx^D} \quad (4.04)$$

Substituting this equation into the ODE defined by equation (4.01), we obtain:

$$\frac{-xy(ABx^{B-1} + CDx^{D-1})}{Ax^B + Cx^D} + ay + bx^n y^2 = 0 \quad (4.05)$$

From equation (4.02):

$$y = \frac{-E}{Ax^B + Cx^D} \quad (4.06)$$

Substituting this equation into equation (4.05) and simplifying the results we arrive at the following general expression:

$$A(a - B)x^B + C(a - D)x^D - bEx^n = 0 \quad (4.07)$$

Based **purely on empirical observations only**, we can conjecture from entry 1, 2, 3 and 5 of table (4.2) and equation (4.02) that:

$$B = a \quad (4.08)$$

and:

$$D = n \quad (4.09)$$

Under this purely hypothetical assumption based entirely on the empirical data obtained, equation (4.07) can now be rewritten as:

$$A(a - B)x^a + C(a - D)x^n - bEx^n = 0 \quad (4.10)$$



Since the initial condition of the ODE is always known in advance, we can also include the following additional equation by substituting " $B = a$ ", " $D = n$ ", " $x = x_0$ " and " $y = y_0$ " into equation (4.02).

The results are:

$$(Ax_0^a + Cx_0^n)y_0 + E = 0 \quad (4.11)$$

The complete system of nonlinear simultaneous equations to solve for where the unknown coefficients are now reduced to "A", "C" and "E", can now be obtained by equating like terms to zero in equation (4.10) and by including equation (4.11) for satisfying the initial condition of the ODE.

The results are:

$$A(a - B) = 0 \quad (4.12)$$

$$C(a - n) - bE = 0 \quad (4.13)$$

$$(Ax_0^a + Cx_0^n)y_0 + E = 0 \quad (4.14)$$

One complete solution set to this system of three nonlinear equations in four unknowns is:

$$A \neq 0 \quad (4.15)$$

$$B = a \quad (4.16)$$

$$C = \frac{-Abx_0^a y_0}{a + bx_0^n y_0 - n} \quad (4.17)$$

and from equation (4.13):

$$E = \frac{(a - n)C}{b} \quad (a \neq n) \quad (4.18)$$

where in equation (4.02), we can set the "A" coefficient as arbitrary defined provided that it is not equal to zero. Note that the expression for "C" in equation (4.17) was derived by multiplying both sides of equation (4.14) with "b", adding the result with equation (4.13) and finally solving for "C".

If for example, we select " $A = -2.75$ ", " $a = 1.5$ ", " $b = 2$ ", " $n = 3$ ", " $x_0 = 1$ " and " $y_0 = -2$ " then using equation (4.15) through (4.18) we find that:

$$B = a = 1.5 \quad (4.19)$$

$$C = \frac{-Abx_0^a y_0}{a + bx_0^n y_0 - n} = \frac{-(-2.75)(2)1^{1.5}(-2)}{1.5 + 2(1)^{3.0}(-2) - 3} = \frac{-11}{-5.5} = 2 \leftarrow \quad (4.20)$$

$$D = n = 3 \quad (4.21)$$

and:

$$E = \frac{(a - n)C}{b} = \frac{(1.5 - 3)(2)}{2} = -1.5 \leftarrow \quad (4.22)$$

Substituting these coefficient values into equation (4.02), we arrive at the same expression as the one defined in the fifth entry of table (4.2).

Equation (4.02) represents an exact solution that appears to only satisfy a limited range of values for the coefficients present in the ODE. However, evidence suggests from table (4.2) that the exact solution obtained in entries 4, 6 and 7 are not in the same format as in equation (4.02).

As a result of this observation, more digging is required before a more complete general exact solution satisfying all the initial conditions and the coefficients present in the ODE is obtained.

For the second candidate " $f_2(x, y) = 0$ " as defined by equation (4.03), a relationship for the coefficients "A", "B", "C" and "D" expressed in terms of the coefficients "a", "b", "n" and the initial condition of the ODE can be determined by simply substituting equation (4.03) into the ODE and equating like terms to zero.

The first derivative of equation (4.03) is defined as:

$$\frac{dy}{dx} = -\frac{\partial f_2}{\partial x} / \frac{\partial f_2}{\partial y} = \frac{-y(ABx^{A-1} + ACx^{A-1} \ln(x) + Cx^{A-1})}{x^A(B + C \ln(x))} \quad (4.23)$$

Substituting this equation into the ODE we get:

$$\left[ \frac{-x^A y (AB + AC \ln(x) + C)}{x^A (B + C \ln(x))} \right] + ay + bx^n y^2 = 0 \quad (4.24)$$

From equation (4.03):

$$y = \frac{-D}{x^A (B + C \ln(x))} \quad (4.25)$$

Substituting this equation into equation (4.24) and simplifying the results we arrive at the following general expression to solve for:

$$y[(-AB + aB - C)x^A + C(a - A)x^A \ln(x) - bDx^n] = 0 \quad (4.26)$$

based purely on empirical observations, we can conjecture from the fourth, sixth and seventh entry of table (4.2) that:

$$A = n \quad (4.27)$$

Thus, on the basis of this purely hypothetical assumption, equation (4.26) becomes:

$$y[(-nB + aB - C - bD)x^n + C(a - n)x^n \ln(x)] = 0 \quad (4.28)$$

Since the initial condition of the ODE is always known in advance, we can also include the following additional equation by substituting " $x = x_0$ ", " $y = y_0$ " and " $A = n$ " into equation (4.03).

The results are:

$$x_0^n y_0 (B + C \ln(x_0)) + D = 0 \quad (4.29)$$

The complete system of nonlinear simultaneous equations to solve for where the unknown coefficients are B, C and D can be obtained by equating like terms to zero in equation (4.28) and by including equation (4.29) for satisfying the initial condition of the ODE.

The results are:

$$B(a - n) - C - bD = 0 \quad (4.30)$$

$$C(a - n) = 0 \quad (4.31)$$

$$x_0^n y_0 (B + C \ln(x_0)) + D = 0 \quad (4.32)$$

Equation (4.31) is a critical equation that specifies under which condition for the parameters in the ODE are " $f_1(x, y) = 0$ " and " $f_2(x, y) = 0$ " a valid exact solution. This condition is clearly visible since from equation (4.31), we know that " $C \neq 0$ " which ultimately leads us to conclude that " $n = a$ ".

Thus as a result of equation (4.31),

$$f_2(x, y) = 0 = x^A y (B + C \ln(x)) + D \quad (4.33)$$

satisfies the ordinary differential equation IF AND ONLY IF " $n = a$ ".

By extending table (4.2) to include additional exact solutions corresponding to a different set of values for the initial conditions and the coefficients present in the ODE, we can easily deduce that:

$$f_1(x, y) = 0 = (Ax^B + Cx^D)y + E \quad (4.34)$$

satisfies the ODE when " $n \neq a$ ".

The complete solution set of this system of three equations in three unknowns is:

$$D \neq 0 \quad (4.35)$$

$$C = -bD \quad (4.36)$$

$$B = \frac{-D}{x_0^n y_0} - C \ln(x_0) = \frac{-D - C x_0^n y_0 \ln(x_0)}{x_0^n y_0} \quad (4.37)$$

where from equation (4.33), we can set the "D" coefficient as arbitrary defined provided that it is not equal to zero.

If we select for example, "D = -1", "a = 2", "b = -1", "n = a = 2", " $x_0 = 1$ " and " $y_0 = 2$ " then using equation (4.27), (4.36) and (4.37), we find that:

$$A = n = 2 \leftarrow \quad (4.38)$$

$$C = -bD = -1(-1)(-1) = -1 \leftarrow \tag{4.39}$$

and:

$$B = \frac{-D - Cx_0^n y_0 \ln(x_0)}{x_0^n y_0} = \frac{-(-1) - (-1)(1)^2(2) \ln(1)}{(1)^2(2)} = 0.50 \leftarrow \tag{4.40}$$

Substituting these coefficient values into equation (4.33), we arrive at the same expression as the one defined in the fourth entry of table (4.2).

**T**he results of having performed such an indebt computational analysis from the application of a unified theory of integration on this particular general first order ODE has provided a very substantial amount of detailed information. In fact, this would go much further beyond the capability of any traditional *non-universal* method of computational analysis.

A typical report that a numerical analyst might be presenting to management would appear as follow:

*"... thus, our empirical findings has indicated to us that for this first order ODE there are two recognizable general exact solutions. The first one is for the case when "n = a" and the other is when "n ≠ a". The general exact solutions obtained can be expressed as a combination of algebraic and elementary basis functions defined only in explicit form. Furthermore, we have established that there is according to the empirical results presented in table (4.2) an explicit relationship involving the initial condition (x<sub>0</sub>, y<sub>0</sub>) of the ODE, the coefficients (a, b, n) of the ODE and the coefficients in our two initially assumed general exact solutions."*

It is expected that many such reporting systems applied on a very large variety of DEs and systems of DEs would inevitably lead to the discovery of many new fundamental theorems similar to the superposition theorem.

### 5. A universal system of implicit numerical interpolation

Finite and infinite expansion series were traditional used for many centuries as a means of approximating certain types of functions. Many forms of approximation were developed in the past but the Taylor's and Fourier's expansion series still remain the most widely used today.

We have described an entirely new *universal* differential expansion form capable of representing far more complex mathematical functions than what is possible under the Taylor's and Fourier's expansion series method. It now becomes a matter of much further and deeper investigation to determine how well can such a type of new differential expansion form succeed in approximating a general mathematical equation.

There are two major requirements for an initially assumed *Multivariate Polynomial Transform* to be used as a practical method of approximation. The first, is of course that there must be some type of DE or system of DEs associated in the process of completely defining the mathematical equation that is being approximated. The second, is that the *Secondary Differential Expansion* must become completely integrable upon having successfully arrived at some fairly good approximate numerical solution set of the relevant system of *Nonlinear Simultaneous Equations*.

When both of these conditions are met then this could potentially open the door for achieving a far more complex system of approximations than what other traditional methods can offer in mathematics. In our case, we would go much beyond the use of the more conventional types of approximation series by allowing only the *computational aspect* of our initially assumed

*Multivariate Polynomial Transform* decide what basis functions are contained in the approximation solution and also whether it is explicit or implicit by nature. This would also include *computationally* arriving at the correct combination of composite functions without imposing any limits whatsoever on each of their degree of composition.

We demonstrate in the following example a case by which a simple exponential function was being successfully approximated by a very complex implicitly defined mathematical equation consisting of at least one high degree composite function. It must be emphasized that the exact nature of the composite function and the very implicit nature of the entire approximation solution obtained were entirely established purely my method of computational analysis only.

**Example (5.1).** If we substitute the following *initially assumed Multivariate Polynomial Transform*:

(1). Primary Expansion:

$$Y = \frac{a_1 W_1 + a_2}{a_3 W_1 + a_4} \quad (5.01)$$

(2). Secondary Differential Expansion:

$$dX = \frac{b_1 W_1 + b_2}{b_3 W_1 + b_4} dW_1 \quad (5.02)$$

into the first order ODE that define the following exponential function:

$$y = 1.5e^{-0.5x} \quad (5.03)$$

then by solving for the relevant system of *nonlinear simultaneous equations*, we arrive at the following initially assumed *Multivariate Polynomial Transform* to invert:

(1). Primary Expansion:

$$Y = \frac{-0.16301958 W_1 + 0.26986711}{1.82320996 W_1 + 0.07715033} \quad (5.04)$$

(2). Secondary Differential Expansion:

$$dX = \frac{0.83816740 W_1 + 1.31793167}{0.64312753 W_1 + 0.0271196} dW_1 \quad (5.05)$$

The complete inverse *Secondary Differential Expansion* can be obtained by first integrating both sides of equation (5.05) for " $W_1(x)$ " using the following general integral formula for partial fractions:

$$\int \frac{au + b}{pu + q} du = \frac{au}{p} + \left[ \frac{bp - aq}{p^2} \right] \ln(pu + q) \quad (5.06)$$

The next step afterwards is to substitute the expression obtained for " $W_1(x)$ " into the *Primary Expansion* defined by equation (5.04).

The solution of equation (5.05) using the above integral formula is written as:

$$x = \left[ \frac{0.838W_1}{0.6431} \right] + \left[ \frac{\{1.318(0.643) - 0.838(0.0271)\}}{0.643^2} \right] \ln(0.643W_1 + 0.0271) + K \quad (5.07)$$

$$= 1.303W_1 + 2.00 \ln(0.643W_1 + 0.0271) + K \quad (5.08)$$

where:

$$K = x_0 - 1.303W_{01} - 2.00 \ln(0.643W_{01} + 0.0271) \quad (5.09)$$

It can be shown that if:

$$y = \frac{A_1W_1 + A_2}{A_3W_1 + A_4} \quad (5.10)$$

then:

$$W_1 = \frac{-A_4y + A_2}{A_3y - A_1} \quad (5.11)$$

so that from our *Primary Expansion* as defined by equation (5.04), we can directly express " $W_1$ " as a function of " $y$ " to obtain

$$W_1 = W_1(y) = \frac{-0.0771503y + 0.26986711}{1.82320996y + 0.16301958} \quad (5.12)$$

where:

$$\begin{aligned} A_1 &= -0.16301958, & A_2 &= 0.26986711 \\ A_3 &= 1.82320996, & A_4 &= 0.07715033 \end{aligned}$$

It follows that:

$$W_{01} = W_1(y_0) = \frac{-0.0771503y_0 + 0.26986711}{1.82320996y_0 + 0.16301958} \quad (5.13)$$

$$= \frac{-0.077(1.5) + 0.270}{1.823(1.5) + 0.163} = \frac{0.1545}{2.8975} \quad (5.14)$$

$$= 0.0533 \quad (5.15)$$

The constant of integration defined by equation (5.09) may now be evaluated as:

$$K = 0 - 1.303(0.0533) - 2.00 \ln[(0.643)(0.0533) + 0.0271] \quad (5.16)$$

$$= 0 - 0.0694 - 2.00 \ln(0.06137) = 5.512 \quad (5.17)$$

By substituting equation (5.12) into equation (5.08) and simplifying the results, we arrive at an implicitly defined equation in the form of:

$$f(x, y) = 0 = 1.303 \left[ \frac{-0.077y + 0.270}{1.823y + 0.163} \right] +$$

$$+ 2.0 \ln \left[ \frac{0.178}{1.823y + 0.163} \right] - x + 5.512 \quad (5.18)$$

$x$	$y_{exact}$	$w_1$	$f(x, y_{exact})$
-5.0	18.273741	-0.034049	-5.861676E-003
-4.5	14.231604	-0.031716	-5.583780E-003
-4.0	11.083584	-0.028729	-5.233661E-003
-3.5	8.631904	-0.024910	-4.795080E-003
-3.0	6.722534	-0.020031	-4.249876E-003
-2.5	5.235514	-0.013808	-3.579069E-003
-2.0	4.077423	-0.005885	-2.765262E-003
-1.5	3.175500	0.004179	-1.797217E-003
-1.0	2.473082	0.016924	-6.780194E-004
-0.5	1.926038	0.033003	5.611024E-004
0.0	1.500000	0.053192	1.838294E-003
0.5	1.168201	0.078390	2.985896E-003
1.0	0.909796	0.109606	3.697588E-003
1.5	0.708550	0.147920	3.452581E-003
2.0	0.551819	0.194418	1.412990E-003
2.5	0.429757	0.250076	-3.704934E-003
3.0	0.334695	0.315614	-1.377577E-002
3.5	0.260661	0.391310	-3.140153E-002
4.0	0.203003	0.476811	-5.999799E-002
4.5	0.158099	0.570992	-1.037754E-001

Table 5.1

Many of the numerical solution sets obtained not shown here satisfied the relevant system of *Nonlinear Simultaneous Equations* to a fairly high degree of accuracy. In fact so much so that we decided to conduct a more indebt numerical analysis by comparing the results with the implicitly defined equation obtained from having inverting the corresponding initially assumed *Multivariate Polynomial Transform*. This has created just the perfect environment by which an implicitly defined analytical solution was able to approximate to a fairly reasonable level of accuracy the simple ordinary exponential function defined by equation (5.03).

### 6. *Mathematica's own approach to analytical integration*

*Mathematica* is a very popular software package that maintains a collection of symbolic and numerical methods for dealing with the entire aspect of differentiation and integration. Their general approach to integration is nowhere near the one described in this article which is based entirely on the application of multivariate polynomials as well as the differential of multivariate polynomials for finding analytical solutions to all types of DEs that would also include systems of DEs as well. Their online documentation does not present a single instance by which multivariate polynomials and the differential of multivariate polynomials have ever being applied for solving any particular type of DE or system of DEs.

Wolfram's general symbolic approach to solving DEs has the greatest drawback that it cannot be applied universally right across all types of DEs and systems of DEs. Under the new proposed unified theory of integration presented in this article, all DEs and systems of DEs are first subjected to a very rigorous computational process designed specifically for acquiring the type of data that would be transformed in terms of analytical solutions involving the algebraic and elementary basis functions only. Depending on the nature of the data acquired, each analytical solution obtained would be expressed in either explicit or in implicit form involving the use of composite functions with no limit whatsoever on each of their degree of composition.

The exact computational process involved would be the result of substituting the initially assumed *Multivariate Polynomial Transform* described by equation (1.002) through (1.006) into a DE or a system of DEs and afterwards solving for the relevant system of *Nonlinear Simultaneous Equations* that are generated from this process. Each numerical solution set of the *Nonlinear Simultaneous Equations* become the data by which all analytical solutions are constructed from.

As part of the general procedure, this would always involve the *exact* integration of a series of *first order ODEs* that are present in the *Secondary Differential Expansion* of an initially assumed *Multivariate Polynomial Transform* as described by equation (1.003) through (1.006). They will always appear as first order ODEs *regardless* of the type of DE or system of DEs that is being solved for. Since only first order ODEs are always involved then each are subjected to passing the fundamental test of exactness for determining whether or not any one of them is an exact differential. If so, then the integration process becomes considerably simplified for all those differentials that succeed in passing the critical test of exactness.

The final stage of the process would require that exact analytical solutions obtained from this unique integration process be substituted into the *Primary Expansion* as defined by equation (1.002). It is at this point that the various boundary conditions of the original DE or system of DEs are being matched with the ones that are naturally present throughout the complete integration process of the *Secondary Differential Expansion* of an initially assumed *Multivariate Polynomial Transform*.

With *Mathematica* you cannot just simply enter *any type* of DE or system of DEs, especially of the PDE type and expect that an analytical solution whether exact or approximate be returned to you in either explicit or in implicit form. Also, you cannot expect an analytical solution to be constructed entirely from composite functions with no limits on each of their degree of composition just from the use of the algebraic and elementary basis functions. "*That is only possible under a true unified analytical theory of integration which is currently not present anywhere within all of Mathematica*". So in no way does Wolfram appear to follow this type of ideology in mathematics mainly because the computational complexities involved would also have been far too overwhelming for execution on just a regular PC.

By writing a general computer program for implementing such a proposed unified theory of integration in a *complete automated setting* would represent a far better alternative than using *Mathematica's general non-universal approach to analytical integration*.



### ***7. The development of a new type of physics for maintaining uniform continuity throughout***

The new proposed mathematical ideology restricts all analysis on mathematical equations at the differential level in order to insure that the concept of continuity be always maintained throughout. This would suggest that the application of a true unified theory of integration for solving *any type* of DEs or system of DEs could hypothetically lead us towards the creation of some infinitely perfect universe over its entire composition. This is provided of course that we are able to maintain complete continuity in mathematical equations throughout the entire process of finding analytical solutions to DEs and systems of DEs. Such an infinitely perfect and continuous universe would be quite feasible to construct but only on the general assumption that "*the mathematical properties of a straight line equation will always remain the same regardless of your exact physical location inside this perfect universe and regardless to what time frame you are specifically referring to*".

A true unified analytical theory of integration will guarantee that every type of DE or system of DEs has some analytical solution behind it whether considered as being exact or approximate. Furthermore, if the theory is to retain all the basic features of universality then it must be applicable to all cases involved without any exceptions whatsoever. The only way for this to be entirely possible is that such a unified analytical theory of integration must *absolutely* be "**computationally-base**" for arriving at complete *analytical solutions* to any type of DE and systems of DEs. So at this point there can be no doubt that the new proposed mathematical ideology being presented in this article does indeed appear to define some sort of a unified theory of integration.

This very powerful assertion made about analytical integration in general has mutated itself into a new kind of physics that I would like to introduce everyone as being an "**idealistic physics**".

The fundamental principle behind this new type of physics is that we can use an infinitely perfect universe for modelling our own *imperfect* physical universe as long as we are able to maintain complete continuity in mathematical equations by solving all DEs and systems of DEs under a single unified theory of integration. Other imperfect physical universes similar to our own may be modeled like clay from the same infinitely perfect controlled universe. Each would then differ from one another only in terms of some mathematical variation representing a measure on how energy is being distributed within the basic atomic structure of matter.

Without some way of maintaining complete continuity in mathematical equations it would virtually become impossible to establish some very fundamental links that can exist between mathematical equations. It's only through the complete consolidation of each of these fundamental links between mathematical equations that in the end would play a vital role for arriving at some unified theory of physics. All of this of course becomes absolutely invisible under any form of experimentally based theory.

In an idealistic physics, discrete variables would have no meaning whatsoever since everything would exist inside an infinitely perfect dynamical structure involving infinitesimal measurements of space and time. All forms of navigation inside this perfect universe would be moving along a pathway of DEs with the new mathematical ideology acting as the main propulsion engine. The only access entry point inside such an infinitely perfect universe is by computation and not based entirely on the use of our imperfect sense of human physical observation that everyone was expecting to succeed during the complete historical development of classical and modern physics.

***" To always remain a part of this reality, we need to listen very attentively to what mathematics is telling us and not what we always want to hear. "***

The complete understanding of our own imperfect physical universe could never become reality unless we take advantage of the basic tools offered by the new proposed mathematical and physical ideology being introduced in this article. Under this new system of logic, all references made from within this infinitely perfect universe would be driven strictly by computation which would virtually eliminate any risk of encountering the type of contradictions that today are so prevalent everywhere in classical and quantum physics.

If we were to succeed in arriving at some unified theory of physics then no doubt we would have at our fingertips a complete and very detailed understanding of our own physical universe that maybe one day might bring us one step closer to its original creator.

*and so ...*

*"what we are able to understand could give us the capacity to change it for the better."*

### **8. The complete unification of all of physics under one computer software development**

A unified theory of physics has true meaning only in relation to some unified theory of analytical integration. It is based on the general assumption that everything in this physical universe can be described by the use of DEs and systems of DEs. They in turn would be completely solvable as some exact or approximate algebraic combination of elementary and algebraic basis functions by following a very unique system of computational logic such as the one being introduced in this article.

Such a grand theory of physics would be constantly referring to the existence of some type of a gigantic universal algebraic system, the very same one in which *Albert Einstein* himself always believed had to exist for completely describing reality. It would stand up at the very top of the hierarchy of all other known existing theories of physics that would include the theory of general relativity, quantum physics and including string theory as well.

All traditional theories in physics lack a great deal of universality, the type that can only lead to the unification of all physics under a single unified theory of integration. By following the same common mathematical ideology that would be entirely based on the fundamental continuity property of all mathematical equations, there would be no risk of encountering any type of contradictions whatsoever. That is because everything would be presented on a computational platform driven entirely from the relentless application of the fundamental laws of differentiation from which the proposed unified theory of integration is entirely based on.

Methods of computation are so important in our everyday lives. The current existing global monetary structure which drives our entire world economy completely depends on it just as much as our technology could not exist without it. None of this would be possible without the use of some form of a "*system of computational logic*" applied to mathematical equations that would have originated from the application of some type of a mathematical ideology.

A *highly automated* computer software program can always be written for the complete implementation of the process involved in solving for any type of DEs and systems of DEs that would be entirely based on the application of the new proposed mathematical ideology. Such a new type of software development would undoubtedly be regarded as being "*the complete unified theory of physics*" but only in its most raw state. Human intervention would then only be necessary for complete translation of all computer results that would appear in the form of *exact* numerical computations into practical decipherable mathematical equations. They in turn would be used exclusively for describing the very fundamental structure of our entire physical universe.

Everyone would have complete access to this computer software over the internet for execution on the most advanced super computers of our time. This software would then be regarded as the main pillar by which all of theoretical physics may now be reconstructed *without leaving the impression that we are attempting to reinvent the wheel*. This I believe is possible since we would be finding ourselves moving along a pathway that would be describing an entirely new ideology in both mathematics and physics, the type that has never been investigated by anyone in the past. Much along the same line of reasoning as *CERN* was built around every part of *experimental* nuclear physics. In our case, we would be implementing a very unique technology by which every part of *theoretical* physics would now be investigated under a single common unified theory of integration.

This may perhaps one day have a very profound effect in the manner by which the prestigious *Nobel Prize* would be presented for major contributions into physics. There would be two such major prizes offered instead of one. The first, would be for exceptional contributions to all aspects

of *experimental* physics while the other, for outstanding new contributions into all aspects of the new proposed *idealistic* physics under a *complete unified theory of integration*. Eventually at some point in time, both types of physics will be expected to intersect at the same common point of intersection by which a theory of everything may one day become reality for all of mankind.

### ***9. Engineering science under one universal system of computational logic***

The new proposed mathematical ideology can also be transformed into a very unique method of engineering analysis by which all DEs and systems of DEs may now be more closely scrutinized for arriving at a much greater variety of analytical solutions. This would not be feasible by following any other existing traditional methods of analysis since the vast majority of analytical solutions obtained are generally limited to very simple functional expressions that are mostly expressed in explicit form.

Today, methods of solving for DEs and systems of DEs particularly of the PDE type are mostly based on the use of various forms of finite element methods of computational analysis. Under the new proposed mathematical ideology, all forms of engineering analysis would be initiated from the direct application of the initially assumed *Multivariate Polynomial Transform* that was defined by equation (1.002) through (1.006) above.

So rather than presenting a solution to a particular physical problem as a part of some traditional numerical database, our very unique approach would consist of building an entirely new different type of database that would have been constructed on the principle of substituting an initially assumed *Multivariate Polynomial Transform* into any type of DEs and system of DEs. The same computer program described earlier as representing the complete unified theory of physics in raw computational form would also be applicable for solving those well know DEs of engineering science that have proven very similar in appearance to those encountered in theoretical physics. In both cases involved, the proposed initially assumed *Multivariate Polynomial Transform* would become the main center stage by which all forms of theoretical analysis would be conducted in the future.

Most particularly important to the engineering science are the need for approximation methods of analysis that are based on the use of highly imperfect control volumes. For these types of engineering problems, we would then be adopting a more *approximate* analytical method of analysis that would be sharing the same common principles as those introduced in section (5).

It is expected that the same computer program originally built for handling all problems in *theoretical physics* would no doubt provide us with the greatest opportunity yet for revisiting all those problems in engineering science that have remained in cold storage. They all have remained there for quite some time now mainly due to a lack of a unified theory of analytical integration

### **10. Conclusions**

You have now all witness a very unique circumstance by which a new mathematical ideology has mutated itself into some form of a new ideology for the physical sciences. The new proposed mathematical ideology is entirely *computational-based* so that the entire process of arriving at some *analytical* solution for resolving *any type* of DEs and systems of DEs can be entirely automated through the development of a *unified computer program*. The proposed initially assumed universal differential expansion as described by equation (1.002) through (1.006) is a testament that all forms of pure *analytical* integration may now be handled under one gigantic unified *computational-based algebraic theory*. The development of such a unified theory of integration would not have been possible without the complete preservation of the fundamental continuity property of all mathematical equations. It is only through the use of differential expansion forms defined in the very special format as described by equation (1.002) through (1.006) that we are able to maintain complete continuity of all mathematical equations throughout the entire process of solving for any type of DEs and systems of DEs. Since virtually all of theoretical physics is founded on mathematical equations, it would be safe to assume that a universal computer program that would be build around such a proposed unified theory of integration would have to be regarded as being some sort of a "*unified theory of physics*" in its most raw numerical state. Human

intervention would then only be necessary for translating all computer results that would appear in the form of *exact* numerical computations into practical decipherable mathematical equations. The very unique computational structure of our standard initially assumed differential expansion form would offer an unlimited variety of mathematical equations for conducting all forms of *exact* theoretical analysis not only in the field of theoretical physics but also in the engineering and biological sciences as well.

It would be conducted on a scale never imagined possible under any other known traditional methods of analysis. Such a new exact method of analysis could one day offer the best hope yet for arriving at some unified theory of physics without the risk of incurring any form of contradictions that are so prevalent in modern physics today. Also, by introducing such a unified theory of integration into the physical sciences, it is expected in the long term that both physicists and engineers would become much less dependent on pure experimental method of analysis for achieving much greater design reliability of commercial products.

---

*Mathematics has no boundaries; its really our inability to understand it that creates such boundaries (12/14/97).*

### **References:**

- Adby and Dempster (1974) Introduction to Optimization Methods, John Wiley & Sons, NY.
- H.T. Davis (1962) Introduction to nonlinear differential and integral equations, Dover publications.
- Einstein, A. (1974) The meaning of Relativity, Princeton Paperback, Fifth Edition.
- Guillen, M. (1995) Five equations that changed the world, Hyperion, NY.
- J.K. Hale and J.P. LaSalle (1967) Differential Equations: Linearity vs. Non-Linearity, The Chauvenet Papers.
- M. Mikalajunas (1981) New algorithm for integrating DEs, SIAM 1981 Fall meeting, Netherland Hilton Hotel Cincinnati. OH.
- M. Mikalajunas (1983) Representing a PDE in terms of an infinite variety of integrable and non-integrable systems of ODEs, Abstracts of papers presented to the AMS Vol. 4, Number 5, 87<sup>th</sup> summer meeting, Albany, NY.
- M. Mikalajunas (1983) On the use of Multivariate Polynomials for integrating ODEs, NY State Mathematical Assoc. of Two Year Colleges, Seaway Section MAA, Spring 1983 meeting, Utica NY.
- G.M. Murphy (1960) Ordinary differential equations and their solutions, D. Van Nostrand Company.

## A Framework for Multiscale Modeling of Warfighter Blast Injury Protection

†Raj K. Gupta<sup>1\*</sup> and Andrzej J. Przekwas<sup>2</sup>

<sup>1</sup> DoD Blast Injury Research Program Coordinating Office HQ, US Army Medical Research and Materiel Command, 504 Scott Street, Fort Detrick, MD 21702, USA

<sup>2</sup> CFD Research Corporation, 701 McMillian Way NW, Huntsville AL 35806, USA,

†\*Presenting and corresponding author: raj.k.gupta.civ@mail.mil, mrmc.amedd.army.mil

### Abstract

Improvised explosive devices (IED) have become the predominant weapon used in recent military conflicts against the coalition forces and in terror acts against civilian populations. IED technology and deployment methods have increased in complexity, and so have the resultant injuries, which require an increase in the sophistication of protection countermeasures and efficacy of medical interventions. Computational injury biomechanics, complemented with model-guided experimental testing can provide valuable support, not only in better understanding of blast injury mechanics, but also in the development of personnel protective armor, injury diagnostics, combat casualty care, and rehabilitation. Contemporary models of injury biomechanics typically focus on a specific loading mechanism and injury type to a single organ (e.g., bone, brain, or lung). IED injuries, however, typically result in polytrauma caused not only by the primary blast wave but also by associated penetrating injures caused by the debris and ejecta from buried IEDs. Computational modeling of blast wave injury poses significant challenges as it involves several physical disciplines such as blast wave gas dynamics, human body biodynamics, injury biomechanics, and trauma pathophysiology as well as a range of spatial and temporal scales. US DoD, in collaboration with academia and industry, is developing computational models and tools for various injury types caused by IEDs. Research projects are exploring experimental and computational tools dedicated to blast wave-induced traumatic brain injury, lung injury, hearing loss, extremity injury, and soldier protection. This paper presents a novel concept and prototype implementation of a multiscale, multiresolution computational framework for modeling human body injury caused by IED blast wave and fragmentation/debris loads. The overall architecture of the framework, major components, and example simulation results of blast injury mechanisms are discussed.

**Keywords:** Blast Waves, Brain Injury, Military Medicine, Injury Biomechanics, Multiscale Models

### Introduction

Improvised explosive devices (IED) have become the predominant weapon in recent military conflicts. IEDs proved to be particularly effective against softer military targets, such as Humvees, trucks and foot patrols. Blast events accounted for nearly 70% of injuries in wounded Service Members in both Iraq and Afghanistan, and are the main cause of traumatic brain injury (TBI) (Eskridge et al., 2012, Heltemes et al., 2012). IEDs have caused even more devastating casualties among civilian populations exposed to terrorist bombing events. Over the last decade, IEDs have evolved to be a formidable threat not only to dismounted soldiers but also to armored vehicles. Buried IEDs are particularly lethal as they not only launch the blast wave toward the target but also projectiles and soil ejecta. Such IEDs typically result in polytrauma involving non-penetrating injuries caused by the primary blast wave, penetrating injures and amputations caused by the debris and ejecta, and behind armor trauma in areas protected by the personal protective equipment (PPE).

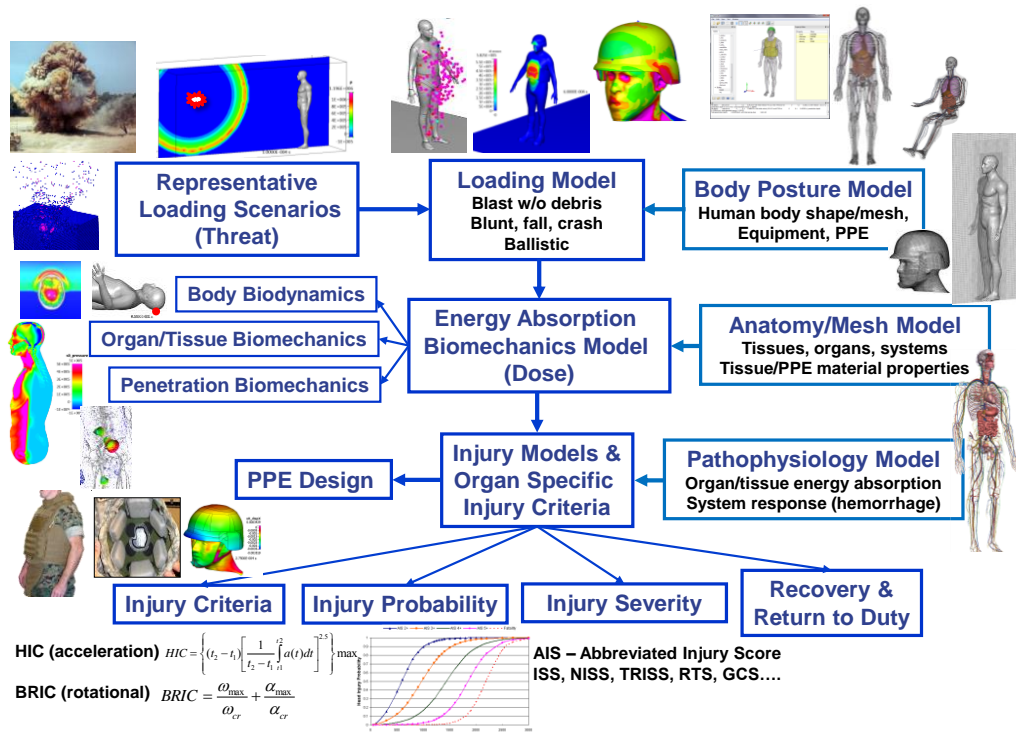
Compared to impact-related injury, the mechanisms involved in blast injury are much less understood. Recent improvements in PPE have reduced the severity of thoracic and brain ballistic injuries, but other sensitive organs, such as the brain, eye, ear, groin, genitals, and extremities, remain vulnerable to blast injuries [Curley et al., 2011]. Protection against blast wave TBI is

particularly challenging because, in spite of the protective helmet, a significant part of the soldier's head is still exposed to the blast. Until recently, it was not clear how a blast wave penetrates the cranium and causes brain injury and if military helmets protect against it [Zhang et al., 2011, Przekwas et al., 2011, Gupta and Przekwas 2013].

A better understanding of blast wave injury mechanisms can be achieved with a complementary experimental and computational modeling approach. However, computational modeling of blast wave injury poses significant challenges as it involves several physical disciplines, such as blast wave gas dynamics, human body biodynamics, injury biomechanics and trauma pathophysiology, as well as a range of spatial and temporal scales. US Department of Defense (DoD) in collaboration with academia and industry, is developing computational models and tools for various injury types caused by IEDs. This paper presents a novel multiscale computational framework for modeling human body injury caused by IED blast wave and fragmentation/debris loads. Our goal is to design a software tool for modeling IED blast threats and resultant human injuries and to assess biomechanical injury criteria and injury severity scores from injury biomechanics simulation results.

**Computational Framework for Multiscale Modeling of Blast Injury**

Design of a simulation framework for multiscale modeling of blast injuries involves not only mathematical modeling and numerical algorithms, but also technical specifications for software user scenarios and data analysis. The target applications include: prediction of injury mechanisms, development of improved PPE, evaluation of injury severity, and trauma care. Figure 1 presents the overall architecture of the designed simulation framework. The framework is built around existing CoBi tools [Gupta and Przekwas, 2013].



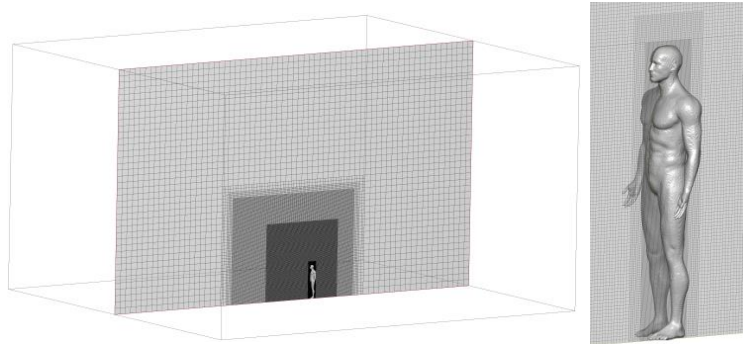
**Figure 1. Schematic representation of the computational framework for multiscale modeling of blast wave soldier injury mechanisms, severity, and protection.**

The Threat module (top left, Figure 1) selects representative threat scenarios with an IED located above, on or under the ground, and specifies the type and the volume of the explosive charge. A set of test cases is being collected, including problems from benchmark experimental tests and from

IED threats seen in real combat and terrorist explosion events. The Body Posture module (top right, Figure 1) positions anatomical geometry models of the human body in the field of threat. The user can setup both high-fidelity and reduced order human body models with injury susceptible organs, including skin, bones, joints, spine, muscles, brain, lungs, liver, spleen, kidneys, vasculature, and lower extremities. The soldier can be equipped with protective armor (e.g., helmet, vest, boots, soft armor, backpack). An adaptive octree mesh is automatically generated with mesh refinements close to the IED and the human body. The Loading module (top center, Figure 1) computes blast wave pressure and blast wind loads on the entire human body surface, as well as impact loading from IED projectiles and ejecta. The time-dependent loading forces on the human body are used as boundary conditions by the Energy Absorption Biomechanics module, which computes three types of biomechanical responses: whole body biodynamics, biomechanics of internal organs and tissues, and penetration mechanics of fragmentation and debris. The Injury module uses the organ/tissue biomechanics results to simulate two aspects of the injury: primary injury caused by tissue mechanical damage followed by secondary physiological and biological mechanisms. The injury results could be used to calculate injury criteria and injury severity scores. The following sections briefly describe the computational methods, present example results, and identify outstanding challenges.

### Human Body Anatomical Geometry and Computational Mesh

The anatomical geometry of a human body was used to generate computational meshes of the outside and inside of the body for blast and biomechanics simulations. Anatomical models can be generated using whole body imaging data (e.g., “Visible Human”) [Spitzer, 2006, Segars et al., 2010]. To simulate blast wave dynamics and its interaction with the body we used an adaptive octree mesh refined around the human body surface. The mesh extends throughout the entire blast scene, including the IED and the human body. Figure 2 presents an example octree mesh for modeling blast wave physics using Computational Fluid Dynamics (CFD) tools. Human anatomical geometry was also used to construct the articulated body model for modeling body biodynamics and a computational mesh within the body to simulate organ/tissue biomechanics. Several anatomical geometric models have been developed to study human body/head impact injury biomechanics [Yasuki, 2011, Gayzik et al., 2011, Tan et al., 2011].

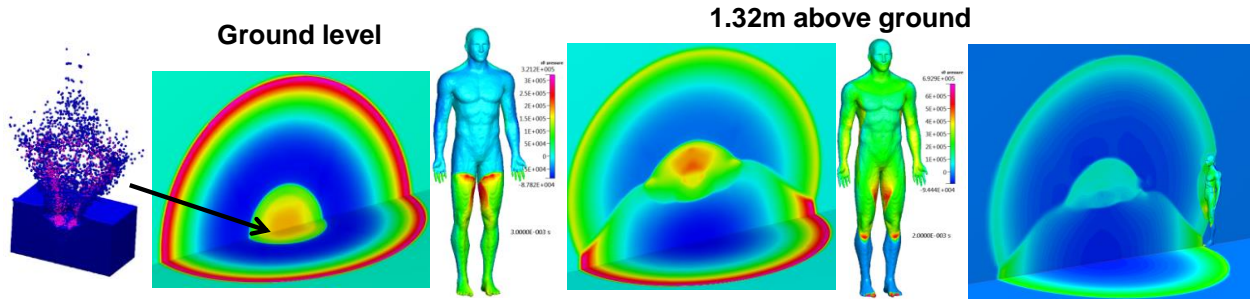


**Figure 2. Zonal adaptive octree grid for CFD simulations of blast wave impact on a human body.**

### Computational Models of Blast Wave and Human Body Blast Loading

A computational model of an IED explosion is often considered as a hallmark multiscale multiphysics problem as it involves four distinct elements: 1) the physics of charge detonation and energy release, 2) the detonation front interaction with soil mechanics and the generation of debris, 3) the gas dynamics of the blast wave, and 4) the blast wave interaction with the human body. We have employed the coupled Eulerian-Lagrangian approach for the interaction between the high explosive detonation products, fragments and soil. A discrete element/particle method (DEM) was used to model the casing fragments and soil through contact and collisions of rigid spherical particles, in addition to the developed Eulerian-based compressible flow solver for gaseous media. This sub-millisecond event is followed by high-fidelity CFD simulations of blast wave propagation toward the human body using CoBi tools with domain decomposition and parallel distributed computing [Harrand et al., 2013]. In the simulations we assumed that the human body inertia is high

enough to represent the human as a rigid body [Tan et al., 2012a, Gupta and Przekwas, 2013]. A Lagrangian ballistics model was used to simulate projectile and debris trajectories and their impact on the human body. Figure 3 presents typical simulation results of an IED detonation, debris generation and blast wave interaction with the human body for explosions on and above the ground as well as contour maps of the pressure loading on the human body surface [Tan et al., 2012a].



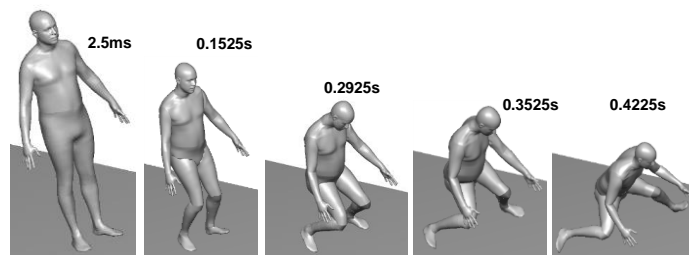
**Figure 3. FEM model simulation results of a buried IED detonation, debris generation, and blast wave interaction with the human body for explosions on and above the ground.**

The blast wave pressure loads and the forces generated by impacting debris on the entire human body are recorded for the duration of the blast event (10-20 milliseconds) and stored for consecutive human body biodynamics and biomechanics simulations.

### Models of Human Body Biodynamics and Biomechanics

The Loading module (Figure 1) generates external forces on the human body for three types of simulations: 1) articulated human body biodynamics modeling of flexible human body translocation in air and body impact on the ground, 2) biomechanical propagation of the primary pressure and shear waves within the body organs and tissues, and 3) terminal ballistics of IED fragments and debris penetrating the human body.

The articulated human body model represents a 1.78m, 84.6kg male with the anatomy divided into 17 body segments, such as the head, neck, upper/lower torso, abdomen, and extremities, connected by flexible anatomically consistent joints. Conic joints constrain the range of joint rotation to imitate contact of the spinous processes. These joints have been defined in terms of locations and shape based on the joint rotations of an average human. To apply the blast loading accurately and efficiently, a time step increment comparable to that in the CFD simulation (10 microseconds) was used to run the multibody dynamics for a period of 6.5 milliseconds. For computational efficiency and to take advantage of the unconditional stability of the implicit multi-body solver, the simulation was restarted at the end of 6.5 milliseconds with a much larger time step of one millisecond, and ran for an additional second. Details of the body model and the Finite Element Method (FEM) biodynamic formulation and its implementation of implicit multi-body dynamics and model validation results can be found in Tan and Przekwas [2011]. Figure 4 presents typical simulation results of a human body response to a shallow buried IED showing human body postures at several time instances.

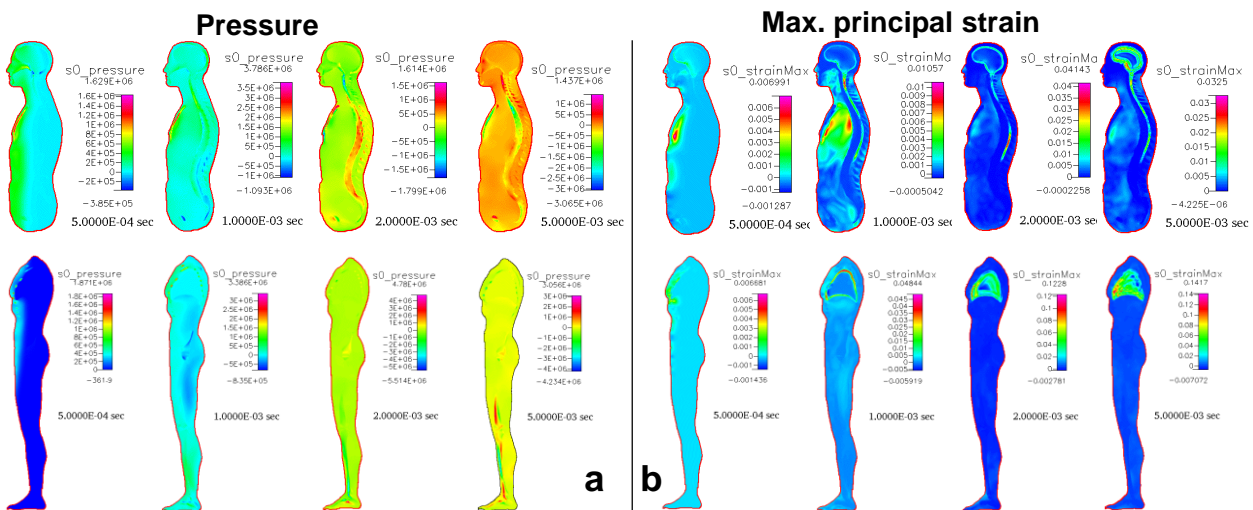


**Figure 4. FEM model of articulated human body biodynamics due to blast wave and fragment loadings at different times.**



Because of the large inertia of the human body, the body has not started to move yet at 2.5ms after the front impact blast on the body. The visible movement starts approximately 100ms post blast. This justifies our assumption of a rigid body during CFD blast simulations, presented in the previous section. The time required for the human body to impact the ground is on the order of one to two seconds, an order of magnitude longer than the blast loading duration. The predicted human body accelerations during free translocation and decelerations during the ground impact are used to calculate potential injury indicators, such as head injury criteria (HIC). The forces during the body impact on the ground are used as additional loads in the human injury biomechanics model.

By applying the blast loading resulting from the CFD simulation, we could model the biomechanical response of human body organs and tissues during the pressure wave propagation within the body. To simulate these human body biomechanics we used the CoBi FEM solver. Because of the time step increment constraints in the present transient simulations, an explicit time marching solution algorithm was used. Details of the biomechanics model can be found in Tan et al., [2012]. Figure 5 presents the predicted pressure fields and maximum principal strains in two sagittal planes (mid plane and mid left extremity) inside the body at four time instances during the propagation of stress waves through the body, following a free field blast impacting the body from the front.



**Figure 5. FEM simulation results of biomechanical stress wave propagation through the human body loaded by a blast wave from the front at four time instances after the blast impact (t=0.5ms, 1ms, 2ms, and 5ms): a) pressure profiles and b) maximum principal strain.**

The simulations demonstrated that a stiffer material, like the skeleton, has high pressure, while the pressure in soft materials, like the lung, is much lower (Figure 5a). A relatively high level of strain appears at the interface between the brain and the skull due to the low shear resistance of the brain (Figure 5b). The maximum strain occurs in the lungs because of their softness compared to their surrounding tissue and skeleton. A detailed description and discussion of blast injury biomechanics of the whole body and of the brain has been presented by Tan et al., [2011] and Gupta and Przekwas [2013]. These macro-scale simulations provide the stress and strain fields in all organs and tissues in the body and can be used as a starting point for micro-scale analyses of injury biomechanics of selected tissue structures, such as brain axonal tracks, spinal cord injury, cochlear injury, vascular injuries.

The simulation results of macro- and micro-scale injury biomechanics could also be used to evaluate biomechanical injury criteria (IC) and injury severity scores. Injury scores (IS), such as the Abbreviated Injury Scale (AIS), were developed to classify the type and severity of injury to aid in medical diagnosis. ISs are typically declared as a single number to represent complex and varying

degrees of critical illness. Despite this ambiguity, ICs are commonly used in critical medical care. This paper proposes that mathematical models may be used for more rigorous calculation of injury criteria and injury severity scores, including AIS, ISS, TRISS, and RTS, as well as injury scores for specific organs and polytrauma scores. Ultimately a fast running mathematical model of patient specific trauma trajectory may replace conventional single number ISs. The overall procedure for model-based calculation of ICs and ISs, estimation of injury severity probability, as well as prediction of trauma trajectory outcome, is shown in Figure 6.



**Figure 6. A procedure to calculate ICs and ISs from computational biomechanics results.**

## Conclusions

This paper presented a novel concept and prototype implementation of a multiscale computational framework for modeling human body injury caused by IED blast wave and fragmentation/debris loads. The overall architecture of the framework, the major components and example simulation results of blast injury mechanisms were described. A framework has been developed with the interaction between the whole human body anatomical geometry model, the mesh generation, and the individual computational tools, including the CFD blast dynamics and FEM biomechanics. This framework provides a foundation for the development of a user-friendly framework, not only for a better understanding of blast injury mechanisms, but also for the development of personal protective armor, injury diagnostics, combat casualty care, and rehabilitation. Several challenges remain, such as better material properties for high strain rate tissue biomechanics, coupling between macro- and micro-scale tissue biomechanics, tissue damage models for the estimation of the injury location and severity (such as axonal injury and bone fracture), accurate models of penetrating injury, and model validation.

## Disclaimer

The views expressed in this paper are those of the authors and may not necessarily be endorsed by the US Army or US Department of Defense.

## References

- Curley, K.C, Leggieri, M., Jaffee, M.S., and Moore, D.F. (2011). Opening Editorial, International State of the Science Meeting on Non-Impact Blast-Induced Mild Traumatic Brain Injury1, *Herndon, VA, USA*, May 12–14, 2009, *NeuroImage* 54, S14–S20
- Eskridge, S.L., Macer, C.A., Galarneau, M.R. et al (2012) Injuries from combat explosions in Iraq: Injury type, location, and severity, *Injury, Int. J. Care Injured* v43, 1678–1682
- Gayzik, F.S., Moreno, D.P., Geer, C.P., Wuetzer, S.D., Martin, R.S, and Stitzel, J.D. (2011). Development of a Full Body CAD Dataset for Computational Modeling: A Multi-modality Approach. *Annals of Biomed. Eng.* v39, 10, 2568-2583.
- Gupta, R.K. and Przekwas, A. (2013) Mathematical models of blast-induced TBI: current status, challenges, and prospects, *Frontiers in Neurology*, 4:59
- Harrand, V. Kannan, R., Tan, X.G., and Przekwas, A (2013) Highly-Scalable Computational Algorithms for Solving Aerostructural FSI Problems on Emerging Parallel Machine Architectures, Test and Evaluation Conference, Eglin Air Force Base, Niceville, Apr 16-19, 2013
- Heltemes, K.J., Holbrook, T.L., MacGregor, A.J., and Galarneau, M.R. (2012). Blast-related mild traumatic brain injury is associated with a decline in self-rated health amongst US military personnel. *Injury*, v43(12):1990-5.
- Przekwas, A., Tan, X.G., Harrand, V., Reeves, D., Chen, Z.J. Sedberry K., and Chancey, V.C. (2011). Integrated Experimental and Computational Framework for the Development and Validation of Blast Wave Brain Biomechanics and Helmet Protection", Proc. HFM-207 NATO Symposium on a Survey of Blast Injury Across the Full Landscape of Military Science, *Halifax NS, Canada*, Oct. 3-5, 2011.

- Segars, WP, Sturgeon G, Mendonca S., Grimes J., Tsui BM. (2010) 4D XCAT phantom for multimodality imaging research *Med Phys.* v37(9):4902-15.
- Spitzer, V.M., Scherzinger, A.L. (2006). Virtual anatomy: An anatomist's playground, *Clinical Anatomy. Special Issue: Visible Human Projects*, v19, 3, 192-203.
- Tan, XG and A.J. Przekwas, (2011) A computational model for articulated human body dynamics, *Int. J. of Human Factors Modeling and Simulation*, Vol. 2, pp. 85-110, 2011.
- Tan, XG., Przekwas, A. Kannan, R., Ott, K., Harrigan, T., Roberts, J. Merkle, A. (2012a) An Enhanced Articulated Human Body Model under C4 Blast Loadings, *ASME Int. Mech. Eng. Congress and Exposition, IMECE2012-89072*, Houston, TX, Nov 9-15, 2012
- Tan, XG., Kannan, R. and Przekwas, A. (2012b) A Comparative Study of the Human Body Finite Element Model under Blast Loadings, *ASME Int. Mech. Eng. Congress and Exposition, IMECE2012-89072*, Houston, TX, Nov 9-15, 2012
- Yasuki, T. (2011). Development of Total Human Model for Safety Version 4 Capable of Internal Organ Injury Prediction , in *Computational Biomechanics for Medicine, Soft Tissues and the Musculoskeletal System*, Ed by Wittek A, Nielsen MF., and Miller K., Springer Verlag 2011
- Zhang, L., Makwana, R., and Sharma, S. (2011). Comparison of the Head Response in Blast Insult with and without Combat Helmet. *Proc. HFM-207 NATO Symp. on a Survey of Blast Injury Across the Full Landscape of Military Science, Halifax NS, Canada*, Oct. 3-5, 2011

## Validation of a Fast Transient Solver based on the Projection Method

Darrin W. Stephens <sup>\*1</sup>, Chris Sideroff<sup>2</sup>, and Aleksandar Jemcov<sup>3</sup>

<sup>1</sup>Applied CCM Pty Ltd, Dandenong, Victoria 3175, Australia.

<sup>2</sup>Applied CCM Canada, Ottawa, ON K1J 6K3, Canada.

<sup>3</sup>Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.

### Abstract

This paper presents a fast transient solver suitable for the simulation of incompressible flows. The main characteristic of the solver is that it is based on the projection method and requires only one pressure and momentum solve per time step. Furthermore, advantage of using the projection method in the formulation is the particularly efficient form of the pressure equation that has the Laplacian term depending only on geometric quantities. This form is highly suitable for the high performance computing that utilises the Algebraic Multi-grid Method (AMG) as the coarse levels produced by the algebraic multi-grid can be stored if the grid is not changing. Fractional step error near the boundaries is removed by utilising the incremental version of the algorithm. The solver is implemented using version 5.04 of the open source library, Caelus [Applied CCM (2015)]. Accuracy of the solver was investigated through several validation cases. The results indicate the solver is accurate and has good computational efficiency.

**Keywords:** Projection Method; Fractional Step Algorithm; Algebraic Multi-grid Solver; Transient solutions; Validation

### Introduction

Fast transient solution of incompressible turbulent flows still occupy a dominant place in engineering computations. Majority of the solvers that use finite volume method on unstructured meshes with the arbitrary number of faces use either transient SIMPLE [Patankar and Splading (1972)] or PISO algorithm [Issa (1985)]. While these algorithms are known to produce spatially and temporally accurate solutions, they are not particularly efficient due to algorithmic constraints. Typically they require a multiple solution of the pressure equation (PISO algorithm) or multiple momentum and pressure solves per time step (SIMPLE algorithm). Multiple solutions of the pressure pressure and/or momentum equation per time step removes the fractional step error due to splitting of equations and recovers the

---

\*Corresponding author: d.stephens@appliedccm.com.au

time accuracy. Unlike SIMPLE and PISO algorithms, the projection algorithm introduced originally by [Chorin (1967)], does not require multiple pressure and momentum solves per time step. In this paper we describe an efficient implementation of the projection algorithm that utilises the Algebraic Multi-grid Method for the pressure equations suitable for high fidelity transient solutions.

## Governing equations and boundary conditions

Incompressible Navier-Stokes equations are given by the following set of partial differential equations:

$$\partial_t u_i + u_j \partial_j u_i = -\partial_i \hat{p} + \nu \partial_{ii} u_i. \quad (1)$$

It should be noted that the density is absorbed in the pressure field since it is considered to be constant in incompressible flow, i.e.  $\hat{p} = p/\rho$ . Discretization in time of Eq. (1) yields time discretized Navier-Stokes equations with ignored pressure term

$$\delta_t^{*,n} u_i = -u_j^n \partial_j u_i^n + \nu \partial_{ii} u_i^n. \quad (2)$$

Here symbol  $\delta_t^{*,n}$  is a finite difference applied in time coordinate

$$\delta_t^{*,n} u_i = \frac{u^* - u^n}{\Delta t}, \quad (3)$$

where  $u^*$  is the intermediate velocity field from which the pressure effects were removed, and superscript  $n$  denotes the previous time level. This step is called a momentum predictor step or simply predictor step in the language of fractional step methods. Introduction of the intermediate velocity  $u^*$  is the part of the algorithm in which the Hodge decomposition was performed in order to compute solenoidal free velocity field. The Hodge decomposition in this case becomes

$$u^* = u_{sol} + u_{irrot} = u^{n+1} + \nabla \phi, \quad (4)$$

where  $u_{sol} = u^{n+1}$  is the incompressible flow field we are solving for and  $u_{irrot} = \nabla \phi$  is the irrotational part of the flow field that is computed from the scalar potential.

The scalar potential  $\phi$  is obtained using the second part of the fractional step algorithm in which pressure gradient is brought to bare on the intermediate velocity field:

$$\delta_t^{n+1,*} u_i = \partial_i \hat{p}^{n+1}, \quad (5)$$

where  $\delta_t^{n+1,*}$  denotes the following finite difference in time

$$\delta_t^{n+1,*} u_i = \frac{u^{n+1} - u^*}{\Delta t}. \quad (6)$$

The Hodge decomposition now becomes

$$u^* = u^{n+1} + \Delta t \partial_i \hat{p}^{n+1}. \quad (7)$$

Since the field  $u^{n+1}$  is solenoidal, application of the divergence operator to Eq. (7) produces the equation for the scalar potential

$$\partial_{ii}\hat{p}^{n+1} = \frac{1}{\Delta t}\partial_i u_i^*. \quad (8)$$

Eq. (8) is used to compute the pressure field that can be used in the corrector step to produce the solenoidal field  $u^{n+1}$ :

$$u^{n+1} = u^* - \Delta t \partial_i \hat{p}^{n+1}. \quad (9)$$

Therefore, the scalar potential  $\phi$  in the Hodge decomposition in Eq. (7) is identified to be

$$\phi = \Delta t \hat{p}^{n+1}. \quad (10)$$

The fractional step error near the boundaries of the computational domain is removed by reintroducing the pressure in the discretized momentum equation:

$$\delta_t^{*,n} u_i = -u_j^n \partial_j u_i^n + \nu \partial_{ii} u_i^n - \partial_i \hat{p}^n. \quad (11)$$

This practice is commonly called incremental pressure method and is shown to remove the fractional step errors near the boundaries [Strikwerda and Lee (1999)].

## Method of solution

With the fractional steps identified above, the solution algorithm takes the following form:

1. Compute intermediate velocity field  $u_i^*$  according to Eq. (2) (momentum predictor step)
2. Compute pressure field  $\hat{p}$  from Eq. (8)
3. Compute divergence free velocity field  $u_i^{n+1}$  according to Eq. (9)
4. Repeat the process in the next time step

The most expensive part of the projection algorithm corresponds to the linear solution of the pressure equation given by Eq. (8). The structure of the pressure equation is strongly elliptic thus requiring the multi-level solver such as Algebraic Multi-grid Method (AMG) for efficiency reasons. The expensive part of the algorithm corresponds to the agglomeration procedure in which the coarse matrix hierarchy is created using the heuristic rules for determining the link strengths among fine equations. One advantage of the projection algorithm is that the coefficients in the Laplacian in Eq. (8) are purely geometrical. Discretization coefficients entering the discrete matrix on fixed grids are constant thus allowing the coarse hierarchy to be computed only once. Since the geometric coefficients do not depend on any other variables in the discretization process, this does not constitute any approximation beyond the usual approximations associated with the discretization process.

Given the geometric nature of the Laplacian in Eq. (8), the algorithm can be modified as follows:

1. Compute coarse mesh hierarchy for Eq. (8) and store it for the duration of the simulation
2. Compute intermediate velocity field  $u_i^*$  according to Eq. (2) (momentum predictor step)
3. Compute pressure field  $p$  from Eq. (8) by using the stored hierarchy of coarse matrices. The only change from time step to time step is in the right-hand-side of Eq. (8).
4. Repeat the process in the next time step starting from the step 2

The above algorithm termed the Semi-Linear Implicit Method (SLIM) is implemented as a solver using version 5.04 of the Caelus [Applied CCM (2015)] library. The origin of the name can be traced to the fact that the coefficients of the Laplacian are always stored and only the right-hand-side changes.

Furthermore, additional efficiency can be gained if the momentum predictor equations are solved using explicit time stepping thus avoiding any need for creating and storing matrices for the momentum equations. With the stored hierarchy for the Laplacian and explicit time stepping for the momentum equations, the algorithm would completely avoid creation of matrices and computationally would perform close to the fully explicit algorithms that are pervasive in compressible flow computations.

## Results and discussions

In the following section, several validation cases are presented for conditions ranging from attached to separated flow. Whilst the SLIM algorithm is by nature a time accurate solver, both steady state and transient cases have been considered under laminar and turbulent flows to fully investigate the performance of the algorithm and its implementation.

### *Laminar flat plate*

In this case, steady, incompressible, laminar flow over a two-dimensional sharp-leading edge flat-plate at zero angle of incidence was investigated. The flow generates a laminar boundary layer and the computational results are compared with the Blasius solution for incompressible flow. Blasius, in his work [Blasius (1908)] obtained the solution to the Boundary Layer Equations using a transformation technique. Here, equations of continuity and momentum in two-dimensional form are converted into a single ordinary differential equation (ODE). The solution to this ODE can be numerically obtained and is regarded as the exact solution to the boundary layer equations. One of the highlights of Blasius solution is the analytical expression for the skin friction coefficient ( $c_f$ ) distribution along the flat-plate given by

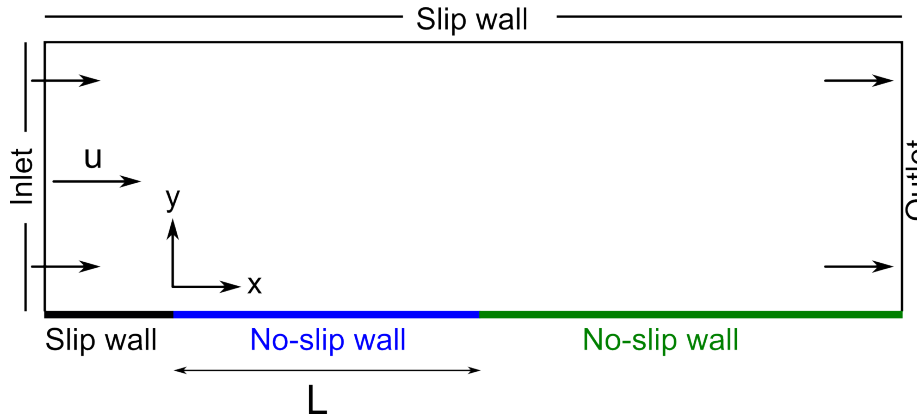
$$c_f \approx \frac{0.644}{\sqrt{Re_x}}, \quad (12)$$

where  $Re_x$  is the local Reynolds number defined as

$$Re_x = \frac{ux}{\nu}, \quad (13)$$

$u$  is the freestream velocity,  $x$  is the distance starting from the leading edge and  $\nu$  is the kinematic viscosity.

This case is based on the validation work carried out by the [NASA NPARC Alliance (2008)] for flow over a flat plate using the same conditions in the incompressible limit. A schematic of the geometric configuration is shown in Fig. 1. The length of the plate is  $L = 0.3048 \text{ m}$  wherein,  $x = 0$  is the leading edge, the Reynolds number of the flow based on the length of the plate is 200,000 and  $u$  is the velocity in the  $x$ -direction. Assuming the inlet flow is at a temperature of 300 K, the kinematic viscosity ( $1.58963 \times 10^{-5} \text{ m}^2/\text{s}$ ) can be determined from dynamic viscosity and density of the fluid. The value of dynamic viscosity is obtained from the Sutherland viscosity formulation [Sutherland (1893)]. Using the Reynolds number, plate length and kinematic viscosity, the freestream velocity evaluates to  $u = 10.4306 \text{ m/s}$ . As we have assumed the flow incompressible, the density ( $\rho$ ) remains constant. In addition, since the fluid temperature is not considered, the viscosity remains constant. For incompressible flows, the kinematic forms of pressure and viscosity are always used in Caelus 5.04.



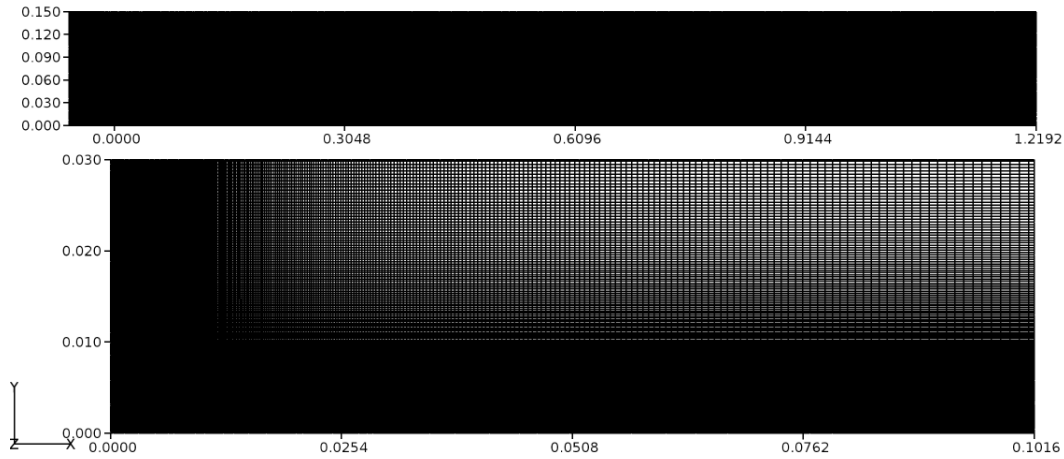
**Figure 1: Flat plate computational domain**

The computational domain is a rectangular block encompassing the flat-plate. Fig. 1 shows the details of the boundaries used in two-dimensions ( $x - y$  plane). The region of interest extends between  $0 \leq x \leq 0.3048 \text{ m}$  and has a no-slip boundary condition. Upstream of the leading edge, a slip boundary is used to simulate freestream uniform flow approaching the flat-plate. However, downstream of the plate, there is an additional no-slip wall a further three plate lengths. This ensures that the boundary layer in the vicinity of the trailing edge is not influenced by the outlet boundary. Since the flow is subsonic, disturbances cause the pressure to propagate both upstream and downstream. Therefore, placement of the inlet and outlet boundaries were chosen to have minimal effect on the solution. The inlet boundary is placed at start of the slip-wall ( $x = -0.06 \text{ m}$ ) and the outlet at the end of the second no-slip wall ( $x = 1.2192 \text{ m}$ ). Both inlet and outlet boundaries are between  $0 \leq y \leq 0.15 \text{ m}$ . A slip-wall condition is used for the entire top boundary. At the inlet a fixed uniform velocity  $u = 10.4306 \text{ m/s}$  in  $x$  direction and zero pressure gradient is applied and at the outlet a fixed uniform pressure



$p = 0 \text{ m}^2/\text{s}^2$  and zero gradient velocity are applied. The simulation is initialised with a constant velocity  $u = 10.4306 \text{ m/s}$  in the  $x$  direction and uniform zero pressure field.

A 3D hexahedral mesh was generated using [Pointwise (2014)] by extruding the 2D quadrilateral mesh in the  $x - y$  plane. The two  $x - y$  planes obtained as a result of grid extrusion need boundary conditions to be specified. As the flow over a flat-plate is generally 2D, we do not need to solve the flow in the third dimension. This is achieved in Caelus by specifying an empty boundary condition for each plane. Although, no flow is computed in the  $z$  direction, a velocity of  $w = 0$  has to be specified for the velocity boundary condition as indicated above.



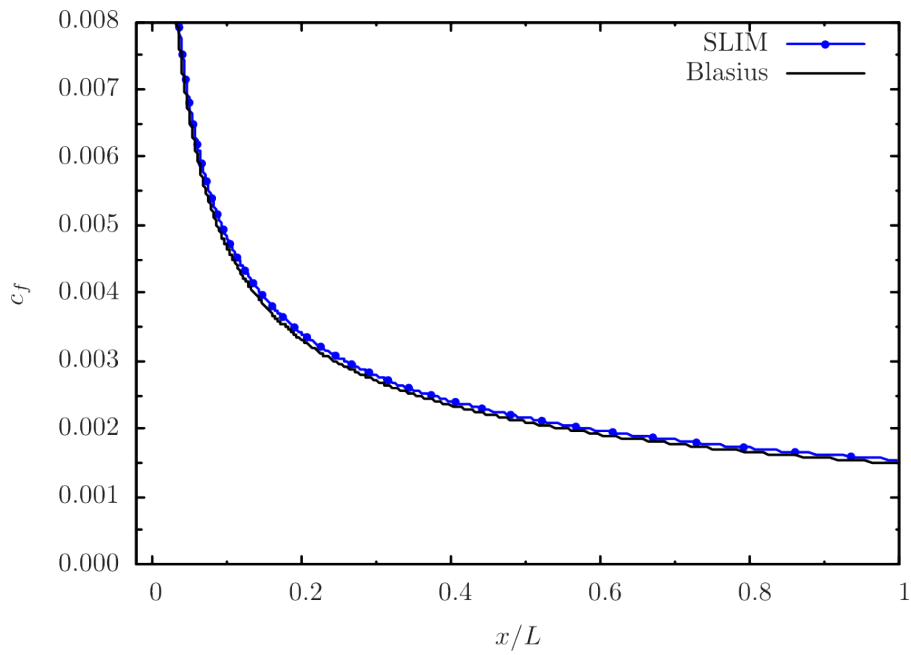
**Figure 2: Computational mesh for the flat plate domain.**

Fig. 2 shows the mesh in the  $x - y$  plane. As can be seen, the grid is refined perpendicular to the wall in order to resolve the viscous effects. To ensure that the gradients in boundary layer are well resolved, about 50 grid nodes are placed between the wall and the boundary layer edge. Grid refinement is also added at the leading edge so that the growth of the boundary layer is also well resolved. In this particular case, 399 cells were used in the streamwise ( $x$ ) direction ( $x \leq 0 \leq 0.3048 \text{ m}$ ) and 297 in the wall normal ( $y$ ) direction. For the no-slip wall beyond  $x > 0.3048$ , a similar distribution is used, resulting in a total cell count of 217,998.

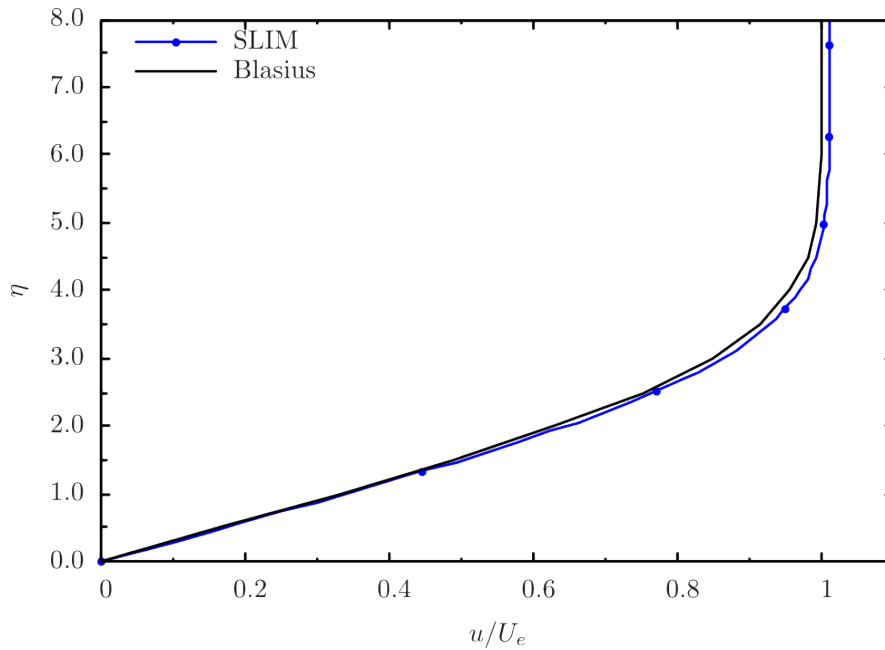
A time-dependent solution was obtained using the SLIM solver. The flow was simulated sufficiently long (several plate length flow times) such that steady flow was established. For the discretization of time-dependent terms, the first-order Euler scheme was used. Pressure and velocity gradients were calculated using the Gauss method. A  $2^{\text{nd}}$  order linear upwind discretization was used for the divergence of velocity.

In Fig. 3, the skin-friction distribution along the flat-plate obtained from the SLIM solver is compared with that of the Blasius analytical solution. Here, the distance  $x$  is normalised with the length of the plate ( $L$ ). Excellent agreement is observed along the entire length of the flat-plate. At the exit plane of the flat-plate at  $x = 0.3048 \text{ m}$ , velocity data was extracted across the boundary layer and compared with the Blasius analytical solution. This is shown in Fig. 4 where the velocity profile is plotted using similarity variables from the Blasius solution.

Here,  $\eta$  is the non-dimensional distance from the wall to the boundary layer edge and  $U_e$  is the velocity at the boundary layer edge. Similar to skin-friction, the velocity profile also exhibits excellent agreement with the Blasius solution.



**Figure 3: Skin-friction comparison between SLIM and Blasius solutions.**

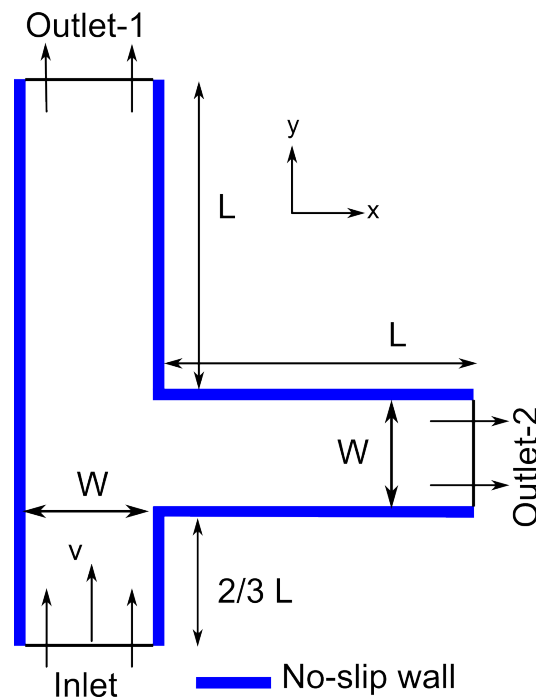


**Figure 4: Non-dimensional velocity profile comparison between SLIM and Blasius solutions.**

### Tee junction

In this validation case, laminar, incompressible flow through a two-dimensional  $90^\circ$  tee junction was investigated. Due to the presence of the side branch, the flow separates and forms a recirculation region. The recirculating region influences the mass flow through the main and side branches. The numerically computed mass flow ratio was calculated and compared with experiment. A comprehensive study of flow through planar branches has been carried out by [Hayes et al. (1989)] due to its prevalence in the bio-mechanical industry.

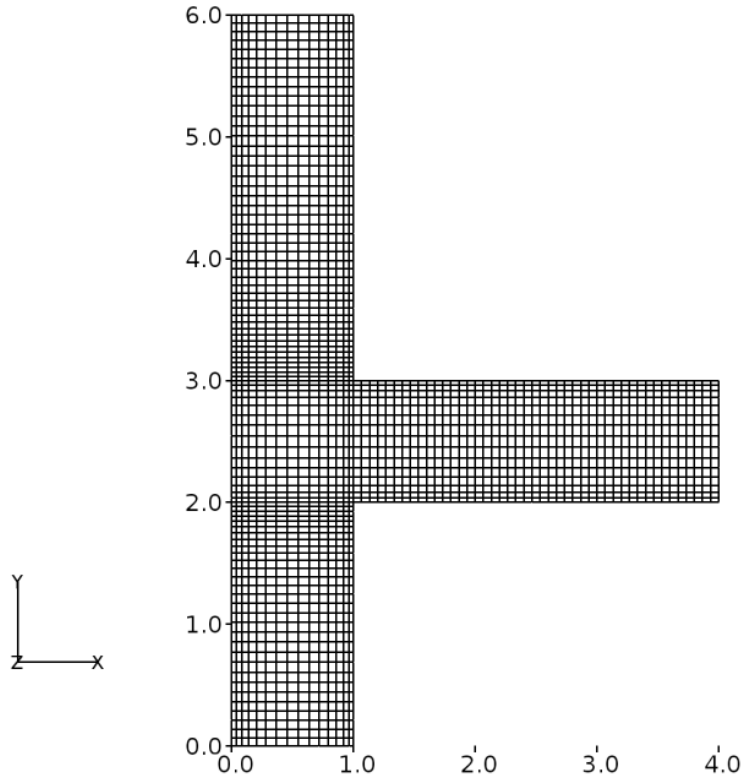
Fig. 5 shows the schematic of the tee-junction. Here,  $L = 3.0\text{ m}$  and  $W = 1.0\text{ m}$  respectively, the Reynolds number based on the width is 300, and  $v$  is the velocity in the  $y$ -direction. For simplicity, we have assumed the velocity,  $v = 1\text{ m/s}$ . Using these values the resulting kinematic viscosity was  $0.00333\text{ m}^2/\text{s}$ . Since this is an internal flow problem, the computational domain is contained within tee-junction geometry. All tee-junction walls have a no-slip boundary condition. At the inlet, a fully developed laminar flow parabolic profile is applied with a mean velocity  $v = 1.0\text{ m/s}$ , otherwise a much longer main branch would be required for the flow to develop. The domain has two outlets, one at the end of the main channel and the other at the end of side branch. Exit pressures at the two outlets are equal ( $p = 0\text{ m}^2/\text{s}^2$ ) and a zero gradient condition is applied to the velocity. The simulation is initialised with uniform zero velocity and pressure fields.



**Figure 5: Computational domain representing tee-junction.**

The computational mesh is shown in Fig. 6. The two  $x - y$  planes obtained as a result of grid extrusion have empty boundary conditions applied to achieve 2D flow. A total of 2,025 hexahedral cells comprise the tee-junction of which, 90 cells are distributed along the height of the main channel, and 45 along the length of the side branch. The distribution is such that a dimensional length of  $L = 1\text{ m}$

has a total of 45 cells, giving a distribution of 30 cells for the  $(2/3)L$  segment of the main channel. The width,  $W$ , consists of 15 cells.



**Figure 6: Structured grid for tee-junction domain.**

A time-dependent solution was obtained using the SLIM solver. The flow was simulated sufficiently long such that steady separated flow was established. To ensure this, the shear-stress distribution was monitored on the lower wall of the side branch. For the discretization of time-dependent terms, the first-order Euler scheme was used. Pressure and velocity gradients were calculated using the Gauss method. A  $2^{nd}$  order linear upwind discretization was used for the divergence of velocity.

The mass flow rate was calculated at the inlet and at the main outlet (outlet-1) and the ratio was subsequently calculated. Table 1 compares the SLIM result with the experimental value. As can be noted, the agreement between the two is excellent.

**Table 1: Comparison of mass flow rate split for SLIM and experiment.**

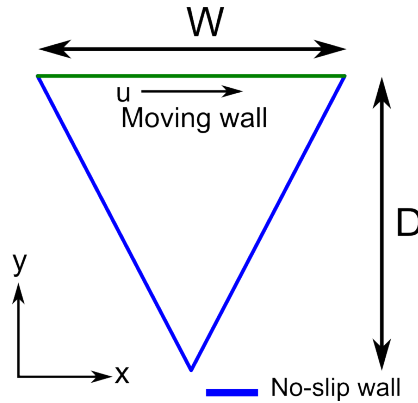
	Experimental	SLIM	Percentage Difference
Flow Split	0.887	0.886	0.112 %

*Triangular Cavity*

This validation study concerns the laminar, incompressible flow inside a lid driven triangular cavity. Here, the top wall of the cavity moves at a constant velocity

initiating a recirculating motion within the cavity. Experiments on this configuration have been reported in [Jyotsna and Vanka (1995)] for a Reynolds number of 800. The main objective of this validation case was to compare the  $x$  velocity distribution against experimental data.

A schematic of the triangular cavity is presented in Fig. 7 where the depth of the cavity  $D = 4\text{ m}$  and the width  $W = 2\text{ m}$ . The Reynolds number based on the cavity depth is 800 and the wall velocity is  $u = 2\text{ m/s}$ . Using the Reynolds number,  $u$ , and  $D$ , kinematic viscosity was calculated to be  $0.01\text{ m}^2/\text{s}$ . The side walls of the cavity have a no-slip boundary condition while the top wall, has a uniform velocity in the  $x$  direction. The simulation is initialised with zero velocity and pressure fields.

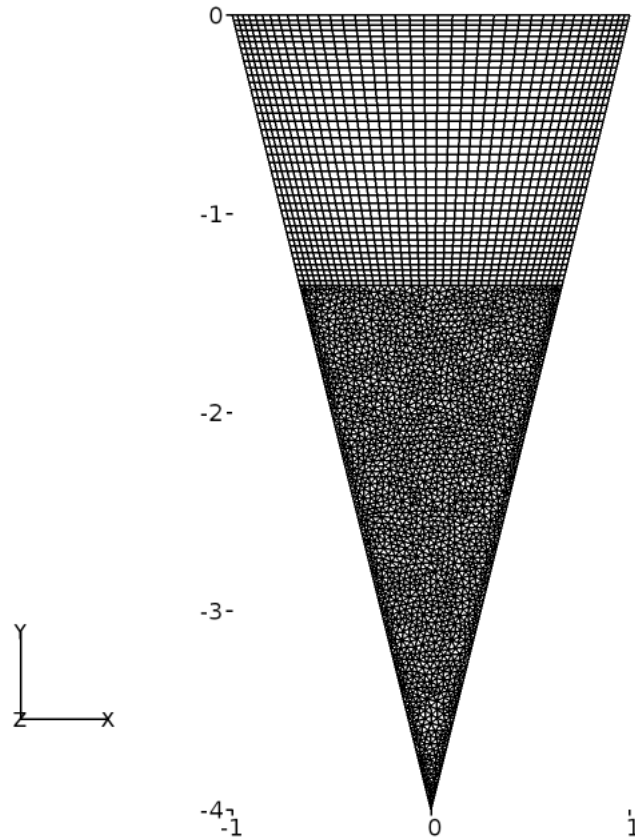


**Figure 7: Computational domain of the triangular cavity.**

The mesh in  $x - y$  plane is shown in Fig. 8. A hybrid mesh is employed for this case with a total of 5,538 cells. Up to a depth of  $D = 1.35\text{ m}$  hexahedral elements are used while below that value triangular prisms are used. The triangular prisms are used in the bottom portion because they resulted in better cell quality. For the hexahedral region, 39 cells are distributed across the width of the cavity and 39 along the depth. The cavity walls in the prism region have 100 cells along each. The interface of the two regions is node matched and has 39 cells across the width. The mesh close to the cavity lid was refined to better capture the shear layer. The flow characteristics in the cavity can be assumed to be 2D and here it has been solved with the same assumption.

A steady solution to the cavity was obtained using the SLIM solver. While a time-dependent approach was used, the solution was simulated sufficiently long so that steady flow was achieved. To determine when this occurred the velocity distribution along the cavity centre-line was monitored with respect to time. For the discretization of time-dependent terms, the first-order Euler scheme was used. Pressure and velocity gradients were calculated using the Gauss method. A  $2^{\text{nd}}$  order linear upwind discretization was used for the divergence of velocity. A linear corrected scheme was used for Laplacian discretization to account for mesh non-orthogonality.

In Fig. 9, the  $x$  velocity distribution along the cavity centre-line is compared with that of the experimental data [Jyotsna and Vanka (1995)]. The  $y$  distance is normalised with the cavity depth ( $D$ ) which gives  $y/d = 0$  at the cavity lid and  $y/d = -1$  at the bottom vertex. Similarly, the  $u$  velocity is normalised with



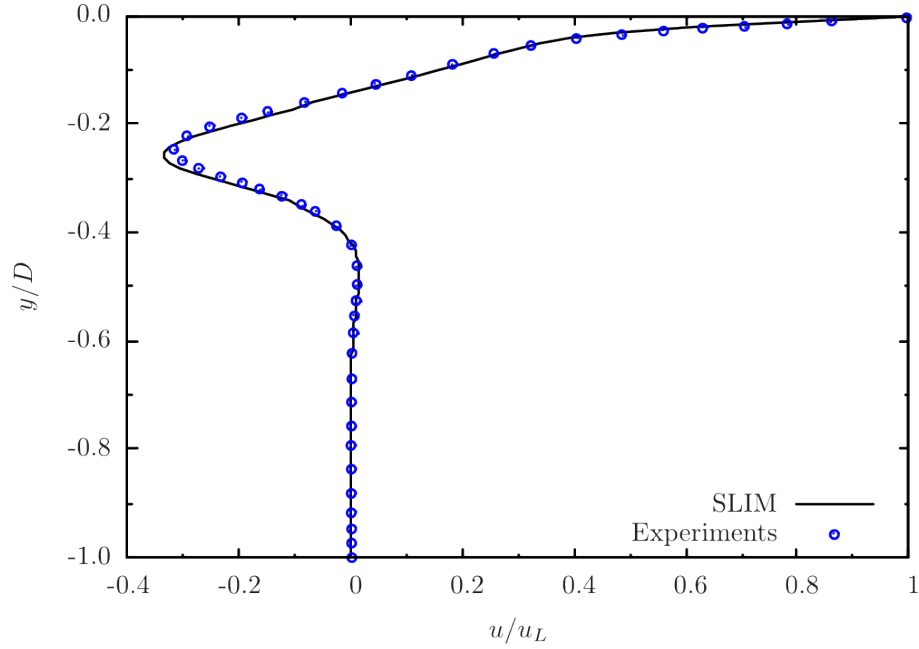
**Figure 8: Hybrid grid for the triangular cavity.**

the velocity of the cavity lid ( $u_L$ ). As seen in Fig. 9 the comparison with the experiment is excellent.

### *2D Circular Cylinder*

In this validation study, laminar incompressible flow over a 2D circular cylinder is investigated at a Reynolds number of 100. This classical configuration represents flow over a bluff body dominated by a wake region. For flows having a low Reynolds number ( $40 \leq Re_D \leq 150$ ), periodic vortex shedding occurs in the wake. The phenomenon of vortex shedding behind bluff bodies is referred to as the Karman Vortex Street [Roshko (1954)] and provides a transient case for CFD code validation.

In his work, [Roshko (1954)] experimentally studied wake development behind 2D circular cylinders from Reynolds number ranging from 40 to 10000. For Reynolds numbers of 40 to 150, the so called the stable range [Roshko (1954)], regular vortex streets are formed with no evidence of turbulence motion in the wake. Therefore, at a Reynolds number of 100, the vortex shedding exhibits smooth, coherent structures making it ideally suited for validating laminar CFD calculations. The frequency associated with the oscillations of the vortex streets can be characterised by the Strouhal Number ( $St$ ). The Strouhal Number is a non-dimensional number defined as



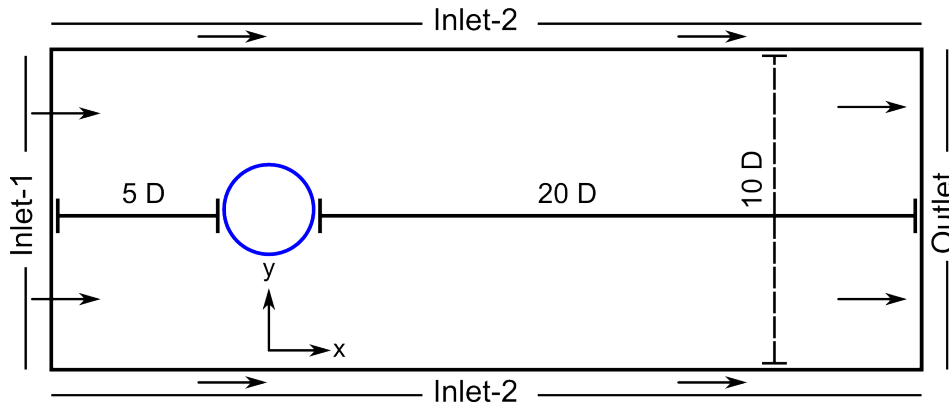
**Figure 9: Comparison of experimental and computational  $x$  velocity distribution along the cavity's centre-line.**

$$St = \frac{fD}{u} \quad (14)$$

where,  $f$  is the frequency of oscillations of vortex shedding,  $D$  is the diameter of the cylinder and  $u$  is the freestream velocity of the flow. Experimentally [Roshko (1954)], it has been determined that for a Reynolds number based on the diameter of the cylinder of 100, the Strouhal number  $St \approx 0.16 - 0.17$ . The main objective of this study was to compare the  $St$  for the SLIM calculation to the experimental data of [Roshko (1954)]. Provided the cylinder has a sufficient span length, the flow characteristics can be assumed to be 2D as the experiments suggest.

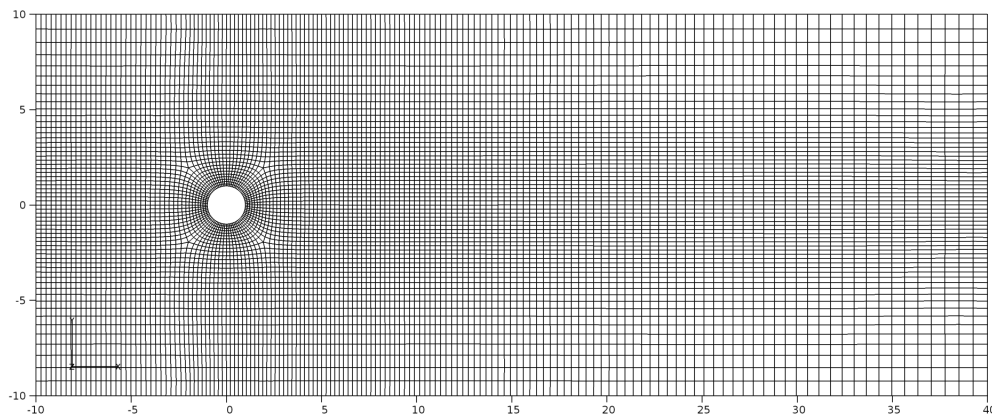
Fig. 10 shows the schematic of the 2D circular cylinder. Here, the diameter  $D = 2 m$  and is the characteristic length for the Reynolds number. For simplicity, the freestream velocity was taken to be  $u = 1 m/s$  in the  $x$ -direction. Using these values the kinematic viscosity was calculated to be  $0.02 m^2/s$ . The domain extends by 5 diameters upstream and 20 diameters downstream. In the  $y$  direction, the domain extends 5 diameters on either side. From the figure, multiple inlet boundaries to this domain can be seen, one at the upstream boundary and the other two for the top and bottom boundaries. This type of configuration is needed to appropriately model the inflow, similar to an undisturbed flow in an experimental set-up. It is noted that for top and bottom boundaries, the flow is in the  $x$  direction. The outlet is located at the downstream boundary. The cylindrical wall is a no-slip boundary condition. The solution is initialised with a uniform velocity  $u = 1.0 m/s$  and uniform zero pressure field.

The computational mesh was generated using Pointwise in the  $x - y$  plane and subsequently extruded one cell in the  $z$  direction. The mesh consisted of 9,260 cells. An O-grid topology was constructed around the cylinder with 10 cells in the radial direction and 84 cells in the circumferential direction. 31 cells were used



**Figure 10: Computational domain of a circular cylinder.**

upstream of the O-grid, in the  $x$  direction while 100 cells were used downstream. The region of interest is about 10 diameters downstream, where the grids are refined. In the  $y$  direction, 21 cells were used above and below the O-grid region.



**Figure 11: O-grid around the circular cylinder.**

A time-dependent simulation was carried out using the SLIM solver. For the discretization of time-dependent terms, the 2<sup>nd</sup> order backward scheme was used. Pressure and velocity gradients were calculated using the Gauss method. A 2<sup>nd</sup> order linear upwind discretization was used for the divergence of velocity. A linear corrected scheme was used for Laplacian discretization to account for mesh non-orthogonality.

To capture the transient start-up process, the simulation was started from time  $t = 0$  s and was simulated up to  $t = 360$  s, while lift and drag forces over the cylindrical surface were monitored at a frequency of 2 Hz. It was found that the on-set of vortex shedding occurred after about  $t = 90$  s which was then followed by a steady shedding process. A Fast Fourier transformation (FFT) was carried out on the lift force data and the peak frequency of vortex shedding occurred at  $f = 0.0888$  Hz. Based on this value, it takes about 7.8 cycles for the shedding to start. Table 2 compares the computed value from SLIM with that of the experiment. The agreement is good given that experimental uncertainty can be



relatively high at low Reynolds numbers.

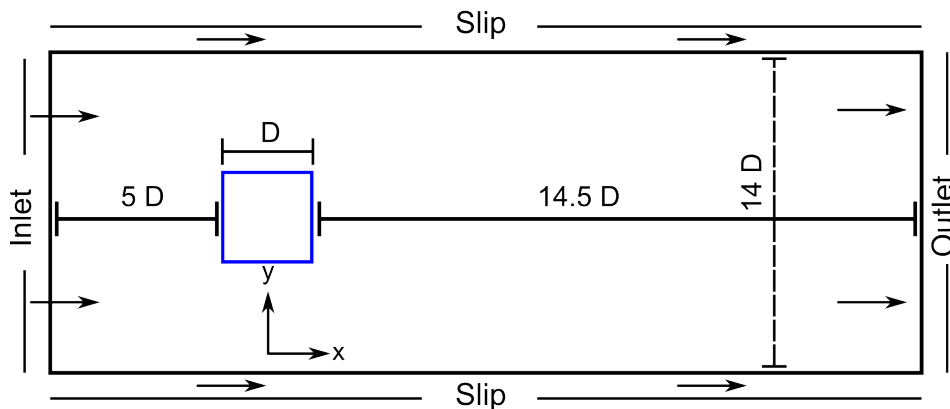
**Table 2: Comparison of experimental and numerical frequency results for the circular cylinder.**

	Frequency (Hz)	Strouhal Number
Experimental	0.0835	0.167
SLIM	0.0888	0.177

### *Square cylinder*

This case considers the turbulent, incompressible flow around a square cylinder, as studied experimentally by [Lyn and Rodi (1994) and Lyn et al. (1995)]. The side of the square cylinder ( $D$ ) is 1  $m$  and it extends along the width ( $4D$ ) of the channel. All distances are made non-dimensional with reference to  $D$ . The mean velocity at the inlet,  $u$ , is assumed to be 0.214  $m/s$ . All velocities are made non-dimensional using the inlet velocity. The Reynolds number, based on  $u$  and  $D$  is 21,400.

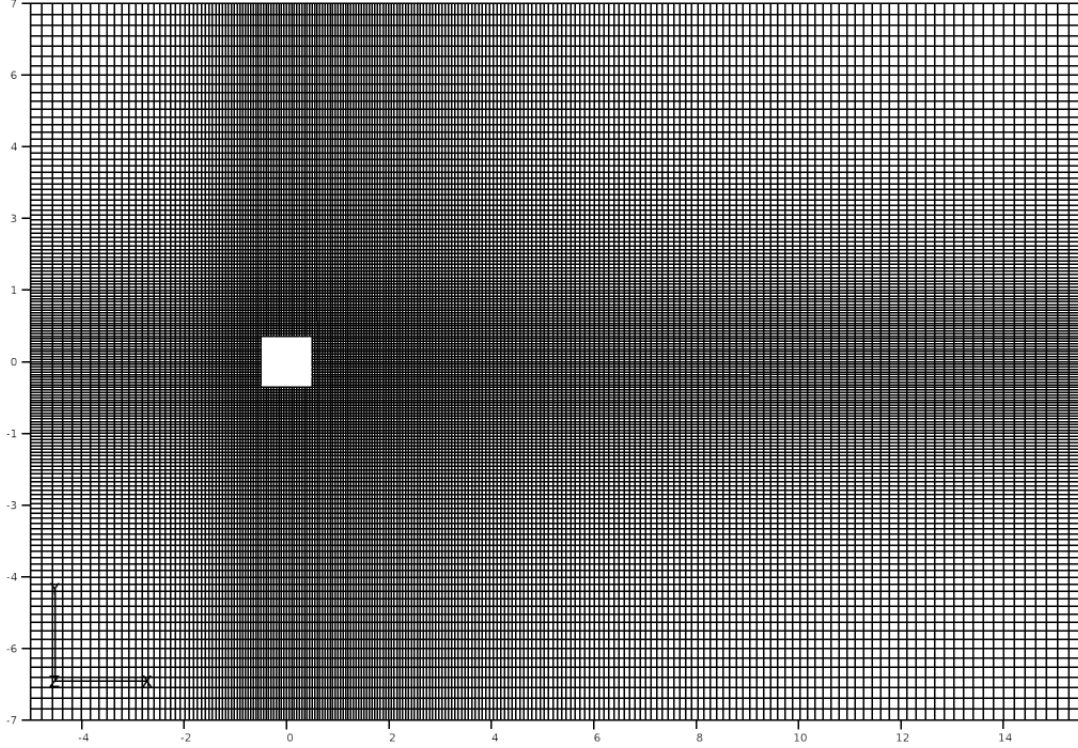
A rectangular computational domain in the  $x - y$  plane was constructed surrounding the square cylinder as shown in Fig. 12. The domain extends in the  $z$  direction a length of  $4D$ . The domain extends by 5 diameters upstream and 14.5 diameters downstream. In the  $y$  direction, the domain extends 7 diameters on either side. At the inlet a uniform velocity  $u = 0.214 m/s$  in  $x$  direction and a zero pressure gradient are applied. At the outlet a uniform zero pressure and zero velocity gradient conditions are applied. The top and bottom boundaries have a slip condition applied to velocity and a zero gradient condition applied to pressure. Cyclic boundaries conditions are used in the span wise direction and a no-slip boundary on the cylinder. The simulation is initialised with a uniform velocity  $u = 0.214 m/s$  in  $x$  direction and zero pressure field.



**Figure 12: Computational domain of the square cylinder.**

The non-uniform computational mesh shown in Fig. 13 was generated using Pointwise and consisted of  $216 \times 150 \times 21$  (668,850) cells (being denser near the cylinder). For the discretization of time-dependent terms, the  $2^{nd}$  order backward

scheme was used. Pressure and velocity gradients were calculated using the Gauss method. A  $2^{nd}$  order linear upwind discretization with multidimensional interpolation linear scheme utilising Barth-Jespersen limiter [Berger et al. (2005)] was used for the divergence of velocity.



**Figure 13: Square cylinder mesh.**

Large Eddy Simulation (LES) computes the large-scale motions of the flow directly. The small-scale, dissipative motions of turbulence tend to be more amenable to modelling because of their more uniform character, whereas the large-scale motions contain the majority of the energy and anisotropy. As a result, LES is expected to be more accurate, particularly in complex flows where the assumptions inherent to RANS models rarely exist. The drawback is that LES simulations are always three-dimensional and unsteady. For this particular case, the Smagorinsky SGS model [Smagorinsky (1963)] was chosen. The model is based on the SGS kinetic energy  $k = \frac{1}{2} (\overline{\mathbf{u}^2} - \tilde{\mathbf{u}}^2)$  where  $\tilde{\mathbf{u}}$  is the filtered velocity. The following assumptions for the SGS stress tensor ( $\mathbf{B}$ ) and the filtered deviatoric part of the rate of strain tensor ( $\tilde{\mathbf{D}}_D$ ) in incompressible flows are used

$$\mathbf{B} = \left[ \frac{2}{3} k \mathbf{I} - 2\nu \tilde{\mathbf{D}}_D \right], \quad (15)$$

$$\tilde{\mathbf{D}}_D = \left[ \tilde{\mathbf{D}} - \frac{1}{3} (tr \tilde{\mathbf{D}}) \mathbf{I} \right], \quad (16)$$

$$\tilde{\mathbf{D}} = \frac{1}{2} [\nabla \tilde{\mathbf{u}} + \nabla \tilde{\mathbf{u}}^T], \quad (17)$$

where  $\mathbf{I}$  is the unit tensor. The Smagorinsky model is an algebraic model for the SGS viscosity  $\nu_{SGS}$ . Caelus 5.04 implements the SGS viscosity as

$$\nu_{SGS} = c_k \sqrt{(k_{SGS})} \Delta \quad (18)$$

where the  $k_{SGS}$  is given by

$$k_{SGS} = \frac{(c_k \Delta^2)}{c_\epsilon} \|\widetilde{\mathbf{D}}\|^2. \quad (19)$$

and  $\Delta$  represents the top-hat filter with a characteristic filter width estimated as the cubic root of the cell volume. The relationship between the classical  $C_s$  constant and the constants  $c_k$  and  $c_\epsilon$  from the Smagorinsky model implementation in Caelus 5.04 is

$$C_s = \left( \frac{c_k^3}{c_\epsilon} \right)^{0.25}. \quad (20)$$

In this work the value used for  $C_s$  was 0.1. In some circumstances the turbulent viscosity near the wall is over predicted by the LES SGS model due to a lack of constraint on the turbulent viscosity. To remedy this, damping is added to the length scale ( $\Delta$ ) using the [Van Driest (1956)] formulation. In Caelus 5.04 the implementation takes the form

$$\Delta = \min \left[ \Delta_{cr}, \frac{\kappa y}{C_\Delta} \left( 1 - e^{-\frac{y}{y^* A^+}} \right) \right], \quad (21)$$

where  $\Delta_{cr}$  is the filter width calculated from the cubic root of the cell volume,  $y$  is the distance from the wall,  $C_\Delta$ ,  $A^+$ ,  $\kappa$  are constants and  $y^*$  is given by the formula  $y^* = \nu/u_\tau$ . Here  $u_\tau$  is the wall friction velocity.

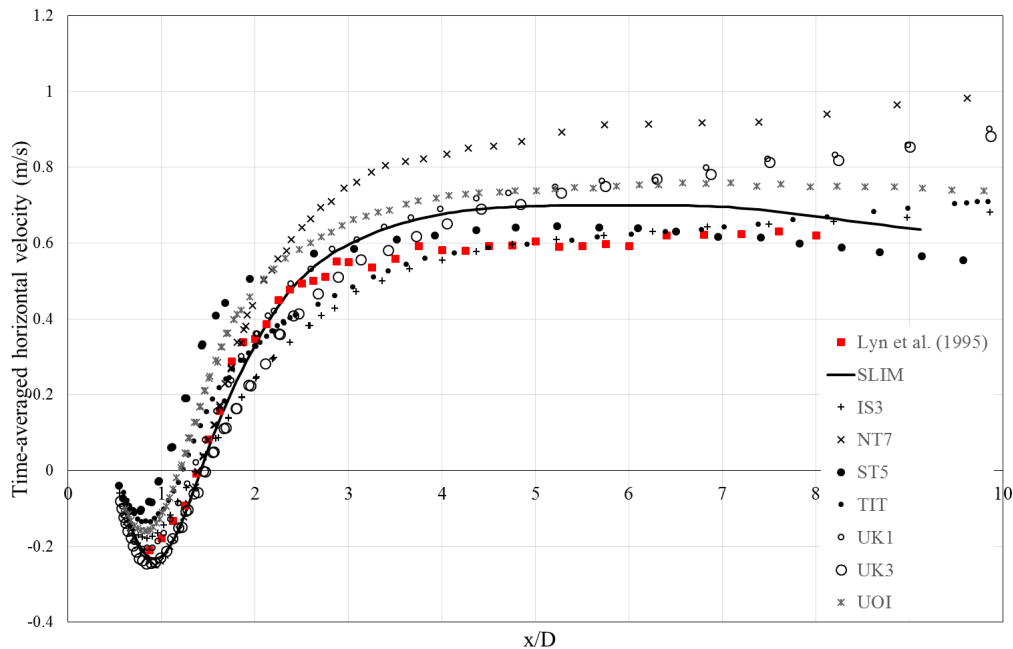
The simulation was started from time  $t = 0$  s and was simulated up to  $t = 1000$  s using a fixed Courant number of 1, while lift and drag forces over the cylindrical surface were monitored. It was found that the on-set of vortex shedding occurred after about  $t = 100$  s which was then followed by a steady shedding process. The velocity field was spanwise and time averaged for the period between  $200 \leq t \leq 1000$  s. A Fast Fourier transformation (FFT) was carried out on the lift force data and the peak frequency of vortex shedding occurred at  $f = 0.028$  Hz. Based on this value, it takes about 2.7 cycles for the shedding to start. The most important time-averaged parameters are presented in Table 3, where they are compared to experimental and numerical data, available from several authors from the ERCOFTAC database. The labels used are the same as in [Voke (1997)]. The parameters compared are the recirculation length, Strouhal number, the drag and lift coefficients and the R.M.S variation of the drag and lift coefficients.

The values predicted by SLIM agree reasonably well with the corresponding experimental and numerical data. The drag and lift coefficients are slightly greater than experimental data but agree as well as other numerical data. Fig. 14 shows the normalised time-averaged horizontal velocity along the centreline 10 diameters downstream of the cylinder. The experiment shows the velocity reaching approximately 0.6 of the freestream value and then leveling off thereafter. It can be observed from this figure that there is a fair amount of disparity among the numerical results. Some of the LES simulations, in particular UK1, UK3 and

**Table 3: Comparison among time-averaged squared cylinder data. The labels are the same as used in [Voke (1997)].**

Set	$l_r$	St	$C_D$	$C_L$	$\overline{C_D}$	$\overline{C_L}$
IS3	1.24	0.133	2.79	-0.125	0.36	1.68
NT7	1.39	0.131	2.05	-0.050	0.12	1.39
ST3	1.24	0.150	2.66	-0.005	0.27	1.33
TIT	1.23	0.131	2.62	0.0093	0.23	1.39
UK1	1.32	0.130	2.20	-0.020	0.14	1.01
UK3	1.44	0.130	2.23	-0.050	0.13	1.02
UOI	1.20	0.130	2.03	0.0400	0.18	1.29
[Lyn et al. (1995)]	1.38	0.132	2.1			
SLIM	1.41	0.131	2.44	0.076	0.33	1.29

NT7, show the velocity approaching the free stream value; others such as ST5 and UOI show a distinct decline beyond  $x/D = 5$ . The results from this work agree reasonably well with the experimental data.



**Figure 14: Streamwise distribution of the normalised time-averaged horizontal velocity along the centreline  $y = 0$  for the experimental and numerical data sets.**

The prediction of fluctuating horizontal and vertical velocities along the centre

line are shown in Figs. 15 and 16 respectively. These figures also display disparity amongst the numerical simulations and reveal that no simulation matches closely. The result from this work shows a slightly higher peak in the horizontal velocity fluctuation just behind the cylinder but does a much better job than most of the other simulations at matching the experimental data beyond this point. The fluctuations in the vertical velocity show reasonable agreement with the experimental data over the entire measurement region.

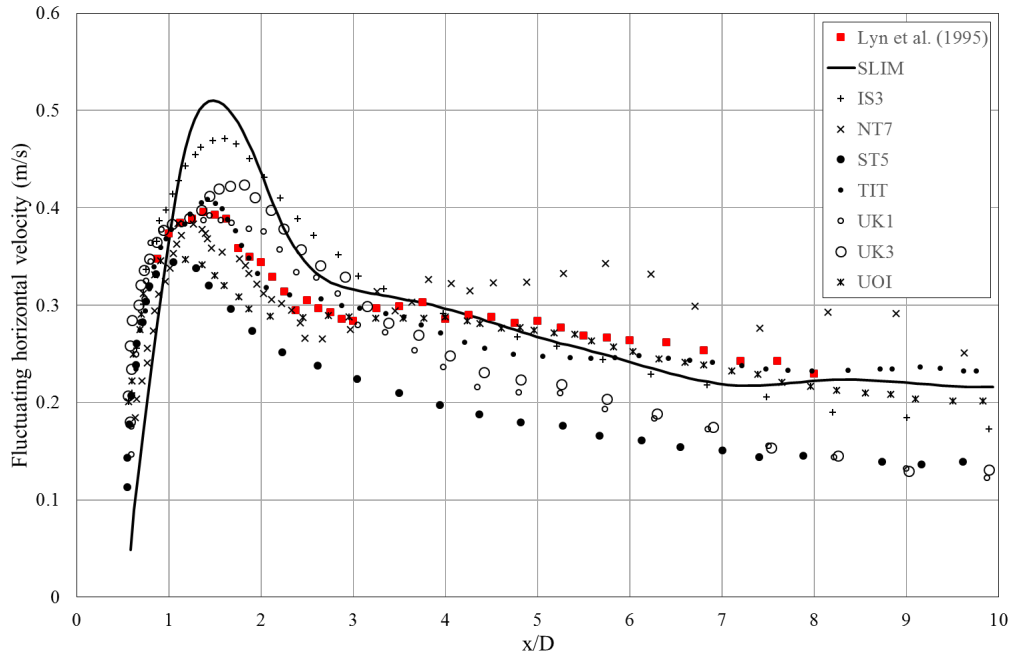
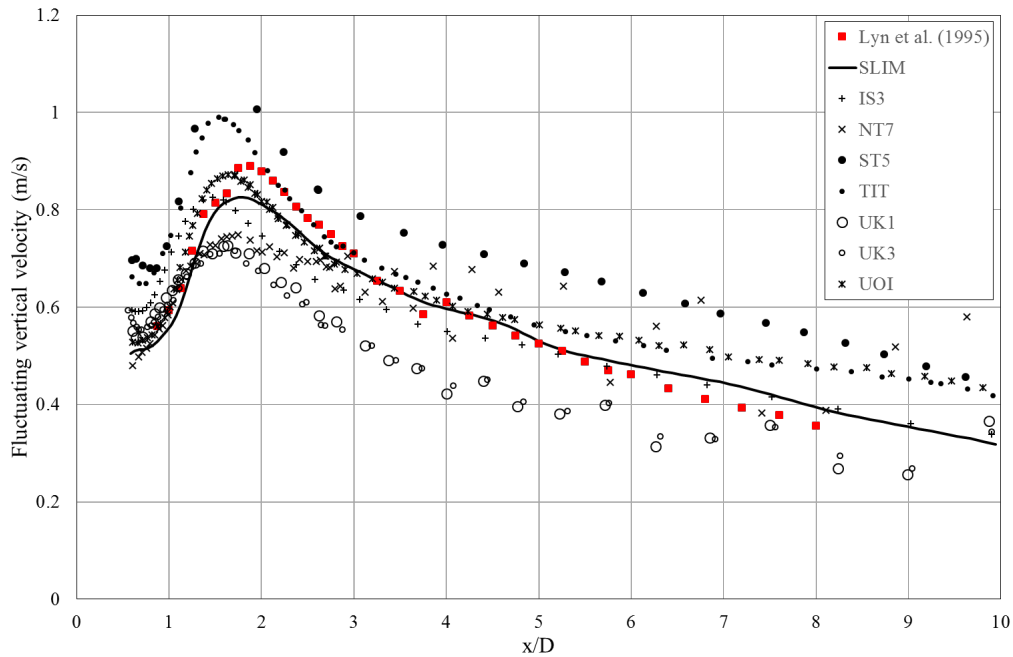


Figure 15: Streamwise distribution of the normalised horizontal velocity fluctuation along the centreline  $y=0$  for the experimental and numerical data sets.



**Figure 16: Streamwise distribution of the normalised vertical velocity fluctuation along the centreline  $y = 0$  for the experimental and numerical data sets.**

## Conclusions

In this paper, the Semi-Linear Implicit Method (SLIM) algorithm was described. The algorithm has many benefits including the exact satisfaction of the continuity equation for each time step. In addition, the algorithm is computationally efficient due to the geometric nature of the pressure Laplacian and the hyperbolic nature of the momentum equations. The algorithm was implemented as a solver using version 5.04 of the Caelus library and its accuracy tested through several validation cases comprising steady and transient laminar problems along with a transient turbulent case.

## References

- Applied CCM (2015). Caelus: Computational Mechanics Library (Version 5.04) [Software]. Available at <http://www.caelus-cml.com>.
- Berger, M., Aftosmis, M.J., and Murman, S.M. (2005). Analysis of Slope Limiter on Irregular Grids. *43rd AIAA Aerospace Sciences Meeting*, Jan. 10-13, 2005, Reno, NV
- Blasius H. (1908). Grenzschichten in Flüssigkeiten mit kleiner Reibung. *Z. Math. Phys.*, **56**, 1–37.
- Chorin A.J. (1967). A numerical method for solving incompressible viscous flow problems. *J. Comp. Phys.*, **2**, 12–26.
- Hayes, R.E., Nandkumar, K. and Nasr-El-Din, H. (1989). Steady Laminar Flow in a 90 Degree Planar Branch. *Computers and Fluids*, **17**(4), 537–553.
- Issa, R. (1985). Solution of the implicitly discretised fluid flow equations by operator splitting. *J. Comp. Phys.*, **61**, 40.

- Jyotsna, R. and Vanka, P. (1995). Multigrid Calculation of Steady, Viscous Flow in a Triangular Cavity. *J. Comp. Phys.*, **122**, 107–117.
- Lyn, D.A. and Rodi, W. (1994). The flapping shear layer formed by flow separation from the forward corner of a square cylinder. *J. Fluid Mech.*, **267**, 353.
- Lyn, D.A., Einav, S., Rodi, W. and Park, J.H. (1995). A laser-Doppler velocimetry study of ensemble averaged characteristics of the turbulent near wake of a square cylinder. *J. Fluid Mech.*, **304**, 285.
- NPARC Alliance. (2008). NPARC Alliance Validation Archive. [ONLINE] Available at <http://www.grc.nasa.gov/WWW/wind/valid/fplam/fplam.htm>. [Accessed 13 May 15].
- Patankar, S. V., and Spalding, D.B. (1972). A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows. *Int. J. of Heat and Mass Transfer*, **15(10)**: 1787–1806.
- Pointwise Inc. (2014). Pointwise, Mesh Generation Software for CFD, Version 17.2R2. Fort Worth, TX. <http://www.pointwise.com/>.
- Roshko, A. (1954). On the Development of Turbulent Wakes from Vortex Streets. Technical Report TR-1191, National Advisory Committee for Aeronautics.
- Smagorinsky, J. (1963). General Circulation Experiments with Primitive Equations, *Month. Weath. Rev.*, **93**, 99-165.
- Strikwerda, J.C., and Lee, Y.S. (1999). The accuracy of the fractional Step method. *SIAM J. Numer. Anal.*, Vol. **37**, No. **1**, 37–47.
- Sutherland, W. (1893). The viscosity of gases and molecular force. *Philosophical Magazine*, **S. 5**, 507–531.
- Van Driest, E.R. (1956). On the turbulent flow near a wall, *J. Aeronautical Sci.*, **23**, 1007-1011.
- Voke, P. (1997) *Direct and Large-Eddy Simulation II*, chapter Flow past a Square Cylinder: Test Case LES2. Kluwer Academic.

## **Novel multimedia architecture design pattern using Audio joiner prototype with delay remover**

†Sachin M. Narangale<sup>1</sup> and \*G. N. Shinde<sup>2</sup>

<sup>1</sup>School of Media Studies, Swami Ramanand Teerth Marathwada University,  
Nanded-431606, India

<sup>2</sup>Indira Gandhi College, Nanded-431603, India

\*Presenting author: shindegn@yahoo.co.in

†Corresponding author: sachin.narangale@gmail.com

### **Abstract**

In this paper a novel multimedia architecture design pattern (MADP) is proposed which uses controlled synchronization signals. The architecture can be tested for multimedia production. It is expected that the synchronization signals will play a vital role in multimedia production. The advantages of design pattern and design formation are clearly used. The prototype presented in this paper focuses on control signal with the feature of delay remover. The multimedia architectures used for framing has importance over synchronization; control signal with delay remover presents a novel architecture for audio joiner prototype. The problems of journalism mass communication for audio transfer, broadcast and presentation are affected at the time of audio mixing by the problems of synchronization and delay occurred in multimedia formation. This algorithm will set new benchmark in commercial, educational, communication, entertainment multimedia products.

**Keywords:** Audio, Delay remover, Binary Joiner, Multimedia, Synchronization

### **Introduction**

Multimedia plays a vital role in different sectors like groupware, video on demand services, video conference, electronic shopping systems or entertainment systems. Even Multimedia enabled learning is very important in modern education. Teaching and learning process is nowadays dependent on interactive whiteboards, multimedia projectors, e-presentations. Modern Education policies are becoming increasingly dependent on Multimedia with quality. This society is in need of better Multimedia Architectures to fulfill the need of new emerging, real world problem oriented curriculum, effective analysis system and strong backbone to education system. Patterns are simple and elegant explanations. Patterns capture solutions that have developed and evolved over time. They are abstractions being used to increase reuse and quality in a variety of fields including architecture, software engineering, hypermedia, and teaching/learning [Jones, D. & Stewart, S. (1999)]. A design pattern is a three-part rule, which expresses a relation between a certain context, a problem, and a solution [Christopher Alexander (1979)]. The Multimedia Architecture revolves around the three-part rule. One has to define the relation between multimedia objects (text, graphics, animation, audio and video). Then identify the tool/ module that are expected. The tool/ module are the outcome of the exercise. After visualizing the relationship between multimedia objects and the tool/ module it is the time to work out for a solution. Mere investigation of solution to the multimedia tool will not solve everybody's problem. More need of the Multimedia Tools is prone to defects, limitation of use, non-relationship with real world problems, and redundancy of development efforts. There has to be a design pattern, which will trigger the generation of multimedia architecture and promote multimedia tool/ module reuse. Design Patterns can and should be used to help develop advanced Multimedia Architecture and implement the concept of modular objects.



## **Multimedia**

Multimedia is two or more media run continuously. Multimedia is a combination of more than two media such as text, graphics, animation, audio and video. The production of multimedia components varies from component to component. When written text, speech, photography, music, video and graphics are combined and integrated in digital texts, we are dealing not only with the convergence of media forms. On a more fundamental level it involves a convergence of semiotic systems, reading conventions and rhetorical patterns [Martin Engebretsen (2006)]. Today, certain media formats are of particular interest. When we are dealing with written text, graphics, sound produced, animation we are not only concerned with media objects; rather we are keen to bring a scenario into highlight where a complete multimedia will satisfy the zeal. Multimedia information is rapidly growing both in importance and in diversity [Jack Meadow (1998)].

## **Multimedia Architecture**

The Multimedia comprises of various media. The collection, production with synchronization can cause the overhead to the production systems. The production systems, therefore, should have a mechanism to overcome the problem of overhead elimination. Following are few architecture designs with specific features. CSI (Complex Streamed Instruction) eliminates overhead instructions (such as instructions for data sectioning, alignment, reorganization, and packing/unpacking) often needed in applications utilizing MMX -like extensions and accelerates key multimedia kernels [Cheresiz, D., Juurlink, B., Vassiliadis, S., Wijshoff, H.A.G. (2005)]. CCM (Community coordinated multimedia) envisions the paradigm of consuming multiple media via diversity display devices, converged networks, and heterogeneous platforms within a virtual, open and collaborative community.

Multimedia Application development industry must demand for specialization in multimedia architecture during the recruitment of multimedia-application-developers [Katre Dinesh (2005)].

## **Design Pattern**

The question is whether to find a solution or to find a method that find solution to a problem? A pattern for software architecture describes a particular recurring design problem. The problem arises in specific design contexts. So the pattern presents a well-proven generic scheme for its solution. The solution scheme is specified by describing its constituent components, their responsibilities and relationships, and the ways in which they collaborate. A design pattern is a method of using the knowledge about problem and its solution continuously. A pattern describes the problem and need of its solution. Design pattern straight-a-way tackles about the method of finding the solution of the problem.

## **Controlled Synchronization Signal Algorithm for Multimedia Architecture Design Pattern:**

### **a. Introduction to Controlled Synchronization**

In this paper we propose a new algorithm with controlled synchronization signals. The MADP with controlled signals will help in reducing efforts, redundancy. Multimedia requirement analysis will start with object identification. Identification of size and parameters of modular objects will help in basic architecture definition. The component requirement validity will remove redundancy. The selective approach to High Level Features and Low Level Features of MARS (Multimedia Analysis and Retrieval System) for indexing of objects will resolve the complexities of object modularization [G.N.Shinde, S. B. Kurumbhatte (2002)]. The object interval, features of objects: discrete and continuous should go under temporal synchronization

specifications with logical mapping [Saul E. Pomares Hernandez, Luis A. Morales Rosales, Jorge Estudillo Ramirez, Gustavo Rodriguez Gomez (2008)]. The object precedence based on their status: Active, Passive has to be identified. There should be a clear notational way to design the object interaction [Rhan Jung, and Soung Won Kim (2009)]. As multimedia components are collection of features, and a query based approach seeks the information out of multimedia components, a refined query fine tunes the weight best suitable to user perception [G.N.Shinde, S. B. Kurumbhatte (2002)].

The synchronization model should possess the support for the generation of synchronization specifications. This ability concerns three aspects: specification maintainability, specification reusability and consistency checking [Saul E. Pomares Hernandez, Luis A. Morales Rosales, Jorge Estudillo Ramirez, Gustavo Rodriguez Gomez (2008)].

In an object identified, the color of a pixel generally cannot be independent from its neighbours. The objects of a same region will have texture effects [Mina Koleini, S. Amirhassan Monadjemi, Payman Moallem (2009)]. The modular objects should contain the database for pixel information (size, shape, color, hue, luminance, etc.). The algorithm sets and retrieves the values as required.

In the audio visual synchronization Processor Scheduling for reducing wait time will be controlled.

wait time ( $t_{wa}$ ) for audio

& wait time ( $t_{wv}$ ) for video

Processor Scheduling ( $tP$ )  $\rightarrow$  ( $t_{wa}$ ) ( $t_{wv}$ ) and Control Signal

Here, we propose a Controlled Synchronization Signal Algorithm for multimedia architecture design pattern.

## 1. Requirement Analysis

### a. Object Identification

#### i. Object Specification

Specify the objects/ components that comprise the tool/ module. The complete listing of the objects/ components will enable the categorization of them and their classification.

#### ii. Object Occurrence

The object occurrence and duration in the story is to be known well in advance.

The occurrence will decide whether to make the appearance of the object/ component to be visible or not.

#### iii. Object modularity

#### iv. Integration Parameters

### b. Object Synchronization

#### i. Object Selection

#### ii. Component Precedence Algorithm

#### iii. Component Relay (Baton Relay)

#### iv. Control signal

#### v. Verification

## 2. Design Formation

### a. Script writing

### b. Multimedia Aspect Preparation

#### i. Component Tokenization and Injection

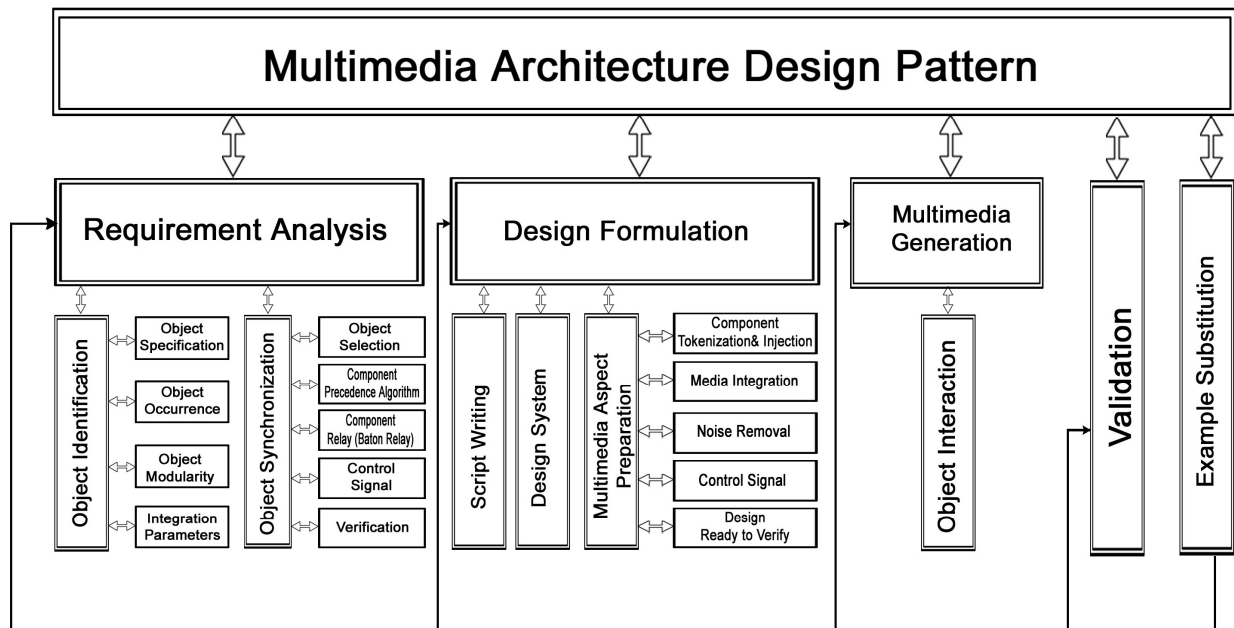
#### ii. Media Integration

#### iii. Noise Removal

- iv. Control Signal
- v. Design-ready-to-verify
- c. Design System
- 3. Multimedia Generation
  - a. Object Interaction
 

The objects are now ready for interaction. Control signals will mix the objects in desired sequence.
- 4. Example Substitution
- 5. Validation

**Flowchart**



**Conclusion**

This design pattern helps multimedia designers to manage the complexity of the multimedia tool/module. When redundancy in problem is identified, the generalized algorithm can be implemented productively to provide a reliable multimedia tool/module. The architecture design presented in this paper with delay remover is effective. The processor scheduling and control signals of delay remover are a part of a novel design in this prototype. This design pattern also describe about the multimedia components, the communication between these components, and the mechanism of building these components for obtaining a real world multimedia tool/module. This pattern defines a way of communication using delay remover technology between multimedia-application-developers. The architecture is useful for communication industry for removing delay problems.

**References**

Jones, D. & Stewart, S., “The case for patterns in online learning”, In Proceedings of WebNet World Conference on the WWW and Internet, 1999, pp. 592-597.

Christopher Alexander, “The Timeless Way of Building”, Oxford University Press, 1979

Martin Engebretsen, “Making sense with multimedia. A text theoretical study of a digital format integrating writing and video” Seminar.net - International journal of media, technology and lifelong learning, 2006, Vol. 2 – Issue 1

Jack Meadow, “Multimedia Information and Library & Information Staff”, DESIDOC Bulletin of Information Technology, Nov 1998, Vol. 18, No. 6, pp.51-54

Cheresiz, D., Juurlink, B., Vassiliadis, S., Wijshoff, H.A.G., “The CSI multimedia architecture”, Very Large Scale Integration (VLSI) Systems, JAN 2005, Vol. 13, Issue 1, pp. 1–13.

Katre Dinesh, “Multimedia Education: Current Trends and Future Potential”, Sampada Special Issue on Education, Mahratta Chamber of Commerce, Industries and Agriculture, July 2005, pp. 27-29.

- G.N.Shinde, S. B. Kurumbhatte, "Features of Multimedia Analysis and Retrieval System (MARS)", *Library Progress (International)*, 2002, Vol. 22, No.1, pp. 77-80.
- Rhan Jung, and Soung Won Kim, "Multimedia Real Time Systems using CBD", *International Journal of Software Engineering and Its Applications*, APR 2009, Vol.3, No.2 pp. 60-79.
- Saul E. Pomares Hernandez, Luis A. Morales Rosales, Jorge Estudillo Ramirez, Gustavo Rodriguez Gomez, "Logical Mapping: An Intermedia Synchronization Model for Multimedia Distributed Systems", *Journal of Multimedia*, DEC 2008, VOL. 3, NO. 5, pp. 33-41.
- Mina Koleini, S. Amirhassan Monadjemi, Payman Moallem, "Film Colorization Using Texture Feature Coding and Artificial Neural Networks", *Journal of Multimedia*, VOL. 4, NO. 4, AUGUST 2009, pp 240-247.

## **Innovative development of the RP assisted customized surgical guides in various surgeries**

†Sandeep W. Dahake<sup>1</sup>, \*Abhaykumar M. Kuthe<sup>2</sup>, Mahesh B. Mawale<sup>3</sup>, Ashutosh D. Bagde<sup>4</sup>

<sup>1,2,3,4</sup>Department of Mechanical Engineering, Visvesvaraya National Institute of the Technology, Nagpur, India.

\*Presenting author: amkme2002@yahoo.com

†Corresponding author: sandeepdahake@students.vnit.ac.in

### **Abstract**

The unique capabilities of the Computer Aided Design (CAD) and rapid prototyping (RP) have been useful for designing and manufacturing of the customized surgical guides (CSGs). In complex surgeries the CSGs have been shown to provide an accurate means of transferring virtual surgical planning (VSP) to actual surgery. The main purpose of this study is to highlight the different areas where RP assisted CSGs have been prominently use. This paper describes the successful application of RP for the production of CSGs in twelve different areas of complex surgeries. Rapid manufacturing of CSGs using RP technique and their applications; for drilling, cutting and positioning of the implant in various complex surgeries is given in this paper. Otherwise these types of complex surgeries could not be accurately treated through traditional surgeries. Different case reports are noted in this study to know the various innovative ideas related to the RP assisted CSGs. The detail methodology for development of CSGs is also described in this paper. The applications reported here have demonstrated that RP is a viable process for the manufacturing of CSGs. This approach shows good results in designing and manufacturing of the CSGs. Using advanced tools; VSP, CAD and RP for development of CSGs for the accurate surgery are possible easily. The main aim of this paper is to represent the applications of RP technology to direct manufacturing of CSGs in various complex surgeries.

**Keywords:** Virtual surgical planning, Computer aided design, Rapid prototyping

### **Introduction**

The CSG is a jig designed and manufactured by an engineer with surgeon input in complex surgeries using CAD/CAM technology and RP. Designing and manufacturing of CSGs are possible according to the pre-operative VSP of the engineer and surgeon with the aim to transfer this plan into the operating theatre [Oliveira et al., (2008); Fantini et al., (2013)]. CAD/CAM and VSP has improve preoperative planning and understanding of the surgical anatomy and has facilitate the digital preparation of CSGs. It consists of intraoperative instruments that transfer the VSP to the operating field for the exact resection, drilling or accurate placement of the implant in the surgery site. This process have the potential of reducing the major negative changes in the patient's quality of life that can lead to low self confidence and negative self perception [Logan et al. (2013)]. In this paper the twelve different applications of CSGs in various complex surgeries and manufacturing methodology for the innovative development is discussed.

### **Methodology for development of CSGs-**

#### *Radiology*

The scan data of CT, CBCT or MRI in DICOM format are generally useful for the development of the CSGs. These are the inputs to develop the CSGs.

#### *Image processing*

By using various advance image processing software like Mimics (Materialise NV, Leuven, Belgium) the collected data of CT, CBCT or MRI scan in DICOM format have to be convert into RP compatible STL format for the development of the CSGs. A thresholding technique is useful in order to highlight the areas of interest using exact threshold level for bone. A 3D model of diseased

part can quickly and easily create with the segmentation done. In the case of the diseased anatomy, reconstruction times ranged from 2 minutes for basic models where minor details were ignored, to half an hour for model where all details were included such as surface indentations. The sophisticated measurement tools in software provide accurate dimensions of complex geometries, allowing for accurate measurements and greater insight into the morphology of the diseased anatomy. The precise measurements which can be determined using software aid in the choice of the most appropriate CSGs design and surgery planning. Finally obtained 3D model of diseased anatomy have to export into STL format for the designing of CSGs.

#### *Virtual surgical planning*

The advance tools in software has capable to eliminate the inaccuracies commonly associated with traditional surgery planning and simplify the execution by eliminating surgical steps such as intraoperative measurement, marking of implant positioning site, etc. [Polley and Figueroa (2013)]. The computerized plan can be easily transfer accurately to the surgical site to position the implant, taking cut or drilling at the time of surgery using VSP [Hsu et al. (2013)]. Before design the CSGs the VSP plays vital role to achieve the accuracy in proper fitting of the implant, drilling or cutting on accurate site. VSP decides the size of implant, thickness of implant, screw positioning of implant, shape of implant, exact match of the implant, also it plan for design of CSGs as per the need like drilling, cutting or positioning of the implant. The main advantage of VSP compared with conventional planning is that it significantly reduces the laborious manual steps [Scheepers et al. (2012)].

#### *CAD of CSG*

Based on 3D model of the affected anatomy the optimal CSGs design generally plan. Surgery simulation on the 3D CAD model of affected anatomy makes it possible to begin designing CSGs in currently available advance CAD software like 3Matics. To create these CSGs, the bone surface was inversed which assured of a perfect fit. From a surgical point of view, the CSGs facilitates the identification of the most appropriate site with regard to bone volume, blood supply, and precisely assists in resection of the desired bony segment in a correct angle at the recipient site [Dérاند et al. (2012)]. The wrap tool is used to design the CSGs for the accurate transfer of contour of affected anatomy on the guide. The designing procedure of the CSGs is given in “Table 1”.

**Table 1. Design methodology**

S.N.	Design methodology
1	Radiology
2	Image processing
3	Virtual surgical planning
4	CAD
5	Export into RP compatible STL format

#### *Manufacturing of CSGs*

After completion of the design of the CSGs, it has to save in STL i.e. RP compatible format for manufacturing. Various authors used various RP techniques for the fabrication of the CSGs. Among them FDM, SLA, SLS, 3DP are the common RP techniques use to easily manufacture the CSGs in plastic material. As per requirement DMLS, SLM, EBM are some RP techniques are used by different researchers to fabricate CSGs in metallic form.

Even though the virtual model shows that there is a perfect fit, the researchers will have to play it safe, for that mock test of fitting and cutting with their own hands is important. In order to do so

prior to a real surgery, the RP plays dominant role to produce the diseased model and corresponding CSGs. If surgeon satisfied with the mock test the CSG can use in the surgery after sterilization. The manufacturing procedure of the CSGs is given in “Table 2”.

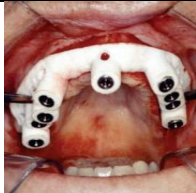



**Table 2. Methodology for the manufacturing of the CSGs**








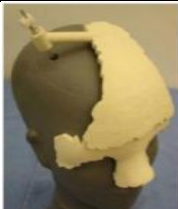
S.N.	Manufacturing process steps
1	Slice STL file using RP software and link with RP machine
2	Fabrication of CSG using RP machine
3	Post processing for finishing
4	Try in before surgery on RP assisted diseased model

**Applications of RP assisted CSG for various surgeries**

Author noted the twelve different areas where the RP assisted CSGs are prominently use in the complex surgeries. Among them dentistry and craniomaxillofacial (CMF) are the common fields where CSGs have been used tremendously. The various application of the RP assisted CSGs in various surgeries are given in the “Table 3”.

**Table 3. Application of RP assisted CSG for various surgeries**

S.N	Area	Application	RP assisted CSG	Researchers
1	Dentistry	<b>CSG for drilling-</b> Drilling and placement of the implant at exact position and orientation		Giacomo et al. 2014
2	CMF	<b>CSG for resection-</b> Accurate resection of the bone tumor		St-Hilaire et al. 2012
3	PSI	<b>CSG for drilling-</b> Accurate insertion of the screws in the pedicles		Porada et al.
4	THA	<b>CSG for resection-</b> Accurate resection of the femur head		Drstvensek et al. 2013

5	TKN	<b>CSG for resection-</b> Accurate resection of the knee		Bagaria et al. 2011
6	CO	<b>CSG for resection-</b> Accurate resection and placement of the plate		Dobbe et al. 2013
7	BRAR	<b>CSG for resection-</b> Accurate bone resection and allograft reconstruction		Bellanova et al. 2013
8	OO	<b>CSG for resection-</b> Exact cutting of the cancer affected part		Khan et al. 2013
9	MTPJ	<b>CSG for drilling-</b> Exact placement of k wires for twisting		Hirao et al. 2014
10	OIP	<b>CSG for drilling-</b> Placement of implants at accurate location in the orbit		Unknown author
11	TMJ	<b>CSG for resection-</b> Exact resection and positioning of the scaffold		Ciocca et al. 2009
12	NS	<b>CSG for drilling-</b> Used in the deep brain stimulation		Rajon et al. 2006

CMF- craniomaxillofacial, PSI- pedicle screw insertion, THA- total hip arthroplasty, TKR- total knee replacement, CO- corrective osteotomy, BRAR- bone resection and allograft reconstruction, OO- orthopedic oncology, MTPJ- metatarsophalangeal joint surgery, OIP- orbital implant placement, TMJ- temporomandibular joint surgery, and NS- neurosurgery

## Discussion

In this paper the development of the RP assisted CSGs and their applications in twelve different areas of medical in complex surgeries have reported. According to the researchers, it is really very



easy to find the best fit for positioning the CSGs manually, because of no significant free motion while the placement in position with slightly pressing against the affected bone. Although the technique requires a very clean preparation of the bone surface, including removal of the attached muscle and fat tissue, it still proves to be better than other techniques reported in the literature. There are in fact several distinct advantages of using advance software to design the CSGs. The surgeon can decide on the location, orientation, and size of the customized implant based on the unique morphology of the patient prior to the surgery. Also it helps to take proper cut, drill holes in the complex sites in the surgery. This technique is very simple to use, so it does not require a specific expertise on the surgeon's part. It eliminates the need for complex equipment and time consuming procedures in the operation theatre, thereby reducing the operation time considerably. Screws for fixation implant can be accurately place without perforating the nerves. The need for fluoroscopy during implant placement and screw insertion is eliminated, which considerably reduces the radiation exposure to the patient and the members of the surgical team [Lu et al. (2012); Cansiz et al. (2013)]. This is important as surgeons these days are tackling more challenging anatomy than ever so the valuable information provided by the 3D reconstructions ensures that the surgeons don't encounter any unforeseen problems during surgery; this in turn increases confidence in the procedure. Using this technique the 3D reconstruction and manufacturing of CSGs is quick to perform and also aid surgeons in improving the treatment. With CSG's wide applicability, high accuracy, proven safety and cost-effectiveness, RP assisted CSGs will likely enjoy widespread use in the future in various complex surgeries.

## Conclusions

This study shows that RP is a viable rapid manufacturing method for the direct production of CSGs for various complex surgeries. The time and cost for development of CSGs using RP is comparable to other manufacturing processes. The CSGs produced using the RP process have been shown to be acceptable in terms of accuracy, quality of fit and function to surgical guides in twelve different prominent surgical areas which could not be otherwise possible easily using traditional manufacturing methods.

## References

- Bagaria, V., Rasalkar, D., Bagaria, S. and Ilyas, J. (2011) Medical applications of rapid prototyping - a new horizon, *INTECH*,1-20.
- Bellanova, L., Paul, L. and Docquier, P. (2013) Surgical guides (patient-specific Instruments) for pediatric tibial bone sarcoma resection and allograft reconstruction, *Hindawi Publishing Corporation Sarcoma* 2013,1-7.
- Cansiz, E., Arslan, Y., Turan, F. and Atalay, B. (2013) Design and production of a novel computer assisted patient specific sagittal split osteotomy guide and soft tissue retractor, *J. Med. Biol. Eng.*
- Ciocca, L., De Crescenzo, F., Fantini, M. and Scotti, R. (2009) CAD/CAM and rapid prototyped scaffold construction for bone regenerative medicine and surgical transfer of virtual planning: A pilot study, *Computerized Medical Imaging and Graphics* 33, 58-62.
- Dérand, P., Rännar, L. and Hirsch, J. (2012) Imaging, virtual planning, design, and production of patient-specific implants and clinical validation in craniomaxillofacial surgery, *Craniomaxillofacial Trauma and Reconstruction* 5, 137-144.
- Dobbe, J., Vroemen, J., Strackee, S. and Streekstra, G. (2013) Patient-tailored plate for bone fixation and accurate 3D positioning in corrective osteotomy, *Med Biol Eng Comput* 51, 19-27.
- Drstvensek, I., Brajljih, T., Paulič, M., Balic, J. and Tomazic, T. (2013) Patient specific instruments for total hip replacement surgery, *Academic journal of manufacturing engineering* 11, 6-9.
- Fantini, M., Crescenzo, F. and Ciocca, L. (2013) Design and manufacturing of customized surgical devices for mandibular rehabilitation, *Int J Interact Des Manuf* 7, 227-237.
- Giacomo, G., Silva, J., Martines, R. and Ajzen, S. (2014) Computer designed selective laser sintering surgical guide and immediate loading dental implants with definitive prosthesis in edentulous patient: A preliminary method, *European Journal of Dentistry* 8,100-106.
- Hirano, S. and Minakuchi, S. (2013) Factors affecting accuracy of implant placement with mucosa-supported stereolithographic surgical guides in edentulous mandibles, *Computers in Biology and Medicine*, (<http://dx.doi.org/10.1016/j.combiomed.2013.07.029>).

- Hsu, S., Gateno, J., Bell, R., Hirsch, D., Markiewicz, M., Teichgraber, J., Zhou, X. and Xia, J. (2013) Accuracy of a computer-aided surgical simulation protocol for orthognathic surgery: A prospective multicenter study, *J Oral Maxillofac Surg* 71, 128-142.
- Khan, F., Lipman, J., Pearle, A., Boland, P. and Healey, J. (2013) Computer-generated Custom Jigs Improve Accuracy of Wide Resection of Bone Tumors, *Clinical Orthopaedics and Related Research* 471, 2007–2016.
- Logan, H., Wolfaardt, J., Boulanger, P., Hodgetts, B. and Seikaly, H. (2013) Exploratory benchtop study evaluating the use of surgical design and simulation in fibula free flap mandibular reconstruction, *Journal of Otolaryngology - Head and Neck Surgery* 42, pp. 1-9.
- Lu, S., Zhang, Y., Shi, J., Chen, Y. and Xu, X. (2012) Accuracy and efficacy of thoracic pedicle screws in scoliosis with patient-specific drill template, *Med Biol Eng Comput* 50, 751–758.
- Oliveira, M., Hussain, N., Dias, A., Lopes, M., Azevedo, L., Zenha, H., Costa, H. and Santos, J. (2008) 3-D biomodelling technology for maxillofacial reconstruction, *Materials Science and Engineering C* 28,1347–1351.
- Polley, J. and Figueroa, A. (2013) Orthognathic positioning system: intraoperative system to transfer virtual surgical plan to operating field during orthognathic surgery, *J Oral Maxillofac Surg* 71, 911–920.
- Porada, S., Millner, P., Chiverton, N., Berry, E. and Seedhom, B. Computer aided surgery with lumbar vertebral drill-guides, using computer aided planning, design and visualization.
- Rajon, D., Bova, F., Bhasin, R. and Friedman, W. (2006) An investigation of the potential of rapid prototyping technology for image-guided surgery, *Journal of Applied Clinical Medical Physics* 7, 1-16.
- Schepers, R., Raghoobar, G., Lahoda, L., Meer, W., Roodenburg, J., Vissink, A., Reintsema, H. and Witjes, M. (2012) Full 3-D digital planning of implant-supported bridges in secondary mandibular reconstruction with prefabricated fibula free flaps, *Head and Neck Oncology* 4, pp. 1-6.
- St-Hilaire, H. (2012) Reconstruction of a mandibular osteoradionecrotic defect with fibula osteocutaneous flap using synthes proplan CMF, patient specific plate contouring (PSPC) and the matrixmandible plating system.
- Unknown author, Fabrication of a Custom fitting surgical template for orbital implant placement: A case report. (<http://www.yuhuihuang.com/SurgicalTemplateCaseReport.pdf> Date-12/8/2014)

## Higher-order two-point efficient family of Halley type methods for simple roots

†Ramandeep Behl and S. S. Motsa

School of Mathematics, Statistics and Computer Sciences, University of KwaZulu-Natal,  
Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa

†Corresponding author: ramanbehl87@yahoo.in

December 9, 2015

### Abstract

In this paper, we proposed a new highly efficient two-point sixth-order family of Halley type methods that do not require any second-order derivative evaluation for obtaining simple roots of nonlinear equations, numerically. In terms of computational cost, each member of the family requires two function and two first-order derivative evaluations per iteration. On the account of the results obtained, it is found that our proposed methods are efficient and show better performance than existing sixth-order methods available in the literature. Further, it is also noted that larger basins of attraction belong to our methods although the others methods are slow and has darker basins while some of the method are too sensitive upon the choice of the initial value.

**Keywords:** Nonlinear equations, Simple roots, Halley's method, Basins of attractions, Order of convergence

### Introduction

Efficient solution techniques are required for finding simple roots of nonlinear equation of the form

$$f(x) = 0, \quad (1)$$

where  $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$  is a nonlinear sufficiently differentiable function in an interval  $D$ , which partake of scientific, engineering and various other models. One of the best known one-point optimal second-order method based on two functional evaluations is the classical Newton's method [Traub (1964); Petković et al. (2012)]. Many methods have been developed which improve the convergence rate of the Newton's method or Newton like at the expense of additional evaluations of functions or derivatives.

Halley's method [Traub (1964); Petković et al. (2012)] is the third-order modification of Newton's method, which is defined as follows:

$$x_{n+1} = x_n - \frac{2f(x_n)f'(x_n)}{f(x_n)f''(x_n) - 2\{f'(x_n)\}^2}. \quad (2)$$

Despite the cubic convergence, this method is considered less practical from a computational point of view because of the costly second-order derivative evaluation. Therefore, researchers introduced multi-point methods and the primary aim of these methods is to achieve as high as possible order of convergence using a fixed number of function evaluations. However, multi-point methods do not use higher order derivatives and has great practical importance since they overcome the theoretical limitations of one-point methods regarding their convergence order and computational efficiency.

Therefore, a number of sixth-order methods are also appearing as the extensions of Newton’s method or Newton like method to solve nonlinear equation (1). In [Neta (1979)] given a three-point sixth-order general iteration scheme for obtaining simple roots of nonlinear equations, which is defined as follows:

$$\left\{ \begin{array}{l} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \frac{1 + \beta u}{1 + (\beta - 2)u} \frac{f(y_n)}{f'(x_n)} = y_n - G_f(u) \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = z_n - \frac{1 - u}{1 - 3u} \frac{f(z_n)}{f'(x_n)}, \end{array} \right. \quad (3)$$

where  $u = \frac{f(y_n)}{f(x_n)}$ ,  $G_f(u) = \frac{u(1+\beta u)}{1+(\beta-u)}$ ,  $\beta \in \mathbb{R}$ .

In [Sharma and Ghua (2011)], proposed three-point family of sixth-order methods based on fourth-order Ostrowski’s method [Kanwar et al. (2011)], which is given by

$$\left\{ \begin{array}{l} y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ z_n = y_n - \frac{f(x_n)}{f(x_n) - 2f(y_n)} \frac{f(y_n)}{f'(x_n)}, \\ x_{n+1} = z_n - \frac{f(x_n) + af(y_n)}{f(x_n) + (a - 2)f(y_n)} \frac{f(z_n)}{f'(x_n)}, \quad a \in \mathbb{R}. \end{array} \right. \quad (4)$$

On the other hand, [Wang and Liu (2009)] have constructed two three-point sixth-order families of Jarratt’s method [Petković et al. (2012); Behl et al. (2013)] requiring two of functions and two of first-order derivative evaluations per iteration, one of them is defined as follows:

$$\left\{ \begin{array}{l} y_n = x_n - \frac{2}{3} \frac{f(x_n)}{f'(x_n)}, \\ z_n = x_n - \frac{9 - 5w}{10 - 6w} \frac{f(x_n)}{f'(y_n)}, \quad w = \frac{f'(y_n)}{f'(x_n)} \\ x_{n+1} = z_n - \frac{a + bw}{c + dw + rw^2} \frac{f(z_n)}{f'(x_n)}, \end{array} \right. \quad (5)$$

where  $a = \frac{5c+3d+r}{2}$ ,  $b = \frac{r-3c-d}{2}$ ,  $c + d + r \neq 0$ ,  $a, b, c, d, r \in \mathbb{R}$ .

But, the body structures of above mentioned three-point sixth-order methods are more complicated as compared with two-point methods [Kanwar et al. (2011); Behl et al. (2013)]. Further, it is very rare to find two-point methods whose order of convergence higher than four [Guem et al. (2015)]. Nowadays, obtaining new two-point methods of order six not requiring the computation of a

second-order derivative, is very important and interesting task from the practical point of view,

Therefore, the principle aim of this manuscript is to provide a new highly efficient two-point sixth-order class of Halley type methods, that do not require any second-order derivative evaluation for obtaining simple roots of nonlinear equations, numerically. It is also observed that the body structures of our proposed families of methods are simpler than the existing three-point families of sixth-order methods. Further, our proposed methods are more effective in all the tested examples to the existing three-point sixth-order methods available in the literature. Further, we also compare them with two-point sixth-order methods that is very recently proposed by [Guem et al. (2015)] and it is found that our methods are better than these methods. Furthermore, the dynamic study of our methods also supports the theoretical aspects.

### Development of two-point sixth-order methods

In this section, we intend to develop many new families of sixth-order Halley type methods, not requiring the computation of second-order derivative. For this purpose, we consider  $w_n = x_n - \frac{f(x_n)}{f'(x_n)}$ , a Newton's iterate. With the help of Taylor series, we expand the function  $f(w_n)$  about a point  $x = x_n$  as follows:

$$f(w_n) \approx f(x_n) + f'(x_n)(w_n - x_n) + \frac{1}{2}f''(x_n)(w_n - x_n)^2, \text{ which further implies}$$

$$f''(x_n) \approx \frac{2\{f'(x_n)\}^2 f(w_n)}{\{f(x_n)\}^2}. \tag{6}$$

Similarly, expanding the function  $f'(w_n) = f'\left(x_n - \frac{f(x_n)}{f'(x_n)}\right)$  about a point  $x = x_n$  by Taylor series expansion, we have  $f'(w_n) \approx f'(x_n) + f''(x_n)(w_n - x_n)$ , which further yields

$$f''(x_n) \approx \frac{f'(x_n)(f'(x_n) - f'(w_n))}{f(x_n)}. \tag{7}$$

Now, taking the arithmetic mean of two equations (6) and (7), we get another approximation of  $f''(x_n)$  as follows:

$$f''(x_n) \approx \frac{\frac{2\{f'(x_n)\}^2 f(w_n)}{\{f(x_n)\}^2} + \frac{f'(x_n)(f'(x_n) - f'(w_n))}{f(x_n)}}{2}. \tag{8}$$

Inserting this approximate value of  $f''(x_n)$  in scheme (2), and using the weight function on the second step, we get

$$\begin{cases} w_n = x_n - \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} = x_n - \frac{4\{f(x_n)\}^2}{3f(x_n)f'(x_n) + f(x_n)f'(w_n) - 2f'(x_n)f(w_n)} M_f(h, k), \end{cases} \tag{9}$$

where the weighting function  $M_f$  is a sufficient differential function with  $h = \frac{f'(x_n)(f(x_n)+f(w_n))}{f(x_n)f'(w_n)}$  and  $k = \frac{f(w_n)}{f(x_n)}$ . Theorem 1 indicates that under what choices on the weight function (9), the order of convergence will reach at six without using any more functional evaluations.

### Convergence analysis

**Theorem 1** *Let a sufficiently smooth function  $f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$  has a simple zero  $\xi$  in the open interval  $D$ . Assume that initial guess  $x = x_0$  is sufficiently close to  $\xi$ . Then, the iterative scheme defined by (9) has sixth-order convergence when*

$$M_{00} = 1, M_{01} = \frac{3}{4}, M_{10} = -\frac{1}{4}, M_{20} = \frac{1}{2}, M_{11} = \frac{1}{2}, M_{02} = -\frac{5}{2}, M_{03} = 9(6M_{30} + 3M_{21} + 5), \\ M_{12} = -3(3M_{30} + 2M_{21} + 3), M_{04} = -3\{4M_{13} + 9(2M_{22} + 4M_{31} + 3M_{40} - 8)\}, \quad (10)$$

where  $M_{ij} = \frac{\partial^{i+j}}{\partial h^i \partial k^j} M_f(h, k)|_{(h=1, k=0)}$ . It satisfies the following error equation

$$e_{n+1} = -\frac{c_2}{12} \left[ 12(M_{13} + 18M_{21} + 9M_{22} + 54M_{30} + 27M_{31} + 27M_{40} - 10)c_2^4 - \{2M_{13} + 3(24M_{21} + 6M_{22} + 72M_{30} + 18M_{31} + 18M_{40} - 5)\}c_2^2c_3 + 6(4 + M_{21} + 3M_{30})c_3^2 - 12c_2c_4 \right] e_n^6 + O(e_n^7). \quad (11)$$

**Proof** Let  $\xi$  be a simple zero of  $f(x)$ . With the help of Taylor's series, we get the following expansion of  $f(x_n)$  and  $f'(x_n)$  around  $x = \xi$

$$f(x_n) = f'(\xi) \left( e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + e_n^5c_5 + e_n^6c_6 + O(e_n^7) \right), \quad (12)$$

and

$$f'(x_n) = f'(\xi) \left( 1 + 2e_nc_2 + 3e_n^2c_3 + 4e_n^3c_4 + 5e_n^4c_5 + 6e_n^5c_6 + 7e_n^6c_7 + O(e_n^7) \right), \quad (13)$$

respectively. By using equations (12)–(13), we get

$$f(w_n) = f'(\xi) \left( c_2e_n^2 - 2(c_2^2 - c_3)e_n^3 + (5c_2^3 - 7c_2c_3 + 3c_4)e_n^4 - 2(6c_2^4 - 12c_2^2c_3 + 3c_3^2 + 5c_2c_4 - 2c_5)e_n^5 + (28c_2^5 - 73c_2^3c_3 + 34c_2^2c_4 - 17c_3c_4 + c_2(37c_3^2 - 13c_5) + 5c_6)e_n^6 + O(e_n^7) \right). \quad (14)$$

and

$$f'(w_n) = f'(\xi) \left( 1 + 2c_2^2e_n^2 - 4(c_2^3 - c_2c_3)e_n^3 + c_2(8c_2^3 - 11c_2c_3 + 6c_4)e_n^4 - 4c_2(4c_2^4 - 7c_2^2c_3 + 5c_2c_4 - 2c_5)e_n^5 + 2(16c_2^6 - 34c_2^4c_3 + 6c_3^3 + 30c_2^3c_4 - 13c_2^2c_5 - 8c_2c_3c_4 + 5c_2c_6)e_n^6 + O(e_n^7) \right). \quad (15)$$

By using equations (12)–(15), we obtain

$$h = \frac{f'(x_n)(f(x_n) + f(w_n))}{f(x_n)f'(w_n)} = 1 + 3c_2e_n + (5c_3 - 3c_2^2)e_n^2 - 7(c_2c_3 - c_4)e_n^3 + (6c_2^4 - 3c_2^2c_3 - 2c_3^2 - 10c_2c_4 + 9c_5)e_n^4 + \{31c_2^3c_3 - 12c_2^5 - c_2^2c_4 - 5c_3c_4 - c_2(11c_3^2 + 13c_5) + 11c_6\}e_n^5 + O(e_n^6). \quad (16)$$

and

$$k = \frac{f(w_n)}{f(x_n)} = c_2e_n + (-3c_2^2 + 2c_3)e_n^2 + (8c_2^3 - 10c_2c_3 + 3c_4)e_n^3 + (-20c_2^4 + 37c_2^2c_3 - 8c_3^2 - 14c_2c_4 + 4c_5)e_n^4 + (48c_2^5 - 118c_2^3c_3 + 51c_2^2c_4 - 22c_3c_4 + c_2(55c_3^2 - 18c_5) + 5c_6)e_n^5 + O(e_n^6). \quad (17)$$

Since it is clear from equations (16) – (17),  $h = 1 + u$  and  $k = O(e_n)$ . Then, from these equations, we get the remainder  $u = h - 1$  and  $k$  are infinitesimal with the same order of  $e_n$ . Therefore, we can expand weight function  $M(h, k)$  in the neighborhood of  $(1, 0)$  by Taylor series expansion up to fourth-order terms as follows:

$$\begin{aligned}
 M_f(h, k) = & M_{00} + M_{10}u + M_{01}k + \frac{M_{20}u^2 + 2M_{11}uk + M_{02}k^2}{2!} + \frac{1}{3!}(M_{30}u^3 + 3M_{21}u^2k \\
 & + 3M_{12}uk^2 + M_{03}k^3) + \frac{M_{40}u^4 + 4M_{31}u^3k + 6M_{22}u^2k^2 + 4M_{13}uk^3 + M_{04}k^4}{4!} \\
 & + O(e_n^5).
 \end{aligned}
 \tag{18}$$

Using equations (12) – (18), in scheme (9), we obtain

$$e_{n+1} = (1 - M_{00})e_n - c_2(M_{01} + 3M_{10})e_n^2 + \sum_{l=3}^6 H_l e_n^l,
 \tag{19}$$

where  $H_l = H_l(c_2, c_3, \dots, c_6)M_{ij}$ , for  $0 \leq i, j \leq 4$ .

From the equation (19), it is clear that by substituting the following values

$$M_{00} = 1, \quad M_{01} = -3M_{10},
 \tag{20}$$

we get at least third-order convergence. Further, using (20) into  $H_3 = 0$ , we find two independent relation as follows:

$$(1 + 4M_{10}) = 0, \quad (M_{02} + 12M_{10} + 6M_{11} + 9M_{20} - 2) = 0
 \tag{21}$$

After some simplification, we get

$$M_{10} = -\frac{1}{4}, \quad M_{02} = (5 - 6M_{11} - 9M_{20}).
 \tag{22}$$

By substituting equations (20) and (22) into  $H_4 = 0$ , we have

$$(M_{11} + 3M_{20} - 2) = 0, \quad (M_{03} + 9(4M_{11} + M_{12} + 12M_{20} + 3M_{21} + 3M_{30} - 4)) = 0.
 \tag{23}$$

Solving the above equation (23) for  $M_{11}$  and  $M_{03}$ , which further yields

$$M_{20} = -\frac{1}{3}(M_{11} - 2), \quad M_{03} = -9(4 + M_{12} + 3M_{21} + 3M_{30}).
 \tag{24}$$

By substituting equations (20), (22) and (24) into  $H_5 = 0$ , we obtain

$$\begin{cases}
 (2M_{11} - 1) = 0, \\
 (4M_{11} - M_{12} - 6M_{21} - 9M_{30} - 11) = 0, \\
 [M_{04} + 3(168 - 48M_{11} + 24M_{12} + 4M_{13} + 144M_{21} + 18M_{22} + 216M_{30} + 36M_{31} + 27M_{40})] = 0.
 \end{cases}
 \tag{25}$$

Solving the above equation for  $M_{20}$ ,  $M_{12}$  and  $M_{04}$ , we get

$$\begin{cases}
 M_{11} = \frac{1}{2}, \\
 M_{12} = -3(3 + 2M_{21} + 3M_{30}), \\
 M_{04} = -3\{4M_{13} + 9(2M_{22} + 4M_{31} + 3M_{40} - 8)\}.
 \end{cases}
 \tag{26}$$

We obtain the following error equation, by using equations (20), (22), (24) and (26) into (19)

$$e_{n+1} = -\frac{c_2}{12} \left[ 12(M_{13} + 18M_{21} + 9M_{22} + 54M_{30} + 27M_{31} + 27M_{40} - 10)c_2^4 - \{2M_{13} + 3(24M_{21} + 6M_{22} + 72M_{30} + 18M_{31} + 18M_{40} - 5)\}c_2^2c_3 + 6(4 + M_{21} + 3M_{30})c_3^2 - 12c_2c_4 \right] e_n^6 + O(e_n^7). \tag{27}$$

This reveals that the modified family of Halley type methods (9) reaches the order of convergence six by using only four functional evaluations (viz  $f(x_n)$ ,  $f'(x_n)$ ,  $f(w_n)$  and  $f'(w_n)$ ) per full iteration. This completes the proof.  $\square$

**Special cases**

In this section, we discuss some interesting special case of weight function defined in (9) by inserting the values of free disposable parameters and different forms of weight functions  $M_f(h, k)$ .

(1) For  $M_{40} = 0$ ,  $M_{31} = 0$ ,  $M_{13} = 0$ ,  $M_{22} = 4$  and  $M_{30} = 0$ , in (18), we get the following weight-function

$$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{3 + 2u + 2M_{21}u^2}{4}k - \left( \frac{5}{4} + \frac{3}{2}(3 + 2M_{21})u - u^2 \right) k^2 + \frac{3}{2}(5 + 3M_{21})k^3, \tag{28}$$

where  $M_{21}$  is a free variable and for the sake of simplicity  $u = h - 1$ . This is a new two-point sixth-order family of methods. For different specific values of  $M_{21}$ , we get various cases as well as two-point methods but some of the important cases describes in the following table 1.

**Table 1: Sub cases of weight function (28) and their error equations**

Particular values of $M_{21}$	Sub cases and their error equation
$M_{21} = 0$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{3+2u}{4}k + \left(u^2 - \frac{5}{4} - \frac{9u}{2}\right)k^2 + \frac{15k^3}{2},$ $\left(-26c_2^5 + \frac{19}{4}c_2^3c_3 - 2c_2c_3^2 + c_2^2c_4\right) e_n^6 + O(e_n^7).$
$M_{21} = -\frac{13}{9}$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{27+18u-26u^2}{36}k + \left(u^2 - \frac{5}{4} - \frac{u}{6}\right)k^2 + k^3,$ $\left(-\frac{47}{12}c_2^3c_3 - \frac{23}{18}c_2c_3^2 + c_2^2c_4\right) e_n^6 + O(e_n^7).$
$M_{21} = -\frac{5}{3}$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{9+6u-10u^2}{12}k + \left(u^2 - \frac{5}{4} + \frac{u}{2}\right)k^2,$ $\left(4c_2^5 - \frac{21}{4}c_2^3c_3 - \frac{7}{6}c_2c_3^2 + c_2^2c_4\right) e_n^6 + O(e_n^7).$

(2) For  $M_{40} = 0$ ,  $M_{31} = 0$ ,  $M_{13} = 0$ ,  $M_{22} = 4$  and  $M_{30} = -\frac{3M_{21}+5}{6}$  in (18), we obtain

$$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} - \frac{3M_{21} + 5}{36}u^3 + \frac{3 + 2u + 2M_{21}u^2}{4}k + \frac{4u^2 - 5 - 3(1 + M_{21})u}{4}k^2, \tag{29}$$

where  $M_{21}$  is a free variable. Therefore, some of the special cases given in the following table 2.



**Table 2: Sub cases of weight function (29) and their error equations**

Particular values of $M_{21}$	Sub cases and their error equation
$M_{21} = 0$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} - \frac{5u^3}{36} + \frac{3+2u}{4}k + \left(-\frac{5}{4} - \frac{3u}{4} + u^2\right)k^2,$ $\frac{1}{4}c_2(76c_2^4 - 41c_2^2c_3 - 3c_3^2 + 4c_2c_4)e_n^6 + O(e_n^7).$
$M_{21} = -\frac{19}{9}$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{u^3}{27} + \frac{27+18u-38u^2}{36}k + \left(-\frac{5}{4} + \frac{5u}{6} + u^2\right)k^2,$ $\left(-\frac{47}{12}c_2^3c_3 - \frac{23}{18}c_2c_3^2 + c_2^2c_4\right)e_n^6 + O(e_n^7).$
$M_{21} = -\frac{5}{3}$	$M_f(h, k) = 1 - \frac{u}{4} + \frac{u^2}{4} + \frac{9+6u-10u^2}{12}k + \left(-\frac{5}{4} + \frac{u}{2} + u^2\right)k^2,$ $\left(4c_2^5 - \frac{21}{4}c_2^3c_3 - \frac{7}{6}c_2c_3^2 + c_2^2c_4\right)e_n^6 + O(e_n^7).$

(3) We consider following weight function, that satisfies all the conditions which are mention in theorem 1

$$M_f(h, k) = \frac{1}{16} \left( 5 + \frac{27}{1+2h} + 4k - 20k^2 - 24k^3 + h(2+8k) \right). \quad (30)$$

(4) We consider another weight function, which is given by

$$M_f(h, k) = \frac{8+k-5k^2-6k^3+h^2(1+4k)+h(3+4k-10k^2-12k^3)}{4+8h}. \quad (31)$$

(5) We consider one more weight function, which is defined as follows:

$$M_f(h, k) = \frac{6+16k-11k^2-12k^3+45k^4+h^2(1-9k^2)+h(-3-k+24k^2)}{4+12k}. \quad (32)$$

### Numerical experiments

In this section, we apply new methods for ( $M_{21} = -\frac{13}{9}$ ) in scheme (28), for ( $M_{21} = -\frac{19}{9}$ ) in scheme (29), denoted by  $OM_6^1$  and  $OM_6^2$  respectively, to solve some nonlinear equations given in table 3, which serve to check the validity and efficiency of theoretical results. These methods are compared with method (5) for ( $c = 0, d = 1, r = 0$ ), proposed by [Wang and Liu (2009)], (called  $WM_6$ ) and method (3) for ( $\beta = 2$ ), proposed by [Neta (1979)], denoted by ( $NM_6$ ). Finally, we will also compare our schemes with a two-point family of sixth-order methods that is very recently proposed by [Guem et al. (2015)], between them we will choose their best expression (3.4, 3.8 and 3.12) denoted by ( $GM_6^1, GM_6^2$  and  $GM_6^3$ ), respectively. For better comparisons of our proposed methods, we have given three comparison tables in each example: one is corresponding to absolute error, the second one is with respect to number of iterations and third one is regarding their computational error in table 4, 5, 6, respectively. All computations have been performed using the programming package *Mathematica* 9 with multiple precision arithmetic. We use  $\epsilon = 10^{-34}$  as a tolerance error. The following stopping criteria are used for computer programs:

(i)  $|x_{n+1} - x_n| < \epsilon$  and (ii)  $|f(x_{n+1})| < \epsilon.$



**Table 6: (Computational order of different sixth-order methods )**

$f(x)$	$x_0$	$WM_6$	$NM_6$	$GM_6^1$	$GM_6^2$	$GM_6^3$	$OM_6^1$	$OM_6^2$
1.	0.85	5.999	6.000	5.999	5.999	6.000	6.000	6.000
	1.6	6.003	6.000	6.000	5.993	5.996	6.000	5.997
2.	4.5	6.000	6.000	5.956	5.952	5.994	6.010	6.001
	6.5	6.000	6.000	6.000	5.994	6.000	6.000	6.000
3.	-2.0	6.970	6.997	8.985	7.005	7.010	7.000	7.000
	-1.9	6.982	7.000	9.000	7.003	7.006	7.000	7.000
	1.9	6.982	7.000	9.000	7.003	7.006	7.000	7.000
	2.0	6.970	6.997	8.985	7.005	7.010	7.000	7.000
4.	0.86	6.000	C	C	C	6.000	6.005	6.002
	1.4	6.000	6.000	6.000	6.000	6.000	6.001	6.011
5.	1.2	5.984	5.980	5.988	5.990	5.994	5.996	6.000
	2.25	6.000	D	6.000	6.000	5.992	6.013	6.003

### Attractor basins in the complex plane

We here investigate the comparison of the attained simple root finders in the complex plane using basins of attraction. It is known that the corresponding fractal of an iterative root-finding method is a boundary set in the complex plane, which is characterized by the iterative method applied to a fixed polynomial  $p(z) \in \mathbb{C}$ , see e.g. [Scott et al. (2011); Neta et al. (2012); Behl and Motsa (2012)]. The aim herein is to use basin of attraction as another way for comparing the iteration algorithms.

From the dynamical point of view, we consider a rectangle  $D = [-3, 3] \times [-3, 3] \in \mathbb{C}$  with a  $400 \times 400$  grid, and we assign a color to each point  $z_0 \in D$  according to the simple root at which the corresponding iterative method starting from  $z_0$  converges, and we mark the point as black if the method does not converge. In this section, we consider the stopping criterion for convergence to be less than  $10^{-4}$  wherein the maximum number of full cycles for each method is considered to be 100. In this way, we distinguish the attraction basins by their colors for different methods.

**Test problem 1.** Let  $p_1(z) = (z^4 + 1)$ , having simple zeros  $\{-0.707107 - 0.707107i, -0.707107 + 0.707107i, 0.707107 - 0.707107i, 0.707107 + 0.707107i\}$ . It is straight forward to see from Fig. 1 – 2 that our methods, namely  $OM_6^1$  and  $OM_6^2$  contain lesser number of non convergent points, have a larger and brighter basin of attraction in comparison to the methods, namely  $WM_6$ ,  $NM_6$ ,  $GM_6^1$ ,  $GM_6^2$  and  $GM_6^3$ . Further, our methods do not show an chaotic behavior as method  $WM_6$ .

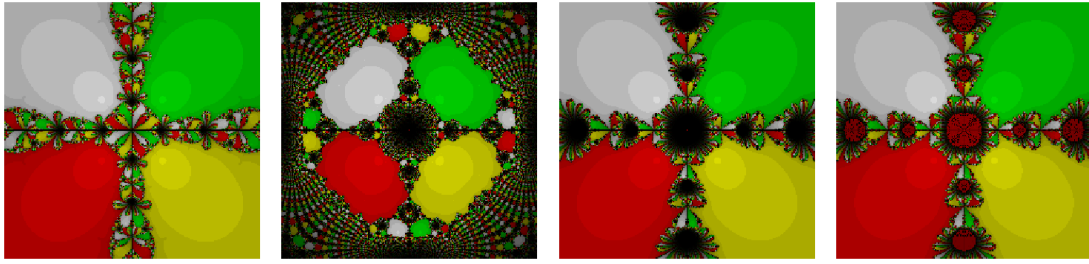


Figure 1: The methods  $WM_6$ ,  $NM_6$ ,  $GM_6^1$  and  $GM_6^2$ , respectively for test problem 1.

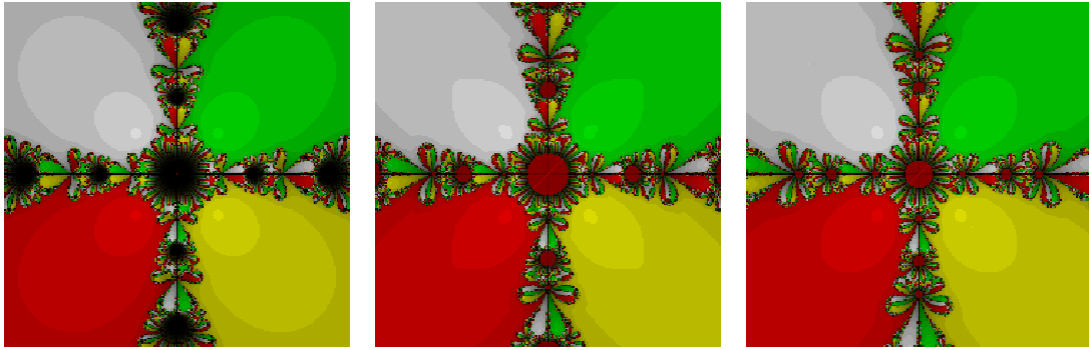


Figure 2: The methods  $GM_6^3$ ,  $OM_6^1$ , and  $OM_6^2$ , respectively for test problem 1.

**Test Problem 2.** Let  $p_2(z) = (z^6 + z)$ , having simple zeros  $\{-1, -0.309017 - 0.951057i, -0.309017 + 0.951057i, 0, 0.809017 - 0.587785i, 0.809017 + 0.587785i\}$ . We can easily observe from Fig. 3 – 4 that our proposed methods have larger and brighter basin of attraction in comparison to the mentioned methods.

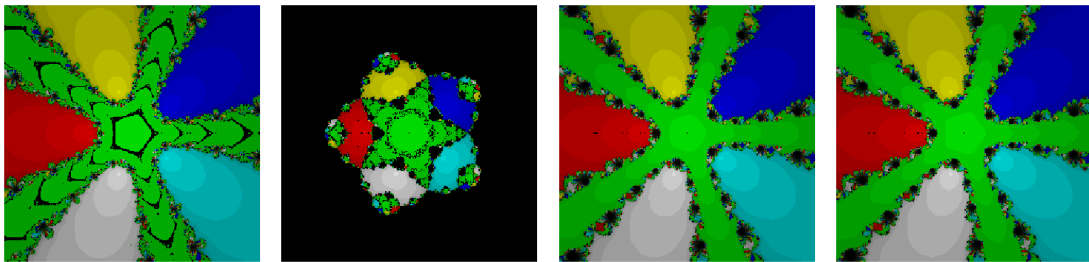
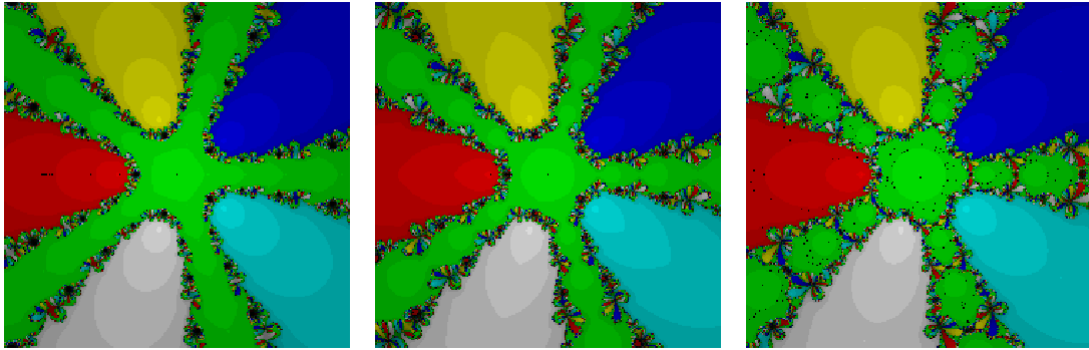


Figure 3: The methods  $WM_6$ ,  $NM_6$ ,  $GM_6^1$  and  $GM_6^2$ , respectively for test problem 2.



**Figure 4: The methods  $GM_6^3$ ,  $OM_6^1$  and  $OM_6^2$ , respectively for test problem 2.**

## Conclusions

In this paper, we proposed several second-derivative free families of Halley type methods based on weight function and arithmetic means of the approximated value of the second-order derivative. We can easily get several new methods by choosing different values of the disposable parameter  $M_{21}$  in schemes (28) and (29). Further, we can also obtain several families of sixth-order Halley-type method by considering different kind of weight functions which satisfy the conditions mentioned in Theorem 1. Each member of the proposed family requires two evaluations of the function  $f$  and two of its first-order derivative  $f'$  per full step. Our proposed iterative methods are compared in their efficiency and performance to various other multi-point methods in Table 4, 5, 6 and it is observed from these tables that our proposed methods are efficient and perform better than existing methods available in the literature. Based on Figs. 1 – 4, we conclude that larger basins of attraction belong to our methods namely,  $OM_6^1$  and  $OM_6^2$  although the others methods are slow and has darker basins while some of the method are too sensitive upon the choice of the initial value.

## References

- Behl, R., Kanwar, V. and Sharma, K.K. (2013). Optimal equiscaled families of Jarratt's method. *Int. J. Comput. Math.*, **90**: 408–422.
- Behl R. and Motsa, S.S. (2012). Geometric Construction of Eighth-Order Optimal Families of Ostrowskis Method. *Sci. W. J.*, Article ID 614612.
- Geum, Y.H., Kim, Y.I. and Neta, B. (2015). On developing a higher-order family of double-Newton methods with a bivariate weighting function. *Appl. Math. Comput.*, **254**: 277–290.
- Kanwar, V., Behl, R. and Sharma, K.K. (2011). Simply constructed family of a Ostrowski's method with optimal order of convergence. *Comput. Math. Appl.*: **62**, 4021–4027.
- Neta, B. (1979). A sixth-order family of methods for nonlinear equations. *Int. J. Comput. Math.*, **7**: 157–161.
- Neta, B., Scott, M. and Chun, C. (2012). Basins of attraction for several methods to find simple roots of nonlinear equations. *Appl. Math. Comput.*, **218**: 10548–10556.
- Petković, M.S., Neta, B., Petković, L. D. and Džunić J. (2012] *Multipoint methods for solving nonlinear equations* (Academic Press).

- Sharma, J.R. and Guha, R.K. (2011). A family of modified Ostrowski methods with accelerated sixth-order convergence. *Appl. Math. Comput.*, **190**: 111–115.
- Scott, M., Neta, B. and Chun, C. (2011). Basins attractors for various methods. *Appl. Math. Comput.*, **218**: 2584–2599.
- Traub, J.F. (1964). Iterative methods for the solution of equations. (Prentice-Hall, Englewood Cliffs).
- Wang, X. and Liu, L. (2009) Two new families of sixth-order methods for solving non-linear equations. *Appl. Math. Comput.*, **213**: 73–78.

# Numerical Study of Interaction between Waves and Floating Body

## by MPS Method

Y.L. Zhang<sup>1,2</sup>, Z.Y. Tang<sup>1,2</sup>, D.C. Wan<sup>1,2\*</sup>

1. State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University,  
2. Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

\*Corresponding author: dcwan@sjtu.edu.cn

### Abstract

In the present study, interaction between regular waves and free roll motion of a 2D floating body is investigated by our in-house particle solver MLParticle-SJTU based on modified Moving Particle Semi-Implicit (MPS) method. Numerical wave tank is developed to calculate the interaction between waves and floating body, including wave-maker module and free roll motion module. The comparison between the numerical wave elevation and analytical solution shows that the MLParticle-SJTU can provide acceptable accuracy of wave making. Roll motion and force acting on the floating body in waves are in good agreement with experimental results. Profiles of the wave surface surrounding floating body are presented.

**Keywords:** Particle method; MPS (Moving Particle Semi-Implicit); Wave Floating body Interaction; Wave making; Roll motion

### Introduction

Recent years, a variety of floating structures, such as ships, offshore platforms, floating-breakwater, fish-farms, floating-airports, play a crucial role in coastal and ocean engineering. It's common for floating structures to suffer from loadings under waves, and responses of these structures mounted in ocean or coastal environments have significant relation to the wave impacts. The interaction between free-surface waves and floating body is one of the key aspects in ship design or offshore structure design to increase performance and efficiency.

In the past decades, both theory and experimental analyses methods have been used by many researchers to investigate the interaction problem. The early established theoretical methods are mainly based on potential flow theories and limited to solve the motion of floating body with simple shape. Chahine, et al. (1999) developed a free surface hydrodynamics code based on three-dimensional Boundary Element Method and then they modeled the nonlinear evolution of waves as they progress along a shallow sloping bottom in the presence of a floating body that is free to rotate and translate. Bai and Eatock Taylor (2006) studied the radiation and diffraction problem of vertical circular cylinders in a fully nonlinear numerical wave tank based on the boundary element method (BEM). You and Faltinsen (2012) developed a 3D fully nonlinear time-domain Rankine source code. A numerical wave tank with a piston wave maker and a numerical damping zone is applied to simulate the interaction between moored floating bodies and waves. Jung et al. (2004a) experimentally studied waves impacting on a fixed rectangular structure. PIV technique is used to obtain the mean velocity and turbulence properties of water around structure. The generation and evolution of vortexes of a barge in beam sea condition is simulated. Subsequently, Jung et al. (2004b, 2005) investigated the two-dimensional flow characteristics of interactions between waves

and freely rolling rectangular structures. Results between the roll motion and the fixed condition were compared. Ren et al. (2015) studied the motions of a freely floating body under nonlinear waves.

Besides, a wide variety of nonlinear numerical models based on the NS equations in time domain have been developed to study the interaction problem. The finite difference method or the finite volume method (FVM) is typically used for spatial discretization. And various techniques are used for interface capturing, such as the Level Set method and the Volume of Fluid method. In Boo's work (2002), a numerical tank was constructed, the linear and nonlinear irregular wave diffraction forces acting on a submerged structure was predicted. In Li's work (2010), a 2-D numerical regular wave tank was built, which mainly based on the spatially averaged Navier- Stokes equations and the k- $\epsilon$  model was used to simulate the turbulence of flow. The fully nonlinear wave-body interactions between a surface piercing body in finite water depth and flat/slop bottom topography were also investigated. Ye et al. (2012) constructed a three-dimensional numerical wave tank with a newly developed solver naoe-FOAM-SJTU based on the open source code library OpenFOAM, and a S-175 container ship sailing in regular heading waves was simulated. Zha, et al. (2013) studied the motion responses of heave and pitch of a ship in different wave conditions. Numerically simulation of the motion response of a moored floating pier in regular waves was described in Liu and Wan (2013).

The above researches are based on Eulerian methods and grids are necessary for spatial discretization. It's difficult and inaccurate to obtain free surface with large deformation. Recently, Lagrangian particle methods draw much attention of researchers and are seen as promising numerical approaches for free surface flows. For example, Moving Particle Semi-implicit (MPS), originally proposed by Koshizuka and Oka (1996) for incompressible flow. In the present study, a particle solver, MLParticle-SJTU based on modified Moving Particle Semi-Implicit (MPS) method, is used for all simulation works. Some improved schemes are used in this solver to suppress numerical unphysical pressure oscillation usually observed in traditional MPS method. These improvements include: (1) momentum conservative pressure gradient model; (2) modified kernel function [Zhang et al., 2011b]; (3) mixed sourced term method for Poisson equation of pressure [Tanaka et al., 2010]; (4) surface detection method based on asymmetry of neighbor particles [Zhang et al., 2011a]. The MLParticle-SJTU was applied in many large free-surface deformation problems, such as dam breaking flow [Zhang, et al., (2011c, 2014)], liquid sloshing in LNG tank [Zhang, et al., 2012; YANG, et al.,2014], Floating Body Interacting with Solitary Wave [Zhang, et al., 2011b].

This paper is organized as follow: Firstly, the MPS method for incompressible fluid is described. Numerical approach to solve the motion of floating body is introduced. Then, numerical wave tank is developed to calculate the interaction between waves and floating body, including wave-maker module and free roll motion module. Time history of wave propagation is measured and compared with the analytical solution to validate the accuracy of wave making. At last, roll motion and force acting on the floating body in waves is calculated and compared with experimental results. Profiles of the wave surface surrounding floating body are also presented.

## **Numerical Scheme**

### *Governing Equations*



Governing equations are the continuum equation and the momentum equation. These equations for incompressible viscous fluid are represented as:

$$\nabla \cdot \mathbf{V} = 0 \quad (1)$$

$$\frac{D\mathbf{V}}{Dt} = -\frac{1}{\rho} \nabla P + \nu \nabla^2 \mathbf{V} + \mathbf{g} \quad (2)$$

where  $\mathbf{V}$  is the velocity vector,  $t$  is the time,  $\rho$  is the density,  $P$  is the pressure,  $\nu$  is the kinematic viscosity,  $\mathbf{g}$  is the gravity acceleration.

### Particle Interaction Models

#### Kernel Function

In particle method, governing equations are transformed to the equations of particle interactions. The particle interactions are based on the kernel function. In traditional MPS method, the kernel function is expressed as follow (Koshizuka, 1996):

$$W(r) = \begin{cases} \frac{r_e}{r} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (3)$$

A drawback of the above kernel function is that it becomes singular at  $r=0$ . This may cause unreal pressure between two neighboring particles with a small distance, and affect the computational stability. To overcome this, an improved kernel function is used in this paper (Zhang and Wan, 2011b):

$$W(r) = \begin{cases} \frac{r_e}{0.85r + 0.15r_e} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (4)$$

The above kernel function has a similar form with the original kernel function Eq. (3), but without singularity.

#### Gradient Model

Gradient operator is modeled as a local weighted average of the gradient vectors between particle  $i$  and its neighboring particles  $j$ :

$$\langle \nabla \phi \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{\phi_j + \phi_i}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (5)$$

where  $\phi$  is an arbitrary scalar function,  $D$  is the number of space dimensions,  $n^0$  is the initial particle number density for incompressible flow. The particle number density in MPS method is defined as:

$$\langle n \rangle_i = \sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (6)$$

#### Laplacian Model

Laplacian operator is derived by Koshizuka et al. (1998) from the physical concept of diffusion as:

$$\langle \nabla^2 \phi \rangle_i = \frac{2D}{n^0 \lambda} \sum_{j \neq i} (\phi_j - \phi_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (7)$$

$$\lambda = \frac{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|) \cdot |\mathbf{r}_j - \mathbf{r}_i|^2}{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|)} \quad (8)$$

where  $\lambda$  is a parameter, introduced to keep the variance increase equal to that of the analytical solution. Both viscous force  $\nabla^2 V$  in Eq. (2) and  $\nabla^2 P$  in the right hand side of the PPE (Eq. 9 and Eq. 10) are discretized by Eq. (7).

#### *Model of incompressibility*

The incompressible condition in traditional 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 in traditional MPS method is defined as (Koshizuka et al., 1998):

$$\langle \nabla^2 P^{k+1} \rangle_i = -\frac{\rho}{\Delta t^2} \frac{\langle n^* \rangle_i - n^0}{n^0} \quad (9)$$

where  $n^*$  is the particle number density in temporal field.

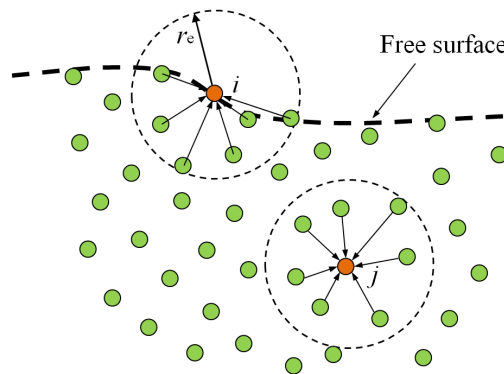
The source term of the Poisson equation in Eq. (9) is solely based on the deviation of the temporal particle number density from the initial value. As the particle number density field is not smooth, the pressure obtained from Eq. (9) is prone to oscillate in spatial and temporal domain. To suppress such unphysical oscillation of pressure, Tanaka and Masunaga (2010) proposed a mixed source term for PPE, which combines the velocity divergence and the particle number density. The main part of the mixed source term is the velocity divergence, while the particle number density is used to keep the fluid volume constant. This improved PPE is rewritten by Lee et al. (2011) as:

$$\langle \nabla^2 P^{k+1} \rangle_i = (1-\gamma) \frac{\rho}{\Delta t} \nabla \cdot V_i^* - \gamma \frac{\rho}{\Delta t^2} \frac{\langle n^k \rangle_i - n^0}{n^0} \quad (10)$$

where  $\gamma$  is a blending parameter with a value between 0 and 1. The value of  $\gamma$  has large effect on the pressure field. In particular, the larger  $\gamma$  produces smoother pressure field. However, the volume of fluid cannot be constant while  $\gamma=0$ . The effects of  $\gamma$  have been investigated by Tanaka, et al. (2010) and Lee, et al. (2011).  $\gamma=0.01$  is used in this paper.

#### *Free Surface boundary condition*

On the surface particles, the free surface boundary conditions, including kinematic and dynamic boundary condition, are imposed. The kinematic condition is directly satisfied in Lagrangian particle method, while the dynamic condition is implemented by setting zero pressure on the free surface particles. So the accuracy of surface particle detection has significant effect on pressure field.



**Figure 1. Description of particle interaction domain**

The interaction domain is truncated in the free surface (Fig. 1), so the particle number density near the free surface is lower than that in the inner field. In traditional MPS method, particle satisfying (Koshizuka et al., 1998):

$$\langle n \rangle_i^* < \beta \cdot n^0 \quad (11)$$

is considered as free surface particle, where  $\beta$  is a parameter, can be chosen between 0.80 and 0.99.

The traditional detection method (Eq. 11) is based on the particle number density. However, inner particles with small particle number density may be misjudged as free surface particles, thus unreal pressure around the misjudged particles occur. This usually causes nonphysical pressure oscillation. To improve the accuracy of surface particle detection, we employ a new detection method in which a vector function is defined as follow (Zhang and Wan, 2011c):

$$\langle \mathbf{F} \rangle_i = \frac{D}{n^0} \sum_{j \neq i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} (\mathbf{r}_i - \mathbf{r}_j) W(r_{ij}) \quad (12)$$

The vector function  $\mathbf{F}$  represents the asymmetry of arrangements of neighbor particles.

Particle satisfying:

$$\langle |\mathbf{F}| \rangle_i > \alpha \quad (13)$$

is considered as free surface particle, where  $\alpha$  is a parameter, and has a value of 0.9  $|\mathbf{F}|^0$  in this paper,  $|\mathbf{F}|^0$  is the initial value of  $|\mathbf{F}|$  for surface particle.

It should be specially noted that the Eq. (13) is only valid for particles with number density between  $0.8n^0$  and  $0.97n^0$  since particles with number density lower than  $0.8n^0$  is definitely surface particles, while those with number density higher than  $0.97n^0$  should get pressure through Poisson equation.

#### *Motion of floating body*

The motion of the floating body is governed by the equations of rigid body dynamics, following the Newton's law of motion. The translation motion of the center of gravity and the rotation of the rigid body are given in a simple 2-D framework by

$$\begin{cases} M \frac{d\mathbf{V}_G}{dt} = M\mathbf{g} + \mathbf{F}_{fluid-solid} \\ I_G \frac{d\boldsymbol{\Omega}_G}{dt} = \mathbf{T}_{fluid-solid} \end{cases} \quad (14)$$

where  $M$  and  $I_G$  are the mass and the moment of inertia of the floating body around the center of gravity, respectively.  $\mathbf{V}_G$  and  $\boldsymbol{\Omega}_G$  are the linear velocity of the center of gravity and the angular velocity of the body, respectively.  $\mathbf{F}_{fluid-solid}$  is the hydrodynamic force acting on the body,  $\mathbf{T}_{fluid-solid}$  is the hydrodynamic torque with the direction normal to the plane.

## **Numerical Simulations**

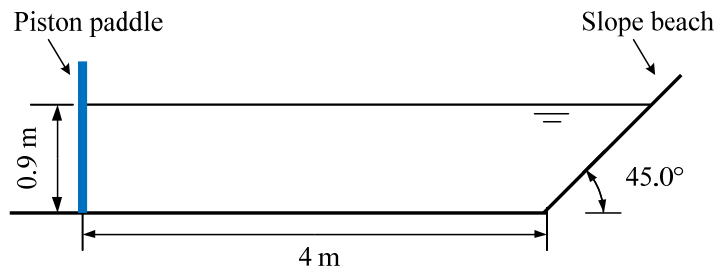
### *Test of wave making*

In present work, a piston-type wave generator was incorporated in the left side of 2D numerical wave tank. A slop beach was installed at the end of the wave tank to absorb waves and avoid reflection. Sketch of the numerical setup is shown in Fig. 2. The numerical wave tank is 5.5 m

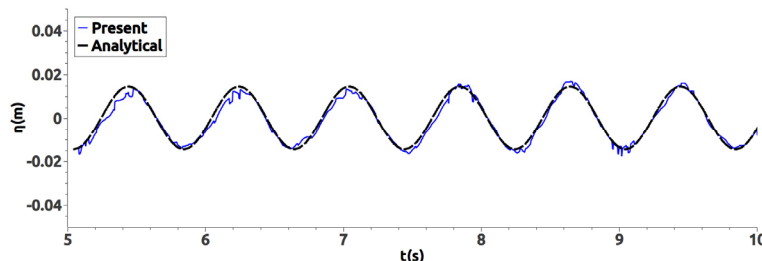
width and 1.5 m height with initial water depth 0.9 m. Wave conditions used in present numerical test is shown in Table 1, and travelling waves were generated based on linear wave theory.

**Table1. Parameters of wave making**

Parameters	Values
Water density	1000(kg/m <sup>3</sup> )
Water height	0.9(m)
Wave length	1(m)
Wave height	0.029
Wave period	0.8(s)
Fluid spacing	0.004 (m)
No. of fluid particles	132750
No. of total particles	138130



**Figure 2. Sketch of the 2-D wave tank**

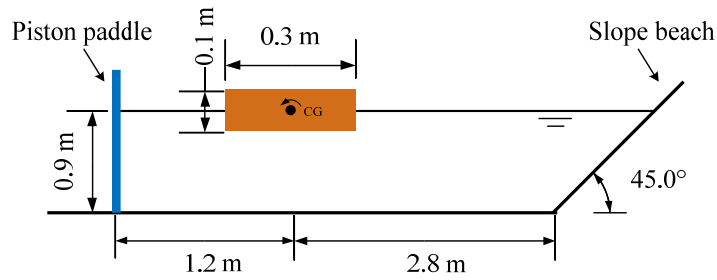


**Figure 3. Comparison between numerical wave elevation and analytical solution at location 1m from the piston paddle**

Fig. 3 shows a comparison between numerical wave elevation and analytical solution at the location 1 m from the initial position of piston paddle. The trend of numerical free surface height is in agreement with analytical solution except that the former is less smoother than the later. The difference can be improved by reducing the particle space.

#### *Simulation of freely rolling body*

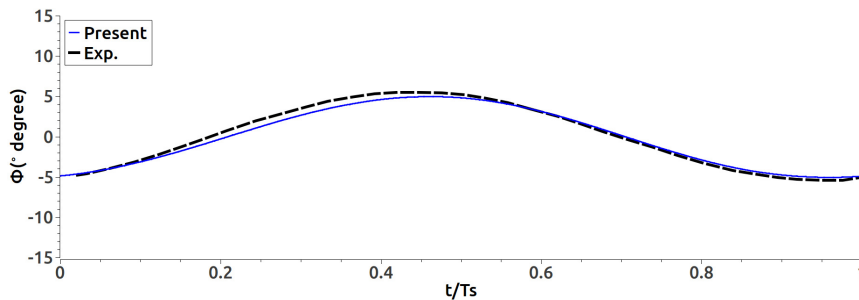
In this section, the roll motion of a 2D floating rectangular structure in a numerical wave tank was investigated in time domain. The wave generator and wave absorbing manner here are same as that in previous section. The width and height of the rectangular floating body are 0.3 m and 0.1 m, respectively. The structure was installed at the point 1.2 m from the wave maker and 0.9 m above the bottom of tank, fixed at the center of its gravity but free in the degree of roll. The initial geometry and set-up are shown in Fig. 4.



**Figure 4. Sketch of the freely rolling body**

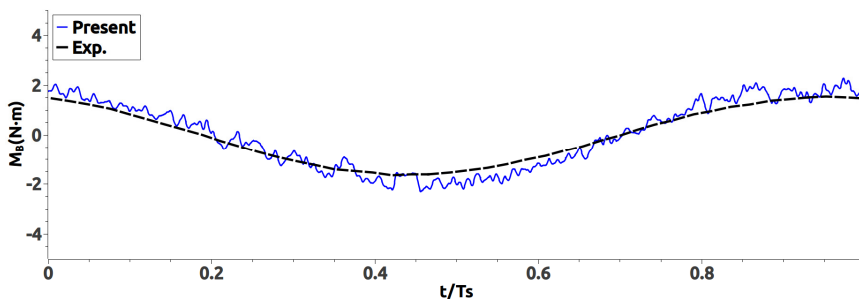
In present simulation, the distance between particles is 0.004 m, the total number of particles is 137718 while the number of fluid particles is 131762. The gravitational acceleration and water density are  $9.8\text{m/s}^2$  and  $1000\text{kg/m}^3$ . The kinematic viscosity of water is given by  $1.01 \times 10^{-6} \text{ m}^2/\text{s}$ . The time step size is 0.0004s and the total computational time is 10s. Waves with period of 0.8 s, was generated in the present study. The wave conditions used in present computation are same as shown in Table 1.

In the free rolling test of rectangular structure, angles of roll motion about the center of gravity were measured when the regular roll motion of rectangular body can be obtained.



**Figure 5. Rolling motion of the rectangular body within wave period (solid line: result of simulation; dashed line: result of JUNG)**

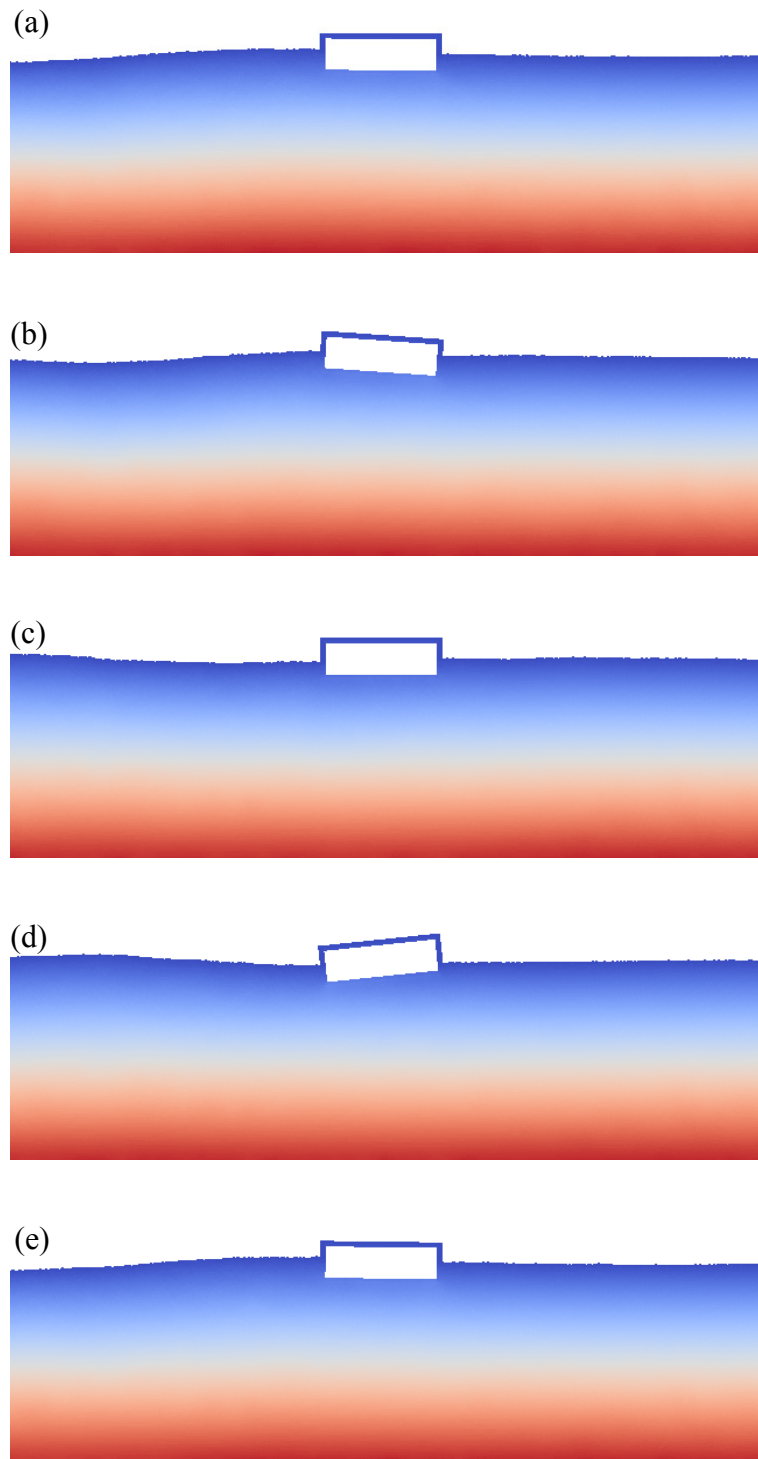
Fig. 5 shows the inclined angle of the rectangular structure over one period of the regular wave. Results about the roll motion of floating body is compared between present simulation and experiment by Jung(2004). It can be found that both the pattern of curves and amplitude of roll angles are in good agreement.



**Figure 6. Time history of buoyancy restoring moment of the rectangular body within wave period (solid line: result of simulation; dashed line: result of JUNG)**

Details about buoyancy restoring moment ( $M_B$ ) of the freely rolling body should be noteworthy, because the roll motion is closely related with the change of  $M_B$  in time domain. Fig. 6 shows the time history of buoyancy restoring moment in one wave period. It can be seen that the calculated

results of  $M_B$  agree fairly well with experimental results, though the calculated curve about moment has a little nonphysical fluctuations.



**Figure 7. Rotary positions and wave surfaces around floating box, (a)  $t=t_0$ , (b)  $t=t_0+T/4$ , (c)  $t=t_0+T/2$ , (d)  $t=t_0+3T/4$ , (e)  $t=t_0+T$**

Fig. 7 shows the evolution process of rotation of the floating body. It can be seen that positions of the floating box is strong influenced by the propagation of incident wave through five snapshots of representative-instants ( $t_0$ ,  $t_0+T/4$ ,  $t_0+T/2$ ,  $t_0+3T/4$  and  $t_0+T$ ) in a wave period. Firstly, the body rotates clockwise until the value of  $M_B$  climbs up to the maximum with the coming wave from left.

After that, the crest of the wave transfers from left to right of the floating box. At the same time, buoyancy restoring moment of the body decreases. As a result, box rotates anti-clockwise and returns to horizontality at the instant of  $t_0+T/2$ . With the propagating of the wave, water surface falling on the left and rising on the right from  $t_0+T/2$  to  $t_0+3T/4$ . Box keeps on rotating anti-clockwise until the value of  $M_B$  declines to the minimum. From the instant of  $t_0+3T/4$ , it begins to rotate clockwise again, and returns to nearly horizontality at  $t_0+T$  finally. The rotation of floating body will repeat with the next coming wave from left.

## Conclusions

In this paper, interaction between regular waves and free roll motion of a 2D floating body is investigated by our in-house particle solver MParticle-SJTU based on modified Moving Particle Semi-Implicit (MPS) method. Four improvements, including nonsingular kernel function, momentum conservative pressure gradient model, mixed source term for PPE and an accurate surface detection method, are employed in this solver. Numerical wave tank is developed to calculate the interaction between waves and floating body, including wave-maker module and free roll motion module. The comparison between the numerical wave elevation and analytical solution shows that the MParticle-SJTU can provide acceptable accuracy of wave making. Numerical roll motion and force acting on the floating body in waves are in good agreement with experimental results. At last, the evolution process of rotation of the floating body was shown through five snapshots of representative-instants ( $t_0$ ,  $t_0+T/4$ ,  $t_0+T/2$ ,  $t_0+3T/4$  and  $t_0+T$ ) in a wave period. It can be seen that positions of the floating box are strong influenced by the propagation of incident wave. According to the results present in previous sections, the solver can be used to deal with waves floating body interaction problems.

## Acknowledgement

This work is supported by National Natural Science Foundation of China (Grant Nos. 51379125, 51490675, 11432009, 51411130131), The National Key Basic Research Development Plan (973 Plan) Project of China (Grant No. 2013CB036103), High Technology of Marine Research Project of The Ministry of Industry and Information Technology of China, Chang Jiang Scholars Program (Grant No. T2014099) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (Grant No. 2013022), to which the authors are most grateful.

## References

- Bai, W., and Taylor, R. E. (2008) Fully nonlinear simulation of wave interaction with fixed and floating flared structures, *Ocean Engineering*, 36(3), 223-236.
- Boo, S.Y., (2002) Linear and nonlinear irregular waves and forces in a numerical wave tank, *Ocean Engineering*, 29, 475-493.
- Jung, K.H., Chang, K.A. and Huang, E.T. (2004a) Two-dimensional flow characteristics of wave interactions with a fixed rectangular structure. *Ocean Engineering*, 31, 975-998.
- Jung, K. H. (2004b) Experimental study on rectangular barge in beam sea, Ph.D. Texas A&M University.
- Jung, K.H., Chang, K.A. and Huang, E.T. (2005) Two-dimensional flow characteristics of wave interactions with a free-rolling rectangular structure. *Ocean Engineering*, 32(1), 1-20.
- Kalumuck, K. M., Chahine, G. L., and Goumilevski, A. G. (1999) *BEM modeling of the interaction between breaking waves and a floating body in the surfzone*, 13th ASCE Engineering Mechanics, Baltimore, Maryland.
- Koshizuka, S., and Oka, Y. (1996) Moving-particle Semi-implicit Method for Fragmentation of Incompressible Fluid, *Nuclear Science and Engineering*, 123, 421-434.
- Koshizuka, S., Obe, 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, 751-769.
- Li, Y., and Lin, M. (2010) Wave-body interactions for a surface-piercing body in water of finite depth, *Journal of hydrodynamics, Ser. B*, 22 (6), 745-752.

- Liu, Y. C., and Wan, D. C. (2013) *Calculation of motion response for moored floating pier in waves*, Proc 25rd Chinese National Conf Hydrodynamics, ZhouShan, China.
- Ren, B., He, M., Dong, P., and Wen, H. j. (2015) Nonlinear simulations of wave-induced motions of a freely floating body using WCSPH method, *Applied Ocean Research*, 50 (2015) 1-12.
- Tanaka, M., and Masunaga, T. (2010) Stabilization and smoothing of pressure in MPS method by Quasi-Compressibility, *J Comp Phys*, 229, 4279-4290.
- YANG Y. Q., ZHANG Y. X., Tang Z. Y., and WAN D. C. (2014) *Numerical study on liquid sloshing in horizontal baffled tank by MPS method*, Proc 26rd Chinese National Conf Hydrodynamics, QingDao, China.
- You, J., Faltinsen, and O. M. (2012) *A 3D Fully Nonlinear Numerical Wave Tank with a Moored Floating Body in Shallow Water*, Proceedings of the Twenty-second International Offshore and Polar Engineering Conference Rhodes, Greece.
- Ye, H. X., Shen, Z. R., and Wan, D. C. (2012) *Numerical analysis of large amplitude motion responses for a container ship in waves*, Proc 24rd Chinese National Conf Hydrodynamics, Wuxi, China.
- Zha, R, S, Ye, H. X., and Wan, D. C. (2013) *Numerical investigation of ship sea-keeping performances in head regular waves*, Proc 25rd Chinese National Conf Hydrodynamics, ZhouShan, China.
- Zhang, Y. X., and Wan, D. C. (2011a) *Application of Improved MPS Method in Sloshing Problem*, Proc 23rd Chinese National Conf Hydrodynamics, Xi'an, China.
- Zhang, Y. X., and Wan, D. C. (2011b) *Apply MPS Method to Simulate Motion of Floating Body Interacting with Solitary Wave*, Proc 7th Int Workshop Ship Hydrodynamics, IWSH, Shanghai, China, 275-279.
- Zhang, Y. X., and Wan, D. C. (2011c) Application of MPS in 3D Dam Breaking Flows, *Sci Sin Phys Mech Astron*, 41, 140-154.
- Zhang, Y. X., and Wan, D. C. (2012) *Apply MPS Method to Simulate Liquid Sloshing in LNG Tank*, Proc 22nd Int Offshore and Polar Eng Conf, Rhodes, Greece, 381-391.
- Zhang, Y. X., Wan, D. C., Hino, and Takanori. (2014) Comparative study of MPS method and level-set method for sloshing flows, *Journal of hydrodynamics*, 26(4), 577-585.
- Zhang, Y. X., TANG, Z. Y., and Wan, D. C. (2013) *Simulation of 3D dam break flow by parallel MPS method*, Proc 25rd Chinese National Conf Hydrodynamics, ZhouShan, China.



## Development of Microsystems Analysis ( $\mu$ sys) Software Using Hybrid Finite Elements and Direct Solution of Coupled Equations

†\*A. Roychowdhury<sup>1,2</sup>, Kunal D. Patil<sup>1,2</sup>, Arup Nandy<sup>1</sup>, C.S. Jog<sup>1,2</sup>, Rudra Pratap<sup>1,2</sup> and G. K. Ananthasuresh<sup>1,2</sup>

<sup>1</sup>Department of Mechanical Engineering, Indian Institute of Science (IISc), Bangalore

<sup>2</sup>Computational Nano engineering (CoNe) group, Center for Nano Science and Engineering, Indian Institute of Science (IISc), Bangalore

\*Presenting author: anishroy@mecheng.iisc.ernet.in

†Corresponding author: anishroy@mecheng.iisc.ernet.in

### Abstract

In this work, we describe the development and implementation of a hybrid finite element based software “ $\mu$ sys” for the analysis of coupled multi-physical phenomena encountered in Micro-Electro-Mechanical Systems (MEMS) or microsystems. The developed computational strategy entails the use of hybrid finite elements for modeling structures, which overcomes some of the inherent shortcomings associated with the interpolation incompatibility of the finite elements. As a result, a coarse mesh with a single type of high aspect ratio element can be used for computationally efficient solution of many kinds of coupled partial differential equations. The same element is used to interpolate multiple state variables including displacement, temperature, fluid pressure and electric potential etc. Additionally, we use direct, rather than staggered, approach to solve coupled PDEs. The  $\mu$ sys software has modular architecture to seamlessly interface the analysis codes with various packages having pre and post-processing capabilities. This helps the user to work with familiar and accessible pre- and post-processing environments while using  $\mu$ sys as the versatile finite element analysis engine. We illustrate the architecture and working of  $\mu$ sys with two coupled analysis modules, viz., coupled structure-electrostatics and coupled structure-squeeze film. Pre- and post-processing capabilities of a commercial finite element software are used to demonstrate working with  $\mu$ sys as an example.

**Keywords:** Coupled PDEs, MEMS, Multi-physics, Squeeze-film, Pull-in

### Introduction.

Micro-Electro-Mechanical Systems (MEMS) or microsystems are ubiquitous in many consumer products and industrial applications today. Simulation and design of microsystem components and devices is indispensable to achieve improved performance and cost-effective manufacturing. Various computational platforms are available to perform numerical modeling and finite element analysis of multi-physical phenomena associated with the microsystems [1-4]. Solving coupled partial differential equations in continuous domain and ordinary differential equations in the reduced order models is unique to microsystems.

The earliest published work [5] advocated the strategy of staggered coupling of PDE solvers and focused on 'wrappers' that exchanged data between different modules during simulation. For example, in MEMCAD (which is now known as Coventorware), commercial finite element

software ABAQUS was used for elastic analysis and in-house developed FASTCAP was used for electrostatic analysis. MEMCAD's wrapping code called ABAQUS and FASTCAP alternately to find self-consistent solution of PDEs of elastic and electrostatic domains. The main focus of MEMCAD then was on emulating microfabrication process and building CAD models as well as post-processing. A comprehensive review of developments related to process modeling and visualization tools for MEMS can be found in [6]. The staggered approach of alternating between solvers of different physical domains is prevalent even now.

In  $\mu$ sys we take a different approach. We have a single integrated analysis engine that can be interfaced with any pre- and post-processing modules of other software. This is motivated by two commonly encountered problems in existing microsystems simulator software. First, the user should be familiar with types of finite elements to be used for meshing the CAD model. It is very common to have many types of elements to be used for meshing the CAD model also it is common to have different element types, one for each type of coupled simulation, in the existing softwares. The users should also be aware of the dangers of coarse meshing, distortion in the elements in the mesh and aspect ratio of the elements (which should ideally be close to unity) to avoid converged but inaccurate results. Interdisciplinary users of microsystem cannot be expected to have intimate knowledge of finite elements and PDEs of many physical domains. Second, the staggered approach is computationally inefficient and leads to prolonged simulation time.

The two shortcomings of existing microsystems software are overcome in  $\mu$ sys by using hybrid finite elements. In hybrid finite elements, displacement and stress (i.e; the state variable and its derivative) are independently interpolated to avoid the incompatibility in the interpolation functions and ensuing problem of "locking" phenomena [7,8]. In a hybrid finite element, aspect ratio need not be close to unity and the element need not be very small to give accurate results. A single 3D element type can cater to solve multiple PDEs. Only one element across the thickness of a MEMS component is enough and this element can be very large in the in-plane dimensions. This makes it possible to get accurate results with coarse meshes [9-11]. The single 3D hybrid finite element type can be endowed with multiple state variables, namely, displacement, pressure, temperature, electric potential etc. This enables simultaneous solution of multiple coupled PDEs directly [12-15] rather than in the staggered manner of most existing software today. Based on the foregoing, the focus in  $\mu$ sys is on a single analysis engine that is developed in-house. This analysis engine is easily linked with any pre- and post-processors.

In the next section, we discuss overall architecture of  $\mu$ sys. Analysis module of  $\mu$ sys is discussed later by illustrating briefly implementation and working of two analysis sub-modules, coupled structure-electrostatics and coupled squeeze film, respectively. The use of pre- and post-processing capabilities of finite element software NISA [16], are demonstrated by use of appropriate data parsing.

### **Architecture of $\mu$ sys**

The overall system architecture of  $\mu$ sys is shown in Figure 1. As is common with simulation softwares,  $\mu$ sys is divided into three modules namely, pre-processing, analysis and post-processing. Our focus is only in the development of  $\mu$ sys on finite element analysis module by, relying on commercially available platforms for pre and post-processing through appropriate data parsing. Pre-

processors are used to create CAD models, finite element meshing, specification of material data, loading data, boundary and initial conditions, etc. Pre-processing can be done using popular commercially available platforms such as NISA, ABAQUS, ANSYS etc. The task of developing an independent pre-processor is very involved, and also, as the users are generally accustomed to using at least one of the available pre-processing platforms. The input data to the analysis module of  $\mu$ sys can be generated by parsing appropriate information from the input file generated using pre-processor. Thus, the user models using accessible pre-processor, and parsers take care of input data format required for  $\mu$ sys. Any additional analysis dependent data specification can be done using data input section of  $\mu$ sys.

After the analysis, the output file generated by  $\mu$ sys, which consists of field variables such as stress, potential, displacement, etc., sought at different nodes, is parsed into appropriate post-processing format. Again, the commercially available post-processing tools can be used for visualization and graph plotting. The visualization of output using post-processing section of  $\mu$ sys is also another alternative. This development is partially completed.

As mentioned earlier, the primary focus of  $\mu$ sys development is on the finite element analysis section. Currently, structural, electrostatics and coupled analysis modules (see, Figure 1) are part of the  $\mu$ sys. The characteristic features and working with the two coupled modules, is the subject of discussion of the next section. In doing so, we have selected NISA [16] as a pre and post-processing platform of  $\mu$ sys as an example.

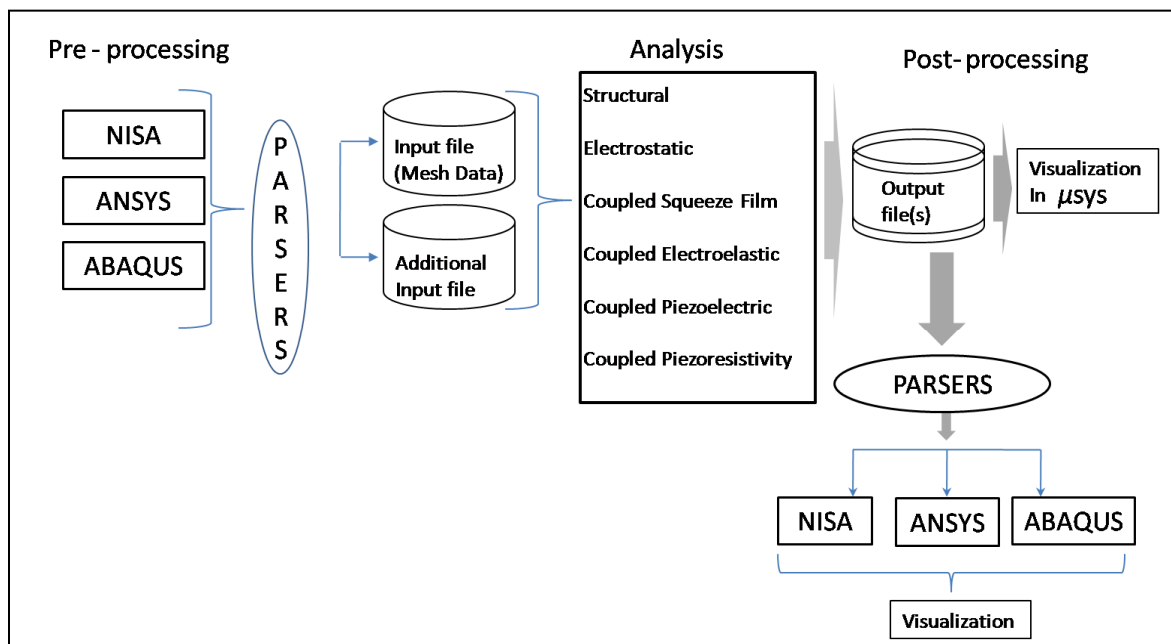


Figure 1: Architecture of  $\mu$ sys

Currently, six different analysis modules are implemented in  $\mu$ sys. It includes:

1. Structural
2. Electrostatics
3. Coupled structure-electrostatics or electroelastics

4. Coupled squeeze film
5. Coupled piezoelectric
6. Coupled piezoresistivity

In this section, we discuss coupled structure-electrostatics and coupled squeeze film modules as illustrative examples to demonstrate solver capabilities of  $\mu\text{sys}$ . Rest of the coupled modules are extension of the coupled structure-electrostatics with addition of analysis dependent data specification [12, 13].

### ***Coupled structure-electrostatics module***

Coupled structure-electrostatics is a very important module in the analysis of microsystems structures. The knowledge of static and dynamic pull-in voltages, capacitance, etc. is important for the design of electrostatic actuators and sensors.

In this section we briefly describe the finite element formulation of the elastic-electrostatic coupled problem. Detailed implementation of nonlinear structural hybrid finite element development can be found in [7] and [8]. For coupling with electrostatics [9, 10] and [13] may be consulted.

Strong form of equations for the structural problem in the reference configuration  $\Omega$  (total domain) is given as

$$\nabla \cdot (FS) + \rho_0 b^0 = 0, \quad \text{on } \Omega \quad (1)$$

$$E = \bar{E}(u) := \frac{1}{2} [(\nabla u) + (\nabla u)^T], \quad \text{on } \Omega \quad (2)$$

$$S_{\text{mech}} = \tilde{S}(C), \quad \text{on } \Omega \quad (3)$$

$$t^0 = FS n^0, \quad \text{on } \Omega \quad (4)$$

$$t^0 = \bar{t}^0, \quad \text{on } \Gamma_t \quad (5)$$

$$u = 0. \quad \text{on } \Gamma_u \quad (6)$$

where  $F(X) := \nabla \chi = I + \nabla u$ , is the deformation gradient defined with respect to material coordinates  $X$ ,  $u$  the displacement field,  $E$  the Lagrangian measure of strain,  $C = F^T F$  the right Cauchy-Green strain tensor,  $S$  the second Piola-Kirchhoff stress,  $t^0 = \|JF^{-T} n^0\| t$  the tractions defined on the reference configuration in terms of tractions on the deformed configuration,  $\rho_0 = J\rho$  the density in the reference configuration in terms of the density in the deformed configuration,  $J = \det(F)$ ,  $n^0$  the outward normal to  $\Gamma$ ,  $S_{\text{mech}}$  is the constitutive equation of solid. The surface  $\Gamma$ , is the boundary between solid and surrounding consisting of open disjoint union of regions,  $\Gamma = \overline{\Gamma_u} \cup \Gamma_t$  for structural problem.

For electrostatics, the governing equations are given by

$$\nabla \cdot d = 0, \quad \text{on } \Omega \quad (7)$$

$$d = \sigma J C^{-1} e, \quad \text{on } \Omega \quad (8)$$

$$e = -\nabla_X \phi, \quad \text{on } \Omega \quad (9)$$

$$\phi = \bar{\phi}, \quad \text{on } \Gamma_\phi \quad (10)$$

$$d_n^0 = \bar{d}_n^0, \quad \text{on } \Gamma_d \quad (11)$$

$$S_M = \varepsilon J \left[ C^{-1} (e \otimes e) C^{-1} - \frac{(e \cdot C^{-1} e)}{2} C^{-1} \right]. \quad (12)$$

where  $S_M$  is the Maxwell stress,  $e$  and  $d$  the Lagrangian electric field and electric displacement vectors, respectively,  $\phi$  the electric potential or voltage,  $d_n^0$  the normal electric displacement vector defined on the reference configuration,  $\sigma$  and  $\varepsilon$  are the conductivity and permittivity, respectively. The surface  $\Gamma$ , is the boundary between solid and surrounding consisting of open disjoint union of regions,  $\Gamma = \overline{\Gamma_d \cup \Gamma_\phi}$  for electrostatic problem.

By enforcing Eqns. (1) and (7) in a weak sense, we get

$$\int_{\Omega} S : \bar{E}_\delta \, d\Omega = \int_{\Omega} \rho_0 u_\delta \cdot b^0 \, d\Omega + \int_{\Gamma_t} u_\delta \cdot \bar{t}^0 \, d\Gamma, \quad (13)$$

$$\int_{\Omega} \nabla \phi_\delta \cdot (\sigma J C^{-1} \nabla \phi) \, d\Omega = - \int_{\Gamma_d} \phi_\delta \bar{d}_n \, d\Gamma \quad \forall \phi_\delta, \quad (14)$$

where  $u_\delta$ ,  $E_\delta$  and  $\phi_\delta$  denotes variation of displacement, strain and potential fields, respectively.

In a hybrid formulation, strain-displacement relation Eq. (2), is also enforced in a weak sense. Thus, we get

$$\int_{\Omega} (S)_\delta : \left[ \bar{E}(u) - \hat{E}(S) \right] \, d\Omega = 0, \quad (15)$$

where  $S_\delta$  denotes the variation of stress field.

After linearization of Eqns. (13), (14) and (15) at the reference configuration, we get incremental formulations [9]. Introducing following interpolation

$$u = N \hat{u}, \quad \phi = N \hat{\phi}, \quad S_{\text{mech}} = P \beta \quad (16)$$

and using the same shape functions  $N$  and  $P$  for increments, we get following finite element equations,

$$Q \hat{u}_\Delta + K_{u\phi} \hat{\phi}_\Delta + (G_0 + G^T) \beta_\Delta = (f_u)_\Delta, \quad (17)$$

$$K_{\phi S} \beta_\Delta + K_{\phi\phi} \hat{\phi}_\Delta = (f_\phi)_\Delta, \quad (18)$$

$$G \hat{u}_\Delta - H \beta_\Delta = (f_S)_\Delta. \quad (19)$$

Expressions for the matrices  $Q$ ,  $G$ ,  $H$ ,  $G_0$ ,  $K_{u\phi}$ ,  $K_{\phi\phi}$ ,  $K_{\phi S}$  and the force vectors  $f_u$ ,  $f_\phi$ ,  $f_S$  can be found in [9].

**Example**

The working of coupled elastic-electrostatics module is demonstrated using an example of static pull-in of a cantilever beam. Pre- and post-processing is done using the commercial FEM package NISA. Figure 2 shows the block-diagram of working with “NISA-  $\mu$ sys” system for coupled elastic-electrostatics problems.

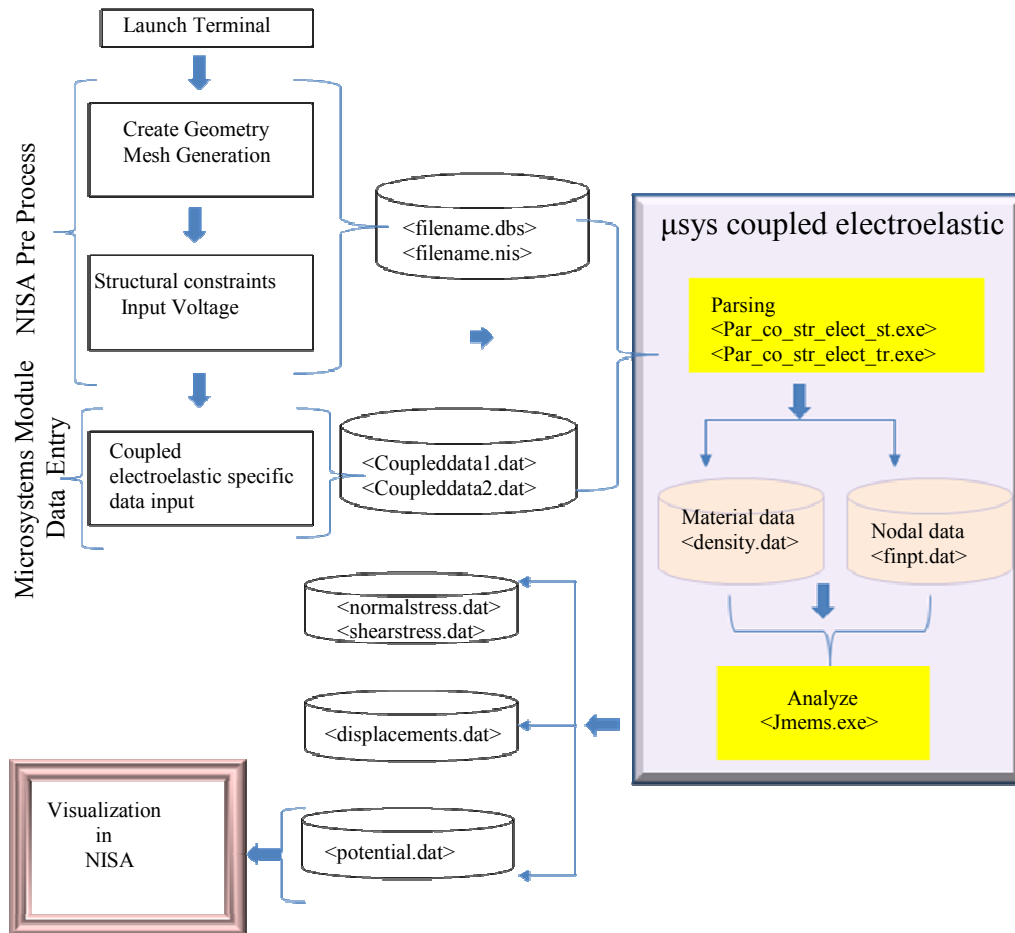


Figure 2: Block-diagram of NISA-  $\mu$ sys coupled elastic-electrostatics module

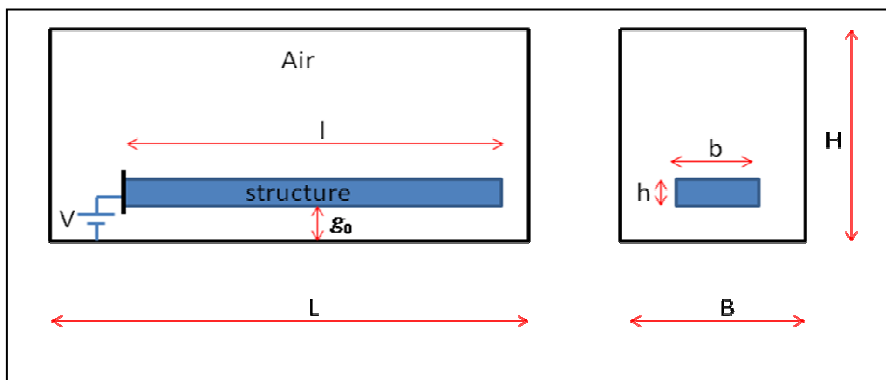


Figure 3: A Cantilever beam, ground, and the surrounding air domain

Figure 3 shows the schematic for a cantilever beam subjected to gradual increase in the voltage from 0 to static pull-in value. In the direct solution strategy, the air domain is modeled as a weak elastic dielectric material. The dimensions and properties for modeling elastic and air domains are given in Table 1.

<b>Data</b>	<b>Values</b>
Beam length (l)	100 $\mu$ m
Beam width (b)	10 $\mu$ m
Beam height (h)	0.5 $\mu$ m
Surrounding domain box length (L)	140 $\mu$ m
Surrounding domain box width (B)	60 $\mu$ m
Surrounding domain box height (H)	52 $\mu$ m
Initial gap (g)	1 $\mu$ m
Young's modulus (beam)	169 GPa
Poisson's ratio (beam)	0.3
Density (beam)	2231 Kg/m <sup>3</sup>
Relative permittivity (beam)	11.7
Electrical conductivity (beam)	1.56e-03 S/m
Young's modulus (air)	1.0e07 Pa
Poisson's ratio (air)	0.0
Density (air)	0.0 Kg/m <sup>3</sup>
Relative permittivity (air)	1
Electrical conductivity (air)	5.5e-15 S/m

Table 1: Dimensions and material data for cantilever pull-in problem

Figure 4 shows the tip-deflection of cantilever beam as the voltage is increased from 0 till pull-in. Accurate calculation of deflection significantly affects determination of the pull-in voltage. Close to pull-in phenomena, the error in the deflection as computed by  $\mu$ sys and COMSOL is close to 16%. Table 2 gives comparison of the values for static pull-in obtained from  $\mu$ sys and COMSOL with analytical calculations. Figures 5 and 6 show the displacement and potential fields visualization in NISA. More comparative studies can be found in [13].

<b>Method</b>	<b>Static pull-in voltage (V)</b>
$\mu$ sys	2.55
COMSOL	2.6
Analytical [19]	2.53
Analytical [20]	2.55

Table 2: Static pull-in calculation for a cantilever beam

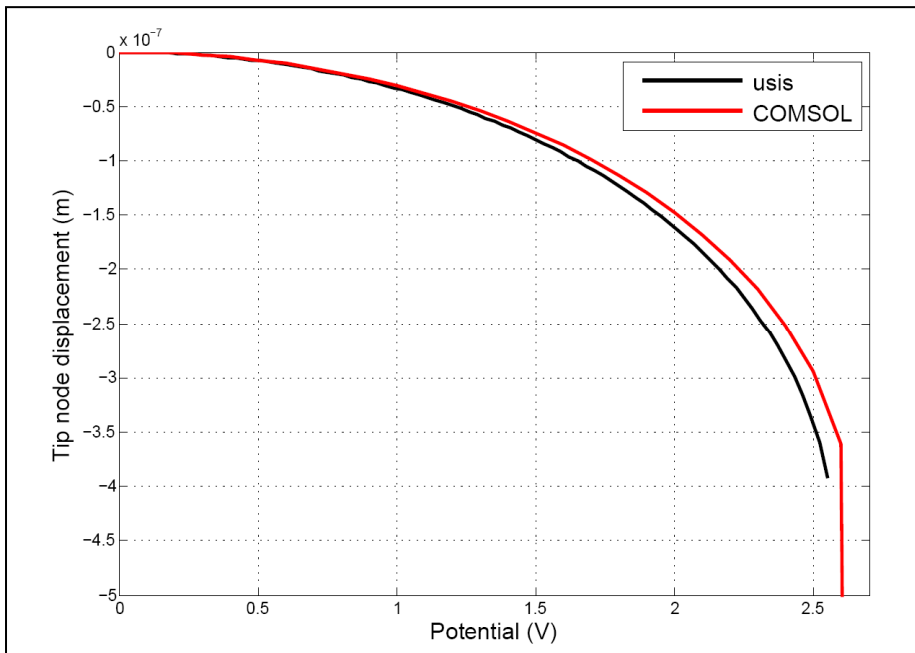


Figure 4: Tip deflection of a cantilever beam with increase in voltage

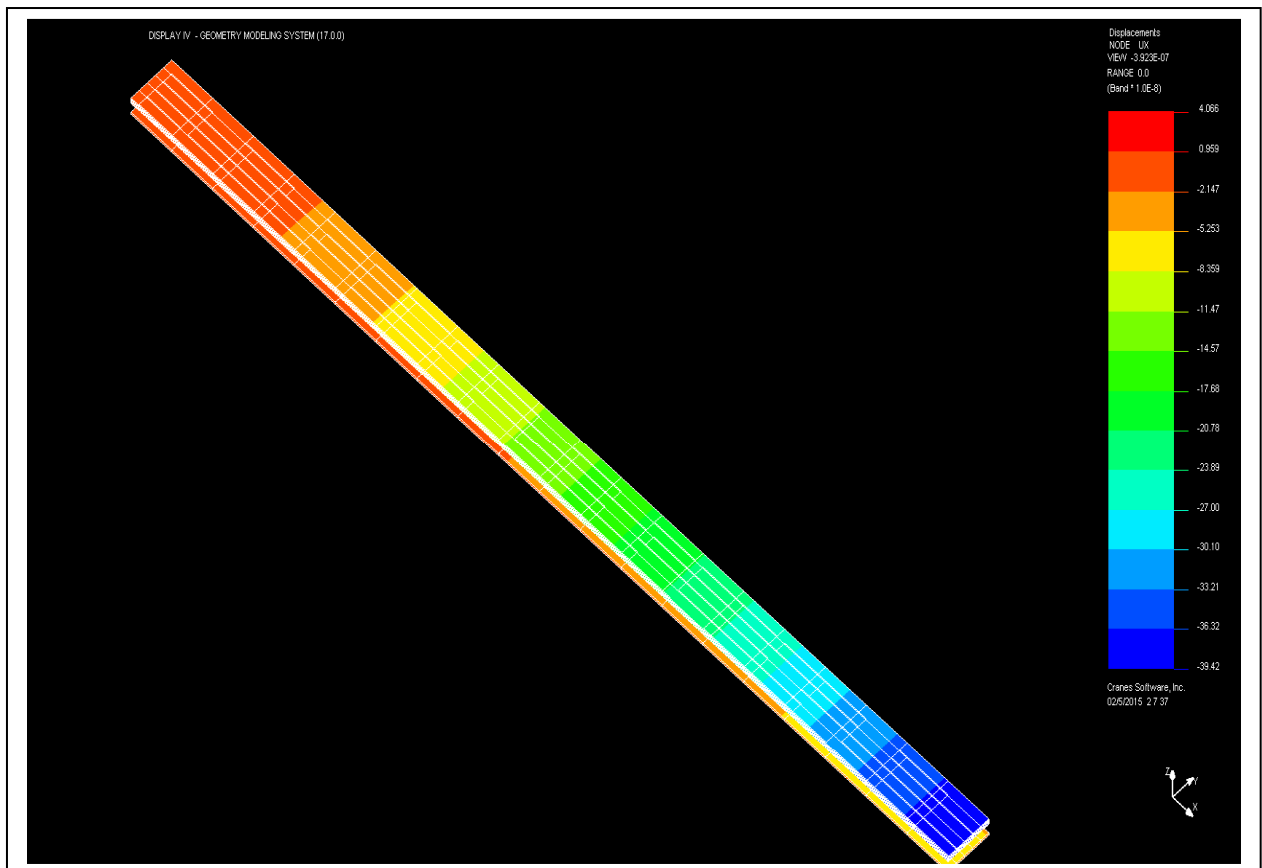


Figure 5: Displacement field visualization in NISA



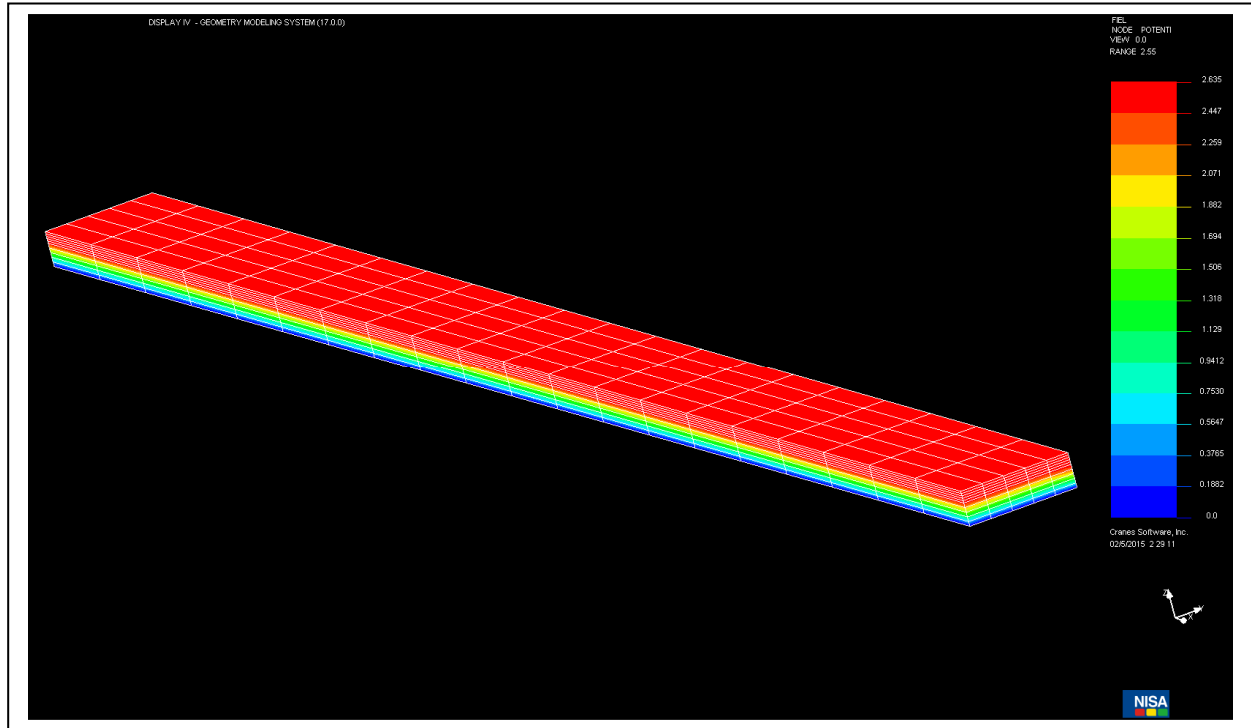


Figure 6: Potential field visualization in NISA

### *Coupled Squeeze-Film Analysis*

The squeeze film effect is prevalent in vibratory MEMS devices where a thin film of air trapped between a fixed substrate and a plate vibrating normally to the fixed plate behaves like both a viscous damper and as an air spring [18]. Various attempts have been made to model the coupled fluid structure problem of squeeze film. A thorough review of the prior works can be found in [14] where we discussed a coupled finite element based methodology to solve the fluid-structure squeeze film problem. In the present software developed, we have implemented a coupled hybrid monolithic formulation. Hybrid elements are known to overcome locking issues faced with displacement based formulations in modeling high aspect ratio structures (as present in MEMS devices with squeeze film). Thus, our present formulation is able to show good results even for those mesh generators which are limited to lower order meshes and with far less number of elements compared to displacement based formulations.

In implementing the analysis module for squeeze film, we solve the 3D elasticity equation and the 2D Reynolds equation for squeeze film flow in a coupled method following the procedure outlined in [14]. For the dynamic structural problem without any body force, we have the following governing equation in a weighed integral sense for hybrid elements [15].

$$\int_{\Omega} u_{\delta} \cdot \left( \nabla \cdot \tau - \rho \frac{\partial^2 u}{\partial t^2} \right) d\Omega + \int_{\Gamma} u_{\delta} \left( \bar{t} - t \right) d\Gamma + \int_{\Gamma} \tau_{\delta} : \left[ \bar{\epsilon}(u) - C^{-1} \tau \right] d\Omega = 0 \quad (20)$$

where  $u$  is the displacement,  $\tau$  the stress,  $t$  the traction,  $u_\delta$  and  $\tau_\delta$  the variations in displacement and stress field respectively,  $C$  the constitutive matrix,  $\bar{t}$  prescribed traction and  $\bar{\epsilon}(u) = (\nabla u) + (\nabla u)^T$ . For coupled squeeze film problem with structural interaction, the wet surface (the surface in contact with air) is subject to the fluid pressure through traction  $\bar{t} = -\tilde{p}\hat{n}$ . We discretize Eqn. (20) using finite element interpolations for displacements and stress fields and their variations as,  $u = N_u \hat{u}$ ,  $u_\delta = N_u \hat{u}_\delta$ ,  $\bar{\epsilon}(u) = B_u \hat{u}$ ,  $\tau = P \hat{\beta}$ ,  $\tau_\delta = P \hat{\gamma}$  and  $\tilde{p} = N_p \hat{p}$  where  $P$  is the stress interpolation function, the choice of which is described in [8], and  $N_u$ ,  $N_p$  and  $B_u$  are as described in [14]. After substituting the interpolations and following the procedure outlined in [17] and considering a harmonic solution, we arrive at the following form of the equation,

$$[K_{uu}] \hat{u} + [K_{up}] \hat{p} = [f_u] \quad (21)$$

where

$$[K_{uu}] = -\omega^2 \int_{\Omega} \rho N_u^T N_u d\Omega + [G]^T [H]^{-1} [G], \quad (22)$$

$$[K_{up}] = \int_{\Gamma_{wet}} N_u^T \hat{n} N_p d\Omega, \quad (23)$$

$$[f_u] = \int_{\Gamma} N_u^T \bar{t} d\Gamma. \quad (24)$$

where  $H$  and  $G$  are as described in [15], and  $\omega$  is the frequency of harmonic vibration. The fluid domain is modeled using the linearized Reynolds equation given by

$$\frac{h_0^3}{12\mu_{eff}} \left( \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} \right) = \frac{h_0}{P_a} \frac{\partial P}{\partial t} + \frac{\partial H}{\partial t} \quad (25)$$

Using the process outlined in [14] we arrive at the following discretized form of Eqn. (25)

$$[K_{pp}] \hat{p} + [K_{pu}] \hat{u} = 0 \quad (26)$$

where

$$[K_{pp}] = \frac{h_0}{12\mu_{eff}} \int_{\Gamma_{wet}} B_p^T B_p d\Gamma + \frac{j\omega h_0}{P_a} \int_{\Gamma_{wet}} N_p^T N_p d\Gamma, \quad (27)$$

$$[K_{pu}] = j\omega \int_{\Gamma_{wet}} N_p^T N_{u_z} d\Gamma. \quad (28)$$

and  $N_p$ ,  $N_{u_z}$  and  $B_p$  are as described in [14]. Now, combining Eqns. (21) and (26) we get the coupled form:

$$\begin{bmatrix} K_{uu} & K_{up} \\ K_{pu} & K_{pp} \end{bmatrix} + \begin{bmatrix} \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} f_u \\ 0 \end{bmatrix} \quad (29)$$

### ***Implementation of Squeeze film module***

Following the implementation architecture discussed in this paper, we discuss, as a representative example, the implementation of our squeeze film analysis module into an existing commercial FEA suite, i.e. NISA [16]. The implementation architecture is shown in Figure 7.

As a first step we did a 'gap' analysis where we identified the additional information needed for squeeze film analysis that is not already a part of NISA input. Thus, an additional input GUI screen with squeeze film parameters, air gap, air viscosity, ambient pressure, frequency range of operation was designed. The NISA suite generates a raw data file in a particular format based on input from user. With our special requirements, two data files in NISA specific format gets created. A parser program written in 'C' language was developed to convert the NISA raw data files into formats compatible with our FORTRAN analysis module. As the current NISA version does not allow the choice of "face id" selection by the user for denoting the face subject to squeeze film pressure, the logic was incorporated in the parser. The nodal connectivity from NISA had to be converted to the connectivity logic used in our analysis module, the conversion logic was also built into the parser.

We have incorporated a special feature "node of interest", wherein the NISA GUI was modified to allow the use to select a particular node of choice for which nodal displacement data will be generated. The output from our module consists of nodal solutions for pressure and displacement which may be visualized using NISA post processing capabilities. A special feature "frequency sweep" which is incorporated in our squeeze film module allows for a tabular output of nodal displacement, squeeze film damping and stiffness forces and the corresponding frequency used in the computation. The data is output in a tabulated format so as to allow the use to further plot frequency response graphs or force vs frequency plots as maybe desired. The user may also compute stiffness and damping coefficients from the displacement and the force values computed at each frequency as follows:  $(K = \frac{F_s}{u_z}, C = \frac{F_d}{u_z})$  where ' $F_s$ ' is the spring force, ' $F_d$ ' is the damping force and ' $u_z$ ' is the 'z' direction displacement. Figure 7. shows the process schematic for implementation of the squeeze film module into NISA.

### ***Example***

We present a test case to validate of our squeeze film solver implemented in NISA. We model a cantilever resonator and compare the first three Q factors with that from published literature. We choose to model a Si, cantilever beam as described in [21]. The simulation parameters are shown in Table 3. The beam is meshed with a converged mesh of ( $N_x = 40, N_y = 6, N_z = 4$ ) elements. The simulations are run for the frequencies between 1e4 Hz to 1e5 Hz in small incremental steps of 1e3 Hz. The vertical displacement ' $u_z$ ' for the tip node of the beam is noted for the range of frequencies. The corresponding value of velocity is obtained as  $V_{tip} = frequency * u_z$ . From the plot of normalized velocity (with input voltage) Vs frequency the Q factor is obtained using 3dB method. The computed Q factor is compared with reported values from experiments and from ANSYS [21] (See Table 4). Thus we see our squeeze film module computed Q factor compares well with experimental and numerical data from published literature. Figure 8, shows the post-processing capabilities of NISA for visualization of pressure field distribution.

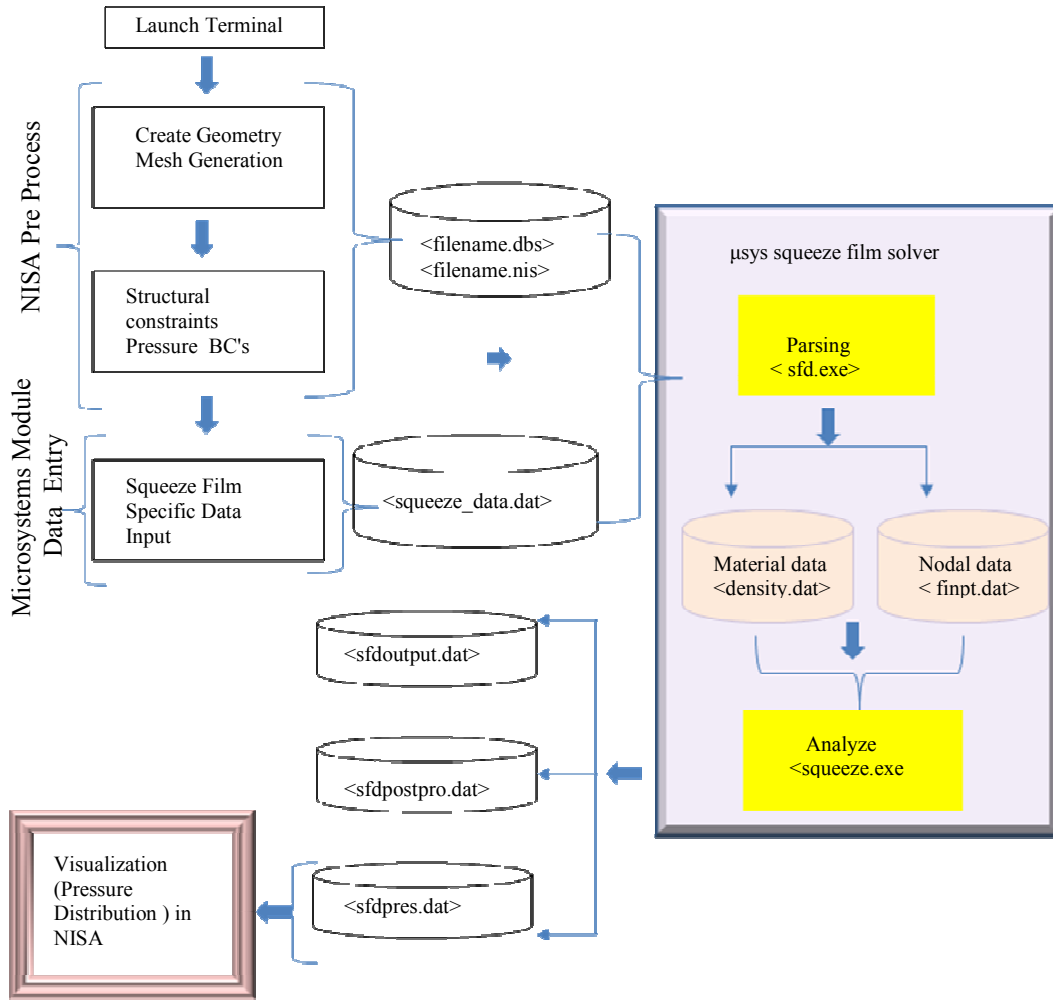


Figure 7. Squeeze film analysis module in NISA

Simulation Parameters	Values
Young's Modulus (Si)	160GPa
Density	2230 Kg/m <sup>3</sup>
Poisson's Ratio	0.22
Density (air)	1.2Kg/m <sup>3</sup>
Viscosity	1.8e-05 Ns/m <sup>2</sup>
Air Gap	1.4 μm
Length	350 μm
Breadth	22 μm
Thickness	4 μm
Actuation Voltage	1.5 Volts

Table 3: Simulation parameters for squeeze film test case

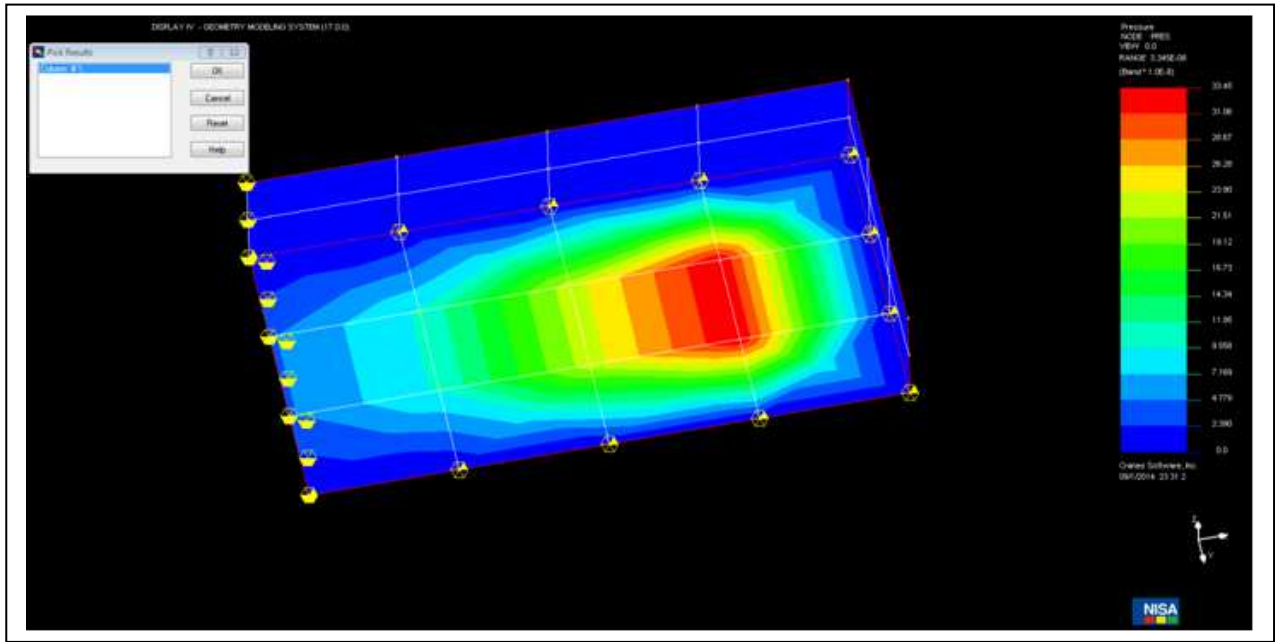


Figure 8. Squeeze film pressure distribution (on the lower surface of a cantilever) visualization in NISA

Q-NISA	Q-Exp	Q-ANSYS
1.13	1.20	1.11

Table 4: Q factor from 1st mode of a cantilever due to squeeze film

### Conclusions

We described the development of hybrid finite element-based direct solution strategy in software  $\mu$ sys, which is capable of doing multi-physical computational analysis of microsystems. The novel features of  $\mu$ sys are the use of hybrid finite elements which alleviates certain short comings of conventional finite elements, and the monolithic implementation of governing equations. Implementation and working of two coupled modules, coupled structure-electrostatics and coupled squeeze film are discussed using pre- and post-processing capabilities of NISA. Currently, the major focus is on the development of analysis module of  $\mu$ sys and using other commercial pre- and post-processing capabilities through appropriate data parsing. Development of post-processing tool of  $\mu$ sys is also underway.

### Acknowledgements

The authors would like to thank the National Program for Micro and Smart Systems (NPMASS) for providing generous support for activities related to the development of microsystems analysis module at CoNe (Computational Nano engineering) Lab at IISc Bengaluru. The authors would also like to acknowledge Vasanthi Chayapathi of Cranes Software for her help with GUI customizations.

## References

- [1] CoventorWare software, [www.coventor.com/mems-solutions/products/coventorware](http://www.coventor.com/mems-solutions/products/coventorware)
- [2] IntelliSuite software, [www.intellisense.com](http://www.intellisense.com)
- [3] COMSOL Multiphysics, MEMS module
- [4] ANSYS Multiphysics, [www.ansys.com](http://www.ansys.com)
- [5] Senturia S.D., Harris R.M., Johnson B.P., Kim S., Nabors K. and White J.K. (1992) A Computer- Aided Design System for MicroelectroMechanical Systems (MEMCAD), *Journal of Microelectromechanical Systems*, 1(1), 3-13.
- [6] Senturia (1998) CAD Challenges for Microsensors, Microactuators, and Microsystems, *Proceedings of the IEEE*, 86(8), 1611-1626.
- [7] C. S. Jog and P. P. Kelkar (2006), Nonlinear analysis of structures using high performance hybrid elements, *International Journal for Numerical Methods in Engineering*, 68, 473-501.
- [8] C.S. Jog (2010) , Improved hybrid elements for structural analysis, *Journal of Mechanics of Materials and Structures*, 5, 507-528
- [9] K. D. Patil, C.S. Jog and G. K. Ananthasuresh (2014), Monolithic hybrid finite element strategy for coupled structure-electrostatic analysis of micromechanical structures, *ISSS International Conference on Smart Materials, Structures and Systems*, Bangalore.
- [10] K. D. Patil, S. Balakrishnan, C.S. Jog and G.K. Ananthasuresh (2014), A simulation module for microsystems using hybrid finite elements: an overview, *Micro and Smart Devices and Systems*, Springer, 355-373.
- [11] M. Sundaram, R. Ganesh, K. Pavan, B. Varun, C. S. Jog, and G. K. Ananthasuresh (2012), Static elastic simulation of micromechanical structures using hybrid finite elements, *ISSS International Conference on Smart Materials, Structures and Systems*, Bangalore.
- [12] S. Balakrishnan, K. Deshpande and G. K. Ananthasuresh (2012), *National Conference of Smart Materials Structures and Systems*, Coimbatore.
- [13] C. S. Jog and K. D. Patil , Monolithic hybrid fem strategy for coupled electro-mechanics of microsystems, *in preparation*.
- [14] Roychowdhury A., Nandy A., Jog C.S. and Pratap R. (2013) A monolithic FEM based approach for the coupled squeeze film problem of an oscillating elastic microplate using 3D 27-node elements, *Journal of Applied Mathematics and Physics*, 1, 20-25.
- [15] Roychowdhury A., Nandy A., Jog C.S. and Pratap R. (2015) A monolithic FEM based approach for the coupled squeeze film problem of an oscillating elastic microplate using 3D 27-node elements, *CMES: Computer Modeling in Engineering and Sciences*, Accepted for publishing.
- [16] NISA software, [www.nisasoftware.com](http://www.nisasoftware.com)
- [17] Keating, D. J., & Ho, L. (2001) Effects of squeezed film damping on dynamic finite element analyses of MEMS. *In Design, Test, Integration, and Packaging of MEMS/MOEMS 2001* 226-236. International Society for Optics and Photonics.
- [18] Blech J.J. (1983) On Isothermal Squeeze Films, *Journal of Lubrication Technology*, 105(4), 615-620.
- [19] P. M. Osterberg and S. D. Senturia (1997), M-test: A test chip for MEMS material property measurement using electrostatically actuated test structures, *Journal of Microelectromechanical systems*, 6(2).
- [20] R. K. Gupta (1997), Electrostatic pull-in structures design for in-situ mechanical property measurements of microelectromechanical systems (MEMS), Ph.D. thesis, MIT.
- [21] A. K. Pandey, and R. Pratap (2007), Squeeze Film Damping in Perforated MEMS Torsion Mirror, *International conference on Emerging Mechanical Technology MACRO TO NANO (EMTM2N-2007)*, BITS Pilani

## Osteoporotic bone fracture healing under the locking compression plate system

†L. Zhang<sup>1</sup>, S. Miramini<sup>1</sup>, P. Mendis<sup>1</sup>, M. Richardson<sup>2</sup>

<sup>1</sup> Department of Infrastructure Engineering, The University of Melbourne, Australia.

<sup>2</sup> The Epworth Hospital, Richmond, Australia.

†Corresponding author: [lih Zhang@unimelb.edu.au](mailto:lih Zhang@unimelb.edu.au)

### Abstract

Osteoporosis is highly prevalent and a costly disease predicted to affect 1,555 million worldwide by 2050, and the total cost of osteoporotic fractures worldwide could reach US\$131 billion by 2025. These statistics clearly affirm the significant economic burden of osteoporotic fractures to the community, and the need for the development of improved fracture treatments. Studies over the last decade reveal that, even though osteoporosis may not necessarily lead to non-union, it is associated with delayed fracture healing due to impaired mechano-regulation and angiogenesis in osteoporotic condition. Despite the advances in locking compression plate (LCP) technology, the operative treatment in osteoporotic fractures remains a challenge for an orthopaedic surgeon, often with unpredictable outcomes. Therefore, it becomes necessary to bridge the ‘information gap’ between osteoporosis and its effect on fracture healing, and so enables healing progression prediction under different fracture geometries and fixation configurations. By using a computational model of fracture healing, this paper demonstrates that fracture healing can be significantly delayed due to impaired mechano-regulation as a result of osteoporosis, and the impact of osteoporosis on fracture healing can be mitigated by adjusting the configuration of the LCP system to allow a certain degree of interfragmentary movement (IFM) without compromising overall fixation stability.

**Keywords:** Osteoporosis; Fracture healing; Interfragmentary movements (IFM); Locking compression plate (LCP).

### Introduction

Osteoporotic fractures lead to chronic pain, disability, loss of independence and even premature death. The total cost of fracture management in community health programs, aids and appliances, and indirect costs, such as lost earnings, is projected to rise to US\$131 billion worldwide by 2050 (Harvey et al., 2010; Watts et al., 2013). Osteoporosis is associated with delayed fracture healing (Augat et al., 2005; Namkung-Matthai et al., 2001), even though it is not a risk factor for non-union (Wunnik et al., 2011). Although it still remains an open question whether fracture repair is impaired in osteoporosis, the mechanical and biological factors involved in the healing process are certainly affected (Augat et al., 2005; Nikolaou et al., 2009). Firstly the reduction in number of mesenchymal stem cells and their impaired response to mechanical stimuli in osteoporotic condition may lead to a delayed fracture healing. Secondly osteoporosis related the impaired growth factor expression and abnormalities in endothelial cells could result in impair angiogenesis progression. Current therapies for osteoporotic fractures focus on prevention, however, little emphasis has been given to the study of the fracture healing process itself in osteoporotic bone.

The initial phase of healing is especially sensitive to mechanical conditions and influences the course of healing (Klein et al., 2003). It is widely believed that callus cell differentiation and proliferation are closely regulated by the magnitude of the so-called ‘stimulus index ( $S$ )’, which is determined by the octahedral shear strain and the interstitial flow velocity within a fracture callus. Mesenchymal stem cells differentiate into chondrocytes, osteoblasts and fibroblasts depending on their biomechanical microenvironment. However, fracture healing in osteoporotic bone may display

an impaired response to these mechanical stimuli (Sterck et al., 1998) due to the presence of fewer mesenchymal stem cells in osteoporotic bone, and a relatively lower proliferative response (Bergman et al., 1996; Giannoudis and Schneider, 2006). For example, an osteoporotic rat model showed a 40% reduction in the cross-sectional area of callus and a 23% reduction in bone mineral density in healing rate femurs (Namkung-Matthai et al., 2001).

The mechanical microenvironment in a fracture callus is greatly influenced by the IFM which is largely dependent on the mechanical stiffness of a LCP. The stiffness of an osteoporotic fracture fixation construct, such as the locking compression plate (LCP) system, has normally to be increased as osteoporosis decreases the mechanical properties of bone. However, an overly stiff fixation construct may lead to impaired IFM at the fracture site which inhibits callus formation, and may lead to potential delayed healing or non-union (Gardner et al., 2010). Clinically, the flexibility of LCP could be enhanced by adjusting the working length (WL), bone-plate (BPD) and number of screws (Claes, 2011). However, the effect of the flexibility of LCP on osteoporotic bone healing has not been fully investigated. The problem becomes further complicated as fixation failure occurs as a consequence of reduced bone density (Barrios et al., 1993).

Thus, the key questions are:

1. How is the mechao-regulation altered in osteoporosis?
2. How can the fixation design be strategically modified to achieve improved fracture healing outcomes in osteoporosis?

To address these questions, the development to computational models for osteoporotic fracture healing becomes necessary.

**Methods**

**Error! Reference source not found.**Figure 1 shows the impact of both osteoporosis and fixation configuration on healing outcomes. Fracture healing is a time-dependent process which is closely regulated by the changes in mechanical microenvironment of fracture site as a result of a change of interfragmentary motion at fracture site as healing progresses. We have recently brought together our previously developed poroelastic large deformation model of biological soft tissues and tissue/cell-level mechano-regulation model of fracture callus to gain a new insight into the early stage bone healing under different configurations of LCP (Miramini et al., 2014; Miramini et al., 2015; Zhang, 2015; Zhang et al., 2013; Zhang et al., 2012). The mechanical behavior of fracture callus is described by using a consolidation approach (Zhang et al., 2007, 2008; Zhang et al., 2009; Zhang et al., 2010) which treats callus as a fluid-saturated porous medium comprising an intrinsically solid phase (i.e. collagen-proteoglycan matrix) and a fluid phase. Mass conservation equations are written for each phase, while empirical laws (e.g. Darcy’s law) are used to describe relative velocity of each phase. Finally, momentum conservation laws can capture the interaction of the mechanical

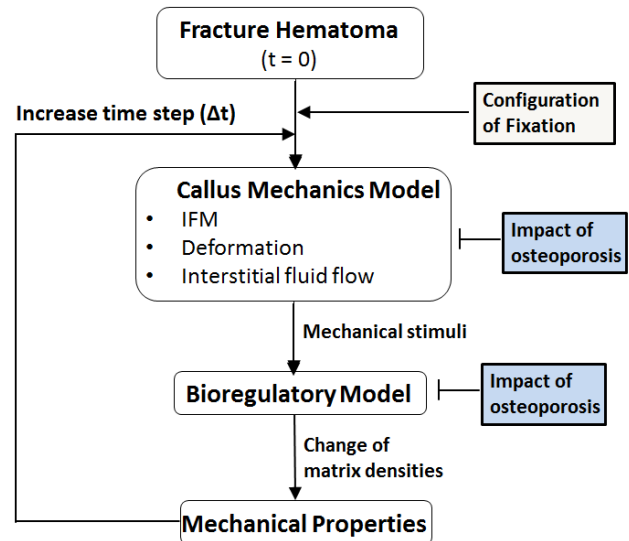


Figure 1 The schematic diagram shows the impact of osteoporosis and fixation configuration on healing outcomes.



quantities in each component. The model can be easily further extended by incorporating the impact of osteoporosis on fracture healing.

The framework of fracture healing model is shown in Figure 2. Our developed model allows the high resolution 3D modelling of a tibia containing a fracture from the 2D computed tomography (CT) image data (Miramini et al., 2015; Zhang, 2015; Zhang et al., 2013; Zhang et al., 2012). The 3D meshed part of the fractured tibia is then exported into finite element applications. The model enables the determination of the change of mechano-regulation distribution profiles (e.g. interstitial fluid flow and deformation) in fracture callus arising from external mechanical loading. Finally, the model predicts the tissue development regulated by callus cells that respond to their mechanical microenvironments. Furthermore, our developed model is capable of taking into consideration the fracture conditions (e.g. size, shape and position), fixation treatment selection, and loading regimes resulting from patient-specific physiological movement. In this paper, our developed model is implemented to investigate the effect of impaired mechanical stimuli due to osteoporosis on the healing outcomes at early stage of healing.

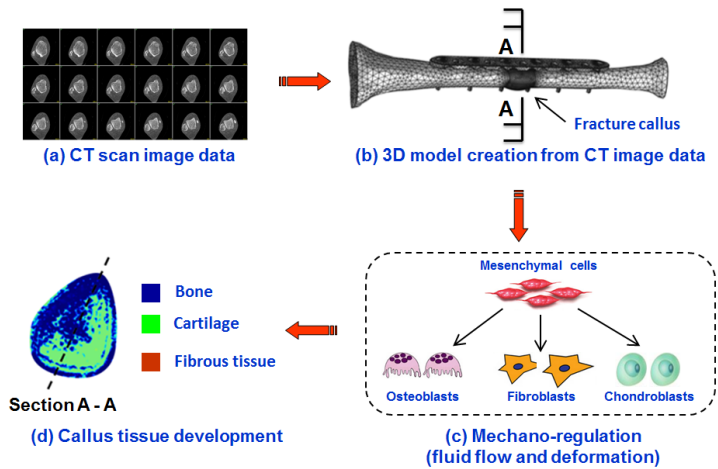


Figure 2 The framework of fracture healing model. (a) Computed tomography (CT) image data. (b) 3D model created from the CT images. (c) In vivo, cells in fracture callus regulate the healing processes by responding to changes in their mechanical microenvironments. (d) Mechanical stimuli mediated tissue differentiation.

### Results and Discussion

As shown in Figure 3, the effect of 25%, 50% and 75% impaired mechanical response of callus cells resulting from osteoporosis on tissue differentiation outcomes is investigated respectively. The results show that at the same time point after surgery, the impaired mechanical stimuli due to osteoporosis could potentially delay the healing process by inhibiting cartilage tissue development at the early stage of healing. However, by allowing certain degree of flexibility of LCP system without compromising the overall stability of the construct (i.e. increasing BPD from 0mm to 2mm), the impact of osteoporosis on fracture healing can be significantly mitigated.

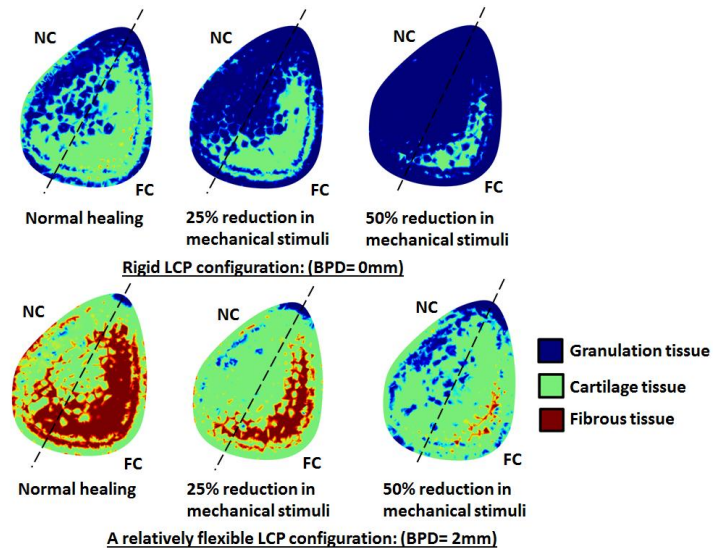


Figure 3 The investigation of the effects of impaired mechano-regulation as a result of osteoporosis and the flexibility of LCP configuration on tissue differentiation at early stage of fracture healing.

## Conclusions

This paper presents an osteoporotic fracture healing computational model, which could bridge the ‘information gap’ between osteoporosis and its effect on fracture healing, and so enable healing progression predictions under different fracture geometries and fixation configurations. The developed model will allow orthopaedic surgeons to design patient-specific surgical solutions by establishing a rigorous scientific relationship between the configuration of the fixation system and the biological processes of healing in patients with osteoporosis, and thereby achieve optimal fracture healing outcomes.

## Acknowledgments

The authors would like to thank Johnson & Johnson Medical, The Pierce Armstrong Foundation (2015-2016), AOTRAUMA Asia Pacific (AOTAP14-02), Victorian Orthopaedic Research Trust (2014-2015), Epworth HealthCare and the University of Melbourne for their support.

## References

- Augat, P., Simon, U., Liedert, A., Claes, L., 2005. Mechanics and mechano-biology of fracture healing in normal and osteoporotic bone. *Osteoporosis International*.
- Barrios, C., Broström, L.-Å., Stark, A., Walheim, G., 1993. Healing complications after internal fixation of trochanteric hip fractures: the prognostic value of osteoporosis. *Journal of Orthopaedic Trauma* 7, 438-442.
- Bergman, R.J., Gazit, D., Kahn, A.J., Gruber, H., McDougall, S., Hahn, T.J., 1996. Age-related changes in osteogenic stem cells in mice. *Journal of Bone and Mineral Research* 11, 568-577.
- Claes, L., 2011. Biomechanical principles and mechanobiologic aspects of flexible and locked plating. *Journal of orthopaedic trauma* 25, S4-S7.
- Gardner, M.J., Nork, S.E., Huber, P., Krieg, J.C., 2010. Less rigid stable fracture fixation in osteoporotic bone using locked plates with near cortical slots. *Injury* 41, 652-656.
- Giannoudis, P., Schneider, E., 2006. Principles of fixation of osteoporotic fractures. *Journal of Bone & Joint Surgery, British Volume* 88, 1272-1278.
- Harvey, N., Dennison, E., Cooper, C., 2010. Osteoporosis: impact on health and economics. *Nature Reviews Rheumatology* 6, 99-105.
- Klein, P., Schell, H., Streitparth, F., Heller, M., Kassi, J.-P., Kandziora, F., Bragulla, H., Haas, N.P., Duda, G.N., 2003. The initial phase of fracture healing is specifically sensitive to mechanical conditions. *Journal of Orthopaedic Research* 21, 662-669.
- Miramini, S., Zhang, L., Richardson, M., Mendis, P., 2014. Computational simulation of mechanical microenvironment of early stage of bone healing under locking compression plate with dynamic locking screws. *Applied Mechanics and Materials* 53, 281-286.
- Miramini, S., Zhang, L., Richardson, M., Pirpiris, M., Mendis, P., Oloyede, K., Edwards, G., 2015. Computational simulation of the early stage of bone healing under different Locking Compression Plate configurations. *Computer Methods in Biomechanics and Biomedical Engineering* 18, 900-913.
- Namkung-Matthai, H., Appleyard, R., Jansen, J., Lin, J.H., Maastricht, S., Swain, M., Mason, R.S., Murrell, G.A.C., Diwan, A.D., Diamond, T., 2001. Osteoporosis influences the early period of fracture healing in a rat osteoporotic model. *Bone* 28, 80-86.

- Nikolaou, V.S., Efstathopoulos, N., Kontakis, G., Kanakaris, N.K., Giannoudis, P.V., 2009. The influence of osteoporosis in femoral fracture healing time. *Injury* 40, 663-668.
- Sterck, J.G.H., Klein-Nulend, J., Lips, P., Burger, E.H., 1998. Response of normal and osteoporotic human bone cells to mechanical stress in vitro. *American Journal of Physiology - Endocrinology and Metabolism* 274, E1113-E1120.
- Watts, J.J., Abimanyi-Ochom, J., Sanders, K.M., 2013. Osteoporosis costing all Australians. *Osteoporosis Australia*.
- Wunnik, B.P.W.v., Weijers, P.H.E., Helden, S.H.v., Brink, P.R.G., Poeze, M., 2011. Osteoporosis is not a risk factor for the development of nonunion: A cohort nested case-control study. *Injury* 42, 1491-1494.
- Zhang, L., 2015. Computational Modeling of Bone Fracture Healing by Using the Theory of Porous Media, in: Liu, Z. (Ed.), *Frontiers in Applied Mechanics*. Imperial College Press
- Zhang, L., Gardiner, B.S., Smith, D.W., Pivonka, P., Grodzinsky, A.J., 2007. The effect of cyclic deformation and solute binding on solute transport in cartilage. *Archives of Biochemistry and Biophysics* 457, 47-56.
- Zhang, L., Gardiner, B.S., Smith, D.W., Pivonka, P., Grodzinsky, A.J., 2008. A fully coupled poroelastic reactive-transport model of cartilage. *Molecular & Cellular Biomechanics* 5, 133-153.
- Zhang, L., Gardiner, B.S., Smith, D.W., Pivonka, P., Grodzinsky, A.J., 2009. Integrated model of IGF-I mediated biosynthesis in deforming articular cartilage. *Journal of Engineering Mechanics (ASCE)* 135, 439-449.
- Zhang, L., Gardiner, B.S., Smith, D.W., Pivonka, P., Grodzinsky, A.J., 2010. The Transport of Insulin-like Growth Factor through Cartilage, in: Vafai, K. (Ed.), *Porous Media: Applications in Biological Systems and Biotechnology*. Taylor & Francis Group.
- Zhang, L., Miramini, S., Richardson, M., Pirpiris, M., Mendis, P., Oloyede, K., 2013. The effects of flexible fixation on early stage bone fracture healing. *International Journal of Aerospace and Lightweight Structures* 3, 181-189.
- Zhang, L., Richardson, M., Mendis, P., 2012. The role of chemical and mechanical stimuli in mediating bone fracture healing. *Clinical and Experimental Pharmacology and Physiology* 39, 706-710.

## Chord rotation demand for Effective Catenary Action under Monotonic

### Loadings

†\*Meng-Hao Tsai

Department of Civil Engineering, National Pingtung University of Science and Technology, Taiwan.

\*Presenting author: mhtsai@mail.npust.edu.tw

†Corresponding author: mhtsai@mail.npust.edu.tw

#### Abstract

In the past decade, several experimental and numerical studies were conducted with reinforced concrete (RC) beam-column sub-assemblages to investigate the progressive collapse resistance of frame structures under column loss. Most of the studies suggested that the catenary action could be used as the final defensive mechanism against collapse. However, it was observed from the load-deflection curves that there was a strength-decreased transition phase between the peak arch resistance and the commencement of catenary action. This transition region may imply an unstable snap-through behavior under a real dynamic column loss scenario. Hence, the chord rotation demands for effective catenary action of RC beams were investigated in this study. The nonlinear static load-deflection response of RC beam-column sub-assemblages under gravitationally monotonic loadings was idealized as a piecewise linear curve divided by the yield strength, peak arch resistance, leveled-off strength, and peak catenary resistance before bar fracture. The corresponding pseudo-static load response was then analytically derived for each linear region. Based on the analytical formulation, numerical analyses were carried out to understand the variation of the chord rotation demand with some key parameters related to the collapse-resistant performance. Parametric study results indicated that smaller peak-arch rotation and larger catenary stiffness could induce less rotation demand for the effective catenary action. This implies that RC beams with a deep section require larger rotation capacity for the effective catenary action. Since RC members with a deep section are usually responsible of large shear and/or moment, it is suggested that their peak arch strength is used as the collapse resistance for the sake of safety. Also, a peak-arch strength recovery in the nonlinear static response curve does not always guarantee a similar strength recovery in the pseudo-static counterpart. Complementary pseudo-static analysis is advised to verify the effective catenary action under realistic dynamic column loss.

**Keywords:** Progressive collapse, Effective catenary action, Pseudo-static response, Chord rotation

#### Introduction

Progressive collapse vulnerability of building structures has become an active research topic since the 9/11 terrorism attack of the World Trade Center in 2001. As stated in the ASCE 7-10 Standard [ASCE (2010)], progressive collapse is defined as “the spread of an initial local failure from element to element, resulting eventually in the collapse of an entire structure or a disproportionately large part of it”. Therefore, it is sometime indicated as “disproportional collapse”. Several experimental studies were performed with beam-column sub-assemblages, as shown in Fig.1, to investigate the progressive collapse resistance of frame structures in the past decade. Su et al. (2009) performed static vertical loading tests on twelve longitudinally restrained RC beams with varied steel ratios and span-to-depth ratios. The tested specimens generally reached peak compressive arch strength at a deflection ranging from 16% to 34% of section depth. For some specimens, the load resistance in catenary stage may be lower than the arch strength. Sasani et al. (2011) adopted a 3/8 scaled sub-assemblage to evaluate the column-loss response of an RC beam bridging over the removed column. Choi and Kim (2011) performed static loading tests on reduced-scale RC sub-assemblages designed with and without seismic detailing and concluded that significant catenary action may be activated for seismically detailed beams. Some dynamic loading tests have been

carried out to investigate the dynamic column-loss response of RC beam-column sub-assemblages and structural frames [Tian and Yu (2011), Qian and Li (2012), Orton and Kirby (2014)]. Yu and Tan (2013) designed eight RC sub-assemblage specimens with varied steel ratios and span-to-depth ratios to study the ultimate catenary resistance under column loss.

Most of the experiments suggested that the catenary action could be used as the final defensive mechanism against collapse. They also revealed that development of the catenary action was dependent on the beam-end rotational capacity. Hence, the design guidelines issued by the US General Service Administration [GSA (2003)] and the Department of Defense [DoD (2005), (2009)] have proposed an acceptance criterion of 0.20 radians for the chord rotation of the two-span beams, as defined in Fig. 1, in nonlinear progressive collapse analysis of RC frames. This threshold was specified independent of structural parameters. However, from those referred experimental studies, the rotational demands corresponding to the peak arch strength, commencement of catenary action, and bar-fracture strength were varied. Hence, analytical resolution of the rotational demands for effective catenary action of progressive collapse is proposed in this study. Piecewise linear curves divided by the yield strength, peak arch resistance, leveled off strength, and peak catenary resistance were used at first to idealize the general nonlinear static response of RC beam-column sub-assemblages under gravitationally monotonic loadings. Then, the corresponding pseudo-static load response was analytically derived for each linear region. A definition for the effective catenary action was described. The analytical expressions were used to carry out numerical investigations on the chord rotation demands of the effective catenary action and associated snap-through response. Practical implications were drawn based on the numerical analysis results.

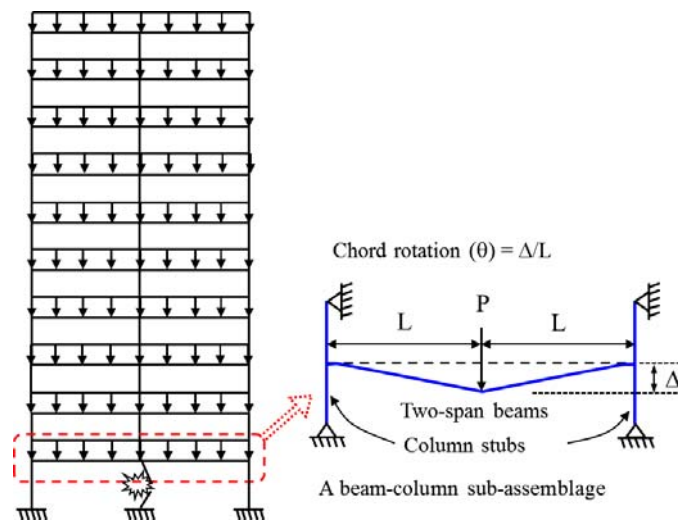


Figure 1 The definition of a beam-column sub-assemblage

### Idealization of Static Response

From the results of most monotonic static loading tests, it was revealed that the load-deflection responses of RC beam-column sub-assemblage specimens were highly nonlinear. The nonlinear static response was initiated at the tensile cracking of concrete and grew significantly as the tensile reinforcement yielded. Along with the flexural yielding, the load response gradually increased to the so-called peak arch resistance, as shown in Fig. 2(a). This load-deflection range may be defined as the “compressive arch” phase. In this phase, compressive axial force is developed in the beam members of the RC sub-assemblage due to the restraint provided by the end columns. From most of the experimental and numerical studies, it was observed that there is a strength-decreased transition phase after the peak arch resistance. It is defined as the “transition phase”. This strength-decreased

region may induce snap-through response under a real dynamic column loss scenario [Tsai (2012), Orton and Kirby (2014)]. Analytical and numerical studies [Tsai and Lin (2008), Tsai (2010), Tsai and You (2012)] have indicated that pseudo-static response obtained from the nonlinear static load-deflection curve may be used to predict the maximum dynamic response under column loss. As the supported loading is larger than the dynamic peak arch resistance, the beam-column sub-assembly will be loaded directly into the catenary phase and significantly large deformation could be induced under column loss. This reveals that the peak arch resistance is an important threshold for the snap-through behavior. In this phase, the load resistance may gradually decrease and level off at  $P_c$ , where the catenary action is activated. The load resistance may be regained under the catenary action until any of the steel bars fails in rupture and this region is thus defined as the “catenary phase”. Axial tension is developed in the beam members during the catenary phase and provides the collapse resistance for the two-span beams.

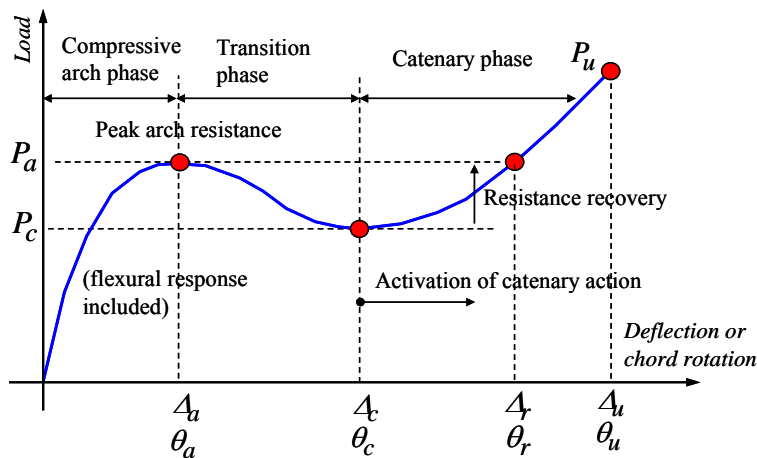


Figure 2(a) Static load-deflection curve under gravitational monotonic loadings

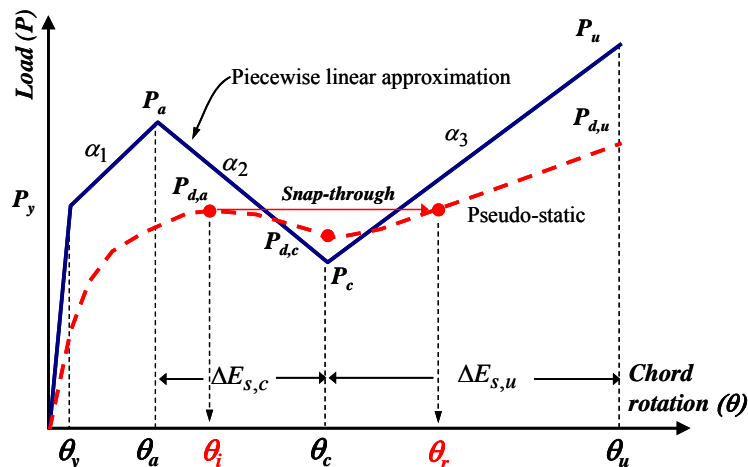


Figure 2(b) Idealized piece-wise linear curve and pseudo-static response

Although the general static load-deflection curve is nonlinear, as a rule of thumb, it can be approximated as piece-wise linear response with four threshold points, as shown in Fig. 2(b). The four threshold points are corresponding to the yielding strength, peak arch resistance, levelled off strength, and peak catenary resistance before bar fracture. As shown in the figure, the stiffness

ratios of the three post-yield regions to the elastic stiffness are designated by  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ .  $\alpha_1$  is defined as the arch stiffness ratio ranging from the yield point to the peak arch resistance.  $\alpha_2$  is defined as the softening stiffness ratio ranging from the peak arch resistance to the leveled off point.  $\alpha_3$  is defined as the catenary stiffness ratio for the catenary phase. With these parameters, the analytical pseudo-static response in each phase can be obtained for the idealized nonlinear static curve.

### Analytical pseudo-static response

The pseudo-static loading may be numerically obtained from dividing the accumulated area under the nonlinear static load-displacement curve by the corresponding displacement of the column-loss point. Mathematically, it may be expressed as

$$P_{CC}(u_d) = \frac{1}{u_d} \int_0^{u_d} P_{NS}(u) du \quad (1)$$

where  $P_{CC}(u)$  and  $P_{NS}(u)$  are, respectively, the pseudo-static loading and the nonlinear static loading at the displacement demand  $u$ . For the idealized nonlinear static response, the pseudo-static force in the elastic range may be written as

$$P_{d,0} = P/2, \quad 0 \leq P \leq P_y \quad (2)$$

where  $P_y$  is the static yielding force. From the yield point to the peak arch resistance, it is derived that the pseudo-static force  $P_{d,1}$  can be expressed as

$$P_{d,1} = \frac{P_y[\alpha_1(\mu-1)^2 + 2(\mu-1) + 1]}{2\mu}, \quad 1 \leq \mu \leq \mu_a \quad (3)$$

where the ductility,  $\mu$ , is the chord rotation divided by the yield rotation  $\theta_y$ .  $\mu_a = \theta_a/\theta_y$ , which denotes the ductility demand at the peak arch resistance  $P_a$  (Fig. 2(b)). Similarly, the pseudo-static forces in the transition and catenary phases are respectively derived as

$$P_{d,2} = \frac{P_{d,a}\mu_a}{\mu} + \frac{P_y[-\alpha_2(\mu-\mu_a)^2 + 2(\mu-\mu_a)[1 + \alpha_1(\mu_a-1)]]}{2\mu}, \quad \mu_a \leq \mu \leq \mu_c \quad (4)$$

and

$$P_{d,3} = \frac{P_{d,c}\mu_c}{\mu} + \frac{P_y[\alpha_3(\mu-\mu_c)^2 + 2(\mu-\mu_c)[1 + \alpha_1(\mu_a-1) - \alpha_2(\mu_c-\mu_a)]]}{2\mu}, \quad \mu_c \leq \mu \quad (5)$$

where  $P_{d,a} = P_{d,1}(\mu = \mu_a)$  and  $P_{d,c} = P_{d,2}(\mu = \mu_c)$ .  $\mu_c = \theta_c/\theta_y$ , which represents the ductility demand at the end of transition phase. A general form for the pseudo-static force in the  $i$ -th linear region and  $i \geq 2$  may be deduced from the above equations as

$$P_{d,i} = \frac{P_{d,i-1}(\mu_{i-1}) \cdot \mu_{i-1}}{\mu} + \frac{P_y[\alpha_i(\mu - \mu_{i-1})^2 + 2(\mu - \mu_{i-1})[1 + \sum_{j=1}^{i-1} \alpha_j(\mu_j - \mu_{j-1})]]}{2\mu}, \quad \mu_{i-1} \leq \mu \quad (6)$$

where  $\mu_{i-1}$  is the ductility demand of the previous turning point and the sign of stiffness is included in the ratios  $\alpha_i$  and  $\alpha_j$ .

As shown in Fig. 2(b), the pseudo-static peak arch resistance, denoted as  $P_{p,a}$ , does not occur at the chord rotation  $\theta_a$  corresponding to its static counterpart. Instead, it happens during the transition phase, ie. in the range from  $\theta_a$  to  $\theta_c$ . From setting the derivative of Eq.(4) equal to zero, it can be obtained that  $P_{p,a}$  occurs at

$$\mu \equiv \mu_i = \sqrt{\frac{(\alpha_2 + 2\alpha_1)\mu_a^2 + 2(1 - \alpha_1)\mu_a}{\alpha_2}} \quad (7)$$

The value of  $P_{p,a}$  is then calculated as  $P_{d,2}(\mu = \mu_i)$ . The chord rotation at  $\mu_i$  is denoted as  $\theta_i$  in Fig. 2(b). This rotation is defined as the snap-through prevention limit. It can be regarded as an index to judge the importance of the catenary action. If  $\theta_i$  is larger than the expected beam-end rotation, the catenary action will be minor under the column loss. Moreover, from the comparison of the nonlinear static and pseudo-static load-deflection curves, it is clear that if the static leveled off rotation  $\theta_c$  is less than  $\theta_i$ , then the pseudo-static response shall be a non-degrading curve with non-negative tangent stiffness. In such a case, there will be no snap-through response under dynamic column loss [Tsai (2012)] and the catenary action is always effective in enhancing the collapse resistance. However, as  $\theta_c$  is larger than  $\theta_i$ , the pseudo-static resistance may be lower than  $P_{p,a}$  and the snap-through response will be induced consequently. Once it happens, the dynamically falling behavior can be arrested only if the resistance of  $P_{p,a}$  may be regained in the catenary phase. Otherwise, dynamic collapse will happen. Therefore, an effective catenary action is defined as the capability of recovering the strength of  $P_{p,a}$  in the catenary phase. The chord rotation demand for the effective catenary action is then determined from  $P_{d,3} \geq P_{p,a}$ , which leads to

$$A\mu^2 + B\mu + C \geq 0 \quad (8)$$

where  $A = \alpha_3$ ,  $B = 2[1 + \alpha_1(\mu_a - 1) - \alpha_2(\mu_c - \mu_a) - \alpha_3\mu_c - 2P_{d,i}/P_y]$ , and

$$C = \alpha_3\mu_c^2 - 2\mu_c[1 + \alpha_1(\mu_a - 1) - \alpha_2(\mu_c - \mu_a) - 2P_{d,c}/P_y] \geq 0.$$

Two values of ductility demand,  $\mu$ , can be resolved from Eq.(8). The one larger than  $\mu_c$ , denoted as  $\mu_r$ , is the ductility demand for the effective catenary action. Its corresponding rotation is designated as  $\theta_r$  in Fig. 2(b).

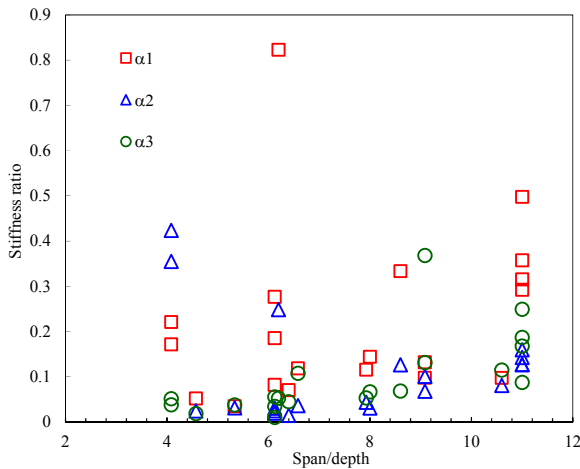


### Parametric study

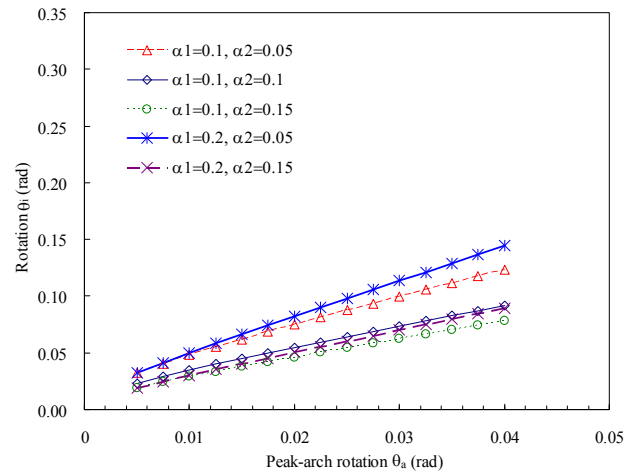
From the previous derivation, it was observed that the chord rotation demand for the effective catenary action was involved with several parameters, which included the three stiffness ratios and chord rotations of the static arch resistance and leveled off point. The stiffness ratios were influenced by the span-to-depth ratio, reinforcement details, material strength, and boundary constraints of the members bridging the removed column. Fig. 3 shows the variations of  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  with the span-to-depth ratio, which were estimated from several published test results [Su et al. (2009); Yu and Tan (2013); Lew et al. (2014); Tsai et al. (2013, 2014)]. They were calculated by using the static peak arch, leveled off, and maximum catenary response prior to bar fracture. It is observed that most of the stiffness ratios varied from 0.1 to 0.3 for  $\alpha_1$ , from 0.05 to 0.15 for  $\alpha_2$ , and from 0.05 to 0.2 for  $\alpha_3$ . Therefore, the stiffness ratios considered in this study were determined as shown in Table 1. In the referred experimental results, a larger span-to-depth ratio generally led to a larger  $\alpha_1$ . However, the corresponding chord rotation of peak arch resistance decreased with increased span-to-depth ratio.

**Table 1 Stiffness ratios of the three phases for the parametric study**

$\alpha_1$	$\alpha_2$	$\alpha_3$
0.1	0.05 ~ 0.15 @ 0.05	0.05 ~ 0.2 @ 0.05
0.2	0.05 ~ 0.15 @ 0.05	0.05 ~ 0.2 @ 0.05



**Figure 3 Variation of stiffness ratios estimated from published test results**



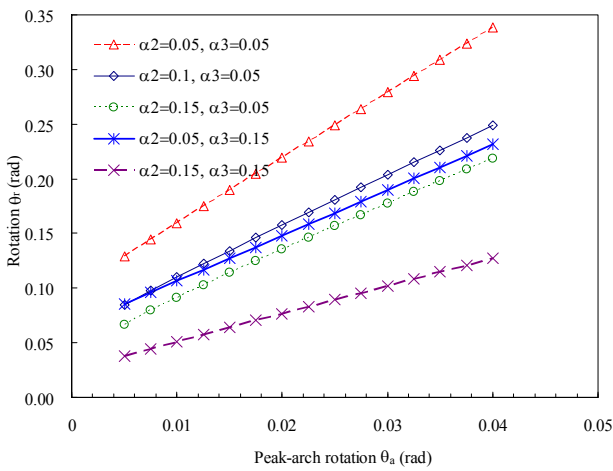
**Figure 4 Variation of snap-through prevention rotations**

In order to investigate the chord rotation demand for the effective catenary action, it was assumed that the static leveled-off strength  $P_c = 0.5P_y$ . This assumption was made to confirm that the pseudo-static peak arch resistance ( $P_{p,a}$ ) occurred by the end of the transition phase. Then the leveled-off ductility was obtained as

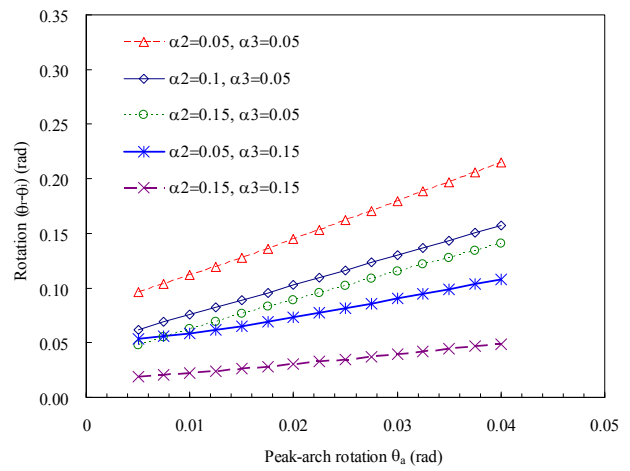
$$\mu_c = \frac{0.5 + \alpha_1(\mu_a - 1) + \alpha_2\mu_a}{\alpha_2} \quad (9)$$

The corresponding leveled-off rotation is determined from the product of the yield rotation  $\theta_y$  and  $\mu_c$ . The yield rotation was assumed as 0.005 rads [FEMA (2000)] in this study. According to the selected stiffness ratios in Table 1, the variations of the snap-through prevention rotation ( $\theta_i$ ) under

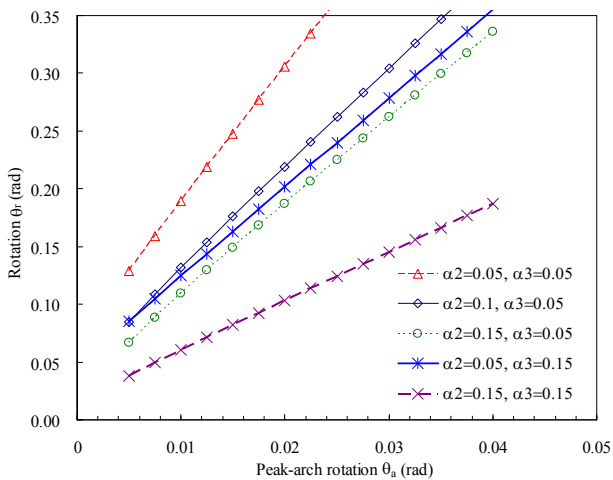
five different combinations of  $\alpha_1$  and  $\alpha_2$  are shown in Fig. 4. As implied in Eq.(7), it is observed that an increased  $\theta_a$  could induce a larger snap-through prevention limit. Effect of the softening stiffness ratio is opposite to that of the arch stiffness ratio on  $\theta_i$ , which appears more sensitive to the former. Fig. 5(a) shows the comparison of the minimum rotation demands ( $\theta_r$ ) for effective catenary action under five different parametric combinations with  $\alpha_1 = 0.1$ . The corresponding snap-through response, which was obtained from  $(\theta_r - \theta_i)$ , is shown in Fig. 5(b). These figures reveal that the rotation demands of both the effective catenary action and snap-through response decreased with increased softening and catenary stiffness ratios. However, they increased with the peak arch rotation. This means that if the snap-through behavior is delayed, more plastic deformation must be developed for the effective catenary action.



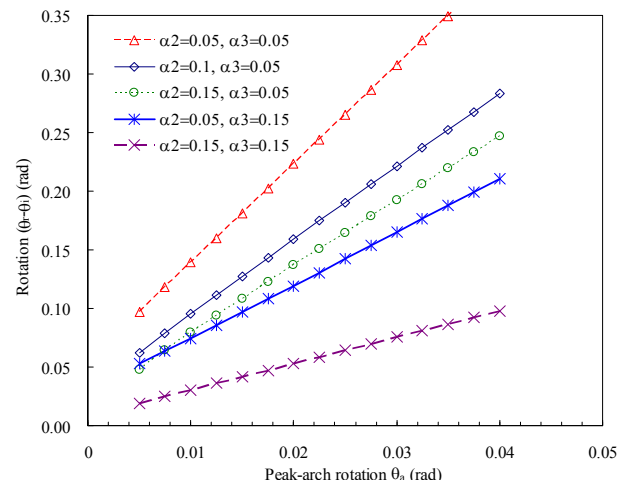
**Figure 5(a) Minimum rotation demands ( $\theta_r$ ) for effective catenary action with  $\alpha_1 = 0.1$**



**Figure 5(b) Snap-through rotation ( $\theta_r - \theta_i$ ) with  $\alpha_1 = 0.1$**



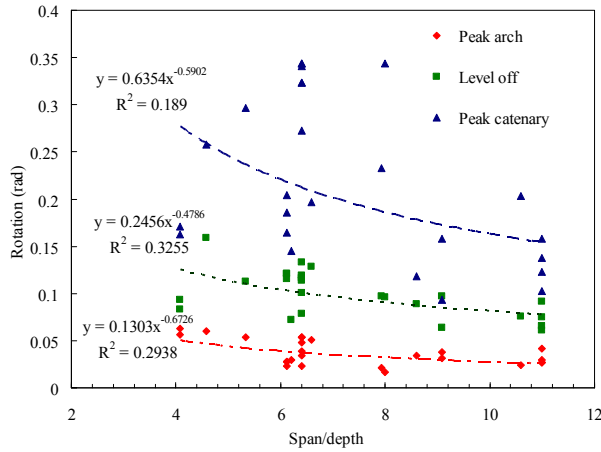
**Figure 6(a) Minimum rotation demands ( $\theta_r$ ) for effective catenary action with  $\alpha_1 = 0.2$**



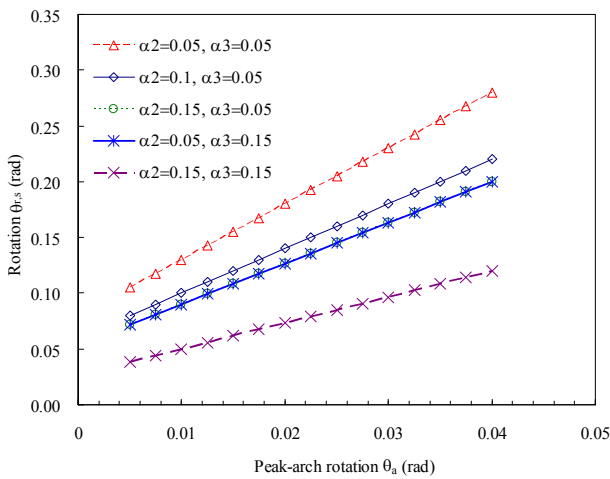
**Figure 6(b) Snap-through rotation ( $\theta_r - \theta_i$ ) with  $\alpha_1 = 0.2$**

It is noted that a large  $\alpha_2$  represents swifter stiffness degradation in the transition phase and thus an earlier activation of the catenary action. It may occur in members suffered from shear failure during the gravitational monotonic loading process [Tsai et al. (2013)]. Also, it is observed that both  $\alpha_2$

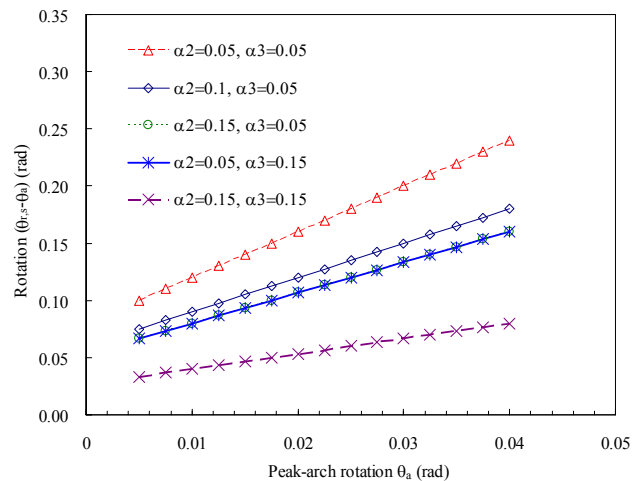
and  $\alpha_3$  had similar influence in the rotation demands of effective catenary action and snap-through response. Either increasing  $\alpha_2$  from 0.05 to 0.15 with  $\alpha_3 = 0.05$  or  $\alpha_3$  from 0.05 to 0.15 with  $\alpha_2 = 0.05$  could result in approximate rotational reduction. As indicated in some test results, larger peak-arch rotation and smaller  $\alpha_2$  were generally resulted from specimens with a deeper section. This implies that more critical chord rotation demand may be advised under such conditions. The acceptance criterion of 0.20 radians, which is regarded as the minimum demand for catenary development as recommended in the UFC guidelines, may not be always conservative for the RC beams with a deep section.



**Figure 7 Experimental variations of the threshold chord rotations (estimated from the referred test results)**



**Figure 8(a) Static rotation demands for the effective catenary action**



**Figure 8(b) Static rotation demands for the snap-through response**

Figs. 6(a) and 6(b) show the chord rotations for effective catenary action and snap-through response with a larger arch stiffness ratio,  $\alpha_1 = 0.2$ . Compared with the corresponding Figs. 5(a) and 5(b), it is observed that a larger arch stiffness ratio may increase the rotation demands for effective catenary action, snap-through prevention limit, and snap-through deformation. From the referred test results, a larger span-to-depth ratio generally led to increased arch and catenary stiffness ratios, as observed from Fig. 3. Thus, the adverse influence of the increased arch stiffness on the chord rotation demand may be partially mitigated by the increased catenary stiffness. Also, from the

experimentally observed variations of the threshold chord rotations, as shown in Fig. 7, the rotation at peak arch resistance decreased with increased span-to-depth ratios. This could further reduce the rotation demands of effective catenary action and snap-through response for RC beams with a larger span-to-depth ratio. These observations explain why RC beams with shallower sections usually present earlier triggered catenary actions.

Although dynamic tests may reflect the realistic column-loss scenarios, the experimental costs for test setup and instrumentation are usually larger than conventional static tests. Hence, there were more experimental studies conducted with static tests. From the idealized piece-wise linear curve, a ductility demand for statically effective catenary action may be obtained from  $P_a = P_u$ . The resulting static rotation demand is written as

$$\theta_{r,s} = \theta_y \frac{(\alpha_2 + \alpha_3)\mu_c - \alpha_2\mu_a}{\alpha_3} \quad (10)$$

Different from the pseudo-static rotation demand obtained from Eq.(8), the static rotation is independent of the arch stiffness ratio. Figs. 8(a) and 8(b) show the static rotation demands for the effective catenary action and snap-through response based on Eq.(10). The comparison of Figs. 8 and Figs. 5 indicates that the rotation demand, either for the effective catenary action or the snap-through response, may be underestimated if based on the nonlinear static response only. Hence, if the static monotonic test results of RC beams present a recovery of the static peak arch resistance in the catenary phase, complementary pseudo-static analysis is necessary to verify the effective catenary action under realistic dynamic column loss.

## Conclusions

Because of the widespread attention paid to the progressive collapse resistance of building structures under column loss, chord rotation demands for the effective catenary action of RC beams were investigated in this study. The nonlinear static response of RC beams under monotonic pushdown loadings was idealized as a piece-wise linear curve with four threshold points. Based on the idealized static response, analytical formulations were derived to determine the pseudo-static chord rotation for the effective catenary action and accompanied snap-through response. The numerical parametric studies revealed that the minimum rotation demand of 0.20 radians recommended in the UFC guidelines for catenary development was not always conservative. Larger peak-arch rotation and smaller catenary stiffness may increase the rotation demand for effective catenary action. This implied that RC beams with a deep section may need higher rotation capacity for catenary development. Since RC members with a deep section are usually responsible of large shear and/or moment, it is suggested to adopt their peak arch strength as the collapse resistance for the sake of safety. Also, a peak-arch strength recovery in the nonlinear static response curve does not always guarantee a similar strength recovery in the pseudo-static counterpart. Complementary pseudo-static analysis is advised to verify the effective catenary action under realistic dynamic column loss. In general, a combination of smaller arch stiffness, larger softening stiffness, and larger catenary stiffness may lead to an earlier strength recovery in the catenary phase. Since the arch, softening, and catenary stiffness are involved with the span-to-depth ratio, main reinforcement ratio and layout, shear reinforcement, and boundary constraint of the RC members, it will be an important task to clarify their relationships in future studies.

## References

- ASCE (2010) *Minimum Design Loads for Buildings and Other Structures*, ASCE/SEI 7-10, American Society of Civil Engineers, Reston, Virginia.
- Choi, H. and Kim, J. (2011) Progressive collapse-resisting capacity of RC beam-column sub-assembly, *Magazine of Concrete Research* **63**, 297-310.

- DoD (Department of Defense) (2005) *Design of Buildings to Resist Progressive Collapse, UFC 4-023-03*, Washington, DC, USA.
- DoD (Department of Defense) (2009) *Design of Buildings to Resist Progressive Collapse, UFC 4-023-03*, Washington, DC, USA.
- FEMA (2000) *Prestandard and Commentary for the Seismic Rehabilitation of Buildings, FEMA-356*, Federal Emergency Management Agency, Washington, DC, USA.
- GSA (General Service Administration) (2003) *Progressive Collapse Analysis and Design Guidelines for New Federal Office Buildings and Major Modernization Projects*, Washington, DC, USA.
- Lew, H. S., Bao, Y., Pujol, S., and Sozen, M. A. (2014) Experimental study of reinforced concrete assemblages under column removal scenarios, *ACI Structural Journal* **111**(4), 881-892.
- Orton, S. and Kirby, J. (2014) Dynamic response of an RC frame under column removal, *Journal of Performance of Constructed Facilities* **28**(4), 04014010-1-8.
- Qian, K. and Li, B. (2012) Dynamic performance of reinforced concrete beam-column substructures under the scenario of the loss of a corner column — experimental results, *Engineering Structures* **42**, 154-167.
- Sasani, M., Werner, A., and Kazemi, A. (2011) Bar fracture modelling in progressive collapse analysis of reinforced concrete structures, *Engineering Structures* **33**, 401-409.
- Su, Y., Tian, Y., and Song, X. (2009) Progressive collapse resistance of axially-restrained frame beams, *ACI Structural Journal* **106**(5), 600-607.
- Tian, Y. and Su, Y. (2011) Dynamic response of reinforced concrete beams following instantaneous removal of a bearing column, *International Journal of Concrete Structures and Materials* **5**(1), 19-28.
- Tsai, M. H. and Lin, B. H. (2008) Investigation of progressive collapse resistance and inelastic response for an earthquake-resistant RC building subjected to column failure, *Engineering Structures* **30**(12), 3619-3628.
- Tsai, M. H. (2010) An analytical methodology for the dynamic amplification factor in progressive collapse evaluation of building structures, *Mechanics Research Communications* **37**, 61-66.
- Tsai, M. H. and You, Z. K. (2012) Experimental evaluation of inelastic dynamic amplification factors for progressive collapse analysis under sudden support loss, *Mechanics Research Communications* **40**, 56-62.
- Tsai, M. H., Lu, J. K., and Chang, Y. T. (2013) Experimental investigation on the collapse resistance of RC beam-column sub-assemblages, *Proceeding the 6th Civil Engineering Conference in Asia Region*, Jakarta, Indonesia, August 20-22, TS8-35-41.
- Tsai, M. H., Lu, J. K., and Huang, B. H. (2014) Column-loss response of RC beam-column sub-assemblages with different bar-cutoff patterns, *Structural Engineering and Mechanics*, **49**(6), 775-792.
- Yu, J. and Tan, K. H. (2013) Structural behaviour of RC beam-column sub-assemblages under a middle column removal scenario. *Journal of Structural Engineering* **139**(2), 233-250.

## **Evaluation of the Singular Stress Field of the Crack Tip of the Infinite Elastic Plate**

**† \* J.Y. Liu<sup>1,2</sup>, G. Lin<sup>1</sup>, Z.C.Sun<sup>3</sup>, P. Zhang<sup>2</sup> and Y.Y. Wang<sup>2</sup>**

<sup>1</sup> Faculty of Infrastructure Engineering, Dalian University of Technology, Dalian 116024, China.

<sup>2</sup> School of Architecture & Civil Engineering, Shenyang University of Technology, Shenyang 110870, China

<sup>3</sup> School of Civil & Environmental Engineering, University of New South Wales, Sydney NSW 2031, Australia

†\*Liu junyu: [liujunyu2003@163.com](mailto:liujunyu2003@163.com); Lin Gao: [Gaolin@dlut.edu.cn](mailto:Gaolin@dlut.edu.cn); Sun zhicheng: [815487759@qq.com](mailto:815487759@qq.com); Wang yuyang: [1019016079@qq.com](mailto:1019016079@qq.com); Zhangping: [1063929049@qq.com](mailto:1063929049@qq.com);

### **Abstract**

In the Scaled Boundary Finite Element Method (abbr. SBFEM), the analytical advantage of the solution in the radial direction allows stress intensity factors (SIFs) to be determined directly from its definition. Therefore no special crack-tip treatment is necessary. In addition, the stiffness of infinite domain can be solved analytically. In present paper, the asymptotic fields of the crack tip of the infinite plate subjected to the different loadings are computed based on the SBFEM combining the sub-structuring technique (or super-element), extracting the SIFs, some of higher order terms. The essential calculating formula of SBFEM is derived. The numerical results are compared with those in the literature, and the results show that SBFEM can evaluate the asymptotic fields of the crack tip with higher efficiency and accuracy. In addition, some of the higher order terms may provide evidence for the further research on the fracture characteristics of the mass concrete materials and structures.

**Keywords:** stress intensity factors; the scaled boundary finite element method; high order singular terms; infinite plate ;

### **Introduction**

The analytical expression of stress fields of crack tip is presented by Williams[Williams (1957)]. It includes stress intensity factors  $K_I$ ,  $K_{II}$ , T-stress and higher order terms which is named crack-tip parameters. Theoretical analysis and test can not determine parameters of the complex case, which is concerned on the material properties of structure, load form, and the initial angle of crack. Some of the numerical methods are applied to determine these parameters of practical and complex problem. The mainly numerical methods which can evaluate the singular fields of crack tip, include the finite element method(FEM), the boundary element method(BEM), the weight function method, the finite difference method and the scaled boundary finite element method(SBFEM) [(Deeks and Chidzey(2005); Song(2005 )]. FEM need discretizing crack tip area, and BEM need discretizing crack surface area. Furthermore, they can't gain analytical solution around singularity points in general case, because of their dependence on the piecewise smooth functions [Song ( 2002)]. The most important advantage of SBFEM is the stress singularity along the radial direction emanating from the crack tip is represented analytically. No analytical asymptotic expansion or enrichment is required. SIFs, T-stress and higher order terms are extracted directly [(Deeks and Chidzey(2005); Song(2005 )]. According to the advantages of SBFEM, it has been applied to evaluating the dynamic stress intensity factors [Song(2004;2008)], the SIF of orthotropic material[Song ( 2002)], the SIF under the temperature load [Song(2006)], and a unified definition of GSIFs was proposed in [Song et al (2010)]. Crack propagation was modelled in [Yang (2006);Yang and Deeks (2007); Ooi and Yang (2009; 2011a; 2011b); Bird et al (2010); Shi (2013) and Zhu(2014)]. Other applications in fracture mechanics include [Li (2014);Liu(2008)].

In the paper, the asymptotic fields of the crack tip of the infinite plate subjected to the different loadings are computed based on the SBFEM combining the sub-structuring technique (or super-element), extracting the SIFs, some of higher order terms of crack tip are provided. Numerical examples are provided to demonstrate the effectiveness and accuracy. The results are compared with those of analytical solutions and numerical solutions. The comparison shows that SBFEM can calculate the asymptotic field of crack tip with accuracy.

## 2 The Fundamental Equations of Scaled Boundary Finite Element Method

With emphasis placed on the two-dimensional problems the concept of the scaled boundary finite element method and the necessary equations for a bounded medium are summarized.

The governing equations of SBFEM without dynamic problems is the following

$$\left[ E^0 \right] \xi^2 \{u(\xi)\}_{,\xi\xi} + \left( \left[ E^0 \right] - \left[ E^1 \right] + \left[ E^1 \right]^T \right) \xi \{u(\xi)\}_{,\xi} - \left[ E^2 \right] \{u(\xi)\} = 0 \quad (1)$$

where  $\left[ E^0 \right]$ ,  $\left[ E^1 \right]$  and  $\left[ E^2 \right]$  are the coefficient matrices on the boundary [Song and Wolf(1997)].

The solution of displacement field is expressed as the following

$$\{u(\xi)\} = \sum_{i=1}^n c_i \xi^{-\lambda_i} \{\phi\}_i \quad (2)$$

where, n is the dimension of eigen value vector matrix  $[\Phi_{11}]$  whose meaning can be seen in the following formula.  $\{\phi\}_i$  is i column of the matrix  $[\Phi_{11}]$ ,  $c_i$  is the i element of the integration constant vector  $\{c_i\}$ . Then the radial displacement field within the sub-structure (super-element) can be obtained by interpolation through the function  $[N^u(\eta)]$ , therefore

$$\{u(\xi, \eta)\} = \sum_{i=1}^n c_i \xi^{-\lambda_i} [N^u(\eta)] \{\phi\}_i \quad (3)$$

the stress field within the sub-structure

$$\{\sigma(\xi, \eta)\} = \sum_{i=1}^n c_i \xi^{-\lambda_i-1} \{\psi(\eta)\}_i \quad (4)$$

where,  $\{\psi(\eta)\}_i$  is the stress mode of local coordinate  $\eta$ , it can be calculated by the corresponding displacement mode  $\{\phi\}_i$ ,

$$\{\psi(\eta)\}_i = [D] \left( -\lambda_i [B^1(\eta)] + [B^2(\eta)] \right) \{\phi\}_i \quad (5)$$

The solutions of displacement and stress in the sub-structure can be expressed as polar-coordinate form in order to calculate SIF. Radial coordinate can be expressed as

$$\hat{r}(\xi, \eta) = \xi r(\eta) \quad (7)$$

where,  $r(\eta) = \sqrt{x^2(\eta) + y^2(\eta)}$  is the radial coordinate on the boundary of sub-structure. Angle  $\theta$  is only related to  $\eta$ ,

$$\theta(\eta) = \arctan \frac{y(\eta)}{x(\eta)} \quad (8)$$

Eq.(7) is substituted into Eq.(4), we can get

$$\{\sigma(\hat{r}, \eta)\} = \sum_{i=1}^n c_i \hat{r}^{-\lambda_i-1} (r(\eta))^{\lambda_i+1} \{\psi(\eta)\}_i \quad (9)$$

Eq.(8) and Eq.(9) constitute the stress field similar to Williams which is expressed by the series coordinate. The expression of the singular fields of crack tip of Williams is used.

### 3 Numerical Examples

#### 3.1 Edge-cracked semi-infinite plate under uniform tension tractions

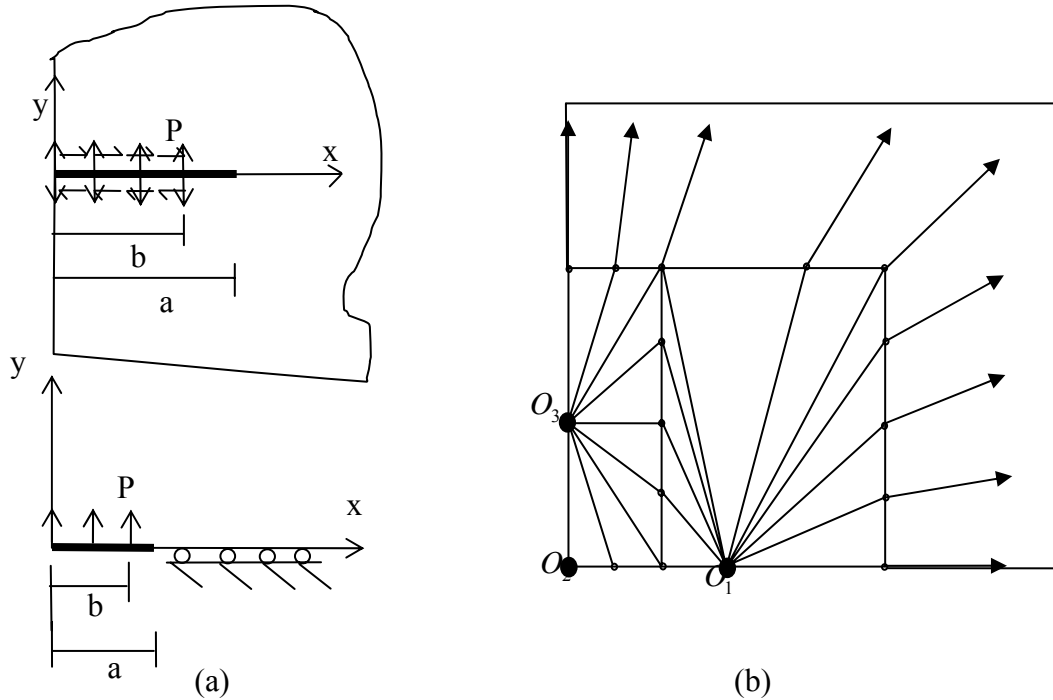


Fig.3a Edge-cracked semi-infinite plate under uniform tension traction  
 Fig.3b Boundary discretization of edge-cracked semi-infinite plate

First considering an edge-cracked semi-infinite plate subjected to uniform distributing tension  $p_0$  loaded on the crack faces as shown in Fig.3a. The semi-infinite plate is divided into three blocks (Fig.3b), and the corresponding scaling centers are  $O_1(0,0)$ ,  $O_2(-1,0)$  and  $O_3(-1,2)$ , respectively. The geometrical parameter of the first block is:  $W=1.5, H=4$ , the crack length is  $a=1$ . The length of loading for the crack face is  $b=0.35, 0.5, 0.75, 0.85$ . The material properties elastic modulus  $E=1$ , Poisson's ratio  $\nu=0.25$ . The Fig.3b gives the discretization model, and the total elements is 16 3-node line elements. The analytical solutions of SIF are shown in Reference[Hiroshi(2000)], and the present compared results and relative difference (RD) are shown in Table 1. The higher order terms  $a_2, a_3$  are shown in Table 1. They are in good agreement and the maximum difference is less than 3%. No analytical solutions or numerical results for the  $a_2, a_3$  are compared.

Table 1 Results of edge-cracked semi-infinite plate under uniform tension tractions (  $W=1.5$  ,  $N=16$  )

The length of loading: b	KI	Analytical solutions	(RD)(%)	a2	a3
0.35	0.5130	0.4998	-2.6342	0.2728	-0.2698
0.5	0.7317	0.7142	-2.4510	0.3474	-0.4386
0.75	1.1385	1.1103	-2.5378	0.4376	-0.9263
0.85	1.2698	1.3100	3.0722	0.6191	-1.3251

Table 2 Results of edge-cracked semi-infinite plate under uniform shear tractions

Loading length b	KII	Analytical solution	RD(%)
0.65	0.9763	0.9415	-3.7035
0.70	1.0401	1.0233	-1.6435
0.75	1.1049	1.1103	0.4871
0.80	1.1803	1.2046	2.0241
0.85	1.2879	1.3100	1.6887
0.90	1.4542	1.4336	-1.4335

#### 3.3 Edge-cracked semi-infinite plane under uniform shear tractions

Fig.3a shows the Edge-cracked semi-infinite plate under uniform shear tractions model. The geometrical parameter is  $W=1.5, H=4$ . The length of crack  $a=1$ , the loading length of uniform shear  $b=0.65, 0.70, 0.75, 0.80, 0.85, 0.90, 0.95$ . The material properties elastic modulus  $E=1$ , Poisson's ratio



$\nu=0.25$ . The total number of elements is 8. Reference [Hiroshi(2000)] gives the analytical solutions. The comparison of the present results with analytical solutions are shown in Table 3. Both the numerical and analytical results are in good agreement. The maximum difference is less than 4%.

**3.3 Edge-cracked semi-infinite plane under concentrated shear tractions**

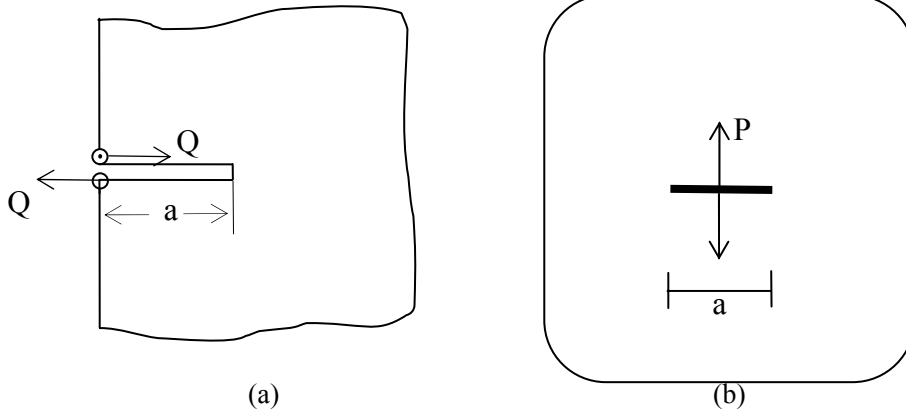


Fig.4a Edge-cracked semi-infinite plate under concentrated shear tractions

Fig.4b Edge-cracked infinite plate under concentrated tension tractions

Fig.4 shows the Edge-cracked semi-infinite plane model under concentrated shear tractions. The geometrical parameter is  $W=2, H=4$ . The length of crack is  $a=0.7,0.8,0.9,1.0,1.1,1.2,1.3$ . The material properties elastic modulus  $E=1$ , Poisson's ratio  $\nu=0.25$ . The discretization is shown in Fig.3b and the total number of elements is 19. The analytical solutions of SIF are shown in Reference[Hiroshi(2000)], and the comparison of the present results with analytical solutions are shown in Table 2. Both the numerical and analytical results are in good agreement. The maximum difference is less than 2%. The higher order terms  $a_3$  are also shown in Table 2. No analytical solutions or numerical results are compared.

Table 3 Results of edge-cracked semi-infinite plate under concentrated shear tractions

Crack length a	KII	Analytical solution	RD (%)	a3
0.7	1.7371	1.7492	0.6917	-2.1328
0.8	1.6263	1.6363	0.6108	-1.9439
0.9	1.5353	1.5427	0.4795	-1.5322
1.0	1.4598	1.4635	0.2522	-1.1236
1.1	1.3968	1.3954	-0.0990	-0.8287
1.2	1.3438	1.3360	-0.5827	-0.6297
1.3	1.2988	1.2836	-1.1865	-0.4953

Table 4 Results of SIF,  $a_1, a_2$  and  $a_3$  of the edge-cracked infinite plate under concentrated tension tractions

Number of elements	KI	RD (%)			
		(Analytical solution 0.5642)	$a_1$	$a_2$	$a_3$
5	0.5605	0.6532	0.2236	0.0263	
7	0.5603	0.6864	0.2235	0.0131	-0.2904
11	0.5680	-0.6725	0.2266	0.0111	-0.2889
19	0.5707	-1.1562	0.2277	0.011	-0.2885

**3.4 Edge-cracked infinite plane under concentrated tension tractions**

Edge-cracked infinite plane under concentrated tension tractions model is studied, as shown in Fig.4b. The geometrical parameter is  $W=1.5, H=4$ . The length of crack  $a=1$ . The material properties elastic modulus  $E=1$ , Poisson's ratio  $\nu=0.25$ . The analytical solutions of SIF are shown in Reference [Hiroshi(2000)]. The comparison of the present results with analytical solutions are shown in Table 4. The results of higher order terms  $a_2$  and  $a_3$  for the different element number are shown in Table 5. It shows that the numerical and analytical results are in good agreement. The maximum difference is less than 2%.

## Conclusions

In the paper, the singular stress fields of edge-cracked infinite plate are computed based on the SBFEM, extracting the SIF and the coefficients of higher order terms. The expression of displacement field and stress field of SBFEM and the equation of asymptotic field of crack tip of fracture mechanics are derived. The results are compared with the ones of analytical solutions and some numerical results, which shows that the SBFEM can calculate the asymptotic field of crack tip with accuracy. In addition, the results of T-stress and the coefficients of higher order terms have certain significance on determining crack stability and studying the fracture characteristics of crack tip.

## Acknowledgements

This work was financially supported by the National Natural Science Youth Foundation of China (51109134, 51208310, 11102118), China Postdoctoral Science Foundation (2013T60283, 2011M500557) and Program for Liaoning Excellent Talents in University (No. LJQ2011009). Thanks the Professor of the University of New South Wales, Australia, Song Chongmin, for his valuable assistance. Thanks for the International Postdoctoral Exchange Fellowship Program 2013 by the Office of China Postdoctoral Council

## References

- Deeks, A. J. & Chidgzy, S. R. 2005, Determination of coefficients of crack tip asymptotic fields using the scaled boundary finite element method, *Engineering Fracture Mechanics*, Vol. 72, 2019-2036.
- Hiroshi T, Paul C P, George R I. The Stress Analysis Of Cracks Handbook(Third Edition)[M]. Third Edition ed. New York: ASEM Press, 2000.
- Li Chao. Fracture analysis of piezoelectric composites using scaled boundary finite element method [D].Sydney: University of New South Wales, 2014
- LIU Junyu. Effect of the Water Pressure inside the Crack on the Fracture Behavior of Concrete Gravity Dam [D]. Da Lian: Dalian University of Technology, 2008. (in Chinese)
- Ooi, E. & Yang, Z. 2009, Modelling multiple cohesive crack propagation using a finite element scaled boundary finite element coupled method, *Engineering Analysis with Boundary Elements*, Vol. 33, No. 7, 915-929.
- Ooi, E. T. & Yang, Z. J. 2010, Efficient prediction of deterministic size effects using the scaled boundary finite element method, *Engineering Fracture Mechanics*, Vol. 77, 985-1000.
- Ooi, E. T. & Yang, Z. J. 2011a, Modelling crack propagation in reinforced concrete using a hybrid finite element-scaled boundary finite element method, *Engineering Fracture Mechanics*, Vol. 78, 252-273.
- Ooi, E. T. & Yang, Z. J. 2011b, Modelling dynamic crack propagation using the scaled boundary finite element method, *International Journal of Numerical Methods in Engineering*, Vol. 88(4), 329-349.
- Shi Mingguang. Fracture simulation and its engineering application based on Scaled Boundary Finite Element Method [D].BeiJing: Tsinghua University, 2013(in Chinese)
- Song, C. 2005, Evaluation of power-logarithmic singularities, T-stresses and higher order terms of in-plane singular stress fields at cracks and multimaterial corners, *Engineering Fracture Mechanics*, Vol. 72, 1498-1530.
- Song, C. & Vrcelj, Z. 2008, Evaluation of dynamic stress intensity factors and T-stress using the scaled boundary finite-element method, *Engineering Fracture Mechanics*, Vol. 75, 1960-1980.
- Song, C. & Wolf, J. P. 1997, The scaled boundary finite element method – alias consistent infinitesimal finite element cell method – for elasto dynamics, *Computer Methods in Applied Mechanics and Engineering*, Vol.147, 329-355.
- Song, C. & Wolf, J. P. 2002, Semi-analytical representation of stress singularity as occurring in cracks in anisotropic multi-materials with the scaled boundary finite-element method, *Computers & Structures*, Vol. 80, 183-197.
- Song, C., Tin-Loi, F. & Gao, W. 2010, A definition and evaluation procedure of generalized stress intensity factors at cracks and multi-material wedges, *Engineering Fracture Mechanics*, Vol. 77(12),2316-2336.
- Williams, M. L. 1957, “On the stress distribution at the base of a stationary crack”, *Journal of Applied Mechanics*, ASME, Vol. 24, 109-114.
- Wolf J P, Song C M. Finite-Element Modelling of Unbounded Media. New York: Wiley, 1996
- Yang, Z. J. 2006, Fully automatic modelling of mixed mode crack propagation using scaled boundary finite element method, *Engineering Fracture Mechanics*, Vol.73, 1711-1731.
- Yang, Z. J. & Deeks, A. J. 2007, Fully-automatic modelling of cohesive crack growth using a finite element-scaled boundary finite element coupled method, *Engineering Fracture Mechanics*, Vol. 74, 2547-2573.
- Zhu Chaolei. Static and Dynamic Fracture Simulation of Concrete Structures Based on Scaled Boundary Finite Element Method [D].Da Lian: Dalian University of Technology, 2014(in Chinese)

# Application of Least Squares Stress Stabilization Method on Nodal Integration of RPIM with Tetrahedral Background Cells

\*M.M. Yavuz<sup>1</sup> and †B. Kanber<sup>2</sup>

<sup>1</sup> *Energy Systems Engineering Department, Osmaniye Korkut Ata University, Osmaniye, TURKEY.*

<sup>2</sup> *Department of Mechanical Engineering, University of Gaziantep, Gaziantep, TURKEY.*

\*Presenting author: muratyavuz@osmaniye.edu.tr

†Corresponding author: kanber@gantep.edu.tr

## Abstract

The authors previously show that NI-RPIM solution of 3D tetrahedral cells [Yavuz and Kanber (2015)] includes unstable stress distributions. Therefore, in this study, the fluctuations in stresses are attempted to reduce using two different algorithms; Average stress distribution (ASD) in local support domain and LSS (least square stabilization) algorithms. NI-RPIM is improved with these algorithms to solve 3D solid mechanics problems with higher order Taylor series terms in nodal integration. Integration methodology is developed based on the study of [Liu et al. (2007)]. Tetrahedral integration cells are used in the solution of 3D elasto-static problems. Order of Taylor series terms is used up to 4<sup>th</sup> order for ASD. Also effects of shape parameters ( $\alpha_c$  and  $q$ ) in RBF and support domain size (sd) are searched. The results are compared with available analytical solutions and discussed in detail.

**Keywords:** Nodal Integration, RPIM, Tetrahedral Background Cells, Least-Squares Stabilization

## Introduction

Strain and stress formations are one of the critical design factors and always considered in machine design. Besides the working performance, strength and service life are also important in a machinery system and must be adjusted in a range of safety. A well designed mechanical system has enough strength against working loads. It been also avoided from the usage of excessive parts. For this reason, the best and optimum structural designs should be created.

It is needed to investigate the initial designs on stress analysis for optimizations. The critical locations must be determined and fixed. In general, the formation of stresses cannot be detected easily, especially for complex and non-uniform shapes. Geometry simplifications can be sometimes used with respect to analysed model. However, they include differences between analysed real model and simplified model, which causes errors. Thus, different numerical and experimental techniques have been developed for investigation with including less simplification, especially for complex shapes. Experimental techniques show responses of forces directly and reflect considerable results. However, investigation of each component or effect of parameter causes to repeat the experiment again and again.

Numerical methods can be another alternative way for investigation. On the contrary to requirements of experimental setup and apparatus in experimental investigation, numerical techniques do not need any experimental preparation, which can give fast and sensible results with respect to developed computer technology and used numerical methods. Even if it is mainly used approximate solutions, precise results can be achieved with a well-defined model. Different

numerical techniques have been developed. FDM (finite difference method), FEM (finite element method) and BEM (boundary element method) are widely used.

Unlike other methods, FEM has been very popular and used especially in the solutions of solid mechanics problems. The analysed model is divided to small elements, which are called finite elements and all solution procedure continues based on these elements. In most of the cases, formation of elements has poor shapes especially for complex geometries, which decreases accuracy of solutions. Construction of finite elements in smooth structure can consume most of analysing time. For this reason, meshfree techniques have been developed for decreasing numerical modelling time and effort. It is mainly used in early stages as SPH (smoothed particle hydrodynamics) [Gingold and Monaghan (1977); Lucy (1977)] on the solutions of astrophysical problems. DEM (diffuse element method) [Nayroles et al. (1992)] is developed at the further development stages. Mesh generation is not required, at least for interpolation. It includes MLS (moving least square) functions, which is used further development stage of EFG (element-free Galerkin) [Belytschko et al. (1994)] method. Partitions of unity [Babuska and Melenk (1996)] are developed by using MLS functions. Reproducing kernel particle method (RKPM) [Liu et al. (1996)], meshless local Petrov-Galerkin (MLPG) [Atluri and Zhu (1998)], point interpolation method (PIM) [Liu and Gu (2001)] and radial point interpolation method (RPIM) [Wang and Liu (2002a)] have been developed. Radial basis functions (RBF) are widely used in meshfree approximations, which are also used for development of BKM (boundary knot method) [Chen and Tanaka (2002)]. Also some of these methods are used in different integration methods. PIM and RPIM are further used with different methods and integration schemes; linearly conforming point interpolation method (LC-PIM) [Liu et al. (2005)], nodal integration radial point interpolation method (NI-RPIM) [Liu et al. (2007)], NS-PIM (node-based smoothed point interpolation method) [Zhang et al. (2007)], edge-based smoothed point interpolation method (ES-PIM) [Liu and Zhang (2008)] and CS-RPIM (cell-based smoothed radial point interpolation method) [Liu and Zhang (2009)] are available in literature.

Different techniques and methods can be combined for increasing the applicability of these methods. One of the integration schemes; nodal integration schemes can increase applicability and be effective for compact usage inside of RPIM, which is explained in detail in the study of Liu et al. (2007). This method uses Taylor series expansion in integration and the solution results are mainly affected with the used order of Taylor expansion. It is used for tetrahedral background cells [Yavuz and Kanber (2015)] on the solution of 3D elasto-static problems for increasing applicability. Each node in the model has its own tetrahedral integration cells. However it is observed that stress results include fluctuations and high values when compared with available analytical and FEM solutions. In literature, behaviour of nodal integration includes instabilities in some cases. Initially Biessel and Belytschko (1996) include additional a stabilization term on potential energy function of element-free Galerkin (EFG) method for nodal integration. It achieves instability problems of nodal integration in weak form formulations. Chen et al. (2001) also focus on instability of direct nodal integration and observe that integration constraints include errors in direct nodal integration and they cannot be satisfied enough in Gaussian integration at irregular discretization. They eliminate this problem with including a strain smoothing stabilization. Zhou et al. (2003) include square residual of equilibrium equation into potential energy function in stabilized nodal integration scheme. The used Voronoi diagram with supporting Delaunay triangles increases accuracy of volume assignment of nodal integration. Van et al. (2007) use conforming nodal integration into finite element formulation of laminate plates, which prevents shear locking. Han (2010) uses stabilized conforming nodal integration method on elasto-plastic analysis of metal forming process. Nodal integration with strain smoothing stabilization is used, which prevents instabilities of integration of Galerkin weak form formulations. Elmer et al. (2012) use a stable nodal integration

method on nearly incompressible solids. Xu (2014) uses stabilized nodal integration for cracking particles method (CPM), which supports stability and computational efficiency. Duan and Belytschko (2008) mention and compare 3 different stabilization techniques; which are least-square stabilization (LSS) [Biessel and Belytschko (1996); Fries and Belytschko (2007)], Taylor series expansion based stabilization (TEBS) [Liu et al. (1985); Nagashima (1999); Liu et al. (2007)] and the finite increment gradient stabilization (FIG) [Bonet and Kulasegaram (2000)], which can adjust stress instabilities.

In the previous study of Yavuz and Kanber (2015), it is observed that the nodal integration of radial point interpolation method (RPIM) based on Taylor series terms with tetrahedral background cells causes some fluctuations in stress results. Therefore, in this study, it is attempted to stabilize stresses using average stress distributions for local support domain of each node and least square stress stabilization method. RBF includes different shape parameters and their effects [Wang and Liu (2002b); Kanber et al. (2013); Bozkurt et al. (2013)] are also searched in some studies. They can affect solution results and suitable parameters must be searched before their usage. Effects of shape parameters ( $\alpha_c$  and  $q$ ) in RBF and support domain size (sd) are searched. The results are compared with available analytical solutions.

### RPIM Shape Functions

Construction of shape functions forms an important section for FEM and meshfree methods. High qualified shape functions gives better results and includes less errors. The shape functions are mainly occurred with interpolations and a suitable interpolation method must be used for increasing accuracy.

Relation between nodes in meshfree methods is mainly obtained with interpolations. Various interpolation methods are available in literature. RPIM [Wang and Liu (2002a)] is one of interpolation technique and widely used. It includes PIM [Liu et al. (2001)] with radial basis functions. When shape functions have been constructed, a field function  $u(x)$  is created and given as in Eq. (1), which is consist of polynomial and basis functions.

$$u(x) = \sum_{i=1}^n R_i(x) \times a_i + \sum_{j=1}^m P_j(x) \times b_j = R^T(x) \times a + P^T(x) \times b \quad (1)$$

$R_i(x)$  and  $P_j(x)$  represent radial basis and polynomial basis functions, respectively.  $a_i$  and  $b_j$  are related constants,  $n$  is the number of field nodes in the local support domain and  $m$  is the number of polynomial terms. Interpolations between nodes are mainly accomplished within the local support domain for each node or point of interests.

Various RBF [Liu and Gu (2005); Liu (2009)] are available in literature, like multi-quadrics (MQ), the Gaussian (Exp), the thin plate spline (TSP) and logarithmic radial basis functions. MQ is [Liu and Gu (2005); Dinis et al. (2007)] used as radial basis function in Eq. (2).

$$R_i(x) = \left( r_i^2 + (\alpha_c \times d_c)^2 \right)^q \quad (2)$$

where  $d_c$  is usually set as average nodal spacing near the point of interest at  $x$ ;  $\alpha_c$  and  $q$  are shape parameters. It is [Liu and Gu (2005); Liu (2009)] recommended to use  $q$  as 1.03 and  $\alpha_c$  as

3.00 for MQ basis function. The radial distance is given in Eq. (3) for 3D cases. Also the polynomial terms are given in Eq. (4) which are mainly derived from binomial expansion.

$$r_i(x) = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2} \quad (3)$$

$$p^T(x) = \{1, x, y, z, x^2, xy, y^2, yz, z^2, zx, \dots\} \quad (4)$$

Interpolation is applied in a support domain for point of interest. Different support domain geometries can be used, like circular, elliptical, triangular or rectangular. A circular local support domain is used and its covered area is given by radius of circle ( $d_s$ ), which is given in Eq. (5).

$$d_s = \alpha_s \times d_c \quad (5)$$

where  $d_c$  is average nodal spacing and  $\alpha_s$  is a positive real number of dimensionless size of the local support domain. Its value [Liu and Gu (2005); Dinis et al. (2007)] is commonly used between 2.00 and 3.00. The unknown constants of field function of  $a_i$  and  $b_j$  in Eq. (1) can be determined by enforcing the field function to pass through all  $n$  field nodes in the local support domain. At the  $k^{\text{th}}$  point or last point in a local support domain, field function can be written as:

$$u(x_k, y_k, z_k) = \sum_{i=1}^n R_i(x_k, y_k, z_k) \times a_i + \sum_{j=1}^m P_j(x_k, y_k, z_k) \times b_j \quad k=1,2,\dots,n \quad (6)$$

The matrix form of Eq. (6) can be expressed as

$$U_e = R_q \times a + P_m \times b = \{u_1 \ u_2 \ \dots \ u_n\}^T \quad (7)$$

where  $U_e$  is the vector of function values at the nodes in the local support domain.  $R_q$  is the moment matrix of RBF and  $P_m$  is the polynomial moment matrix [Liu and Gu (2005)], which are given in Eq. (8) and (9), respectively.

$$R_q = \begin{bmatrix} R_1(r_1) & R_2(r_1) & \dots & R_n(r_1) \\ R_1(r_2) & R_2(r_2) & \dots & R_n(r_2) \\ \dots & \dots & \dots & \dots \\ R_1(r_n) & R_2(r_n) & \dots & R_n(r_n) \end{bmatrix}_{(n \times n)} \quad (8)$$

$$P_m = \begin{bmatrix} 1 & x_1 & y_1 & z_1 & \dots & P_m(r_1) \\ 1 & x_2 & y_2 & z_2 & \dots & P_m(r_2) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_n & y_n & z_n & \dots & P_m(r_n) \end{bmatrix}_{(n \times m)} \quad (9)$$

$a$  and  $b$  are vectors of unknown coefficients for radial and polynomial basis functions respectively. They are given in Eq. (10) and (11).

$$a^T = \{a_1 a_2 \dots a_n\} \quad (10)$$

$$b^T = \{b_1 b_2 \dots b_n\} \quad (11)$$

For solution of field function, unknown parameter  $a$  in Eq. (7) must satisfy in polynomial function,

$$\sum_{j=1}^n p_j(x_i) \times a_i = P_m^T \times a = 0 \quad j=1,2,\dots,m \quad (12)$$

Combination of Eq. (7) and Eq. (1) yields the following equations in the matrix form:

$$\tilde{U}_e = \begin{bmatrix} U_e \\ 0 \end{bmatrix} = \begin{bmatrix} R_q & P_m \\ P_m^T & 0 \end{bmatrix} \begin{Bmatrix} a \\ b \end{Bmatrix} = G \times a_0 \quad j=1,2,\dots,m \quad (13)$$

where

$$\tilde{U}_e = \begin{bmatrix} U_e \\ 0 \end{bmatrix} = \{a_1 a_2 \dots a_n 0 0 \dots 0\}^T \quad (14)$$

Unique solution is obtained if inverse of matrix  $G$  exists:

$$a_0 = \begin{Bmatrix} a \\ b \end{Bmatrix} = G^{-1} \times \tilde{U}_e \quad (15)$$

Substituting Eq. (15) into Eq. (1), interpolation with respect to field function can be expressed as,

$$u(x) = \{R^T(x) \times P^T(x)\} \times G^{-1} \times \tilde{U}_e = \tilde{\varphi}(x) \times \tilde{U}_e \quad (16)$$

Finally [Liu and Gu (2005); Dinis et al. (2007)], RPIM shape functions for the corresponding  $n$  field nodes can be obtained as

$$\varphi^T(x) = \{\varphi_1(x) \varphi_2(x) \dots \varphi_n(x)\} \quad (17)$$

The approximation function can be written as

$$u(x) = \varphi^T(x) \times U_e = \sum_{i=1}^n \varphi_i \times u_i \quad (18)$$

The derivatives of  $u(x)$  can be easily obtained as

$$u_{i,k}(x) = \varphi_{i,k}^T(x) \times U_e \quad (19)$$

where  $k$  denotes the coordinates  $x$ ,  $y$  or  $z$ . Partial differentiation is taken with respect to that defined coordinated by  $k$ .

### 3D Nodal Integration with Taylor Series Expansion

The approximated solution must be adapted to equilibrium equation with respect to applied boundary conditions. General form of equilibrium equation [Liu and Gu (2005); Liu et al. (2007)] and natural and essential boundary conditions are given in Eq. (21), (22) and (23), respectively.

$$L^T \times \sigma + b = 0 \quad (20)$$

$$\sigma \cdot n = \bar{t} \text{ on } \tau_i \quad (21)$$

$$u = \bar{u} \text{ on } \tau_u \quad (22)$$

$L^T$  is [Liu and Gu (2005)] differential operator,  $\sigma$  is the stress vector,  $u$  is the displacement vector,  $b$  is the body force vector,  $\bar{t}$  is prescribed traction on the natural boundaries,  $\bar{u}$  is prescribed displacement on the essential boundaries and  $n$  is the vector of unit outward normal on the natural boundary.  $L^T$ ,  $\sigma$ ,  $u$  and  $b$  are given in Eq. (23), (24), (25) and (26) respectively.

$$L^T = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \\ 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{bmatrix} \quad (23)$$

$$\sigma^T = [\sigma_{xx} \quad \sigma_{yy} \quad \sigma_{zz} \quad \tau_{yz} \quad \tau_{xz} \quad \tau_{xy}] \quad (24)$$

$$u = \begin{bmatrix} u \\ v \\ w \end{bmatrix} \quad (25)$$

$$b = \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} \quad (26)$$

Equilibrium equation and natural and essential boundary conditions are represented with respect to weak form formulation in RPIM. Hence, the equilibrium equation, Eq. (20) can be defined as in Galerkin weak formulation in Eq. (27),

$$\int (L \times \delta u)^T \times (D \times L \times u) d\Omega - \int (\delta u^T \times b) d\Omega - \int (\delta u^T \times t) d\Gamma = 0 \quad (27)$$

$D$  matrix is material coefficient matrix and it is given for isotropic solids as;



$$D = \frac{E \times (1 - \nu)}{(1 + \nu) \times (1 - 2\nu)} \times \begin{bmatrix} 1 & \frac{\nu}{1 - \nu} & \frac{\nu}{1 - \nu} & 0 & 0 & 0 \\ \frac{\nu}{1 - \nu} & 1 & \frac{\nu}{1 - \nu} & 0 & 0 & 0 \\ \frac{\nu}{1 - \nu} & \frac{\nu}{1 - \nu} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1 - 2\nu}{2 \times (1 - \nu)} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1 - 2\nu}{2 \times (1 - \nu)} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1 - 2\nu}{2 \times (1 - \nu)} \end{bmatrix} \quad (28)$$

where  $E$  is Young's modulus and  $\nu$  is Poisson's ratio [Liu and Gu (2005)]. When substituting the approximated function in Eq. (18), into Eq. (27), general form of stiffness and force matrices are obtained.

$$K \times u = f \quad (29)$$

In Eq. (30), stiffness matrix formulation is represented. Force matrix formulation is given in Eq. (31). In addition, strain matrix, which includes derivatives of shape functions, is given in Eq. (32).

$$K_{ij} = \int B_i^T \times D \times B_j d\Omega \quad (30)$$

$$f_i = \int \varphi_i \times td\Gamma + \int \varphi_i \times bd\Omega \quad (31)$$

$$B_i = \begin{bmatrix} \varphi_{i,x} & 0 & 0 \\ 0 & \varphi_{i,y} & 0 \\ 0 & 0 & \varphi_{i,z} \\ 0 & \varphi_{i,z} & \varphi_{i,y} \\ \varphi_{i,z} & 0 & \varphi_{i,x} \\ \varphi_{i,y} & \varphi_{i,x} & 0 \end{bmatrix} \quad (32)$$

A suitable integration method is needed to solve these equations. Various integration techniques are available. In NI-RPIM [Liu et al. (2007)], a series integration scheme is used with Taylor series. Series are widely used in mathematical operations, especially in numerical studies. An unknown value of a valid function can be estimated with a known value with series operations. One of the series is Taylor series and it is widely used in computational fluid dynamics with respect to finite difference method (FDM). An example of value estimation from  $x_0$  to  $x_0 + h$  can be defined serial expansion of functions and it is given in Eq. (33).  $R_n$  is the total error between value of  $f(x_0 + h)$  and its Taylor expansion results. In general, the degree of used terms in FDM increases the accuracy. In Eq. (33), Taylor series are expanded and derived to  $n^{\text{th}}$  order.

$$f(x_0 + h) = f(x_0) + \frac{f'(x_0)}{1!} \times h + \frac{f''(x_0)}{2!} \times h^2 + \frac{f'''(x_0)}{3!} \times h^3 + \dots + \frac{f^n(x_0)}{n!} \times h^n + R_n \quad (33)$$

Equations of integrations are expanded with respect to Taylor series expansion as Eq. (33). Stiffness matrix is set as approximate function as  $f(x, y, z)$ , which is given in Eq. (34),

$$\begin{aligned} f(x, y, z) = & f(x_0, y_0, z_0) + (x - x_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial x} + (y - y_0) \times \\ & \frac{\partial f(x_0, y_0, z_0)}{\partial y} + (z - z_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial z} + \frac{1}{2!} \times (x - x_0)^2 \times \\ & \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x^2} + \frac{1}{2!} \times 2 \times (x - x_0) \times (y - y_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x \partial y} + \\ & \frac{1}{2!} \times (y - y_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y^2} + \frac{1}{2!} \times 2 \times (y - y_0) \times (z - z_0) \times \\ & \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y \partial z} + \frac{1}{2!} \times (z - z_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z^2} + \frac{1}{2!} \times 2 \times \\ & (z - z_0) \times (x - x_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z \partial x} + \frac{1}{3!} \times (x - x_0)^3 \times \\ & \frac{\partial^3 f(x_0, y_0, z_0)}{\partial x^3} + \frac{1}{3!} \times (y - y_0)^3 \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial y^3} + \frac{1}{3!} \times (z - z_0)^3 \times \\ & \frac{\partial^3 f(x_0, y_0, z_0)}{\partial z^3} + \frac{1}{3!} \times 3 \times (x - x_0)^2 \times (y - y_0) \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial x^2 \partial y} + \dots \end{aligned} \quad (34)$$

The nodal integration of Eq. (34) can be written as Eq. (35),

$$\begin{aligned} \int f(x, y, z) d\Omega = & \int \left( f(x_0, y_0, z_0) + (x - x_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial x} + (y - y_0) \times \right. \\ & \frac{\partial f(x_0, y_0, z_0)}{\partial y} + (z - z_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial z} + \frac{1}{2!} \times (x - x_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x^2} + \\ & \frac{1}{2!} \times 2 \times (x - x_0) \times (y - y_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x \partial y} + \frac{1}{2!} \times (y - y_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y^2} + \\ & \frac{1}{2!} \times 2 \times (y - y_0) \times (z - z_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y \partial z} + \frac{1}{2!} \times (z - z_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z^2} + \\ & \frac{1}{2!} \times 2 \times (z - z_0) \times (x - x_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z \partial x} + \frac{1}{3!} \times (x - x_0)^3 \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial x^3} + \\ & \frac{1}{3!} \times (y - y_0)^3 \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial y^3} + \frac{1}{3!} \times (z - z_0)^3 \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial z^3} + \frac{1}{3!} \times 3 \times \\ & \left. (x - x_0)^2 \times (y - y_0) \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial x^2 \partial y} + \dots \right) d\Omega \end{aligned} \quad (35)$$

If Eq. (35) is separated and arranged, the following form can be obtained,

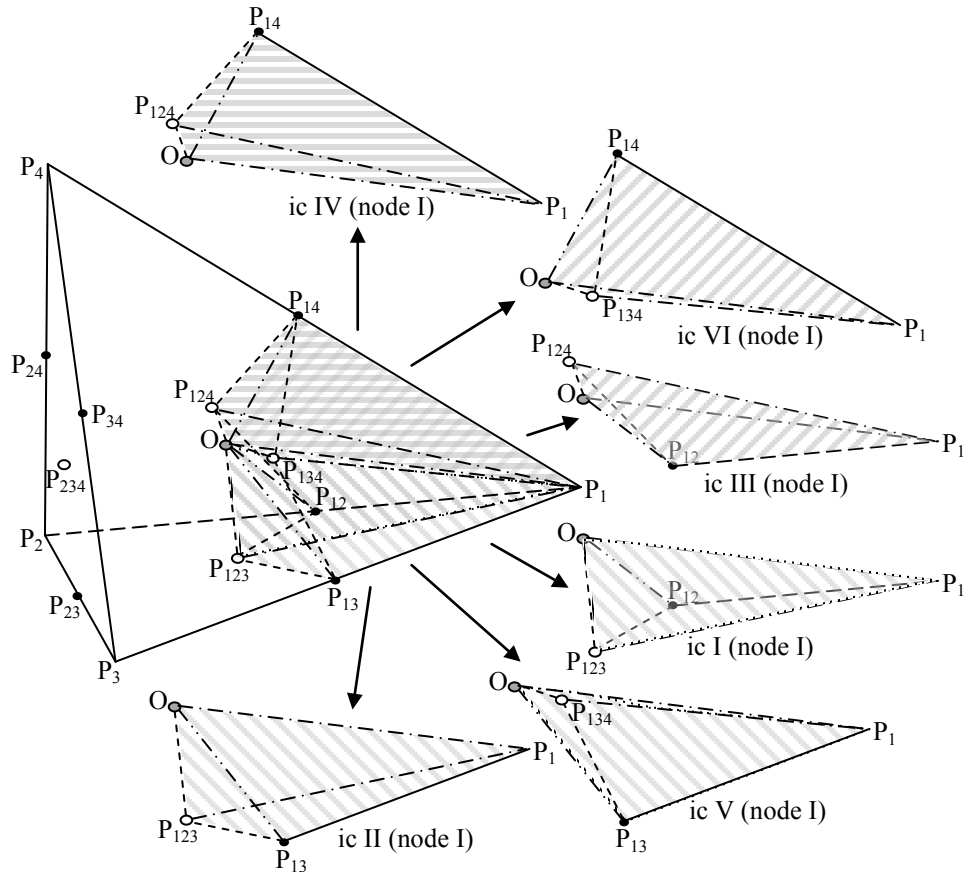
$$\begin{aligned}
 \int f(x, y, z)d\Omega = & \int f(x_0, y_0, z_0)d\Omega + \int (x - x_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial x} d\Omega + \int (y - y_0) \times \\
 & \frac{\partial f(x_0, y_0, z_0)}{\partial y} d\Omega + \int (z - z_0) \times \frac{\partial f(x_0, y_0, z_0)}{\partial z} d\Omega + \int \frac{1}{2!} \times (x - x_0)^2 \times \\
 & \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x^2} d\Omega + \int \frac{1}{2!} \times 2 \times (x - x_0) \times (y - y_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial x \partial y} d\Omega + \\
 & \int \frac{1}{2!} \times (y - y_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y^2} d\Omega + \int \frac{1}{2!} \times 2 \times (y - y_0) \times (z - z_0) \times \\
 & \frac{\partial^2 f(x_0, y_0, z_0)}{\partial y \partial z} d\Omega + \int \frac{1}{2!} \times (z - z_0)^2 \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z^2} d\Omega + \int \frac{1}{2!} \times 2 \times \\
 & (z - z_0) \times (x - x_0) \times \frac{\partial^2 f(x_0, y_0, z_0)}{\partial z \partial x} d\Omega + \int \frac{1}{3!} \times (x - x_0)^3 \times \frac{\partial^3 f(x_0, y_0, z_0)}{\partial x^3} \\
 & d\Omega + \dots
 \end{aligned} \tag{36}$$

$$\begin{aligned}
 \int f(x, y, z)d\Omega = & f(x_0, y_0, z_0) \times \int 1 d\Omega + f_x(x_0, y_0, z_0) \times \int (x - x_0) d\Omega + \\
 & f_y(x_0, y_0, z_0) \times \int (y - y_0) d\Omega + f_z(x_0, y_0, z_0) \times \int (z - z_0) d\Omega + \\
 & f_{xx}(x_0, y_0, z_0) \times \frac{1}{2!} \times \int (x - x_0)^2 d\Omega + f_{xy}(x_0, y_0, z_0) \times \\
 & \int (x - x_0) \times (y - y_0) d\Omega + \dots
 \end{aligned} \tag{37}$$

In Eq. (37),  $x - x_0$  represents  $\Delta x$ ,  $y - y_0$  represents  $\Delta y$  and  $z - z_0$  represents  $\Delta z$ . The distance is calculated from midpoint of related field node and integration cell, which is called Taylor integration cells.  $d\Omega$  is equal to volume of Taylor integration cell. Each Taylor integration cell for each field node includes a volume that does not interact with other field nodes of Taylor integration cells.

### Integration with Tetrahedral Shaped Taylor Integration Cells

It is required in NI-RPIM [Liu et al. (2007)] to construction of integration domains for each node in the analysed model. These domains should not coincide with each other and must be fully integrated. Taylor integration domains are constructed with irregular tetrahedral geometries for irregular distributed nodes. These irregular domain construction supports to more flexible domains, rather than hexahedral cells. Delaunay triangulation method is used for creating tetrahedral cells between nodes by using MATLAB.



**Figure 1. Sub-division of tetrahedral background cell for each node.**

In Fig. 1, a tetrahedral background cell is given. The cell which is formed by field nodes of  $P_1 (x_1, y_1, z_1)$ ,  $P_2 (x_2, y_2, z_2)$ ,  $P_3 (x_3, y_3, z_3)$  and  $P_4 (x_4, y_4, z_4)$  is further divided into 6 different integration cells (ic I, II...). So, the total number of integration cells for a field node can be calculated by multiplying 6 by the number of tetrahedral cells that are connected to the field node.  $P_{12}$ ,  $P_{13}$ ,  $P_{14}$ ,  $P_{23}$ ,  $P_{24}$  and  $P_{34}$  are centre point of each side line of tetrahedral cell.  $P_{123}$ ,  $P_{124}$ ,  $P_{134}$  and  $P_{234}$  are centre points of surface areas. O is the centre point.

**Table 1. Components of edges for subdivided Taylor Integration cells for each node.**

	integration cell (ic)	components of edges					integration cell (ic)	components of edges			
Node I (P1)	I	$P_1$	$P_{12}$	$P_{123}$	O	Node III (P3)	I	$P_3$	$P_{13}$	$P_{123}$	O
	II	$P_1$	$P_{123}$	$P_{13}$	O		II	$P_3$	$P_{123}$	$P_{23}$	O
	III	$P_1$	$P_{124}$	$P_{12}$	O		III	$P_3$	$P_{134}$	$P_{13}$	O
	IV	$P_1$	$P_{14}$	$P_{124}$	O		IV	$P_3$	$P_{34}$	$P_{134}$	O
	V	$P_1$	$P_{13}$	$P_{134}$	O		V	$P_3$	$P_{23}$	$P_{234}$	O
	VI	$P_1$	$P_{134}$	$P_{14}$	O		VI	$P_3$	$P_{234}$	$P_{34}$	O
Node II (P2)	I	$P_2$	$P_{123}$	$P_{12}$	O	Node IV (P4)	I	$P_4$	$P_{124}$	$P_{14}$	O
	II	$P_2$	$P_{23}$	$P_{123}$	O		II	$P_4$	$P_{24}$	$P_{124}$	O
	III	$P_2$	$P_{12}$	$P_{124}$	O		III	$P_4$	$P_{14}$	$P_{134}$	O
	IV	$P_2$	$P_{124}$	$P_{24}$	O		IV	$P_4$	$P_{134}$	$P_{34}$	O
	V	$P_2$	$P_{234}$	$P_{23}$	O		V	$P_4$	$P_{234}$	$P_{24}$	O
	VI	$P_2$	$P_{24}$	$P_{234}$	O		VI	$P_4$	$P_{34}$	$P_{234}$	O

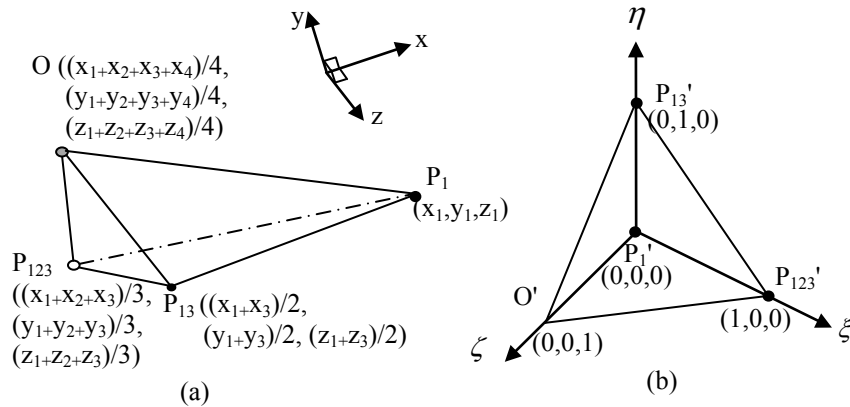
Subdivision of tetrahedral integration cell components for each node is tabulated in Table 1. Subdivision of tetrahedral cell can cause negative volumes [Toron (2004); Kovalev (2005)]. Therefore, the orientation of subdivided cell nodes is placed with an order of preventing negative volume results.

In Eq. (37), the integrations  $\int 1d\Omega$ ,  $\int(\Delta x)d\Omega$ ,  $\int(\Delta y)d\Omega$ ,  $\int(\Delta z)d\Omega$ ,  $\int(\Delta x^2)d\Omega \dots$  must be calculated for tetrahedral background cells. The first integration represents the volume of tetrahedron and can also be calculated as determinant of edge distances [Bhowmick and Shontz (2012)] in Eq. (38).

$$Volume = \frac{1}{6} \cdot \begin{vmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_3 - y_1 & y_4 - y_1 \\ z_2 - z_1 & z_3 - z_1 & z_4 - z_1 \end{vmatrix} \quad (38)$$

However, integration of volume in terms of  $x, y, z, x^2, \dots$  includes complex operations for irregular tetrahedral shapes. Hence, geometry can be transformed into natural coordinates by using Jacobian transformation. In Fig. 2, the transformation of tetrahedron geometry from global to natural coordinates is shown for ic II of field node I.  $P_1', P_{123}', P_{13}'$  and  $O'$  represent the transformed natural coordinates from global coordinates of  $P_1, P_{123}, P_{13}$  and  $O$  points, respectively. Also for other integration cells,  $P_1', P_{123}', P_{13}'$  and  $O'$  in Fig. 2 correspond to second, third, fourth and fifth columns in Table 1 respectively.

The bounds of integral starts from zero to upper natural coordinate. Hence  $P_1, P_2, P_3$  and  $P_4$  are placed at  $P_1'$  in integral calculations. This condition can provide to no usage of parallel axis theorem to carry integration results to related edge node. Hence,  $x_1, y_1$  and  $z_1$  in Eq. (39), (40) and (41) are assigned  $x, y$  and  $z$  coordinates of  $P_1, P_2, P_3$  and  $P_4$  in related integral node calculation.



**Figure 2. Transformation of an integration cell (ic II in Fig.1) from global (a) to natural (b) coordinates.**

The transformation of coordinates is given in Eq. (39), (40) and (41) [Bhowmick and Shontz (2012)].  $\eta, \xi$  and  $\zeta$  are natural coordinates.

$$x = x_1 + (x_2 - x_1) \cdot \xi + (x_3 - x_1) \cdot \eta + (x_4 - x_1) \cdot \zeta \quad (39)$$

$$y = y_1 + (y_2 - y_1) \cdot \xi + (y_3 - y_1) \cdot \eta + (y_4 - y_1) \cdot \zeta \quad (40)$$

$$z = z_1 + (z_2 - z_1) \cdot \xi + (z_3 - z_1) \cdot \eta + (z_4 - z_1) \cdot \zeta \quad (41)$$

Determinant of Jacobian (J) transformation matrix are given in Eq. (42).

$$DET(J) = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{vmatrix} = 6 \cdot Volume \quad (42)$$

The transformation of integration of uniform geometry in natural coordinates is given in Eq. (43). Its application to tetrahedron geometry is given in Eq. (44). Hence integration in terms of  $x, y, z, x^2, \dots$  can be easily calculated as follows:

$$\int f(x, y, z) d\Omega = \int f[x(\xi, \eta, \zeta), y(\xi, \eta, \zeta), z(\xi, \eta, \zeta)] * |DET(J)| d\Omega' \quad (43)$$

$$\int_D f(x, y, z) dD = |DET(J)| * \int_0^1 d\xi \int_0^{1-\xi} d\eta \int_0^{1-\xi-\eta} f[x(\xi, \eta, \zeta), y(\xi, \eta, \zeta), z(\xi, \eta, \zeta)] d\zeta \quad (44)$$

Even if suitable orientation is used in Table 1, in some cases node coordinate orientations can be changed with respect to construction of tetrahedral cells in Delaunay triangulation technique. Its symptoms can be determined with negative value of determinant of Jacobian matrix. For this reason, upper integration bounds in Eq. (44) are updated with respect to where the subdivided tetrahedral cell exists, which are given in Table 2.

**Table 2. Direction of tetrahedral cell and used bounds of integral in natural coordinates.**

	$\Delta x_{12}$	$\Delta y_{13}$	$\Delta z_{14}$	$\int_0^1 f[x, y, z] d\zeta$	$\int_0^1 f[x, y, z] d\eta$	$\int_0^1 f[x, y, z] d\xi$
I	+	+	+	$1 - \xi - \eta$	$1 - \xi$	+1
II	+	+	-	$-(1 - \xi - \eta)$	$1 - \xi$	+1
III	+	-	+	$1 - \xi + \eta$	$-(1 - \xi)$	+1
IV	+	-	-	$-(1 - \xi + \eta)$	$-(1 - \xi)$	+1
V	-	+	+	$1 + \xi - \eta$	$1 + \xi$	-1
VI	-	+	-	$-(1 + \xi - \eta)$	$1 + \xi$	-1
VII	-	-	+	$1 + \xi + \eta$	$-(1 + \xi)$	-1
VIII	-	-	-	$-(1 + \xi + \eta)$	$-(1 + \xi)$	-1

Where  $\Delta x_{12}$  is the sign of difference between  $x_2 - x_1$ ,  $\Delta y_{13}$  is the sign of difference between  $y_3 - y_1$  and  $\Delta z_{14}$  is the sign of difference between  $z_4 - z_1$  for each sub-divided integration cell. The signs of  $\eta$ ,  $\xi$  and  $\zeta$  in Eq. (39), (40) and (41) is also regulated with respect to signs of  $\Delta x_{12}$ ,  $\Delta y_{13}$  and  $\Delta z_{14}$ .

These calculations are applied for each tetrahedral cell and the results are summed for the corresponding field node. Hence, a node includes more than one tetrahedral integration cells, which looks like a polyhedron geometry. However, all main structure of integrations is carried on each subdivided tetrahedral cells for related nodes.

### Application of Stress Stabilizations

The application of nodal integration is fast and suitable for complex geometries. However, in the previous study of Yavuz and Kanber (2015), the stress results of pure/unstabilized nodal integration include high fluctuations. Even if two close nodes are selected, their stress results can be highly different. Hence application of different methods is investigated for trying to decrease the stress fluctuations.

#### *Average Stress Distribution for Each Local Support Domain*

In the previous study [Yavuz and Kanber (2015)], it is observed in nodal integration that high stress fluctuations occur at far away from application locations of force and boundary conditions. It is not expected that formation of these high fluctuations occur at that conditions with respect to Saint Venant's principle.

For decreasing the fluctuations, average stress distributions are taken for each local support domain. The stresses of nodes in a local support domain of a related node is summed and averaged for number of nodes in that support domain.

#### *Stiffness Effect on Boundary Locations*

In the previous study [Yavuz and Kanber (2015)], some stress peak points are observed, especially on the application location of boundary conditions. Hence local stiffness matrix of nodes on boundary condition locations are changed for decreasing stress peak points. Its effects are investigated.

#### *Least Square Stabilization*

In order to decrease fluctuations, least square stabilization method (LSS) [Biessel and Belytschko (1996), Fries and Belytschko (2007), Duan and Belytschko (2008)] is used in nodal integration scheme. This method adds two equations ( $K^\gamma$  and  $f^\gamma$ ) in Galerkin weak form of governing equation for providing stress stabilization. The main governing equation is previously given in Eq. 20. Addition of stress stabilization equations and detail transformation operations about governing equations can be reached from the study of Duan and Belytschko (2008). The simplified equation is given in Eq. 45.  $K^\gamma$  and  $f^\gamma$  equations are given in Eq. 46 and 47.

$$(K + K^\gamma + \beta K^p)d = (f + f^\gamma + \beta f^p) \quad (45)$$

$$K^\gamma = \int_{\Omega} \gamma (L^T DLN)^T (L^T DLN) D\Omega \quad (46)$$

$$f^\gamma = -\int_{\Omega} \gamma (L^T DLN)^T b D\Omega \quad (47)$$

$\gamma$  is the stabilization parameter and it is given in Eq. 48.

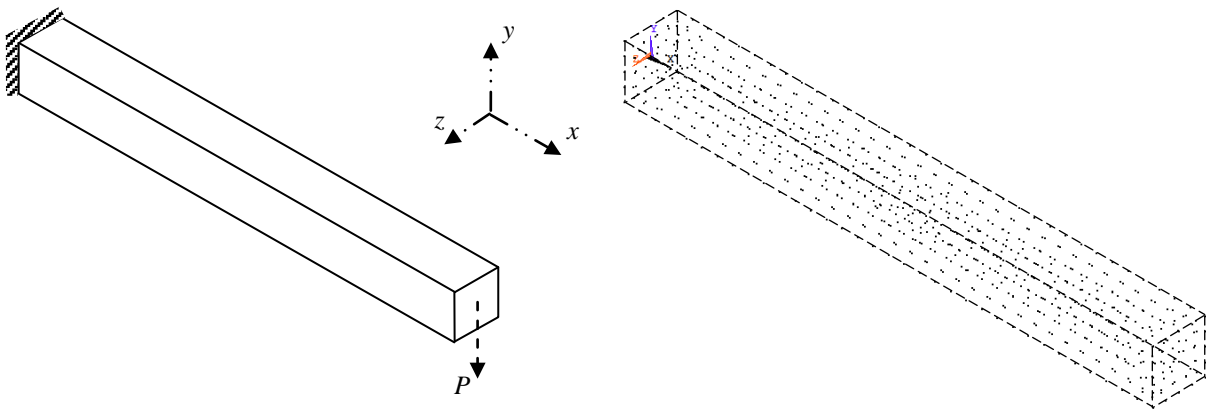
$$\gamma = \frac{2\alpha_s l_c^2}{E} \quad (48)$$

$\alpha_s$  is the dimensionless stabilization parameter and used as 0.3 in this study.  $l_c$  is used as nodal spacing, which is nearly equal to  $0.04 \cdot L/100$ .

### Solution and Discussion

A cantilever beam problem is examined for investigation of stress stabilization methods. The nodal integration of RPIM is mainly prepared with respect to the study of Liu. et al. (2007). The solutions of formulations, creation of tetrahedral integration cells and subdivision calculations are accomplished with MATLAB.

The geometry and meshfree model is given in Fig. 3. The used beam has a length of 1.0 m with square section of  $0.1\text{m} \times 0.1\text{m}$ . The used meshfree model has 878 irregularly distributed nodes.



**Figure 3. The used cantilever beam model geometry and meshfree model**

The used material properties have linear elastic behaviour with a Young's Modulus of 200 GPa and Poisson's ratio of 0.0. The selection of Poisson's ratio as 0.0 is aimed for providing similar conditions as analytical solutions. Applied force ( $P$ ) is used as 29000 N and other side of force application location is determined as application of boundary condition location. This location is fixed.

Vertical deflection equation [Beer et al. (2009)] of cantilever beam is given in Eq. 49. Where  $P$  is applied force,  $x$  is distance from fixed support location,  $E$  is Young's Modulus and  $I$  represents inertia of beam.

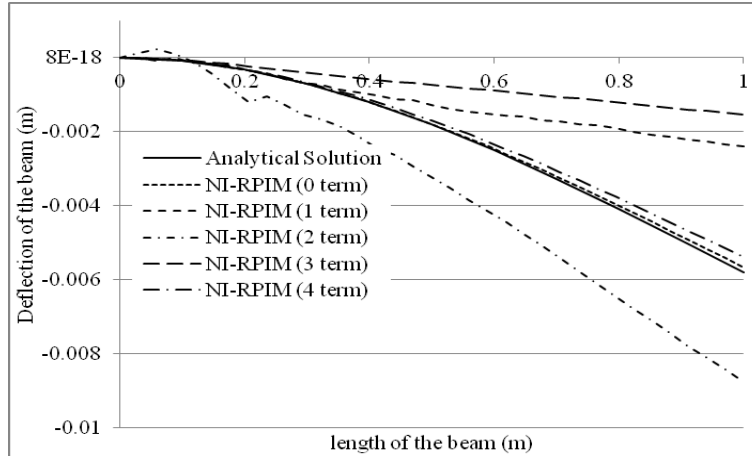
$$y = \frac{P \times x^2}{6 \times E \times I} (3 \times L - x) \quad (49)$$

The bending stress at upper and lower surfaces of beam is given as;



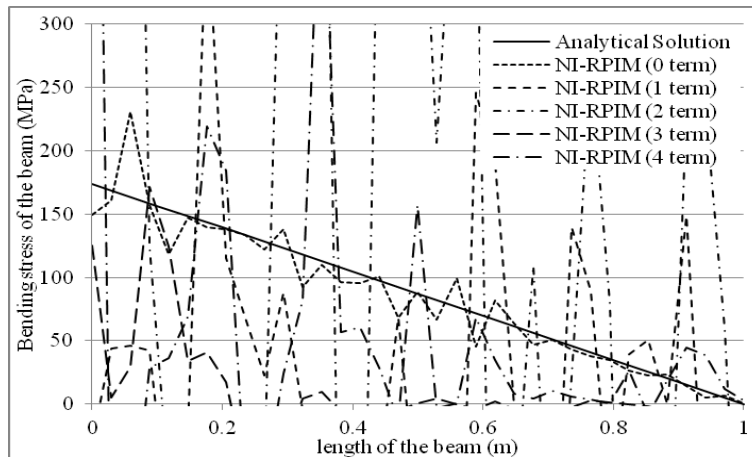
$$\sigma = \frac{M \times c}{I} \tag{50}$$

$c$  is the distance between upper/lower surface and natural axis of the beam. In meshfree solutions,  $\alpha_c$  is used as 3.00, dimensionless support domain size ( $\alpha_s$ ) parameter is used as 1.30 and  $q$  is used as 1.03 as default parameters.



**Figure 4. Comparison of deflection results of cantilever beam for analytical and pure nodal int. technique.**

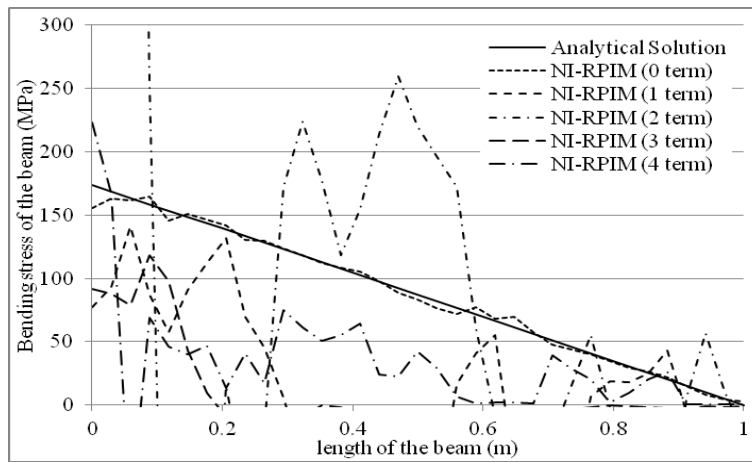
Deflection results of cantilever beam is shown in Fig. 4 for NI-RPIM with 0<sup>th</sup>, 1<sup>st</sup>, 2<sup>rd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> order Taylor terms. It is observed that 0<sup>th</sup> and 4<sup>th</sup> order terms give the best results when they are compared with analytical solution. There is no fluctuation occurs on deflection results of these order terms. However the used other terms includes less accuracy and small distortions at the results.



**Figure 5. Comparison of bending stress results of cantilever beam for analytical and pure/unstabilized nodal int. technique.**

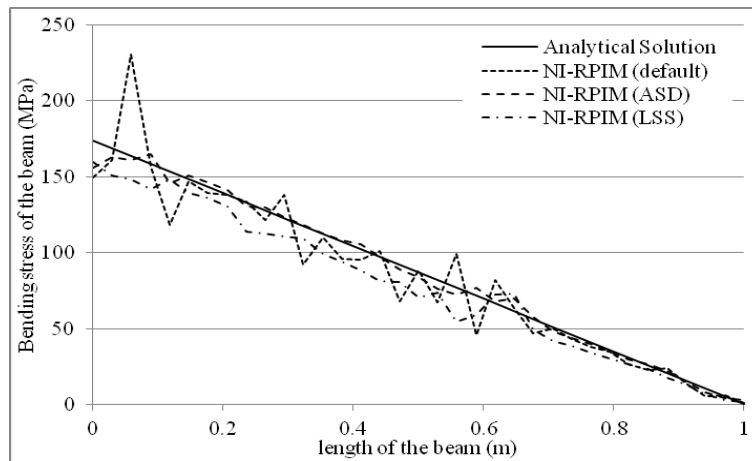
The bending stress results are given in Fig. 5 for pure/unstabilized nodal integration, which include highly fluctuation results. The best results achieved with 0<sup>th</sup> order terms, when results of all terms

are compared with each other. The stress results of 1<sup>st</sup> and 4<sup>th</sup> order terms include high stress values at the application locations of boundary conditions.



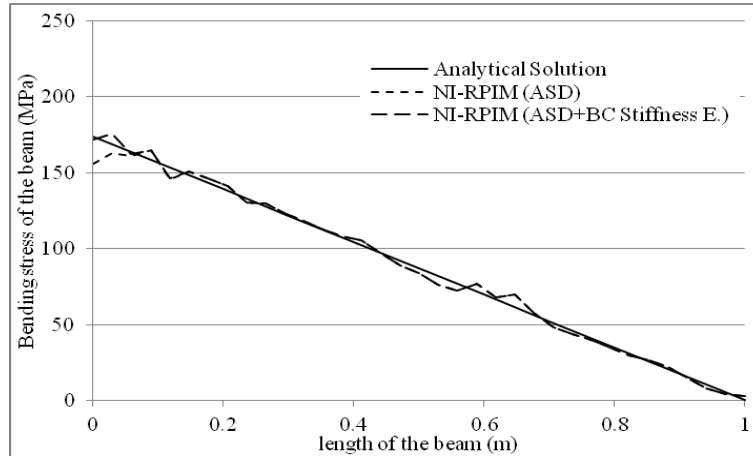
**Figure 6. Comparison of bending stress results of cantilever beam for analytical and nodal int. technique with ASD.**

The formation of stress results are given in Fig. 6, when averaged stress distribution (ASD) of local support domain method is applied for each node. It is observed that fluctuations of stress results are decreased. The best accuracy is achieved with the usage of 0<sup>th</sup> order term. However, results of higher terms without 0<sup>th</sup> order term have poor accuracy and include fluctuations.



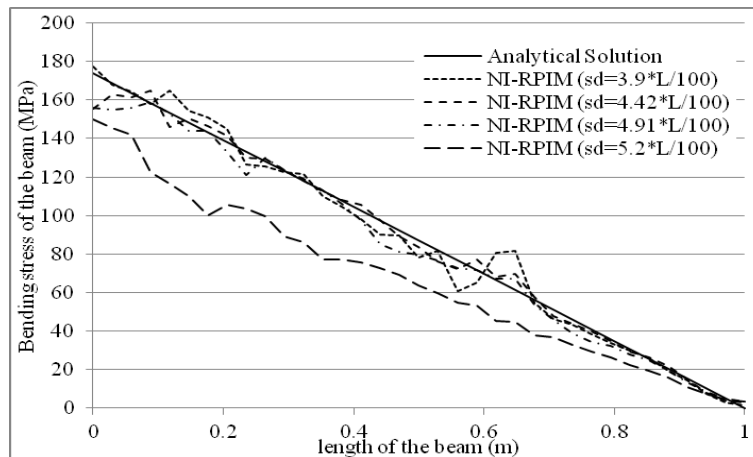
**Figure 7. Comparison of bending stress results of cantilever beam for analytical and nodal int. technique of 0<sup>th</sup> order term with ASD and LSS.**

In Fig.7, only the usage of 0<sup>th</sup> order term results of NI-RPIM are compared with unstabilized, ASD and LSS methods. It is observed that the usage of both stress stabilization methods decreases fluctuations in the results.



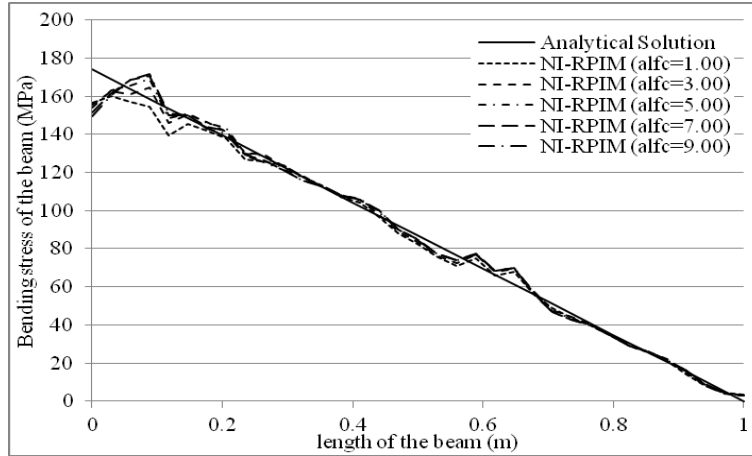
**Figure 8. Comparison of bending stress results of cantilever beam for analytical and nodal int. technique of 0<sup>th</sup> order term with ASD/ASD+BC Stiffness E.**

NI-RPIM results of 0<sup>th</sup> order term have a difference from analytical solution at the application location of boundary conditions (BC). NI-RPIM results have less stress values than analytical solution at BC. Hence, the stiffness values are decreased about %30 at this location for achieving same results at BC.



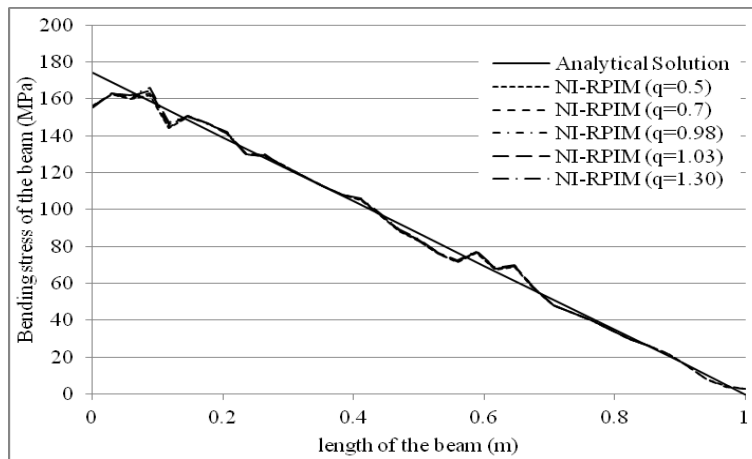
**Figure 9. Comparison of effect of sd on bending stress results of cantilever beam for nodal int. with ASD (0<sup>th</sup> order term).**

Effect of support domain (sd) size is shown for 0<sup>th</sup> order term results of NI-RPIM with ASD. When sd value of 3.9\*L/100 is used, there is high fluctuation occurs about at a beam length of 0.6 m. sd size of 4.42\*L/100 gives better stress results, which is used default in other solutions in this study. The average number of nodes in the local support domains is approximately equal to 44. When sd is increased, the accuracy begins decreasing.



**Figure 10. Comparison of effect of  $\alpha_c$  on bending stress results of cantilever beam for nodal int. with ASD (0<sup>th</sup> order term).**

Effect of  $\alpha_c$  results are given in Fig. 10. The only dominant effect is shown at BC and its near locations. Different  $\alpha_c$  results have similar characteristics.



**Figure 11. Comparison of effect of q on bending stress results of cantilever beam for nodal int. with ASD (0<sup>th</sup> order term).**

Effect of q results is given in Fig. 11. Nearly all the results of different q values have same values.

### Conclusions

Tetrahedral background cells are used for the nodal integration of RPIM. The fluctuation problem in stress results of nodal integration is investigated with application of different methods. Averaged stress distribution (ASD) on local support domain and least square stabilization (LSS) methods are used. Effect of orders of Taylor terms, support domain size and RBF terms are investigated at NI-RPIM with ASD.

Results can be summarized as follows:

- The usage of various orders of Taylor terms directly affects deflection and stress results.

- There is no fluctuation on deflection results of 0<sup>th</sup> and 4<sup>th</sup> order terms, which gives good accuracy with respect to analytical solution.
- Less fluctuated stress values are obtained with the usage of 0<sup>th</sup> order Taylor series terms in NI-RPIM when no stabilization method is used.
- Fluctuations on stress results are decreased with the usage of ASD and LSS methods.
- 0<sup>th</sup> order term results of NI-RPIM with ASD gives the best stress results and less fluctuation values in the solutions.
- High stress fluctuations on boundaries where essential boundary conditions are applied can be decreased with changing local stiffness values of corresponding nodes.
- Support domain size (sd) can affect stress results of NI-RPIM with ASD method.
- Changes of  $\alpha_c$  with ASD method do not have significant effect on NI-RPIM stress results in the regions where the essential boundary conditions are not included. Different values of q show similar stress distribution characteristics on NI-RPIM with ASD method.

## References

- Atluri, S. N. and Zhu, T. (1998) A new meshless local Petrov-Galerkin (MLPG) approach in computational mechanics, *Computational Mechanics* **22**, 117-127.
- Babuska, I. and Melenk, J. (1996) The partition of the unity finite element method: basic theory and applications, *Comput. Methods Appl. Mech. Engrg.* **139**, 289-314.
- Belytschko, T., Lu, Y. Y. and Gu, L. (1994) Element-free Galerkin methods, *Int J Numer Methods Eng* **37**, 229-256.
- Beer, F. P., Johnston-Jr, E. R., Dewolf, J. T. and Mazurek, D. (2009). *Mechanics of Materials*. (5<sup>th</sup> edn.) McGraw-Hill Press, Singapore.
- Bhowmick, S. and Shontz, S. M. (2012) Towards high-quality, untangled meshes via a force-directed graph embedding approach. *Procedia Computer Science*, **1**, 357–366.
- Biessel, S. and Belytschko, Y. (1996) Nodal integration of the element-free Galerkin method, *Comput. Methods Appl. Mech. Engrg.* **139**, 49-74.
- Bonet, J. and Kulasegaram, S. (2000) Finite increment gradient stabilization of point integrated meshless methods for elliptic equations, *Communications in Numerical Methods in Engineering* **16**, 475–483.
- Bozkurt, O. Y., Kanber, B. and Asik, M. Z. (2013) Assessment of RPIM shape parameters for solution accuracy of 2D geometrically nonlinear problems, *International Journal of Computational Methods* **10(3)**, 1-26.
- Chen, J., Wu, C., Yoon, S. and You, Y. (2001) A stabilized conforming nodal integration for Galerkin mesh-free methods, *International Journal for Numerical Methods in Engineering* **50**, 435-466.
- Chen, W. and Tanaka, M. (2002) Meshless, integration-free, and boundary-only rbf technique, *Computers and Mathematics with Applications* **43**, 379-391.
- Duan, Q. and Belytschko, T. (2008) On the stabilization of stress-point integration in the Element Free Galerkin method, *Meshfree Methods for Partial Differential Equations IV Lecture Notes in Computational Science and Engineering* **65**, 47-68.
- Elmer, W., Chen, J. S., Puso, M. and Tacioglu, E. (2012) A stable meshfree nodal integration method for nearly incompressible solids, *Finite Elements in Analysis and Design* **51**, 81–85.
- Fries, T. and Belytschko, T. (2007) Convergence and stabilization of stress-point integration in mesh-free and particle methods, *International Journal for Numerical Methods in Engineering* **74(7)**, 1067-1087.
- 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.
- Han, K. (2010) Efficient meshfree analysis Using stabilized conforming nodal integration for metal forming simulation, *Journal of the Korean Society of Marine Engineering* **34(7)**, 943-950.
- Kanber, B., Bozkurt, O. Y. and Erklig, A. (2013) Investigation of RPIM shape parameter effects on the solution accuracy of 2D elastoplastic problems, *International Journal for Computational Methods in Engineering Science and Mechanics* **14**, 354–366.
- Kovalev, K. (2005) *Unstructured hexahedral non-conformal mesh generation*. PhD thesis, Faculty of Engineering Vrije Universiteit Brussel, Belgium.
- Liu, G. R. and Gu, Y. T. (2001) A point interpolation method for two-dimensional solids, *International Journal for Numerical Methods in Engineering* **50**, 937-951.
- Liu, G. R. and Zhang, G. Y. (2008) Edge-based smoothed point interpolation methods, *International Journal of Computational Methods* **5(4)**, 621-646.
- Liu, G. R. and Zhang, G. Y. (2009) A normed G space and weakened weak (W2) formulation of cell-based smoothed point interpolation method, *International Journal of Computational Methods* **6(1)**, 147-179.

- Liu, G. R., Zhang, G. Y., Dai, K. Y., Wang, Y. Y., Zhong, Z. H., Li, G. Y. and Han, X. (2005) A linearly conforming point interpolation method (LC-PIM) for 2D solid mechanics problems, *International Journal of Computational Methods* **2(4)**, 645–665.
- Liu, G. R., Zhang, G. Y., Wang, Y. Y., Zhong, Z. H., Li, G. Y. and Han, X. (2007) A nodal integration technique for meshfree radial point interpolation method (NI-RPIM), *International Journal of Solids and Structures* **44**, 3840-3860.
- Liu, W. K., Chen, Y., Chang, C. T. and Belytschko, T. (1996) Advances in multiple scale kernel particle methods, *Comput Mech* **18**, 73-111.
- Liu, W. K., Ong, J. S. and Uras, R. A. (1985) Finite element stabilization matrices - a unification approach, *Computer Methods in Applied Mechanics and Engineering* **53**, 13–46.
- Lucy, L. B. (1977) A numerical approach to the testing of the fission hypothesis, *The Astronomical Journal* **82(12)**, 1013-1024.
- Nagashima, T. (1999) Node-by-node meshless approach and its applications to structural analysis, *Int. J. Numer. Meth. Engrg.* **46**, 341–385.
- Nayroles, B., Touzot, G. and Villon, P. (1992) Generalizing the finite element method: diffuse approximation and diffuse elements, *Comput Mech* **10**, 307-318.
- Nguyen-Van, H., Duy, N. M. and Tran-Cong, T. (2007) A simple and accurate four-node quadrilateral element using stabilized nodal integration for laminated plates, *CMC* **6(3)**, 159-175.
- Tonon, F. (2004). Explicit exact formulas for the 3-D tetrahedron inertia tensor in terms of its vertex coordinates. *Journal of Mathematics and Statistics*, **1(1)**, 8-11.
- Wang, J. G. and Liu, G. R. (2002a) A point interpolation meshless method based on radial basis functions, *International Journal for Numerical Methods in Engineering* **54**, 1623-1648.
- Wang, J. G. and Liu, G. R. (2002b) On the optimal shape parameters of radial basis functions used for 2-D meshless methods, *Comput Methods Appl Mech Engrg* **191**, 2611–2630.
- Xu, S. (2014) Stable cracking particles method based on stabilized nodal integration and updated Lagrangian kernel, *Mathematical Problems in Engineering* **2014**, 1-10.
- Yavuz, M. M. and Kanber, B. (2015) On the usage of tetrahedral background cells in nodal integration of RPIM for 3D elasto-static problems, *International Journal of Computational Methods* (Accepted for publication).
- Zhang, G. Y., Liu, G. R., Wang, Y. Y., Huang, H. T., Zhong, Z. H., Li, G. Y. and Han, X. (2007) A linearly conforming point interpolation method (LC-PIM) for three dimensional elasticity problems, *International Journal for Numerical Methods in Engineering* **72**, 1524-1543.
- Zhou, J. X., Wen, J. B., Zhang, H. Y. and Zhang, L. (2003) A nodal integration and post-processing technique based on Voronoi diagram for Galerkin meshless methods, *Comput. Methods Appl. Mech. Engrg.* **192**, 3831–3843.

## Imaging and characterizing damages in metallic plates using Lamb waves

\*C.T. Ng

School of Civil, Environmental & Mining Engineering, The University of Adelaide, SA, Australia.

\*Presenting author: alex.ng@adelaide.edu.au

### Abstract

The paper presents a study for quantitative imaging of damage in metallic plates using the fundamental anti-symmetric mode of ( $A_0$ ) Lamb wave. The study proposed a two-stage approach, in which the damage location is first determined in stage-one through analyzing the cross-correlation of the excitation pulse and scattered wave signal, and the damage is then characterized in stage-two using the Mindlin plate theory based Lamb wave diffraction tomography. The damage considered in this study is an elliptical shape of plate thickness reduction, which is a simplified representation of corrosion damage in metallic plates. The two-stage approach is employed to quantitatively image the plate thickness reduction, i.e. determine the location, size and shape of the thickness reduction. Finite element simulation of a circular transducer network with eight transducers is used to demonstrate the capability of the two-stage approach in characterizing the damage. The results show that the two-stage approach is able to accurately identify the damage location and provide a reasonable estimation of the size and shape of the damage.

**Keywords:** Lamb wave, diffraction tomography, scattering, imaging, damage characterization, structural health monitoring

### Introduction

Structural health monitoring (SHM) is a process of monitoring the performance and evaluating the state of health of structures based on measurements. It can be used to ensure the safety and sustainability of structures and have been widely employed in different engineering fields, such as civil, mechanical and aerospace engineering. Lamb wave based approach has been widely recognized as one of the promising techniques for damage detection, and hence, to ensure the structural safety (Alleyne et al 2001; Veidt and Ng 2011; Ng and Veidt 2012). The advantages of using Lamb wave in damage detection are its high sensitive to most types of damage, efficient in detecting small and subsurface damage, and capable for inspecting large structure areas. In the last decade, different Lamb wave based methods have been developed for damage inspection. Apart from the essential requirement that damage inspection systems must be reliable, two highly desirable features are the graphic representation and quantitative damage identification.

Lamb wave based tomographic approach is one of the techniques that can achieve the aforementioned desirable features. Early developments focused on the use of the time-of-flight information to reconstruct an image for damage identification (Jansen and Hutchins 1990; Malyarenko and Hinder 2001; Leonard and Hinders 2005; Belanger et al. 2010; Huthwaite and Simonetti 2013). A Mindlin plate theory based Lamb wave diffraction tomography framework was proposed by Wang and Rose (2003). The framework reconstructs the damage image through Born approximation of the Lamb wave scattering from the damage and the inversion is solved by direct Fourier inversion approach. Rohde et al. (2009) then demonstrated the damage imaging reconstruction could be achieved using eight transducers through the far-field Born approximation. Recently Rose and Wang (2010) proposed a filtered back-propagation algorithm to solve the inverse problem in the Lamb wave diffraction tomography. They demonstrated that the filtered back-propagation algorithm in reconstructing the damage image is more robust and computationally efficient.

This study employs a two-stage approach to reconstruct the elliptical shape of plate thickness reduction in metallic plates, which demonstrates the capability of the approach in imaging the damage size and shape. The paper is organized as follows. The proposed two-stage approach is first described in the first section. The numerical verification using a three-dimensional (3D) explicit finite element (FE) simulation is then described and the results are discussed in detail. Finally, conclusions are drawn in the last section.

## Two-stage imaging approach

The two-stage imaging approach relies on collecting Lamb wave data through a distributed transducer network with  $N$  transducers. Each of these transducers can act as actuator and sensor for actuating and sensing the Lamb waves. In the damage detection process, the network of transducers is used to sequentially scan the structure before and after the presence of the damage by transmitting and receiving Lamb waves. In each scan, one of the transducers is used to excite the  $A_0$  Lamb wave and the rest of the transducers are used to measure the Lamb wave signals. In each sequential scan process, there are in total of  $N(N-1)$  actuator-sensor signal paths. Once the data of before and after the presence of the damage is collected, a baseline subtraction process is then used to extract the scattered wave signals (Ng and Veidt 2011)

$$\mathbf{u}^S = \mathbf{u}^D - \mathbf{u}^{UD} \quad (1)$$

where  $\mathbf{u}^{UD}$  and  $\mathbf{u}^D$  are the signals measured before and after the presence of the damage.  $\mathbf{u}^S$  is the scattered signal extracted using the baseline subtraction process. The scattered signals are then used in the two-stage approach to reconstruct the damage.

In the stage-one, the targeted inspection area is discretized into a set of image pixels located at  $(x,y)$ . Considering two of the transducers in the transducer network as an example, the intensity of the image pixel ( $I_{ab}$ ) at  $(x,y)$  contributed by the actuator/sensor signal path  $a$ - $b$  (two of the transducers in the network) can be calculated as (Wang et al. 2005; Ng and Veidt 2009)

$$I_{ab}(x,y) = C_{ab} \left( \frac{\sqrt{(y_b - y)^2 + (x_b - x)^2} + \sqrt{(y - y_a)^2 + (x - x_a)^2}}{c_g} \right) \quad (2)$$

where  $(x_a, y_a)$  and  $(x_b, y_b)$  is the location of transducer  $a$  and  $b$ , respectively.  $(x, y)$  is the location of image pixel.  $c_g$  is the group velocity of the fundamental asymmetric mode of Lamb wave.  $C_{ab}$  is the cross-correlation between the incident pulse generated by transducer  $a$  and the scattered wave signal measured by transducer  $b$ . In this study the wavelet coefficient calculated using the continuous Gabor wavelet transform (Kishimoto et al. 1995; Ng et al. 2009) is used in the cross-correlation to provide a reliable extraction of the scattered wave signal components at the excitation frequency. Once the image of each actuator-sensor signal path is reconstructed using Eq. (2), the image indicating the estimated damage location can then be reconstructed by superimposing the power flux of all images as

$$I(x,y) = \sum_{a=1}^N \sum_{b=1, b \neq a}^N A_{ab} I_{ab}^2 \quad (3)$$

where  $A_{ab}$  is a weighting factor used to account for varying sensitivities of individual transducer and is equal to unity for uniform aperture weighting.

After the damage location is estimated in stage-one, the Mindlin plate theory based Lamb wave diffraction tomography is then employed to reconstruct the size and shape of the plate thickness reduction in stage-two. It is assumed that the damage is a weak inhomogeneity with a finite region  $\Sigma$ . The scattered flexural wave can be represented as follow using the Born approximation (Wang and Chang 2005)

$$\mathbf{u}_B^S(\mathbf{x}, \omega) = \iint_{\Sigma} \left[ s_1 D \Gamma_{\beta\alpha}^I g_{3\alpha,\beta} + s_2 \kappa^2 G h (\mathbf{u}_{\alpha}^I - \Omega_{\alpha}^I) (g_{3\alpha} + g_{33,\alpha}) + s_3 \rho I \omega^2 \Omega_{\alpha}^I g_{3\alpha} + s_4 \rho h \omega^2 g_{33} \right] d^2 \xi \quad (4)$$

where  $\omega$  is the excitation frequency of the incident wave.  $D = Eh^3 / 12(1-\nu)^2$  is the flexural stiffness of the plate, where  $E$ ,  $h$  and  $\nu$  are the Young's modulus, thickness and Poisson's ratio, respectively.  $\rho$  and  $G$  are the effective shear modulus and the density of the plate material.  $I = h^3 / 12$  is the through-thickness area moment.  $\kappa = \pi / \sqrt{12}$  is the shear correction factor for accurate representation of the low frequency behavior.  $s_n$ , for  $n=1, \dots, 4$ , are parameter perturbations and have non-zero value for  $\mathbf{x} \in \Sigma$  but vanish for  $\mathbf{x} \notin \Sigma$ .  $\mathbf{x}$  is a position vector and  $\xi$  is an arbitrary point within the region  $\Sigma$ .  $\mathbf{u}_{\alpha}^I$  and  $\Omega_{\alpha}^I$  are the plate-normal displacement and rotary deflections.  $\Gamma_{\beta\alpha}^I$  is the plate-theory strain. The comma indicates the differentiation between subscripts  $[\alpha, \beta] = 1, 2$ .  $g_{\beta\alpha} = g_{\beta\alpha}(\xi | \mathbf{x})$  is the dynamic Green's functions solution for wave scattering (Rose and Wang 2004; Ng et al. 2012). The scattered waves can be simplified using the far-field asymptotic expression of the Hankel function as



$$u_B^S(r_s, \theta^S, \theta^I, \omega) = \sum_{n=1}^4 \left[ P_n(k_1, \theta^S - \theta^I) \hat{S}_n(\mathbf{k}_1^S - \mathbf{k}_1^I) \sqrt{\frac{2}{\pi k_1 r_s}} e^{i(k_1 r_s - \frac{\pi}{4})} \right] \quad (5)$$

where  $r_s = \sqrt{(x-\xi)^2 + (y-\eta)^2}$ .  $\theta^I$  and  $\theta^S$  are the angle of the incident and scattered waves, respectively.  $\mathbf{k}_1^I = k_1 [\cos \theta^I, \sin \theta^I]$  and  $\mathbf{k}_1^S = k_1 [\cos \theta^S, \sin \theta^S]$  denote the incident and scattered wave vector.  $k_1$  is the wavenumber evaluated using Mindlin plate theory.  $P_n$  for  $n=1,2,3,4$  are defined as

$$P_1(k_1, \theta) = -\frac{i\gamma k_1^2}{4(k_1^2 + k_2^2)} (\cos^2 \theta + \nu \sin^2 \theta) \quad (6)$$

$$P_2(k_1, \theta) = -\frac{i\kappa^2 G h k_1^2 (1-\gamma)^2}{4D\gamma(k_1^2 + k_2^2)} \cos \theta \quad (7)$$

$$P_3(k_1, \theta) = \frac{i\rho I \omega^2 \gamma k_1^2}{4D(k_1^2 + k_2^2)} \cos \theta \quad (8)$$

$$P_4(k_1, \theta) = \frac{i\rho h \omega^2}{4D\gamma(k_1^2 + k_2^2)} \quad (9)$$

where  $k_2$  is the wavenumber of the second possible wave mode in Mindlin plate theory.  $\gamma = i/4D(k_1^2 - k_2^2)$ .  $\hat{S}_n(\mathbf{k})$  is the two-dimensional Fourier transform of the perturbation functions and is defined as

$$\hat{S}_n(\mathbf{k}) = \iint_{\Sigma} s_n(\xi) e^{-i\mathbf{k}\cdot\xi} d\xi \quad (10)$$

In this study the focus is on the reduction of the plate thickness due to corrosion in metallic plate, and hence,  $3s_1(\xi) = s_2(\xi) = 3s_3(\xi) = s_4(\xi)$  is used in the image reconstruction. Once the  $\hat{S}_n(\mathbf{k})$  is calculated, the perturbation function  $s_n(\xi)$  can be obtained by carrying out the two-dimensional inverse Fourier transform as

$$s_n(\mathbf{x}) = \frac{1}{(2\pi)^2} \iint_{|\mathbf{k}| \leq k_1} \hat{S}_n(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} dk_x dk_y \quad (11)$$

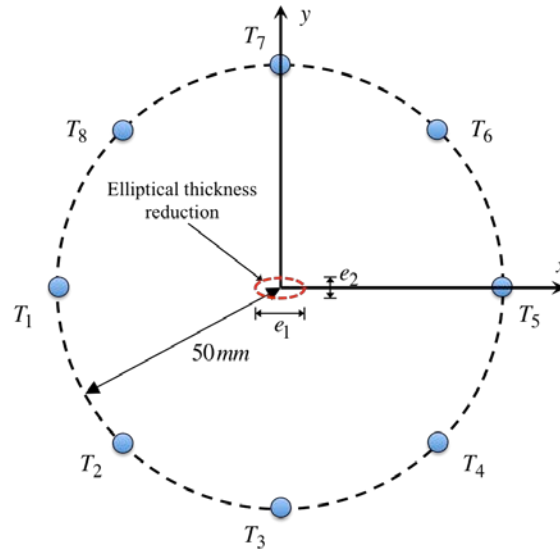
where  $\mathbf{k} = [k_x, k_y] = [k_1(\cos \theta^S - \cos \theta^I), k_1(\sin \theta^S - \sin \theta^I)]$ . The obtained perturbation function can then provide a quantitative imaging of the damage.

### Numerical case studies

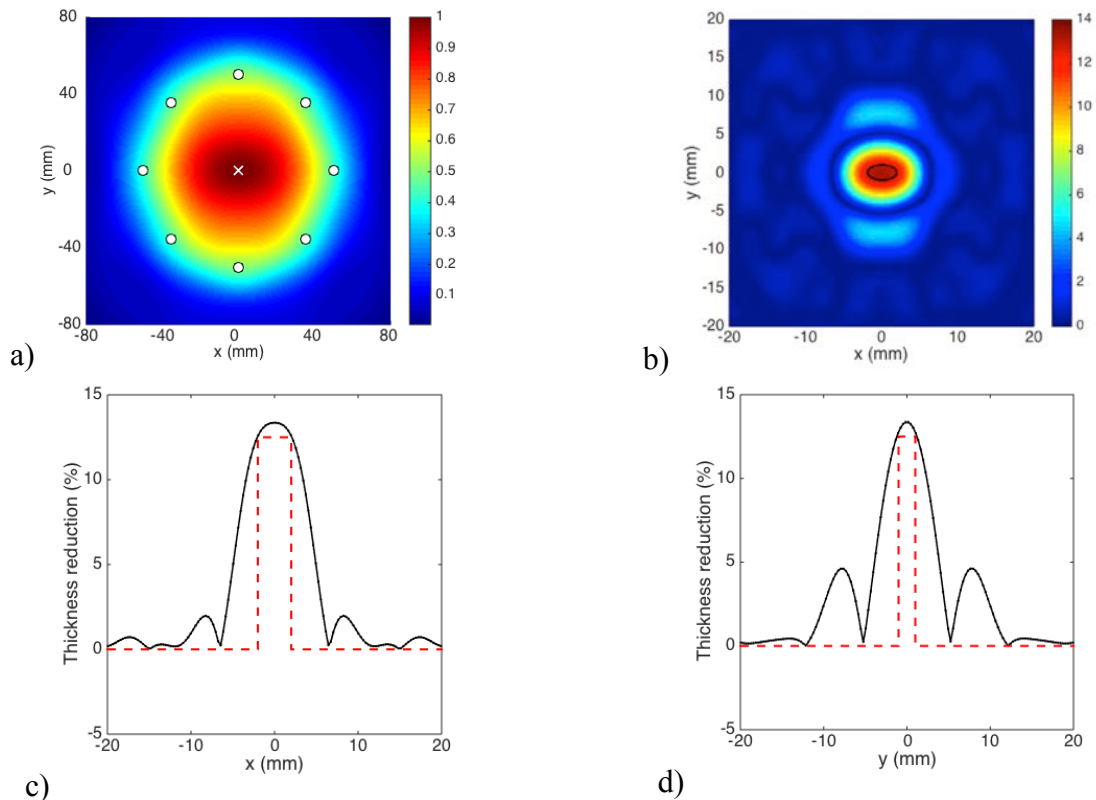
A  $300 \times 300 \times 6 \text{ mm}^3$  steel plate was modeled using 3D explicit FE simulation. The commercial FE software LS-DYNA was used in this study to carry out the FE simulation. The material properties of the steel are  $E = 203 \text{ GPa}$ ,  $\nu = 0.29$  and  $\rho = 7800 \text{ kg/m}^3$ . The steel plate was modeled using eight-noded 3D reduced integration solid brick elements with hourglass control. It is assumed that a 50mm radius circular transducer network was installed on the steel plate as shown in Fig. 1. The transducer network has eight piezoceramic transducers ( $T_1 - T_8$ ) with 5mm diameter. The excitations generated by the transducers were modeled through applying the out-of-plane displacement to the surface nodal point covered by the piezoceramic transducers. The excitation signal is a 200kHz narrow band six-cycle sinusoidal tone burst pulse modulated by a Hanning window. In this study the  $A_0$  Lamb wave was excited and the wavelength was 11.99mm. The solid brick elements had in-plane square shape with dimension  $0.75 \times 0.75 \text{ mm}^2$  and the thickness was 0.375mm, and hence, the plate was modeled using 16 layers of solid element in the thickness direction.

Three damage cases, Cases E1 – E3, were considered in this study. Case E1 considered an elliptical thickness reduction with major axis length  $e_1 = 4 \text{ mm}$  and minor axis length  $e_2 = 2 \text{ mm}$ , respectively. In this case the major axis is along the  $x$ -axis as shown in Fig. 1. The thickness reduction was generated by removing a top and bottom layer of the solid element in the elliptical thickness reduction area. The total percentage of thickness reduction is 12.5%. As the thickness reduction was symmetrical about the mid-plane of the plate, there was no mode conversion effect in

the scattered waves. Cases E2 and E3 consider the same elliptical thickness reduction but the major axis direction was rotated  $45^\circ$  and  $90^\circ$  in anti-clockwise direction, respectively. All cases considered the center of the elliptical thickness reduction located at the center of the circular transducer network.



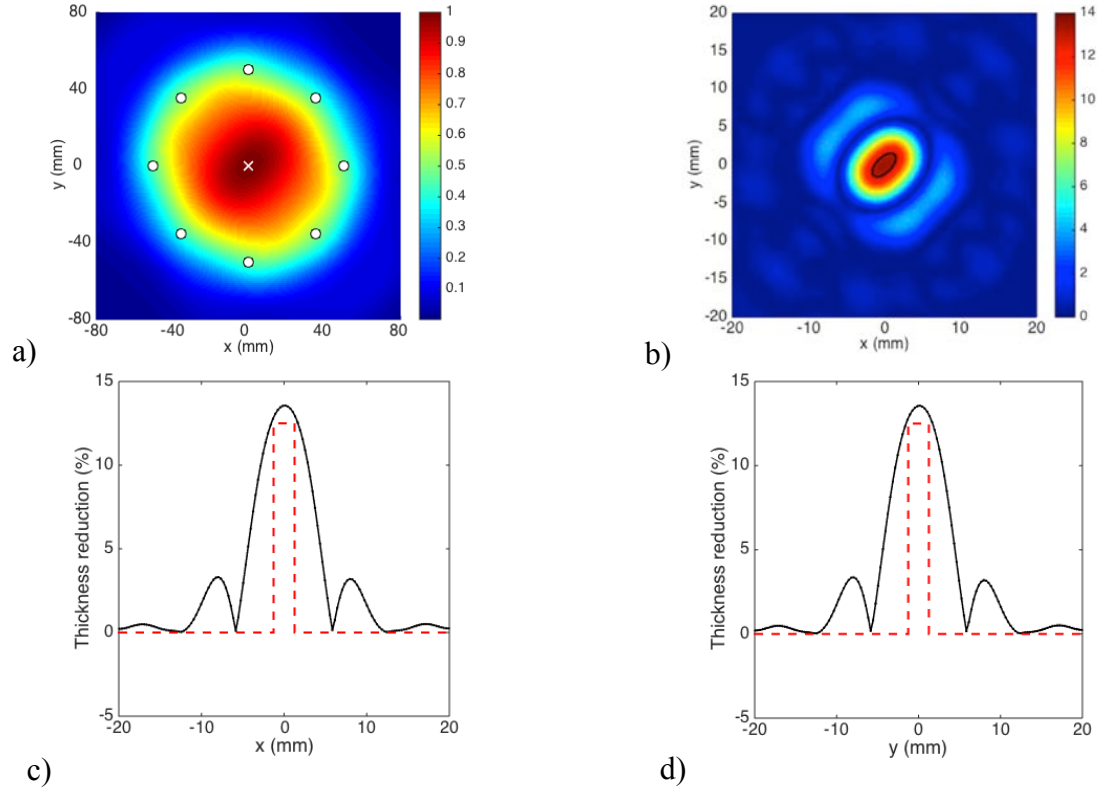
**Figure 1. Transducer network and elliptical thickness reduction in the numerical case studies**



**Figure 2. a) Reconstructed damage location image (white cross: actual damage location), b) reconstructed plate thickness reduction percentage image (close up view at the damage region (black ellipse: actual size and shape of the damage), and profile across the damage in c)  $x$ -axis and d)  $y$ -axis (black solid line: reconstructed thickness reduction percentage; red dashed line: actual thickness reduction percentage) for Case E1**

Using the stage-one of the proposed methodology, a reconstructed damage location image for Cases E1, E2 and E3 are shown in Figs. 2a, 3a and 4a, respectively. The white circles indicate the transducer locations and the white cross indicate the center of the actual thickness reduction. The reconstructed damage location image correctly determines the actual damage location for all three

cases. Figs. 2b, 3b and 4b show the reconstructed plate thickness reduction percentage image using stage-two of the proposed methodology, i.e. the Lamb wave diffraction tomography. In these figures, the actual thickness reduction size and shape are indicated by black ellipses. The images show that there is a good agreement between the reconstructed and actual plate thickness reduction size and shape.

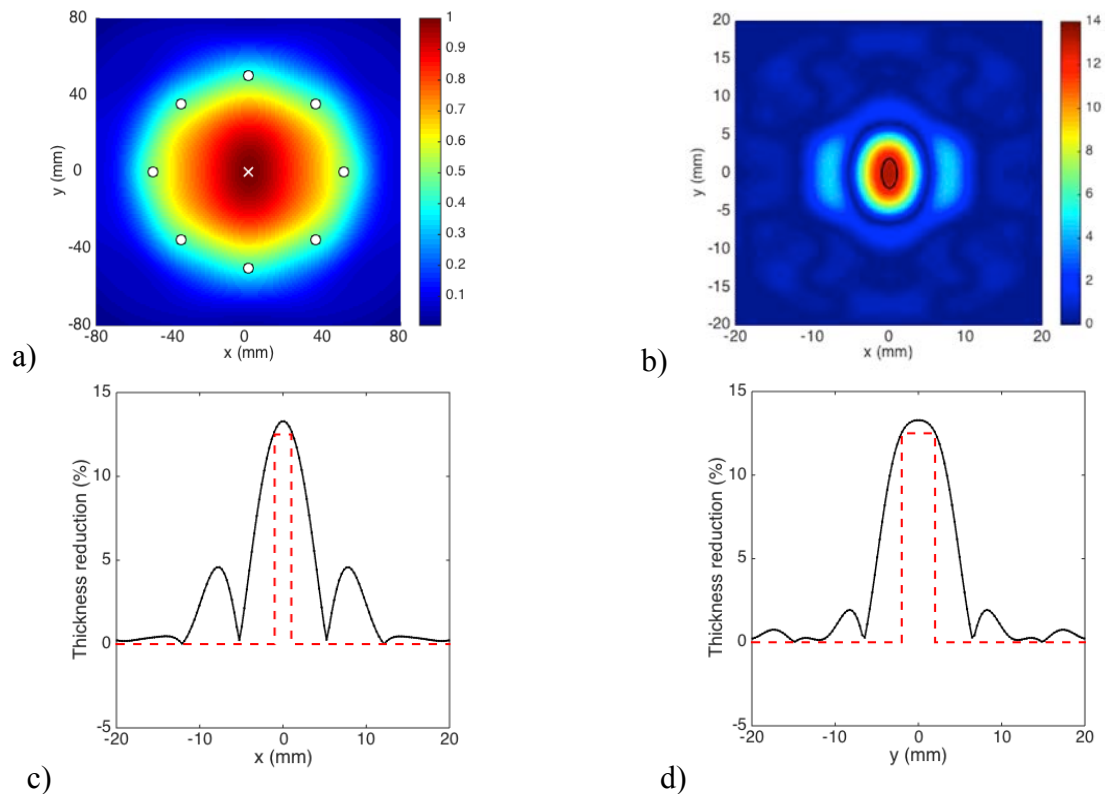


**Figure 3. a) Reconstructed damage location image (white cross: actual damage location), b) reconstructed plate thickness reduction percentage image (close up view at the damage region (black ellipse: actual size and shape of the damage), and profile across the damage in c)  $x$ -axis and d)  $y$ -axis (black solid line: reconstructed thickness reduction percentage; red dashed line: actual thickness reduction percentage) for Case E2**

Figs. 2c, 2d, 3c, 3d, 4c and 4d show the thickness reduction percentage profile across the damage for Cases E1 to E3. The black solid line is the reconstructed thickness reduction percentage and the red dashed line is the actual thickness reduction percentage. These figures show that the reconstructed profiles are able to provide a reasonable estimation of the actual thickness reduction percentage.

## Conclusions

This paper has presented study of employing a two-stage approach for quantitative imaging of damage in metallic plates. The damage type considered in this study is an elliptical thickness reduction. In stage-one of the approach, the location of the thickness reduction was first determined. The size and shape of the thickness reduction were then determined in stage-two using the Lamb wave diffraction tomography. Numerical cases studies have been carried out using the 3D explicit FE simulation. Three cases were considered in the study, which considered the same size of the elliptical thickness reduction but with different orientations of the major axis. The results show that the two-stage approach could accurately determine the damage location and provide a reasonable estimation of the size and shape of the thickness reduction.



**Figure 4. a) Reconstructed damage location image (white cross: actual damage location), b) reconstructed plate thickness reduction percentage image (close up view at the damage region (black ellipse: actual size and shape of the damage), and profile across the damage in c)  $x$ -axis and d)  $y$ -axis (black solid line: reconstructed thickness reduction percentage; red dashed line: actual thickness reduction percentage) for Case E3**

## Acknowledgements

The research described in this paper was financially supported by the Australian Research Council under grant number DE130100261. The support is greatly appreciated.

## References

- Alleyne D., Pavlakovic B., Lowe M. and Cawley P., Rapid, long range inspection of chemical plant pipework using guided waves, *Insight*, **43**(2): 93-96, 2001
- Belanger P., Cawley P. and Simonetti F., Guided wave diffraction tomography within the Born approximation, *IEEE Transactions of Ultrasonics, Ferroelectrics and Frequency Control*, **57**(6): 1405-1418, 2010
- Huthwaite P. and Simonetti F., High-resolution guided wave tomography, *Wave Motion*, **50**: 979-993, 2013
- Jansen, D.P. and Hutchins, D.A. (1990), "Lamb wave tomography", *IEEE Ultrasonics Symposium Proceedings*, Honolulu, HI, December 1017-1020.
- Kishimoto K., Inoue H., Hamada M. and Shibuya T., Time frequency analysis of dispersive waves by means of wavelet transform, *Journal of Applied Mechanics*. **62**: 841-848, 1995
- Leonard K.R. and Hinder M.K., Lamb wave tomography of pipe-like structures, *Ultrasonics*, **43**: 574-583, 2005
- Malyarenko E.V. and Hinder M.K., Ultrasonic man wave diffraction tomography, *Ultrasonics*, **39**(4): 269-281, 2001
- Ng C.T. and Veidt M., A Lamb-wave-based technique for damage detection in composite laminates, *Smart Materials and Structures*, **18**(7)074006: 1-12, 2009
- Ng C.T. and Veidt M., Scattering characteristics of Lamb waves from debondings at structural features in composite laminates, *The Journal of the Acoustical Society of America*, **132**(1): 115-123, 2012
- Ng C.T. and Veidt M., Scattering of the fundamental anti-symmetric Lamb wave at delaminations in composite laminates, *The Journal of the Acoustical Society of America*, **129**(3): 1288-1296, 2011
- Ng C.T., Veidt M. and Rajic N., Integrated piezoceramic transducers for imaging damage in composite laminates, *Proceedings of SPIE*, **7493M**: 1-8, 2009

- Ng C.T., Veidt M., Rose L.R.F. and Wang C.H., Analytical and finite element prediction of Lamb wave scattering at delaminations in quasi-isotropic composite laminates, *Journal of Sound and Vibration*, **331**(22): 4870-4883, 2012
- Rohde A.H., Rose L.R.F., Veidt M. and Wang C.H., Two inversion strategies for plate wave diffraction tomography, *Materials Forum*, **33**: 489-495, 2009
- Rose J.L., A baseline and vision of ultrasonic guided wave inspection potential, *Journal of Pressure Vessel Technology*, **124**: 273-282, 2002
- Rose L.R. and Wang C.H., Mindlin plate theory for damage detection: imaging of flexural inhomogeneities, *The Journal of the Acoustical Society of America*, **127**(2): 754-763, 2010
- Rose L.R. and Wang C.H., Mindlin plate theory for damage detection: source solutions, *The Journal of the Acoustical Society of America*, **116**: 154-171, 2004
- Veidt M. and Ng C.T., Influence of stacking sequence on scattering characteristics of the fundamental anti-symmetric Lamb wave at through holes in composite laminates, *The Journal of the Acoustical Society of America*, **129**(3): 1280-1287, 2011
- Wang C.H. and Chang F.K., Scattering of plate waves by a cylindrical inhomogeneity, *Journal of Sound and Vibration*, **282**: 429-451, 2005
- Wang C.H. and Rose L.R.F., Plate-wave diffraction tomography for structural health monitoring, *Review of Progress in Quantitative Nondestructive Evaluation*, **22**: 1615-1622, 2003
- Wang C.H., Rose J.T. and Chang F.K., A synthetic time reversal imaging method for structural health monitoring, *Smart Materials and Structures*, **13**: 415-423, 2004

# A Non-intrusive Polynomial Chaos For Uncertainty Quantification On

## Numerical Simulation Of Flows Around Cylinder<sup>†</sup>

Yanjin Wang\*, Shudao Zhang

Institute of Applied Physics and Computational Mathematics,  
No. 2, Fenghao-Donglu, Haidian, Beijing, 100094, China

\*Corresponding author: wang\_yanjin@iapcm.ac.cn

### Abstract

Uncertainty quantification (UQ) aims at a meaningful characterization of uncertainties in stochastic systems and efficient propagation of these uncertainties for quantitative validation of model predictions from available measurements. In this paper, UQ of flows around cylinder is studied by non-intrusive polynomial chaos (NIPC) method. The flow is described under the incompressible Navier-Stokes equations as

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u}$$

$$\nabla \cdot \mathbf{u} = 0$$

Where  $\mathbf{u} = (u, v)$ , and  $p$  are the  $x$ -component of the velocity, the  $y$ -component of the velocity, and the scalar pressure field, respectively, with at the initial time, zero flow is assumed, the inflow boundary conditions for this case are given by a parabolic inflow profile, modulated by a sine function depending on time as

$$u(x, y, t) = 0.41^2 \sin\left(\frac{\pi}{8}t\right) 6(y + 0.2)(0.21 - y),$$

$$v(x, y, t) = 0,$$

no slip boundary conditions are imposed on the cylinder and channel walls; at the outflow, a zero pressure is assumed and natural boundary conditions for velocity. We here assume that the inflow  $x$ -component of velocity  $u$  and viscosity coefficient  $\nu$  are random parameters. By the initial velocity of inflow, we let its simulating terminated time be  $T = 8$ . In this study we assume that their PDF are uniform distribution. So their bases of polynomial chaos are Legendre polynomials.

Based on the validation with benchmark results, discussions are mainly focused on the statistic properties of horizontal and vertical components of the total force on the cylinder. The influence of random variables on flow parameters decreases with increasing of viscosity coefficient. The results indicate the effect of NIPC method on the simulation of propagation of uncertainty in the flow field.

**Keywords:** Non-intrusive Polynomial Chaos, Uncertainty Quantification, Flows around Cylinder, Numerical Simulation.

### References

- Weiner, N. (1938) The Homogenous Chaos, *American Journal of Mathematics*, Vol. 60, pp.897-936.
- Xiu D, Karniadakis G E. (2002) The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J Sci Comput*, 24: 619–644.
- Schofer M., Turek S., The benchmark problem flow around a cylinder. In Flow Simulation with High-Performance Computers II, E.H., Hirschel (Eds.). *Notes on Numerical Fluid Mechanics* 52. Vieweg: Braunschweig, 1996, 547-566.
- Wang X D, Kang S.(2010) Application of polynomial chaos on numerical simulation of stochastic cavity flow. *Sci China Tech Sci.*, 53: 2853 2861.
- Jones B.A, Doostan A.(2013) Satellite collision probability estimation using polynomial chaos expansions. *Advances in Space Research*, 52:1860-1875.
- Hesthaven Jan S., Warburton T., Nodal Discontinuous Galerkin Methods: Algorithm, Analysis, And Applications, *Texts in Applied Mathematics* 54, 2008, Springer Publisher.

<sup>†</sup> This work is supported by NSF(11371069,11372052) of China

**Appendix:**

The spatial domain for this problem is shown in Fig. 1. In the simulation of Fig. 2-4., the mean value of viscosity coefficient is  $10^3$ ; its standard deviation is 10% of the mean value. In the simulation of Fig. 5-7., the mean value of velocity  $u$  is the boundary condition of inflow; its standard deviation is also 10% of the mean value.

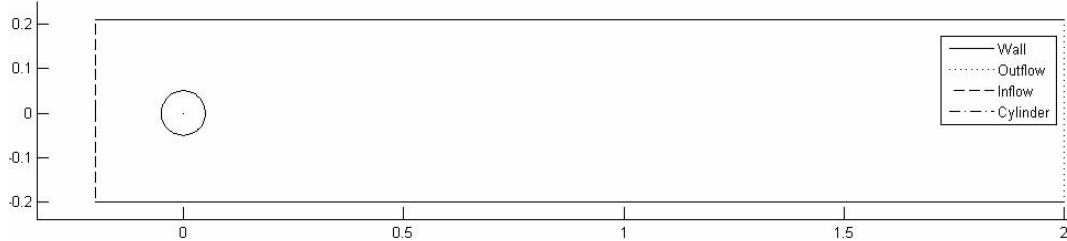


Fig. 1. The spatial domain of the problem.

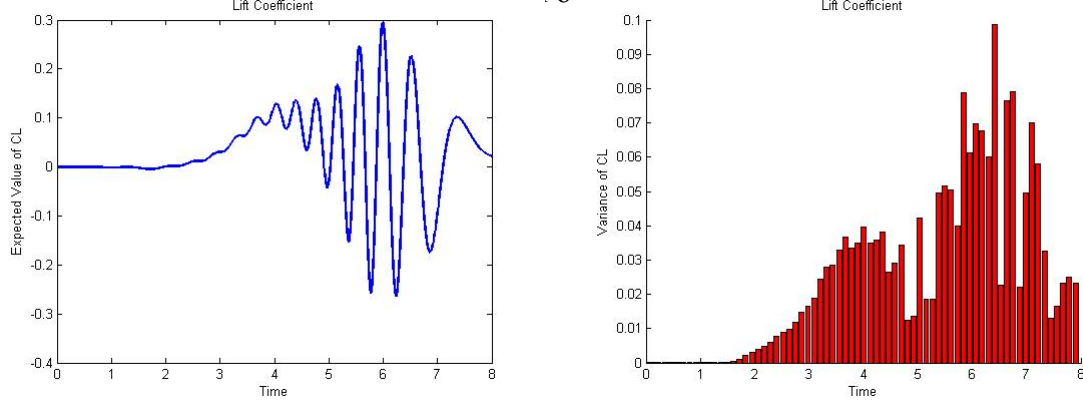


Fig. 2. Vertical components of the total force on the cylinder with mean value  $10^3$  of  $\nu$

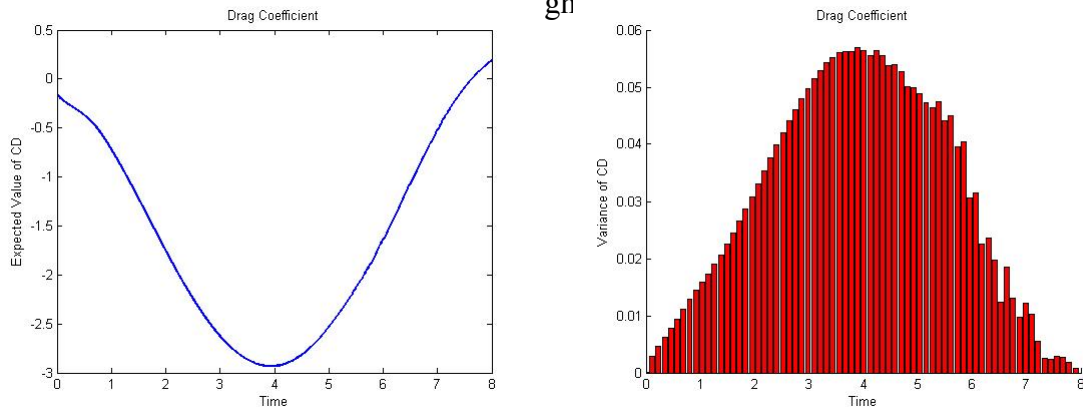


Fig. 3. Horizontal components of the total force on the cylinder with mean value  $10^3$  of  $\nu$  (Left: mean values; Right: standard deviations)

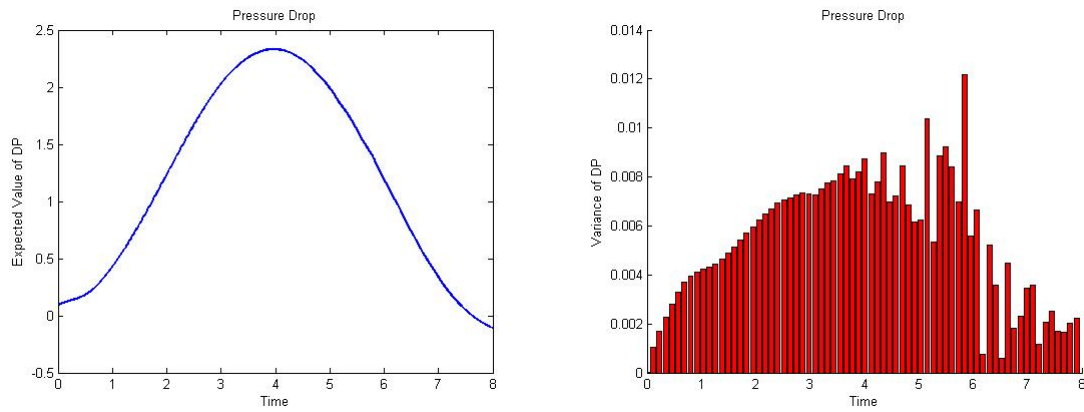


Fig. 4. The drop in pressure between the leading point of the cylinder and trailing point with mean value  $10^3$  of  $\nu$  (Left: mean values; Right: standard deviations)

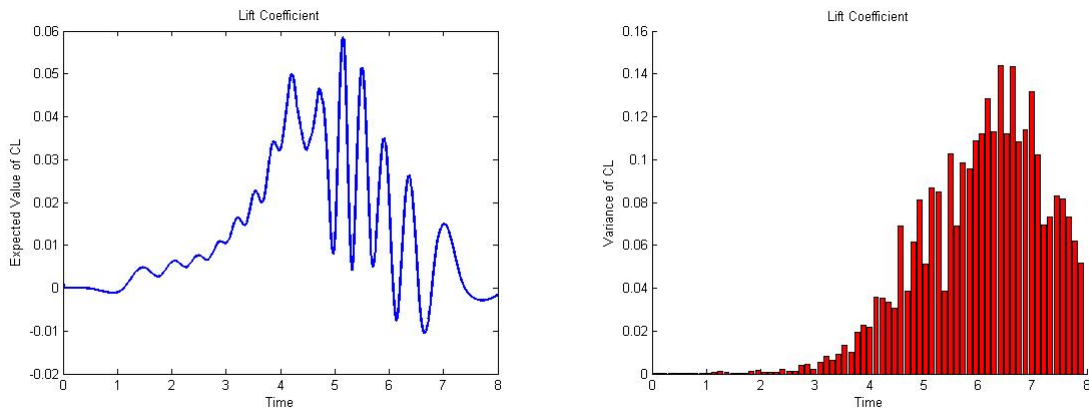


Fig. 5. Vertical components of the total force on the cylinder with random parameter  $u$  (Left: mean values; Right: standard deviations)

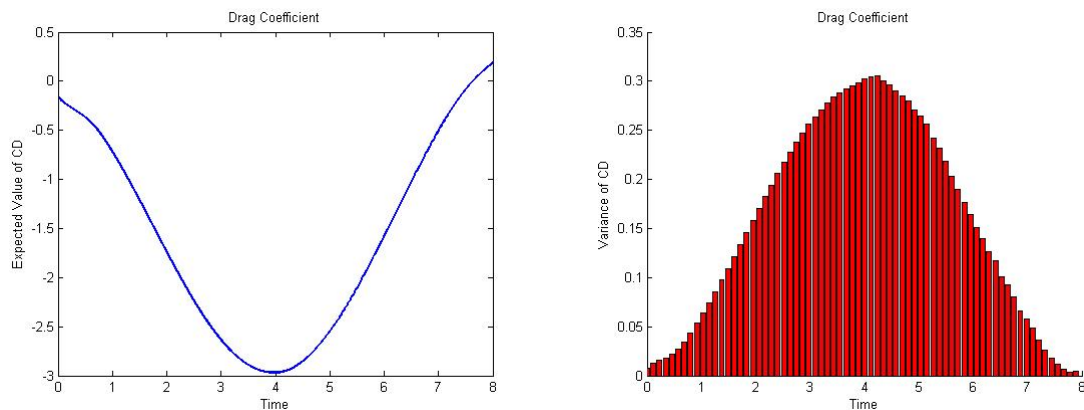


Fig. 6. Horizontal components of the total force on the cylinder with random parameter  $u$  (Left: mean values; Right: standard deviations)



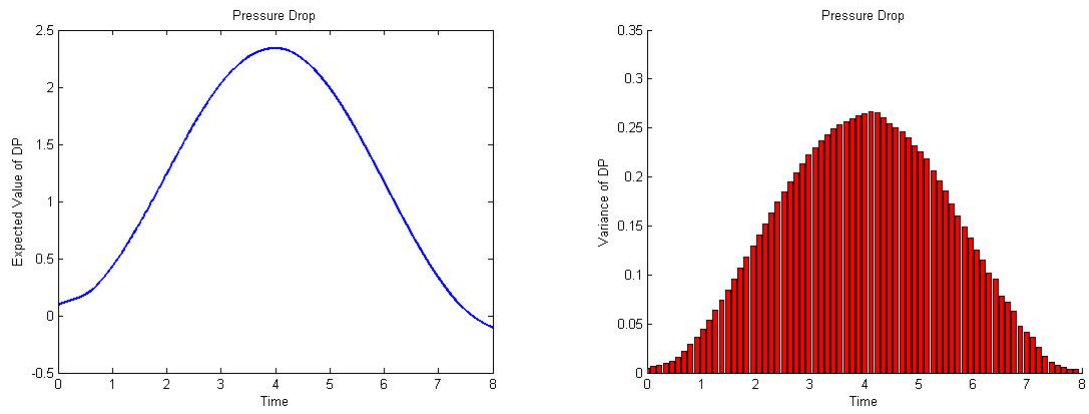


Fig. 7. The drop in pressure between the leading point of the cylinder and trailing point with random parameter  $u$  (Left: mean values; Right: standard deviations)

## Multi-resolution MPS method for 2D free surface flows

Z.Y. Tang<sup>1,2</sup>, Y.L. Zhang<sup>1,2</sup> and D.C. Wan<sup>1,2\*</sup>

1. State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University,  
2. Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

\*Corresponding author: dcwan@sjtu.edu.cn

### Abstract

A multi-resolution MPS (Moving Particle Semi-implicit) method is applied into 2D free surface flows based on our in-house particle solver MLParticle-SJTU in the present work. The basic idea of the present MPS method is to distribute high-resolution particles in the local concern region and low-resolution particles in the other region, such that both the number of particles and the computational cost can be reduced without sacrificing the corresponding accuracy. Considering the effect of different size particles, the kernel function is modified for gradient and Laplacian models and the incompressible condition between different size particles is enforcing by increasing the weight of divergence of velocity in the mixed source term of PPE (Poisson Pressure Equation). In order to validate the present MPS method, two cases are carried out. Firstly, a hydrostatic case is performed. The results show that the contour of pressure field by multi-resolution MPS is quite agreement with that by single resolution MPS. Especially, the multi-resolution MPS can still provide a relative smooth pressure together with single resolution MPS in the vicinity of the interface between the high-resolution and low-resolution particles. For a long time simulation, the kinetic energy of particles by multi-resolution MPS can decrease quickly to the same level as that of single resolution MPS. In addition, a 2D dam breaking flow is simulated and the multi-resolution case can run stably during the whole simulation. The pressure by the multi-resolution MPS is in agreement with experimental data together with single resolution MPS. The contour of pressure field by the former is also similar to that by the later. In addition, the simulation by multi-resolution MPS is as accurate as the traditional MPS with fine particles distributed in the whole domain and the corresponding CPU time can be reduced.

**Keywords:** Multi-resolution method, MPS (Moving Particle Semi-Implicit), dam breaking, free surface flows

### Introduction

In recent years, meshfree particle methods have been developed widely and applied successfully into practical engineering. Unlike the mesh-based method, the fluid is presented as a set of Lagrangian particles in the meshfree method and there is no constant topology relationship between these particles. Thanks to the Lagrangian nature, the particle method is very suitable to deal with flow with largely deformed free surface and moving boundaries [Liu (2008); Khayyer (2008); Tanaka (2010)]. MPS (Moving Particle Semi-implicit) is one of such meshfree methods, which is first proposed by Koshizuka [(1996; 1998)] and then improved by numerous MPS practitioners [Tanaka and Masunaga (2010); Khayyer and Gotoh (2012); Kondo and Koshizuka (2011); Zhang and Wan (2012a)]. Up to now, this method has been applied into a wide variety of violent free surface flows, such as liquid sloshing [Zhang and Wan (2012b; 2014)], dam breaking [Khayyer and Gotoh (2012); Shakibaeinia and Jin (2011)], wave breaking [Gotoh and Sakai (2006); Khayyer and Gotoh (2008); Tang et al. (2014)], green water [Zhang et al. (2013)] and ship-wave interaction [Shibata et al. (2012)]. Despite being an excellent method for solving the largely deformed free surface problems, it still suffers from high computational cost. Especially when applied into 3D simulations, a great number of particles are necessary and the required CPU time can increase sharply. To overcome this problem, multi-resolution simulation are introduced to accelerate the computation. In the framework of SPH, Vacondio et al. [(2012; 2013)] presented a dynamic particle refine algorithm based on particle merging and coalescing during the simulation. Omidvar et al. [(2013)] studied the wave body interaction using variable particle mass distribution. Most of these

works are carried out based on explicit algorithm. Unlike the SPH, the pressure field is obtained based on semi-implicit algorithm in the MPS method. The consuming time per one time level in the later method is much more than that in the former. Many MPS practitioners attempt to introduce the multi-resolution simulation in the MPS. Shibata et al. [(2012)] proposed an overlapping particle technique (OPT) and applied it into a 2D green water. Tanaka [(2009)] presented a multi-resolution method based on traditional MPS. Unfortunately, validation is not given in the article.

In the present work, the multi-resolution technique is applied into 2D free surface flows based on modified MPS [Zhang and Wan (2012a); Zhang et al. (2014)]. Considering the effect of different size particles, the kernel function is modified for gradient and Laplacian models. The incompressible condition between different size particles is enforcing by increasing the weight of divergence of velocity in mixed source term of PPE. In order to validate the present MPS method, two cases are carried out. Firstly, a hydrostatic case is performed. The pressure field and the kinetic energy for fluid particles by multi-resolution MPS are compared with that by single resolution MPS with fine particle distributed in the entire domain. In addition, a 2D dam breaking flow is simulated and the multi-resolution case can run stably during the whole simulation. The comparison among the pressure by single resolution MPS and multi-resolution MPS and experimental data is also made.

## NUMERICAL SCHEME

### *Governing Equations*

In the MPS method, governing equations are the mass and momentum conservation equations. They can read as:

$$\frac{1}{\rho} \frac{D\rho}{Dt} = -\nabla \cdot \mathbf{V} = 0 \quad (1)$$

$$\frac{D\mathbf{V}}{Dt} = -\frac{1}{\rho} \nabla P + \nu \nabla^2 \mathbf{V} + \mathbf{g} \quad (2)$$

Where:  $\rho$  is the fluid density,  $\mathbf{V}$  is the velocity vector,  $P$  presents the pressure,  $\nu$  is the kinematic viscosity, and  $\mathbf{g}$  is gravitational acceleration vector,  $t$  indicates the time. It is noted that Eq. (1) is only available for incompressible fluid.

### *Particle Interaction Models*

#### *Kernel Function*

In the original MPS method, the kernel function (Eq. (3)) first proposed by Koshizuka [(1998)] is usually employed by MPS researchers. However, it has a drawback due to its singular at  $r=0$ . Conversely, we adopt an improved kernel function introduced by Zhang [(2012a)]:

$$W(r) = \begin{cases} \frac{r_e}{r} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (3)$$

$$W(r) = \begin{cases} \frac{r_e}{0.85r + 0.15r_e} - 1 & 0 \leq r < r_e \\ 0 & r_e \leq r \end{cases} \quad (4)$$

Where  $r_c$  denotes the radius of the particle interaction. According to Koshizuka's suggestion, the radius adopted in particle number density and the gradient model is  $r_c = 2.1l_0$ , while  $r_c = 4.0l_0$  is used for the Laplacian model, where  $l_0$  is the initial distance between two adjacent particles.

### Gradient Model

In MPS, the gradient operator is discretized as weighted average of the gradient vector between particles  $i$  and its neighboring particle  $j$ , it can be given as:

$$\langle \nabla P \rangle_i = \frac{d}{n^0} \sum_{j \neq i} \frac{P_j - P_i}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (5)$$

Eq. (5) cannot conserve the linear and angular momentum of the system, and a conservative form is introduced as following [Tanaka and Masunaga (2010)]:

$$\langle \nabla P \rangle_i = \frac{d}{n^0} \sum_{j \neq i} \frac{P_j + P_i}{|\mathbf{r}_j - \mathbf{r}_i|^2} (\mathbf{r}_j - \mathbf{r}_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (6)$$

Where  $n^0$  is the initial particle number density,  $d$  indicates the number of space dimensions,  $\mathbf{r}$  presents the coordinate vector of fluid particle.

### Divergence Model

Similar to gradient model, the divergence model for vector  $\mathbf{v}$  can be formulated as [Shakibaenia and Jin (2012)]:

$$\langle \nabla \cdot \mathbf{v} \rangle_i = \frac{d}{n^0} \sum_{j \neq i} \frac{(\mathbf{v}_j - \mathbf{v}_i) \cdot (\mathbf{r}_j - \mathbf{r}_i)}{|\mathbf{r}_j - \mathbf{r}_i|^2} W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (7)$$

### Laplacian Model

The Laplacian operator is modeled by weighted average of the distribution of a quantity  $\phi$  from particle  $i$  to its neighboring particle  $j$ , it can read as the following equations:

$$\langle \nabla^2 \phi \rangle_i = \frac{2d}{n^0 \lambda} \sum_{j \neq i} (\phi_j - \phi_i) \cdot W(|\mathbf{r}_j - \mathbf{r}_i|) \quad (8)$$

$$\lambda = \frac{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|) \cdot |\mathbf{r}_j - \mathbf{r}_i|^2}{\sum_{j \neq i} W(|\mathbf{r}_j - \mathbf{r}_i|)} \quad (9)$$

Where, the parameter  $\lambda$  is introduced to keep the variance increase equal to the analytical solution.

### *Model of Incompressibility*

In the MPS method, the incompressibility is assured by keeping the particle number density constant. There are two stages in each time step when incompressible condition is enforced: firstly, temporal velocity of particles is calculated explicitly under the action of viscous and gravitational forces, and particles move to intermediate location; secondly, pressure fields are obtained implicitly through solving the Poisson Pressure Equation (PPE), and the velocity and position of particles are updated based on the obtained pressure.

The Poisson Pressure Equation in MPS method is first derived by Koshizuka [(1998)] as following:

$$\langle \nabla^2 P^{k+1} \rangle_i = \frac{\rho}{\Delta t^2} \frac{\langle n^* \rangle_i - n^0}{n^0} \quad (10)$$

Where  $n^*$  is the intermediate particle number density,  $\Delta t$  denotes the calculation time step.

Eq. (10) is commonly employed by many MPS practitioners. However, the source term of the PPE only consists of the deviation of the temporal particle number density from the initial value, and this may lead to high oscillation pressure field in spatial and temporal domain because of unsmooth particle number density. To stabilize the pressure calculation, a mixed source term method combining the velocity divergence-free and constant particle number density is investigated by Tanaka et al. [(2010)] and rewritten by Lee et al. [(2011)] as:

$$\langle \nabla^2 P^{k+1} \rangle_i = (1-\gamma) \frac{\rho}{\Delta t} \nabla \cdot V_i^* - \gamma \frac{\rho}{\Delta t^2} \frac{\langle n^k \rangle_i - n^0}{n^0} \quad (11)$$

Where the subscript  $k$  and  $k+1$  indicate the physical quantity in the  $k$  th and  $k+1$  th time step,  $\gamma$  is the weight of the particle number density in the source term and is assigned a value between 0 and 1. In this paper,  $\gamma = 0.01$  is selected.

### *Detection of Free Surface Particles*

In the original MPS method, zero pressure boundary condition are enforced to the free surface particles. Free surface particles are recognized as the particle number density satisfying the following condition [Koshizuka (1998)]:

$$n_i^* < \beta \cdot n^0 \quad (12)$$

Where  $n^0$  is the initial particle number density, and  $\beta$  indicates a threshold parameter and can be chosen between 0.8 and 0.99. However, misjudgment may occur for inner particles with small particle number density, and imposed unphysical false zero pressure may increase the oscillation frequency and amplitude of pressure field. To overcome this, some approaches have been developed to improve the accuracy of surface particle detection. Tanaka et al. [(2010)] judged the surface particle by using number of neighbor particles. This approach is further improved by Lee et al. [(2011)]. Khayyer et al. [(2009)] proposed a new criteria based on asymmetry of neighboring particles in which particles are judged as surface particles according to the summation of x-coordinate or y-coordinate of particle distance. In the present study, we employ a detection method [Zhang and wan (2012b)] which is also based on the asymmetry arrangement of neighboring

particles, but uses different equations, aiming at describing the asymmetry more accurately, as following:

$$F_i = \frac{D}{n^0} \sum_{j \neq i} \frac{1}{|r_i - r_j|} (r_i - r_j) W(r_{ij}) \quad (13)$$

$F$  is a vector with a large value near the free surface where the neighboring particles distribute largely asymmetry. In  $F$  function, the nearer neighboring particles have larger contribution, while further neighboring particles have smaller effect. This make  $F$  function not sensitive to the neighboring particles locating near the boundary of interaction domain.

If the absolute of the function  $F$  at particle  $i$  is more than a threshold  $\alpha$ , then particle  $i$  is considered as free surface particle. Where  $\alpha$  is assigned to  $0.9|F|^0$ ,  $|F|^0$  is the initial value of  $|F|$  for surface particle.

#### *Modified Gradient and Laplacian Model*

In the single resolution MPS, the interaction radius for each particle is the same as its neighbor particles. However, this condition cannot be ensured since both low-resolution particles with larger interaction radius and high-resolution particles with smaller interaction radius are distributed in the computational domain. This may lead to situations where two interaction particles  $i$  and  $j$  with different interaction radiuses. In the other words, the influence domain of particle  $i$  contains particles  $j$  but not vice versa. When calculating the force between particle  $i$  and its neighbor particle  $j$ , a violation of Newton's third law may occur. In the present work, the supported domain for two neighbor particles  $i$  and  $j$  is modified using the arithmetic mean. In particular, the cut-off radiuses for gradient and Laplacian models are presented as following respectively:

$$r_e = \frac{(r_{ei} + r_{ej})}{2} \quad (14)$$

$$r_{e\_lap} = \frac{(r_{ei\_lap} + r_{ej\_lap})}{2} \quad (15)$$

Define  $L$  is the particle diameter, the modification for gradient [Tanaka (2009)] and Laplacian models can be expressed as following:

$$\langle \nabla P \rangle_i = \frac{D}{n^0} \frac{P_j + P_i}{|r_j - r_i|^2} (r_j - r_i) W(|r_j - r_i|) \frac{L_j}{L_i} \quad (16)$$

$$\langle \nabla^2 \phi \rangle_i = \frac{2D}{n^0} (\phi_j - \phi_i) W_{ij} \frac{\left( \frac{m_j}{L_i} / \lambda_j + \frac{m_j}{L_j} / \lambda_i \right)}{\frac{m_i}{L_i} + \frac{m_j}{L_j}} \quad (17)$$

## TEST CASES

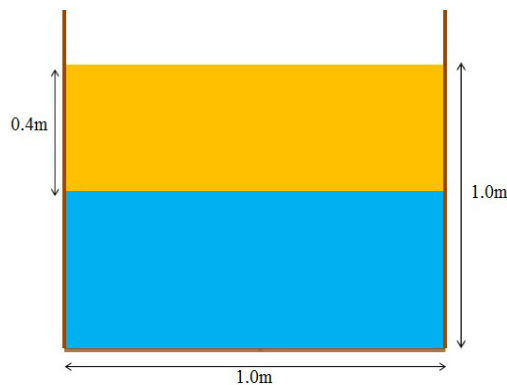
### *The hydrostatic case*

In this section, the hydrostatic problem is carried out by the employment of the single resolution and multi-resolution MPS. A schematic view of the computational domain for this test is shown in

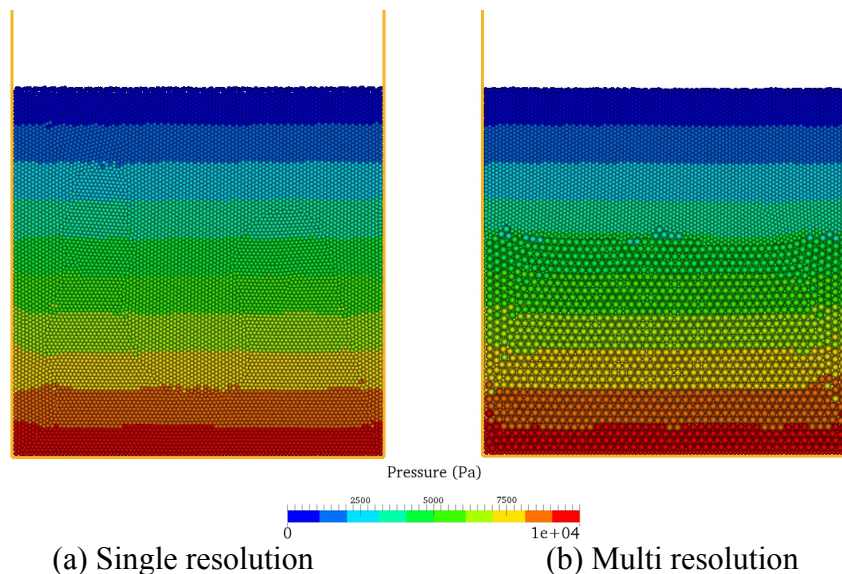
Fig. 1, where both the width of the water tank and the height of fluid are  $H = 1.0$  m. The computational parameters are summarized in Table 1. In case A1, the entire computational domain is discretized by high resolution particles and the corresponding initial particle space is  $d = 0.01$  m. In case A2, two kinds of particle size are employed where the yellow and blue regions are presented by high resolution ( $H/d=100$ ) and low resolution ( $H/d=50$ ) particles respectively. The height of the high resolution region in Case A2 is 0.4 m as the yellow region in Fig. 1. Specially note that the fact fluid height in Case A1 and A2 are 1.0 m and 0.995 m respectively.

Fig. 2 shows the pressure field after a long time evolution of the hydrostatic test. The contour of pressure field in the fine region by multi-resolution MPS is quite similar to that by single resolution MPS with fine particles. Furthermore, in the vicinity of the interface between the high and low resolution particles, a relative smooth pressure field can also be predicted by multi-resolution MPS.

Fig. 3 shows the comparison between the kinetic energy predicted by multi-resolution MPS and single resolution MPS, where the entire domain is represented by fine particles in the later, while the fine particles are only distributed in the yellow region in the former. In Fig. 3, the kinetic energy by multi-resolution MPS quickly decreases to the same level as that by single resolution MPS. This means that the disturbance produced in the interface between different particle sizes is not large, and can be reduced quickly as the initial disturbance in the uniform particle size simulation in this case.



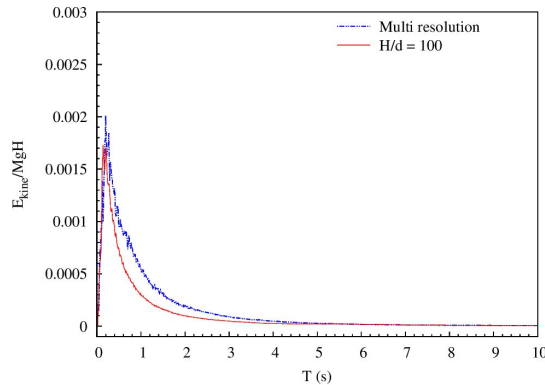
**Figure 1. A schematic sketch of the computational domain for hydrostatic problem**



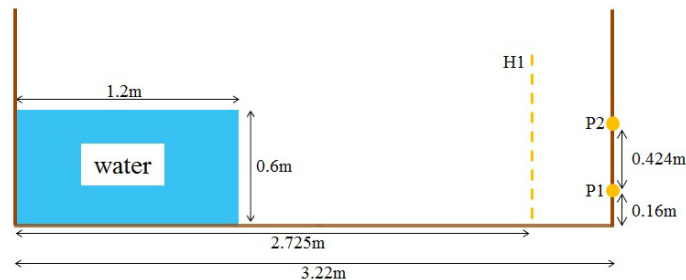
**Figure 2. The pressure field predicted in Case A1 and A2**

**Table 1. Computational parameters in the simulations**

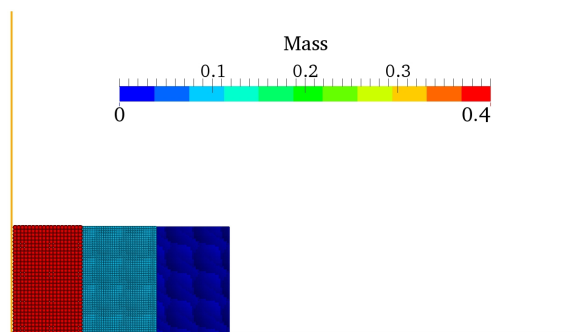
Cases	Initial particle space (H/d)	Description
A1	100	Single resolution
A2	50, 100	Multi resolution



**Figure 3. The comparison between the kinetic energy predicted by single resolution and multi-resolution MPS**



**Figure 4. A schematic view of the computation domain for dam breaking**



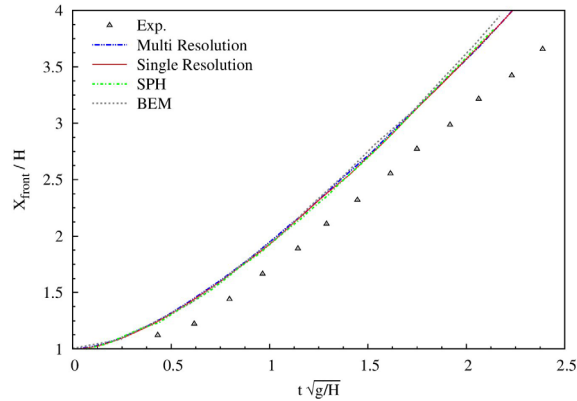
**Figure 5. Initial particle mass distribution for 2D dam breaking problem**

*Dam Break Flow*

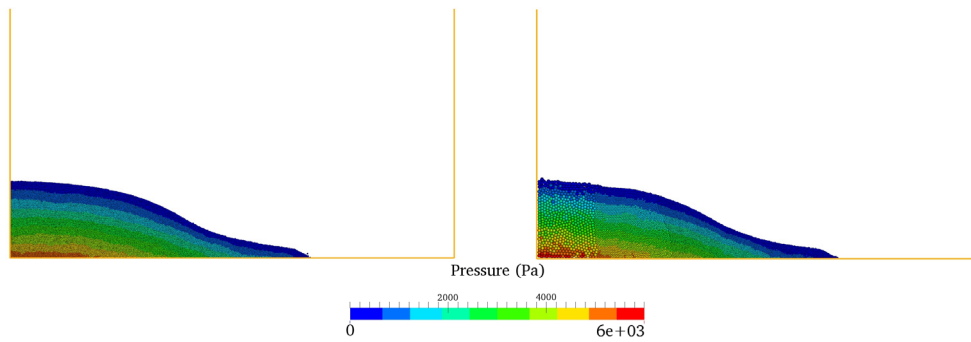
Dam breaking is commonly computed as benchmark case for validation of CFD method in violent flows. In this paper, a dam break is simulated to verify the validation and efficiency of the Multi-resolution MPS method. A schematic view of the computational domain is shown in Fig. 4. The tank is 3.22 m long, 2.0 m high. Initial water column is 1.2 m long and 0.6 m high. The water is



initially constricted by a removable door which is picked up suddenly at  $t=0$  s. A wave height probe is placed at 2.725 m from the left boundary and two pressure probes are placed on the right wall. The initial particle mass distribution for multi-resolution simulation is depicted in Fig. 5, where three kinds of particle size are selected, including  $H/d=30, 60, 120$ , and corresponding masses are 0.4, 0.1 and 0.025 respectively.



**Figure 6. Propagation of the surge front after dam gate removal compared to literature data**

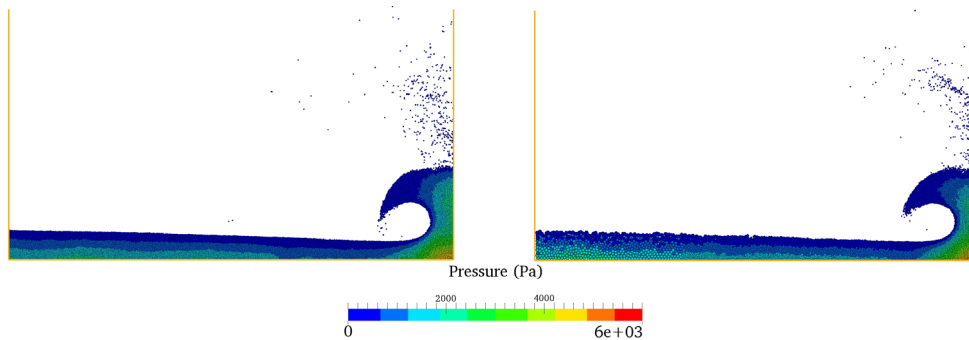


(a) Single resolution

(b) Multi-resolution

**Figure 7. Comparisons of dam-break flows using Single resolution and Multi-resolution MPS**

at  $t\sqrt{g/H}=1.45$



(a) Single resolution

(b) Multi-resolution

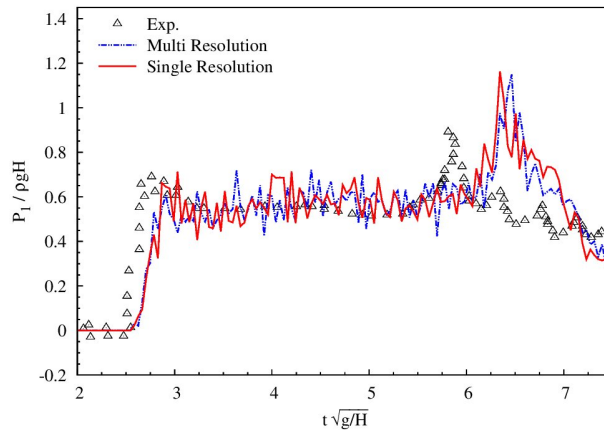
**Figure 8. Comparisons of dam-break flows using Single resolution and Multi-resolution MPS**

at  $t\sqrt{g/H}=5.7$

The wave front propagation along the downstream horizontal dry bed after the dam door release are shown in Fig. 6. The multi-resolution result is quite agreement with that of single resolution MPS and is also quite similar to that of SPH [Ferrari et al. (2009)] and BEM [Colagrossi and Landrini

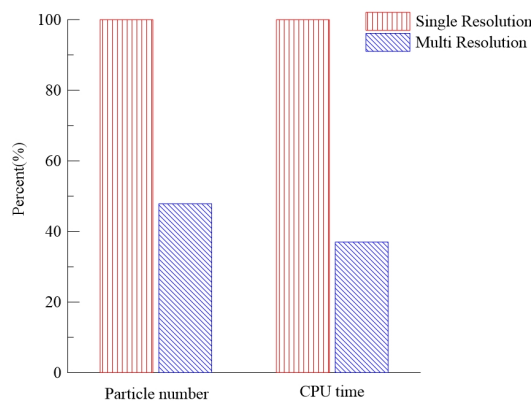
(2003)]. However, the speed of the leading edge by these numerical results is quick than that of the experiment. Similar results can also be reported in literatures [Koshizuka and Oka (1996); Rogers et al. (2010); Abdolmaleki (2004)].

In Fig. 7 and Fig. 8, the pressure fields by single resolution and multi-resolution MPS are depicted for comparison. It can be seen that the computed pressure fields are both relative smooth throughout the time of flow propagation, free surface overturning and impacting the underline water. The contour of the pressure field by multi-resolution MPS is also similar to that by single resolution MPS.



**Figure 9. Time variations of dimensionless pressure at the bottom of the probe P1**

The detailed comparisons of time variations of pressure results at the bottom of the probe P1 are shown in Fig. 9, it can be seen that the overall tendency of pressure variation by both the single resolution and multi-resolution MPS is quite in agreement with experimental data except a clear discrepancy between the position of second pressure peak by the numerical results and experimental data, which is also reported by many researchers employing the single phase model [Marrone et al. (2011); Khayyer et al. (2009)]. However, the pressure variation by multi-resolution MPS is quite close to that of single resolution MPS, including the first impact time and the position of the second pressure peak.



**Figure 10. The number of particles and required CPU time for flowing 3 seconds**

Fig. 10 shows the total number of particles and required CPU time for flowing 3s by conventional single resolution MPS and multi-resolution MPS. Both of these two cases are carried out on personal computer with Intel i7-3770. From the Fig. 10, the number of particles by multi-resolution MPS is nearly half of that by single resolution MPS, while the consuming CPU time in the later is

about two times and a half than that in the former. As presented by Koshizuka [(1988)], the consuming time for solving PPE is proportional to  $N^{1.5}M^{0.5}$ , where  $N$  is the number of particles and  $M$  presents the average number of neighboring particles. The number of particles in multi-resolution MPS is less than that in single resolution MPS, which means that multi-resolution MPS can reduce the number of particles and further decrease the required CPU time. Therefore, the multi-resolution method can be an alternative way to reduce the required computational time if one only concern the local region. Furthermore, considering the pressure variation and the contour of pressure field between single and multi-resolution MPS, the multi-resolution MPS can reduce the CPU time without sacrificing the accuracy.

## Conclusions

In this paper, the multi-resolution MPS method is applied into 2D free surface flows based on in-house particle solver MLParticle-SJTU. In particular, the entire computation domain is discretized with both the low-resolution and high-resolution particles, where only the high-resolution particles are distributed in the concerned local region. Considering the effect of different size particles, the kernel function is modified for gradient and Laplacian models and the incompressible condition between different size particles is enforcing by increasing the weight of divergence of velocity in the mixed source term of PPE. To verify the availability and efficiency of the multi-resolution MPS, two cases are carried out. Firstly, a hydrostatic case is performed. The results show that the contour of pressure field by multi-resolution MPS is nearly the same as that of single resolution MPS. Especially, the multi-resolution MPS can still provide a relative smooth pressure in the vicinity of the interface between the high-resolution and low-resolution particles. For a long time simulation, the kinetic energy of particles by multi-resolution MPS can decrease quickly to the same level as that of single resolution MPS with fine particles distributed in the entire domain. In addition, a 2D dam break flow is carried out and the multi-resolution case can run stably during the whole simulation. Both the pressure variation at the measuring position and the contour of the pressure field at different times by multi-resolution MPS are quite in agreement with that of single resolution MPS. Considering the required CPU time of these two methods, multi-resolution MPS can reduce the computational time without sacrificing its accuracy.

From the above mentioned, it can be seen that the multi-resolution MPS can reduce the required number of particles and further decrease the computational cost. When the traditional MPS is applied into 3D free surface flows such as ship-wave interaction, a great number of particles are necessary and the computational cost can increase sharply. The multi-resolution MPS can be an alternative method to solve this problem and relative work is ongoing.

## Acknowledgement

This work is supported by National Natural Science Foundation of China (Grant Nos. 51379125, 51490675, 11432009, 51411130131), The National Key Basic Research Development Plan (973 Plan) Project of China (Grant No. 2013CB036103), High Technology of Marine Research Project of The Ministry of Industry and Information Technology of China, Chang Jiang Scholars Program (Grant No. T2014099) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (Grant No. 2013022), to which the authors are most grateful.

## References

- Abdolmaleki, K., Thiagarajan, K. P. and Morris-Thomas, M. T. (2004) *Simulation of the dam break problem and impact flows using a Navier-Stokes solver*, Proc 15th Australasian Fluid Mechanics Conference, Sydney.
- Colagrossi, A., and Landrini, M. (2003) Numerical simulation of interfacial flows by smoothed particle hydrodynamics, *Journal of Computational Physics* **191**, 448-475.
- Ferrari, A., Dumbser, M., Toro, E. F., and Armanini, A. (2009) A new 3D parallel SPH scheme for free surface flows, *Computers and Fluids* **38**, 1203-1217.
- Gotoh, H., and Sakai, T. (2006) Key issues in the particle method for computation of wave breaking, *Coastal Engineering* **53**, 171-179.
- Hori, C., Gotoh, H., Ikari, H., and Khayyer, A. (2011) GPU-acceleration for moving particle semi-implicit method. *Computers and Fluids* **51**, 174-183.
- Khayyer, A., and Gotoh, H. (2008) Development of CMPS method for accurate water-surface tracking in breaking waves, *Coastal Engineering Journal* **50**, 179-207.
- Khayyer, A., Gotoh, H., and Shao, S. D. (2008) Corrected incompressible SPH method for accurate water-surface tracking in breaking waves, *Coastal Engineering* **55**, 236-250.
- Khayyer, A., Gotoh, H., and Shao, S. (2009) Enhanced predictions of wave impact pressure by improved incompressible SPH methods, *Applied Ocean Research* **31**, 111-131.
- Khayyer, A., and Gotoh, H. (2012) A 3D higher order Laplacian model for enhancement and stabilization of pressure calculation in 3D MPS-based simulations. *Applied Ocean Research* **37**, 120-126.
- Kondo, M., and Koshizuka, S. (2011) Improvement of stability in moving particle semi - implicit method, *International Journal for Numerical Methods in Fluids* **65**, 638-654.
- 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**, 751-769.
- Koshizuka, S., and Oka, Y. (1996) Moving-particle semi-implicit method for fragmentation of incompressible fluid, *Nuclear science and engineering* **123**, 421-434.
- 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**, 1113-1125.
- Liu, M. B., Liu, G. R., and Zong, Z. (2008) An overview on smoothed particle hydrodynamics, *International Journal of Computational Methods* **5**, 135-188.
- Marrone, S., Antuono, M., Colagrossi, A., Colicchio, G., Touzé, D. L., and Graziani, G. (2011)  $\delta$ -SPH model for simulating violent impact flows, *Computer Methods in Applied Mechanics and Engineering* **200**, 1526-1542.
- Omidvar, P., Stansby, P. K., and Rogers, B. D. (2013) SPH for 3D floating bodies using variable mass particle distribution, *International Journal for Numerical Methods in Fluids* **72**, 427-452.
- Rogers, B. D., Dalrymple, R. A., and Crespo, A. J. C. (2010) State of the art of classical SPH for free surface flows, *Journal of Hydraulic Research* **48**, 6-27.
- Shakibaeinia, A., and Jin, Y. C. (2011). A mesh-free particle model for simulation of mobile-bed dam break, *Advances in Water Resources* **34**, 794-807.
- Shakibaeinia, A., and Jin, Y. C. (2012) MPS mesh-free particle method for multiphase flows, *Computer Methods in Applied Mechanics and Engineering* **229**, 13-26.
- Shibata, K., Koshizuka, S., Sakai, M., and Tanizawa, K. (2012) Lagrangian simulations of ship-wave interactions in rough seas, *Ocean Engineering* **42**, 13-25.
- Shibata, K., Koshizuka, S., and Tamai, T. (2012). *Overlapping particle technique and application to green water on deck*, Proceedings of 2nd International Conference on Violent Flows, Nantes, France.
- Tanaka, M., and Masunaga, T. (2010) Stabilization and smoothing of pressure in MPS method by quasi-compressibility, *Journal of Computational Physics* **229**, 4279-4290.
- Tanaka, M., Masunaga, T., and Nakagawa Y. (2009) Multi-resolution MPS method, *Transactions of JSCEs*.
- Tang, Z. Y., Zhang, Y. X., Li, H. Z., and Wan, D. C. (2014) *Overlapping MPS method for 2D free surface flows*, Proceedings of the 24th International Ocean and Polar Engineering Conference, Busan.
- Vacondio, R., Rogers, B. D., and Stansby, P. K. (2012) Accurate particle splitting for smoothed particle hydrodynamics in shallow water with shock capturing, *International Journal for Numerical Methods in Fluids* **69**, 1377-1410.
- Vacondio, R., Rogers, B. D., Stansby, P. K., Mignosa, P., and Feldman, J. (2013) Variable resolution for SPH: a dynamic particle coalescing and splitting scheme, *Computer Methods in Applied Mechanics and Engineering* **256**, 132-148.
- Zhang, Y. X., and Wan, D. C. (2012a) Numerical Simulation of liquid sloshing in low-filling tank by MPS, *Chinese Journal of Hydrodynamics* **27**, 100-107.
- Zhang, Y. X., and Wan, D. C. (2012b) *Apply MPS method to simulate liquid sloshing in LNG tank*, In Proceedings of the 22nd international offshore and polar engineering conference , Rhodes.
- Zhang, Y. X., and Wan, D. C. (2014) Comparative study of MPS method and level-set method for sloshing flows, *Journal of Hydrodynamics* **26**, 577-585.

- Zhang, Y. X., Wang, X., Tang, Z. Y., and Wan, D. C. (2013) *Numerical Simulation of Green Water Incidents Based on Parallel MPS Method*, In The Twenty-third International Offshore and Polar Engineering Conference, International Society of Offshore and Polar Engineers, Alaska.
- Zhang Y. X., Yang Y. Q., Tang Z. Y., and Wan D. C. (2014) *Parallel MPS method for Three-Dimensional liquid sloshing*, Proceedings of the 24th International Ocean and Polar Engineering Conference, Busan.
- Zhu, X., Cheng, L., Lu, L., and Teng, B. (2011) Implementation of the moving particle semi-implicit method on GPU, *Science China Physics, Mechanics and Astronomy* **54**, 523-532.

# Numerical Investigation of Influence of Eccentricity on the Hydrodynamics of a Ship Maneuvering into a Lock

Qingjie Meng<sup>1,2</sup>, Decheng Wan<sup>1,2\*</sup> and Wenhua Huang<sup>3</sup>

1. State Key Laboratory of Ocean Engineering, School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University,

2. Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration, Shanghai 200240, China

3. School of Science, Huzhou University, Huzhou 313000, China

\*Corresponding author: dcwan@sjtu.edu.cn

## Abstract

By solving the unsteady RANS (Reynolds Averaged Navier–Stokes) equations in combination with the  $k-\omega$  SST turbulence model, the unsteady viscous flow around a 12000TEU ship model while entering the Third Set of Panama Locks with different eccentricity is simulated and the hydrodynamic forces, vertical displacement are predicted and analyzed. Overset grid technology is used to maintain grid orthogonality. The effects of the free surface are taken into account. A benchmark test case is designed first to validate the capability of the present methods in the prediction of the viscous flow around the ship when maneuvering into the lock. The accumulation of water in front of the ship during entry into a lock is noticed, which causes the increase of the velocity of the return flow. A set of systematic computations with different eccentricity are then carried out to examine the effect of eccentricity on the ship-lock hydrodynamic interaction while entering the lock. With higher eccentricity, higher hydrodynamic forces and higher yawing moment is observed, which cause greater risk of contact of a ship with the lock infrastructure.

**Keywords:** lock; 12000TEU, overset grid; ship-lock hydrodynamic interaction; eccentricity

## Introduction

The lock approach will always be accompanied with bank effect on the hydrodynamic forces, since ships can have very small side margins in the lock. Besides, ships are also allowed with a very small under keel clearance to exploit the lock maximally and thus typically a high blockage. The high blockage influences the flow along the ship hull, increasing the relative speed between the ship and inverse flow. Furthermore, the high blockage also causes a so called piston effect, which provokes an accumulation of water during ship's entry into the lock. The frictional resistance increases and water piles up inside the lock resulting in higher resistance. What's more, it is very difficult to keep the ship maneuvering along the lock's centerline. This obviously produces a worse situation for a ship approaching a lock. Overall, a ship will experience a particular hydrodynamic force caused by the hydrodynamic interaction with the lock, during the entering manoeuvre into a lock. This hydrodynamic interaction has a significant influence on the ship navigation safety. The investigation on the effects of eccentricity is of crucial importance for safe operation and effective control of ships passing through a lock.

Several methods have been used to examine the feasibility to use locks for large vessels. Although real scale and model scale tests can be carried out, reliable simulations are necessary to examine infrastructure in an affordable and efficient manner. Vrijburcht (1988) used six-waves-model to calculate the translation waves generated by the lock entry. Vergote (2012) improved the six-waves-model model. Chen (2010) developed a viscous frictional model to calculate dynamical ship-lock interaction problem. Delefortrie et al. (2008, 2009) analyzed the navigation behavior of different ship models in the Third Set of Panama Locks and the influences of approach wall configurations, eccentricities, propeller rates, approaching scenarios and under keel clearances were also discussed. Verwilligen and Richter (2012) investigated the entering manoeuvre of full form ships into the Terneuzen West Lock by means of model testing, full scale trials and real-time simulations. Wang et al. (2014) studied the viscous flow around a ship when entering the Pierre Vandamme Lock based on the CFD package Fluent, but the free surface was neglected.

The objective of this study is to examine the effects of eccentricity on a 12000TEU ship model when entering the Third Set of Panama Locks by predict and analyze the viscous flow and hydrodynamic forces of the ship model maneuvering into the lock with different lateral position. At

first, the capability of the present method for the prediction of the viscous flow around the ship model is validated by the good agreement of the predicted results with the corresponding experimental data. Then a series of systematic computations with different lateral position are carried out and the predicted forces and moments, vertical displacement are analyzed to investigate the viscous flow around the ship and the influence of the eccentricity on the ship-lock hydrodynamic interaction during its entry into a lock.

The computation is carried out by an in-house CFD code based on the Finite Difference Method (FDM). The code was proved to be competent in simulating the unsteady viscous flow around a ship in confined water (Meng et al. 2014). Finite Difference Method (FDM) can be used even when high cell aspect ratio is present and thus is very suitable to solve low speed problems. Refinement grid are used only in the vertical direction to ensure the grid number within an acceptable range as well as the accuracy of the capture of the free surface. When a ship entering into a lock, a mass of water will be pushed into the lock, which might causes the initial bow up and a significantly increased resistance. So, to make our prediction more reasonable, the effects of the free surface is taken into account. Besides, the overset grid technique is introduced to avoid the deterioration of the computational accuracy caused by the mesh distortion near the boundary layer region when the ship moves with large amplitude.

## Mathematical Method

### Governing Equations

The viscous flow around the ship is assumed incompressible and the numerical problem is described by RANS equations coupled with the time-averaged continuity equation in non-dimensional tensor form:

$$\frac{\partial U_i}{\partial t} + U_j \frac{\partial U_i}{\partial x_j} = -\frac{\partial \hat{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left[ Re_{eff}^{-1} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] + S_i \quad (1)$$

$$\frac{\partial U_i}{\partial x_i} = 0 \quad (2)$$

where  $U_i=(u, v, w)$  denote the Reynolds-average velocity components;  $x_i=(x, y, z)$  represent the independent coordinate directions;  $S_i$  stand for a body force due, for instance, to a propeller model, respectively; the piezometric pressure  $\hat{p}$  and the effective Reynolds number  $Re_{eff}$  are:

$$\hat{p} = p_{abs} / \rho U_0^2 + z / Fr^2 + 2k/3 \quad (3)$$

$$Re_{eff}^{-1} = 1 / Re + \gamma_t \quad (4)$$

with  $k$  the turbulence kinetic energy,  $p_{abs}$  the absolute pressure,  $z$  the local vertical elevation and  $\gamma_t$  the non-dimensional turbulent viscosity obtained from a turbulence model. All the variables and properties are nondimensionalized by the reference velocity  $U$ , ship length  $L_{pp}$  and the mass density of the water  $\rho$ .

### Turbulence Model

The  $k-\omega$  SST turbulence model was chosen to close the RANS equations. As the  $k-\omega$  SST turbulence model employs the  $k-\omega$  model near walls and the  $k-\varepsilon$  model away from walls, it gives highly accurate predictions of the onset and the amount of flow separation under adverse pressure gradients.

The turbulence kinetic energy  $k$  is computed using a blended  $k-\varepsilon/k-\omega$  model (Menter, F. R. 1994). In this model, the eddy viscosity  $\nu_t$ , turbulence kinetic energy  $k$  and the turbulence specific dissipation rate  $\omega$  can be computed from:

$$v_i = \frac{k}{\omega} \quad (5)$$

$$\frac{\partial k}{\partial t} + \left( U_j - \sigma_k \frac{\partial v_i}{\partial x_j} \right) \frac{\partial k}{\partial x_j} - \frac{I}{R_k} \nabla^2 k + s_k = 0 \quad (6)$$

$$\frac{\partial \omega}{\partial t} + \left( U_j - \sigma_\omega \frac{\partial v_i}{\partial x_j} \right) \frac{\partial \omega}{\partial x_j} - \frac{I}{R_\omega} \nabla^2 \omega + s_\omega = 0 \quad (7)$$

where the source terms, effective Reynolds numbers, and turbulence production can be described as:

$$S_k = R_k (-G + \beta^* \omega k) \quad (8)$$

$$S_\omega = R_\omega \left( -\gamma \frac{\omega}{k} G + \beta \omega^2 + 2(1 - F_l) \sigma_{\omega 2} \frac{I}{\omega} \frac{\partial k}{\partial x_j} \frac{\partial \omega}{\partial x_j} \right) \quad (9)$$

$$R_{k/\omega} = \frac{I}{I/Re + \sigma_{k/\omega} v_i} \quad (10)$$

$$G = \tau_{ij} \frac{\partial u_i}{\partial u_j} = v_i \left[ (u_y + v_x)^2 + (u_z + w_x)^2 + (v_z + w_y)^2 + 2u_x^2 + 2v_y^2 + 2w_z^2 \right] \quad (11)$$

with  $\beta$ ,  $\beta^*$ ,  $\sigma_k$ ,  $\sigma_\omega$  constants and  $F_l$  the blending function, which is designed to be one in the sub-layer and logarithmic regions of boundary layers and gradually switches to zero in the wake region to take advantage of the strengths of the  $k-\omega$  and  $k-\varepsilon$  models in different position. The model constants are calculated locally as a weighted average, i.e.,  $\phi = F_l \phi_1 + (1 - F_l) \phi_2$  where  $\phi_1$  denote the standard  $k-\omega$  and  $\phi_2$  represent  $k-\varepsilon$  values.

### Free Surface

We capture the location of the transient free surface using level set function  $\phi$  (Osher, S., & Sethian, J. A., 1988; Sussman et al. 1994; Sethian and Smereka. 2003), whose value is related to the distance to the interface. And the value of  $\phi$  is arbitrarily set to positive in water and negative in air and the iso-surface  $\phi=0$  represents the free surface. Since the free surface is considered a material interface, then it should satisfy the kinematic free surface boundary condition and can be described as:

$$\frac{\partial \phi}{\partial t} + v \cdot \nabla \phi = 0 \quad (12)$$

The following conditions for the velocity and pressure should be satisfied:

$$\nabla v \cdot n = 0 \quad (13)$$

$$\hat{p} = \frac{z}{Fr^2} \quad (14)$$

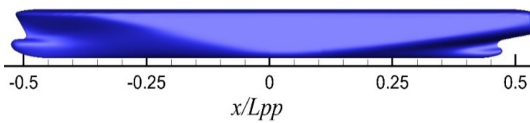
where  $n = -\frac{\nabla \phi}{|\nabla \phi|}$  is the unit normal vector to the free surface, pointing from water to air.

To make sure that the level set function is kept a distance function after the transport step, a reinitialization procedure is used in which the points close to the free surface are reinitialized geometrically, while a transport equation is solved for all other points.

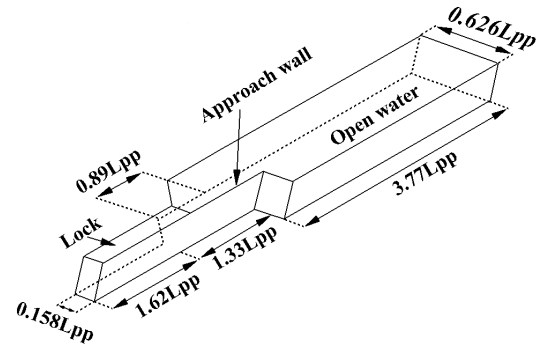


## Simulation Design

The problem under study is a 12000TEU container carrier moving into the Third Set of Panama Lock with different eccentricity. The ship is a 1/80 scale model of 12000TEU. The geometry of the ship model is shown in Fig. 1. A 1/80 scale model of a lock and approach channel is designed according to the preliminary design of the Panama Canal Third Set of Locks. The geometry and the principal dimensions of the lock model is shown in Fig. 2. These geometry is a benchmark on the ship behavior in locks, tested in Flanders Hydraulics Research in 2007-2008. Numerical results are compared with the experimental data by Vantorre, M. et al. (2012).



**Fig.1 Geometry of the 12000TEU hull**



**Fig.2 Geometry of the lock**

All the present work is conducted on a computer cluster which consists of 16 Intel Xeon E5520 (2.27GHz) processors, with 8 cores and 24GB RAM per processor. Each computation is performed using 16 cores and costs about 144 hours of wall clock time.

### Case Conditions

The ship model is  $L_{pp}=4.35$  m long with 0.19 m draft. Table 1 gives the principal dimensions of the ship model. Three cases are studied numerically. Case A, with lateral position  $\Delta y = 0.0$  mm, is designed according to the benchmark test and selected for validation. Numerical tests cases A1 and A2 are also carried out to study the effects of eccentricity on the hydrodynamic forces acting on the ship, with lateral positions:  $\Delta y = 7.5$  mm and  $\Delta y = 19$  mm model scale, respectively. Details of the case conditions are shown in table 2. Free surface is considered for all cases.

**Table 1 Principal dimensions of the 12000TEU model**

Item	Symbol	Unit	Value
Length between perpendiculars	$L_{pp}$	m	4.350
Breadth (molded)	$B$	m	0.613
Draft (molded)	$D$	m	0.190
Blockage coefficient	$C_B$	--	0.65

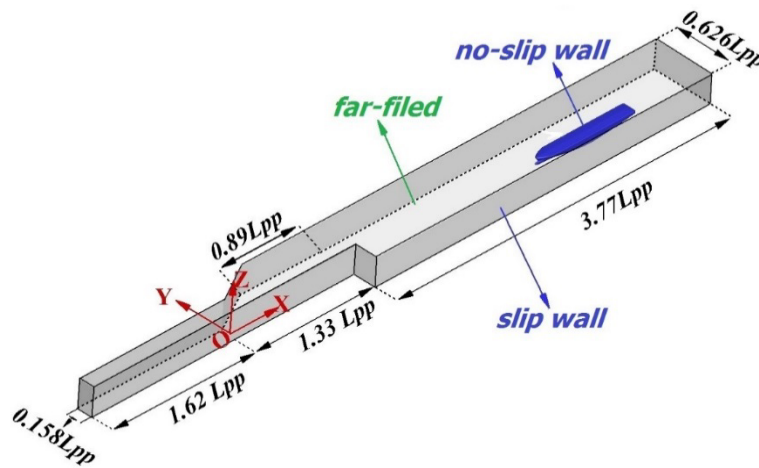
**Table 2 Details of the case conditions**

Conditions			Test No.		
			A	A1	A2
water depth	$h$	m	0.228	0.228	0.228
Froude number	$Fr$	-	0.0176	0.0176	0.0176
Reynolds number	$Re$	$\times 10^5$	4.392	4.392	4.392
Depth/draft ratio	$h/D$	-	1.2	1.2	1.2
lateral position	$\Delta y$	mm	0.0	7.5	19

### Computational Domain, Coordinate System and Boundary Conditions

The computational domain, which extend within  $-1.62 \leq x \leq 5.10$ ,  $-0.0524 \leq z \leq 0.17$ ,  $-0.079 \leq y \leq 0.079$  for the lock and  $-0.313 \leq y \leq 0.313$  for the approach channel, covers the whole ship

considering the asymmetry of the flow field. A right-handed Cartesian coordinate system is located at the door of the lock. The longitudinal  $Ox$ -axis points to the approach channel, the  $Oz$ -axis is vertical and points upward, and the undisturbed free surface is taken as the plane  $z=0$ . The origin of the coordinates is located at the intersection of the waterline, the center plane of the lock and the plane of the lock door. A schematic diagram indicating the coordinate system and the computational domain is given in Fig. 3. The boundary conditions mimic the conditions in the FHR (Flanders Hydraulics Research) towing tank for later comparison of the numerical results with the experimental data. The computational domain is made up of three kinds of boundaries: no-slip wall (hull surface), far-field ( $z=z_{max}$ ) and slip wall (all other boundaries).

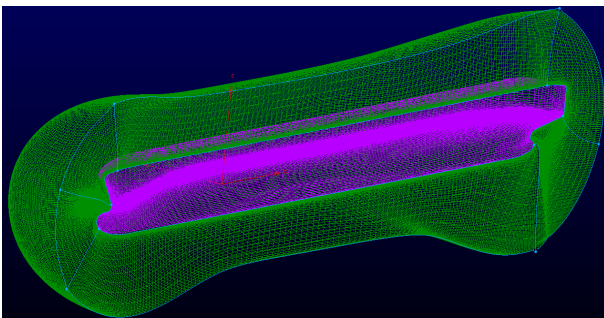


**Fig.3 Computational domain, coordinate system and boundary conditions**

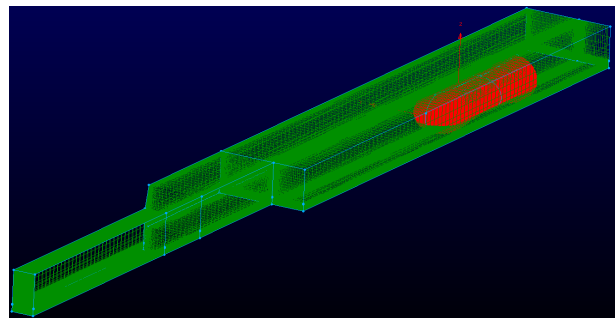
### Grid Design

For all cases, structured grids are used and overset grid technique is utilized to keep the orthogonality of the grid under the consideration of keeping a good computational accuracy.

A sketch of the grid distribution is shown in Figs. 4-5, where the grids are coarsened for clarity. The grid consists of a background orthogonal grid, which mimics the towing tank, and a boundary layer curvilinear grid which conforms to the ship geometry where two clusters of grid points are concentrated around the bow and stern regions. The boundary layer grid is generated with a grid spacing at the hull satisfying the condition  $y^+ < 1$  for all case. All the grids are refined in the vertical direction in  $-0.003 \leq z \leq 0.003$ , where the free surface is expected.



**Fig.4 boundary layer curvilinear grid of the ship model**



**Fig.5 a sketch of the grid distribution**

### Results of Benchmark Test Case

Test case A, designed according to the experiment, is used to confirm the capability of present method in the prediction of the viscous flow around the ship model while maneuvering into the lock.

Comparisons between computed and experimental data is presented for the  $x$ -direction hydrodynamic force  $X$ ,  $y$ -direction hydrodynamic force  $Y$ , the moment around the  $z$ -axis  $N$  and the vertical displacement of the fore and aft perpendiculars  $z_{FP}$  and  $z_{AP}$ .

For a clear insight of the viscous flow around the ship during the maneuvering into the lock, three special instants, shown in Fig. 6, are defined. At time T1, the ship is still moving in the approach channel, while forebody of the ship has entered the lock at time T2. At time T3, the ship has entered the lock.

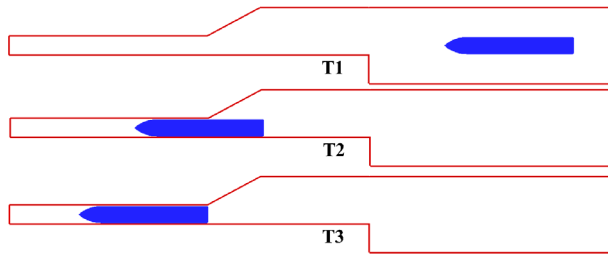


Fig. 6 Definition of three certain times

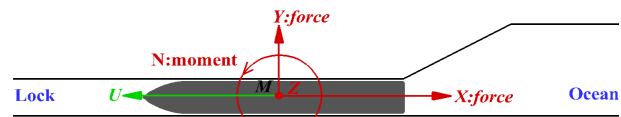


Fig. 7 Coordinate systems for hydrodynamic forces and moment

*Hydrodynamic Forces*

Coordinate systems for hydrodynamic forces and moment is shown in Fig. 7. Fig. 8 presents the comparison of the computed results and experimental data for the time history of hydrodynamic forces and moment, where the  $x$  axis represents the position of the bow. Notice that the  $x$  coordinate values change from 2.5 to -1 when the ship maneuvering from the approach channel into the lock, as the coordinate system, shown in Fig. 3, is located at the door of the lock and the longitudinal  $Ox$ -axis is pointing to the approach channel. The predicted results show good agreement with measured data and exhibits consistent oscillation compared with the experimental data.

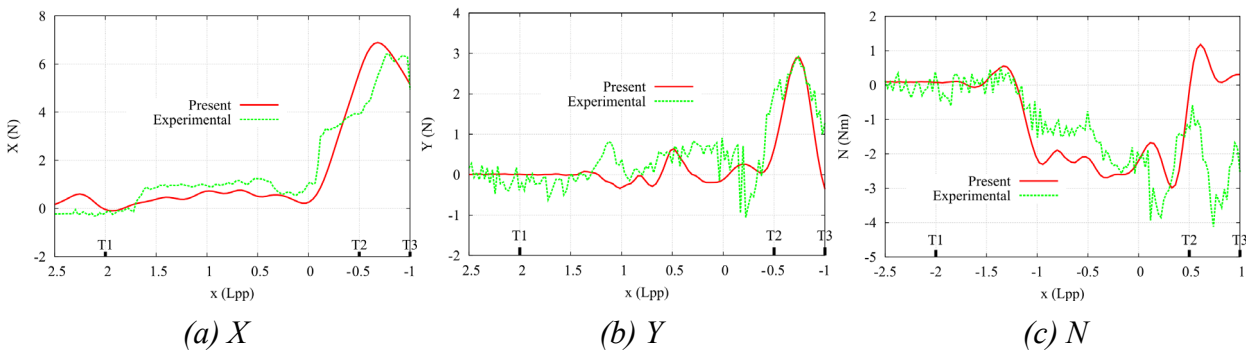


Fig. 8 Time history of the hydrodynamic forces and moment

Taking longitudinal force  $X$  as an example, during maneuverer the ship into the lock, it goes through three stages. In the first stage, the hydrodynamic force  $X$  is very small with insignificant oscillation when the ship maneuvers in the approach channel. This is caused by the extremely low ship speed. Then, a soaring increase of the hydrodynamic force  $X$  is noticed during the ship maneuvering into the lock. During the entry into the lock, an accumulation of water is provoked in front of the ship because of the piston effect. This induces the increase of the bow pressure and the decrease of the stern pressure, which causes a dramatically increase of the viscous pressure resistance. Also, an increased friction resistance will occur because of the higher velocities of the return flow. Accordingly, the longitudinal force  $X$  increases. At the least stage, once the ship has entered the lock, a significant decrease of the resistance force can be noticed. When the ship has entered the lock, water is evacuated out of the lock. As a result, the pressure difference between the fore and aft part of the ship decreases, reducing the oscillations. The hydrodynamic force  $Y$  and the moment  $N$  also shows the same tendency. Bank effect develops in case of an eccentric approach until the ship has entered the lock. The asymmetric flow around the ship induced by the vicinity of

banks causes pressure differences between port and starboard sides. The velocity will increase alongside the nearest bank, causing a pressure decrease. As a result, a lateral force will act on the ship, directed towards the closest bank, as well as a yawing moment pushing the ship bow towards the center of the waterway. Once the ship has entered the lock, the asymmetric flow disappears and the lateral force, as well as the yawing moment falls to zero. This is confirmed by the results shown in Figs. 8b and 8c.

### Vertical Displacements

Piston effect, which provokes an accumulation of water during entry into a lock, causes the increase of the return flow velocity and a sinkage of the water level in the vicinity of the ship, which induces a general sinkage of the ship. The return flow of the ship causes a pressure drop around the ship, according to the Bernoulli principle. As a result, the ship also moves vertically downward. Generally, the pressure drop will not be distributed equally over the ship, causing a trim.

The mean sinkage  $\sigma$  and trim  $\tau$  were determined from the calculated sinkage force  $Z$  and trim moment  $M$  using the formulae (positive sinkage upwards and positive trim bow-up):

$$\sigma = Z / \rho g A_w \quad (15)$$

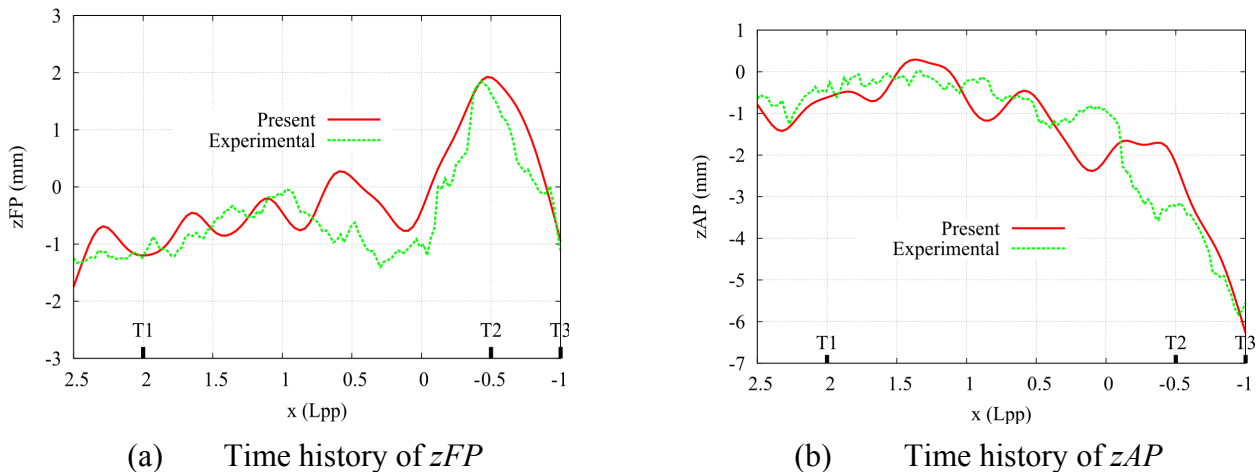
$$\tau = M / \rho g I_w \quad (16)$$

where  $A_w$  denotes the water plane area and  $I_w$  represents the longitudinal moment of inertia of the water plane area about the center of floatation. The vertical displacement of the fore and aft perpendiculars  $z_{FP}$  and  $z_{AP}$  can be computed by:

$$z_{FP} + z_{AP} = 2\sigma \quad (17)$$

$$z_{FP} - z_{AP} = L_{pp}\tau \quad (18)$$

Fig. 9 reports the vertical displacement of the fore and aft perpendiculars. Reasonable correspondence between the computational and experimental data can be noted.



**Fig. 9 Time history of  $z_{FP}$  and  $z_{AP}$**

Three stages in the development of  $z_{FP}$  are shown in Fig. 9a. Firstly, negatively  $z_{FP}$ , which is caused by the shallow water effect, with insignificant oscillation is noticed when the ship maneuvers in the approach channel. This is caused by the low speed. Then, an initial rising of the bow, which might be caused by the accumulation of water in front of the ship, can be noticed when the ship entering the lock. Thirdly, once more than half of the vessel is inside of the lock, the sinkage oscillation of  $z_{FP}$  decreased. This is induced by the evacuation of the water in the lock. When the ship has entered the lock, the sinkage of the bow disappears. Fig. 9b manifest that the vertical displacement of the aft perpendiculars  $z_{AP}$  is increasing as the ship entering the lock. This might be caused by the continuous declination of the surface pressure of the aft-body of the ship.

The increasing of the relative speed between ship and return flow, as well as the sinkage of water level around the stern induce the decrease of the surface pressure of the aft-body of the ship, so that the aft perpendiculars moves vertically downward.

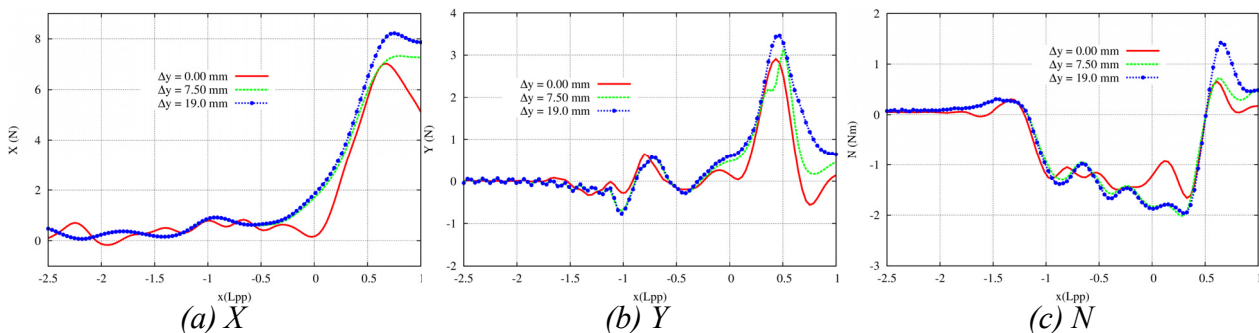
The results of hydrodynamic forces and moments, as well as the vertical displacements show good agreement with the experimental data, which demonstrates that FDM method can accurately simulate the viscous flow around the ship with relatively small amounts grids. The results also indicate that the currently used numerical methods are suitable for studying the viscous flow around the ship during the entering manoeuvre into a lock with different eccentricity, which is presented in the next section.

## Results of Systematic Computations

The effect of eccentricity on the characteristic of the viscous flow around the ship, during the entering manoeuvre of a ship into a lock is significant. In order to give a clear insight into the ship behavior when entry the lock with different eccentricity, case A1 and case A2 were carried out. The details of the test conditions for the systematic study was given in table 2.

### Hydrodynamic Forces

Fig. 10 presents the hydrodynamic forces and moment of the 12000TEU ship model when entry into the lock with different lateral position. All the results show that the eccentricity increases do not alter the general pattern of the forces and moment, but just cause higher extreme values. The results declare that more pronounced prevention of water in the lock from evacuate out can be noticed with increasing eccentricity. So, more water in the lock might be accumulated in the lock, which causes higher velocities of the return flow. Accordingly, the longitudinal force  $X$  increases. Furthermore, the flow around the ship will be more asymmetric with higher eccentricity. As the ships are allowed with very small side margins in the lock, higher eccentricity represents higher risk of collision of ship and side walls. So, protection measures, such as fenders, are suggested.



**Fig. 10 Time history of the hydrodynamic forces and moment for different eccentricity cases**

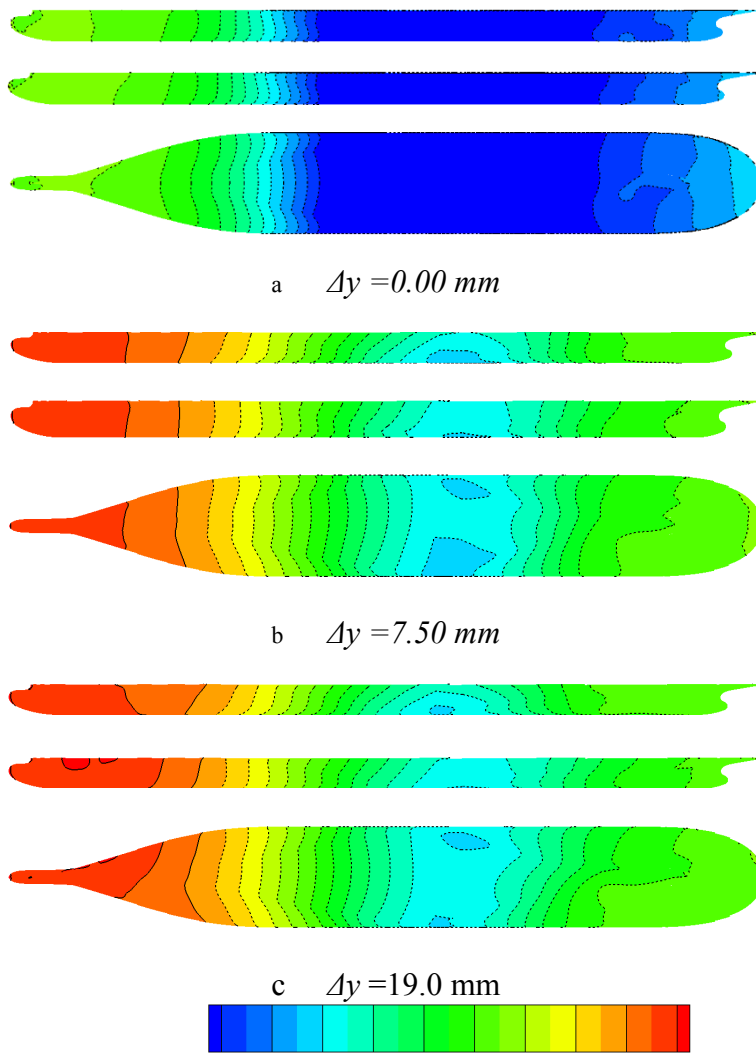
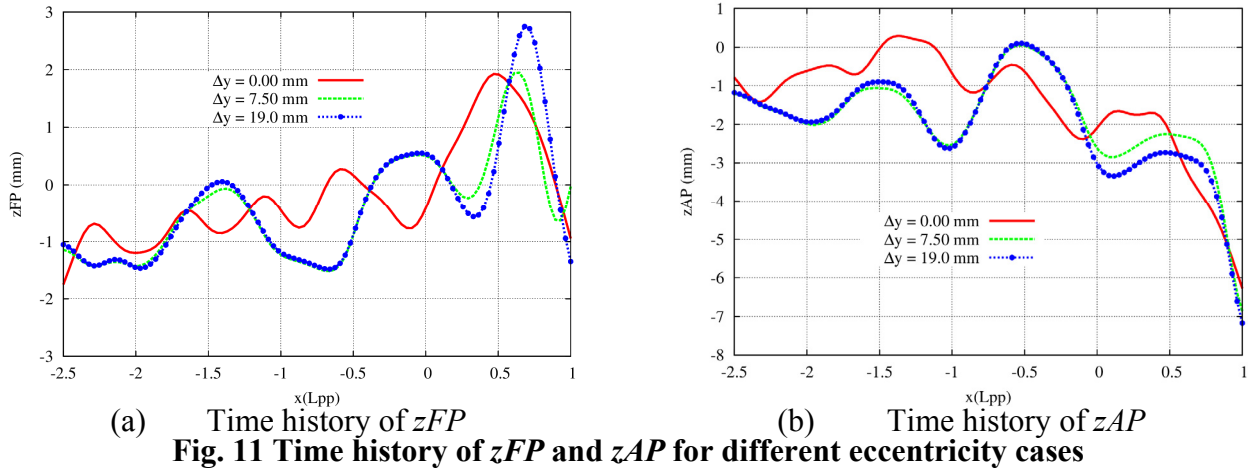
### Vertical Displacements

The vertical displacement of the fore and aft perpendiculars  $z_{FP}$  and  $z_{AP}$  for different lateral position are computed and presented in Fig. 11. Both of  $z_{FP}$  and  $z_{AP}$  increase with increasing eccentricity. This causes higher risk of bottom touch. Due to the high blockage of the lock, especially the shallow water effect, the increase of the risk of bottom touch are noteworthy. According to the results, moving along the lock's centerline is the safest way when maneuvering into a lock.

### Surface pressure distribution

Fig. 12 presents the computed results of the surface pressure at time T3 for different eccentricity cases. Firstly, higher surface pressure around fore-body than that around aft-body is shown in all cases. This is caused by the higher water level elevation in the vicinity of the fore-body. Secondly, the results also manifest that the surface pressure, as well as the asymmetry of the surface pressure increases with increasing eccentricity. The difference of surface pressure between port and starboard sides is mainly concentrated in the aft-body of the hull and the surface pressure on the

starboard side is larger than that on the port side. As a result, a negative yaw moment will turn the hull toward the starboard-side.



**Fig. 12 Pressure (port side view, starboard side view and top view) against different ship speed at time T3**

## Conclusions

Effects of eccentricity on the hydrodynamics of a ship while entering a lock were investigated and reported in this paper. The capability of the present method for the prediction of the viscous flow around the ship model when entering the lock is confirmed by the good agreement of the predicted results with the corresponding experimental data. The hydrodynamic forces, vertical displacements were presented and discussed. A significant effect of high blockage on the hydrodynamic forces and moment, as well as ship sinkage was noticed. Furthermore, FDM method is especially suitable for solving low speed headache with remarkable free surface effects.

A systematic investigations were then made to examine the effects of the eccentricity on the hydrodynamics on the ship when it enter a lock. According to the results, several interesting observations can be made:

1. Hydrodynamic forces moment, as well as the sinkage of the ship are sensitive to the eccentricity.
2. Hydrodynamic forces and moment, as well as the sinkage of the ship, increase with increasing eccentricity. This causes the increases of the difficulty to manoeuvre a ship into a lock, safely.
3. When the eccentricity increases, an increase of the risk of side collision is observed.
4. When entering a lock, moving along the lock's centerline and protection measures is suggested.

## Acknowledgement

This work is supported by National Natural Science Foundation of China (Grant Nos. 51379125, 51490675, 11432009, 51411130131, 11272120), The National Key Basic Research Development Plan (973 Plan) Project of China (Grant No. 2013CB036103), High Technology of Marine Research Project of The Ministry of Industry and Information Technology of China, Chang Jiang Scholars Program(Grant No. T2014099) and the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning(Grant No. 2013022), to which the authors are most grateful.

## References

- Chen, X. N. (2010) From water entry to lock entry, *Journal of Hydrodynamics*, **Ser. B**, 22(5), 885-892.
- Delefortrie, G., Willems, M., et al. (2008) Tank test of vessel entry and exit for third set of Panama locks, *In Proceedings of the International Navigation Seminar following PIANC AGA*, 517-523.
- Delefortrie, G., Willems, M., et al. (2009) Behavior of post panamax vessels in the Third Set of Panama locks, *In International Conference on Marine Simulation and Ship Maneuverability (MARSIM'09)*.
- Meng, Q. J., & Wan, D. C. (2014) Numerical Simulations of Ship Motions in Confined Water by Overset Grids Method, *In The Twenty-fourth International Ocean and Polar Engineering Conference*.
- Menter, F. R. (1994) Two-equation eddy-viscosity turbulence models for engineering applications, *AIAA journal*, **32(8)**, 1598-1605.
- Osher, S., & Sethian, J. A. (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, *Journal of computational physics*, **79(1)**, 12-49.
- Richter, J., Verwilligen, J., et al. (2012) Analysis of full ship types in high-blockage lock configurations, *In MARSIM 2012*, 1-9.
- Sethian, J. A., & Smereka, P. (2003) Level set methods for fluid interfaces. *Annual Review of Fluid Mechanics*, **35(1)**, 341-372.
- Sussman M, Smereka P, Osher SJ. (1994) A level set approach to computing solutions to incompressible two-phase flow, *Journal of Computational physics*, **114(1)**, 146-159.
- Vantorre, M., Delefortrie, G., Mostaert, F. (2012) Behaviour of ships approaching and leaving locks: Open model test data for validation purposes. Version 2\_1, *WL Rapporten*, **WL2012R815\_08e**, Flanders Hydraulics Research and Ghent University - Division of Maritime Technology: Antwerp, Belgium.
- Vergote, T. (2012) Hydrodynamics of a ship while entering a lock. *Eng. thesis*, Ghent Univ., Ghent, Belgium.
- Vrijburcht, A. (1988) Calculations of wave height and ship speed when entering a lock, *Delft Hydraulics Publication* **391**, 1-17.
- Wang, H. Z., & Zou, Z. J. (2014) Numerical study on hydrodynamic interaction between a berthed ship and a ship passing through a lock, *Ocean Engineering*, **88**, 409-425.

## SSRLS based Enhanced Impulsive Noise OFDM Suppressor in AWGN Channel

Alina Mirza, Sumrin Mehak Kabir, and Shahzad Amin Sheikh

Department of Electrical Engineering, College of Electrical & Mechanical Engineering (CEME)  
National University of Sciences and Technology (NUST), Pakistan.

[alina.mirza78@ceme.nust.edu.pk](mailto:alina.mirza78@ceme.nust.edu.pk), [sumrin.mehak75@ee.ceme.edu.pk](mailto:sumrin.mehak75@ee.ceme.edu.pk), [sheikh.shahzadamin@gmail.com](mailto:sheikh.shahzadamin@gmail.com)

### Abstract

Orthogonal Frequency division multiplexing is a popular method for high data rate applications and is corrupted by impulsive noise. Methods must be investigated to mitigate this noise. In this paper, a State Space Recursive Least Square (SSRLS) algorithm based adaptive impulsive noise suppressor ofdm communication is proposed. It gives the better tracking due to its state space model-dependent recursive parameters. The proposed method came out to be very effective in noise cancellation of OFDM signal without requiring reference noise source. The fastest convergence and better tracking characteristics of proposed scheme demonstrated by the simulation results in mean square error (MSE) sense proved to be the effective solution for the impulsive noise cancellation in OFDM signal.

**Keywords:** Impulsive Noise, OFDM, cyclic prefix, NLMS, RLS, SSRLS, MSE

### Introduction

Orthogonal Frequency Division Multiplexing (OFDM), with their ability to mitigate the effects of multipath, makes them most suitable for the transmission over high data rate in wireless communications [Ghorpade et al (2013); Krishnamoorthy et al (2013)]. Therefore numeric wireless standards of audio/video broadcasting fourth generation mobile systems particularly WIMAX and LTE have been used in OFDM systems over the past decade [Zhou and Xie (2007)]. Orthogonal Frequency division multiplexing is suspected to have impulsive noise, which completely destroys the information. Practically, impulsive noise is a non-Gaussian noise generated by human activities and has more catastrophic effects in communication systems[Torio (2011)]. Nowadays active area of research is to inspect the impulsive noise behavior and suggest solutions to improve the performance of systems by suppressing it. For noise cancellation, various techniques are reported in literature which attempt to recover the original transmitted signal by [Jimaa et al (2012); Araj (2011)].

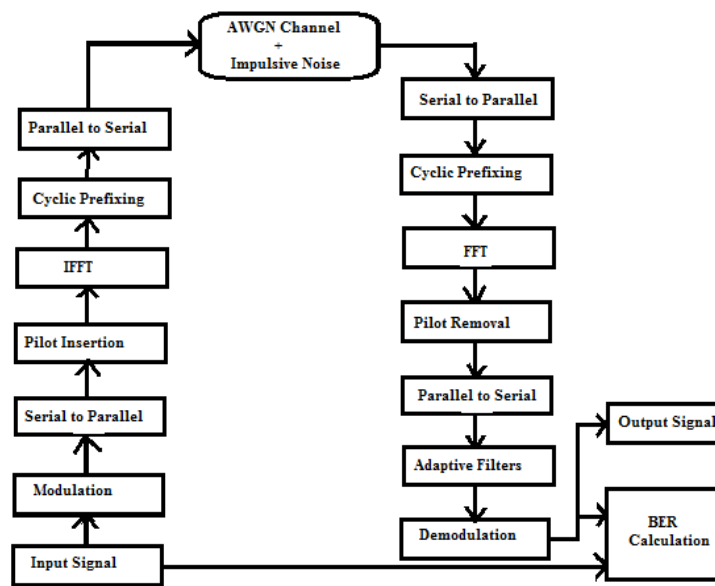
[Mathew and Murukan (2014)] carried out the comparative bit error rate analysis of adaptive notch and Least Mean Square algorithm to suppress the Periodic impulsive noise from OFDM based power line communication systems. An adaptive receiver technique based on NLMS, RLS, VSNLMS filters are used to remove impulsive noise from the MIMO-OFDM system by [Hakam et al, (2012)].An impulsive noise canceller of sinusoidal and ECG signal based on SSRLS filter in time domain has been proposed by [Mirza et al (2015)]. In addition comparison of SSRLS with RLS and NLMS adaptive algorithm was carried out. Motivated by the results obtained by [Mirza et al (2015)] an impulsive noise suppressor for OFDM system is designed and implemented in this paper.



This paper is organized as: Section II briefly explain the basic mechanism of OFDM, Section III gives the review of different adaptive filters which is followed by comparative analysis supported with the simulation results in section-IV. In the end, section V then concludes the paper and is then followed by the references.

**OFDM**

Orthogonal Frequency Division Multiplexing is a multicarrier modulation in which transmission over a dispersive channel is carried out .In OFDM the high data rate streams are splitted into low data rate streams in parallel and modulated separately on different orthogonal sub-carriers. The introduction of pilot insertion and cyclic redundancy at the transmitter reduces the complexity to only Fast Fourier Transform FFT processing on the receiver side.



**Figure 1. Block Diagram of OFDM system**

These subcarriers are multiplexed and passed through the channel, which is responsible for adding impulsive noise and white Gaussian noise in the transmitted OFDM signal. At the receiver side, the signal is demodulated and passed through the adaptive filter block for impulsive noise reduction in the OFDM signal. The parameters used in simulating the OFDM system are tabularized below.

**Table 1: Parameter set for simulation of OFDM system**

Parameters	Values
Modulation technique	QPSK
Number of subcarriers	52
Size of cyclic prefix	16
FFT-length	64
Number of bits generated	52000

## Adaptive Algorithms

There are many adaptive algorithms used for noise removal. The brief summaries of adaptive algorithms which are used in this research are as follow.

### *RLS Algorithm*

The Gauss Recursive least squares (RLS) adaptive filter is the one in which autocorrelation matrix estimation is used to de-correlate the current input data. It recursively finds the filter coefficients that is then used to minimize a weighted linear least squares cost function relating to the deterministic input signals. Also, The RLS exhibits extremely fast convergence over all variants of LMS but with a cost of high computational complexity. The filter weights  $w$  are updated in RLS algorithm by following equations.

$$w(n+1) = w(n) + k(n)x(n) \quad (1)$$

$$k(n) = \frac{\lambda^{-1}\Phi^{-1}(n-1)x(n)}{1 + \lambda^{-1}x^T(n)\Phi^{-1}(n-1)x(n)} \quad (2)$$

$$\Phi^{-1}(n) = \lambda^{-1}\Phi^{-1}(n-1) - \lambda^{-1}k(n)x^T(n)\Phi^{-1}(n-1) \quad (3)$$

Where  $\lambda$  is the forgetting factor.  $\Phi^{-1}$  is the cross correlation matrix. The  $\lambda$  is initialized with 1 and  $\Phi^{-1}$  with  $\delta^{-1}I$ .  $I$  is the identity matrix.

### *SSRLS Algorithm*

State Space Recursive Least Squares or SSRLS algorithm is state space representation of an extension of RLS algorithm. It is used to remove noise and its performance can be evaluated in a non-stationary environment (impulsive noise). The steps of SSRLS form II filter along with sinusoidal model for implementation are given by [Malik (2004)].

$$\hat{x}[n] = \hat{x}[n] + K[n]\varepsilon[n] \quad (4)$$

$$\hat{x}[n] = A\hat{x}[n-1] \quad (5)$$

$$\varepsilon[n] = y[n] - \bar{y}[n] \quad (6)$$

$$\bar{y}[n] = Cx[n] \quad (7)$$

$$\Phi[n] = \lambda(A^{-T}\Phi[n-1])A^{-1} + C^T C \quad (8)$$

$$K[n] = \Phi^{-1}(n)C^T \quad (9)$$

Where  $\hat{x}[n]$  is the input state,  $\varepsilon[n]$  is the prediction error,  $K(n)$  is observer gain,  $n$  is predicted input state,  $\hat{n}$  is estimated state,  $\bar{y}[n]$  is the predicted output state and  $\Phi[n]$  is the correlation matrix.

*NLMS Algorithm*

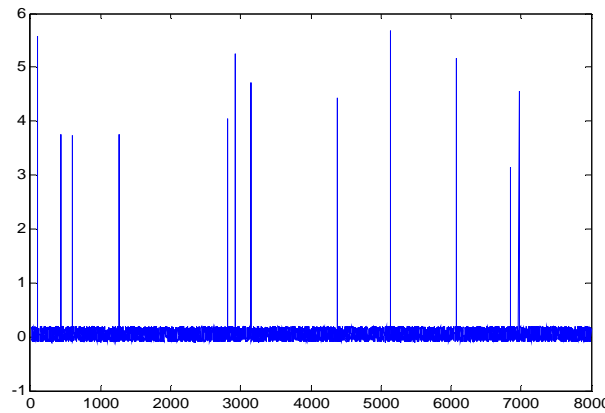
The variant of Least Mean Square (LMS) algorithm is Normalized Least Mean Square which gives faster convergence than LMS. The limitation of the LMS algorithm is its sensitivity to its input signal scaling. The convergence is very slow and step size should be chosen carefully to guarantee algorithm stability. The whole algorithm remains same only filter tap weights are updated by following recursive formula:

$$w(n+1) = w(n) + \frac{\mu e(n)x(n)}{\epsilon + \|x(n)\|^2} \tag{10}$$

Where  $\epsilon$  a small number is added for algorithm stability,  $\mu$  is the step size of filter and  $e(n)$  is error signal.

**Simulation Results**

In this section, we compare the performance of the different adaptive filters in impulsive noise cancellation of OFDM signal by computer simulation using MATLAB version 12.

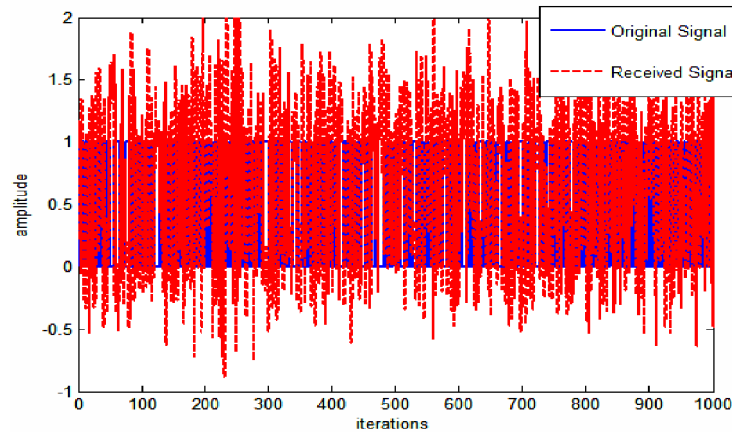


**Figure 2. Impulsive Noise Signal**

**Table 2. Parameter set for simulation of Impulsive Noise**

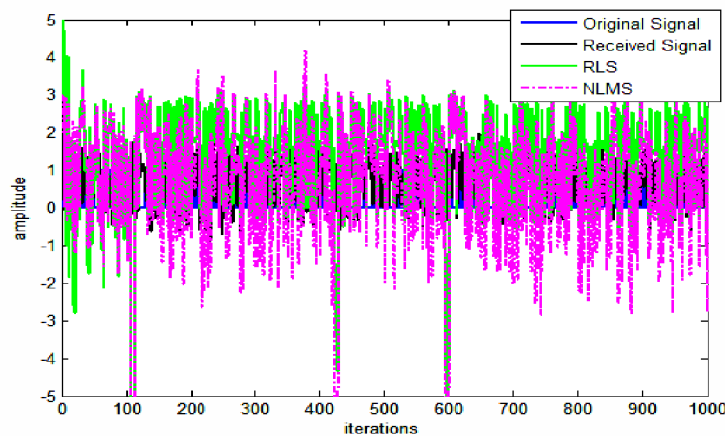
Parameters	Symbol	Value
Sampling Frequency	$F$	10
Total time	$T$	100
Average Time between samples	$\beta$	1s
Mean of log amplitude	$A$	10dB
Standard deviation of log amplitude	$B$	5dB
Mean of Additive Gaussian Noise	$m$	0.1
Standard deviation of Gaussian Noise	$\sigma$	0.4

The generated binary data is passed through the channel responsible for adding impulsive noise and white Gaussian noise are depicted in Fig. 3.



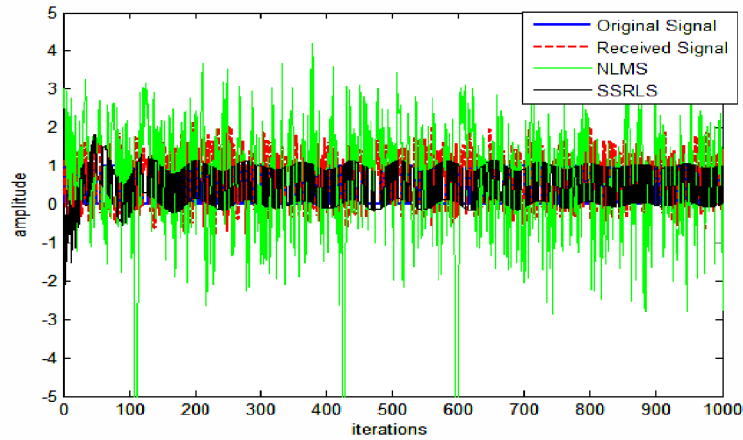
**Figure 3. Original and received signal**

The system output error signal should contain the original signal in an optimum sense. The length of all the three adaptive filters is fixed to 10. The step size parameter for NLMS Algorithm is chosen to be equal to 0.005 and forgetting factor for RLS is 1 and for SSRLS is 0.99. The error signals obtained by above mentioned adaptive filters are compared with one another in Fig. 4-6.



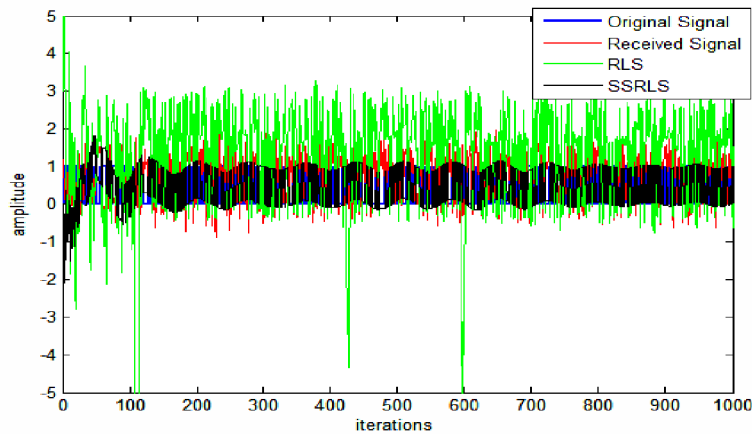
**Figure 4. Comparison of original data, received data and recovered data using NLMS and RLS Filters**

Fig.4 represents that the largest peaks of impulsive noise from the noisy binary signal are not properly removed by the RLS and NLMS algorithm. The RLS filter is removing impulsive noise better than NLMS. The error plots of above mentioned algorithms are also compared with the original signal and received signal.



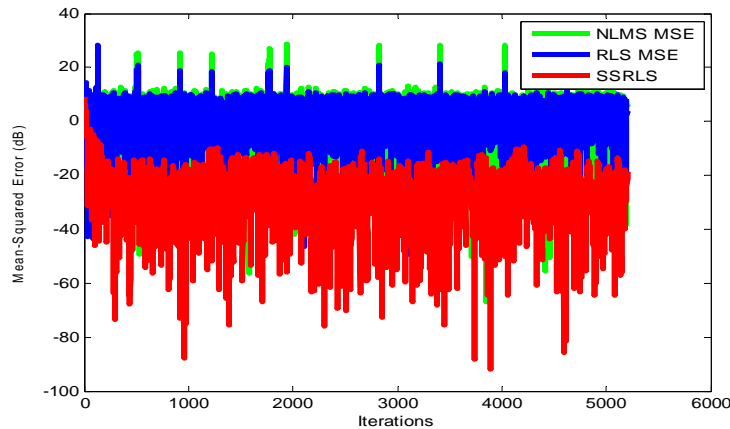
**Figure 5. Comparison of original data, received data and recovered data using NLMS and SSRLS Filters**

The results of Fig. 5 illustrate that superior performance of SSRLS in suppressing impulsive noise from the noisy OFDM signal as compared to the NLMS adaptive algorithm. The error plots of above mentioned algorithms are also compared with the original signal and received signal.



**Figure 6. Comparison of original data, received data and recovered data using RLS and SSRLS Filters**

Similarly the comparison of SSRLS and RLS filter error plots while cancelling impulsive noise from the OFDM signal are shown in Fig.6. It is clear from the above Figures that the SSRLS exhibit better performance in cancelling the largest peaks of impulsive noise from the OFDM signal, while other two investigated algorithms fail to remove the noise with large amplitudes.



**Figure 7. Comparison of MSE (dB) of adaptive filters**

The mean square error in terms of decibel simulation results also confirms that SSRLS give lowest MSE and fastest convergence while cancelling impulsive noise as depicted in Fig. 7. The conclusion drawn from Fig. 4-7 is that SSRLS filter outperforms both NLMS filter and RLS in Impulsive noise cancellation of OFDM signal.

## Conclusion

In this paper, an adaptive impulsive noise suppressor for OFDM system has been proposed that is based on state space recursive least square (SSRLS) algorithm. Due to the state space dependent model, the proposed technique exhibit better impulsive noise cancellation in OFDM signal when compared to normalized Least Square (NLMS) and Recursive Least Square (RLS). The simulation results obtained by proposed enhanced impulsive noise suppressor guarantees the superior performance of SSRLS in terms of convergence speed and lowest MSE.

## References

- S. S. Ghorpade., et al, (2013) Behaviour of OFDM System using MATLAB Simulation, *International Journal of Advanced Computer Research* **3**, issue **10**.
- R. Krishnamoorthy., et al, (2013) Forward Error Correction Code for MIMO-OFDM System in AWGN and Rayleigh Fading Channel, *International Journal of Computer Applications* **69**.
- Wen Zhou., Bing Xie. (2007) Link level simulation and performance estimation of WIMAX IEEE 802.16e, *2<sup>nd</sup> International Conference on Pervasive Computing and Applications*, 667– 671.
- P.Torio., M.G. Sanchez., I. Cuinas. (2011) An algorithm to simulate impulsive noise, *19th International conference on Software, Telecommunications and Computer Networks*, 1 – 4.
- S. A.Jimaa., et al. (2012) Impulsive noise reduction using adaptive receiver structure technique, *IEEE 11<sup>th</sup> International Conference on Signal Processing*, Beijing ,China..
- S. R. Al-Araji., M. A. Al-Qutayri., M. S. Al-Tenaiji, (2011) Impulsive noise reduction using Auto-Gating technique, *IEEE GCC Conference and Exhibition (GCC)*, 104 – 107.
- S.Mathew., et al. (2014) Periodic Impulsive Noise Reduction in OFDM based Power line Communication, *International Journal of Research in Engineering and Technology* **3**, 517 – 522.
- A.Hakam., et al, (2012) Impulsive Noise Reduction in MIMO-OFDM Systems using Adaptive Receiver Structures, *11th International Conference on Signal Processing (ICSP 2012)*, Abu Dhabi, UAE.
- A.Mirza., et al, (2015) Impulsive Noise Cancellation of ECG signal based on SSRLS, *The 2015 International Conference on Soft Computing and Software Engineering (SCSE 2015)*, USA.
- M.B.Malik. (2004) State-space recursive least-squares: part I & II, *Signal Processing Journal* **84**, 1709 – 1728.

## Spatial scaling issues in constitutive modelling of geomaterials

†C.T. Nguyen<sup>1</sup>, \*G.D. Nguyen<sup>1</sup>, H.H.Bui<sup>2</sup>, and V.P. Nguyen<sup>1</sup>

<sup>1</sup>School of Civil, Environmental and Mining Engineering, University of Adelaide, Australia.

<sup>2</sup>Department of Civil Engineering, Monash University, Australia

\*Presenting author: g.nguyen@adelaide.edu.au

†Corresponding author: chi.nguyen@adelaide.edu.au

### Abstract

Most geomaterials exhibit localised modes of failure, which appear in the form of shear bands or cracks. This appearance could be gradual or very abrupt, depending on the type of materials and also loading conditions. In particular, failure of rocks or concrete is very abrupt in tension and indicates the brittle nature of these materials under such conditions, while in compression regime and under increasing confining conditions, their failure gradually becomes more ductile with the onset and propagation of shear bands. The orientation and size of shear bands and/or fracture process zones in localised failure vary with the loading and materials. As a consequence, a correct description of geomaterial behaviour must take into account all these characteristics of failure. This is however not always the case in constitutive modelling of geomaterials, and while working well under homogeneous condition, most (if not all) existing models do not possess details on the size and orientation of localisation zones when localised failure takes place. This prevents them from capturing correctly the failure behaviour of the materials, the most important characteristic of which is size effect. We present in this study a new approach to incorporating both size and orientation of localisation zone in constitutive models for geomaterials. The concepts together with technical details, and preliminary results are presented to show the potentials of the new approach.

**Keywords:** Localisation, Size effects, Geomaterials, Constitutive Modelling

### Introduction

Constitutive models play a key role in correctly predicting load carrying capacity and failure of materials and structures. They give a mathematical description of material behaviour under different loading conditions, and it is essential that this description follows closely the observed behaviour of materials in experiments and/or real life. This requires not only advanced experimental techniques for better understanding of material behaviour, but also a generic and strong theoretical framework to transfer this understanding to a constitutive model. In geomaterial modelling, while the experimental part has been advancing quickly in the last few decades, with sophisticated techniques such as X-Ray and Digital Image Correlation [Alshibli and Sture (1999); Desrues and Viggiani (2004), Réthoré et al (2007)] for the observation of failure initiation at the micro scale, in our opinion, the development of a theoretical framework to accommodate these experimental development is still lagging behind. Despite the developments of more and more advanced constitutive models, one of the most important features of geomaterial failure - localised failure mode and the associated size effect - is still an outstanding issue in constitutive models for geomaterials.

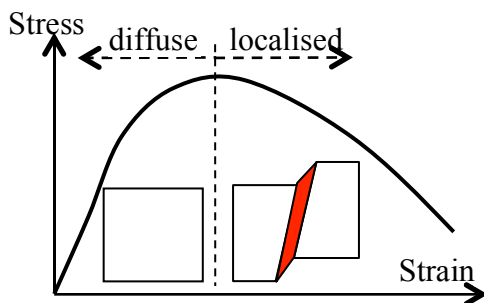
Localisation of deformation appears in the form of thin bands such as shear bands in soils or fracture process zones in concrete/rock under tensile loadings. The material outside a localisation band usually unloads elastically, giving the indication that most (or all) inelastic behaviour happens

inside this thin band. In addition, while the orientation of a localisation band depends on loading conditions [Runesson and Ottosen (1991)], the size and behaviour of the material inside this band are material characteristics. It can be seen that classical continuum models utilise a single stress-strain relationship under both diffuse and localised modes of failure; this is not adequate, due to lack of details on the size and behaviour of the localisation band to capture correctly localised failure of the material. A correct description of failure must, therefore, account for the transition from diffuse to localised mode, and the progression of material failure beyond that. In this sense, most existing models ignore localised failure and size effect issues, either partly or completely, in the model developments and interpretation/mapping of experimental data. In particular, they possess no details on the size, orientation and evolution of the localisation band. As a consequence, their behaviour does not scale with the volume, and hence size effects cannot be captured. This requires ad hoc treatments incorporated later, after the model development, should these models be used in the analysis of failure that involves numerical methods for the solutions of boundary value problems. Typical examples of these treatments are the nonlocal/gradient regularisation [Bažant and Lin (1988); de Borst and Mühlhaus (1992)], and viscous regularisation [Forest, E. Lorentz (2004)], with several corresponding applications in geomaterial modelling [Pedersen et al (2008); Nguyen and Einav (2010); Das et al (2013)]. Other examples to overcome deficiencies of existing continuum models in capturing localised failure and size effects involve enhancements to the discretisation techniques (e.g. finite element methods) to correctly describe the kinematics of localised failure and behaviour.

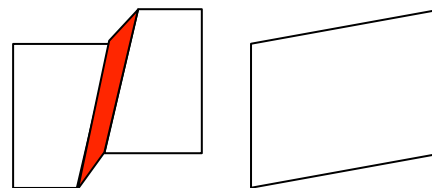
In this study, we start from a key characteristic of localisation in constitutive modelling of geomaterials: kinematics of localised failure, and enhancement is developed after enriching the kinematics of constitutive models. In this sense, an enriched constitutive model will possess more than one stress-strain relationships to correctly describe different material behaviours in- and outside the localisation zone. Size effects are automatically taken into account thanks to the intrinsic material length scale in the framework. A new formulation will be established and applied to different kinds of existing constitutive models. Numerical examples at the constitutive levels are used to illustrate key features of the new theoretical framework.

### Kinematic enrichment at the constitutive level

Localised deformation takes place at certain stage of the deformation process (Fig. 1), with the onset and propagation of localisation band, where inelastic response occurs. This leads to incorrect representation of the deformation if a single homogeneous strain is used, as in classical continuum models (Fig. 2). In other words, the strain in the volume where constitutive model is defined is no longer homogenous, while homogenised deformation is considered only appropriate inside the localisation zone and the bulk material outside it.



**Figure 1. Diffuse and localised stages.**



**Figure 2. Observed localised failure and classical assumption of homogenous deformation.**



In this section, the connection between the averaged (macro) stress  $\boldsymbol{\sigma}$  and averaged (macro) strain  $\boldsymbol{\varepsilon}$  over the domain  $\Omega$  crossed by a localisation band (Figs. 1 & 2) is derived by coupling the response of the material inside and outside the localised region. For the volume of consideration, the stresses and strains for materials inside the localisation band and outside the homogeneous bulk are denoted as  $(\boldsymbol{\sigma}_i, \boldsymbol{\varepsilon}_i)$  and  $(\boldsymbol{\sigma}_o, \boldsymbol{\varepsilon}_o)$  respectively.

As can be seen in Fig. 2, two strain measures and correspondingly two stresses are required to capture correctly localised deformation, while for backward compatibility with existing numerical codes that takes a single stress-strain relationship, a macro stress-strain relationship is needed. For this, the macro strain rate is defined as the volume averaged quantity:

$$\dot{\boldsymbol{\varepsilon}} = f \dot{\boldsymbol{\varepsilon}}_i + (1 - f) \dot{\boldsymbol{\varepsilon}}_o \quad (1)$$

in which  $f$ , the volume fraction of the localisation zone, is defined as the ratio between the width  $h$  of the localisation zone and the effective size  $H$  of the volume element:

$$f = \frac{hA}{\Omega} = \frac{h}{H} \quad (2)$$

where  $A$  is the surface area of the localisation zone. The strain inside the localisation zone takes the following form (Nguyen et al, 2012):

$$\dot{\boldsymbol{\varepsilon}}_i = \frac{1}{h} (\mathbf{n} \otimes [\dot{\mathbf{u}}])^s \quad (3)$$

where  $\mathbf{n}$  denotes the normal vector of the band and  $[\dot{\mathbf{u}}]$  is the relative velocity between opposite sides of the localisation band. We note that homogeneous term was ignored in the above equation, on the assumption that the deformation inside the localisation zone is usually of higher order of magnitude than that in the bulk volume containing it [Vardoulakis et al (1978), Vardoulakis and Graf (1985a; 1985b)]. This simplifies the formulation and also makes it adaptable to other cases of quasi-brittle modelling (Nguyen et al, 2014). The macro stress in this case cannot assume a volume averaged form, but needs to be worked out from the virtual work equation in the following form:

$$\boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} = f \boldsymbol{\sigma}_i : \dot{\boldsymbol{\varepsilon}}_i + (1 - f) \boldsymbol{\sigma}_o : \dot{\boldsymbol{\varepsilon}}_o \quad (4)$$

Substituting (1) and (3) into the above and after some rearrangement, we have:

$$\frac{f}{h} (\mathbf{t} - \mathbf{t}_i) \cdot [\dot{\mathbf{u}}] + (1 - f) (\boldsymbol{\sigma} - \boldsymbol{\sigma}_o) : \dot{\boldsymbol{\varepsilon}}_o = \mathbf{0} \quad (5)$$

in which  $\mathbf{t} = \boldsymbol{\sigma} \cdot \mathbf{n}$  and  $\mathbf{t}_i = \boldsymbol{\sigma}_i \cdot \mathbf{n}$  are the tractions associated with the macro stress and the stress inside the localisation zone, respectively.

Since the strain rate and velocity jumps are arbitrary the following conditions are obtained:

$$\mathbf{t} = \mathbf{t}_i, \text{ and } \boldsymbol{\sigma} = \boldsymbol{\sigma}_o \quad (6a,b)$$

Given

$$\dot{\boldsymbol{\sigma}}_o = \mathbf{D}_o : \dot{\boldsymbol{\varepsilon}}_o, \text{ and } \dot{\boldsymbol{\sigma}}_i = \mathbf{D}_i^T : \dot{\boldsymbol{\varepsilon}}_i \quad (7a,b)$$

as generic constitutive behaviours for the macro volume element and the localisation zone, with  $\mathbf{D}_o$  and  $\mathbf{D}_i^T$  being the corresponding tangent stiffnesses, using equations (1, 3, 6), we can solve for the velocity jump as:

$$[\dot{\mathbf{u}}] = \underbrace{\left[ \frac{f}{h} (\mathbf{n} \cdot \mathbf{D}_o \cdot \mathbf{n}) + \frac{1-f}{h} (\mathbf{n} \cdot \mathbf{D}_i^T \cdot \mathbf{n}) \right]^{-1}}_c \cdot (\mathbf{D}_o : \dot{\boldsymbol{\varepsilon}}) \cdot \mathbf{n} \quad (8)$$

Substituting (1, 3 and 8) into (7a) and rearranging the obtained expression, we get the macro stress-strain relationship in the rate form as:

$$\dot{\boldsymbol{\sigma}} = \frac{1-f}{h} \mathbf{D}_o : \left\{ \dot{\boldsymbol{\varepsilon}} - \frac{1}{H} [\mathbf{n} \otimes (\mathbf{C}^{-1} \cdot (\mathbf{D}_o : \dot{\boldsymbol{\varepsilon}}) \cdot \mathbf{n})]^s \right\} \quad (9)$$

in which tensors  $\mathbf{A}_o = \mathbf{n} \cdot \mathbf{D}_o \cdot \mathbf{n}$  and  $\mathbf{A}_i = \mathbf{n} \cdot \mathbf{D}_i^T \cdot \mathbf{n}$  are the acoustic tensors associated with the tangent stiffness outside and inside the localisation band, respectively. Further details on the approach can be found in Nguyen et al (2012 & 2014).

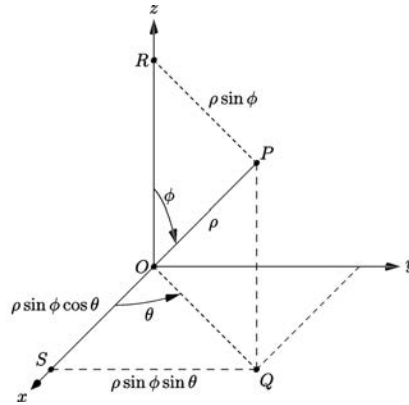
### Numerical implementation

As mentioned in the previous section, the enriched approach focuses on the post-localisation behaviour which means it is active only when localised failure has been detected. Bifurcation analysis, therefore, is of crucial importance and is presented in this section. Later, the implicit stress update algorithms for the enriched approach is given.

#### *Onset and orientation of localisation band*

The deformation behaviour of geomaterials under some specific circumstances can bifurcate from homogeneous to inhomogeneous localisation. There are several equivalent methods used for detecting that bifurcation point [Rudnicki and Rice (1975); Ottosen and Runesson (1991); Nielsen and Schreyer (1993)]. The loss of positive definiteness of the acoustic tensor is used in this paper as the indication of the onset of a localisation band. The normal vector  $\mathbf{n}$  determines the orientation of the localisation band and also indicates the occurrence of the discontinuous bifurcation by the loss of positive definiteness of the acoustic tensor  $\mathbf{A}$  associated with the tangent stiffness  $\mathbf{D}$  [Schreyer and Nielsen (1996); Chambonet al. (2000)]:

$$\det(\mathbf{A}) = \det(\mathbf{n} \cdot \mathbf{D} \cdot \mathbf{n}) \leq 0 \quad (10)$$



**Figure 3. A spherical coordinates, after Wolfram**

Using a standard spherical coordinates and define  $\theta$  to be the azimuthal angle in the  $xy$ -plane from the  $x$ -axis with  $0 \leq \theta \leq 2\pi$  and  $\phi$  the zenith angle from the  $z$ -axis with  $0 \leq \phi \leq \pi$ , Fig. 3, the unit direction vector can be defined as following:

$$\mathbf{n} = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{bmatrix} \sin \phi \cdot \cos \theta \\ \sin \phi \cdot \sin \theta \\ \cos \phi \end{bmatrix} \quad (11)$$

The matrix form of  $\mathbf{n}$ , derived from the equation used to calculate the traction vector from the stress vector, is:

$$\mathbf{n} = \begin{bmatrix} n_x & 0 & 0 \\ 0 & n_y & 0 \\ 0 & 0 & n_z \\ n_y & n_x & 0 \\ 0 & n_z & n_y \\ n_z & 0 & n_x \end{bmatrix} \quad (12)$$

For the case of incrementally multi-linear models such as classical elasto-plastic models, a wide range of orientations can be physically admitted [Ottosen and Runesson (1991); Chambonet al. (2000)] hence it is determined through minimization of  $\det(\mathbf{A})$  provided that  $\det(\mathbf{A}) \leq 0$  [Sanborn and Prevost (2011)].

#### *Implicit stress return algorithm*

After the bifurcation point, a post-localisation analysis is performed with two distinct solutions of stresses and strains for the zones inside and outside the band. In this approach the macro stress coincides with the stress outside the localisation zone, whilst the macro strain is computed from the inside and outside strains and the volume fraction  $f$ . This section presents the stress update algorithm used in the enhanced approach.

Following the implicit algorithm, the traction equilibrium across the discontinuity plane is rewritten in the residual form as following:

$$\mathbf{r} = \boldsymbol{\sigma} \cdot \mathbf{n} - \boldsymbol{\sigma}_i \cdot \mathbf{n} \quad (13)$$

then the iterative scheme is performed so that the norm of this residual is zero, indicating a converged solution in which the equilibrium across the discontinuity plane has been met. A first order Taylor expansion of the above equation at the state of the last iteration can be written as:

$$\mathbf{r}^{new} = \mathbf{r}^{old} + \delta \boldsymbol{\sigma} \cdot \mathbf{n} - \delta \boldsymbol{\sigma}_i \cdot \mathbf{n} \quad (14)$$

After some arrangement the following relation between the residual and the increment displacement jump is obtained:

$$\mathbf{r}^{new} = \mathbf{r}^{old} - \left[ \frac{1-f}{1-f} \frac{f}{h} (\mathbf{n} \cdot \mathbf{D}_o \cdot \mathbf{n}) + \frac{1}{h} (\mathbf{n} \cdot \mathbf{D}_i^T \cdot \mathbf{n}) \right] \cdot \delta[\mathbf{u}] \quad (15)$$

Enforcing the requirement of  $\mathbf{r}^{new} = 0$  leads to:

$$\delta[\mathbf{u}] = \left[ \frac{1}{1-f} \mathbf{C} \right]^{-1} \cdot \mathbf{r}^{old} \quad (16)$$

Once  $\delta[\mathbf{u}]$  is calculated, the iterative strain inside the localisation band is computed as:

$$\delta \boldsymbol{\varepsilon}_i = \frac{1}{h} (\mathbf{n} \otimes \delta[\mathbf{u}])^s \quad (17)$$

then the stress increment  $\delta \boldsymbol{\sigma}_i$  is updated along with internal variables for the material inside the band and the tangent stiffness  $\mathbf{D}_i$ . The convergence criteria is defined as:

$$\frac{\|\mathbf{r}^{old}\|}{\|\boldsymbol{\sigma} \cdot \mathbf{n}\|} \leq \text{TOLERANCE} \quad (18)$$

with TOLERANCE is a sufficiently small positive number.

## A constitutive model based on breakage mechanics

The theoretical aspects and numerical implementation of the enriched approach are given in the previous sections. This section presents a specific constitutive model which will be used with the enriched approach to illustrate its capabilities. A breakage constitutive model [Einav (2007a)] is used to describe the behaviour of the material inside the localisation zone. First, a brief description of the breakage model is given including the stress-strain relationship, yield condition and tangent stiffness tensor; then performances of the model with and without using the enriched approach are given. For the sake of simplicity, the responses obtained without using the enriched approach are named “homogeneous”, whereas the responses obtained using the enriched approach are named “localised” indicating that the localisation behaviour has been taken into account.

### Model description

The breakage mechanics theory for crushable granular materials [Einav (2007a)] is built on the micromechanics of grains, using statistical homogenisation to upscale the grain-scale energy potential to obtain the macro energy potential of the continuum model. A simple breakage model proposed by Einav [2007b] and later improved by Nguyen and Einav [2009] is considered in this study. The following standard notations are used: bulk and shear modulus  $K$  and  $G$ , mean effective stress and deviatoric shear stress  $p$  and  $q$ , total and elastic volumetric strain  $\varepsilon_v$  and  $\varepsilon_v^e$ , total and elastic deviatoric shear strain  $\varepsilon_s$  and  $\varepsilon_s^e$ , total and deviatoric stress tensor  $\sigma_{ij}$  and  $s_{ij}$ , and total, elastic and plastic strain tensors  $\varepsilon_{ij}$ ,  $\varepsilon_{ij}^e$  and  $\varepsilon_{ij}^p$ , the Kronecker delta  $\delta_{ij}$ .

The stress-strain relationship in the triaxial stress space is:

$$p = (1 - \vartheta B)K\varepsilon_v^e \quad (19)$$

$$q = 3(1 - \vartheta B)G\varepsilon_s^e \quad (20)$$

The above relationships can also be written in tensorial form as:

$$\sigma_{ij} = (1 - \vartheta B)D_{ijkl}\varepsilon_{kl} \quad (19)$$

in which  $D_{ijkl}$  is the linear elastic stiffness tensor. The grading index  $\vartheta$  is a result of the statistical homogenisation, and can be obtained from the initial and ultimate grain size distributions (gsd). Physically  $\vartheta$  is related to the crushing potential of the materials and admits values from 0 to 1. The breakage variable  $B$  represents the degree of grain crushing, and is used to linearly interpolate the current *gsd* from the initial and ultimate *gsd*'s [Einav, 2007a].

Einav [2007b] derived an elastic-plastic-breakage yield criterion, written in the mixed stress-breakage energy space as:

$$y_B = \frac{(1-B)^2 E_B}{E_C} + \left(\frac{q}{Mp}\right)^2 - 1 \leq 0 \quad (22)$$

where  $M$  is the slope of the critical state line in the  $p - q$  space;  $E_C$  is the critical breakage energy which is computed from the critical crushing pressure during isotropic loading condition  $P_{cr}$  as:

$$E_C = \frac{P_{cr}^2 \vartheta}{2K} \quad (23)$$

$E_B$  is the breakage energy, the thermodynamical conjugate to the breakage variable  $B$ , and has the form:

$$E_B = \frac{\vartheta}{2(1-\vartheta B)^2} \left( \frac{p^2}{K} + \frac{q^2}{3G} \right) \quad (23)$$

The evolution laws for the breakage and plastic strains can be written as:

$$dB = d\lambda \frac{2(1-B)^2 E_B \cos^2 \omega}{E_C} \quad (24)$$

$$d\varepsilon_{ij}^p = d\lambda \left( -\frac{2(1-B)^2 E_B \sin^2 \omega}{p E_C} \frac{\delta_{ij}}{3} + \frac{3s_{ij}}{M^2 p^2} \right) \quad (25)$$

in which  $d\lambda$  is the nonnegative breakage/plasticity multiplier;  $\omega$  is a parameter that couples the plastic volumetric deformation with grain crushing. Physically  $\omega$  represents the pore collapse of the material due to grain crushing and reorganisation [Einav (2007a; 2007b)].

The incremental stress-strain relationship can be written as:

$$d\sigma_{ij} = L_{ijkl} d\varepsilon_{kl} \quad (26)$$

where the elastic-plastic-breakage tangent stiffness tensor  $L_{ijkl}$  has the form [Nguyen & Einav, 2009]:

$$L_{ijkl} = (1 - \vartheta B) D_{ijkl} - \frac{(1-\vartheta B) X_{ij} Y_{mn} D_{mnkl}}{Y_{ij} X_{ij} - \frac{\partial y}{\partial B}} \quad (27)$$

in which:

$$X_{ij} = (1 - \vartheta B) D_{ijkl} \frac{d\varepsilon_{ij}^p}{dB} + \frac{\vartheta \sigma_{ij}}{1-\vartheta B} \quad (28)$$

and

$$Y_{ij} = \frac{\partial y}{\partial \sigma_{ij}} \quad (29)$$

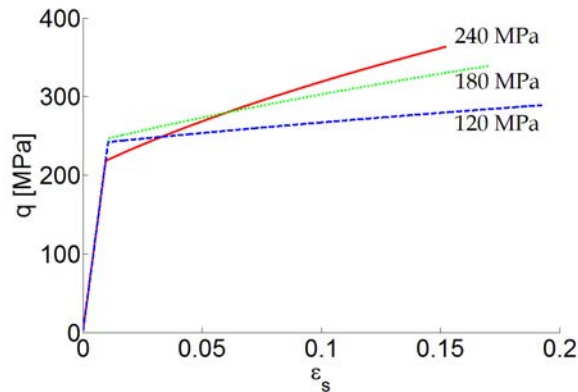
#### *Behaviour of the breakage model*

For the illustration of the behaviour of the breakage model, a triaxial drained test simulation is performed at the constitutive level on the highly porous Bentheim sandstone. In the context of the triaxial test, the confining pressure  $p_0$  means the initial isotropic confining pressure and this confining pressure is maintained horizontally during a strain-controlled axial loading.

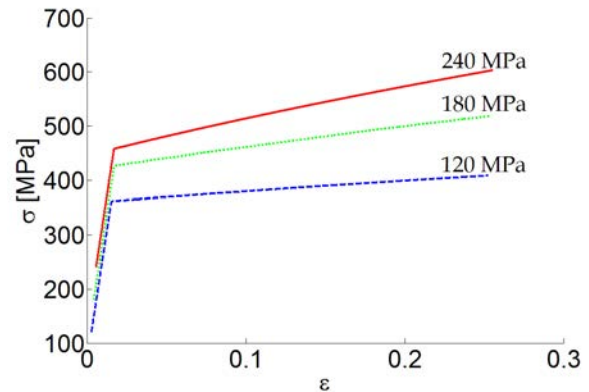
Bentheim sandstone exhibits significant compactive behaviour and distinct compaction bands in experiments [Wong et al. (2001)]. Material parameters of Bentheim sandstone are taken from the reported experimental data and calibrated parameters [Wong et al. (2001), Das et al. (2011)]: bulk stiffness  $K = 138333$  MPa, shear stiffness  $G = 7588$  MPa, critical breakage energy  $E_C = 4.67$  MPa, critical state parameter  $M = 1.7$  MPa, grading index  $\vartheta = 0.85$  and  $\omega = 70^\circ$ .

Fig. 4(a, b) presents the homogeneous responses of the breakage model in the drained triaxial test with different initial confining pressures. The increasing hardening of the breakage model is consistent with the experimental data [Baud et al (2004)] and represents the compactive cataclastic flow regime in which the stress-strain relationship is hardening and the porosity decrease monotonically with deformation [Wong et al. (2001)]. At the onset of inelastic deformation, localisation analysis is carried out following the classical bifurcation condition, Eq. (10-13). It is revealed that for all three confining pressures of 120, 180 and 240 MPa, bifurcation points exist and the corresponding inclination angles  $\phi$  of the localisation surfaces are  $35^\circ$ ,  $15^\circ$  and  $0^\circ$ , respectively.

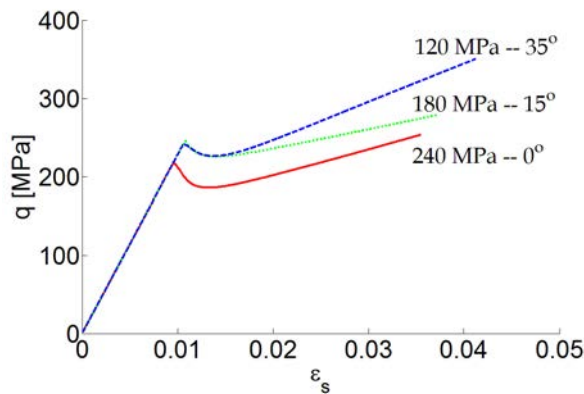
Because of the symmetry in the triaxial loading condition, the angle  $\theta$  (Fig. 3) is  $90^\circ$  in all three cases.



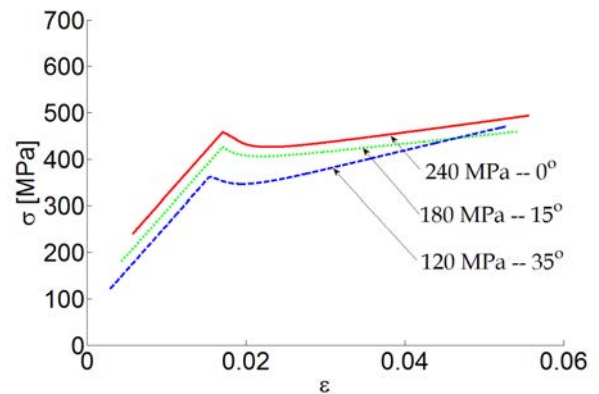
(a) Homogeneous response,  $\varepsilon_s - q$



(b) Homogeneous response, axial  $\sigma - \varepsilon$



(c) Localised response,  $f = 0.25$ ,  $\varepsilon_s - q$



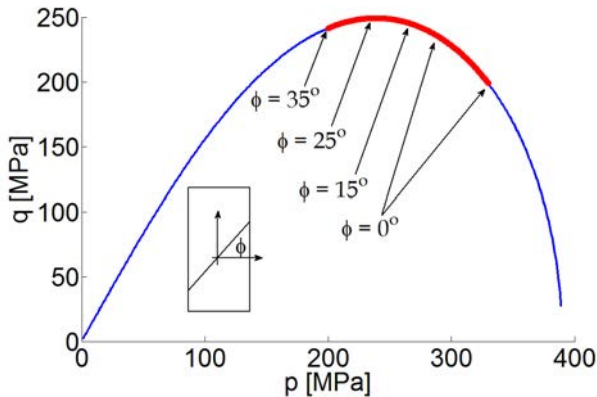
(d) Localised response  $f = 0.25$ , axial  $\varepsilon - \sigma$

**Figure 4. Triaxial loading at different confining pressure**

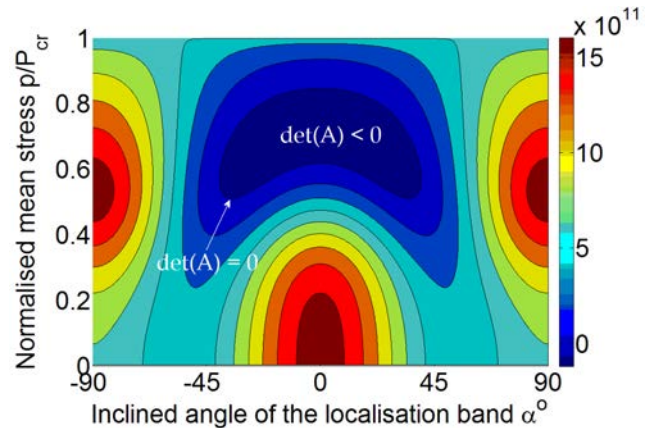
Fig. 4(c, d) depicts the localised responses following the enriched framework with the volume fraction  $f = h/H = 0.25$  in the same loading conditions and given parameters. It is clear from Fig. 4 that localisation changes the material behaviour not only quantitatively but also qualitatively. The homogeneous behaviour is hardening but the localised behaviour, following the enriched framework, first exhibits softening associated with the activation and evolution of a localisation band, followed by hardening with the slope of the macro hardening much larger than that of the homogeneous hardening. Details of the localised behaviour are analysed in the next section.

Fig. 5 presents the initial yield envelopes obtained from the given constitutive parameters. The range of stress state along that initial yielding surface, in which the localisation bands most likely occur, is predicted numerically. In very high pressure regime, no localisation is observed which may be due to diffuse compaction band formation over the entire sample under high confining pressure [Wong et al. (2001)]. In such cases, the deformation may evolve into cataclastic flow without localisation [Das et al. (2011)]. However, as pointed out by Besuelle [2001], extension/dilation shear bands can form in Bentheim sandstone at low confining pressures. Note that the material behaviour at low confining pressures is not yet captured by this version of the breakage model. Fig. 5 also indicates that the inclination angle of the compact band varies from  $0^\circ$ - $35^\circ$  with the variation of the pressure regime, which implies that the localisation band is compactive dominant. In the high pressure regime, the angle is  $0^\circ$  which implies a pure compaction localisation

band; whereas in the medium pressure regime, the angle is about  $30^\circ$ - $35^\circ$  which implies a compactive shear localisation band.



**Figure 5. Predicted zone of localisation**



**Figure 6. Contours of the determinant of the acoustic tensor**

Fig. 6 presents the contours of the determinant of the acoustic tensor against the band orientation angle and the normalised mean stress ( $p/P_{cr}$  in which  $P_{cr} = 389.87$  MPa) at the onset of crushing. In Fig. 6, the inner zone denotes negative determinant of the acoustic tensor which implies that localisation is possible only within this zone. It is clear from Fig. 6 that for a given stress state there is a set of inclination angles that are physically admissible in forming a localisation band. In the post-localisation analysis, only the angle with the minimum value of the determinant of the acoustic tensor is chosen as the inclination angle of the localisation band.

### Size dependent behaviour

In this section, the performance of the enriched approach is investigated in details. The post-bifurcation responses of the Bentheim sandstone during the drained triaxial test are given in the first section; in the second section, the size dependent behaviour is proven by studying the influence of the localisation band width  $h$  on the macro responses.

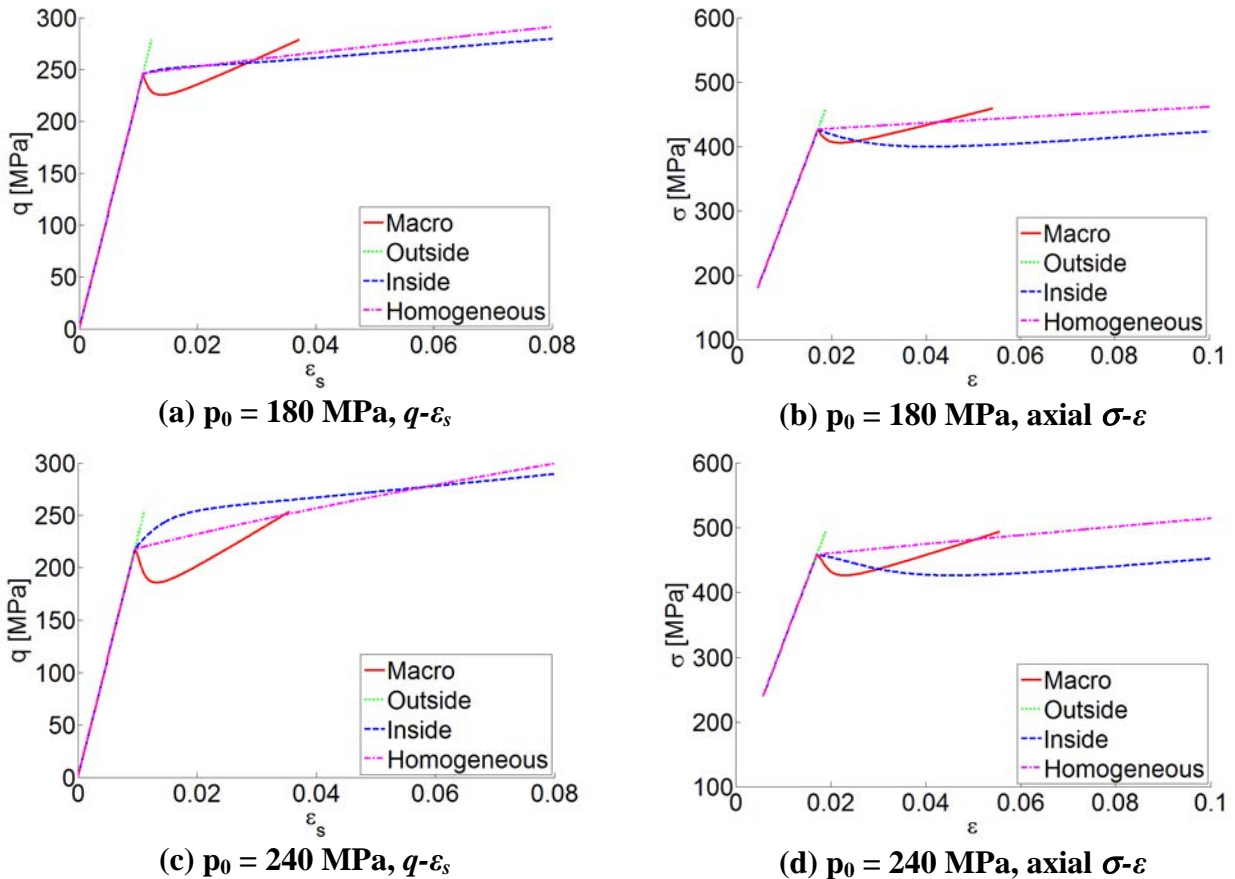
#### Post-bifurcation responses

The stress-strain responses during the drained triaxial test at different confining pressures obtained by using the enhanced framework are given in Fig. 7; the general responses ( $q-\varepsilon_s$ ) are given in Fig. 7(a, c) whereas the responses on the loading direction – in short: the axial response – ( $\sigma-\varepsilon$ ) are given in Fig. 7(b, d). Two confining pressures of 180 and 240 MPa, corresponding to two inclination angles of  $15^\circ$  and  $35^\circ$  of localisation bands, as calculated in the previous section from the bifurcation analysis, are applied. In these analyses, the volume fraction  $f$  is chosen to be 0.25.

It is clear from Fig. 7 that the general and axial homogeneous behaviours are always hardening whereas the macro localised behaviours are more complex. Initially after localisation, the mechanism of elastic unloading for the material outside the band and plastic hardening for the material inside the band dominates, leading to a steep softening behaviour for the macro material. However when the load reaches a minimum, the outside material starts behaving elastically, the macro softening terminates and the macro response is hardening. This hardening phase is then dominated and the localised hardening slope is much larger than the homogeneous hardening slope.

The general response of the material inside of the band is hardening with the long-term localised hardening slope is smaller than the homogeneous one, Fig. 7(a, c). This can be explained from the

fact that in the localised situation, only a band with the volume fraction  $f = 0.25$  exhibit a hardening behaviour whereas in the homogeneous situation the whole sample is undergoing hardening deformation. It is also noted that the initial hardening slope in the localised response is larger with the larger confining pressure. This can originate from different types of the localisation bands at different confining pressures which changes from a compactive shear band, Fig. 7(a) with  $\phi = 35^\circ$  to a pure compaction band, Fig. 7(c) with  $\phi = 0^\circ$ . On the loading direction, Fig. 7(b, d), the axial response of the inside material is first softening and then hardening which is analogous to how the macro behaviour is.



**Figure 7. Responses at different confining pressure,  $f = 0.25$**

Since the homogeneous behaviour of the breakage model in drained triaxial tests is hardening, the initial strong softening behaviour originates from the kinematics enhancement occurred with strain localisation. Gajo et al [2004] also observed this saturation process and suggested that another mechanism of strain hardening is occurring inside the band, induced by the contractive behaviour of the material. It is worth mentioning that as the hardening phase is reached – or the band saturation process continues, at a certain point the material outside the localisation zone will reach the yield condition. At this point, a localised inelastic deformation occurs outside the band while the material inside the band starts elastic unloading, as suggested by Gajo et al [2004]. This behaviour is, however, not considered in this paper.

#### *Effects of localisation band width*

There is considerable experimental evidence on the relation between the shear band thickness  $h$  and the mean grain size  $d_{50}$  [Scarpelli and Wood (1982); Vardoulakis et al. (1985a; 1985b); Muhlhaus and Vardoulakis (1987)], usually given in the form:



$$h = (10 \div 15)d_{50} \quad (30)$$

This band width  $h$  is of a great importance in understanding and numerical modelling of progressive failure in granular materials [Gudehus (1978); Muhlhau and Vardoulakis (1987)]. In this section, the size dependent behaviour in the drained triaxial test is investigated by varying the value of  $h$  whilst keeping  $H$  constant which means the volume fraction  $f$  is changed. Three values of  $f$  are chosen 0.25, 0.50 and 0.75, and two confining pressures of 180 and 240 MPa are applied.

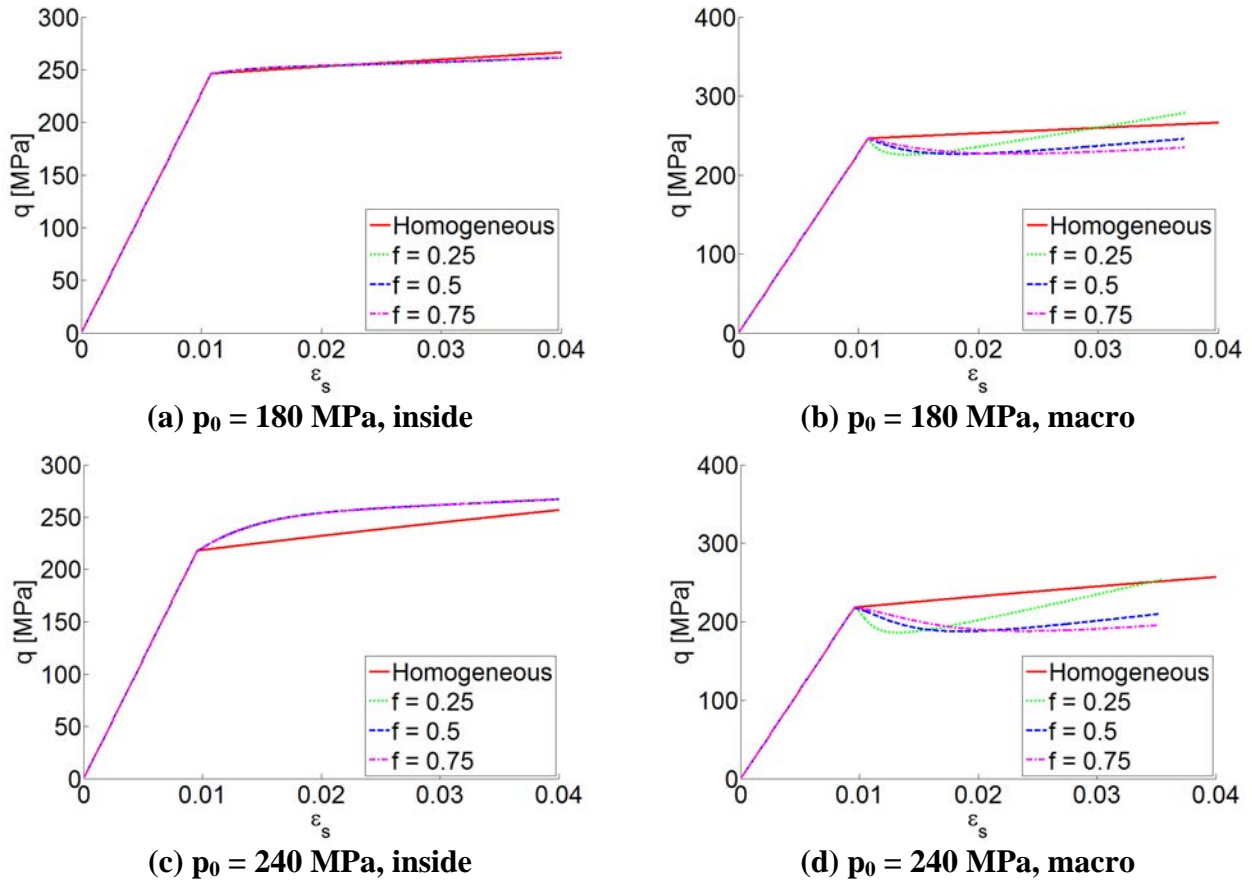
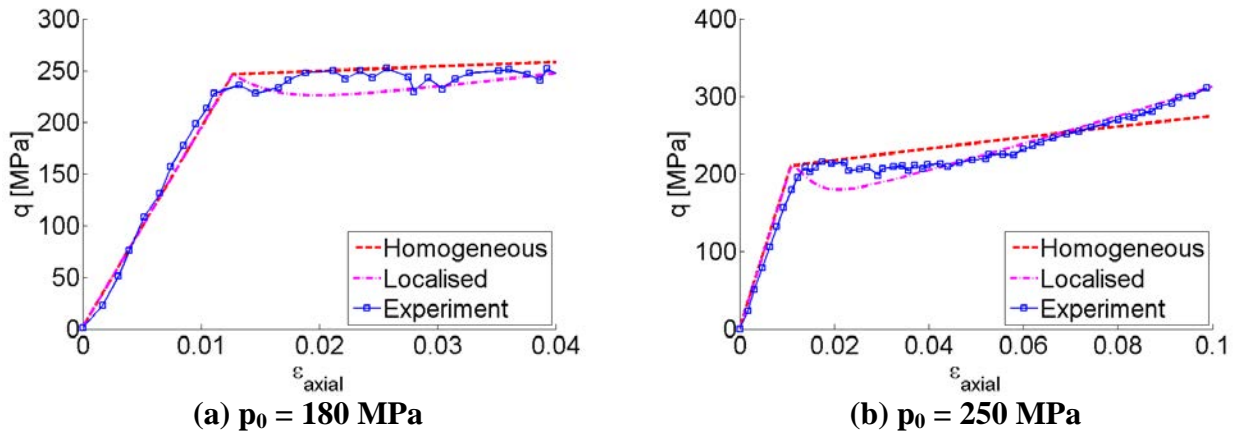


Figure 8. Size dependent behaviour on  $q$ - $\varepsilon_s$  response

It is clear from Fig. 8 (a, c) that the response of the material inside the localisation band is independent of the value of  $f$ . This independence implies that the macro dissipation energy, which only takes place inside the localisation band, scales linearly with the volume fraction of the band. Fig. 8 (b, d) illustrates that a size effect is exhibited since the macro response is strongly affected by the value of  $f$ . Since all of the dissipation energy occurs only inside the localisation band, apparently the smaller the value of  $f$ , the steeper the macro softening experiences initially. It can also be seen from Fig. 8 (b, d) that the band saturation process occurs at larger strain magnitude for the larger value of  $f$  which is explained by the fact that the localisation area is larger. Moreover, with a larger value of  $f$ , the influence of the elastically loading behaviour of the outside material on the subsequent hardening of the macro becomes smaller which is indicated by the smaller value of the hardening slope, Fig. 8 (b, d).

The experimental validation of the proposed constitutive model is shown in Fig. 9. The same set of parameters for the breakage model obtained from experimental data on Bentheim sandstone (see the preceding section on the breakage model) was used. A volume fraction  $f=0.35$  needed for the enriched model was estimated from the experimental figures in [Baud et al (2004)]. We know that  $f$

is evolving during failure of the material, so the fixed value of  $f=0.35$  is used merely as a rough estimation to explore the trend of the material behaviour.



**Figure 9. Validation of model response under different confining pressures**

As can be seen in Fig. 9, the softening behaviour at the onset of localised failure, followed by hardening due to an increase of the density as a result of compaction, can be captured by the enriched model, while the original breakage model always exhibits a hardening trend. This preliminary result should serve as a basis for future improvements of the both the breakage model and the enrichment framework.

## Conclusions

Theoretical development, numerical implementation and performance of a new constitutive modelling framework were presented in conjunction with the use of a model based on breakage mechanics for exploring localised failure in sandstone. The results explain the response of the material inside and outside the localisation band and how they influence the macro response. This helps improve both the performance and capacity of constitutive model in capturing post-bifurcation behaviour of geomaterials. Although the enriched approach provides a good understanding on the influence of the localised deformation, it is not adequate yet to capture correctly the material behaviour with oscillation in the macro response due to the gradual formation of several compaction bands during loading [Wong and Baud (2012); Das (2012)]. With that evolution of the localisation behaviour, inelastic behaviour outside the band must be also taken, and this is a subject of an on-going research.

## Acknowledgements

Funding support from the Australian Research Council via projects DP110102645 and FT140100408 (Giang D. Nguyen) and LP13010088 (Ha Bui) is gratefully acknowledged.

## Reference

- A. Das, G. Nguyen, I. Einav (2011), Compaction bands due to grain crushing in porous rocks: a theoretical approach based on breakage mechanics, *Journal of Geophysical Research* 116 1–14
- A. Das (2012), *A theoretical study of grain crushing induced compaction localization in porous sandstones*, PhDthesis - University of Sydney
- A. Das, G. D. Nguyen, I. Einav (2013), The propagation of compaction bands in porous rocks based on breakage mechanics, *Journal of Geophysical Research: Solid Earth* 118 (5) 2049–2066
- A. Gajo, D. Bigoni and D. M. Wood (2004), Multiple shear band development and related instabilities in granular materials, *Journal of the mechanics and physics of solids* 52 2683-2724

- D. Kolymbas (1981), Bifurcation analysis for sand samples with a non-linear constitutive equation, *Ingenieur Archive* 50 (2) 31–40
- S. Forest, E. Lorentz (2004), Localization phenomena and regularization methods. *Local Approach to Fracture. Paris, Les Presses de l'Ecole des Mines* 311-371.
- G. D. Nguyen, I. Einav (2009), The energetics of cataclasis based on breakage mechanics, *Pure and Applied Geophysics* 166 (10-11) 1693–1724
- G. D. Nguyen, I. Einav, A. M. Korsunsky (2012), How to connect two scales of behaviour in constitutive modelling of geomaterials, *Geotechnique Letters* 2, 129-134.
- G. D. Nguyen, A. M. Korsunsky, I. Einav (2014), A constitutive modelling framework featuring two scales of behaviour: Fundamentals and applications to quasi-brittle failure, *Engineering Fracture Mechanics* 115, 221–240.
- G.D. Nguyen, I. Einav (2010). Nonlocal regularisation of a model based on breakage mechanics for granular materials. *International Journal of Solids and Structures* 47(10) 1350-1360.
- G. Scarpelli, D.M. Wood (1982), Experimental observations of shear band patterns in direct shear tests. *Proc. IUTAM Conf. Deformation and Failure of Granular Materials, Delft* 473–484
- G. Gudehus (1978), Engineering approximations for some stability problems in geomechanics, *Advances in analysis of geotechnical instabilities, SM Study* 13 1–24
- H.-B. Muhlhaus, I. Vardoulakis (1987), The thickness of shear bands in granular materials, *Geotechnique* 37(3) 271–283
- H. Schreyer, M. Neilsen (1996), Analytical and numerical tests for loss of material stability, *International Journal for Numerical Methods in Engineering* 39 (10) 1721–1736.
- I. Einav (2007a), Breakage mechanics-part i: Theory, *J Mech Physics Solids* 55 (6) 1274 – 1297
- I. Einav (2007b), Breakage mechanics-part ii: Modelling granular materials, *J Mech Physics Solids* 55 (6) 1298 – 1320
- I. Vardoulakis, M. Goldscheider, G. Gudehus (1985a), Formation of shear bands in sand bodies as a bifurcation problem, *International journal of numerical and analytical methods in Geomechanics* 2 (0) (1978) 99–128
- I. Vardoulakis, B. Graf, Calibration of constitutive models for granular materials using data from biaxial experiments, *Geotechnique* 35(3) 299–317
- I. Vardoulakis, B. Graf, A. Hettler (1985b), Shear band formation in a fine-grained sand, *Proc. 5<sup>th</sup> Int. Conf. Numer. Meth. Geomech., Nagoya* 1 517–521
- J. Desrues, G. Viggiani (2004), Strain localization in sand: an overview of the experimental results obtained in grenoble using stereophotogrammetry, *International Journal for Numerical and Analytical Methods in Geomechanics* 28 (4) 279–321.
- J. Rudnicki, J. Rice (1975), Conditions for the localization of deformation in pressure-sensitive dilatant materials, *Journal of the Mechanics and Physics of Solids* 23 (6) 371 – 394
- J. Rice (1976), The localization of plastic deformation, in: *Theoretical and Applied Mechanics, 14th IUTAM Congress, North-Holland, Amsterdam*
- J. Réthoré, F. Hild, S. Roux (2007), Shear-band capturing using a multiscale extended digital image correlation technique, *Computer Methods in Applied Mechanics and Engineering*, 196(49–52) 5016-5030
- K. Runesson, N. S. Ottosen, P. Dunja (1991), Discontinuous bifurcations of elastic-plastic solutions at plane stress and plane strain, *International Journal of Plasticity* 7 (1-2) 99–121.
- K. A. Alshibli, S. Sture (1999), Sand shear band thickness measurements by digital imaging techniques, *Journal of Computing in Civil Engineering* 13 (2) 103–109
- M. Neilsen, H. Schreyer (1993), Bifurcations in elastic-plastic materials, *International Journal of Solids and Structures* 30 (4) 521 – 544
- N. S. Ottosen, K. Runesson (1991), Properties of discontinuous bifurcation solutions in elasto-plasticity, *International Journal of Solids and Structures* 27 (4) 401 – 421
- P. Baud, E. Klein and T.-f. Wong (2004), Compaction localisation in porous sandstones: spatial evolution of damage and acoustic emission activity, *Journal of Structural Geology*, 26(2004), 603-624
- P. Besuelle (2001), Compacting and dilating shear bands in porous rocks: theoretical and experimental conditions, *Journal of Geophysical Research* 106 (-) 13435–13442
- P. Bažant, F.-B. Lin (1988), Non-local yield limit degradation, *International Journal for Numerical Methods in Engineering* 26 (8) 1805–1823.
- R. R. Pedersen, A. Simone, and L. J. Sluys (2008), An analysis of dynamic fracture in concrete with a continuum visco-elastic visco-plastic damage model. *Engineering Fracture Mechanics* 75(13) 3782-3805.
- R. de Borst, H.-B. Muhlhaus (1992), Gradient-dependent plasticity: Formulation and algorithmic aspects, *International Journal for Numerical Methods in Engineering* 35 (3) 521–539.
- R. Chambon, S. Crochepeyre, J. Desrues (2000), Localisation criteria for non-linear constitutive equations of geomaterials, *Mechanics of Cohesive-frictional Materials* 5 (-) 61–82
- S. Sanborn, J. Prevost (2011), Frictional slip plane growth by localisation detection and the extended finite element method (xfem), *International journal of numerical and analytical methods in Geomechanics* 35 (-) 1278–1298
- T.-f. Wong, P. Baud, and E. Klein (2001), Localised failure modes in a compactant porous rock, *Geophys. Res. Lett.*, 28(13), 2521–2524

T.-f. Wong and P. Baud (2012), The brittle-ductile transition in porous rock: A review, *Journal of structural geology*, 44, 25–53

## Multiphase model for thermal activity in a multistory residential home

\*C. De Costa<sup>1</sup> and †Satoru Ushijima<sup>2</sup>

<sup>1</sup>PhD student at Department Civil and Earth resources Engineering, Kyoto University, Kyoto, Japan.

<sup>2</sup> Prof. Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan.

\*Presenting author: de.chamika.83m@st.kyoto-u.ac.jp

†Corresponding author: ushijima@media.kyoto-u.ac.jp

### Abstract

This paper discusses the indoor climate in a full-scale three-story house using the MICS model, which is a multiphase model that is able to predict the velocity and temperature profiles in a 3D domain, as well as the interaction of temperature between air and different solid materials. The paper first compares the numerical results with existing experimental results done for similar multistory structures. Further more, numerical experiments are conducted with different heating methods: floor heaters and panel heaters, in order to find the optimum conditions required to heat the building efficiently. As a result, the suitable heating method is successfully suggested in the multistory building by the multiphase model.

**Keywords:** Indoor climate, multiphase model, CFD, multi-story house

### Introduction

It is an important engineering subject to keep the comfortable indoor climate of multi-story houses located in cold districts with efficient heating system to save energy. For this purpose, many experimental investigations have been conducted within experimental rooms and in actual outside fields. On the other hand, the numerical studies to predict indoor climate have relatively few in particular for the actual multi-story houses due to the difficulty in dealing with the complicated-shaped and multiple materials of houses, which have different thermal conductivities.

Thus, the numerical predictions were carried out for a full-scale three-story house [1] with a multiphase model, MICS proposed by Ushijima [2]. The MICS model is useful to deal with thermal interactions between air and the solid materials having different thermal conductivities. As a result of computations, it was shown that the temperatures on the walls between the experiment and simulation are in good agreement in first and second floors. In addition, numerical experiments were conducted for the house with different heating elements: floor heaters and wall attached panel heaters. It was shown that the floor heaters provide a better heating in the house compared with the panel heaters.

## Computational method

The governing equations of the multiphase model MICS model [2] are shown as follows:

$$\frac{\partial u_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \frac{1}{\rho} \frac{\partial \tau_{ij}}{\partial x_j} + g_i \quad (2)$$

where  $\tau_{ij}$  is given by Equation (4)

$$\tau_{ij} = \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (3)$$

$$\frac{\partial T}{\partial t} + \frac{\partial (T u_i)}{\partial x_i} = \frac{1}{\rho S_p} \frac{\partial}{\partial x_i} \left( \lambda \frac{\partial T}{\partial x_i} \right) + S_c \quad (4)$$

In the above equations,  $t$  is time,  $x_i$  is the orthogonal coordinate,  $S_c$  is the heat source term,  $g_i$  is acceleration due to gravity and  $u_i$  is the mass-averaged velocity between gas and solid phases defined by the following equation:

$$u_i = \frac{\sum_k \rho_k V_k U_{ki}}{\sum_k \rho_k V_k} \quad (5)$$

In Equations (2) to (4), temperature  $T$ , density  $\rho$ , specific heat  $S_p$ , Pressure  $p$  and viscosity coefficient  $\mu$  are volume-averaged variables defined by

$$\Psi = \frac{\sum_k V_k \Psi_k}{\sum_k V_k} \quad (6)$$

In Equations (5) and (6),  $V_k$  represents volume of the phase- $k$ . In this study, the non-isotropic characteristics for thermal conductivity  $\lambda$  are taken into account [3].

The governing equations are discretized with FVM. The numerical algorithm for solving the set of equations is divided into three separate stages: prediction stage, pressure calculation stage and correction stage. In the prediction stage, in which convection and diffusion equations are solved, the implicit method is used on a collocated grid system, which is called C-ISMAL method [4]. In the C-ISMAL method, tentative velocity  $u_i^*$  is calculated from Equation (2) as follows:

$$\frac{u_i^* - u_i^n}{\Delta t} = g_i - \frac{1}{\rho} \frac{\partial p^n}{\partial x_i} - \gamma_1 \frac{\partial}{\partial x_j} (u_i^* u_j^n) - (1 - \gamma_1) \frac{\partial}{\partial x_j} (u_i^n u_j^n) + \frac{\gamma_2}{\rho} \frac{\partial \tau_{1,ij}}{\partial x_j} + \frac{1 - \gamma_2}{\rho} \frac{\partial \tau_{2,ij}}{\partial x_j} \quad (7)$$

with

$$\tau_{1,ij} = \frac{\partial}{\partial x_j} (\mu u_i^*) + \frac{\partial}{\partial x_i} (\mu u_j^*) \quad (8)$$

and

$$\tau_{2,ij} = \frac{\partial}{\partial x_j} (\mu u_i^n) + \frac{\partial}{\partial x_i} (\mu u_j^n) \quad (9)$$

where  $0 \leq \gamma_1, \gamma_2 \leq 1$ . Fifth-order TVD scheme [5] is used to solve the convection terms included on the right hand side of Equation (7).

In the pressure calculation stage, C-HSMAC method [6] is used. It satisfies the incompressible conditions accurately.

The relationship between density and temperature is given by the following equation:

$$\rho = \frac{\rho_0}{1 + \beta(T - T_0)} \quad (10)$$

where  $\beta$  is the coefficient of thermal expansion,  $\rho_0$  and  $T_0$  are the reference density and temperature respectively.

In order to improve computational efficiency, the numerical procedures are parallelized using Message Passing Interface (MPI). Due to this parallelization, computational times are largely reduced in particular to solve the implicit equations as well as the pressure computation stage.

### Application

The multiphase model was applied to the full-scale three-story house, for which experimental study had been conducted by Usami[1] in the Sendai prefecture in Japan. Figure 1 shows the house that was used for the experimental study.



Figure:1 Experimental house [1]

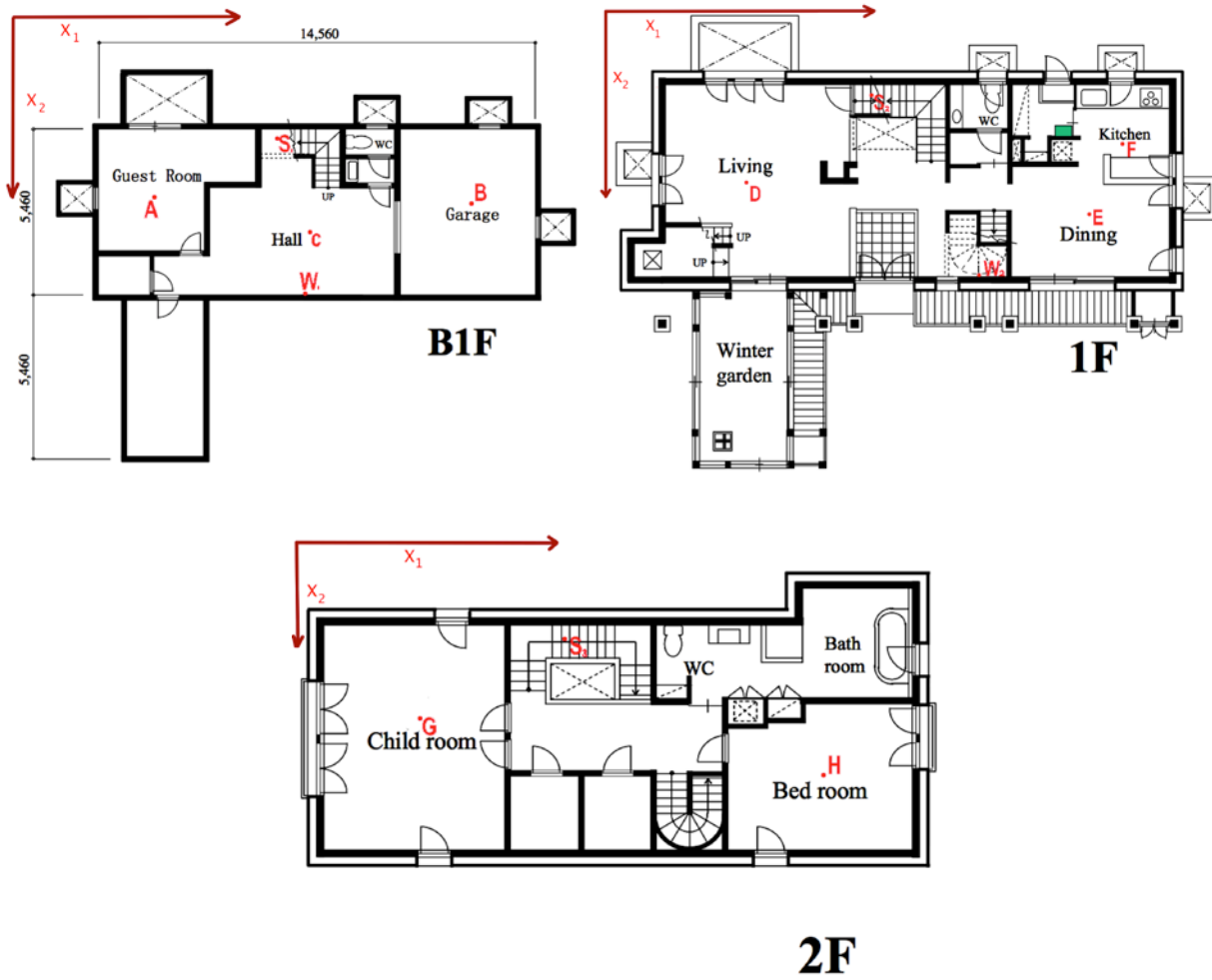


Figure: 2 Blue prints of the house with the locations of measured points [1]

Table 1: Locations of measuring points.

Location for measurement ( $x_1$ [m] * $x_2$ [m])			
A	2 x 2.3	G	2.4 x 2.5
B	12.9 x 2.6	H	11.9 x 3.5
C	7.3 x 3.5	S1	6.1 x 0.5
D	2.6 x 2.7	S2	6.1 x 0.5
E	11.9 x 2.7	S3	6.1 x 0.5
F	12.9 x 1.8	W1	7.3 x 5.5
		W2	8.9 x 5.5

The house used in the experiment is a three-story house complete with minimal furniture. Figure 2 shows the blueprints of each level as well as the measured locations of temperature. The exact locations are given in Table 1. The height of the measured locations is 1.5m above the floor level. The locations were chosen for the center point of major human activity.

In the experiments, the effect of the floor heaters was examined. The thermal conductivity for each material in the house is shown in Table 2. The heat supplied by the floor heater is 7.8 kW.



Table 2: Thermal conductivity for materials in the house

Materials	Thermal conductivity (W/(m.K))
1st floor walls	0.45
2nd floor	0.45
2nd floor walls	0.13
3rd floor walls	0.13
furniture	0.13
2nd floor windows	0.96
3rd floor windows	0.96

In the present numerical study, the solid materials of the house were represented by tetrahedron elements to respond to the complicated-shaped structures as shown in Figures 3 and 4.

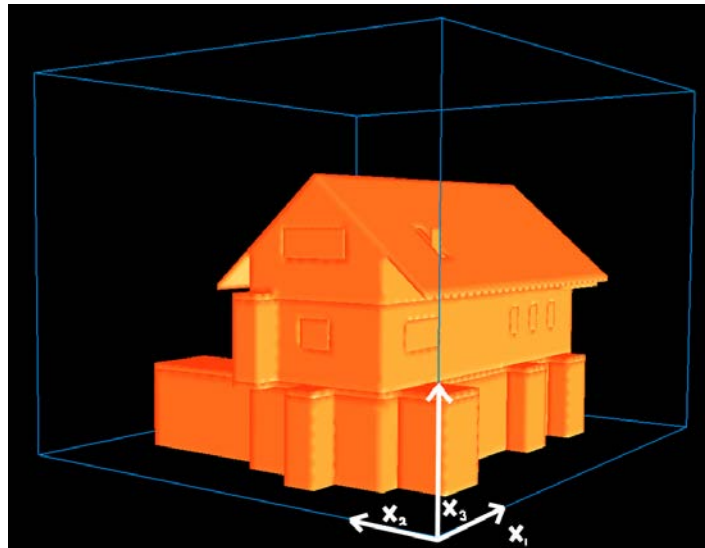


Figure 3: Visualization of the house in 3D and co-ordinate system

Figure 4: Tetrahedron mesh created for the solid parts of the house

The number of computational cells is  $240 \times 240 \times 120$ . The simulation required 26 hours of computational time for 10 actual hours. The computation used 128 processors. The simulations were run in Cray XE6 in Kyoto University.

Figure 5: Cross section on  $x_1 - x_3$  plane at  $x_2 = 4.5$  [m]

Figure 5 shows the  $x_1 - x_3$  plane view of the velocity profile inside the middle of the house. The green and yellow area indicates the location of the floor heaters. The outside boundary condition is set at  $0^\circ\text{C}$  with the no wind, while the adiabatic condition is used for the floor. As shown in Figure 5, the bottom left and bottom right rooms have little disturbance.

Figure 6: Cross section on  $x_2 - x_3$  plane at  $x_1 = 4.5$  [m]

The  $x_2 - x_3$  plane view of the velocity profile is shown in Figure 6. It is noted that the location of the stairs that is on the right side of the Figure 6 plays an important role in drawing out air in the bottom floors, as shown in Figure 7. In Figure 7, the air flows upwards due to buoyancy effects, thus reducing the heat at the first floor.

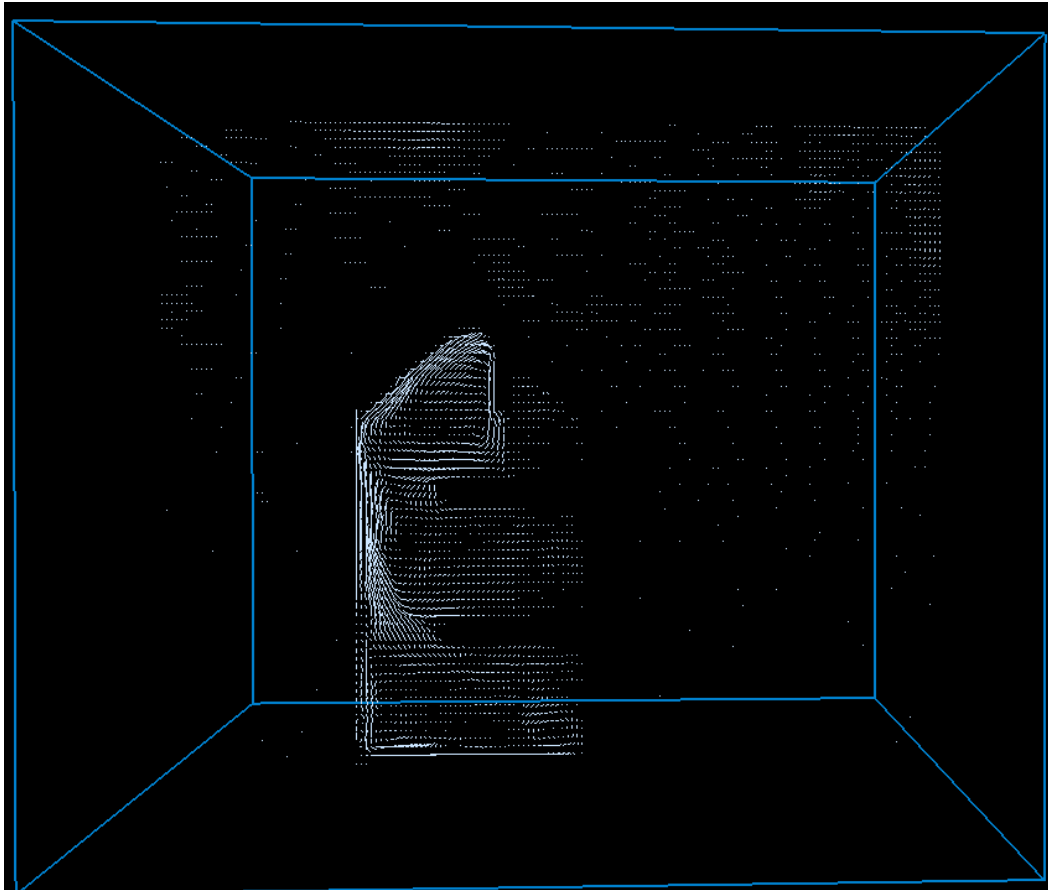


Figure 7: Velocity profile in the stairs on  $x_2 - x_3$  plane at  $x_1 = 8.5$  [m]

**Comparisons with experimental results**

The calculated temperatures on the walls in first floor, W1, and second floor, W2, are shown in Figure 8. Compared with experimental results in Figure 8, it can be seen that they are in good agreement.

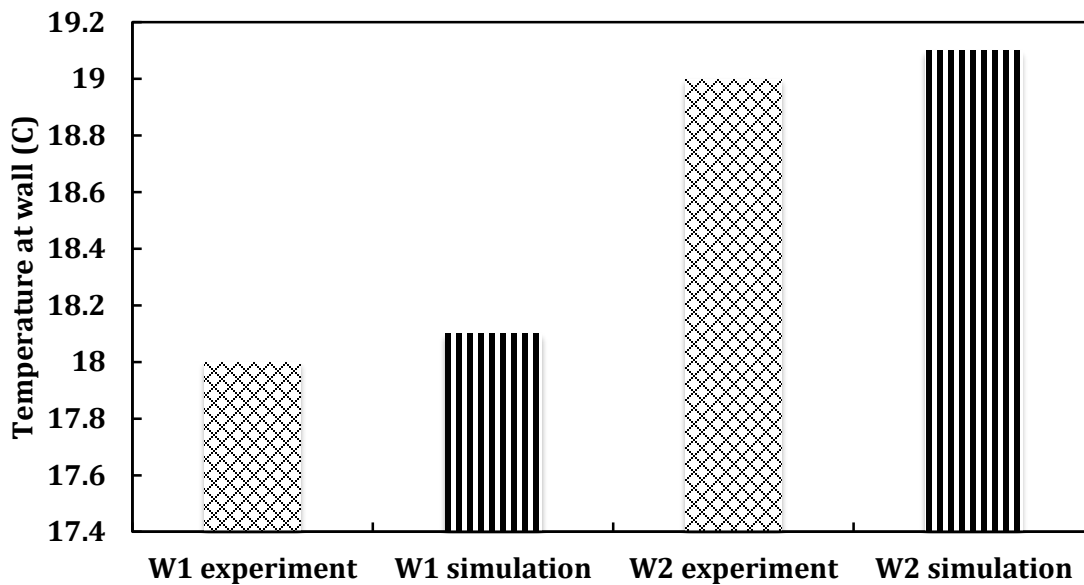


Figure 8 : Temperature at the walls in experiment and current simulation

Figure 9 shows the comparison of the temperatures at the measuring points, A, D and G, in the rooms shown in Figure 2. In general, the calculated results agree with the experimental data. The differences of temperature between the experimental data and the calculated results are less than 0.1 °C.

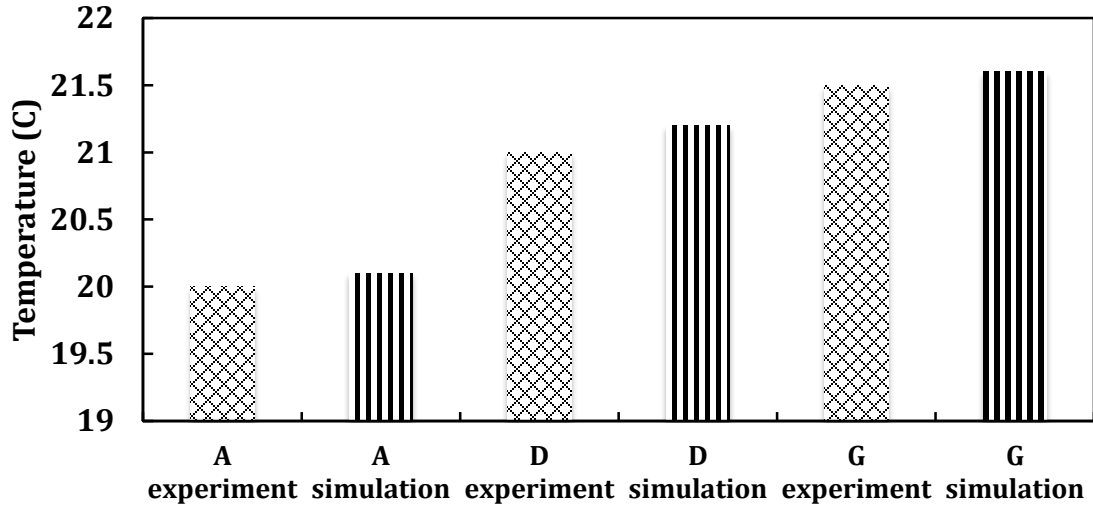


Figure 9: Temperature in the rooms in experiment and simulation

**Numerical experiments**

In addition to the comparisons with experimental results, numerical experiments were conducted to find the effective heating method. Using the same three-story house and boundary conditions, two heating methods were compared: floor heaters and heat panels with same heat source at 7.8 kW. The floor heater configuration is the same as the simulation above. The panel heaters were installed near the floor on the second floor walls of the house. Figure 10 shows the locations of three panel heaters, in blue, as well as the resulting velocity profile. Figure 11 shows the resulting temperature distributions in the house. The measured points are taken from the center of the room at 1.5m in height. The locations of points A to S3 are shown in Figure 2.

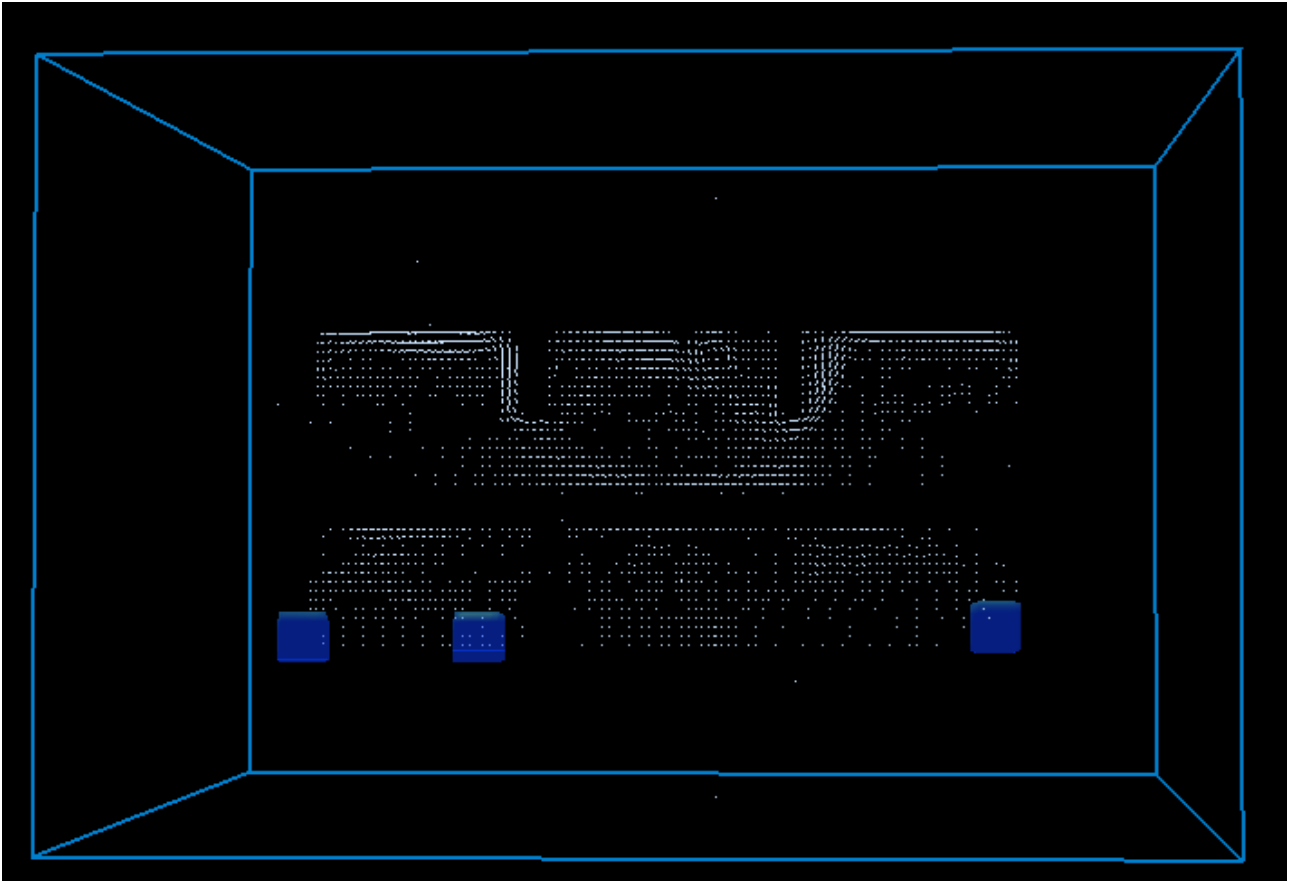


Figure 10: Cross section of the house on  $x_1 - x_3$  plane at  $x_2 = 5$  [m]

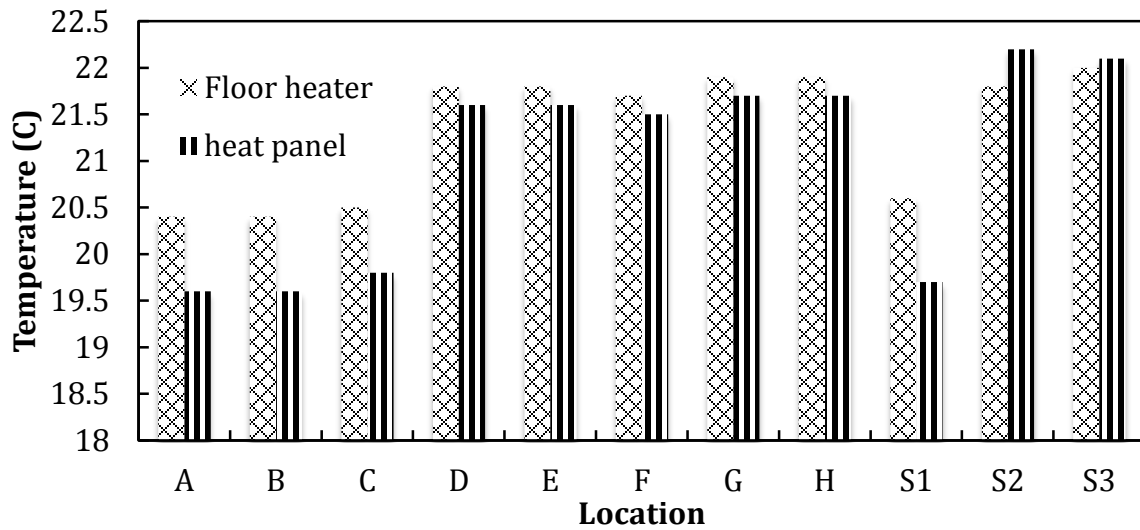


Figure 11: Temperature distribution in house comparing floor heater and heat panel

As shown in Figure 11, it is clear that the floor heater is able to increase temperature in all of rooms, A to H, except the stair regions, S2 and S3. The top two floors, locations D, E, F, G and H, have a less significant difference in the temperature at only less than 0.5 C different results.

Conclusively, it was suggested that the present multiphase model enables us to find optimum heating method with minimum energy. The computational method is a useful tool since the heating method can be easily changed compared with the full-scale experiments.

## References

- [1] Y. Usami, H. Yoshino, H. Sugawara, Indoor Environment and energy consumption pattern of an experimental house with energy efficient design
- [2] S. Ushijima. Multiphase-model approach to predict arbitrarily-shaped objects moving in free surface flows. Proc of APCOM'07 – EPMESC XI, pages MS41–3–1, 2007.
- [3] D. Toriu, S. Ushijima: 3D computation for natural convection in porous media using multiphase model.
- [4] S. Ushijima and I. Nezu. Higher-order implicit (C-ISMAC) method for incompressible flows with collocated grid system. JSCE Journal, (719/II-61):21–30, 2002.
- [5] S. Yamamoto and H. Daiguji. Higher-order-accurate upwind schemes for solving the compressible Euler and Navier-Stokes equations. Computers Fluids, 22(2/3):259–270, 1993.
- [6] S. Ushijima, Y. Okuyama, I. Nezu: Parallel Computational Method for Incompressible Flows with C-HSMAC Method on 3D Unstructured Collocated-Grid System. Japan Society of Civil Engineers, 2004.
- [7] S. Ushijima and N. Kuroda. Multiphase modeling to predict finite deformations of elastic objects in free surface flows. Fluid Structure Interaction V, WIT Press, pages 34–45, 2009.
- [8] B. R. Shin, T. Ikohagi, and H. Daiguji. An unsteady implicit SMAC scheme for two-dimensional incompressible Navier-Stokes equations. JSME International Journal, 36(4):598–606, 1993.
- [9] C. W. Hirt and J. L. Cook. Calculating three-dimensional flows around structures and over rough terrain. J. Comput. Phys., 10:324–340, 1972.
- [10] S. Ushijima, S. Yamada, S. Fujioka, and I. Nezu. Prediction method (3D MICS) for transportation of solid bodies in 3D free-surface flows. JSCE Journal, 810/II-74:79–89, 2006.
- [11] H. A. Van Der Vorst. BI-CGSTAB : A first and smoothly converging variant of BI-CG for the solution of nonsymmetric linear systems. SIAM J. Sci. Stat. Comput., 13:631–644, 1992.
- [12] W. Gropp, E. Lusk, and R. Thakur. Using MPI-2. The MIT Press, 1999.
- [13] J. P. Holman. Heat Transfer. McGRAW-HILL, INC., 1997.

## Numerical simulation of flows around jellyfish in a current

\*T. Inomoto<sup>1</sup>, K. Matsuno<sup>2</sup>, M. Yamakawa<sup>2</sup>, S. Asao<sup>3</sup> and S. Ishihara<sup>1</sup>

<sup>1</sup>Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan.

<sup>2</sup>Department of Mechanical and Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan.

<sup>3</sup>Department of Mechanical Engineering, College of Industrial Technology, 1-27-1 Amagasaki, Hyogo, Japan.

\*Presenting and corresponding author: d3821501@edu.kit.ac.jp

### Abstract

Coupled behaviors of fluid-flow and swimming-jellyfish are necessary for swimming movements and the dynamics governing these coupled behaviors is difficult and complex. A lot of researches investigated coupled behaviors of fluid-flow and jellyfish-swimming by using two-dimensional axisymmetric numerical simulations. However, in order to simulate swimming jellyfish in an asymmetric current, the three-dimensional simulation is necessary. On the other hand, in the simulation of an unsteady flow caused by a moving wall boundary, the Geometric Conservation Law (GCL) is important. In the computational method which does not strictly satisfy GCL, arbitrary grids moving affects the flow field and the physical conservation law is destroyed. Moving-Grid Finite-Volume Method (MGFVM) is suitable for such a flow because GCL is strictly satisfied. In MGFVM, GCL condition is automatically and strictly satisfied by the discretization performed using a four-dimensional control volume in the space and time unified domain  $(x, y, z, t)$ . In this paper, we perform the three-dimensional coupled simulation of fluid-flow and jellyfish-swimming with six degrees of freedom of motion by using MGFVM and investigate the influence of a current on swimming jellyfish.

**Keywords:** Coupled simulation, Swimming, Jellyfish, Incompressible Flow

### Introduction

Coupled behaviors of fluid-flow and swimming-jellyfish are necessary for swimming movements and the dynamics governing these behaviors is difficult and complex. Jellyfish were the earliest animals to evolve muscle-powered swimming in the sea. A swimming of jellyfish consists of contraction, relaxation and inertia. First swimming mechanism is a jet propulsion caused by the subumbrella volume change that occurs during the contraction and the relaxation. In a jet motion, a first vortex ring, which is called the ‘starting vortex’, occurs by the contraction and causes a strong jet propulsion. Second swimming mechanism is a paddling motion on the bell margin and not as simple as jet propulsion. In a paddling motion, a second vortex ring, which is called the ‘stopping vortex’, occurs by the relaxation. The stopping vortex rotates in the direction opposite to the starting vortex and influences the starting vortex. The stopping vortex plays an important role in swimming mechanisms [Colin and Costello (2002); Mchenry and Jed (2003); Dabiri et al. (2005); Costello et al. (2008)].

A lot of researches investigated swimming jellyfish by using computational fluid dynamics (CFD). The dynamics of swimming jellyfish was modeled by using the two-dimensional simulation using the SIMPLE algorithm [Dular et al. (2009)]. The vortex structure caused by a swimming jellyfish was investigated by using the two-dimensional simulation using the arbitrary Lagrangian–Eulerian (ALE) method [Sahin and Mohseni (2009)]. The relationship between kinematics and swimming jellyfish was investigated by using the two-dimensional axisymmetric simulation [Alben et al. (2013)]. The three-dimensional geometry of swimming jellyfish was extrapolated from the two-dimensional axisymmetric simulation [Rudolf and Mould (2010)]. Thus far, there are few investigations into three-



dimensional simulations of swimming jellyfish. Moreover, in these investigations, it was assumed that jellyfish is in stationary fluid. However, in the sea, a current may change the vortex structure caused by swimming jellyfish and affect a swimming of jellyfish. The swimming jellyfish in the current cannot be simulated by using a two-dimensional axisymmetric simulation because the current is not axisymmetric. Thus, a three-dimensional simulation is necessary in order to investigate the influence of the current on swimming jellyfish.

On the other hand, in the simulation of an unsteady flow caused by a moving wall boundary, the computational grid moves and deforms time-dependently. As the computational method for such a moving grid, the method applying discretization of the governing equation on a general body-conforming curvilinear coordinate [Vinokur (1974)], the arbitrary Lagrangian-Eulerian (ALE) method in which the mesh point can be moved independently of fluid motion [Noh (1964)], the space-time finite-element method [Tezduyar et al. (1992)] and so on were suggested. It is most important for the computational method for moving grid that the Geometric Conservation Law (GCL) is satisfied [Thomas and Lombard (1979)]. In the computational method which does not strictly satisfy GCL, arbitrary grids moving may affect the flow field and physical conservation law may be destroyed. Moving-Grid Finite-Volume Method (MGFVM) was suggested as the computational method which strictly satisfy GCL [Mihara (1999)] and its performance was shown in various unsteady flows [Matsuno (2010)]. The GCL condition is automatically and strictly satisfied by the discretization performed using a four-dimensional control volume in the space and time unified domain  $(x, y, z, t)$ . In structured grids, the method was firstly applied to compressible flows [Matsuno (2001)] and extended to incompressible flows [Inomoto (2004)]. In incompressible flows, the couple of pressure and velocity was done by the fractional step method on the four-dimensional domain. In order to apply to the object of complicated shape, Unstructured Moving-Grid Finite-Volume Method, which was MGFVM extended to unstructured grids, was suggested [Yamakawa and Matsuno (2003)]. Moreover, in order to apply to a greatly moving wall boundary, Moving Computational Domain (MCD) approach in which whole of computational region could move was suggested [Watanabe and Matsuno (2009)].

In this paper, we perform the three-dimensional coupled simulation of fluid-flow and jellyfish-swimming in a current with six degrees of freedom of motion by using Moving-Grid Finite-Volume Method. The swimming jellyfish demonstrated and the influence of the current is shown.

## Governing equations for fluid flow

### *Governing Equation*

The governing equations of fluid-flow are the continuity equation,

$$\frac{\partial u}{\partial t} + \frac{\partial v}{\partial x} + \frac{\partial w}{\partial y} = 0, \quad (1)$$

and the Navier-Stokes equations for incompressible flow,

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{E}}{\partial x} + \frac{\partial \mathbf{F}}{\partial y} + \frac{\partial \mathbf{G}}{\partial z} + \frac{\partial \mathbf{P}_1}{\partial x} + \frac{\partial \mathbf{P}_2}{\partial y} + \frac{\partial \mathbf{P}_3}{\partial z} = 0, \quad (2)$$

with

$$\mathbf{E} = \mathbf{E}_a - \mathbf{E}_v, \quad \mathbf{F} = \mathbf{F}_a - \mathbf{F}_v, \quad \mathbf{G} = \mathbf{G}_a - \mathbf{G}_v, \quad \mathbf{q} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}, \quad \mathbf{E}_a = \begin{bmatrix} u^2 \\ uv \\ uw \end{bmatrix}, \quad \mathbf{F}_a = \begin{bmatrix} vu \\ v^2 \\ vw \end{bmatrix}, \quad \mathbf{G}_a = \begin{bmatrix} wu \\ wv \\ w^2 \end{bmatrix},$$

$$\mathbf{E}_v = \frac{1}{\text{Re}} \begin{bmatrix} \partial u / \partial x \\ \partial v / \partial x \\ \partial w / \partial x \end{bmatrix}, \quad \mathbf{F}_v = \frac{1}{\text{Re}} \begin{bmatrix} \partial u / \partial y \\ \partial v / \partial y \\ \partial w / \partial y \end{bmatrix}, \quad \mathbf{G}_v = \frac{1}{\text{Re}} \begin{bmatrix} \partial u / \partial z \\ \partial v / \partial z \\ \partial w / \partial z \end{bmatrix}, \quad \mathbf{P}_1 = \begin{bmatrix} p \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix}, \quad \mathbf{P}_3 = \begin{bmatrix} 0 \\ 0 \\ p \end{bmatrix},$$

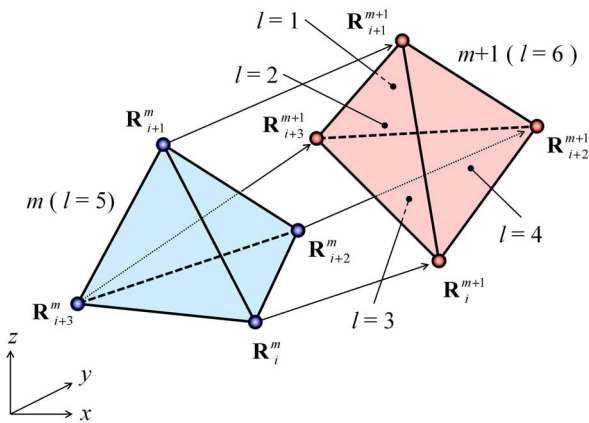
where  $\mathbf{q}$  is the vector of conserved variables,  $\mathbf{E}_a$ ,  $\mathbf{F}_a$  and  $\mathbf{G}_a$  are the convective flux vectors,  $\mathbf{E}_v$ ,  $\mathbf{F}_v$  and  $\mathbf{G}_v$  are the viscous flux vectors,  $u$ ,  $v$  and  $w$  are the fluid velocity,  $p$  is the fluid pressure and  $\text{Re}$  is Reynolds number. The equations are nondimensionalized by

$$x = \frac{\bar{x}}{\bar{L}_0}, \quad y = \frac{\bar{y}}{\bar{L}_0}, \quad z = \frac{\bar{z}}{\bar{L}_0}, \quad u = \frac{\bar{u}}{\bar{U}_0}, \quad v = \frac{\bar{v}}{\bar{U}_0}, \quad w = \frac{\bar{w}}{\bar{U}_0}, \quad p = \frac{\bar{p}}{\bar{\rho}\bar{U}_0^2}, \quad t = \frac{\bar{t}}{\bar{L}_0/\bar{U}_0}, \quad \text{Re} = \frac{\bar{\rho}\bar{L}_0\bar{U}_0}{\bar{\mu}},$$

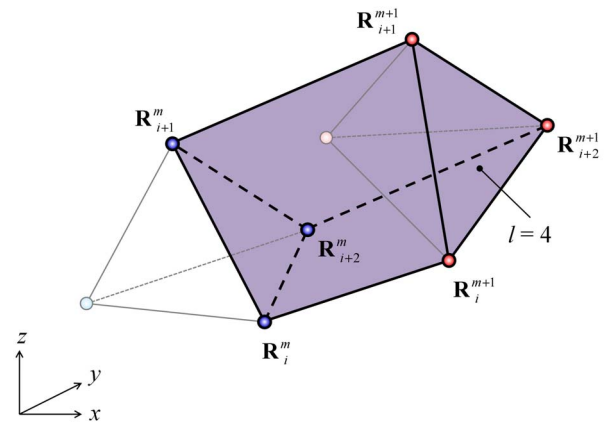
where overline shows the dimensional quantity,  $\bar{L}_0$  is the characteristics length,  $\bar{U}_0$  is the characteristics velocity,  $\bar{\rho}$  is the characteristics density and  $\bar{\mu}_0$  is the viscosity.

### Discretization

In Moving-Grid Finite-Volume Method (MGFVM) is used for discretization. In MGFVM, the control volume in the space-time unified four-dimensional ( $x, y, z, t$ ) domain is used in order to assure both physical and geometrical conservation laws simultaneously. Fig. 1 shows the Unstructured computational cell on three-dimensional ( $x, y, z$ ) domain at  $m$  time step and  $m+1$  time step.  $\mathbf{R} = (x, y, z)^T$ , the subscript  $i$  shows the computational grid number and the superscript  $m$  shows time step. In four-dimensional domain, the blue computational cell is the surface ( $l = 5$ ) perpendicular to  $t$ -axis at  $m$  time step and the red computational cell is the surface ( $l = 6$ ) perpendicular to  $t$ -axis at  $m+1$  time step. The control volume  $\Omega$  is a volume on four-dimensional domain and formed between the lower surface ( $l = 5$ ) and the upper surface ( $l = 6$ ). The control surface is the surface of the control volume on a unified four-dimensional space-time ( $x, y, z, t$ ) domain and corresponds the volume on three-dimensional domain. The control surface is formed by the surface at  $m$  time step and it at  $m+1$  time step ( $l = 1, 2, \dots, 4$ ), corresponds the computational cell at  $m$  time step ( $l = 5$ ) and corresponds the computational cell at  $m+1$  time step ( $l = 6$ ). For example, the control surface  $l = 4$  corresponds the volume on three-dimensional domain shown by the heavy line in Fig. 2.



**Figure 1. Computational cells at  $m$  time step and  $m+1$  time step**



**Figure 2. Control surface  $l = 4$**

Navier-Stokes equations for incompressible flows are discretized with Unstructured Moving-Grid Finite-Volume Method. Eq.(2) is integrated over the control volume  $\Omega$  as

$$\int_{\Omega} \left[ \frac{\partial(\mathbf{E} + \mathbf{P}_1)}{\partial x} + \frac{\partial(\mathbf{F} + \mathbf{P}_2)}{\partial y} + \frac{\partial(\mathbf{G} + \mathbf{P}_3)}{\partial z} + \frac{\partial \mathbf{q}}{\partial t} \right] d\Omega = \mathbf{0}. \quad (3)$$

Eq.(3) is shown in divergence as

$$\int_{\Omega} \left[ \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial t} \right) \{ (\mathbf{E} + \mathbf{P}_1), (\mathbf{F} + \mathbf{P}_2), (\mathbf{G} + \mathbf{P}_3), \mathbf{q} \} \right] d\Omega = \mathbf{0}. \quad (4)$$

Gauss' divergence theorem is applied to four-dimensional domain and Eq.(4) becomes as follows:

$$\sum_{l=1}^6 \{ (\mathbf{E} + \mathbf{P}_1) \tilde{n}_x + (\mathbf{F} + \mathbf{P}_2) \tilde{n}_y + (\mathbf{G} + \mathbf{P}_3) \tilde{n}_z + \mathbf{q} \tilde{n}_t \}_l S_l = \mathbf{0}, \quad (5)$$

where the subscript  $l$  shows the control surface number as Fig. 2. In the control surface,  $(\tilde{n}_x, \tilde{n}_y, \tilde{n}_z, \tilde{n}_t)$  are the outward unit normal vectors,  $S$  is magnitude of the normal vector,  $\tilde{n}_t S$  corresponds to volume in  $(x, y, z)$  domain,  $\tilde{n}_x S$  corresponds to volume in  $(y, z, t)$  domain,  $\tilde{n}_y S$  corresponds to volume in  $(z, t, x)$  domain,  $\tilde{n}_z S$  corresponds to volume in  $(t, x, y)$  domain. In the control surface  $l = 5$  and  $6$ ,  $\tilde{n}_x = \tilde{n}_y = \tilde{n}_z = 0$  and  $\tilde{n}_t S$  corresponds to the volume  $V$  of the computational cell.  $\mathbf{E}$ ,  $\mathbf{F}$ ,  $\mathbf{G}$ ,  $\mathbf{q}$ ,  $\mathbf{P}_1$ ,  $\mathbf{P}_2$  and  $\mathbf{P}_3$  are evaluated at  $m+1/2$  time step ( $l = 1, 2, \dots, 4$ ), at  $m$  time step ( $l = 5$ ) and  $m+1$  time step ( $l = 6$ ).

Therefore Eq.(5) becomes as follows:

$$\mathbf{q}^{m+1} V^{m+1} - \mathbf{q}^m V^m + \sum_{l=1}^4 \{ \mathbf{q}^{m+1/2} \tilde{n}_t + (\mathbf{E} + \mathbf{P}_1)^{m+1/2} \tilde{n}_x + (\mathbf{F} + \mathbf{P}_2)^{m+1/2} \tilde{n}_y + (\mathbf{G} + \mathbf{P}_3)^{m+1/2} \tilde{n}_z \}_l S_l = 0. \quad (6)$$

This equation is the fundamental equation of Unstructured Moving-Grid Finite-Volume Method.

### Fractional Step Method

By fractional step method, Eq.(6) is divided as

1st step:

$$\mathbf{q}^* V^{m+1} - \mathbf{q}^m V^m + \frac{1}{2} \left[ \sum_{l=1}^4 \{ (\mathbf{q}^* + \mathbf{q}^m) \tilde{n}_t + (\mathbf{E}^* + \mathbf{E}^m) \tilde{n}_x + (\mathbf{F}^* + \mathbf{F}^m) \tilde{n}_y + (\mathbf{G}^* + \mathbf{G}^m) \tilde{n}_z \}_l S_l \right] = 0, \quad (7)$$

2nd step:

$$(\mathbf{q}^{m+1} - \mathbf{q}^*) V^{m+1} + \sum_{l=1}^4 (\mathbf{P}_1^{m+1/2} \tilde{n}_x + \mathbf{P}_2^{m+1/2} \tilde{n}_y + \mathbf{P}_3^{m+1/2} \tilde{n}_z)_l S_l = 0, \quad (8)$$

where \* shows intermediate step.

The divergence of Eq.(8) on three-dimensional  $(x, y, z)$  domain at  $m+1$  time step becomes as follows:

$$(D^{m+1} - D^*)V^{m+1} + \sum_{l=1}^4 \left( \frac{\partial p^{m+1/2}}{\partial x^{m+1}} \tilde{n}_x + \frac{\partial p^{m+1/2}}{\partial y^{m+1}} \tilde{n}_y + \frac{\partial p^{m+1/2}}{\partial z^{m+1}} \tilde{n}_z \right)_l S_l = 0, \quad (9)$$

with

$$D^{m+1} = \frac{\partial u^{m+1}}{\partial x^{m+1}} + \frac{\partial v^{m+1}}{\partial y^{m+1}} + \frac{\partial w^{m+1}}{\partial z^{m+1}}, \quad D^* = \frac{\partial u^*}{\partial x^{m+1}} + \frac{\partial v^*}{\partial y^{m+1}} + \frac{\partial w^*}{\partial z^{m+1}}.$$

Here, pay attention that this is the divergence on  $(x, y, z)$  domain at  $m+1$  time step in order to correlate to the continuity equation Eq.(1) at  $m+1$  time step. Assuming that the continuity equation Eq.(1) is satisfied at  $m+1$  time step ( $D^{m+1} = 0$ ), Eq.(9) becomes the pressure equation including the normal vectors on four-dimensional  $(x, y, z, t)$  domain as follows:

$$-D^*V^{m+1} + \sum_{l=1}^4 \left( \frac{\partial p^{m+1/2}}{\partial x^{m+1}} \tilde{n}_x + \frac{\partial p^{m+1/2}}{\partial y^{m+1}} \tilde{n}_y + \frac{\partial p^{m+1/2}}{\partial z^{m+1}} \tilde{n}_z \right)_l S_l = 0, \quad (10)$$

where the differential of pressure on three-dimensional  $(x, y, z)$  domain at  $m+1$  time step is solved with finite-volume-method on the computational cell at  $m+1$  time step.

The computational procedure is as follows:

1.  $q^*$  is calculated from  $q^m$  by Eq.(7).
2.  $p^{m+1/2}$  is calculated from  $q^*$  by Eq.(10).
3.  $q^{m+1}$  is calculated from  $q^*$  and  $p^{m+1/2}$  by Eq.(8).

#### Other Numerical Method

The convective flux vectors ( $\mathbf{E}_a, \mathbf{F}_a, \mathbf{G}_a$ ) are evaluated with second order upwind difference scheme. The viscous flux vectors ( $\mathbf{E}_v, \mathbf{F}_v, \mathbf{G}_v$ ) and the pressure vectors ( $\mathbf{P}_1, \mathbf{P}_2, \mathbf{P}_3$ ) are evaluated with central difference scheme. The iterative method of Eq.(7) is LU-SGS [Yoon and Jameson (1988)] and the iterative method of Eq.(10) is Bi-CGSTAB [van der Vorst (1992)].

### Numerical methods for body motion

#### Governing Equation

The governing equations of body-motion are Newton's motion equation with six degrees of freedom of motion including translation and rotation as follows:

$$\frac{d}{dt} \begin{bmatrix} Mu_B \\ Mv_B \\ Mw_B \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ F_z - Mg \end{bmatrix}, \quad (11)$$

$$\frac{d}{dt} \begin{bmatrix} x_B \\ y_B \\ z_B \end{bmatrix} = \begin{bmatrix} u_B \\ v_B \\ w_B \end{bmatrix}, \quad (12)$$

$$\frac{d}{dt} \begin{bmatrix} I_x \omega_x \\ I_y \omega_y \\ I_z \omega_z \end{bmatrix} + \begin{bmatrix} (I_z - I_y) \omega_z \omega_y \\ (I_x - I_z) \omega_x \omega_z \\ (I_y - I_x) \omega_y \omega_x \end{bmatrix} = \begin{bmatrix} T'_x \\ T'_y \\ T'_z \end{bmatrix}, \quad (13)$$

where  $M$  is the weight of the body,  $u_B$ ,  $v_B$  and  $w_B$  are the body velocity,  $F_x$ ,  $F_y$  and  $F_z$  are the forces to be received by the fluid,  $g$  is gravity acceleration,  $I_x$ ,  $I_y$  and  $I_z$  are the moment of inertia of the body about a center of  $x'$ ,  $y'$  and  $z'$  axis,  $\omega_x$ ,  $\omega_y$  and  $\omega_z$  are the body angle speed about a center of  $x'$ ,  $y'$  and  $z'$  axis,  $T_x$ ,  $T_y$  and  $T_z$  are the torques to be received by the fluid about a center of  $x'$ ,  $y'$  and  $z'$  axis and  $(x', y', z')$  is the cartesian coordinate fixed to the body. The equations are nondimensionalized by

$$\begin{aligned} M &= \frac{\bar{M}}{\bar{\rho}\bar{L}_0^3}, & u_B &= \frac{\bar{u}_B}{\bar{U}_0}, & v_B &= \frac{\bar{v}_B}{\bar{U}_0}, & w_B &= \frac{\bar{w}_B}{\bar{U}_0}, & F_x &= \frac{\bar{F}_x}{\bar{\rho}\bar{U}_0^2/\bar{L}_0^2}, & F_y &= \frac{\bar{F}_y}{\bar{\rho}\bar{U}_0^2/\bar{L}_0^2}, & F_z &= \frac{\bar{F}_z}{\bar{\rho}\bar{U}_0^2/\bar{L}_0^2} \\ g &= \frac{\bar{g}}{\bar{U}_0^2/\bar{L}_0}, & I_x &= \frac{\bar{I}_x}{\bar{\rho}\bar{L}_0}, & I_y &= \frac{\bar{I}_y}{\bar{\rho}\bar{L}_0}, & I_z &= \frac{\bar{I}_z}{\bar{\rho}\bar{L}_0}, & \omega_x &= \frac{\bar{\omega}_x}{\bar{U}_0/\bar{L}_0}, & \omega_y &= \frac{\bar{\omega}_y}{\bar{U}_0/\bar{L}_0}, & \omega_z &= \frac{\bar{\omega}_z}{\bar{U}_0/\bar{L}_0}, \\ T_x &= \frac{\bar{T}_x}{\bar{\rho}\bar{U}_0^2/\bar{L}_0}, & T_y &= \frac{\bar{T}_y}{\bar{\rho}\bar{U}_0^2/\bar{L}_0}, & T_z &= \frac{\bar{T}_z}{\bar{\rho}\bar{U}_0^2/\bar{L}_0}, \end{aligned}$$

where overline shows the dimensional quantity. Quaternion [Yatabe (2007)] is used for the coordinate transform.

The time derivative of Eq.(11) and (13) is discretized by forward Euler method and the time derivative of Eq.(12) is discretized by Crank-Nicolson method as follows:

$$\begin{cases} Mu_B^{m+1} = Mu_B^m + F_x^m \Delta t \\ Mv_B^{m+1} = Mv_B^m + F_y^m \Delta t \\ Mw_B^{m+1} = Mw_B^m + (F_z^m - Mg) \Delta t \end{cases}, \quad (14)$$

$$\begin{cases} x_B^{m+1} = x_B^m + 0.5(u_B^{m+1} + u_B^m) \Delta t \\ y_B^{m+1} = y_B^m + 0.5(v_B^{m+1} + v_B^m) \Delta t, \\ z_B^{m+1} = z_B^m + 0.5(w_B^{m+1} + w_B^m) \Delta t \end{cases}, \quad (15)$$

$$\begin{cases} I_x \omega_x^{m+1} = \omega_x^m + (T_x^m - (I_z - I_y) \omega_z^m \omega_y^m) \Delta t \\ I_y \omega_y^{m+1} = \omega_y^m + (T_y^m - (I_x - I_z) \omega_x^m \omega_z^m) \Delta t. \\ I_z \omega_z^{m+1} = \omega_z^m + (T_z^m - (I_y - I_x) \omega_y^m \omega_x^m) \Delta t \end{cases} \quad (16)$$

### Coupled Procedure of Fluid-Flow and Body-Motion

The computational procedure is as follows:

1. The force and the torque of the body at  $m$  time step are calculated from the pressure and the shear stress of the fluid.
2. The velocity and the angle speed of the body at  $m+1$  time step are calculated by Eq.(14) and Eq.(16).
3. The translation and the rotation of the body at  $m+1$  time step are calculated by Eq.(15) and quaternion.
4. The computational grid at  $m+1$  time step is formed.
5. The velocity and the pressure of the fluid at  $m+1$  time step are calculated by Eq.(7), Eq.(8) and Eq.(10).

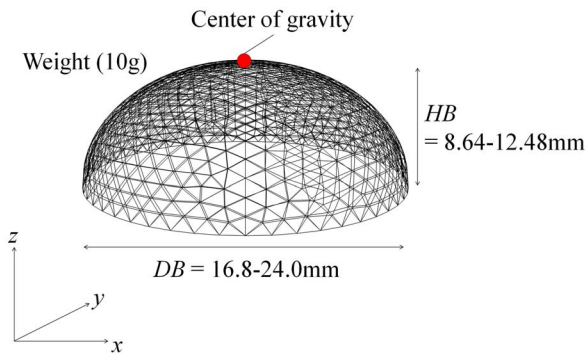
### Numerical simulation

### Computational Condition

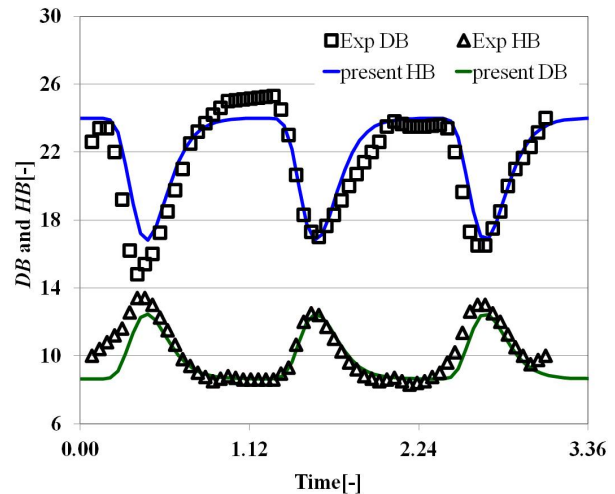
Fig. 3 shows the jellyfish which is the target of this simulation. The jellyfish is 16.8-24.0mm in diameter  $DB$ , 8.64-12.48mm in height  $HB$ , 1mm in thickness, 10g in weight and elliptical cross section. The water is  $1000\text{kg/m}^3$  in density and  $1.0 \cdot 10^{-3}\text{kg/ms}$  in viscosity.  $\bar{L}_0$  is 24.0mm,  $\bar{U}_0$  is 24.0mm/s,  $\bar{\rho}$  is  $1000\text{kg/m}^3$ ,  $\bar{\mu}_0$  is  $1.0 \cdot 10^{-3}\text{kg/ms}$ ,  $Re$  is 576,  $M$  is 10g and  $g$  is  $0.0\text{m/s}^2$  considering buoyancy. Case1 is condition without a current and Case2 is condition with a current. The jellyfish velocity  $(u_B, v_B, w_B)$  is  $(0.0, 0.0, 0.0)$  in the initial condition. The change of diameter  $DB$  and height  $HB$  is decided as follows:

$$\begin{cases} \frac{DB}{\bar{L}_0} = 1.0 - 0.3 \exp\left(-\text{Ln}\left(\frac{[t \bmod 1.12]}{0.44}\right)^2 / 0.15\right) \\ \frac{HB}{\bar{L}_0} = 0.36 - 0.16 \exp\left(-\text{Ln}\left(\frac{[t \bmod 1.12]}{0.44}\right)^2 / 0.15\right) \end{cases} \quad (14)$$

Fig.4 shows time history of diameter  $DB$  and height  $HB$ , where Exp denotes the experimental data [Mchenry and Jed (2003)].

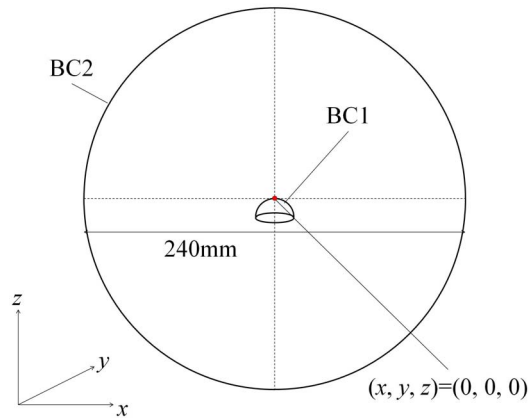


**Figure 3. Jellyfish**



**Figure 4. Time history of diameter  $DB$  and height  $HB$ .**

Fig. 5 shows the computational domain and the boundary conditions. The computational domain is a sphere 240mm in diameter. In BC1 which is the wall boundary of the jellyfish, the velocity is fixed to the velocity of the jellyfish and Neumann boundary condition applies to the pressure. In BC2 which is the external boundary, the inflow velocity is fixed to  $(u, v, w) = (u_i, v_i, w_i)$ , the outflow velocity is calculated by linear interpolation and Neumann boundary condition applies to the pressure. The whole of computational grid moves together with the jellyfish by using Moving Computational Domain (MCD) approach [Watanabe and Matsuno (2009)]. Simulations are performed in three conditions Case1, Case2 and Case3 as shown in Table 1.



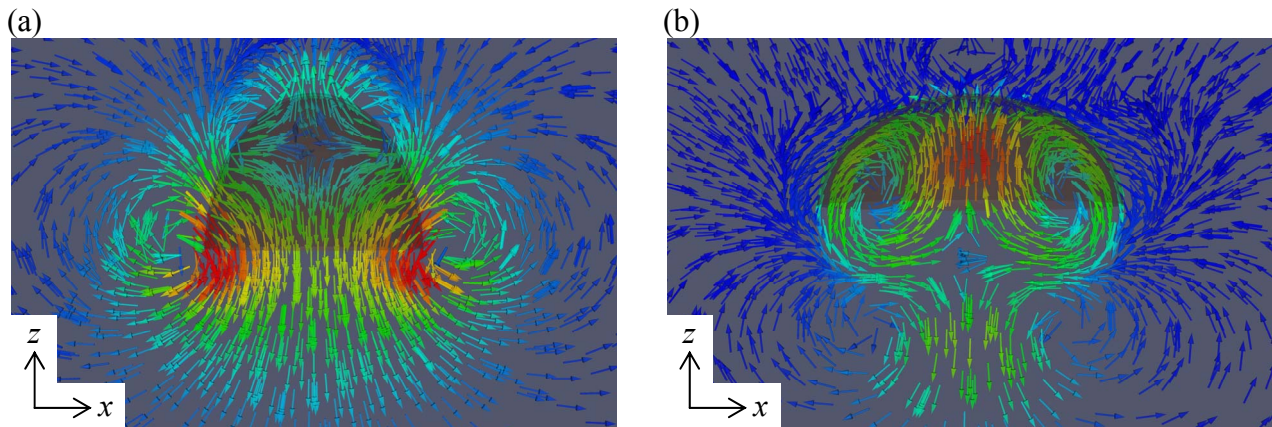
**Figure 5. Computational domain and boundary conditions**

**Table 1. Conditions of simulation**

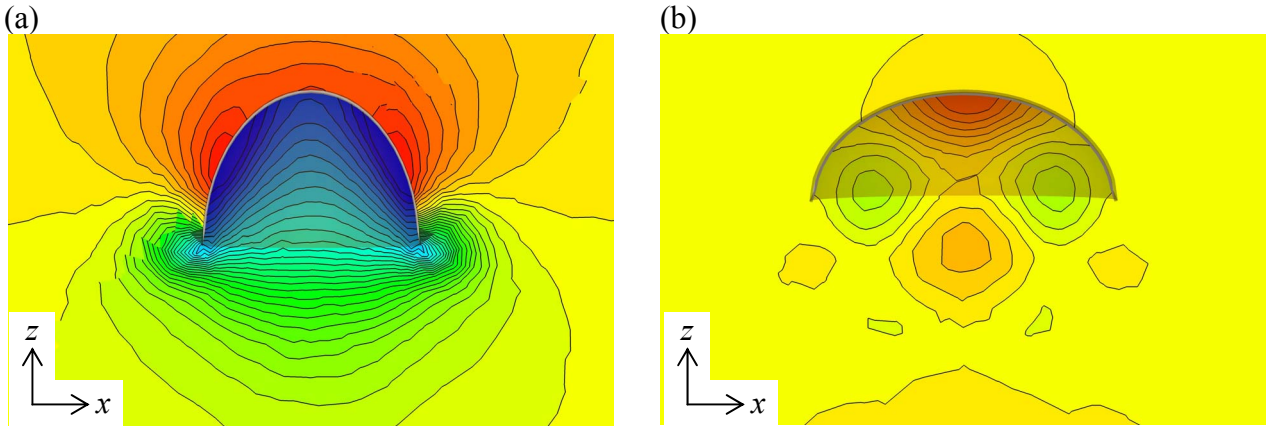
	Case1	Case2	Case3
Current	Without	With	With
$(u_i, v_i, w_i)$	(0.0, 0.0, 0.0)	(0.5, 0.0, 0.0)	(0.5, 0.0, 0.0)
Diameter $DB$	Eq.(14)	Eq.(14)	Constant (24.0mm)
Height $HB$	Eq.(14)	Eq.(14)	Constant (8.64mm)

*Result of Simulation*

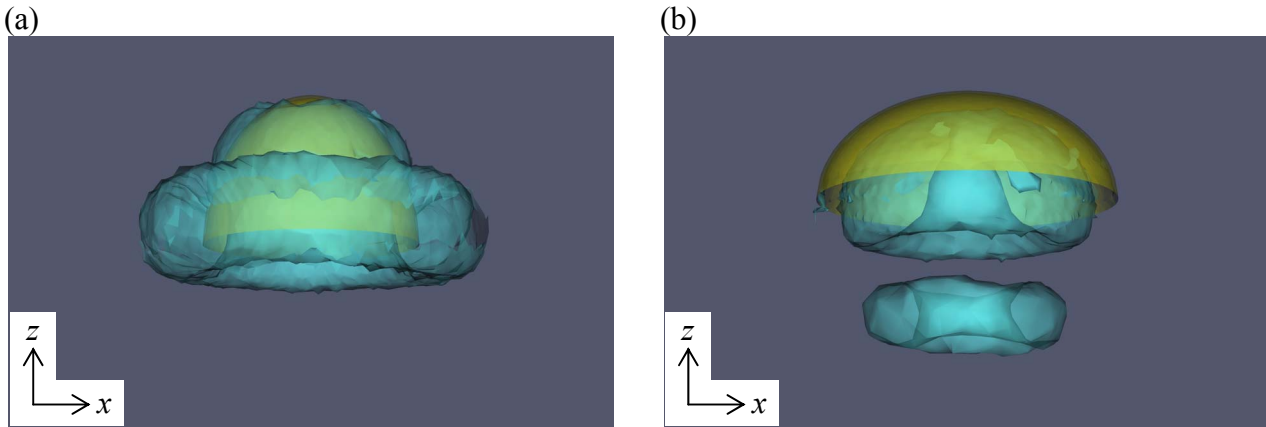
As a result, Fig. 6, 7 and 8 each show velocity vectors, pressure contours and vorticity magnitude contours in Case1 at  $t = 0.40$  (contraction) and  $t = 1.12$  (relaxation). Colors denote magnitude velocity in Fig. 6. The contour denotes 4.0 of vorticity magnitude in Fig. 8. The starting vortex occurs outside the jellyfish at the contraction and the stopping vortex occurs inside the jellyfish at the relaxation. The whole flow field is axisymmetric.



**Figure 6. Velocity vectors in Case1 at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**

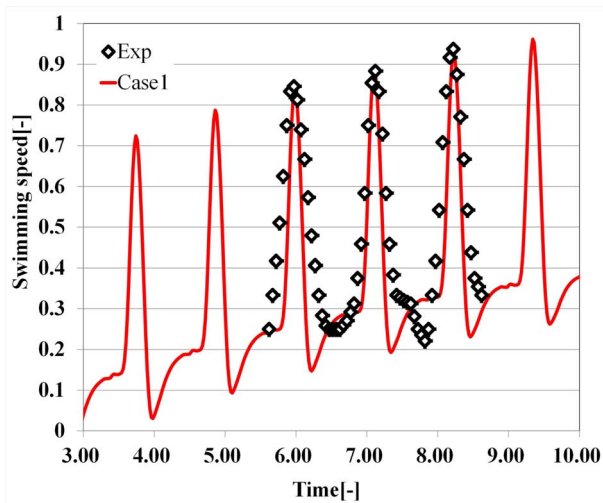


**Figure 7. Pressure contours in Case 1 at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**



**Figure 8. Vorticity magnitude contours in Case 1 at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**

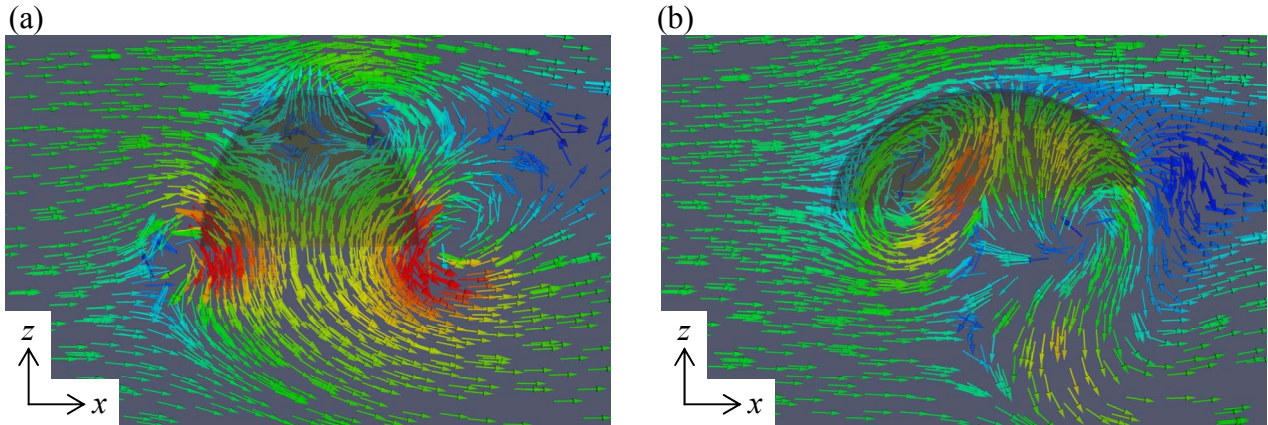
Fig. 9 shows swimming speed  $w_B$  of the jellyfish in Case 1, where Exp denotes experimental result [Mchenry and Jed (2003)]. The jellyfish accelerates at the contraction and decelerates at relaxation. The present result gives good agreement with experimental result.



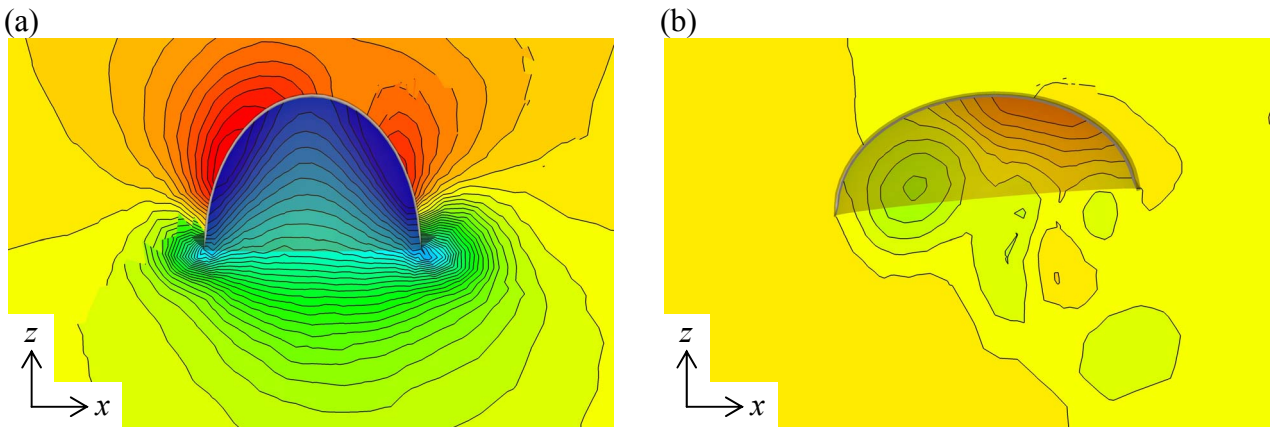
**Figure 9. Time history of swimming speed in Case 1 and experimental result**



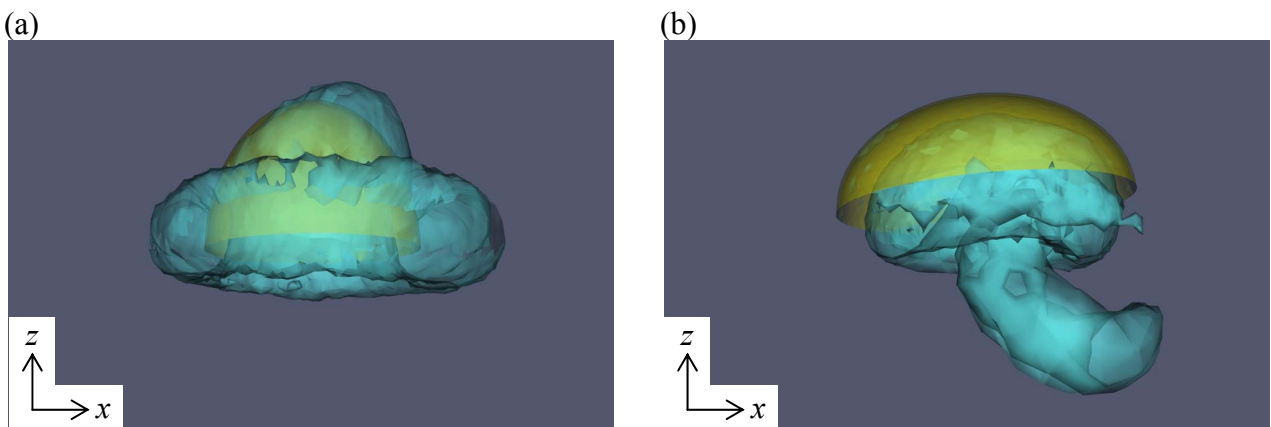
Fig. 10, 11 and 12 each show velocity vectors, pressure contours and vorticity magnitude contours in Case2 at  $t = 0.40$  (contraction) and  $t = 1.12$  (relaxation). Colors denote magnitude velocity in Fig. 10. The contour denotes 4.0 of vorticity magnitude in Fig. 12. The flow field is not axisymmetric because of the current. At relaxation, the stopping vortex is clear not vortex structure in the right side and pressure in the jellyfish is higher in the right side. The asymmetry of the vortex structure inclines the jellyfish to  $x$ -axis negative direction.



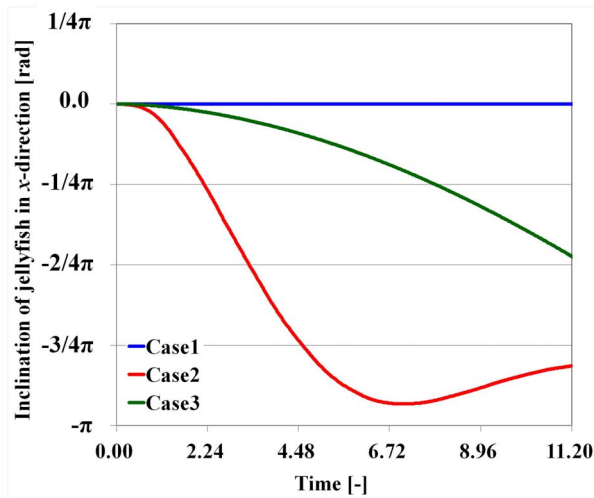
**Figure 10. Velocity vectors in Case2  
at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**



**Figure 11. Pressure contours in Case2  
at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**



**Figure 12. Vorticity magnitude contours in Case2  
at (a)  $t = 0.40$  (contraction) and (b)  $t = 1.12$  (relaxation)**



**Figure 13. Time history of inclination angle of jellyfish in x-direction in Case1 (blue line), Case2 (red line) and Case3 (green line)**

Fig. 13 shows time history of the inclination angle of the jellyfish in  $x$ -direction. In Case1 without a current and with the contraction motion, the jellyfish does not rotate because the whole flow field is axisymmetric, and the inclination angle is almost zero. In Case3 with a current and without the contraction motion, the current itself rotates the jellyfish to  $x$ -axis negative direction, and the inclination angle decreases gradually. In Case2 with a current and the contraction motion, moreover the vortex structure which becomes asymmetric by the current rotates the jellyfish to  $x$ -axis negative direction, and the inclination angle decreases faster than Case3.

## CONCLUSIONS

In this paper, we have performed three-dimensional coupled simulation of fluid-flow and jellyfish-swimming in a current with six degrees of freedom of motion by using Moving-Grid Finite-Volume Method. In the simulation without a current, the jellyfish accelerates at the contraction and decelerates at relaxation. The swimming speed of the jellyfish gave good agreement with experimental result. Moreover, in the simulation with a current, the current makes flow field asymmetric and changes the vortex structure caused by swimming jellyfish. The vortex structure rotates swimming jellyfish.

## References

- Inomoto, T. Matsuno, K. and Yamakawa, M. (2004) A moving-grid finite-volume method for three-dimensional incompressible flows, *Proceedings of JSFM 18th Computational Fluid Dynamics Symposium*, D2-3.
- Mihara, K., Matsuno, K. and Satofuka, N. (1999) An iterative finite-volume scheme on a moving grid (1st report, The fundamental formulations and validation), *Transactions of the Japan Society of Mechanical Engineers, Series B*, Vol.65, No.637, 2945-2953.
- Matsuno, K. (2001) A moving mesh finite-volume scheme for compressible flows, *Computational Fluid Dynamics 2000*, Springer, 705-710. (*Proceedings of 1st International Conference on Computational Fluid Dynamics 2000*, Kyoto, Japan, July, 2000)
- Matsuno, K. (2010) *Developments and Applications in Engineering Computational Technology*, Saxe-Coburg Publications, 103-127.
- Noh, W. F. (1964) A time-dependent, two space dimensional, coupled Eulerian-Lagrange code, *Methods in Computational Physics*, **3**, 117-179.
- Tezduyar, T. E., Behr, M. and Liou, J. (1992) A new strategy for finite element computations involving moving boundaries and interfaces - the deforming-spatial-domain/space-time procedure : I. The concept and preliminary numerical tests, *Computer Methods in Applied Mechanics and Engineering*, **94**, 339-351.
- Thomas, P. D. and Lombard, C. K. (1979) Geometric conservation law and its application to flow computations on moving grids, *AIAA Journal*, **17-10**, 1030-1037.

- van der Vorst, H.A. (1992) Bi-CGSTAB : A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, *SIAM Journal on Scientific and Statistical Computing*, **13-2**, 631-644.
- Vinokur, M. (1974) Conservation equations of gasdynamics in curvilinear coordinate systems, *Journal of Computational Physics*, **14-2**, 105-125.
- Watanabe, K. and Matsuno, K. (2009) Moving computational domain method and its application to flow around a high-speed car passing through a hairpin curve, *Journal of computational Science and Technology*, **3-2**, 449-459.
- Yamakawa, M. and Matsuno, K. (2003) An iterative finite-volume method on an unstructured moving grid (1st report, the fundamental formulation and validation for unsteady compressible flows), *Transactions of the Japan Society of Mechanical Engineers, Series B*, **69-683**, 1577-1582.
- Yatabe, M. (2007) Handy Note for Quaternion, *MSS*, **18**, 29-34.
- Yoon, S. and Jameson, A. (1988) Lower-upper Symmetric-Gauss-Seidel method for the Euler and Navier-Stokes equations, *AIAA Journal*, **26-9**, 1025-1026.
- Costello, J. H., Colin, S. P. and Dabiri, J. O. (2008) Medusan morphospace: phylogenetic contrasts, biomechanical solutions, and ecological consequences, *Invertebrate Biology*, **127**, 265–290.
- Colin, S. P. and Costello, J. H. (2002) Morphology, swimming performance and propulsive mode of six co-occurring hydromedusae. *Journal of Experimental Biology*, **205**, 427–437.
- Mchenry, M. J. and Jed, J. (2003) The ontogenetic scaling of hydrodynamics and swimming performance in jellyfish (*Aurelia aurita*), *Journal of Experimental Biology*, **206**, 4125–4137.
- Dabiri, J. O., Colin, S. P., Costello, J. H. and Gharib, M. (2005) Flow patterns generated by oblate medusa jellyfish: field measurements and laboratory analyses, *Journal of Experimental Biology*, **208**, 1257–1265.
- Dular, M., Bajcar, T. and Sirok, B. (2009) Numerical investigation of flow in the vicinity of a swimming jellyfish, *Engineering Applications of Computational Fluid Mechanics*, **3**, 258–270.
- Sahin, M. and Mohseni, K. (2009) An arbitrary Lagrangian–Eulerian formulation for the numerical simulation of flow patterns generated by the hydromedusa *Aequorea victoria*, *Journal of Computational Physics*, **228**, 4588–4605.
- Alben, S., Miller, L. A. and Peng, J. (2013) Efficient kinematics for jet-propelled swimming, *Journal of Fluid Mechanics*, **733**, 100–133.
- Rudolf, D. and Mould, D. (2010) An interactive fluid model of jellyfish for animation, *Communications in Computer and Information Science*, **68**, 59–72.

## **A reliability optimization allocation method considering differentiation of functions**

**X.J. Yi<sup>1</sup>, Y.H. Lai<sup>1</sup>, \*†H.P. Dong<sup>1</sup>, and P. Hou<sup>2</sup>**

<sup>1</sup>School of Mechatronical Engineering, Beijing Institute of Technology, China

<sup>2</sup>Taiyuan Branch of China CNR Corporation Limited, China

\*Presenting author: reliabilitybit@sina.com

†Corresponding author: reliabilitybit@sina.com

### **Abstract**

The reliability optimization has achieved great concern in recent years. Nowadays, many researchers obtain allocation results which can maximize the system reliability subject to the system budget. In these researches, the effect of system's functions is always neglected or only considering the single main function of system. In addition, there are also not obvious evidences in results to distinguish the importance level of different units. However, complex systems tend to perform multiple functions. What's more, the use frequency of each function and the combinations of units to realize different function is not the same. In addition, the use demand of different function is decided by different task environment, the demand differentiation of functions has led to the use frequency of various functions should have different levels about reliability. Therefore, the reliability optimization allocation that only considering cost constraint conditions is not accuracy and will results in disaccord between the obtained results with actual situation.

Focusing on the problem mentioned above, a reliability optimization allocation method that considering cost constraint and importance factor is proposed. In this paper we consider systems consisting of units characterized by different reliability and importance factors. Such systems are multi-function because they must perform different tasks depending on the combination of units. Different functions may be work simultaneously. Firstly, the concept of importance factor is defined to describe the importance of a unit and the required importance factor level of system functions in the task is also given. To deal with the differentiation of system functions, the corresponding bound about importance factor are executed when looking for the optimal solution. Similarly, the cost constraint is also forced. Finally, in order to reduce the randomness of intelligent algorithm, a number of optimization are conducted and a rule is proposed to select the most optimal solution from all the optimal solutions which are obtained in every iterative loop.

Example of the comprehensive transmission equipment is presented. To begin with, we establish the reliability function of system as the objective optimization function. Then, the restraint of budget and different demands of importance factor of system functions are posed. Furthermore, using a genetic algorithm as the optimization tool, the optimization result can be obtained. Finally, the most optimal solution is selected. The results show that the method we proposed is more correct and more approximate the reality.

The reliability optimization allocation method presented in this paper can not only consider the constraint of cost, but also can consider the diversities of functions, and thus the optimization results will more approximate actual situation. At the same time, this paper can also provide guidance for the similar reliability optimization problem.

**Keywords:** reliability optimization, differentiation, importance factor, system function, genetic algorithm

## Introduction

Reliability is an important index of a system, and system reliability allocation is prerequisite for reliability design and analysis of system. The reliability design of system must be done under several constraints such as cost, volume, weight etc. In engineering, some researchers often make reliability allocation which can maximize the reliability of system under the condition of budget constraint. The early researchers also study the method of reliability allocation for all kinds of system structure, such as series, parallel, series-parallel etc., and the related reliability allocations subject to the budget of system are done under the condition that units of a system are independent of each other: [Chen (1992); Bueno (2005); Levitin and Lisnianski (1999); Ramirez-marquez and Coit (2004)]. In these research work for reliability allocation, people often only consider one single main function of the system and ignore other functions. That may lead to a great error for system reliability allocation. G. Levitin etc. conduct the research about the reliability allocation of multi-state system, and they make a conclusion that all units in the system must be involved in to realize any state of the system while the reliability of system is allocated to all the units: [Levitin and Lisnianski (2003); Levitin (2003); Levitin et al. (2003); Levitin and Lisnianski (2001)]. However, with the development of science and technology, the structure of the system is more and more complex and the system that has multiple main functions is also more and more popular. In these complex systems, the usage frequency of each function and the combination of units to realize different functions are not the same. In addition, different use demand of each system function is determined by different environment of task, and the differentiation of use demand of each system function results in that the reliability level of various functions should be different. But the existing research results cannot solve the reliability optimization allocation problem of systems with differentiation of functions. Meanwhile, as the complexity of reliability model of system and the complex relation between reliability index and constraints, it is difficult to obtain the solution for the reliability optimization allocation problem. A large number of researches can prove that applying artificial intelligence methods (such as neural networks, ant colony algorithm, genetic algorithm, hybrid genetic algorithm, simulated annealing algorithm, etc.) into reliability optimization allocation can achieve good effect: [Altiparmak et al. (2003); Nabil and Mustapha (2005); Chen and You (2005)]. Way Kuo made a good overview about the related content of reliability allocation before 2006: [Kuo and Rin (2007)]. Nevertheless, the artificial intelligence methods mentioned above also have their disadvantages, such as: premature phenomena may occur in the course of using, swinging near optimal solution when approaching optimal solution, slow convergence and easy to fall into local extremum, or optimization result is not ideal, and so on. Therefore, it is meaningful to improve the basic algorithm so that it is applicable for specific problems and can obtain the optimal solution efficiently.

In order to solve the above problem, a reliability optimization allocation method for systems with differentiation of functions is proposed in this paper. Firstly, a constraint considering the differentiation of functions is presented, namely a new index named importance factor is used to measure the usage frequency of each function and the use demand of each function determined by different environment of task. Then, the objective function of optimal allocation of system reliability is built according to the new cost model proposed in this paper, and further put forward the description of reliability optimization allocation of systems with differentiation of functions. In addition, the operation of code, evolution and selection in the genetic algorithm is also transformed to improve the accuracy and efficiency of solving the optimization problem. Finally, the proposed method is applied into the reliability optimization allocation of an integrated transmission device

whose goal is to minimize the cost of system. The result shows that the obtained solution is more in line with the actual situation.

### A reliability optimization allocation method considering differentiation of functions

Assumption that a system that consists of  $m$  units can execute  $n$  functions.

#### Establishment of objective function

The reliability allocation is usually done according to the requirement of task. Not only the final reliability of system must meet a certain standard  $R_s$ , but also the allocation results must minimize the cost of system. In fact, reliability and cost are often nonlinear relationship. A cost model is presented under the assumption that their relationship submits to exponential distribution, it is shown as Eq. (1).

$$c_i(P_i, R_i, R_{i,\min}) = P_i e^{\left(\frac{R_i}{R_{i,\min}} - 1\right)} \quad (1)$$

Where,  $c_i(\cdot)$  represents the cost model of  $i$  th unit,  $P_i$  is the basic price of  $i$  th unit,  $R_i$  is the reliability of  $i$  th unit,  $R_{i,\min}$  is the lower limit value of reliability of  $i$  th unit. The total cost of system is the objective function of reliability optimization allocation, and its mathematical form is shown as Eq. (2).

$$C_s(R) = \sum_{i=1}^m c_i(P_i, R_i, R_{i,\min}) \quad (2)$$

Where,  $C_s(\cdot)$  is the total cost function of system.

#### Constraint considering the differentiation of functions

To deal with the differentiation of functions in the process of reliability allocation, the more reasonable approach is to convert it into a constraint condition and to select an optimal solution in the optimization process. As to it, a new index named importance factor is proposed. Obviously, the units in a system not only have reliability index, but also have the parameter of importance factor. The importance factor is used to measure the importance of a unit or its usage frequency. The importance factor is related to the reliability of the unit. Generally, the higher usage frequency of a unit, the greater importance of the unit and the higher reliability allocated to the unit. Therefore, according to the positive correlation between importance factor and reliability, a model is proposed based on the experience in engineering application, shown as in Eq.(3).

$$g_i(R_i) = e^{(R_i - 1)} \quad 0 < g < 1 \quad (3)$$

Where,  $g_i(\cdot)$  is the importance factor of  $i$  th unit.

When importance factor is applied to deal with the differentiation of functions, if the importance factor of  $j$  th function cannot reach the scheduled request, as shown as in Eq.(4).

$$F_{f_j}(G_{f_j}, W_{f_j}) < 0 \quad (4)$$

The Eq. (4) indicates the allocated result is disqualified, and new iteration has to be done to make it meet the requirement, as shown as in Eq.(5).

$$F_{f_j}(G_{f_j}, W_{f_j}) \geq 0 \quad (5)$$

Where,  $F_{f_j}$  is a function to determine whether  $j$  th function can be realized or not,  $G_{f_j}$  is the importance factor of  $j$  th function,  $W_{f_j}$  is the predetermined level of importance factor of  $j$  th function.

Because a function of system is implemented by combination of units and importance factor reflects the important degree of each unit, the calculation method of importance factor of each function is shown as in Eq. (6).

$$G_{f_j} = \prod_{i=1}^l g_i(R_i) \quad k \in [1, l] \quad l \leq m \quad (6)$$

Where,  $l$  is the number of units involved in  $j$  th function.

Compare  $G_{f_j}$  with  $W_{f_j}$ , the state of  $j$  th function can be determined. When a system has several functions, only all functions can be realized successfully does a system succeed. In other words, the importance factor of each function should satisfy the requirement of importance factor of itself, shown as in Eq. (7).

$$F_{f_j} \geq 0 \quad j \in [1, n] \quad (7)$$

#### *Problem formulation*

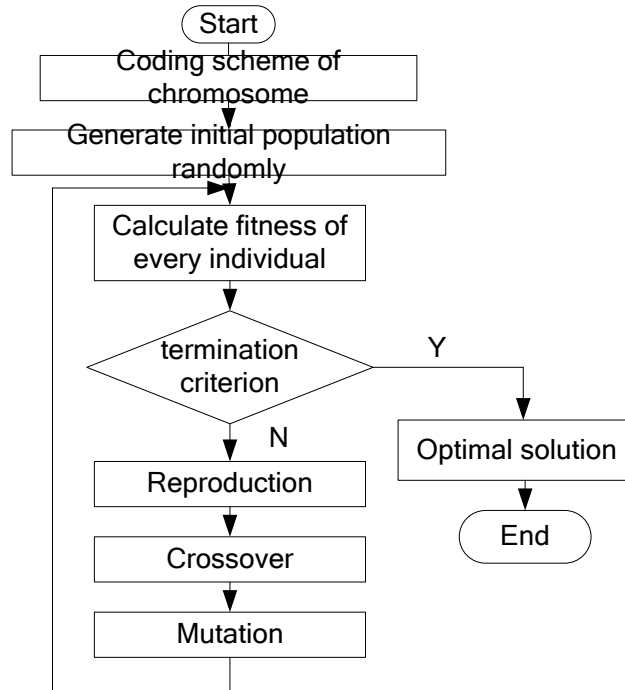
Combining the objective function and constraint conditions mentioned above, the reliability optimization allocation problem with differentiation of functions can be formulated as in Eq. (8).

$$\left\{ \begin{array}{l} \min C_S(R) = \sum_{i=1}^m C_i(R_i) = \sum_{i=1}^m P_i e^{\left(\frac{R_i}{R_{i,\min}} - 1\right)} \\ s.t. \\ R_{i,\min} \leq R_i \leq R_{i,\max} \\ \prod_{i=1}^m R_i \geq R_s \quad i = 1, 2, \dots, m \\ F_{f_j}(G_{f_j}, W_{f_j}) \geq 0 \quad j \in [1, n] \end{array} \right. \quad (8)$$

Where,  $R_s$  is the reliability goal of system which needs to be achieved.

#### *The improvement of genetic algorithm*

The concrete operation process of basic genetic algorithms is shown as in Fig. 1. The operation of code, evolution and selection in the genetic algorithm are also transformed to improve the accuracy and efficiency for solving the optimization problem.



**Figure 1. Operation process of basic genetic algorithm**

### 1. Coding scheme

The real number coding is used for its several advantages compared with binary coding, such as: unnecessary to be transformed from the chromosome to performance value which can improve the efficiency of the genetic algorithm, reduces the requirements on memory and the use of different genetic operator is very free and so on.

### 2. Initial population

The characteristics of the initial population have important influences on the calculated results and efficiency of calculation. To achieve the global optimal solution, initial population should be dispersed in the solution space.

In order to ensure the uniform distribution in the solution space and get good initial population, homogeneous array can be constructed artificially. Firstly, the solution space is divided into  $A$  subspaces. Then,  $B$  chromosomes are produced in each subspace based on the predetermined or random method. Finally,  $C$  chromosomes whose fitness is the largest are selected from all these chromosomes.

### 3. Design of fitness function

Fitness indicates that the ability of an individual to adapt to the environment, and how to select fitness function depends on different situations. The objective function is taken as the fitness function in this paper and the best individual adapt to the environment has the largest value of fitness, which equal to the minimum of objective function. In addition, in order to guarantee that individuals produced randomly are within the above established constraints and remove the unsuitable individuals in the process of searching, penalty strategy is adopted. For the unsuitable individual, the fitness of it is set to be zero. The formulation of penalty function is shown as in Eq. (9).



$$\left\{ \begin{array}{ll} fit(\alpha) = 0 & F_{f_j} < 0 \text{ or } \prod_{i=1}^m R_i < R_s \\ fit(\alpha) = 0.5 + 2 \cdot \frac{S - \alpha}{S - 1} & F_{f_j} \geq 0 \text{ and } \prod_{i=1}^m R_i \geq R_s \end{array} \right. \quad (9)$$

Where,  $fit(\cdot)$  is the function of fitness,  $S$  is the size number of population,  $\alpha$  is the ranking in the descending order of reliability.

#### 4. Reproduction

The value of fitness is used to judge whether the chromosome is good or not. The larger the individual's fitness is, the larger the chance of being selected is. Fitness assignment based on rank and the strategy of roulette are adopted to select the excellent individuals in this paper.

#### 5. Crossover operation

Crossover operation is the main genetic operator in genetic algorithm, which is acting on the individuals selected for reproduction. In this paper, the discrete restructuring and randomly selecting parent individuals with the same probability are selected as the crossover operation to determine the value of offspring individuals.

#### 6. Mutation

Selection and crossover basically complete most work of the search function of genetic algorithm, and the mutation can enhance the ability of genetic algorithm to find the optimal solution. The mutation operation in this paper is that an individual is selected randomly among population and the value of a character in the gene string is changed randomly with a certain probability. Variation can increase the diversity of population, and adaptive mutation is operated to make the individual whose fitness is larger to do searching in the small scope while the individual whose fitness is smaller to do searching in the wide scope.

#### 7. Termination criterion

Because genetic algorithm optimization can be unlimited evolution, and it is often difficult to have a big improvement after a certain degree of evolution, so a specific number of iteration need to be predetermined as the standard to decide whether the optimization process will go on or not. The certain generation of evolution is selected as the optimization criterion in this paper.

#### 8. Optimization of multi-population

Because selection operation may make the population evolution to be controlled by superior individual, and crossover operation might destroy the good genes, so the basic genetic algorithm has the problem of premature convergence. Therefore, the optimization of multi-population is carried out to improve the performance of genetic algorithm.

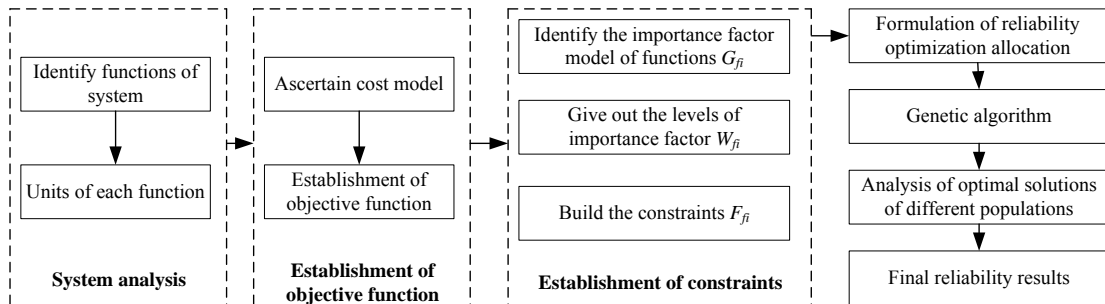
Genetic algorithm of multi-population allows the child population to evolve along different directions, and the outstanding individuals in the solution space of different area can be obtained. It also can prevent the occurrence of premature convergence from searching in expanding scope. On

the other hand, with the migration and spread of excellent individual among different populations, the convergence rate and precision of solution can be improved.

Different optimal solutions can be obtained through optimization of multi-population. According to the actual situation, we can formulate some suitable guidelines to select the final optimal solution from these different optimal solutions.

*Reliability optimization allocation process for systems with differentiation of functions under the goal of minimizing the cost of systems*

For systems with differentiation of functions, the reliability optimization allocation to minimize the cost of systems can be done according to Fig. 2.

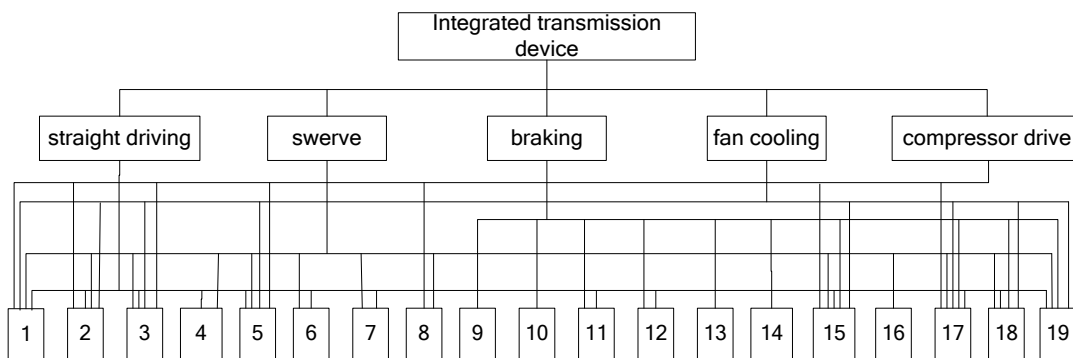


**Figure 2. Reliability optimization allocation process for systems with differentiation of functions under the goal of minimizing the cost of systems**

**Example**

*The analysis of an integrated transmission device*

There are 19 subsystems in an integrated transmission device, and they are 1-overall gearing, 2-body parts, 3-middle bracket, 4-torque converter assembly, 5-transmission assembly, 6-hydraulic torque converter, 7-planet before shift gear, 8-auxiliary drive, 9-hydraulic gear reducer, 10-hydraulic retarder control valves, 11-left side cover, 12-right side cover, 13-fan drive assembly, 14-liquid viscous clutch assembly, 15-oil pump group, 16-couplet of pump motor, 17-oil supply system, 18-hydraulic control system and 19-manipulation of the electronic control system. The integrated transmission device needs to implement five functions: straight driving, swerve, braking, fan cooling and compressor drive, and the reliability requirement of system is 0.7. The relationship among system functions and subsystems is shown as in Fig. 3.



**Figure 3. Relationship among system functions and subsystems**

*Establishment of the objective function*

Compared with the data of engineering, the lower limit of reliability of each subsystem is all equal to 0.8, the upper limit of reliability of each subsystem is all equal to 0.9999, and the basic price of each subsystem is shown as in Table 1.

**Table 1. The basic price of each subsystem**

No.	1	2	3	4	5	6	7	8	9
$P$	320	170	60	200	230	260	80	80	240
10	11	12	13	14	15	16	17	18	19
110	220	220	120	240	540	420	90	270	210

According to Eq. (2, 3), the objective function of integrated transmission device can be established as the sum of cost of each subsystem.

$$C_S(R) = \sum_{i=1}^{19} P_i e^{\left(\frac{R_i}{R_{i,\min}} - 1\right)}$$

*Establishment of constraint condition considering differentiation of functions*

Only the five functions of integrated transmission device all meet the requirements in terms of importance factor, the system can be denoted as success. The values of  $W_f$  about five functions are shown in Table 2.

**Table 2. The values of  $W_f$  about five functions**

No.	straight driving	swerve	braking	fan cooling	compressor drive
$W_f$	0.52	0.39	0.32	0.2	0.35

Naturally, the constraint conditions can be obtained as follows.

$$\begin{cases} R_{i,\min} \leq R_i \leq R_{i,\max} \\ \prod_{i=1}^m R_i \geq 0.7 & i = 1, 2, \dots, 19 \\ F_{f_j}(G_{f_j}, W_{f_j}) \geq 0 & j \in [1, 5] \end{cases}$$

*Problem formulation of reliability optimization allocation of integrated transmission device*

Reliability is allocated for each subsystem under the constraint conditions of reliability index of system with differentiation of functions, so as to make the cost of system to be minimum.

$$\min C_S(R) = \sum_{i=1}^{19} C_i(R_i) = \sum_{i=1}^{19} P_i e^{\left(\frac{R_i}{R_{i,\min}} - 1\right)}$$

$$\left\{ \begin{array}{l} \text{s.t.} \\ R_{i,\min} \leq R_i \leq R_{i,\max} \\ \prod_{i=1}^{19} R_i \geq 0.7 \quad i = 1, 2, \dots, 19 \\ F_{f_j}(G_{f_j}, W_{f_j}) \geq 0 \quad j \in [1, 5] \end{array} \right.$$

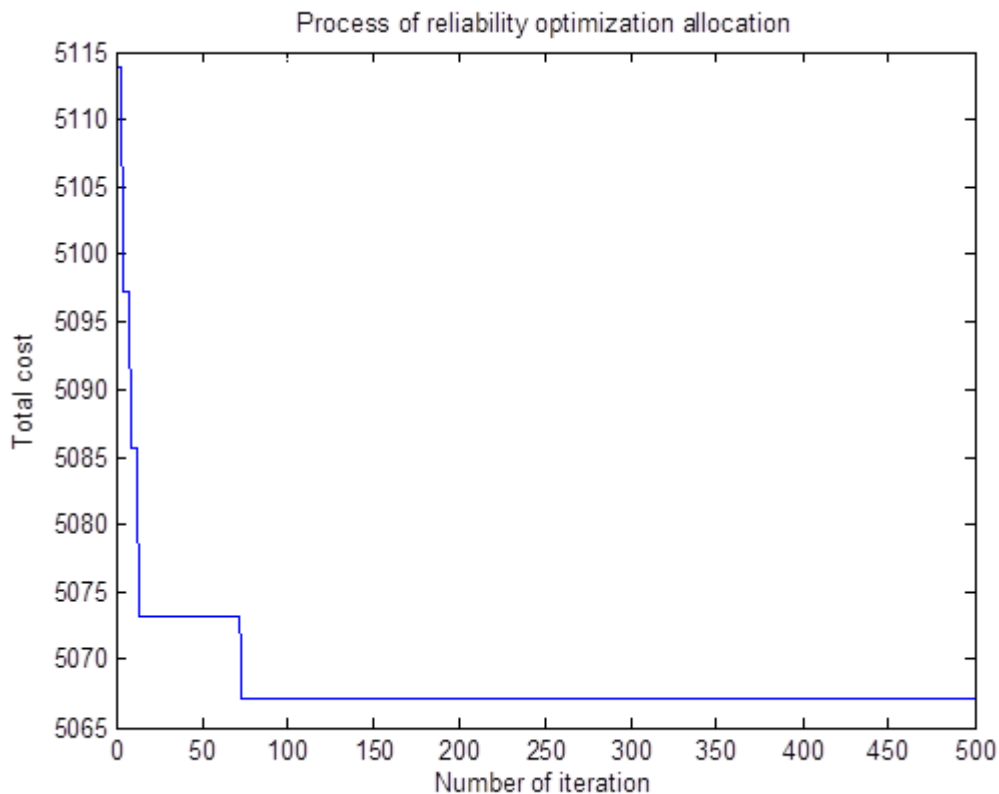
*Genetic algorithm and selection of results of optimization allocation*

The suitable parameters of genetic algorithm are determined and shown as in Table 3.

**Table 3. Parameter of genetic algorithm**

Size number of population	Probability of crossover	Probability of mutation	Maximum generation
120	0.9	0.01	500

The reliability optimization allocation of integrated transmission device is done according to the reliability optimization allocation process in section 2.5. The optimization allocation process with a single population is shown as in Fig. 4.



**Figure 4. Process of reliability optimization allocation with a single population**

At last, the additional 20 populations are selected to do the reliability optimization allocation of the integrated transmission device, and the results are shown in Table 4,  $\beta_i$  indicates the  $i$  th population. The results are shown in Table 4. The obtained reliability of each subsystem is ranked every time. Each line represents a set of optimal solutions, and the number of subsystem whose

reliability is the lowest is set at far left, and the number of subsystem whose reliability is the highest is set at far right. In other words, the degree of reliability value increase gradually from left to right.

**Table 4 The rank of reliability of subsystem**

$\beta$	Degree of reliability (decrease gradually)																		
	D <sub>1</sub>	D <sub>2</sub>	D <sub>3</sub>	D <sub>4</sub>	D <sub>5</sub>	D <sub>6</sub>	D <sub>7</sub>	D <sub>8</sub>	D <sub>9</sub>	D <sub>10</sub>	D <sub>11</sub>	D <sub>12</sub>	D <sub>13</sub>	D <sub>14</sub>	D <sub>15</sub>	D <sub>16</sub>	D <sub>17</sub>	D <sub>18</sub>	D <sub>19</sub>
1	14	15	6	16	1	12	5	18	19	11	9	4	2	13	17	10	8	3	7
2	16	15	1	9	19	18	5	13	6	10	4	12	7	17	3	14	8	2	11
3	1	16	15	18	14	12	2	6	9	5	19	4	7	8	17	13	3	10	11
4	15	16	1	11	18	14	12	9	4	19	6	8	10	5	7	2	17	13	3
5	16	15	14	1	9	6	18	12	5	4	19	2	11	17	10	13	7	8	3
6	15	1	16	18	6	11	12	9	14	10	19	4	5	13	2	8	3	17	7
7	15	18	16	6	9	11	1	5	19	2	14	12	4	13	10	8	7	17	3
8	15	1	11	6	18	16	12	5	19	4	9	14	10	3	7	13	17	2	8
9	16	9	15	1	6	14	18	12	11	4	5	2	19	17	8	10	7	13	3
10	6	15	16	1	5	14	9	11	18	19	3	2	4	13	8	10	7	17	12
11	16	15	1	6	4	9	18	14	12	11	13	2	5	10	19	17	7	3	8
12	16	15	12	11	1	6	9	14	4	18	5	7	2	10	13	17	19	8	3
13	15	9	16	1	5	11	14	6	12	4	19	18	17	10	13	2	3	7	8
14	1	15	18	5	16	9	2	19	6	14	12	11	7	13	10	17	8	3	4
15	16	12	15	18	1	11	19	6	14	9	5	4	2	7	8	10	17	13	3
16	15	1	14	16	6	9	18	11	5	4	12	10	13	19	7	2	8	3	17
17	15	6	14	16	1	18	5	9	19	11	17	12	13	4	10	7	3	8	2
18	15	1	16	14	6	9	18	5	11	12	19	4	13	8	2	10	3	7	17
19	16	6	15	18	1	14	12	9	11	5	19	4	13	17	2	10	3	7	8
20	5	15	16	1	12	19	18	6	11	4	9	7	2	3	14	8	17	10	13

In addition, the total costs with different populations are shown in Table 5.

**Table 5. Total costs with different populations**

NO.	1	2	3	4	5	6	7	8	9	10
Total cost	5073.9	5070.4	5075.5	5067.2	5078.4	5077.5	5077.2	5072.5	5076.3	5074.8
NO.	11	12	13	14	15	16	17	18	19	20
Total cost	5062.9	5076.7	5078.4	5078.6	5066.4	5077.9	5078.3	5069.1	5077.7	5076.0

The improved genetic algorithm is applied in the reliability optimization allocation of the integrated transmission device. We can see from Fig. 4 that the total cost of system become stable after iterating 75 times and the optimal solution is also obtained. What’s more, Table 5 shows that the total costs of different population are very close. This result illustrates the correctness and accuracy of the algorithm proposed in this paper. Observation can be seen from Table 4, the reliabilities assigned to the 15th and 16th subsystems are smaller, and reliabilities assigned to the 3rd, 7th, 17th subsystems are bigger. This shows that the method proposed in this paper can well deal with the issue of the reliability optimization allocation of complex system with differentiation of functions. Based on the correct reliability allocation method, a better and more correct optimal solution can be obtained through several times of reliability optimization allocation, and the randomness is also can be reduced. At last, the rule for selecting the final solution for the reliability optimization allocation can be formulated with reference of the allocation results of different population.

In this paper, the allocation results of the population that its cost of system is less and smaller reliabilities are assigned to the 15th, 16th subsystems and the larger reliabilities are assigned to the 3rd, 7th, 17th subsystems, will be selected to be the final solution. From Table 4 and Table 5, we can see that the allocation value of 19th population can be determined as the final solution for the integrated transmission device. The corresponding cost of system is 5062.9 and the reliability of each subsystem is shown in Table 6.

**Table 6. Final reliability of each subsystem**

No.	1	2	3	4	5	6	7	8	9	10
Reliability	0.9841	0.9972	0.9995	0.9621	0.9911	0.9658	0.9740	0.9610	0.9832	0.9776
No.	11	12	13	14	15	16	17	18	19	
Reliability	0.9890	0.9630	0.9692	0.9885	0.9745	0.9745	0.9889	0.9807	0.9795	

## Conclusion

In this paper a reliability optimization allocation method for systems with differentiation of functions is proposed. Firstly, a cost model is presented, in which the meaning and value of its parameters are easy to ascertain. What's more, the index of importance factor is proposed to deal with the differentiation of functions. Finally, the system model of reliability optimization allocation is built and the improved genetic algorithm is used to solve the problem of system reliability optimization allocation under the target of minimum total cost.

In order to verify the validity and accuracy of the proposed method in this paper, the reliability optimization allocation of an integrated transmission device is done. The results show that the proposed method can solve the problem of differentiation of functions, and the allocated results also indicate the differentiation of reliability requirements. This method has better guidance than traditional reliability allocation method and it is easy to be applied in engineering.

## References

- Chen, M. S. (1992) On the computational complexity of reliability redundancy allocation in a series system," *Operations research letters* **11**, 309–315.
- Bueno, V. C. (2005) Minimal standby redundancy allocation in a k-out-of-n : F system of dependent components, *European Journal of Operational Research* **165**, 786–793.
- Levitin, G. and Lisnianski, A. (1999) Joint redundancy and maintenance optimization for multistate series–parallel systems, *Reliability Engineering and System Safety* **64**, 33–42.
- Ramirez-marquez, J. E. and Coit, D. W. (2004) A heuristic for solving the redundancy allocation problem for multi-state series–parallel system, *Reliability Engineering and System Safety* **83**, 341–349.
- Levitin, G. and Lisnianski, A. (2003) Optimizing survivability of vulnerable series–parallel multi-state systems, *Reliability Engineering and System Safety* **79**, 319–331.
- Levitin, G. (2003) Optimal multilevel protection in series–parallel systems, *Reliability Engineering and System Safety* **81**, 93–102.
- Levitin, G., Dai, Y., Xie, M. and Poh, K. L. (2003) Optimizing survivability of multi-state systems with multi-level protection by multi-processor genetic algorithm, *Reliability Engineering and System Safety* **82**, 93–104.
- Levitin, G. and Lisnianski, A. (2001) Structure optimization of multi-state system with two failure modes, *Reliability Engineering and System Safety* **72**, 75–89.
- Altıparmak, F., Dengiz, B. and Smith, A. E. (2003) Optimal design of reliable computer networks: a comparison of metaheuristics, *Journal of Heuristics* **9**, 471–487.
- Nabil, N. and Mustapha, N. (2005) Ant system for reliability optimization of a series system with multiple-choice and budget constraints, *Reliability Engineering and System Safety* **72**, 1–12.
- Chen, T. C. and You, P. S. (2005) Immune algorithms-based approach for redundant reliability problems with multiple component choices, *Computers in Industry* **56**, 195–205.
- Kuo, W. and Rin, W. (2007) Recent advances in optimal reliability allocation, *Ieee Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans* **37**, 143–156.

# Acoustic resonance scattering of Bessel beam by elastic spheroids in water

\*Z.X. Gong<sup>1,2</sup>, †W. Li<sup>1,2,3</sup>, and Y.B. Chai<sup>1,2</sup>

<sup>1</sup>School of Naval Architecture and Ocean Engineering, Huazhong University of Science and Technology, China

<sup>2</sup>Hubei Key Laboratory of Naval Architecture & Ocean Engineering Hydrodynamics (HUST), Huazhong University of Science and Technology, Wuhan City, Hubei 430074, China

<sup>3</sup>Collaborative Innovation Center for Advanced Ship and Deep-Sea Exploration (CISSE), Shanghai City, 200240, China

\*Presenting author: hustgzx@hust.edu.cn

†Corresponding author: hustliw@hust.edu.cn

## Abstract

The acoustic resonance scattering of a zero-order Bessel beam by elastic spheroids immersed in water and centered along the beam axis is investigated. The T-matrix method is utilized to establish the Bessel acoustic scattering formulas through the harmonic expansion of Bessel beam. For a specific half-conical angle  $\beta$ , the far-field backscattering form functions of spheroids with different materials and a variety of aspect ratios are curved versus the dimensionless frequency  $kL/2$ , where  $k$  is the wave number in water and  $L$  is the length of the spheroid. By subtracting an appropriate background from the total backscattering form functions, the corresponding resonance of the elastic spheroids can be determined. It is concluded that the T-matrix method is effective to calculate both the total and resonance backscattering fields under an end-on incident Bessel beam illumination. Moreover, by selecting appropriate half-conical angles, the excitation of certain resonance of elastic spheroids may be suppressed and this phenomenon may have some potential value in practical applications.

**Keywords:** Acoustic resonance scattering, Zero-order Bessel beam, T-matrix method, Backscattering form function, Elastic spheroid

## 1. Introduction

During the past decades, considerable efforts to analyze acoustic resonance scattering by elastic targets immersed in an ideal fluid illuminated by plane wave has been devoted through experimental and theoretical methods by many authors [Flax et al. (1978); Gaunaurd and Uberall (1983); Werby et al. (1988); Williams and Marston (1986); Bao et al. (1992); Haumesser et al. (2002)]. When the incident acoustic excitation takes Bessel beam into consideration, the corresponding investigation is very limited. Bessel beam was first introduced by Durnin et al [Durnin (1987); Durnin et al. (1987)] in optics and then had a following development in acoustics [Campbell and Soloway (1990); Lu and Greenleaf (1990); Marston (1992)]. Bessel beam is characterized by an important parameter, termed as half-conical angle  $\beta$ , which describes the angle of the planar wave components of the beam relative to the beam axis. Note that in practice, Bessel beam can only propagate over a limited distance without spreading due to the finite width of sources, however, in the

published studies, the ideal Bessel beam is always taken as the incident acoustic excitation. Also, in our work, targets are under the illumination of ideal Bessel beam. Recently, Bessel beam has attracted increasing attention in acoustic aspect because it has demonstrated several advantages over the plane waves, such as the characteristics of non-diffraction and the ability to retain its form without block. Marston first studied the scattering characteristics of both rigid and soft spheres centered on a Bessel beam and curved the form function modulus as a function of scattering angle  $\theta$  [Marston (2007a)]. After that in 2007, Marston continued to investigate the resonance scattering of elastic solid sphere and spherical shell under the Bessel beam illumination and evaluated the influences of several selective half-conical angles on the suppression of backscattering resonance [Marston (2007b)]. Subsequently, further study on acoustic scattering associated with Bessel beam was also reported by others, such as Mitri and Li's research group. There has been considerable interest in Mitri's study of high-order Bessel beams to explore acoustic scattering characteristics by several objects, including rigid (movable and immovable) spheres [Mitri (2009a; 2011); Mitri and Silva (2011)], elastic spheres [Mitri (2008a; 2009b)], and elastic spherical shells [Mitri (2010; 2012)]. In addition, Mitri devoted much efforts to calculate acoustic radiation force of spheres and rigid spheroids [Mitri (2008b; 2009c; 2015); Silva et al. (2013)], which may provide an impetus to design acoustic tweezers. Moreover, our research group has studied the scattering properties of arbitrary-shape rigid scatterer facing the incident Bessel beam. In our work, the backscattering fields of rigid spheroid and finite cylinder with two hemispherical endcaps were investigated, and the peak to peak intervals in backscattering form functions were analyzed both in geometry and using numerical method [Li et al. (2015)].

Despite the recent reports about acoustic Bessel beam reviewed above, to date, it still remains an unsolved problem when calculation of resonance scattering by elastic spheroid placed in Bessel beam is taken into consideration. In previous studies published by Marston and Mitri, the exact scattering by spheres and spherical shells was expressed as a partial wave series. Specifically, Mitri improved the partial-wave series expansion (PWSE) method to calculate acoustic radiation force of rigid spheroid under the illumination of Bessel beam [Mitri (2015)]. Unfortunately, the PWSE method has not yet been further developed to provide theoretical analysis on elastic spheroid interacted with Bessel beams. Instead, the T-matrix method, as originally conceived by Waterman [Waterman (1965; 1971)], has been demonstrated a very efficient tool to handle acoustic scattering problems by elastic targets with arbitrary shape, for instance, cylinders [Varadan 1978], spheres [Pao and Mow (1963)] and spheroids [Flax et al. (1983); Bostrom (1980a)]. The philosophy of the T-matrix method is to expand all field quantities in terms of a set of spherical functions in order to obtain the T matrix (also called transition matrix) that relates the known coefficients of expansion of the incident wave to the unknown expansion coefficients of the scattered field. When all parameters of scatterer and incident wave are provided, the scattered fields can be obtained immediately by using the T-matrix method. To our knowledge, there is no evidence that the Bessel beam scattering by elastic spheroid immersed in an ideal fluid has been studied in previous work published. To this end, in the present study, we aim to extend the application of T-matrix to study the acoustic scattering characteristics of elastic spheroid under the illumination of Bessel beam.

This paper is outlined as follows. In section 2, a brief review of the T-matrix method for acoustic scattering is given and subsequently, we derive the formula of incident coefficients of Bessel beam. In section 3, two numerical examples are carried out to



explore the acoustic resonance characteristics of PMMA sphere and spheroid immersed in an ideal fluid under the illumination of ideal zero-order Bessel beam. Finally, the conclusion of this paper is conducted in section 4.

## 2. Theoretical formulation

In this section, the theoretical formulas of an elastic scatterer immersed in an ideal fluid under the normal illumination of Bessel beam by using the well-known T-matrix method are present. As in the previous work of other authors, the ideal Bessel beam is also considered here.

### 2.1 A brief review of the T-matrix method for the scattering

To determine the transition matrix, which connects the expansion coefficients of the incident and scattered fields, the integral representations for the displacement field, the boundary conditions and expansions of surface fields should be provided explicitly. Now consider an elastic scatterer with its geometry shown in Fig. 1. The boundary surface between fluid and solid is denoted by  $S$ . The host medium is homogeneous water with density  $\rho_f$  and lame parameter  $\lambda_f$ . The properties of elastic scatterer are given by density  $\rho$  and lame parameters  $\lambda$  and  $\mu$ . For convenience, we would rather use the vector formalism than the scalar wave equation for both the fluid (outside  $S$ ) and solid (inside  $S$ ) regions to facilitate the application of the boundary conditions. As given in detail by Bostrom and Peterson [Bostrom (1980b); Peterson et al. (1980)], here, we will only present the most pertinent formulas.

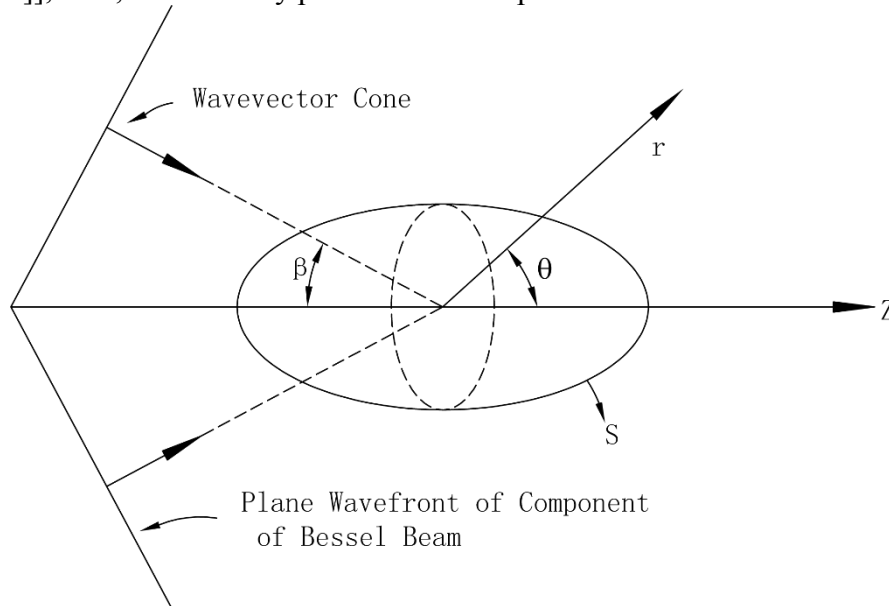


Fig. 1 Elastic scatterer under Bessel beam illumination.

The starting point for the following is the integral representation for the displacement  $\mathbf{u} = \mathbf{u}^i + \mathbf{u}^s$  in the fluid region

$$\int_S \left\{ \mathbf{u}_+ \cdot \left[ n \cdot \boldsymbol{\Sigma}_f(\mathbf{r}, \mathbf{r}') \right] - \mathbf{G}_f(\mathbf{r}, \mathbf{r}') \cdot \mathbf{t}_+ \right\} dS = \begin{cases} \mathbf{u}^s & \mathbf{r} \text{ outside } S \\ -\mathbf{u}^i & \mathbf{r} \text{ inside } S \end{cases} \quad (1)$$

Here  $n$  is the unit normal taken as outward pointing,  $\boldsymbol{\Sigma}_f$  and  $\mathbf{G}_f$  are the free space Green's stress triadic and Green's dyadic, respectively. The surface traction vector  $\mathbf{t}_+$  can be obtain by Hooke's law expressed as

$$\mathbf{t} = n \cdot [\lambda \mathbf{I} \nabla \cdot \mathbf{u} + \mu (\nabla \mathbf{u} + \mathbf{u} \nabla)] \quad (2)$$

For the solid region inside  $S$ , the displacement  $\mathbf{u}$  is governed by the following representation

$$\int_S \left\{ \mathbf{u}_- \cdot [n \cdot \boldsymbol{\Sigma}(\mathbf{r}, \mathbf{r}') - \mathbf{G}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{t}_-(\mathbf{r}')] \right\} dS = \begin{cases} -\mathbf{u} & \mathbf{r} \text{ inside } S \\ 0 & \mathbf{r} \text{ outside } S \end{cases} \quad (3)$$

where  $\mathbf{G}$  is the Green's displacement dyadic and  $\boldsymbol{\Sigma}$  is the Green's stress dyadic related to  $\mathbf{G}$  by Hooke's law [Varadan (1980)]. For an elastic scatterer immersed in water, the relevant boundary conditions on surface are given by

$$\begin{aligned} n \cdot \mathbf{u}_+ &= n \cdot \mathbf{u}_- \\ n \cdot \mathbf{t}_+ &= n \cdot \mathbf{t}_- \\ n \times \mathbf{t}_- &= 0 \end{aligned} \quad (4)$$

Both the integral representations for the displacement fields and the corresponding boundary conditions are presented above. Subsequently, the incident and scattered displacement fields are expanded in vector spherical functions, with the time factor  $e^{-i\omega t}$  being suppressed throughout, as follows:

$$\mathbf{u}^i = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{\sigma} a_{nm\sigma} \text{Re} \phi_{nm\sigma}(\mathbf{r}) \quad (5)$$

$$\mathbf{u}^s = \sum_{n=0}^{\infty} \sum_{m=0}^n \sum_{\sigma} f_{nm\sigma} \phi_{nm\sigma}(\mathbf{r}) \quad (6)$$

Finally, the incident and scattered field coefficients are related through the transition as given by

$$f_{nm\sigma} = \sum_{n'm'\sigma'} T_{nm\sigma, n'm'\sigma'} a_{n'm'\sigma'} \quad (7)$$

where

$$T = -\text{Re} [QR^{-1}P] [QR^{-1}P]^{-1} \quad (8)$$

The detailed expressions of matrices  $Q$ ,  $R$  and  $P$  are given by Bostrom and Peterson [Bostrom (1980b); Peterson et al. (1980)]. Through the formulas described above, the unknown scattered coefficient can be computed immediately, and thus the scattered field can be acquired correspondingly.

## 2.2 Derivation of the incident coefficients of Bessel beam

In this part, the formulas of the incident coefficient  $a_{nm\sigma}$  in the case of Bessel beam will be derived for the combination with original T-matrix method. Note that by using the T-matrix method, all field quantities should be expanded in terms of one series of spherical functions. Here the following scalar basis functions adopted in the computation are defined as

$$\phi_{nm\sigma}(\mathbf{r}) = \xi_{nm} h_n(kr) P_n^m(\cos \theta) \begin{cases} \cos(m\varphi), & \sigma = e \\ \sin(m\varphi), & \sigma = o \end{cases} \quad (9)$$

$$\xi_{nm} = \left( \varepsilon_m \frac{(2n+1)(n-m)!}{4\pi(n+m)!} \right)^{1/2} \quad (10)$$

where  $h_n(kr)$  is the spherical Hankle function of the first kind,  $P_n^m$  is the associated Legendre function and  $\varepsilon_m = 2 - \delta_{m0}$  is the Neumann factor.  $\sigma = e, o$  (even, odd) specifies azimuthal parity,  $m = 0, 1, \dots, n$  specifies rank and  $n = 0, 1, \dots$  specifies order of the spherical harmonics. Also,  $\theta$  and  $\varphi$  are the axial and azimuthal angles, respectively.

The wave function defined above may be used to expand the scattered fields that satisfy radiation conditions at infinity. But to expand the incident fields that are finite at origin, the spherical Hankle function in the wave function in Eq. (9) should be replaced by the spherical Bessel function, which is regular at origin. In the case of the ideal zero-order Bessel beam we are interested, the expression of incident coefficients can be derived by using the foregoing basis functions in the following procedure.

Provided that the complex velocity potential of ideal zero-order Bessel beam could be denoted as follows [Marston (2007a)]

$$\phi_b(z, \rho) = \phi_0 \exp(i\kappa z) J_0(\mu\rho) \quad (11)$$

where  $\phi_0$  stands for the beam amplitude,  $z$  and  $\rho$  specify the axial and radial coordinates,  $\kappa$  and  $\mu$ , satisfying the relation  $\kappa^2 + \mu^2 = k^2$ , represent the axial and radial wavenumbers, and  $J_0$  is a zero-order Bessel beam.

Given by Eq. (B2) in APPENDIX B by Marston [Marston (2007a)], Eq. (11) may be expanded in spherical partial wave as

$$\phi_b(r, \theta) = \phi_0 \sum_{n=0}^{\infty} i^n \times (2n+1) j_n(kr) P_n(\cos\theta) P_n(\cos\beta) \quad (12)$$

By using the addition theorem of the Legendre polynomial [Stratton (2007)], the desired integration can be obtained

$$P_n(\cos\theta) = \sum_{m=0}^n \varepsilon_m \frac{(n-m)!}{(n+m)!} \times P_n^m(\cos\theta_i) P_n^m(\cos\theta_s) \cos m(\varphi_i - \varphi_s) \quad (13)$$

where  $\theta_i$  and  $\theta_s$  represent the incident and scattered axial angles, respectively.

Combining Eq. (9), (10), (12) and (13) gives an expression of the incident coefficients  $a_{nm\sigma}$  under ideal zero-order Bessel beam as follows

$$a_{nm\sigma} = 4\pi \xi_{nm} i^n \times P_n^m(\cos\theta_i) P_n^m(\cos\beta) \begin{pmatrix} \cos(m\varphi), & \sigma = e \\ \sin(m\varphi), & \sigma = o \end{pmatrix} \quad (14)$$

It should be noted that the expression of the incident coefficients depend on incident axial angles, half-conical angle and incident azimuthal angle. The ideal Bessel beam taken as the incident excitation is with unit amplitude ( $\phi_0 = 1$ ) and meanwhile normally incident ( $\theta_i = 0$ ). The half-conical angle  $\beta$  is decided by plane wavefront of component of Bessel beam relative to the beam axis. Specially,  $\beta = 0$  gives the limiting case of an ordinary plane wave. During the derivation process presented above, the azimuthal angle  $\varphi$  is neglected, which is totally feasible for rotationally symmetrical targets considered here.

Substituting Eq. (8) and (14) into the relation given by Eq. (7), the scattered fields of elastic targets under the illumination of ideal zero-order Bessel beam can be obtained and this may help lay a foundation for further exploring the characteristics of acoustic scattering of Bessel beam.

### 3. Numerical results

In order to study the Bessel beam modification to the coupling to resonances of elastic scatterers, several numerical examples are presented by using T-matrix method. In the first part, sphere model is given to verify the validity of the T-matrix method to investigate the acoustic resonance scattering under the illumination of Bessel beam. While in the second part, we extend the range of applicability of T-matrix to further investigate Bessel beam scattering characteristics of spheroid.

#### 3.1 Acoustic Resonance Scattering from Spheres

In this part, acoustic resonance analyses for the case of Bessel beam scattering from polymethylmethacrylate (PMMA) sphere are carried out and the ambient ideal fluid here is considered to be water. The corresponding material parameters in our numerical example are listed explicitly in Table 1. To facilitate the discussion in the following, several half-conical angles are selected specially which are defined  $\beta_n$  as the lowest root of  $P_n(\cos \beta_n) = 0$ . The 6-digit approximations to  $\beta_n$  for  $n = 2, 3, 4$  and  $5$  are  $\beta_2 = 54.7346^\circ$ ,  $\beta_3 = 39.2315^\circ$ ,  $\beta_4 = 30.5556^\circ$  and  $\beta_5 = 25.0173^\circ$ , respectively.

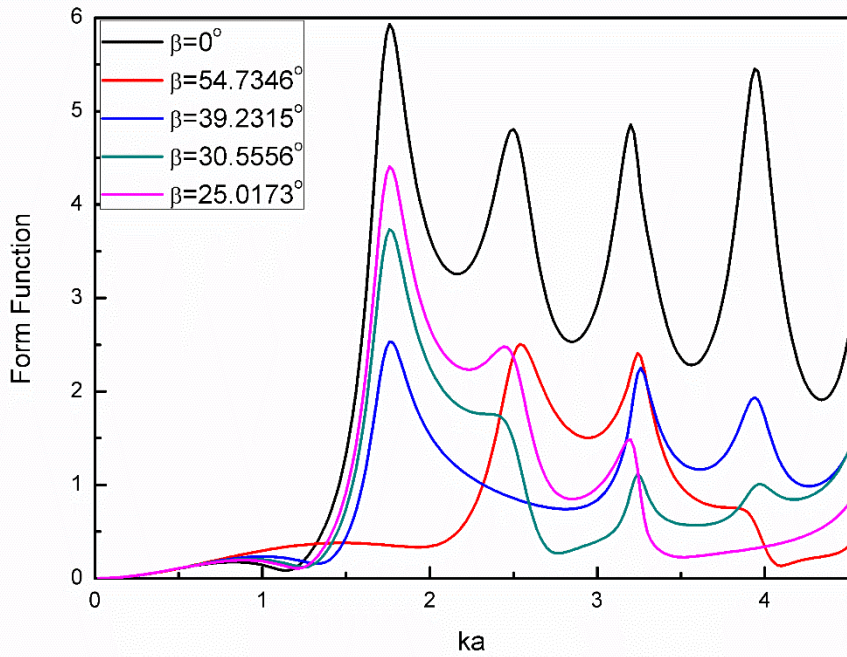
The backscattering form functions with different half-conical angles are calculated and plotted in Fig. 2. The black solid curve shown in Fig. 2 is the backscattering ( $\theta = \pi$  in Eq. (9)) for plane wave illumination ( $\beta = 0$ ). When other approximations ( $\beta_2, \beta_3, \beta_4$  and  $\beta_5$ ) are implemented, it could be found that the  $n$ th order resonance is suppressed. This is most easily seen from the dashed black line which has  $\beta = \beta_2$ . This behavior is in agreement with Eq. (14) because of the dependence on  $P_n(\cos \beta)$ . All of the curves shown in Fig. 2 agree very well with results in Marston's work.

**Table 1 Material parameters**

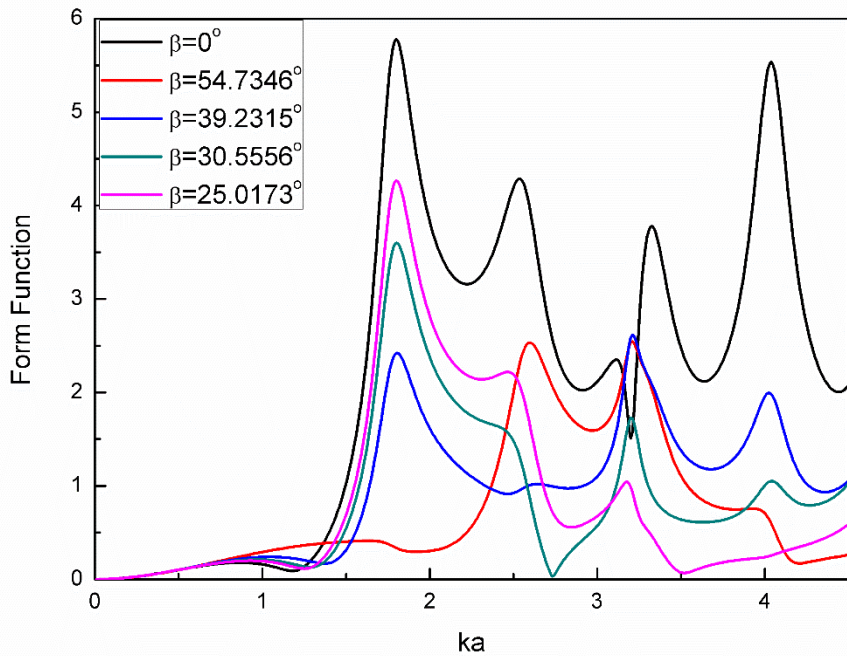
Material	Density (kg/m <sup>3</sup> )	Longitudinal velocity (m/s)	Shear velocity (m/s)
PMMA	1.19	2690	1340
Water	1000	1500	—

#### 3.2 Acoustic Resonance Scattering from Spheroids

In the prior part, backscattering form function of PMMA sphere with different half-conical angles are calculated. In the following part, we will further investigate the influences of different half-conical angles upon the backscattering resonance of PMMA spheroid. The aspect ratio (semi-minor axis/semi-major axis) of spheroid is 0.95. The backscattering form functions of PMMA spheroid under the illumination of Bessel beam with different half-conical angles are plotted in Fig. 3. The plane wave case is also given here by the black solid curve when  $\beta = 0^\circ$ . Similarly, by choosing approximations ( $\beta_2, \beta_3, \beta_4$  and  $\beta_5$ ), the corresponding resonances are suppressed. That is, in the case of  $\beta = \beta_2$ , the 2nd order resonance is suppressed, in the case of  $\beta = \beta_3$ , the 3rd order resonance is suppressed, and for the remaining  $\beta_4, \beta_5$ , the same results can be obtained.



**Fig. 2** Backscattering form function modulus computed as a function of dimensionless frequency  $ka$  for an elastic PMMA sphere in water under the illumination of Bessel beam with different half-conical angles



**Fig. 3** Backscattering form function modulus computed as a function of dimensionless frequency  $ka$  for an elastic PMMA spheroid in water under the illumination of Bessel beam with different half-conical angles

#### 4. Conclusions

In this paper, two numerical examples are carried out to calculate the backscattering form functions of underwater PMMA sphere and spheroid illuminated by zero-order Bessel beam with several selective half-conical angles. The T-matrix method is implemented and thus demonstrated to be an effective tool to compute the scattered fields. Also, the T-matrix method is able to expand its range of applicability through the combination with Bessel beam. With appropriate selection of specific Bessel beam parameters, some resonances can be suppressed and this in turn may provide useful directions on engineering applications.

#### References

- Bao, X. L., Überall, H. and Niemiec, J. (1992) Experimental study of sound scattering by elastic spheroids, and the excitation of their resonances, *The Journal of the Acoustical Society of America* **92**, 2313-2314.
- Bostrom, A. (1980a) Multiple scattering of elastic waves by bounded obstacles. *The Journal of the Acoustical Society of America*, **67**(2), 399-413.
- Bostrom, A. (1980b) Scattering of stationary acoustic waves by an elastic obstacle immersed in a fluid, *The Journal of the Acoustical Society of America* **67**, 390-398.
- Campbell, J. A. and Soloway, S. (1990) Generation of a nondiffracting beam with frequency-independent beamwidth, *The Journal of the Acoustical Society of America* **88**, 2467-2477.
- Durnin, J. (1987) Exact solutions for nondiffracting beams. I. The scalar theory, *JOSA A* **4**, 651-654.
- Durnin, J., Miceli Jr, J. J. and Eberly, J. H. (1987) Diffraction-free beams, *Physical Review Letters* **58**, 1499.
- Flax, L., Dragonette, L. R. and Überall, H. (1978) Theory of elastic resonance excitation by sound scattering, *The Journal of the Acoustical Society of America* **63**, 723-731.
- Flax, L., Dragonette, L. R., Varadan, V. K. and Varadan, V. V. (1982) Analysis and computation of the acoustic scattering by an elastic prolate spheroid obtained from the T - matrix formulation, *The Journal of the Acoustical Society of America* **71**, 1077-1082.
- Gaunaurd, G. C. and Überall, H. (1983) RST analysis of monostatic and bistatic acoustic echoes from an elastic sphere, *The Journal of the Acoustical Society of America* **73**, 1-12.
- Haumesser, L., Décultot, D., Léon, F. and Maze, G. (2002) Experimental identification of finite cylindrical shell vibration modes, *The Journal of the Acoustical Society of America* **111**, 2034-2039.
- Li, W., Li J. and Gong Z.X. (2015) Study on underwater acoustic scattering of a Bessel beam by rigid objects with arbitrary shapes, *Acta Physica Sinica* **64** (Accepted).
- Lu, J. Y. and Greenleaf, J. F. (1990) Ultrasonic nondiffracting transducer for medical imaging, *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on* **37**, 438-447.
- Marston, P. L. (1992) Geometrical and catastrophe optics methods in scattering, *Physical Acoustics* **21**, 1-234.
- Marston, P. L. (2007a) Scattering of a Bessel beam by a sphere, *The Journal of the Acoustical Society of America* **121**, 753-758.
- Marston, P. L. (2007b) Acoustic beam scattering and excitation of sphere resonance: Bessel beam example, *The Journal of the Acoustical Society of America* **122**, 247-252.
- Mitri, F. G. (2008a) Acoustic scattering of a high-order Bessel beam by an elastic sphere, *Annals of Physics* **323**, 2840-2850.
- Mitri, F. G. (2008b) Acoustic radiation force on a sphere in standing and quasi-standing zero-order Bessel beam tweezers, *Annals of physics* **323**, 1604-1620.
- Mitri, F. G. (2009a) Langevin acoustic radiation force of a high-order Bessel beam on a rigid sphere, *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on* **56**, 1059-1064.
- Mitri, F. G. (2009b) Equivalence of expressions for the acoustic scattering of a progressive high-order Bessel beam by an elastic sphere, *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on* **56**, 1100-1103.
- Mitri, F. G. (2009c) Negative axial radiation force on a fluid and elastic spheres illuminated by a high-order Bessel beam of progressive waves, *Journal of Physics A: Mathematical and Theoretical* **42**, 245202.
- Mitri, F. G. (2010) Interaction of a high-order Bessel beam with a submerged spherical ultrasound contrast agent shell—Scattering theory, *Ultrasonics* **50**, 387-396.
- Mitri, F. G. (2011) Acoustic beam interaction with a rigid sphere: The case of a first-order non-diffracting Bessel trigonometric beam, *Journal of Sound and Vibration* **330**, 6053-6060.
- Mitri, F. G. and Silva, G. T. (2011) Off-axial acoustic scattering of a high-order Bessel vortex beam by a rigid sphere, *Wave Motion* **48**, 392-400.
- Mitri, F. G. (2012) Generalized theory of resonance excitation by sound scattering from an elastic spherical shell in a nonviscous fluid *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on* **59**, 1781-1790.

- Mitri, F. G. (2015) Acoustic radiation force on oblate and prolate spheroids in Bessel beams, *Wave Motion* (available online)
- Pao, Y. H. and Mow, C. C. (1963) Scattering of plane compressional waves by a spherical obstacle, *Journal of Applied Physics* **34**, 493-499.
- Peterson, B., Varadan, V. V. and Varadan, V. K. (1980) Scattering of acoustic waves by layered elastic and viscoelastic obstacles in water, *The Journal of the Acoustical Society of America* **68**, 673-685.
- Silva, G. T., Lopes, J. and Mitri, F. G. (2013) Off-axial acoustic radiation force of repulsor and tractor bessel beams on a sphere, *Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on* **60**, 1207-1212.
- Stratton, J. A. (2007) *Electromagnetic theory*, John Wiley & Sons, New York, USA.
- Varadan, V. V. (1978) Scattering matrix for elastic waves. II. Application to elliptic cylinders, *The Journal of the Acoustical Society of America* **63**, 1014-1024.
- Varadan, V. V. (1980) Acoustic, electromagnetic, and elastic wave scattering--focus on the T-matrix approach: *international symposium held at the Ohio State University, Columbus, Ohio, USA*, New York, USA
- Waterman, P. C. (1965) Matrix formulation of electromagnetic scattering, *Proceedings of the IEEE* **53**, 805-812.
- Waterman, P. C. (1971) Symmetry, unitarity, and geometry in electromagnetic scattering, *Physical review D* **3**, 825.
- Werby, M. F., Überall, H., Nagl, A., Brown, S. H. and Dickey, J. W. (1988) Bistatic scattering and identification of the resonances of elastic spheroids, *The Journal of the Acoustical Society of America* **84**, 1425-1436.
- Williams, K. L. and Marston, P. L. (1986) Synthesis of backscattering from an elastic sphere using the Sommerfeld-Watson transformation and giving a Fabry-Perot analysis of resonances, *The Journal of the Acoustical Society of America* **79**, 1702-1708.

# A highly sensitive mass sensor using high-mode resonant cantilever with step change in thickness

Renjing Gao<sup>1</sup>, Mingli Li<sup>1,†</sup>, Jian Zhao<sup>1</sup>

<sup>1</sup>School of Automotive Engineering, State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, Liaoning, China

†Corresponding author: jzhao@dlut.edu.cn

## Abstract

Resonant cantilever mass sensors (RCS) can quantitatively detect the unknown analyte by measuring the frequency shift induced by attached the mass of analyte. With the characteristics of high sensitivity, long term stability and easy interfacing with digital signal processing, such sensors have been successfully applied in various application fields such as proteomics, genomics, gas sensing, food contamination, cancer detecting, chemical or fluidic detection. Recently, a great deal of attention has been paid on developing high sensitivity sensors that can be used outside of laboratory environments. According to the operating principle of the resonant sensor, the mass sensitivity is mainly depending on the ratio of the resonant frequency at mode number  $n$  to the corresponding effective mass of the resonant cantilever. Simultaneously increasing the resonance frequency and reducing the effective mass can effectively improve the mass detection sensitivity. As reported in many references, reducing geometry dimensions to increase the natural frequency by MEMS or NEMS technology is an effective way to improve sensitivity and spatial resolution of resonance mass sensors. Due to the limitation of the equipment for measuring the nano-sized cantilever deflection, the geometrical size reduction method encounters a bottleneck in improving the sensitivity. Different from the method of reducing the cantilever dimension, a new sensitivity improving method was proposed to inspire high-order mode by optimizing the stiffness and mass distribution of the cantilever. Through theoretically analyzing the relationship among structural parameters, vibration modes and the sensitivity, a novel piezoelectric resonant mass sensor working in the fourth order vibration mode is designed and fabricated. The experimental sensitivity of the proposed sensor is  $188.41 \times 10^4$  Hz/g nearly 19.5 times greater than that of the custom rectangular cantilever sensor of  $9.85 \times 10^4$  Hz/g. The simulated sensitivity is nearly consistent with that of the experiment with the deviation of 1.28%. Meanwhile, the quality factor is 82.65, which is about 3.5 times as great as that of the rectangular uniform cantilever sensor with the same size, hence, validates the feasibility and effectiveness of the newly proposed sensitivity improving method.

**Key words:** mass sensor, cantilever, high mode resonance, sensitivity, quality factor

## 1 Introduction

Resonant cantilever mass sensors (RCS) can quantitatively detect the unknown analyte by measuring the induced frequency shift by itself. With the characteristics of high sensitivity, long term stability and easy interfacing with digital signal processing, such sensors have been successfully applied in various application fields such as proteomics [Liu et al.(2007); Ekinici and Roukes(2005); Frederic et al.(2005)], genomics, gas sensing,



food contamination, cancer detecting, chemical or fluidic detection [Murali et al.(2007); Margarita et al.(2009);Degertekin et al.(2001)].

Recently, a great deal of attention has been paid on developing high sensitivity sensors that can be used outside of laboratory environments. According to the operating principle of the resonant sensor, the mass sensitivity is mainly depending on the ratio of the resonant frequency at mode number  $n$  to the corresponding effective mass of the resonant cantilever, which can be indicated by

$$S_n = \frac{\Delta f_n}{\Delta m^e} \cong \frac{f_n}{2M_n^e} \quad (1)$$

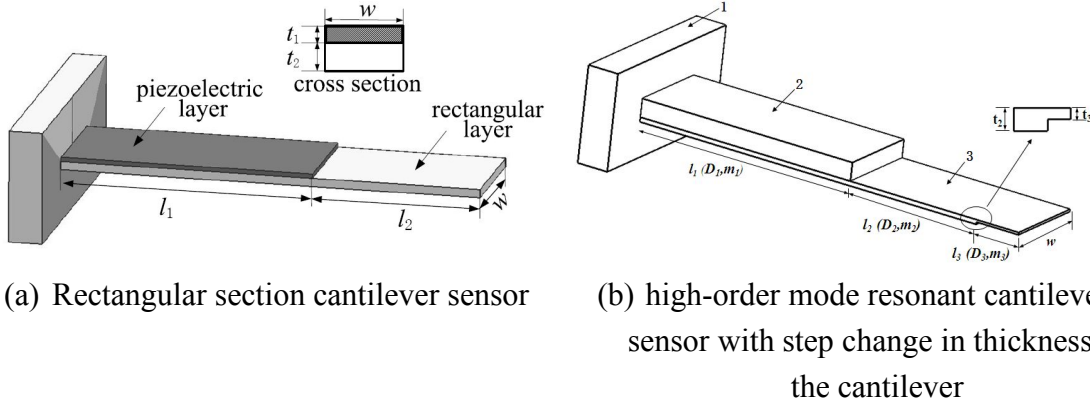
In Eq. (1),  $\Delta f_n$  is the frequency shift for per unit mass change of  $\Delta m^e$ ,  $f_n$  is the  $n$ th flexural-mode resonance frequency for a cantilever of effective mass  $M_n^e$ . From equation (1), it can be seen that simultaneously increasing the resonance frequency and reducing the effective mass can effectively improve the mass detection sensitivity. As reported in many references, reducing geometry dimensions to increase the natural frequency by MEMS or NEMS technology is an effective way to improve sensitivity and spatial resolution of resonance mass sensors[Shen et al.(2006; 2007); Kumar et al.(2010)]. Due to the limitation of the equipment for measuring the nano-sized cantilever deflection, the geometrical size reduction method encounters a bottleneck in improving the sensitivity. In other words, for nano-sized cantilevers, the high sensitivity is achieved at the expense of complex supplementary equipments for accurate detecting and low anti-jamming in vacuum. Alternatively, another method of using high-order resonant modes has been introduced to improve sensitivity and quality factor[Lochon et al.(2005); Maraldo and Mutharasan (2010)]. Unfortunately, as the most often used method to achieve high order mode vibration, the geometric size reduction method will directly reduce the cantilever surface area and make it difficult to manipulate the particles in the detecting system, Actually, the cantilever profile can also affects the vibration modes of the uniform cantilevers by means of changing the cantilever profile as well as by changing the effective stiffness and the mass distribution. Intuitively speaking, the cantilever profiles can be optimized to lightweight in the manner of reducing the stiffness by a smaller factor, thus resulting in higher resonance frequencies than that of the uniform cantilevers.

In this paper, different from the methods of reducing the cantilever dimension, a new sensitivity improving method was proposed to inspire high-order mode by optimizing the resonant cantilever with step change in thickness of the mass sensor. By theoretically analyzing the relationship among structural parameters, vibration modes and the sensitivity, a novel piezoelectric resonant mass sensor inspired with fourth order mode vibration is designed and fabricated. The experimental and simulated sensitivities of the proposed sensor are  $188.41 \times 10^4$  Hz/g and  $192.50 \times 10^4$  Hz/g respectively, which are nearly 19.50 times greater than that of the custom rectangular cantilever sensor of  $9.85 \times 10^4$  Hz/g. Additionally, the simulation results of the sensitivity is nearly consistent with that of the experiments with the deviation within 1.28%. Meanwhile, the quality factor is 82.65, which is about 3.50 times as great as that of the rectangular uniform cantilever sensor with the same size, hence, validates the feasibility and effectiveness of the newly proposed sensitivity improving method. Furthermore, the influence of the length and thickness of the step change section of the resonant cantilever on the sensitivity is analyzed.

## 2 Structure of the high-order mode resonant mass sensor

Commonly, rectangular uniform cantilever is the most widely used spring element in designing resonant mass sensor, as shown in Fig.1(a), and the corresponding elastic constant is determined by the structural geometry and material properties. Actually, the sensitivity is deeply related to the cantilever resonance frequency and the deflection at the free end. Therefore, it is an effective way to reduce the thickness at the free end properly to change both the stiffness and mass distribution, as shown in Fig.1(b), in which, the non-piezoelectric extension is set to be the detection area, and the characters of  $l_1$ ,  $l_2$  and  $l_3$  represent the lengths of the piezoelectric layer, and the cantilever and the step part,  $w$  is the width of the cantilever,  $t_1$ ,  $t_2$  and  $t_3$  are the thicknesses of the piezoelectric layer and the cantilever and the step part.

In Fig.1(b), the cantilever can be divided into three distinguished parts with different cross sections and non-uniform lengths of  $l_1$ ,  $l_2$  and  $l_3$ . The bending vibration equation for constant cross-section beam based on Euler–Bernoulli beam theory is established as follows.



**Fig.1 Comparison between two kinds of cantilever sensors**

$$D_i \frac{d^4 Y(x)}{dx^4} - \omega^2 m_i Y(x) = 0 \quad (2)$$

In Eq. (2),  $D_i$  and  $m_i$  represent the bending modulus and mass per unit length of each part, respectively.  $\omega$  is the angular frequency,  $Y(x)$  is the displacement amplitude. In Eq. (3),  $D_1$  and  $m_1$  are the bending modulus and mass per unit length in section 1,  $D_2$  and  $m_2$  in section 2, and  $D_3$  and  $m_3$  in section 3. Then, the natural frequencies of the cantilever are obtained as

$$f = \frac{k_{1n}^2}{2\pi} \sqrt{\frac{D_1}{m_1}} = \frac{k_{2n'}^2}{2\pi} \sqrt{\frac{D_2}{m_2}} = \frac{k_{3n''}^2}{2\pi} \sqrt{\frac{D_3}{m_3}} \quad (3)$$

where,  $k_{1n}$ ,  $k_{2n'}$  and  $k_{3n''}$  are the wave numbers of the  $n$ th,  $n'$ th and  $n''$ th order mode in section 1, section 2, and section 3.

When single piezoelectric layer is utilized, the effective stiffness and mass of the three lengths ( $D_1$ ,  $m_1$ ,  $D_2$ ,  $m_2$ ,  $D_3$ ,  $m_4$ ) can be expressed as Eq. (4) to Eq. (9).

$$D_1 = \frac{E_1^2 t_1^4 + E_2^2 t_2^4 + 2E_1 E_2 t_1 t_2 (2t_1^2 + 2t_2^2 + 3t_1 t_2)}{12(E_1 t_1 + E_2 t_2)} \quad (4)$$

$$m_1 = \rho_1 t_1 w + \rho_2 t_2 w \quad (5)$$

$$D_2 = \frac{E_2 t_2^3}{12} \quad (6)$$

$$m_2 = \rho_2 t_2 w \quad (7)$$

$$D_3 = \frac{E_3 t_3^3}{12} \quad (8)$$

$$m_3 = \rho_3 t_3 w \quad (9)$$

where,  $\rho_1$ ,  $\rho_2$  and  $\rho_3$  are the mass densities of each section,  $w$  is the width.

If the added mass is a point mass loaded at the cantilever tip, the flexural-mode resonance frequency thus can be approximated in terms of Rayleigh Quotient.

$$f_n' = \frac{1}{2\pi} \sqrt{\frac{K_n^e}{M_n^e + \Delta m \cdot Y_{2n}^2(x')}} \quad (10)$$

In Eq. (10),  $x'$  is the position of the added point mass,  $Y_{2n}$  is the  $n$ th mode shape,  $K_n^e$  and  $M_n^e$  are effective flexible stiffness and effective mass.

For  $\Delta m \ll M_n^e$ , the resonance frequency shift  $\Delta f_n$ , due to the added point mass at the cantilever tip, is thus Eq. (11)

$$\Delta f_n = f_n - f_n' = \frac{1}{2} f_n \frac{\Delta m \cdot Y_{2n}^2(x')}{\sqrt{M_n^e (M_n^e + \Delta m \cdot Y_{2n}^2(x'))}} \cong \frac{1}{2} f_n \frac{\Delta m \cdot Y_{2n}^2(x')}{M_n^e} \quad (11)$$

If the added mass  $\Delta m$  is distributed on the surface of the non piezoelectric extension part, the resonance frequency shift can be expressed as Eq. (12).

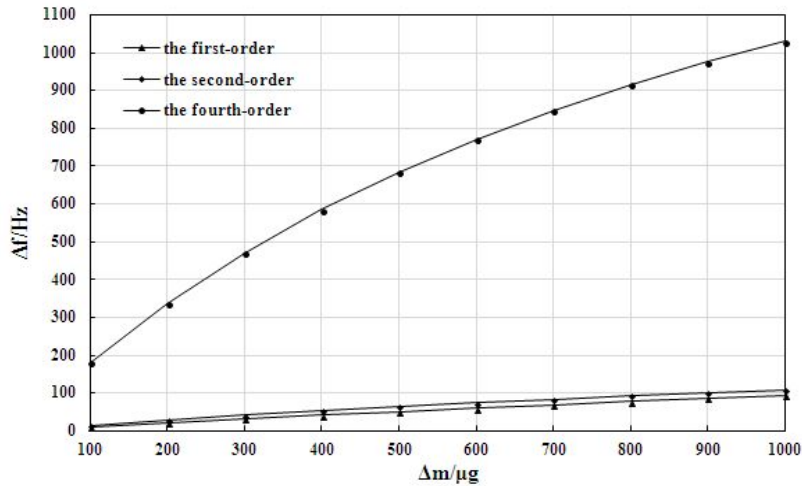
$$\Delta f_n = f_n - f_n' \cong \frac{1}{2} f_n \frac{\Delta m^e}{M_n^e} \quad (12)$$

The detailed expression of the mass sensitivity of the proposed cantilever sensor can be obtained, which can be expressed as Eq. (13).

$$S = \frac{\Delta f}{\Delta m} = \frac{1}{4\pi w} \sqrt{\frac{D_1}{m_1}} \cdot \frac{k_1^2 Y_{1n}^2(x)}{m_1 \int_0^{l_1} Y_{1n}^2(x) dx + m_2 \int_{l_1}^{l_1+l_2} Y_{2n}^2(x) dx} \quad (13)$$

### 3 Simulation and Experiments

Through numerical simulation, the detection sensitivity for different vibration modes can be obtained in Fig.2. It can be seen that the sensitivity of the fourth order vibration mode is much higher than those of the lower order vibration modes. Then, the influences of the step parameters including step length and thickness on the sensitivity were analyzed, as shown in Fig.3 and Fig.4. The output sensitivity is varying with the step length nonlinearly while the thickness and the total length keep constant. When the length  $l_3$  reaches  $6.00 \times 10^{-3}$  m, the sensitivity can achieve the highest value, as shown in Fig.3. Meanwhile, keeping the step length and the total cantilever thickness constant, the sensitivity also varies nonlinearly with the thickness ratio of the step part to the cantilever. In the thickness ratio range from  $3.33 \times 10^{-1}$  to 5.00, the sensitivity decreases with the increasing of the thickness ratio.



**Fig.2 Sensitivity curve of each order with step change in thickness of the cantilever sensor**

More importantly, to evaluate the sensitivity improvement, we fabricated a novel grooved cantilever type mass sensor by using electric discharge wire-cutting technology, as shown in Fig.5. The mass sensor mainly consists of a fixture and a cantilever which was made by bonding a PZT layer with  $4.00 \times 10^{-3}$  m width,  $2.50 \times 10^{-4}$  m thickness,  $1.00 \times 10^{-2}$  m length to a two-layer steel sheet with  $t_2 = 6.00 \times 10^{-5}$  m,  $l_2 = 1.62 \times 10^{-2}$  m,  $t_3 = 3.00 \times 10^{-5}$  m,  $l_3 = 1.80 \times 10^{-3}$  m using conductive epoxy. The experiment results in Fig.6 show that the sensitivity is about  $188.41 \times 10^4$  Hz/g with a quality factor of 82.65, which is much higher than that of the rectangular uniform cantilever sensor. Also, the simulation result is nearly consistent with that of the experiments with a maximum deviation of 1.28%. Finally, the proposed design model for the high order modes resonant sensors is adequately validated.

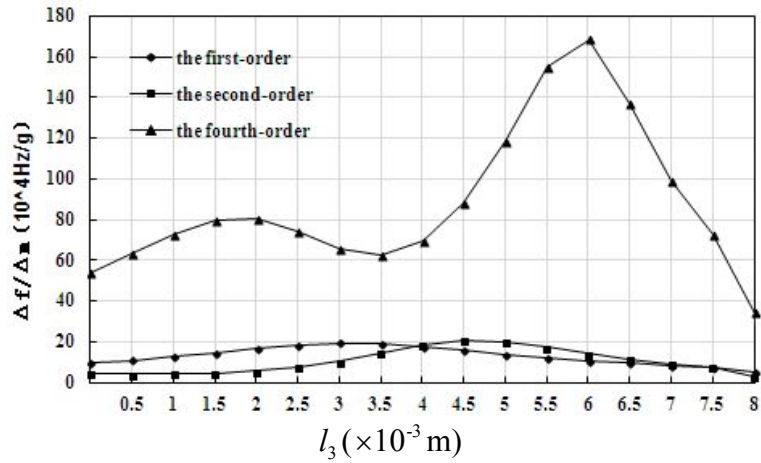


Fig.3 The influence of the length  $l_3$  on the sensitivity of the sensor

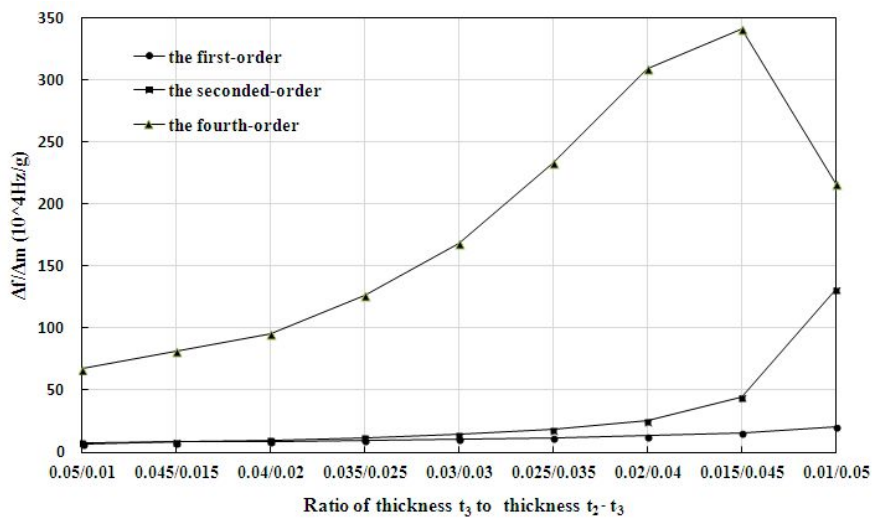


Fig.4 The influence of the ratio of thickness  $t_3 / (t_2 - t_3)$  on the sensitivity

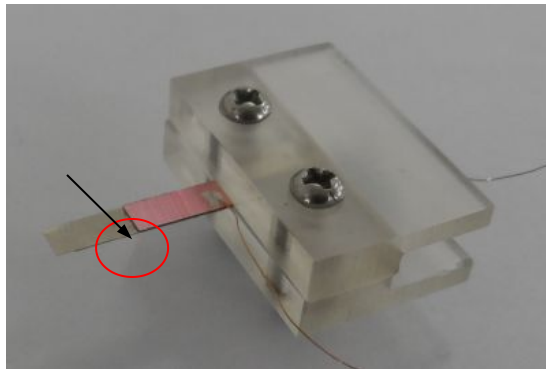


Fig.5 Non-uniform thickness sensor with large thickness ratio



**Fig.6 Test system for sensor sensitivity measurement**

#### **4 Conclusion**

As the key spring element in the resonant mass sensor, a new cantilever profile has been proposed to inspire high-order mode by optimizing the mass and stiffness distribution at the free end. The analytical model was established for theoretically analyzing the relationship among structural parameters, vibration modes and the sensitivity. Through the linear search optimization, a novel piezoelectric resonant mass sensor inspired with fourth order mode vibration is designed and fabricated. The experimental and simulated sensitivities of the proposed sensor are  $188.41 \times 10^4$  Hz/g and  $192.50 \times 10^4$  Hz/g respectively, which are nearly 19.50 times greater than that of the custom rectangular cantilever sensor of  $9.85 \times 10^4$  Hz/g. Additionally, the simulation results of the sensitivity is nearly consistent with that of the experiments with the deviation within 1.28%. Meanwhile, the quality factor is 82.65, which is about 3.50 times as great as that of the rectangular uniform cantilever sensor with the same size, hence, validates the feasibility and effectiveness of the newly proposed sensitivity improving method.

#### **Acknowledgement**

This research is supported by the National Natural Science Foundation of China(11372063, 11172052), the fundamental research funds for the central universities of china (2342013DUT13RC(3)28) and the 111 Project. The financial supports are greatly acknowledged.

#### **References**

- LIU M W, WANG J, WANG L D, et al. Deposition and characterization of Pb(Zr,Ti)O<sub>3</sub> sol-gel thin films for piezoelectric cantilever beams[J]. *Smart Mater Struct.* 2007,16:93-99.
- Ekinci K L, Roukes M L. Nanoelectromechanical systems[J]. *Rev.Sci.Instrum.* 2005,76: 061101-1-061101-12
- Frederic Lochon, Isabelle Dufour, Dominique Rebiere. An alternative solution to improve sensitivity of resonant microcantilever chemical sensors: comparison between using high-order modes and reducing dimensions[J]. *Sensors and Actuators.* 2005, 108:979 -985.
- Murali Krishna Ghatkesar, Viola Barwich, Thomas Braun, et al. Higher modes of vibration increase mass sensitivity in nanomechanical microcantilevers[J]. *Nanotechnology.* 2007, 18:44-52
- Margarita Narducci, Eduard Figueras, Maria Jose Lopez, et al. Sensitivity improvement of a microcantilever based mass sensor[J]. *Microelectronic Engineering.* 2009, 86:1187- 1189.
- Degertekin F L, Hadimioglu B, Sulchek T, et al. Actuation and characterization of atomic force microscope cantilevers in fluids by acoustic radiation pressure[J]. *Applied Physics Letters.* 2001,11:1628-1632.

- Shen Z, Shih W Y, Shih, W H. Mass detection sensitivity of piezoelectric cantilevers with a nonpiezoelectric extension[J]. *Review of Scientific Instruments*, 2006, 77(6):065101.
- [Shen Z, Shih W Y, Shih W H. Flexural vibrations and resonance of piezoelectric cantilevers with a nonpiezoelectric extension[J]. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2007, 54(10):2001-2010.
- Kumar V, Boley J W, Ekowaluyo H, et al.. Linear and nonlinear mass sensing using piezoelectrically-actuated microcantilevers[C]. *MEMS and Nanotechnology*, 2010, 2:57-65.
- Lochon, F., et al. An alternative solution to improve sensitivity of resonant microcantilever chemical sensors: comparison between using high-order modes and reducing dimensions [J]. *Sensors and Actuators B: Chemical*. 2005, 108(1-2): 979-985.
- Maraldo D, Mutharasan R. Mass-change sensitivity of high-order mode of piezoelectric-excited millimeter-sized cantilever (PEMC) sensors: Theory and experiments[J]. *Sensors and Actuators B*, 2010, 143:731-739.

# F-bar aided edge-based smoothed finite element method with tetrahedral elements for large deformation analysis of nearly incompressible materials

\*Y.Onishi<sup>1</sup>, K.Amaya<sup>1</sup>

<sup>1</sup>Department of Mechanical and Environmental Informatics, Tokyo Institute of Technology, 2-12-1-W8-36, O-okayama, Meguro-ku, Tokyo, 152-8552, Japan.

\*Presenting and corresponding author: yonishi@a.mei.titech.ac.jp

## Abstract

A new smoothed finite element method (S-FEM) with tetrahedral elements for finite strain analysis of nearly incompressible solids is proposed. The proposed method is basically a combination of the F-bar method and edge-based S-FEM with tetrahedral elements and is named “F-barES-FEM-T4”. The isovolumetric part of the deformation gradient ( $\mathbf{F}_{iso}$ ) is derived from the  $\mathbf{F}$  of ES-FEM, whereas the volumetric part ( $\mathbf{F}_{vol}$ ) is derived from the cyclic smoothing of  $J$  ( $=\det(\mathbf{F})$ ) between elements and nodes. F-barES-FEM suppresses the pressure oscillation in nearly incompressible materials successfully with no increase in DOF.

**Keywords:** F-bar method, Smoothed finite element method, Tetrahedral element, Finite strain, Nearly incompressible, Pressure oscillation.

## Introduction

In the practical numerical simulation for solids, the use of tetrahedral elements are indispensable due to the complexity of body shapes. The edge-based smoothed finite element method with first-order tetrahedral elements (ES-FEM-T4) [Liu et al. (2010)] is one of the most accurate FE formulations that resolve the shear locking issue of the standard tetrahedral (i.e., constant strain) FE formulation. However, ES-FEM-T4 raises volumetric locking in the analysis of nearly incompressible materials such as rubber. Therefore, development of accurate FE formulations with T4 elements is still in a research stage.

There are already several methods proposed for locking-free analysis with tetrahedral meshes, but they all have some sort of serious drawbacks as follows.

- Hybrid (or mixed) second-order tetrahedral elements: significant increase in DOF; inevitable Lagrange multipliers; convergence problems in contact analysis; accuracy loss in severely large strain analysis.
- F-bar-Patch method [Neto et al. (2005)]: difficulties in pre-definition of good-quality patches; shear locking.
- Selective reduced integration hexahedral elements as the subdivision of tetrahedral elements (4 hexahedrons in a tetrahedron): significant increase in DOF; severe element distortion of initial mesh; pressure oscillation.
- Selective edge/node-based S-FEM-T4 (ES/NS-FEM-T4) [Onishi and Amaya (2014)]: pressure oscillation; locking at corners.
- Bubble-enhanced ES-FEM (bES-FEM-T4) [Ong et al. (2015)]: significant increase in DOF; quick pop out of bubble nodes; pressure oscillation.

In addition, formulations with the selective integration have a drawback that they cannot treat materials with deviatoric/volumetric coupling terms. For these reasons, a new formulation without intermediate nodes, mixed variational principle, pre-defined patches, element subdivision, selective integration, nor bubble nodes is awaited.

In this paper, we propose a new ES-FEM-T4 formulation combined with F-bar method [Neto et al. (1996)] named “F-barES-FEM-T4”. The smoothed deformation gradient and stress in the proposed method are defined at each edge, and the stress integration is performed over the smoothing domains of ES-FEM. The isovolumetric part of the deformation gradient ( $\mathbf{F}_{iso}$ ) is derived from  $\mathbf{F}$  of ES-FEM, whereas the volumetric part ( $\mathbf{F}_{vol}$ ) is derived from the cyclic smoothing of  $J$  ( $=\det(\mathbf{F})$ ) between elements and nodes. Therefore, the proposed method can suppress the pressure oscillation without any unfavorable approaches listed above.



## Method

### Mutli-Smoothed Deformation Gradient

As is well known today, NS-FEM-T4 can avoid volumetric locking and gives solutions that have less pressure oscillation in nearly incompressible cases. It implies that the node-based smoothing of  $J$  ( $\equiv \det(\mathbf{F})$ ) can be regarded as an operation to suppress pressure oscillation to some extent. Here we define the cyclic smoothing of  $J$  as the following operation:

- 
- 1) Calculate  $J$  at each element,  $J^{\text{elem}}$ , in the same way as the standard FEM-T4.
  - 2) Calculate smoothed  $J$  at each node,  $\tilde{J}^{\text{node}}$ , in the same way as NS-FEM-T4:

$$\tilde{J}^{\text{node}} = \frac{\sum_{i \in \mathbb{I}} J_i^{\text{elem}} V_i^{\text{elem}}}{\sum_{i \in \mathbb{I}} V_i^{\text{elem}}}, \quad (1)$$

where  $\mathbb{I}$  is the set of elements attached to the each node and  $V$  denotes the volume.

- 3) Calculate smoothed  $J$  at each element,  $\tilde{J}^{\text{elem}}$ , as the average of  $\tilde{J}^{\text{node}}$ :

$$\tilde{J}^{\text{elem}} = \frac{\sum_{j \in \mathbb{J}} \tilde{J}_j^{\text{node}}}{4}, \quad (2)$$

where  $\mathbb{J}$  is the set of four nodes comprising the each element.

- 4) Repeat 2) and 3) as necessary to calculate multi-smoothed  $J$  at each element,  $\bar{J}^{\text{elem}}$ . ( $\tilde{J}^{\text{elem}}$  is considered as  $J^{\text{elem}}$  in Eq. (1), and the last  $\tilde{J}^{\text{elem}}$  in Eq. (2) is considered as  $\bar{J}^{\text{elem}}$ .)
- 5) Calculate multi-smoothed  $J$  at each edge,  $\bar{J}^{\text{edge}}$ , in the same way as ES-FEM-T4:

$$\bar{J}^{\text{edge}} = \frac{\sum_{k \in \mathbb{K}} \bar{J}_k^{\text{elem}} V_k^{\text{elem}}}{\sum_{k \in \mathbb{K}} V_k^{\text{elem}}}, \quad (3)$$

where  $\mathbb{K}$  is the set of elements attached to the each edge.

---

In F-barES-FEM-T4,  $\bar{J}^{\text{edge}}$  is used to define the volumetric part of the deformation gradient as  $\bar{\mathbf{F}}_{\text{vol}}^{\text{edge}} = (\bar{J}^{\text{edge}})^{1/3} \mathbf{I}$ .

Meanwhile, ES-FEM-T4 can avoid shear locking and gives accurate solutions in the analysis with compressible materials. The smoothed deformation gradient of ES-FEM-T4 is calculated as

$$\tilde{\mathbf{F}}^{\text{edge}} = \frac{\sum_{k \in \mathbb{K}} \tilde{\mathbf{F}}_k^{\text{elem}} V_k^{\text{elem}}}{\sum_{k \in \mathbb{K}} V_k^{\text{elem}}}. \quad (4)$$

In F-barES-FEM-T4,  $\tilde{\mathbf{F}}^{\text{edge}}$  is used to define the isovolumetric part of the deformation gradient as  $\bar{\mathbf{F}}_{\text{iso}}^{\text{edge}} = (\bar{J}^{\text{edge}})^{-1/3} \tilde{\mathbf{F}}^{\text{edge}}$ , where  $\bar{J}^{\text{edge}} \equiv \det(\tilde{\mathbf{F}}^{\text{edge}})$ .

Finally, the multi-smoothed deformation gradient at each edge,  $\bar{\mathbf{F}}^{\text{edge}}$ , is defined in manner of the F-bar method as

$$\bar{\mathbf{F}}^{\text{edge}} = \left( \frac{\bar{J}^{\text{edge}}}{\bar{J}^{\text{edge}}} \right)^{1/3} \tilde{\mathbf{F}}^{\text{edge}}. \quad (5)$$

### Nodal Internal Force and Stiffness

The Cauchy stress at each edge  $\mathbf{T}^{\text{edge}}$  is calculated in the standard way with  $\bar{\mathbf{F}}^{\text{edge}}$ . In contrast to the methods based on the selective integration, F-barES-FEM-T4 has no difficulty in treating dev/vol coupled materials.

The contribution of each edge to the nodal internal force is then calculated in manner of the F-bar method as

$$\{f_{\text{int}}^{\text{edge}}\} = [\tilde{B}^{\text{edge}}]\{T^{\text{edge}}\}V^{\text{edge}}, \quad (6)$$

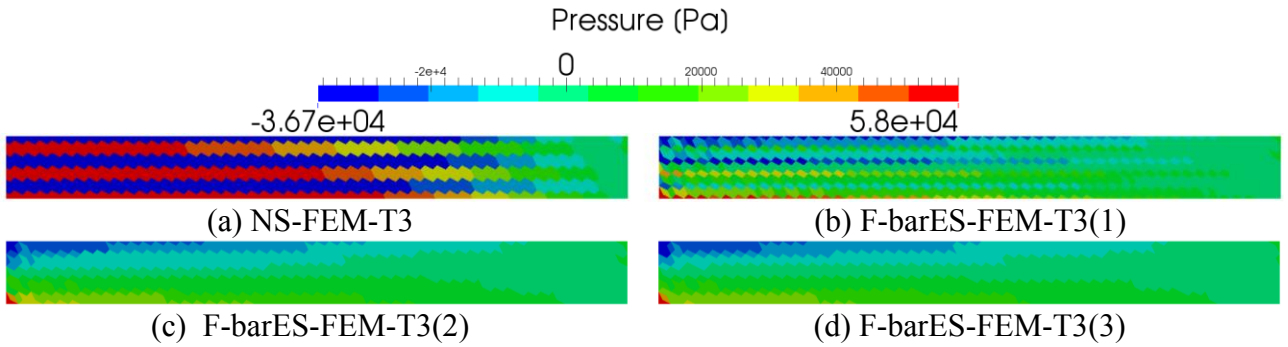
where  $[\tilde{B}^{\text{edge}}]$  is the smoothed  $B$  matrix of ES-FEM-T4 at each edge. In contrast to the methods based on mixed variational formulation, F-barES-FEM-T4 is purely a displacement-based formulation and thus eliminates the use of pressure or volumetric strain unknowns.

The contribution of each edge to the tangent stiffness matrix,  $[K^{\text{edge}}]$ , necessitates a little complicated way of derivation due to the cyclic smoothing, thereby omitting in this paper. As the number of the smoothing cycles is increased, the bandwidth of  $[K]$  is also increased; however, there is no increase in DOF.

## Results

### *Bending of Cantilever*

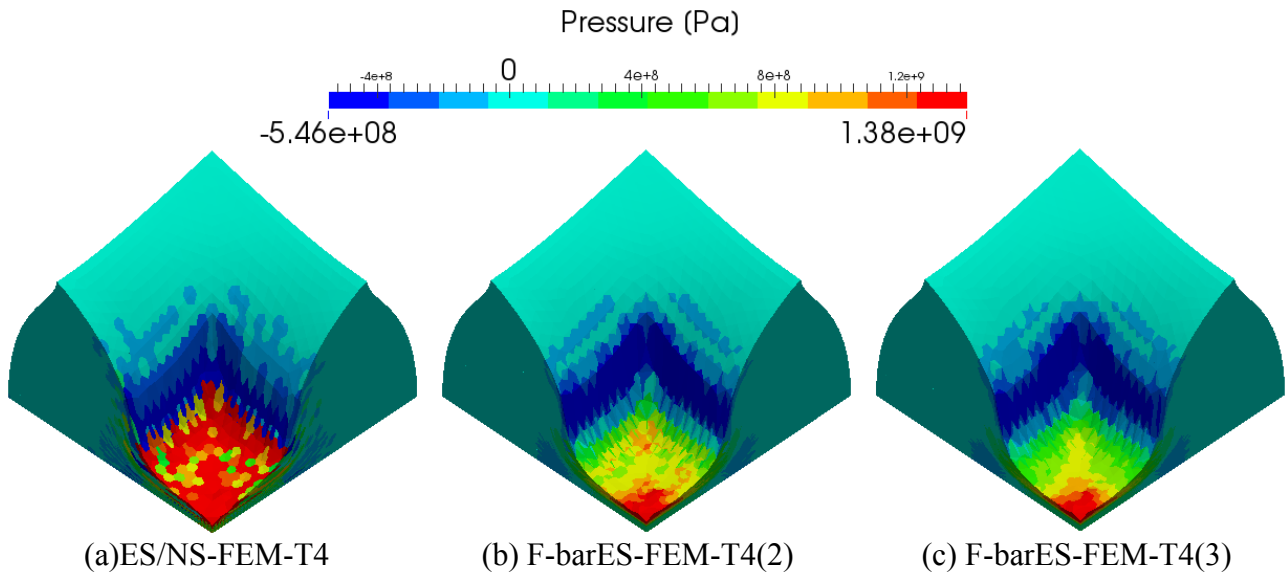
Figure 1 shows the pressure distributions of a small deformation cantilever bending analyses in 2D plane-strain condition. The analysis domain is a  $10 \text{ m} \times 1 \text{ m}$  rectangle discretized with a structural mesh of 500 triangular elements. Its left-hand side is perfectly constrained, and a concentrated bending load is assigned to the top-right node. The material is the neo-Hookean hyperelastic material with an initial Poisson's ratio of 0.499. It is clearly shown that NS-FEM-T3 raises severe pressure oscillation (Fig. 1 (a)), whereas our method with  $k$  times cyclic smoothing, F-barES-FEM-T3( $k$ ), suppresses pressure oscillation to some extent (Fig. 1 (b)-(d)). Our heuristic diagnostics suggest that  $\nu_{\text{ini}} < 0.48 \Rightarrow k = 0$  (i.e., ES-FEM is recommended),  $0.48 \leq \nu_{\text{ini}} < 0.49 \Rightarrow k = 1$ ,  $0.49 \leq \nu_{\text{ini}} < 0.499 \Rightarrow k = 2$ ,  $0.499 \leq \nu_{\text{ini}} < 0.4999 \Rightarrow k = 3$  and so on.



**Figure 1. Pressure distributions of the plane-strain small deformation cantilever bending analysis. The contour range is adjusted to the result of (d). Increase in the number of cyclic smoothing helps suppressing the pressure oscillation in the nearly incompressible material.**

### *Partial Compression of Block*

Figure 2 shows the pressure distributions of a finite strain block compression analyses in 3D. The analysis domain is a  $1 \text{ m}$  cube discretized with an unstructured mesh of 55821 tetrahedral elements. Its  $-x$ ,  $-y$  and  $-z$  faces are constrained in the normal direction of each face, its  $+z$  face is constrained in  $x$  and  $y$  directions, and a pressure is applied to  $1/4$  of  $+z$  face. The material is the Arruda-Boyce hyperelastic material with an initial Poisson's ratio of 0.499. As noted in Introduction, ES/NS-FEM-T4 suffers from pressure oscillation (Fig. 2 (a)). In contrast, our F-barES-FEM-T3(2) and F-barES-FEM-T3(3) suppress pressure oscillation to some extent (Fig. 2 (b), (c)).



**Figure 2. Pressure distributions of the finite strain block compression analysis. The contour range is adjusted to the result of (c). Our method with three-time cyclic smoothing, F-barES-FEM-T4(3), successfully suppresses the pressure oscillation in finite strain analysis of nearly incompressible materials.**

## Conclusions

A novel S-FEM called F-barES-FEM-T4 was proposed. The cyclic smoothing of  $J (\equiv \det(\mathbf{F}))$  between elements and nodes was introduced to combine the idea of F-bar method and ES-FEM-T4 efficiently. Accordingly, F-barES-FEM-T4 suppressed the pressure oscillation in nearly incompressible materials with no increase in DOF.

Meanwhile, the relations among the Poisson's ratio, the number of smoothing cycles, underestimation of stress concentrations, etc. are still unclear. The impact of the increase in the matrix bandwidth on the computational efficiency is also unclear. Further investigation about them is our future work.

## References

- Liu, G. R. and Nguyen, T. T. (2010) Smoothed finite element methods, CRC Press, Boca Raton, USA.
- Neto, E. A. S., Pires F. M. A., and Owen D. R. J. (2005) F-bar-based linear triangles and tetrahedra for finite strain analysis of nearly incompressible solids. Part i: formulation and benchmarking, *International Journal for Numerical Methods in Engineering*, 62(3), 353–383.
- 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.
- Ong, T. H., Heaney, C. E., Lee, C. K., Liu, G. R., 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, 315–345.
- Neto, E. A. S., Pires D., Dutko, M., and Owen D. R. J. (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.

# A multi-level meshless method based on an implicit use of the method of fundamental solutions

C. Gáspár<sup>1</sup>

<sup>1</sup>Department of Mathematics and Computer Sciences, Széchenyi István University,  
Egyetem tér 1, H-9026 Győr, Hungary

gasparcs@sze.hu

## Abstract

A new version of the Method of Fundamental Solutions is proposed. Instead of locating external point sources, an external boundary should be defined, and an extension of the original solution is created by enforcing the original boundary conditions for the extended problem. This leads to a better conditioned problem than the traditional Method of Fundamental Solutions. To numerically solve the extended problem, a quadtree-based multi-level finite volume method is used, which is quite economical from the computational points of view. In addition to it, the problem of large, dense and ill-conditioned systems is completely avoided.

**Keywords:** Meshless Methods, Method of Fundamental Solutions, Multi-level Methods, Quadtrees

## Introduction

Due to its simplicity and accuracy, the Method of Fundamental Solutions (MFS, see e.g. [Alves et al. 2002]) has become quite popular among the meshless methods. This approach can be applied easily, if the fundamental solution of the original partial differential equation is known (or the original problem can be converted to such a problem). In its traditional form, the approximate solution is sought as a linear combination of the fundamental solution shifted to some external source points:

$$u(x) \approx \sum_{j=1}^N \alpha_j \Phi(x - \tilde{x}_j), \quad (1)$$

where  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$  are predefined *source points* located in the exterior of the domain  $\Omega$  of the original partial differential equation,  $\Phi$  denotes the fundamental solution. In case of the familiar second-order partial differential equations,  $\Phi$  has a (weak) singularity at the origin, so that the approximate solution (1) has singularities at the source points but remains smooth inside the domain. The a priori unknown coefficients  $\alpha_1, \alpha_2, \dots, \alpha_N$  can be computed by enforcing the boundary conditions at some *boundary collocation points*  $x_1, x_2, \dots, x_N$ :

$$\begin{aligned} \sum_{j=1}^N \alpha_j \Phi(x_k - \tilde{x}_j) &= u_k & (x_k \in \Gamma_D) \\ \sum_{j=1}^N \alpha_j \frac{\partial \Phi}{\partial n_k}(x_k - \tilde{x}_j) &= v_k & (x_k \in \Gamma_N) \end{aligned} \quad (2)$$

Here  $\Gamma_D$  and  $\Gamma_N$  denote the Dirichlet part (and the Neumann part, respectively) of the boundary of the domain  $\Omega$ ;  $n_k$  is the outward normal unit vector at the point  $x_k$ ;

$u_k, v_k$  are the Dirichlet and Neumann boundary conditions at the boundary collocation points.

Generally speaking, the MFS has excellent convergence properties (see [Li (2005)]); however, it is well known that the matrix of the system (2) is fully populated, non-selfadjoint and severely ill-conditioned, especially when the sources are located far from the boundary. On the other hand, if they are located too close to the boundary, numerical singularities are generated. In addition to it, the proper definition of the location of the sources can hardly be automatized. Therefore it is popular to allow the source points and the boundary collocation points to coincide, which needs special tricks to handle the problem of appearing singularities.

To circumvent these difficulties, a lot of approaches have been developed. In the boundary knot method [Chen (2002)], [Chen et al. (2005)], nonsingular general solutions are used instead of fundamental solutions. It is also possible to use fundamental solutions concentrated to *lines* rather than discrete points, see [Gáspár (2013b)]. In both approaches, the solution is approximated by nonsingular functions, thus, the problem of singularity is avoided. However, though they have especially good accuracy, the resulting linear system is severely ill-conditioned, which can cause serious computational difficulties. The problem is more difficult, if the original fundamental solution is used. Utilizing some boundary mesh structure, the appearing singular integrals can be evaluated analytically, see [Young et al. (2005)]. The desingularization can also be carried out by solving some auxiliary Dirichlet subproblem, as in the Modified Method of Fundamental Solutions (MMFS, see [Šarler (2008; 2009)]), or in the Boundary Distributed Source method (BDS, see [Liu (2010)]). It is also possible to combine the above approach with the use of approximate fundamental solutions which have no singularity at the origin; such an approximate fundamental solution may be the fundamental solution of the fourth-order Laplace-Helmholtz operator  $\Delta(\Delta - c^2 I)$  with a sufficiently large scaling parameter  $c$ , (the Regularized Method of Fundamental Solutions, see [Gáspár (2013a; 2013b)]). A further possibility is that, in contrast to Eq. (1), the approximate solution is sought as a linear combination of the *normal derivatives* of  $\Phi$  (dipole formulation), which can be considered a discretization of an indirect boundary integral equation based on a double layer potential and so forth.

Most of the above methods are indirect in the sense that they convert the original problem to the determination of some coefficients of the linear combination (i.e. the strengths of the point sources in the original MFS-formulation). These coefficients control the values (and their derivatives) at the collocation points, thus control the whole approximate solution inside the domain.

In this paper, we present a technique in which the values at the collocation points are controlled by the values at some external boundaries. In another words, an extension (continuation) of the solution is computed directly by prescribing certain, a priori unknown ‘external boundary conditions’. If the external boundary is located sufficiently close to the original boundary, this results in much better conditioned problem. The external conditions are updated iteratively. During the iteration procedure, the original problem should be solved in a larger domain, which can economically be performed by using quadtree-based multi-level tools. In addition to

this, the problem of large, fully populated and ill-conditioned matrices is completely avoided.

### Approximation of the solution by external boundary conditions

To illustrate the above outlined idea, consider the 2D model problem with pure Dirichlet boundary conditions:

$$\Delta U = 0 \quad \text{in } \Omega, \quad U|_{\Gamma} = u \quad (3)$$

defined on a circle  $\Omega := \Omega_R := \{x \in \mathbf{R}^2 : \|x\| < R\}$  with boundary  $\Gamma_R := \partial\Omega_R$ . Let us express the function  $u$  in terms of (complex) Fourier series:

$$u(t) = \sum_k \alpha_k \cdot e^{ikt}, \quad (4)$$

where, for the sake of simplicity, we used polar coordinates. Then the solution of (3) is expressed as:

$$U(r, t) = \sum_k \alpha_k \cdot \left(\frac{r}{R}\right)^{|k|} \cdot e^{ikt} \quad (5)$$

Now consider a larger circle  $\tilde{\Omega} := \Omega_{R+\delta}$  with boundary  $\tilde{\Gamma} := \partial\Omega_{R+\delta}$ , where  $\delta > 0$ . Then Eq. (5) defines a harmonic continuation of  $U$  to  $\tilde{\Omega}$ . On the boundary this yields:

$$\tilde{u}(t) = \sum_k \alpha_k \cdot \left(1 + \frac{\delta}{R}\right)^{|k|} \cdot e^{ikt}, \quad (6)$$

provided that the series is convergent in a proper function space. (This is sometimes not the case due to the exponentially increasing factor  $\left(1 + \frac{\delta}{R}\right)^{|k|}$ .)

Conversely, if  $U$  is prescribed along the external boundary  $\tilde{\Gamma}$ :

$$\Delta U = 0 \quad \text{in } \tilde{\Omega}, \quad U|_{\tilde{\Gamma}} = \tilde{u}, \quad (7)$$

where  $\tilde{u}(t) = \sum_k \beta_k \cdot e^{ikt}$ , then the restriction of  $U$  to  $\Gamma$  defines a (much more smooth) function:

$$u(t) = \sum_k \beta_k \cdot \left(1 + \frac{\delta}{R}\right)^{-|k|} \cdot e^{ikt} \quad (8)$$

The operator  $\tilde{A} : H^{1/2}(\tilde{\Gamma}) \rightarrow H^{1/2}(\Gamma)$  is always bounded and  $\|\tilde{A}\| < 1$ , but it is not regular, since the inverse operator is not bounded. However, the discretized operators might have uniformly bounded inverses, if the distance  $\delta$  itself depends on the discretization. Let us discretize the above problems by cutting the Fourier series (6) and (8) at a maximal index  $N$ . Define the distance  $\delta$  of the original and the external boundary to be inversely proportional to  $N$ , i.e.  $\delta := \delta_0 \cdot \frac{2R\pi}{N}$ . Then Eqs. (4), (6) are rewritten as:

$$u_N(t) = \sum_{|k| \leq N} \alpha_k \cdot e^{ikt}, \quad \tilde{u}_N(t) = \sum_{|k| \leq N} \alpha_k \cdot \left(1 + \frac{2\pi\delta_0}{N}\right)^{|k|} \cdot e^{ikt} \quad (9)$$

Since  $1 \leq \left(1 + \frac{2\pi\delta_0}{N}\right)^{|k|} \leq \left(1 + \frac{2\pi\delta_0}{N}\right)^N \leq e^{2\pi\delta_0}$ , the operators  $\tilde{A}_N^{-1}$  defined by  $\tilde{A}_N \tilde{u}_N = u_N$  are uniformly bounded, moreover, the condition numbers are also bounded:  $\text{cond}(\tilde{A}_N) \leq e^{2\pi\delta_0}$ , independently of  $N$ . Therefore the ‘discrete’ problem:

- find  $\tilde{u}_N$  in such a way that the solution of the Dirichlet problem

$$\Delta U = 0 \quad \text{in } \tilde{\Omega}, \quad U|_{\tilde{\Gamma}} = \tilde{u}_N, \quad (10)$$

satisfies the original ‘inner’ boundary condition  $U|_{\Gamma} = u_N$ ;

is now a well-conditioned problem, independently on  $N$  (which characterizes the fineness of the discretization).

Based on the above considerations, in order to solve the Dirichlet problem (3), it is sufficient to solve the extended problem

$$\Delta U = 0 \quad \text{in } \tilde{\Omega}, \quad U|_{\tilde{\Gamma}} = \tilde{u}, \quad (11)$$

where the boundary condition  $\tilde{u}$  should be chosen in such a way that the original boundary condition

$$U|_{\Gamma} = u, \quad (12)$$

is satisfied. That is, the approximate solution of the original boundary value problem is controlled by the external boundary condition  $\tilde{u}$  rather than external sources as in the case of the traditional method of fundamental solutions. It is expected that, in order to enforce the boundary condition (12), the simple, traditional iterative methods e.g. the Seidel iteration or the simplest Richardson iteration

$$\tilde{u}_{n+1} := \tilde{u}_n - \omega \cdot (U_n|_{\Gamma} - u) \quad (13)$$

can be applied, where  $\omega > 0$  is a predefined iteration parameter. In the above model problem, this remains the case.

The external boundary  $\tilde{\Gamma}$  can be defined in a flexible way. The approach can easily be generalized for mixed boundary conditions and 3D problems as well.

#### *Solution of the extended problem*

In principle, the extended problem (11) can be solved by an arbitrary method. Since the external boundaries are not predefined, it is worth defining them in such a way that the extended problem (11) could easily be handled. In the next examples, the extended problem is solved by the regularized method of fundamental solutions [Gáspár (2013a)], which does not result in an optimal technique from a computational point of view, however, it demonstrates well the main idea of the approach.

*Example 1.* Let  $\Omega$  be the unit circle and discretize the boundary  $\Gamma$  by the boundary collocation points  $x_1, x_2, \dots, x_N \in \Gamma$  in an equidistant way. Let the external boundary

$\tilde{\Gamma}$  be a concentric circle with radius  $1 + \delta$ , where  $\delta := \delta_0 \cdot \frac{2\pi}{N}$  ( $\delta_0$  is a constant of proportionality). That is, the distance of the original and the external boundary is proportional to the characteristic distance of the boundary collocation points. Consider the pure Dirichlet problem (3) with the simple test solution

$$U(x, y) = y \quad (13)$$

(using the more familiar notations  $x, y$  for the space variables). The extended problem (11) is (approximately) solved by the regularized method of fundamental solution (see [Gáspár (2013a)]), using the truncated fundamental solution:

$$\Phi(x) := \begin{cases} \frac{1}{2\pi} \log \|x\| & \text{if } \|x\| \geq \varepsilon \\ \frac{1}{2\pi} \log \varepsilon & \text{if } \|x\| < \varepsilon \end{cases} \quad (14)$$

In this example, the truncation parameter  $\varepsilon$  was set to the value  $\varepsilon := \frac{1}{4} \cdot \frac{2\pi(1+\delta)}{N}$ .

The collocation points on the external boundary  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N \in \tilde{\Gamma}$  were defined from the original collocation points  $x_1, x_2, \dots, x_N$  by shifting them in the outward normal direction. The extended solution is sought in the form:

$$U(x) := \sum_{j=1}^N \tilde{\alpha}_j \Phi(x - \tilde{x}_j), \quad (15)$$

where the coefficients  $\tilde{\alpha}_j$  are determined by the system of equations:

$$\sum_{j=1}^N \tilde{\alpha}_j \Phi(\tilde{x}_k - \tilde{x}_j) =: \sum_{j=1}^N \tilde{\alpha}_j \tilde{A}_{kj} = \tilde{u}_k \quad (k = 1, 2, \dots, N) \quad (16)$$

In short:

$$\tilde{\mathbf{A}} \tilde{\boldsymbol{\alpha}} = \tilde{\mathbf{u}} \quad (17)$$

(where the matrix  $\tilde{\mathbf{A}}$  and the column vectors  $\tilde{\boldsymbol{\alpha}}, \tilde{\mathbf{u}}$  are assembled from the entries defined in Eq. (16)).

At the original boundary collocation points:

$$\sum_{j=1}^N \tilde{\alpha}_j \Phi(x_k - \tilde{x}_j) =: \sum_{j=1}^N \tilde{\alpha}_j A_{kj} = u_k \quad (k = 1, 2, \dots, N), \quad (18)$$

In short:

$$\mathbf{A} \tilde{\boldsymbol{\alpha}} = \mathbf{u} \quad (19)$$

Eliminating the vector of coefficients  $\tilde{\boldsymbol{\alpha}}$ , we have:

$$\mathbf{A} \tilde{\mathbf{A}}^{-1} \tilde{\mathbf{u}} = \mathbf{u} \quad (20)$$



*Remark:* In the traditional method of fundamental solutions, the coefficients  $\tilde{\alpha}$  are determined by enforcing the original boundary conditions at the original boundary collocation points, i.e. by solving Eq. (19) only. However, the matrix  $\mathbf{A}$  is generally much more ill-conditioned than  $\mathbf{A}\tilde{\mathbf{A}}^{-1}$ . That is, the original boundary conditions are more easily controllable by the external boundary conditions than by the strengths of the external sources. If the distance  $\delta$  is small enough, then the matrix  $\mathbf{A}\tilde{\mathbf{A}}^{-1}$  is diagonally dominant, and the Richardson iteration

$$\tilde{\mathbf{u}}_{n+1} := \tilde{\mathbf{u}}_n - \omega \cdot (\mathbf{A}\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{u}}_n - \mathbf{u}) \quad (21)$$

is convergent for a sufficiently small iteration parameter  $\omega > 0$ .

In this example, the iteration parameter was set to  $\omega := 1$ . Table 1 shows the condition numbers of the matrices  $\mathbf{A}$  and  $\mathbf{A}\tilde{\mathbf{A}}^{-1}$  as well as the relative  $L_2$ -errors of the approximate solution (in %, after 5 Richardson iterations (21)) at the boundary collocation points with different numbers of boundary collocation points ( $N$ ). The constant of proportionality  $\delta_0$  was set to  $\delta_0 := 2$ . The results demonstrate that the system (20) is really much better conditioned than the system (19) obtained by the classical method of fundamental solutions.

**Table 1. Results of Example 1**

$N$	16	32	64	128	256	512
cond( $\mathbf{A}$ )	478	1.1E+3	4.1E+3	1.1E+4	2.8E+4	6.2E+4
cond( $\mathbf{A}\tilde{\mathbf{A}}^{-1}$ )	46	89	139	180	207	226
Rel. $L_2$ -error (after 5 Richardson iterations)	0.4528	0.0317	0.0012	3.25E-5	1.54E-5	2.44E-5

*Example 2.* The difference between the test problems of Example 1 and Example 2 is that now a mixed boundary condition is prescribed:

$$U|_{\Gamma_D} = u, \quad \frac{\partial U}{\partial n}|_{\Gamma_N} = v, \quad (22)$$

A half of the boundary was considered to be the Dirichlet part  $\Gamma_D$ , and the remaining part was treated as a Neumann boundary  $\Gamma_N$ . The boundary conditions were defined to be consistent with the test solution (13). In principle, it is possible to use the same strategy as earlier, i.e. to control the original mixed boundary by pure Dirichlet condition on the external boundary, but this seemed to result in slow convergence. Instead, let us control the original mixed boundary by a similar mixed boundary condition on the external boundary as shown in the followings.

The extended problem:

$$\Delta U = 0 \quad \text{in } \tilde{\Omega}, \quad U|_{\tilde{\Gamma}_D} = \tilde{u}, \quad \frac{\partial U}{\partial n}|_{\tilde{\Gamma}_N} = \tilde{v} \quad (23)$$

is solved again by a version of the regularized method of fundamental solutions, assuming the approximate solution in the following form

$$U(x) := \sum_{j=1}^N \tilde{\alpha}_j \Phi(x - \tilde{x}_j), \quad (24)$$

where  $\Phi$  denotes again the truncated fundamental solution (14). The Dirichlet boundary condition is treated as earlier, but the proper treatment of the Neumann condition needs a desingularization procedure [Šarler (2008; 2009)], [Liu (2010)], [Gáspár (2013a; 2013b).] The normal derivatives of  $U$  are expressed as:

$$\frac{\partial U}{\partial n}(x) = \sum_{j=1}^N \tilde{\alpha}_j \frac{\partial \Phi}{\partial n}(x - \tilde{x}_j) \quad (24)$$

Consequently, the boundary values of the external boundary satisfy

$$\begin{aligned} \sum_{j=1}^N \tilde{\alpha}_j \Phi(\tilde{x}_k - \tilde{x}_j) &=: \sum_{j=1}^N \tilde{\alpha}_j \tilde{A}_{kj} = \tilde{u}_k & (\tilde{x}_k \in \tilde{\Gamma}_D) \\ \sum_{j=1}^N \tilde{\alpha}_j \frac{\partial \Phi}{\partial n_k}(\tilde{x}_k - \tilde{x}_j) &=: \sum_{j=1}^N \tilde{\alpha}_j \tilde{B}_{kj} = \tilde{v}_k & (\tilde{x}_k \in \tilde{\Gamma}_N) \end{aligned} \quad (25)$$

Note that the diagonal entries  $\tilde{B}_{kk}$  should be defined in a special way (by solving a pure Dirichlet subproblem in the extended domain) due to the desingularization procedure, see [Liu (2010)], [Gáspár (2013a)] for details.

The original boundary conditions can be enforced by the following system of equations:

$$\begin{aligned} \sum_{j=1}^N \tilde{\alpha}_j \Phi(x_k - \tilde{x}_j) &=: \sum_{j=1}^N \tilde{\alpha}_j A_{kj} = u_k & (x_k \in \Gamma_D) \\ \sum_{j=1}^N \tilde{\alpha}_j \frac{\partial \Phi}{\partial n_k}(x_k - \tilde{x}_j) &=: \sum_{j=1}^N \tilde{\alpha}_j B_{kj} = v_k & (x_k \in \Gamma_N) \end{aligned} \quad (26)$$

Let us build up the following matrices and vectors:

$$\begin{aligned} \tilde{C}_{kj} &:= \tilde{A}_{kj}, & C_{kj} &:= A_{kj}, & \tilde{w}_k &:= \tilde{u}_k, & w_k &:= u_k & (x_k \in \Gamma_D) \\ \tilde{C}_{kj} &:= \tilde{B}_{kj}, & C_{kj} &:= B_{kj}, & \tilde{w}_k &:= \tilde{v}_k, & w_k &:= v_k & (x_k \in \Gamma_N) \end{aligned} \quad (26)$$

Then we have:

$$\tilde{C}\tilde{\alpha} = \tilde{w}, \quad C\alpha = w \quad (27)$$

Eliminating the vector of coefficients  $\tilde{\alpha}$ , we have:

$$C\tilde{C}^{-1}\tilde{w} = w, \quad (28)$$

Once the external boundary conditions  $\tilde{w}$  have been computed, the coefficients can also be computed by  $\tilde{\alpha} := \tilde{C}^{-1}\tilde{w}$ . Thus, the approximate solution on the original boundary  $A\tilde{\alpha} = \mathbf{u}$ , which makes it possible to directly compute the  $L_2$ -error of approximation at the boundary collocation points (referred to as ‘direct solution’ in

Table 2). The iteration parameter was set again to  $\omega := 1$ . Table 2 shows the condition numbers of the matrices  $\mathbf{C}$  and  $\mathbf{C}\tilde{\mathbf{C}}^{-1}$  as well as the relative  $L_2$ -errors of the approximate solution (in %, after 5 Richardson iterations) at the boundary collocation points with different numbers of boundary collocation points ( $N$ ). The constant  $\delta_0$  was set to  $\delta_0 := 2$ . The results show that the method still works in case of mixed boundary conditions.

Finally note that the extended solution can be computed also in a quite economical way based on a non-uniform cell system and multi-level tools. This is outlined in the next section.

**Table 2. Results of Example 2**

$N$	16	32	64	128	256	512
cond( $\mathbf{C}$ )	322	1.1E+3	4.1E+3	1.2E+4	2.9E+4	6.5E+4
cond( $\mathbf{C}\tilde{\mathbf{C}}^{-1}$ )	65	218	1.0E+3	5.3E+3	2.4E+3	1.0E+5
Rel. $L_2$ -error (direct solution)	0.0407	0.0077	0.0026	0.0014	0.0010	0.0008
Rel. $L_2$ -error (after 5 Richardson iterations)	0.8515	0.2100	0.0526	0.0137	0.0036	0.0012

### Multi-level solution using quadtree-based cell systems

From a computational point of view, the realization of the above methods is far from being optimal. However, if the extended problem (11) - (12) or (23) is handled directly, this makes it possible to use the much more economical multi-level techniques. Here a quadtree- (QT-) based finite volume method is used (see [Gáspár (2000)]). (The natural 3D generalization is based on the octtree cell system.) Strictly speaking, this is a domain type method; however, the generation of the cell system is performed entirely on the basis of the boundary collocation points in a completely automatic way, so that it can be considered a meshless method. The computational cost as well as the memory requirement is typically  $\mathcal{O}(N \cdot \log N)$  only.

Thus, the solution algorithm is as follows.

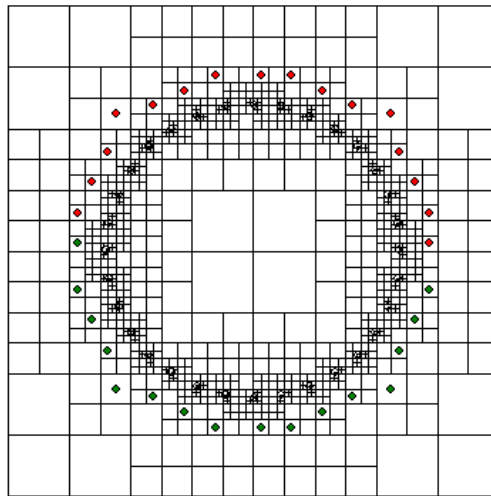
- Generate a quadtree cell system by the boundary collocation points  $x_1, x_2, \dots, x_N$ . This results in a nested cell system with automatically created local refinements at the boundary collocation points. By additional subdivisions, it is possible to assure that the ratio of the sizes of the neighbouring cells is at most 2, i.e. no abrupt changes occur in cell sizes (regularization of the QT-cell system).
- Shift the points  $x_1, x_2, \dots, x_N$  in the outward normal direction with the distance  $\delta$ . Determine the leaf cells of the QT-cell system which contain these points. (These cells have typically larger sizes than the finest cells containing

the boundary collocation points.) Define the external boundary points  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N$  to be these cell centers.

- Using simple cell-centered finite volume schemes, solve the extended problem and update the external boundary condition by e.g. a Richardson iteration. Repeat this step until convergence.

The solution procedure can be embedded in a natural multi-level context (see [Gáspár (2000; 2004)] for details).

*Example 3.* Let  $\Omega$  be a circle contained in the unit square  $[0,1] \times [0,1]$  centered at the point  $[0.5, 0.5]$  with radius  $R := 0.30$ . Discretize the boundary  $\Gamma$  by the boundary collocation points  $x_1, x_2, \dots, x_N \in \Gamma$  in an equidistant way. A regular QT-cell system was generated by recursively subdividing the unit square based on the boundary collocation points. The maximal subdivision level was 8, i.e. the finest cell size was  $1/256$ . The collocation points of the external boundary  $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N \in \tilde{\Gamma}$  were defined as the cell centers of the boundary collocation points shifted in the outward normal direction with distance  $\delta := \frac{2R\pi}{N}$ . Figure 1. illustrates the QT-cell system and the external boundary points generated by 32 boundary collocation points.



**Figure 1. Quadtree cell system and external boundary points generated by 32 boundary collocation points**

The test solution was as follows:

$$U(x, y) = -x + 2y + \frac{1}{2} \quad (29)$$

(using the more familiar notations  $x, y$  for the space variables). Mixed boundary conditions were prescribed: on a half of the boundary, Dirichlet boundary condition was supposed, while the remaining part of the boundary was considered as Neumann boundary. Along the boundary of the initial unit square, a separate boundary

condition can be prescribed independently of the original boundary conditions; in this example, a homogeneous Dirichlet boundary condition was imposed. The method gives the approximate solutions in the interior and the exterior of the original domain at the same time. The Dirichlet data at the external boundary points were updated by Richardson iteration (13).

Another variant of the method was also tested. Here mixed boundary conditions were prescribed also along the external boundary. That is, the original Neumann boundary condition was controlled by an external Neumann boundary condition, which was updated by Richardson iteration as well:

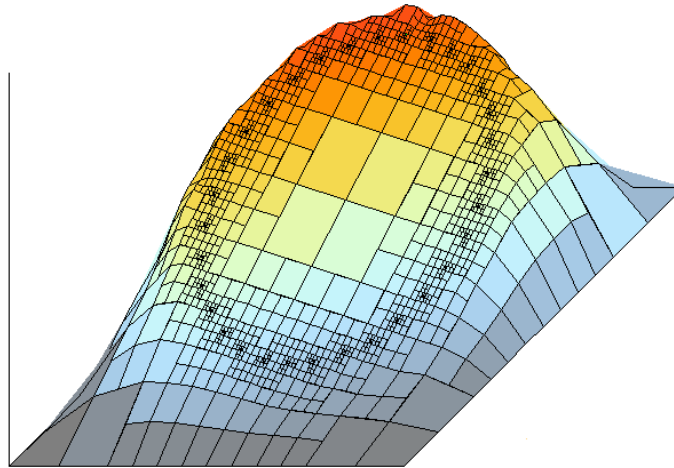
$$\tilde{v}_{n+1} := \tilde{v}_n - \omega \cdot \left( \frac{\partial U_n}{\partial n} \Big|_{\Gamma} - v \right) \quad (30)$$

At the Neumann part, the external boundary was supposed to be the union of circles centered at the external Neumann points; the radii were defined to be proportional to the characteristic distance of the external Neumann boundary points. This boundary condition was implemented on the (coarser) QT-cells containing the external Neumann points only. (The role of external boundaries is only to control the original boundary conditions at the original boundary collocation points, therefore the solution at the external boundaries is allowed to be less exact than at the original boundary.) Table 3 shows the relative  $L_2$ -errors (in %) at the original boundary collocation points in both cases. Here ‘Method 1’ refers to the method which controls the mixed boundary conditions via external Dirichlet boundary condition, while ‘Method 2’ corresponds to the external mixed boundary conditions. It can be clearly seen that the exactness of the two variants is essentially the same: however, the second variant has proved faster.

**Table 3. Relative  $L_2$ -errors using QT-cell system, Example 3**

$N$	16	32	64	128	256
Rel. $L_2$ -error (Method 1)	0.7035	0.1606	0.0918	0.0338	0.005
Rel. $L_2$ -error (Method 2)	0.7036	0.1606	0.0913	0.0337	0.004

For illustration, Figure 2 shows the approximate solution on the QT-cell system with 32 boundary collocation points. Along the boundary of the initial rectangle of the QT-subdivision, a homogeneous Dirichlet boundary condition was prescribed. In the vicinity of the external boundary, the solution is less smooth than the interior of the original domain, similarly to the case of the traditional method of fundamental solutions. However, the irregularity is much less, due to the fact that the solution is controlled by external boundary condition rather than the strengths of the external singularities; moreover, the cell system is allowed to be coarser here than in the vicinity of boundary collocation points.



**Figure 2. Approximate test solution on a quadtree cell system generated by 32 boundary collocation points**

## Conclusions

In this paper, the original idea of the Method of Fundamental Solution has been extended in the sense that the approximate solution was sought as a solution of an extended problem with an external boundary. This often results in a well-conditioned problem provided that the external boundary is located sufficiently close to the original boundary (depending also on the discretization). The method controls the values along the original boundary via the external boundary conditions. These external boundary conditions are adjusted iteratively using familiar, simple iterative techniques. The external boundary can be defined in a flexible way. In the vicinity of the external boundary, the approximate solution is allowed to be less exact than along the original boundary, which makes it possible to apply a coarser discretization at the external boundaries. The numerical benefit of the approach is that the extended problem can be handled by the quite economical quadtree-based multi-level method. Moreover, the problem of large, dense and ill-conditioned systems of equations is also avoided.

## References

- Alves, C. J. S., Chen, C. S. and Šarler, B. (2002) The method of fundamental solutions for solving Poisson problems. In: *Int. Series on Advances in Boundary Elements. (Proceedings of the 24th International Conference on the Boundary Element Method incorporating Meshless Solution Seminar)*. Eds: C.A.Brebbia A.Tadeu, V.Popov, Vol. 13. WitPress, Southampton, Boston, 67–76.
- Chen, W. (2002) Symmetric boundary knot method. *Engineering Analysis with Boundary Elements* **26**, 489–494.
- Chen, W., Shen, L. J., Shen, Z. J. and Yuan, G. W. (2005) Boundary Knot Method for Poisson Equations. *Engineering Analysis with Boundary Elements* **29**, 756–760.
- Gáspár, C. (2000) Multi-level biharmonic and bi-Helmholtz interpolation with application to the boundary element method. *Engineering Analysis with Boundary Elements* **24/7-8**, 559–573.
- Gáspár, C. (2004) A meshless polyharmonic-type boundary interpolation method for solving boundary integral equations. *Engineering Analysis with Boundary Elements* **28/10**, 1207–1216.
- Gáspár, C. (2013a) Some Variants of the Method of Fundamental Solutions: Regularization Using Radial and Nearly Radial Basis Functions. *Central European Journal of Mathematics* **11**, No. 8, 1429–1440.

- Gáspár, C. (2013b) A Regularized Method of Fundamental Solutions Without Desingularization. *Computer Modeling in Engineering & Sciences* **92**, No. 1, 103-121.
- Li, X. (2005) On convergence of the method of fundamental solutions for solving the Dirichlet problem of Poisson's equation. *Advances in Computational Mathematics* **23**, 265–277.
- Liu, Y. J. (2010) A new boundary meshfree method with distributed sources. *Engineering Analysis with Boundary Element* **34**, 914–919.
- Šarler, B. (2008) Desingularised Method of Double Layer Fundamental Solutions for Potential Flow Problems. In: *Boundary Elements and Other Mesh Reduction Methods XXX*. Eds: L.Skerget, C.A.Brebbia, WitPress, Southampton, Boston, 159–168.
- Šarler, B. (2009) Solution of Potential Flow Problems by the Modified Method of Fundamental Solutions: Formulations with the Single Layer and the Double Layer Fundamental Solutions. *Engineering Analysis with Boundary Elements* **33**, 1374–1382.
- Young, D. L., Chen, K. H. and Lee, C.W. (2005) Novel meshless method for solving the potential problems with arbitrary domin. *Journal of Computational Physics* **209**, 290–321.

## Bird strike on an engine primary compressor with high rotating speed: Numerical simulations and parametric study

Jia Huang<sup>1</sup>, \*Yulong Li<sup>1</sup>, Zhixue Zhang<sup>2</sup>, Jun Liu<sup>1</sup>, Zhongbin Tang<sup>1</sup>

<sup>1</sup>School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, P.R. China

<sup>2</sup>AVIC Shenyang Engine Design Institute

\*Corresponding author: [liyulong@nwpu.edu.cn](mailto:liyulong@nwpu.edu.cn)

### Abstract

Bird ingestion by jet engines will lead to serious consequences such as power loss, aircraft fires, and high speed blades debris, which is catastrophic to aviation. The primary compressor is the most vulnerable component during the transient impact. The method of experiment and simulation usually was used to anti-bird strike design on blade. However, simulation and experiment on compressor under bird strike at high rotating speed was rarely carried out.

In present paper, the model of a compressor impacted by bird is established by using commercial explicit code PAM-CRASH. The bird is meshed with smooth particles hydrodynamic (SPH) method and Murnaghan equation of state is employed to describe the fluid behaviour of the bird under high speed impact. The bird model shows no sign of instability and accurately characterized the splashing particles of the bird. Simulation of compressor revolving at speed of 3000r/min and 8000r/min impacted by bird with mass of 250g is preformed according to the experiments. The results show a good agreement between simulation and test, which indicates that the SPH - FE method could provide a very powerful numerical model in predicting the transient dynamic responses of engine structures in bird impact events. Finally, a number of parametric studies are conducted including: influence of the failure stain, the bird impact location, and impact timing.

**Keywords:** Bird impact, Engine blades, SPH method, Constitutive model, Transient finite element analysis

### Introduction

Just about one and a half years after the first airplane was invented, Wright Brothers recorded the first bird-strike on 7 September 1905 [1]. Nowadays, bird-strike puts almost every single flight in risk. According to Federal Aviation Administration (FAA) database [2], the number of reported wildlife strikes of civil aircraft increased dramatically from 1990 to 2013, as is shown in Fig.1. In America, 10,856 bird strikes occurred solely in the year of 2013 and 601 of them caused damage to aircraft components. Besides, over 50 planes of civil aviation have been destroyed and more than 220 people have died due to bird-impact accidents since 1912 [3].

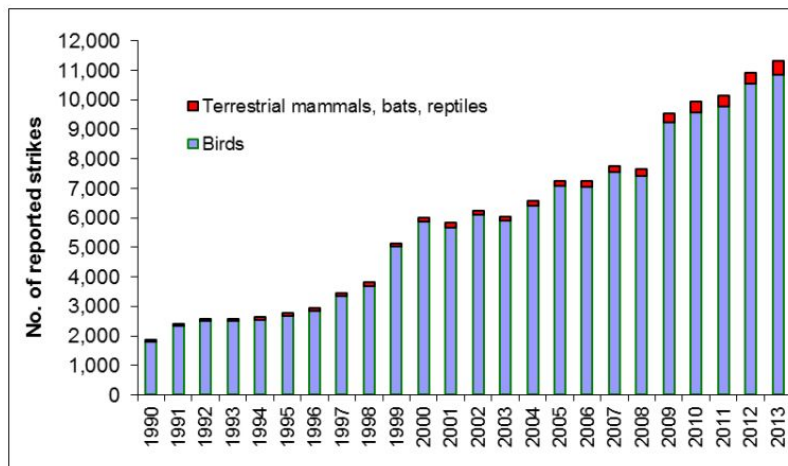


Figure 1. Number of reported wildlife strikes with civil aircraft, USA, 1990 – 2013.



Aircraft engines are most easily damaged by bird strikes (30 percent of all damaged components) compared with other components. They are particularly vulnerable while turning at a very high speed during the takeoff phase when the plane is at a low altitude where birds are commonly found. On 20th December, 2013, a Bombardier Challenger 300 hit 4 or 5 soaring turkey vultures at 1800 feet AGL while climbing from the Florida airport. Pilot declared an emergency and landed safely at an alternate airport. Repair costs for the engine and tail were over \$800,000. The relevant airline closed down for 22 days. Fortunately, no one was injured in the accident.



**Figure 2. The severely damaged engine**

In bird ingestion tests, engines are required to demonstrate their ability to withstand bird ingestion and the following ingestion to produce enough thrust as required by safety regulations. However, full scale of an aircraft engine testing is very expensive. In order to decrease this cost, a range of analysis methods applicable to bird strike simulation were developed. R.H.Mao et al. studied the nonlinear transient response of a bird striking a fan system using LS-DYNA. The bird is modeled as a fluid jet with a homogenized fluidic constitutive relation, using the Brockman hydrodynamics model [4]. They also analyzed the geometry effects of an artificial bird striking an aero engine fan blade, including hemispherical-ended cylinder, straight-ended cylinder, and ellipsoid [5]. Guan Yupu et al. set up a three-blade computational model of a fan rotor with shrouds. They compared the transient response curves from the simulation with that obtained from experiment and found that the variations in measured points and corresponding points of simulation agreed well [6,7]. Rade Vignjevic et al. carried out a number of parametric studies including: influence of the bird shape; the bird impact location and impact timing [8].

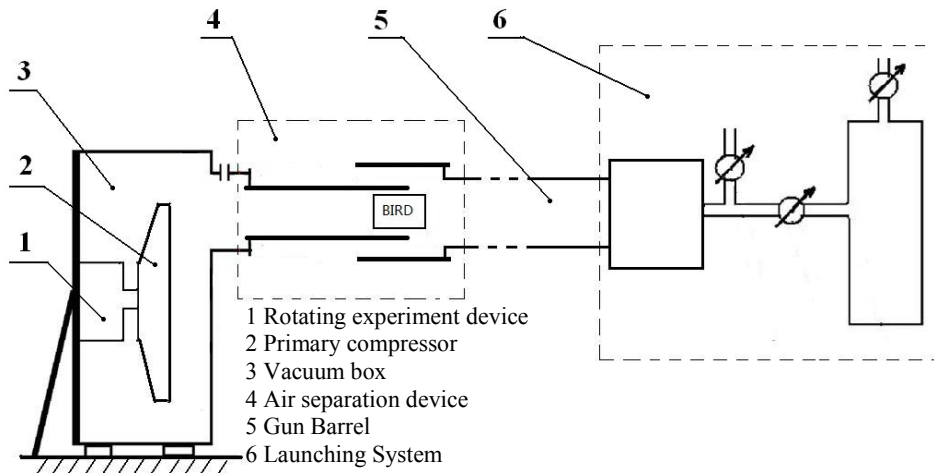
Rich experience has been obtained from the one or several engine blades tests subject to bird impacting. But simulations and experiments of bird strikes on the full scale of compressor rotating at high speed are rarely carried out. In the present paper, a 3D finite element compressor with a failure model is developed using commercial explicit code PAM-CRASH. The bird is modeled with smooth particles hydrodynamic (SPH) method and Murnaghan equation of state (EOS) is employed to describe the constitutive model. Simulations of the compressor rotating at different speeds (3300r/min and 8800r/min respectively) impacted by a 250g bird at velocity of 102 m/s were performed to compare with the experimental results and analyze the influence of parameters used in the simulations. The flexible fan blades adopted in the present paper are typical metallic aero engine fan blades.

### Test of bird impacting on primary compressor

#### *Test methods and apparatus*

The intention of the test is to obtain the dynamic responses of a primary compressor structure under a bird strike. The test results is used to compare with the numerical results presented in Section 3.

Fig.3 illustrates the arrangement of experimental apparatus during the test. The engine test-bed with high-speed rotation is placed in a vacuum box. The launching system accelerates the bird into the vacuum box through an air separation device and the bird will crash on the high-speed rotating blades.

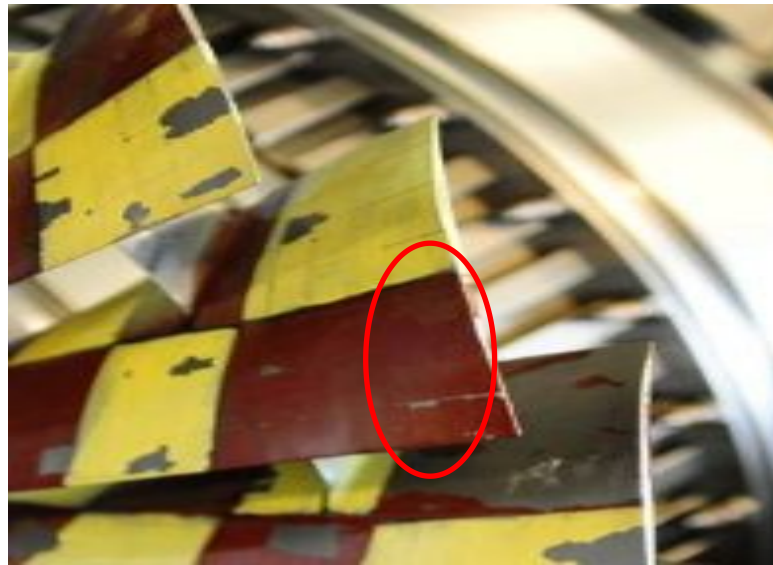


**Figure 3. Experiment rig of bird impact on a rotating compressor**

The geometry shape of the bird in present study is straight-ended cylinder. The ratio of length to diameter is 2 to 1. Gelatin is used to characterize the bird body due to its similarity with the real bird properties. The mass of the bird is 280g. The blades of engine is made of Ti6Al4V and the geometrical details of the primary compressor are showed in the following section.

#### *Test results*

Fig.4 displays the experimental results of a gelatin bird body impacting on a primary compressor in low rotating speed. The bird body is 280g and the velocity of the bird is 110m/s. The rotating speed of the blades ranges from 2900 to 3100r/min. The blades are visually inspected after the test. Slight plastic deformation of 3mm is found in one blade.



**Figure 4. The deformation of blades after impact under low rotating speed**

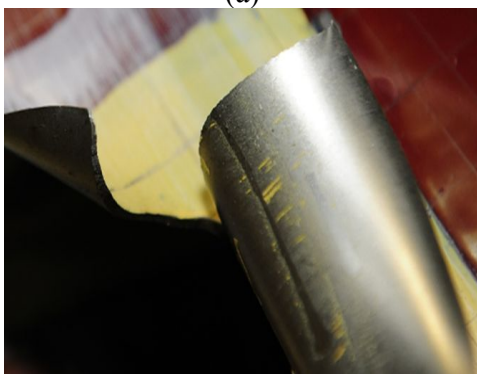
Fig. 5 shows the experimental results of bird body impacting on a primary compressor at high rotating speed. The actual mass and velocity of the bird is 280g and 110m/s respectively. The rotating speed varies from 7000 to 9000r/min. Severe plastic deformation are found in 7 blades, see Fig. 5(a). Fig. 5(b) shows a 40mm by 40mm block missing at a leading edge. A serious laceration is shown in Fig. 5(c). A shroud has a small dislocation accompanied by missing a small piece of block, which is shown in Fig. 5(d).



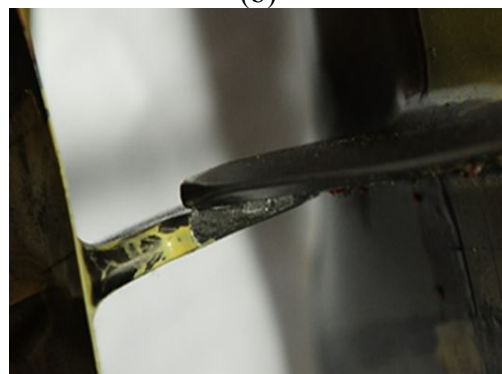
(a)



(b)



(c)



(d)

**Figure 5. The deformation of blades after impact under high rotating speed**

## Numerical simulation

### *Bird modelling*

During the impact at high velocity, the bird crushes into particles and splashes to all directions, which could be modeled by Smoothed Particle Hydrodynamics (SPH) method with the Murnaghan EOS for Solids.

### *SPH method*

The SPH method is a grid-less Lagrange technique which allows for severe distortion. It is introduced by Lucy in the 1970s and first applied to solve hydrodynamic problems in astrophysical environment [9]. In terms of situation of high-velocity impact, the SPH method is very suitable for modeling problems associated with characterizing large displacements, strong discontinuities and complex interface geometries. This method has been implemented in PAM-CRASH [10].

A smooth particle is input like a 3-DOF (degree of freedom) solid element and defined by its center of mass, volume, part number, and domain of influence. It can be used with great advantage to model bulk materials with no cohesion (sand, liquid, gases) or in situations where perforation or mixing is expected. Note however that it can be much more time consuming for computational calculation than a classical solid element. Every smooth particle with its own shape function, which is similar to a finite element, is reconstructed at each iteration from its dynamic connectivity. Localization and information transmission from one particle to another are achieved through the notion of an interpolation distance called the smoothing length. Please see reference [11] for more detailed information.

### *Geometry and material model for bird*

In order to match with the geometry shape of the bird used in the experiment, the straight-ended cylinder is adopted in the following simulation.

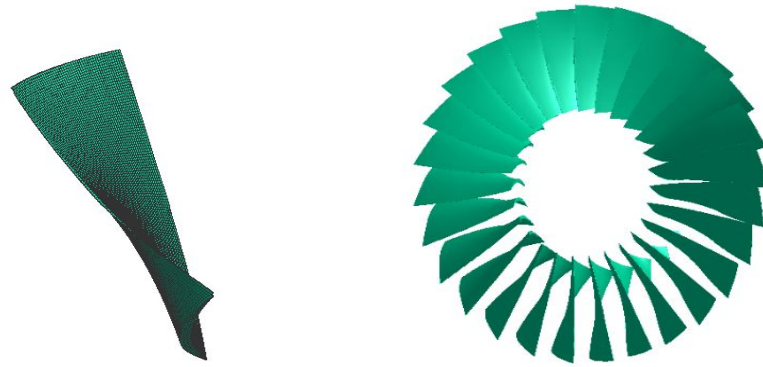
Five impact mechanics for solid materials is list by Wilbeck: elastic, plastic, hydrodynamic, sonic velocity and explosive. The hydrodynamic region shows to be suitable for characterizing the bird in a bird strike. In the case of a bird impacting, the yield stress of the bird is greatly exceeded due to its rapid deceleration and its inhomogeneity becomes increasingly negligible, so the bird can be considered as a homogeneous jet of fluid while impacting on a structure. The Murnaghan EOS implemented in PAM-CRASH (materials type 28) provides a powerful tool to simulate hydrodynamic behavior. The pressure for the Murnaghan EOS is given by:

$$p = p_0 + B \left[ \left( \frac{\rho}{\rho_0} \right)^\gamma - 1 \right] \quad (1)$$

where  $p_0$  is a reference pressure,  $p$  is the current pressure,  $\rho_0$  and  $\rho$  are initial and current mass density respectively, and  $B$  and  $\gamma$  are the materials constants to be determined.  $B$  equals to 128 MPa and  $\gamma$  equals to 7.9 in present paper.

### *Blade and rotor modelling*

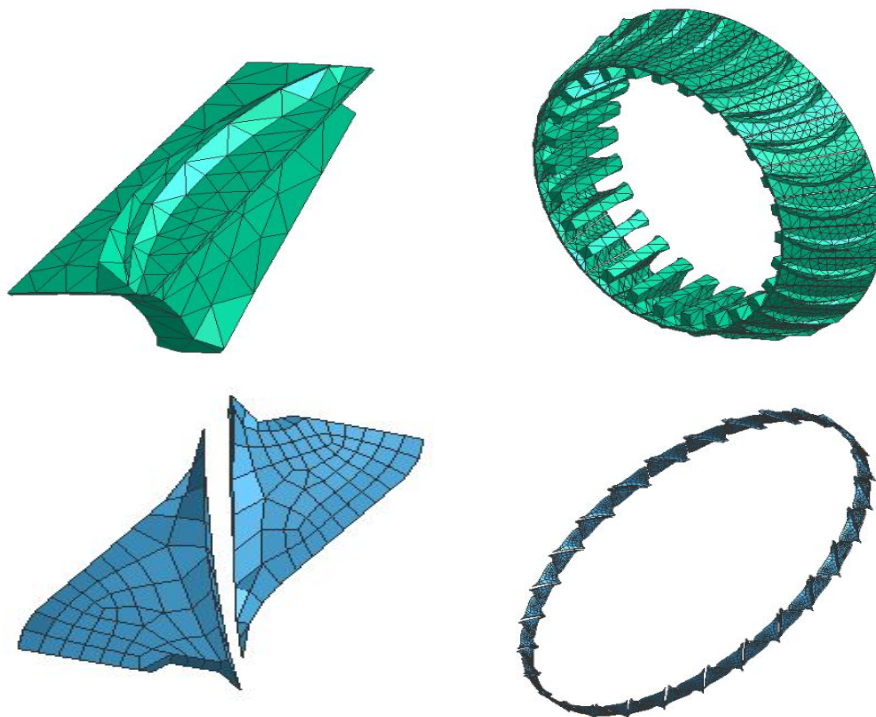
The blade disk assembly considered in this work comes from an aircraft jet engine. It consists of 27 equally spaced blades (Fig. 6) attached to the disk. The external radius is measured to the tip of the blade, the internal radius is measured to the root of the blade and the initial impact location is defined by the initial position of the bird's center of mass.



**Figure 6. Finite element model of the blades**

*FE model of the disk and shrouds*

The blades are meshed using Hypermesh software. Constant stress brick elements are used to mesh the blade and its attachment. Although plastic deformation due to in-plane and bending loads could be modelled by shell elements, the brick elements allows stress along the thickness and wave propagation to be taken into account, which is impossible using shell elements. A number of analyses are performed with different mesh densities, including increasing number of elements through the thickness. However, there is no considerable improvement of the results with the significant growth of the computational time. Hence, the blade is meshed with 21,400 elements: 107 elements along lengthwise, 50 elements width wise and 4 elements through the thickness, see Fig. 7. The disk is modelled as rigid body using tetrahedral elements and the shrouds are modelled using tetrahedral elements as well.



**Figure 7. Finite element model of the disk and shrouds**

*Constitutive model of material*

The disk is assigned with a rigid material, which is used to simulate a structure which is much stiffer than the regions of interest or experiences negligible deformations. In addition, the rigid material is computationally less expensive than other material models. The deformable blades are

made of titanium alloy Ti-6Al-4V and the constitutive model used for this material is elastic-plastic solid with damage and failure which contains elastic behavior, elastic-plastic behavior, strain rate behavior, and damage behavior. The elastic behavior is described by the shear modulus and bulk modulus as follows:

$$\text{Shear Modulus: } G = \frac{E}{2(1+\nu)} \quad (2)$$

$$\text{Bulk Modulus: } K = \frac{E}{3(1-2\nu)} \quad (3)$$

The Johnson-Cook law is applied to describe elastic-plastic behavior and strain rate behavior, formulated as:

$$\sigma(\varepsilon, \dot{\varepsilon}) = \left[ a + b(\varepsilon_p)^n \right] \left[ 1 + \frac{1}{p} \ln \left( \max \left( \frac{\dot{\varepsilon}}{D}, 1 \right) \right) \right] \quad (4)$$

where  $a + b(\varepsilon_p)^n$  denotes the basic elastic-plastic material curve;  $a$ ,  $b$  and  $n$  are material constants which could be obtained by the test results of split Hopkinson bar;  $\varepsilon_p$  indicates the plastic strain; and  $p$  and  $D$  are the control parameters of strain rate which could be obtained by the stress-plastic strain curve results at different strain rates.

For damage and failure behavior, the isotropic damage law acts on the total element stresses in the formula:

$$\sigma = (1 - d(\varepsilon_p)) \sigma_0 \quad (5)$$

where  $\sigma$  is damaged full tensor;  $\varepsilon_p$  is the isotropic function;  $\sigma_0$  is the full stress tensor as calculated from the undamaged elastic-plastic material law. An element is failure as soon as one of the equivalent plastic strain values at its integration points reaches the specified limit of failure.

The blade material properties of titanium alloy Ti-6AL-4V are given in Table 1.

**Table 1. Johnson Cook material properties for Tie6Ale4V**

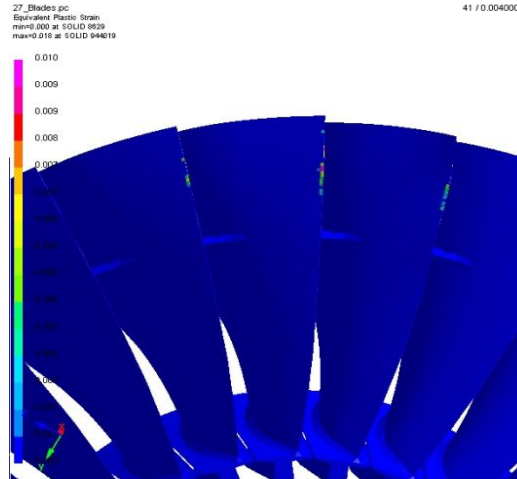
Parameter	Value
Density	$4.4 \times 10^3 \text{ kg/m}^3$
Shear modulus	43GPa
Bulk modulus	110Gpa
Initial yield stress	1070MPa
Strain hardening modulus	850MPa
Strain hardening exponent	0.6
Strain rate dependence coefficient	0.01
Maximum plastic strain for element elimination	0.2

#### *Calculation model*

After establishing the fan FE model and the bird SPH model, the contact type 34 in software is applied to define the contact between the FE model and SPH model as well as between the two tips on adjacent blades. The rotation and displacement of the fan in y and z directions is fixed, and the displacement of the fan in x direction is fixed. Just rotation of the fan around x is free. The rotation speed 3300 r/min and 8880r/min are defined in the present simulation. At last, the bird velocity along negative x direction is set at 112 m/s and 114m/s respectively according to the experiment.

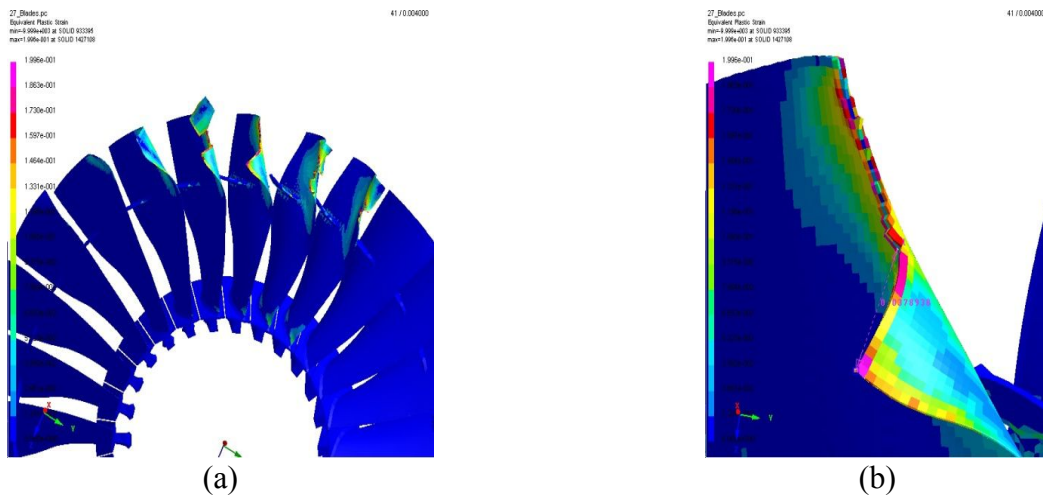
*Simulation results*

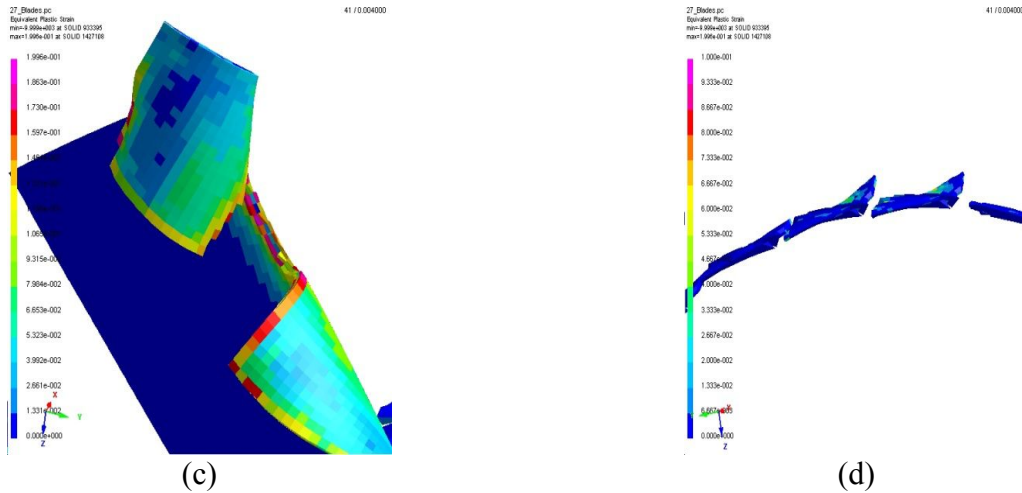
Fig. 8 illustrates the simulation results of bird impacting on the rotating blades of 3000r/min. As shown in the figure, slight plastic deformation take place on three blades' leading edges indicating that the blades are not severely damaged under low rotating speed situation, which shows a good agreement with the experimental results.



**Figure 8. The deformation of blades after impact under low rotating speed**

Fig. 9(a)-(d) are the simulation results of bird striking on high speed rotating blades. Fig. 9(a) shows that 6 blades yield large plastic deformation with 4 of them damaged. Fig. 9(b) shows that the leading edge of a blade is damaged severely with a block of 37mm by 58 mm missing. Fig .9(c) shows a serious laceration in another blade's leading edge. Fig. 9(d) shows dislocation occurred on two shrouds.





**Figure 9. The deformation of blades after impact under high rotating speed**

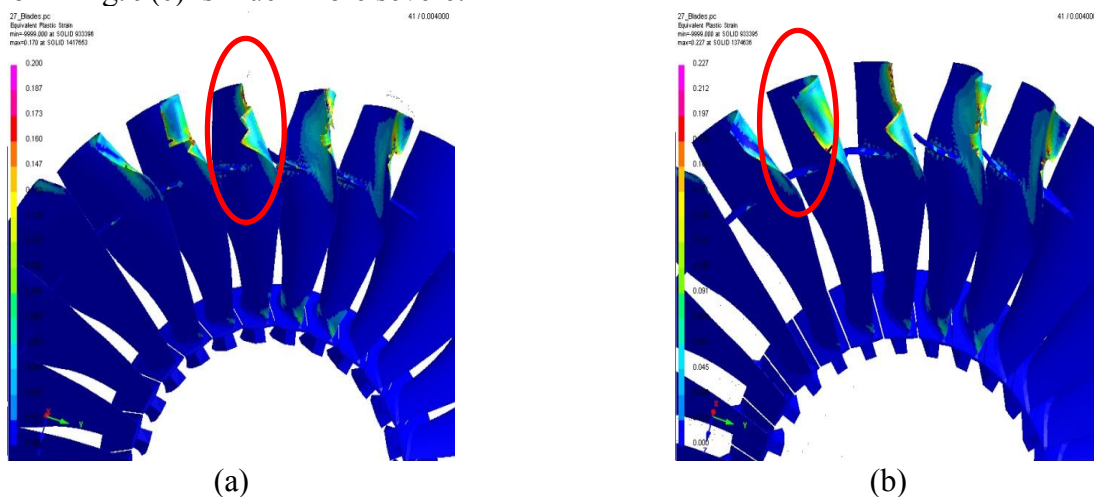
The good agreement of the simulation results and the experimental results suggests that the FE model with Murnaghan EOS for solid is quite suitable for soft body strike situation. This model can predicts the transient response of the blades under bird impact very well.

**Parametric study**

In order to study the influence of the parameters in the model intensively and analyze the stability, a number of parametric studies are performed to assess the influence of different impact conditions on the blade response.

*Influence of the failure stain*

The element elimination is determined by the failure stain. The maximum plastic strain for element elimination in the original model is 0.2. But in the actual situation, the value varies in a certain range. Fig. 10(a) and (b) gives the impacted results of model with different failure stain (0.17 and 0.23 respectively). Fig. 10(a) shows that the missing block area becomes larger than that in Fig. 9(a). Fig. 10(b) shows although the blade deforms a lot, the laceration of the blade is slight, while i the laceration n Fig. 9(c) is much more severe.



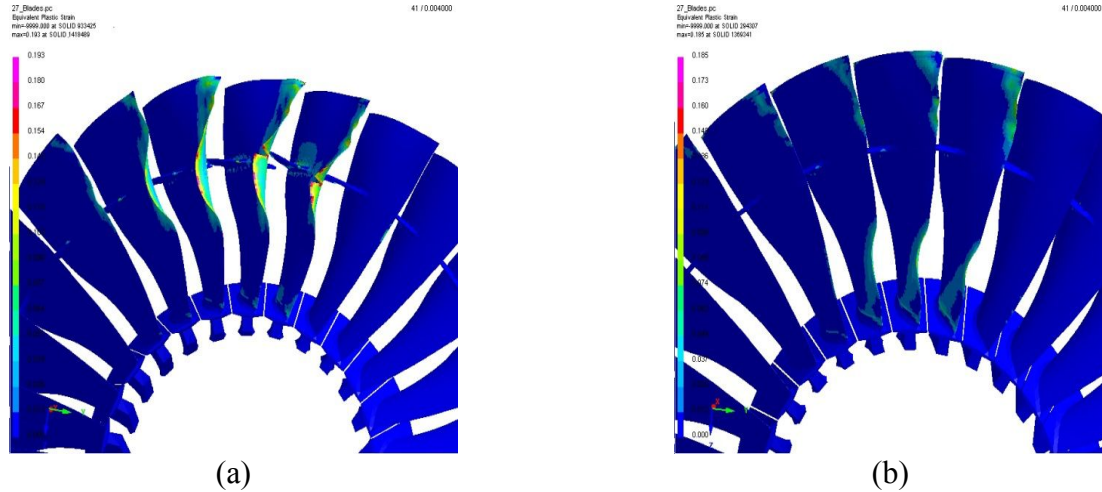
**Figure 10. Comparison of models with different failure stain**

*Bird impact location*

In the actual situation, the bird may strike on almost anywhere of the compressor. To find the most vulnerable striking position is necessary. Fig. 11(a) and (b) show the deformation of blades of



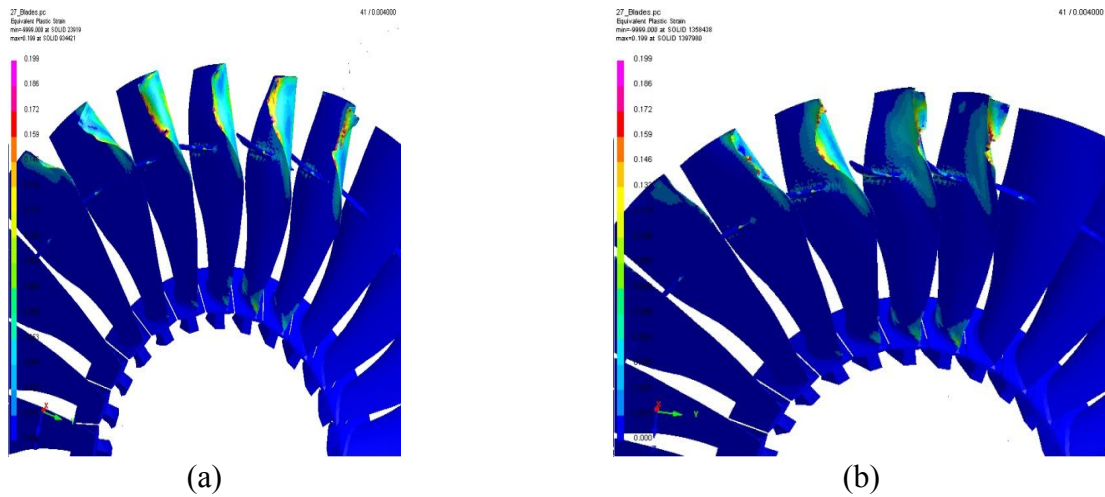
different impact locations (on middle and root of the blades). As shown in Fig. 11(a), 4 blades have large deformation and two of them have slight lacerations, while only a small extent of plastic deformation occurs in Fig. 11(b).



**Figure 11. Comparison of models with different impact location**

*Impact timing*

In the case of high rotating speed, the amount of bird sliced off by the blade in front of the leading blade is random. In this assessment impacts with different levels of interaction between the bird and the leading blade are considered. The impact timings considered are defined by varying initial location of the bird along the x axis and the amount of the bird which is removed from the analysis ( from 5mm to 20mm). The distribution of plastic deformation on the blades varies a lot.



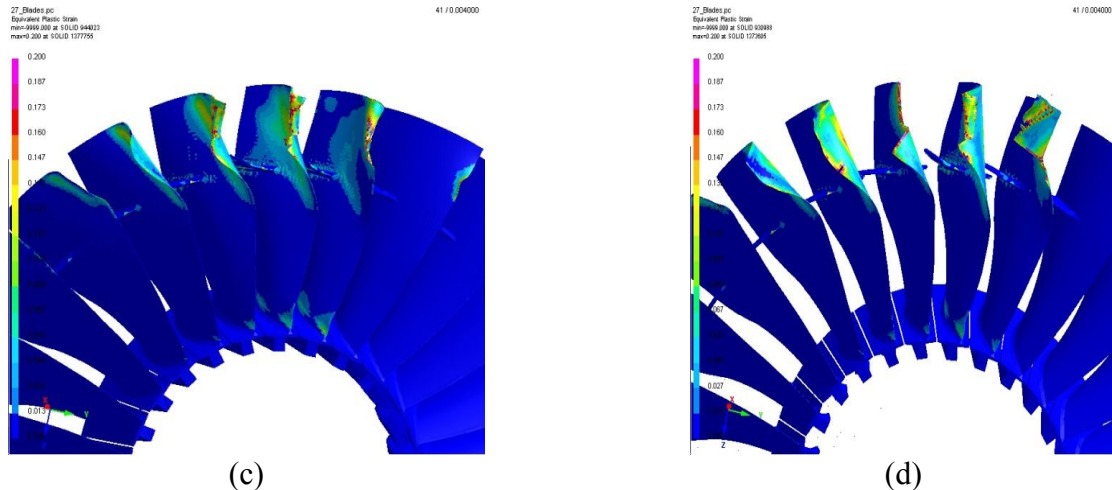


Figure 12. Comparison of models with different impact timing

## Conclusions

The FE model accompanied with SPH method provides accurate predicative numerical tools for simulation. The SPH method can attain high accuracy in process the problems with large displacements and deformation. The numerical results are compared with the blade recovered from the physical experiment, indicating a good reliability of the numerical simulation method on this issue. The simulation indicates that the failure stain in the model somehow influences the result while there are strong dependences for the results on the bird impact timing and deformation location on the blade. This study shows the potential of the finite element method in predicting the deformation of engine primary compressor in bird strike events.

## References

- [1] Sharing the skies manual. TP13549. Transport Canada; 2004.
- [2] Wildlife Strikes to Civil Aircraft in the United States 1990-2013. U.S. Department of Transportation Federal Aviation Administration; 2014.
- [3] Completeness and accuracy of birdstrike reporting in the UK. CAA Paper 2006/05. UK: Civil Aviation Authority; 2006
- [4] R. H. Mao, S. A. Meguid, and T. Y. Ng. Finite Element Modeling of a Bird Striking an Engine Fan Blade. Vol.44, No.2, Journal of Aircraft; 2007.
- [5] S. A. Meguid, R. H. Mao, T. Y. Ng. FE analysis of geometry effects of an artificial bird striking an aeroengine fan blade. International Journal of Impact Engineering, 35, 487-498; 2008
- [6] Guan Yupu, Zhao Zhenghua, Chen wei, Gao Deping. Foreign Object Damage to Fan Rotor Blades of Aeroengine Prat I: Experimental Study of Bird Impact. Chinese Journal of Aeronautics, 20, 408-414; 2007
- [7] Guan Yupu, Zhao Zhenghua, Chen wei, Gao Deping. Foreign Object Damage to Fan Rotor Blades of Aeroengine Prat II: Numerical Simulation of Bird Impact. Chinese Journal of Aeronautics, 20, 408-414; 2007
- [8] Rade Vignjevic, Michal Orłowski, Tom De Vuyst, James C. Campbell. A parametric study of bird strike on engine blades. International Journal of Impact Engineering, 60, 44-57; 2013
- [9] Espan˜ol P, Revenga M. Smoothed dissipative particle dynamics. Phys Rev2003; 67(2):12–25.
- [10] PAM-CRASH-2006 solver reference manual. Paris: ESI Corporation; 2006.
- [11] Vignjevic R, Reveles J, Lukyanof A. Analysis of compressor blade behaviour under bird impact. In: International conference on computational methods for coupled problems in science and engineering; 2005. p. 1–14

## Crashworthiness simulation of multi-cell tubes under oblique impact loads

\*Jianguang Fang, Guangyong Sun, †Qing Li

School of Aerospace, Mechanical and Mechatronic Engineering,  
The University of Sydney, Sydney, NSW 2006, Australia

\*Presenting author: [fangjg87@gmail.com](mailto:fangjg87@gmail.com)

†Corresponding author: [qing.li@sydney.edu.au](mailto:qing.li@sydney.edu.au)

### Abstract

Multi-cell tubes have been drawn increasing attention for their excellent energy-absorbing ability. However, the effect of cell number and oblique loads on crashing behaviors is seldom studied to date. In this paper, a group of multi-cell tubes with different cell numbers are comprehensively investigated under both axial and oblique loads. The finite element models are first established and then validated by experimental tests. The simulation results show that the increase in cell number can be beneficial to the energy absorption ( $EA$ ) but detrimental due to increase in peak force ( $F_{\max}$ ) under axial load. When the oblique loads are taken into account, the tubes could undergo global bending, which is an inefficient deformation mode.

**Keywords:** Multi-cell tube; Oblique impact; Crashworthiness

### 1 Introduction

In automotive engineering, the bumper system requires to endure a load with a  $30^\circ$  loading angle to the longitudinal axis [1]. Under this condition, thin-walled structures always undergo a combined deformation of bending and axial crushing, rather than pure axial collapse that is designed as an efficient deformation mode for energy absorption. Due to the presence of global bending, the energy absorption of thin-walled structures can be reduced dramatically. For this reason, oblique loading has to be considered in crashworthiness design for thin-walled structures. For this purpose, Han and Park [2] investigated the crush behavior of a square column subjected to oblique loads and found that the loading angle has considerable effect on deformation pattern from the axial buckling to global bending. Reyes et al. [1, 3, 4] studied the crashworthiness performance of hollow and foam-filled aluminum tubes under quasi-static oblique loads by using experimental and numerical methods. Zarei and Kröger [5] performed axial and oblique impact tests on empty and aluminum honeycomb filled square tubes and observed a dramatic decrease in first peak load during the oblique impact test. Tarlochan et al. [6] proposed a design process for a thin-walled structure subjected to both axial and oblique loads and found that the hexagonal profile was a better choice for energy absorption application. Yang and Qi [7] developed an optimization procedure for design of the empty and foam-filled square columns under oblique loading and the results showed that compared with the foam-filled tubes, the empty column can behave better under oblique impact but worse under pure axial loading. To be the author's best knowledge, there have been very limited studies of multi-cell tubes to take into account the oblique loads so far. Qi et al. [8] employed LS-DYNA to predict the crashing behaviors of four tubes with different cellular configurations under oblique loads, and they found that multi-cell tapered tube has the best crashworthiness performance. Song and Guo [9] found that the

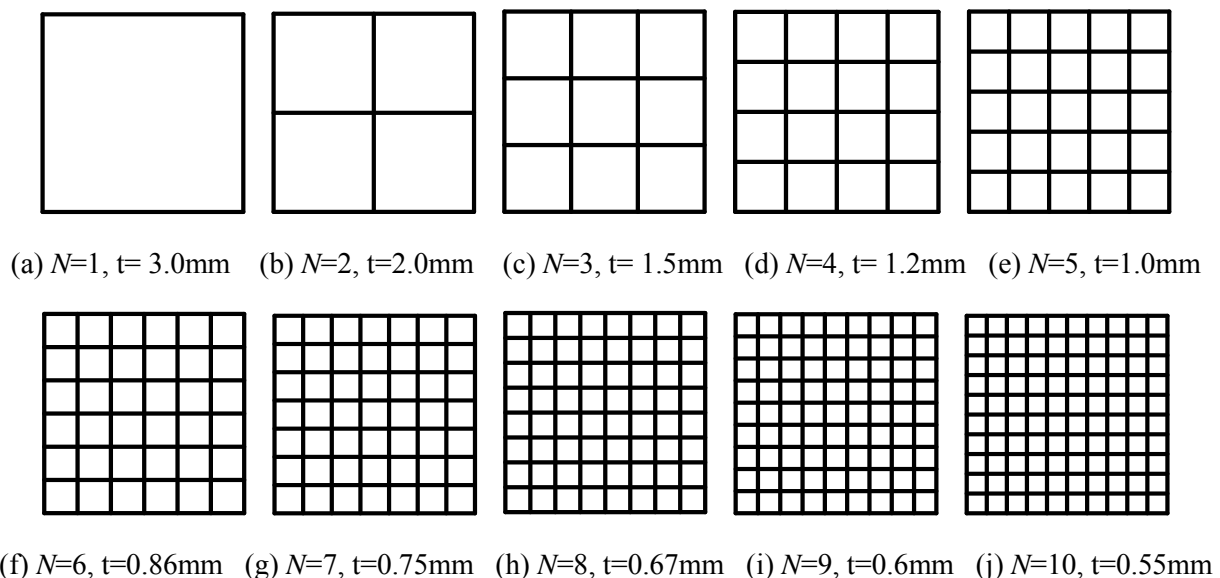
effectiveness of multi-cell configuration for improving tubes' energy absorption reduces as the load angle increases, and multi-cell tubes can even have worse performance if they collapse in global bending mode.

From the abovementioned studies, it is known that comprehensive comparison of the crashworthiness of multi-cell structures with different cell numbers remains under-studied thus far. Furthermore, how does the cell number affect the crashworthiness under oblique loads? This study aims to address these two issues.

## 2 Numerical modeling

### 2.1 Finite element (FE) model

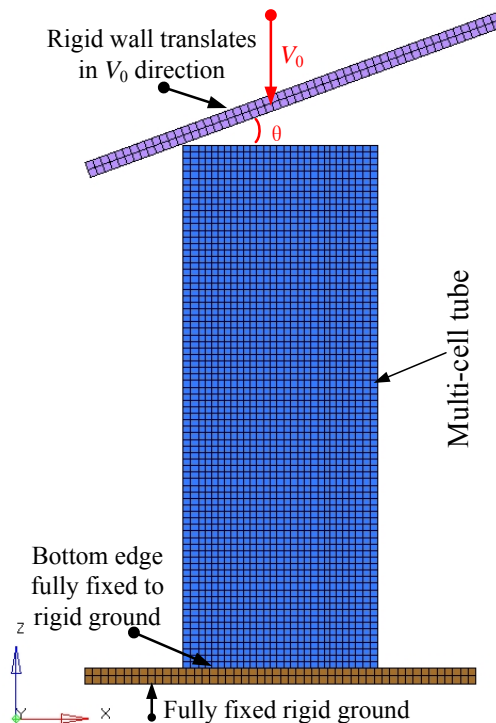
The structures to be analyzed here comprise a group of thin-wall multi-cell square tubes with the same axial length of  $L=200$  mm and the same width of  $b=75$ mm subjected to oblique impact loading (Fig. 1). These square tubes have different cell numbers (i.e.,  $N \times N=1 \times 1, 2 \times 2, 3 \times 3, \dots, 10 \times 10$ ), and are assigned different thicknesses in order to investigate these tubes with the same mass (Fig. 1). As shown in Fig. 2, a rigid wall with a mass of 600 kg and an incident angle  $\theta$  impacts on the top end of the tubes at an initial velocity of  $v=15$  m/s. the bottom ends of the tube is attached to the fixed rigid ground.



**Fig. 1** Cross-sections of multi-cell tubes with the same mass but different wall thicknesses.

The numerical models are developed using explicit non-linear finite element code LS-DYNA. The Belytschko-Lin-Tsay reduced integration shell elements with five integration points through the thickness were employed to model the tubes. Stiffness-based hourglass control was employed to avoid spurious zero energy deformation modes and reduced integration was used to avoid volumetric locking. The interfaces between the tube and rigid wall and between the tube and rigid ground were both modeled as an “automatic node to surface”. “Automatic single surface” contact was also prescribed to the tube to avoid interpenetration during tube folding. For both

static and dynamic friction, the friction coefficient of 0.2 was adopted for all contact conditions [10-12].



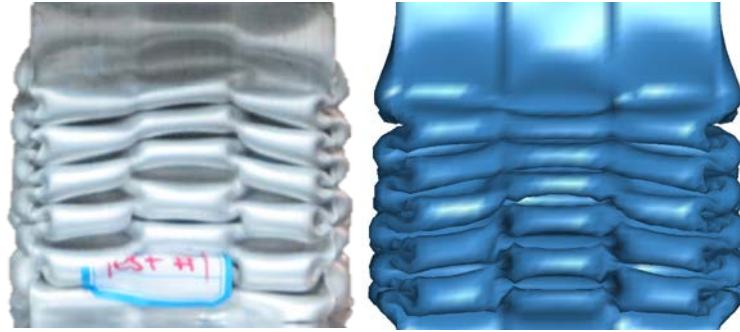
**Fig. 2** Finite element model of multi-cell tube subjected to oblique loading.

The tube was modeled through a piecewise linear elastic-plastic behavior with strain hardening (material model 24 in LS-DYNA). The thin wall material was aluminum alloy AA6063-O with the density  $=2700 \text{ kg/m}^3$ , Poisson's ratio  $=0.3$ , and Young's modulus  $=70 \text{ GPa}$ . The material model was considered insensitive to strain rate but defined as non-linear isotropic work hardening in the plastic region [13, 14].

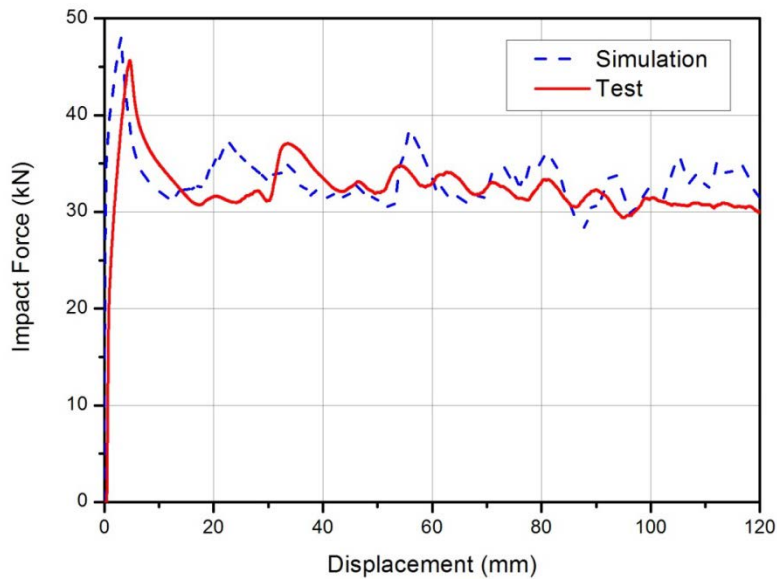
## 2.2 Validation of FE modeling

In order to validate the FE modeling, a tube with  $3 \times 3$  uniform cells was compressed using the available test facilities with a constant axial impact velocity of  $3 \text{ mm/min}$ . The simulation model with the same specimen configuration was established under the same testing conditions. Fig. 3(a) depicts a comparison of the deformed shapes between the FE simulation and experimental tests. The overall profile of FE model in terms of the collapse mode and number of folds is in good agreement with the experiment. Fig. 3(b) illustrates a comparison of the FE and experimental results of the load-displacement responses, in which a fairly good overall agreement between them is observed. From which, it is easily found that the maximum discrepancy of initial peak force is less than 5% and the FE prediction exhibits a slightly higher peak load than the experiment during the subsequent folding. This could be attributed to that the existence of imperfections in the physical specimen, which is difficult to be considered precisely in FE

modeling. Overall, the numerical modeling technique adopted here is considered sufficiently accurate and will be extended to the subsequent study.



(a) Comparison of deformation profile between experimental test (left) and numerical model (right)



(b) Force – deformation curves from experimental test and FE modeling

**Fig. 3** Comparison of experimental and simulation results.

### 3 Crashworthiness of multi-cell tubes under different load angles

To systematically measure the crashworthiness of different structures, many different indicators have been proposed [15, 16], of which the energy absorption ( $EA$ ) denotes the absorbed energy via plastic deformation of structures as

$$EA = \int_0^{\delta} F(\mathbf{x}) dx \quad (10)$$

where  $F(\mathbf{x})$  is the instantaneous crushing force in axial direction.  $\delta$  denotes the crash displacement, which is taken as  $0.6L$  in this study, and  $L= 200\text{mm}$  is the tube length.

In addition to energy absorption, crushing forces designate other key crushing criteria, of which peak crushing force ( $F_{\max}$ ) represents the maximum value of  $F(\mathbf{x})$  and the mean crushing force (MCF) measures level of average crushing force given as

$$F_{avg} = \frac{EA}{\delta} = \frac{\int_0^{\delta} F(\mathbf{x})dx}{\delta} \quad (11)$$

Besides, crushing force efficiency ( $CFE$ ) is also used as an indicator given as:

$$CFE = \frac{F_{avg}}{F_{\max}} \times 100\% \quad (12)$$

Note that a large  $F_{\max}$  often leads to a high deceleration and high risk of severe injury or even death of occupant when impact occurs. Besides, the higher the  $CFE$  value, the better the load uniformity for an energy absorber. In this paper, these abovementioned indicators will be considered under oblique and axial loads.

### 3.2 Results of pure axial crushing

Fig. 4 displays the deformations of multi-cell tubes with the different cell numbers of  $1 \times 1$ ,  $4 \times 4$ ,  $7 \times 7$  and  $10 \times 10$  (i.e.,  $N=1, 4, 7$  and  $10$ ) under pure axial loads respectively, and Fig. 5a depicts the corresponding crushing force curves. For single cell structure ( $N=1$ ), it can be seen that the crushing force first reaches an initial peak, followed by a dramatic drop and then fluctuates at some low values that correspond to the progressive folding. When cell number  $N$  increases, the crush force does not increase noticeably in the initial peak and tends to be stable during the whole crushing displacement. Specifically, the difference between the initial peak and the following peaks becomes small, and the force tends to fluctuate within a narrow range when cell number  $N$  increases. This might be attributed to that the fold length (the distance between two adjacent plastic hinges) decreases and thus the fold number increases with the increase in  $N$  (as shown in Fig. 4).

From Fig. 6(a) it can be seen that under the same mass the  $EA$  increases as the cell number increases, but the increased value become smaller and smaller, which represent that the energy absorption becomes stable gradually. From Fig. 6(b), the value of  $F_{\max}$  increases fairly slowly when the cell number increases, indicating that  $F_{\max}$  is insensitive to the cell number under the pure axial load condition. As for the  $CFE$ , its value increases when  $N < 4$  and then oscillates when  $N > 4$ .

### 3.3 Results of oblique crushing

From Fig. 4, the deformation modes of tubes the oblique impact of  $\theta=10^\circ$  are fairly similar to those of  $\theta=0^\circ$ . That is to say that when the tubes are subjected to the oblique impact with a small incident angle, they are still able to generate proper progressive folding in the axial direction, which is typically regarded as an efficient deformation pattern. Interestingly, the impact forces fluctuate within a fairly small range after the initial peaks when  $\theta$  increases from  $0$  to  $10^\circ$  (Fig. 5(b)). Similarly to the situations of  $\theta=0^\circ$ , the more the cells of the tube has, the larger the force

bears during the impact. As a result, it is observed from Figs. 8(a) and 8(b) that the increase in cell number  $N$  leads to the increase of  $F_{\max}$  and  $EA$ . Besides, the  $F_{\max}$  and  $EA$  of all tubes both decrease when  $\theta$  increases from  $0^\circ$  to  $10^\circ$ ; and the  $F_{\max}$  decreases more than  $EA$  does, which well explains why the  $CFE$  of  $\theta=10^\circ$  is larger than that of  $\theta=0^\circ$  for each tube (Fig. 6(c)), since the  $EA$  is equivalent to  $F_{\text{avg}}$  (see Eq.(11)).

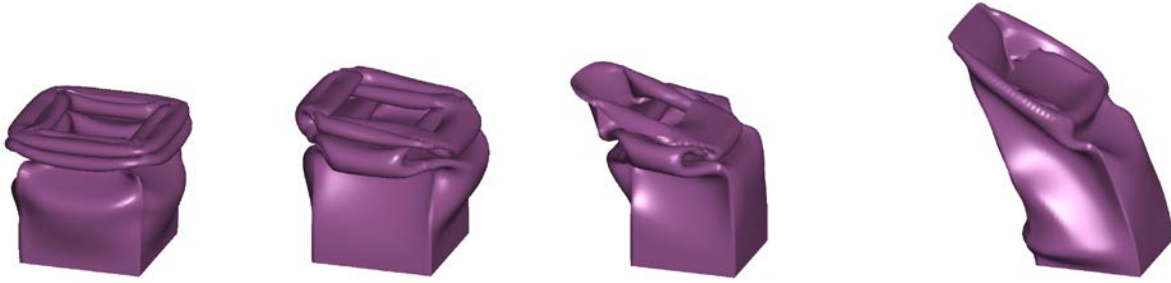
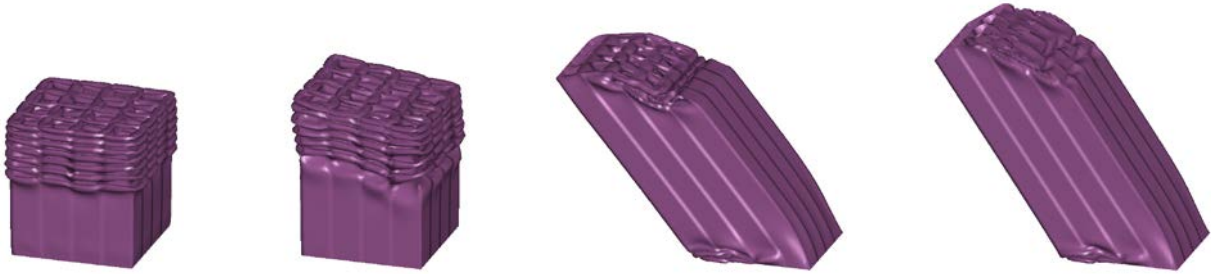
When  $\theta$  increases to  $20^\circ$ , the global bending mode is observed for the tubes with more cells ( $N > 2$ ). The compression flange buckles and develops a plastic hinge close to the clamped distal end (in  $N=4, 7, 10$  in Fig. 4). Then the whole tube bends around this hinge as almost an rigid body motion. On the other hand, the tubes with a smaller cell number (e.g.  $N=1$  and  $2$ ) still succeed to avoid such a global bending. From Fig. 5(c), the tubes of  $N=4, 7$  and  $10$  reach a peak load and then decrease dramatically due to the global bending, weakening the energy absorption and load-carrying capacity. As a result, Figs. 8a and 8b exhibit that the  $EA$  and  $F_{\max}$  of  $N=4, 7$  and  $10$  reduce much more when  $\theta$  increases from  $10^\circ$  to  $20^\circ$ , compared with those of  $N=1$  and  $2$ . When  $\theta$  increases up to  $30^\circ$ , global buckling occurs in all the tubes (Fig. 4) and the  $EA$  and  $F_{\max}$  become even smaller (as Figs. 8(a) and 8(b)). Interestingly, from Fig. 5(d), the tube of  $N=1$  undergoes a longer deformation before the impact force decreases significantly, representing that it bends later than the other three.

Overall, when multi-cell tubes are impacted obliquely with a small loading angle  $\theta$ , all of them collapse progressively in the axial direction, which is very similar to the situation in pure axial loading. When  $\theta$  increases, multi-cell tubes begin to bend globally, which leads to reduction of load-carrying and energy absorption. It can also be seen that the tubes with more cells are more sensitive to oblique loads, although they succeed to absorb more energy in pure axial case and oblique case with small angle  $\theta$ .

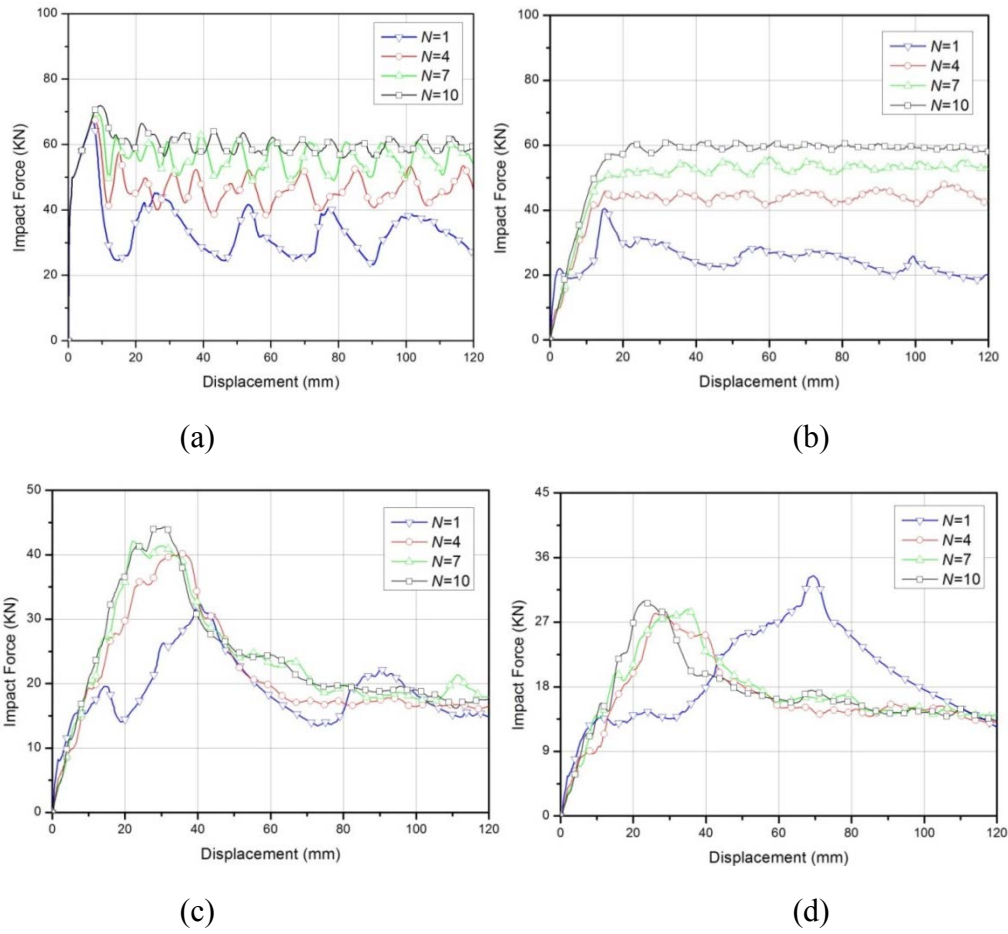
#### 4 Conclusions

In this study, the crashing behaviors of multi-cell tubes with ten different cell numbers have been comprehensively investigated under both axial and oblique loads. The simulation models were established using non-linear finite element code LS-DYNA and then validated by experimental test. It was found that both energy absorption ( $EA$ ) and peak force ( $F_{\max}$ ) of multi-cell tubes increase with the increase in cell number under axial loads. Regarding the oblique loads, the tubes can remain to collapse progressively in axial direction when the incident angle ( $\theta$ ) is small, but they develop global bending when  $\theta$  increases up to  $20^\circ$  and  $30^\circ$ . It was also observed that the tubes with fewer cells are more robust to oblique loads with large angle but perform worse under pure axial load than the tubes with more cells.

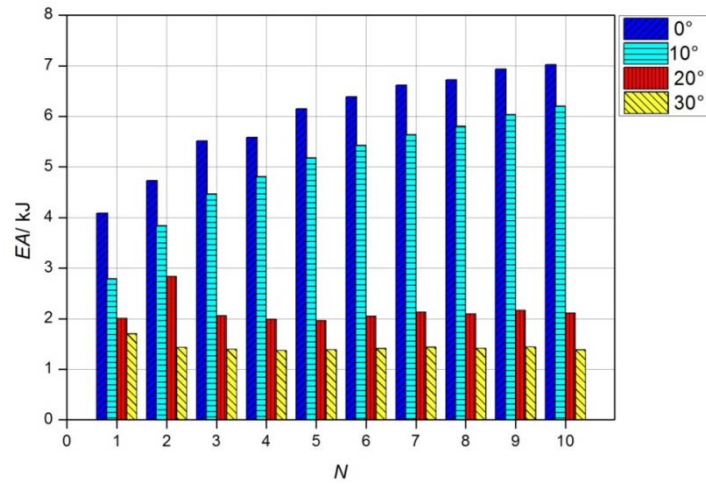


$\theta = 0^\circ$  $\theta = 10^\circ$  $\theta = 20^\circ$  ICAM2015, 14-17 July, 2015, Auckland, New Zealand $N = 1$  $N = 4$  $N = 7$  $N = 10$ 

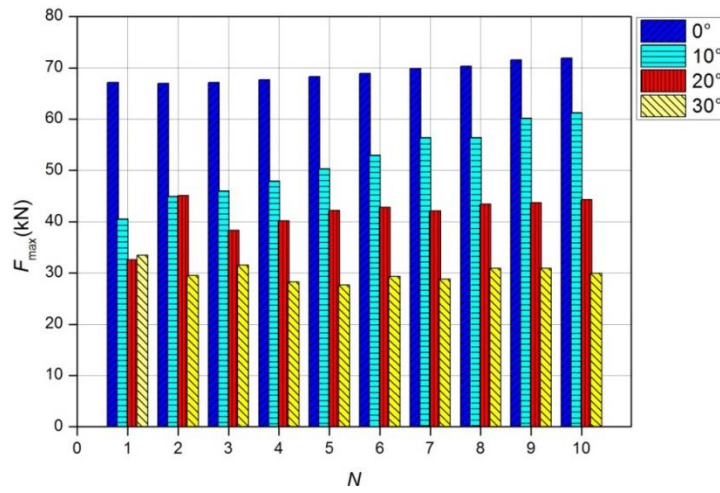
**Fig. 4** Deformation modes of multi-cell tubes with different cell number  $N$



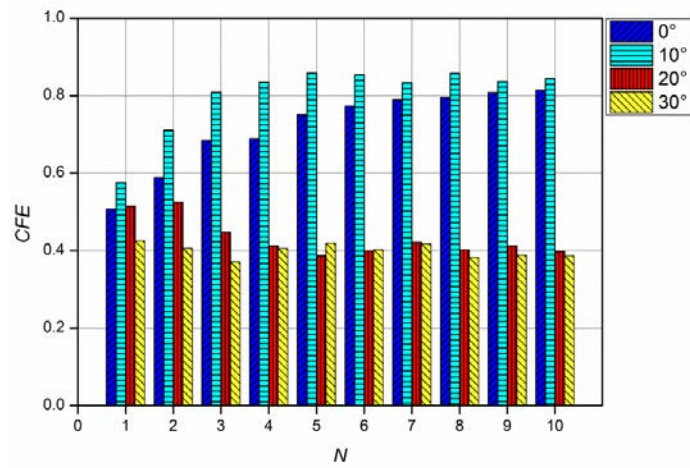
**Fig. 5** Crashing force vs. displacement curve: (a)  $\theta = 0^\circ$ , (b)  $\theta = 10^\circ$ , (c)  $\theta = 20^\circ$ , (d)  $\theta = 30^\circ$ .



(a)



(b)



(c)

**Fig. 6** Performance comparisons under different loading angles: (a)  $EA$ , (b)  $F_{max}$ , (c)  $CFE$ .

## References

- [1] A. Reyes, M. Langseth, O. Hopperstad, Crashworthiness of aluminum extrusions subjected to oblique loading: experiments and numerical analyses, *Int J Mech Sci*, 44 (2002) 1965-1984.
- [2] D. Han, S. Park, Collapse behavior of square thin-walled columns subjected to oblique loads, *Thin Wall Struct*, 35 (1999) 167-184.
- [3] A. Reyes, M. Langseth, O. Hopperstad, Square aluminum tubes subjected to oblique loading, *Int J Impact Eng*, 28 (2003) 1077-1106.
- [4] A. Reyes, O. Hopperstad, M. Langseth, Aluminum foam-filled extrusions subjected to oblique loading: experimental and numerical study, *Int J Solids Struct*, 41 (2004) 1645-1675.
- [5] H. Zarei, M. Kröger, Optimum honeycomb filled crash absorber design, *Mater Design*, 29 (2008) 193-204.
- [6] F. Tarlochan, F. Samer, A.M.S. Hamouda, S. Ramesh, K. Khalid, Design of thin wall structures for energy absorption applications: Enhancement of crashworthiness due to axial and oblique impact forces, *Thin Wall Struct*, 71 (2013) 7-17.
- [7] S. Yang, C. Qi, Multiobjective optimization for empty and foam-filled square columns under oblique impact loading, *Int J Impact Eng*, 54 (2013) 177-191.
- [8] C. Qi, S. Yang, F.L. Dong, Crushing analysis and multiobjective crashworthiness optimization of tapered square tubes under oblique impact loading, *Thin Wall Struct*, 59 (2012) 103-119.
- [9] J. Song, F.L. Guo, A comparative study on the windowed and multi-cell square tubes under axial and oblique loading, *Thin Wall Struct*, 66 (2013) 9-14.
- [10] H.F. Yin, G.L. Wen, Z.B. Liu, Q.X. Qing, Crashworthiness optimization design for foam-filled multi-cell thin-walled structures, *Thin Wall Struct*, 75 (2014) 8-17.
- [11] Z. Zhang, S. Liu, Z. Tang, Crashworthiness investigation of kagome honeycomb sandwich cylindrical column under axial crushing loads, *Thin Wall Struct*, 48 (2010) 9-18.
- [12] Z. Tang, S. Liu, Z. Zhang, Analysis of energy absorption characteristics of cylindrical multi-cell columns, *Thin Wall Struct*, 62 (2013) 75-84.
- [13] J. Fang, Y. Gao, G. Sun, Y. Zhang, Q. Li, Parametric analysis and multiobjective optimization for functionally graded foam-filled thin-wall tube under lateral impact, *Computational Materials Science*, 90 (2014) 265-275.
- [14] J. Fang, Y. Gao, G. Sun, Y. Zhang, Q. Li, Crashworthiness design of foam-filled bitubal structures with uncertainty, *International Journal of Non-Linear Mechanics*, 67 (2014) 120-132.
- [15] G. Sun, G. Li, S. Hou, S. Zhou, W. Li, Q. Li, Crashworthiness design for functionally graded foam-filled thin-walled structures, *Materials Science and Engineering: A*, 527 (2010) 1911-1919.
- [16] M.A. Guler, M.E. Cerit, B. Bayram, B. Gerceker, E. Karakaya, The effect of geometrical parameters on the energy absorption characteristics of thin-walled structures under axial impact loading, *International Journal of Crashworthiness*, 15 (2010) 377-390.

## Topology optimization of anisotropic constrained damping structures based on ESO method

L. Zhao, \*Q. Deng, S.Y.Liu, M.J.Su, and Y.H.Xu

School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, P.R. China

\*Corresponding author: dengqiong24@nwpu.edu.cn

### Abstract

The topology optimization design of structures with anisotropic constrained damping layers (ACDL) is investigated in present paper. The evolutionary structural optimization (ESO) method is employed to find the optimal layout of ACDL with aim to maximum the modal damping loss factor of the sandwich structure. Modal strain energy (MSE) method is used to calculate the damping loss factor of whole sandwich structure, and the sensitivity is analyzed. Optimization topology configuration of ACDL is obtained under the constraint of total amount of ACLD materials in terms of percentage added weight to the base structure. The merit of constrained damping treatment is evaluated quantitatively by damping efficiency formula introduced.

**Keywords:** Evolution Structural Optimization, Modal Strain Energy, Anisotropic, Topology optimization

### Introduction

Passive constrained damping layer (CDL) treatment is one of the common and effective means to reduce vibration and noise of thin-walled structures. However, aero structures require strictly for additional mass quality. On one hand, optimizing the layout of the CDL parts is needed to reduce the amount of material, and obtain a better suppression damping effect; on the other hand, we suppose to employ anisotropic constrained damping layer (ACDL), whose density is smaller and damping loss factor may not be neglected.

Lall<sup>[1,2]</sup> used analytical methods to study the structural characteristics of fragmentation damping structure for beams and plates respectively. Zheng<sup>[3]</sup> has studied the structural characteristics of fragmentation damping structure for cylindrical shell by same methods and explored the fragmentation arrangement of constraint damping layout optimization problem. From their work, we learn that better vibration control effect with less structural mass increment could be achieved by arranging damping structure appropriately without full coverage, which leads to the topology optimization problem of constrained damping structure.

Several topology optimization methods are commonly used, such as homogenization method (Bendsøe and Kikuchi<sup>[4]</sup>), variable density method (Bendsøe<sup>[5]</sup>, Sigmund<sup>[6]</sup>), evolutionary structural optimization method (Xie and Steven<sup>[7]</sup>), level set method (Wang et al.<sup>[8]</sup>, Allaire et al.<sup>[9]</sup>) and so on. Guo<sup>[10]</sup> optimized a square plate with CDL based on ESO method, and obtained a clear topological structure by remove the inefficient damping materials gradually; Zheng<sup>[11]</sup> used SIMP interpolation model and MMA solution algorithm to optimized the layout of constraint damping structure with a target to maximize the modal loss factor. The research of optimization for anisotropic constrained damping layers (ACDL) is rare.

This paper introduces damping efficiency formula to evaluate the merit of constrained damping treatment results quantitatively and then anisotropic material is employed as constraint sheet and the topology optimization problem of structures with ACDL is investigated based on ESO method. Finally, a numerical example is presented to show the superiority of ACDL in lightweight design concept.

### Damping efficiency

The damping loss factor of structure will increase after an additional constrained damping layer treatment. We always want to attach additional mass as little as possible to get a bigger damping loss factor. In other word, a larger ‘damping efficiency’ after damping treatment is expected. To compare the results of damping treatments quantitatively, formula (1) is introduced to calculate the damping efficiency:

$$\eta_{eff} = \frac{\Delta\eta}{\Delta m} \quad (1)$$

In this formula,  $\Delta\eta$  denotes the increment of the structure’s damping loss factor, and  $\Delta m$  denotes the mass of the damping layers, which is also additional mass.

### Topology optimization model

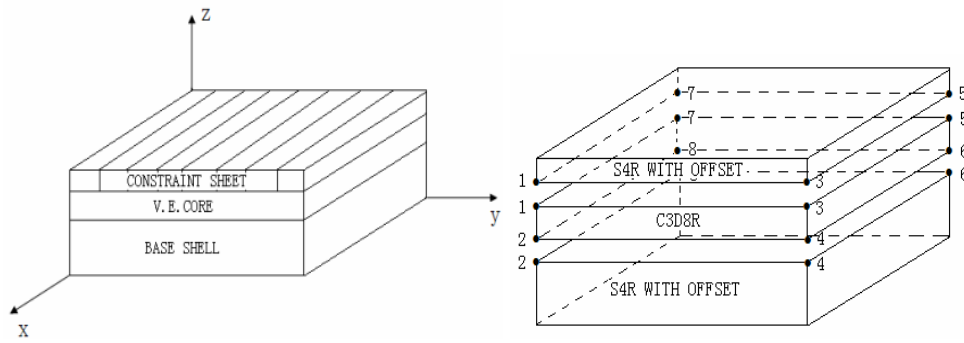
The topology optimization problem of given thickness of both constraint layer and damping core, aiming at the maximum of damping loss factor, under the constraint of additional mass quantity, using element existence state as design variables, is formulated as:

$$\begin{aligned} & \text{Max } \eta \\ & \text{find } x = [x_1, x_2, \dots, x_n] \\ & \text{s.t. } W = W_v + W_c \leq W^* \\ & \quad x_i = \{0, 1\} \end{aligned} \quad (2)$$

Where  $\eta$  is the damping loss factor of the sandwich structure;  $n$  denotes the number of viscoelastic damping core elements;  $x_i$  is design variable with 0 denoting element deleted and 1 denoting element reserved;  $W_v$  and  $W_c$  denote the mass quantity of the viscoelastic damping core and the constraint sheet respectively.  $W^*$  is the constraint mass quantity.

### Finite element model

In a thin-walled structure with anisotropic constraint damping layer, energy consumption comes from both damping layer and constraint layer. The damping layer dissipates energy by shear deformation and the constraint layer dissipates energy by tensile/compressive and bending deformation. Thus, a finite element model reference to Johnson’s<sup>[12]</sup> method is employed using the commercial program package ABAQUS, as shown in figure 1. The constraint sheet and bottom sheet are modeled with quadrilateral shell elements called S4R producing stiffness at two rotational and three translational degrees of freedom per node. The viscoelastic damping core is modeled with solid elements called C3D8R producing stiffness at three translational degrees of freedom per node. All nodes are at element corners. A key feature of this kind of element in the present application is its ability to account for coupling between stretching and bending deformations. This feature allows the plate nodes to be offset to one surface of the plate, coincident with the corner nodes of the adjoining solid elements. In this way, a three-layer plate can be modeled with only two layers of nodes, and finite element analysis will not take too much time.



**Figure 1. Finite element modeling of structure with ACDL**

### Sensitivity analysis of modal damping loss factor

According to modal strain energy (MSE) method, the  $k$ -th modal damping loss factor of the sandwich structure is calculated as follows:

$$\eta_k = \frac{\eta_v E_{vk} + \eta_c E_{ck} + \eta_b E_{bk}}{E_{tk}} \quad (3)$$

where  $\eta_v$ ,  $\eta_c$  and  $\eta_b$  are damping loss factors of damping material, constraint sheet material and bass structure material respectively;  $E_{vk}$ ,  $E_{ck}$ ,  $E_{bk}$  denote the  $k$ -th modal strain energy of damping core, constraint sheet and bass structure respectively, and  $E_{tk} = E_{vk} + E_{ck} + E_{bk}$ .

In a difference method way, when the damping core element  $i$  and the corresponding constraint layer element are deleted, small change of the whole sandwich structure's damping loss factor can be approximately expressed as:

$$\Delta\eta_k = \frac{\eta_v \Delta E_{vk} + \eta_c \Delta E_{ck}}{E_{tk}} - \frac{(\eta_v E_{vk} + \eta_c E_{ck} + \eta_b E_{bk}) \Delta E_{tk}}{E_{tk}^2} \quad (4)$$

In each iteration of the ESO method, the whole structure changes very little when an element  $i$  is deleted, so the following approximation can be adopted:

$$\begin{aligned} \Delta E_{vk} &\approx -E_{vki} \\ \Delta E_{ck} &\approx -E_{cki} \\ \Delta E_{tki} &\approx -E_{vki} - E_{cki} \end{aligned} \quad (5)$$

Where  $E$  always denotes model strain energy. From (3), (4) and (5),  $\Delta\eta_k$  can be calculated as

$$\Delta\eta_k = \frac{(\eta_v E_{vk} + \eta_c E_{ck} + \eta_b E_{bk})(E_{cki} + E_{vki})}{E_{tk}^2} - \frac{\eta_v E_{vki} + \eta_c E_{cki}}{E_{tk}} \quad (6)$$

Formula (6) denotes the change amount of the  $k$ -th modal damping loss factor after the damping core's element  $i$  and its corresponding constraint element are deleted. Thus formula (7) can be defined as the design sensitivity of element  $i$ .

$$S_{ki} = \frac{(\eta_v E_{vk} + \eta_c E_{ck} + \eta_b E_{bk})(E_{cki} + E_{vki})}{E_{tk}^2} - \frac{\eta_v E_{vki} + \eta_c E_{cki}}{E_{tk}} \quad (7)$$

If maximization of the former  $m$ -th model damping loss factors is required, the design sensitivity of element  $i$  is formulated as

$$S = \sum_{k=1}^m \omega_k S_{ki} \quad (8)$$

where  $\omega_k$  denotes the weight coefficients of each model.

### Mesh-independent filtering

Discrete mesh-independent filtering method is employed during the optimization process to avoid checkerboard phenomenon. Referring to Sigmund's work, the new sensitivity of element  $i$  after filtering is:

$$S_i^{new} = \frac{1}{x_i \sum_{i=1}^N H_i} \sum_{j=1}^N H_j x_j S_j \quad (9)$$

Where  $x_j$  is design variable (0 or 1). The weight factor  $H_i$  is written as

$$H_i = \max(0, r - \text{dist}(i, j)) \quad (10)$$

In formula (10) the operator  $\text{dist}(i, j)$  is defined as the distance between center of element  $i$  and center of element  $j$  and operator  $\max(\ )$  is defined to take larger one of two values;  $r$  denotes filtering radius.

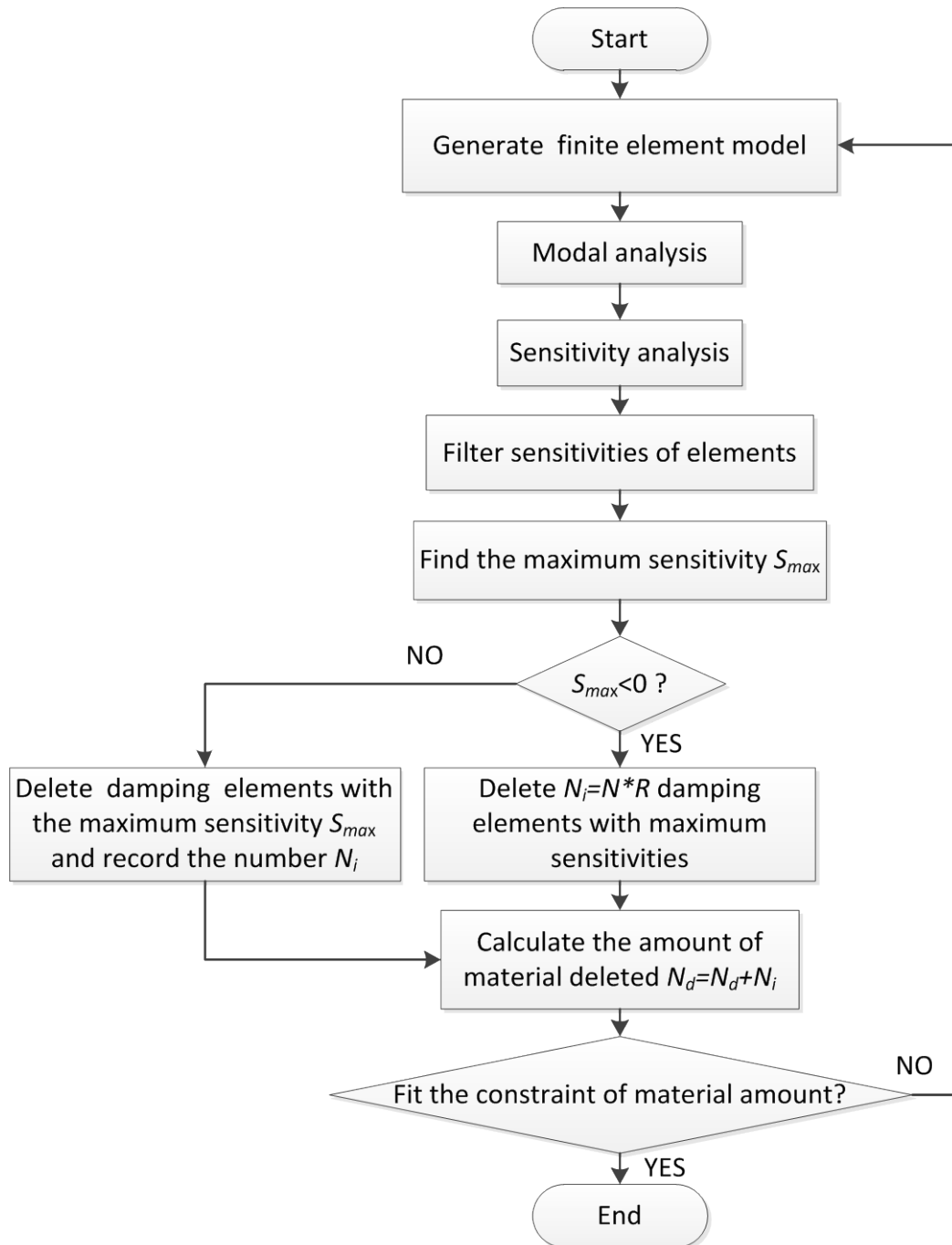
### Optimization process

The initial idea of ESO topology optimization method is the 'survival of the fittest', which will not change when the traditional ESO method is applied to ACDL topology optimization, but its form varies. Design sensitivities of elements could be negative or positive, or they could all be positive. In the former case, elements with maximum positive sensitivities should be removed in each iteration so that the loss factor of whole structure may increase, and in the latter case, elements with minimum absolute sensitivity value should be deleted so that the loss factor may decrease slowly.

The optimization process aiming at the maximization of the modal damping loss factor of constrained damping structure with assigned material properties, boundary conditions, and the single iteration delete ratio  $R$  is shown in figure 2. The specific steps are as follows:

1. Generate the finite element model of the sandwich structure.
2. Execute structural modal analysis and calculate model strain energy.
3. Calculate the sensitivities of elements and filter them.
4. Find the maximum sensitivity.
5. If the maximum value is negative, remove  $N_i = N * R$  ( $N$  denotes the total number of original damping core elements) damping core elements with maximum sensitivities and corresponding constraint sheet elements; If the maximum value is positive, remove damping core elements with maximum value and corresponding constraint layer elements.
6. Judge whether the amount of material constraint conditions is reached, if not, repeat steps 1-5, otherwise finish the result.





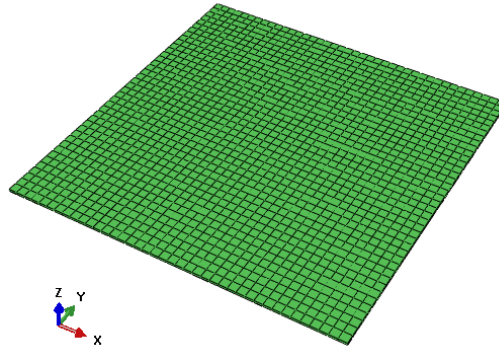
**Figure 2. Topology optimization process based on ESO method**

However, it should be noted that, if the original structure is symmetrical, the number of elements removed will not necessarily equal to  $N * R$  in step 6 to obtain a reasonable symmetrical topology result. The reason may be that the elements in symmetrical positions share equal sensitivities. The deleted method now is to obtain a sensitivity threshold using  $N * R$  and then remove the elements with sensitivities below (or above) the threshold value.

### Numerical examples

Consider a square aluminum plate fully covered with ACDL and four edges clamped, which finite element model is shown in figure 3. Resin-based carbon fiber composite unidirectional laminates is adopted as the constrained sheet and C fiber direction parallels to the x-axis. The viscoelastic

material 3M112D is employed as the damping core. Properties of each material are shown in table 1 and table 2.



**Figure 3. Initial design model**

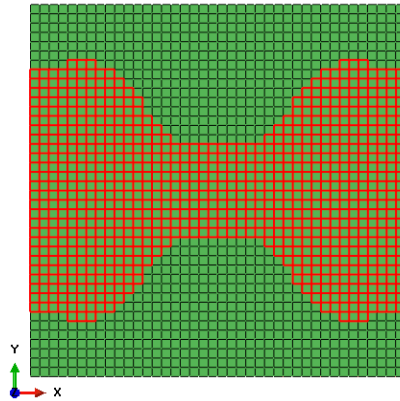
**Table 1. Material properties**

Properties	Bass shell	Damping core	Constraint sheet
Elastic modulus, $E$ (GPa)	70	—	—
Density, $\rho$ (kg/m <sup>3</sup> )	$2.7 \times 10^3$	$0.98 \times 10^3$	$1.64 \times 10^3$
Thickness, $h$ (m)	0.002	0.003	0.001
Shear modulus, $G$ (MPa)	—	10	—
Poisson ratio, $\mu$	0.33	0.49	—
Loss factor, $\eta$	0.0001	1	0.01

**Table 2. Modulus of the constraint sheet**

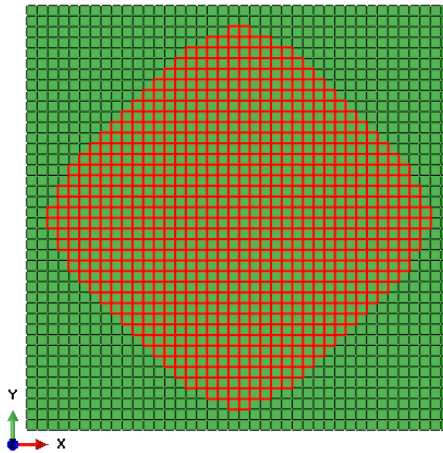
$E_1$ /GPa	$E_2$ /GPa	$\mu_{12}$	$G_{12}$ /GPa	$G_{13}$ /GPa	$G_{23}$ /GPa
125	10.3	0.29	4.3	4.3	4.3

Taking 50% amount of ACDL material as constraint and aiming at the maximization of the first modal damping loss factor, optimization result is shown in figure 4:



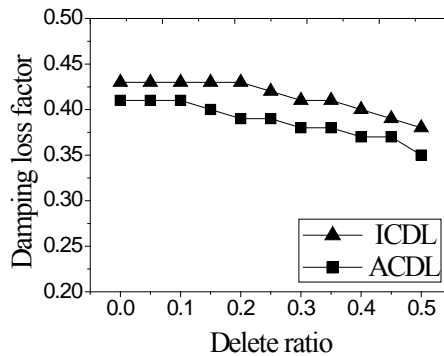
**Figure 4. Topology optimization design of ACDL**

Isotropic material can be treated as a special anisotropic material. The method used here are equally applicable to optimize structure with isotropic constrained damping layers (ICDL). Replace constraint layer material with aluminum in the same finite element model, we can get an optimal result of square plate with traditional ICDL, which is shown in figure 5.



**Figure 5. Topology optimization design of ICDL**

Figure 6 shows damping loss factor's variation history with delete ratio. Loss factors of both model decline slowly when more and more damping elements are deleted. After eliminating 50% constraint damping layer material, structure with ACLD dropped from an initial 0.41 to 0.36, decreased by 12.2%; structure with ICDL dropped from an initial 0.43 to 0.37, decreased by 14.0%.The optimization results are satisfactory.



**Figure 6. History of damping loss factor**

Calculate damping efficiency using formula (1), and damping efficiency was increased greatly after optimization, as shown in table 3. We can also find that ACDL has higher damping efficiency than traditional ICDL.

**Table 3. Damping efficiency of the plates with CDL ( $\text{kg}^{-1}$ )**

	Before optimization	After optimization
ACDL	0.56	0.98
ICDL	0.48	0.82

## Conclusions

The topology optimization problem of thin-walled structure attached ACDL with given thickness is investigated based on ESO method in present paper. The numerical example shows that the damping efficiency increases after optimization, and damping efficiency of ACDL with assigned materials is higher than traditional ICDL due to the non-negligible damping loss factor of composites. The advantage of ACDL is shown in lightweight design of vibration suppression and the approach presented has a strong engineering practicability.

## References

- [1] Lall A K, Asnani N T, Nakra B C. Damping analysis of partially covered sandwich beams[J]. *Journal of Sound and Vibration*, 1988, 123(2): 247-259.
- [2] Lall A K, Asnani N T, Nakra B C. Vibration and damping analysis of rectangular plate with partially covered constrained viscoelastic layer[J]. *Journal of Vibration and Acoustics*, 1987, 109(3): 241-247.
- [3] Zheng H, Cai C, Pau G S, et al. Minimizing vibration response of cylindrical shells through layout optimization of passive constrained layer damping treatments[J]. *Journal of Sound and Vibration*, 2005, 279(3): 739-756.
- [4] Bendsøe M P, Kikuchi N. Generating optimal topologies in structural design using a homogenization method[J]. *Computer methods in applied mechanics and engineering*, 1988, 71(2): 197-224.
- [5] Bendsøe M P. Optimal shape design as a material distribution problem[J]. *Structural optimization*, 1989, 1(4): 193-202.
- [6] Sigmund O. A 99 line topology optimization code written in Matlab[J]. *Structural and Multidisciplinary Optimization*, 2001, 21(2): 120-127.
- [7] Xie Y M, Steven G P. A simple evolutionary procedure for structural optimization[J]. *Computers & structures*, 1993, 49(5): 885-896.
- [8] Wang M Y, Wang X. A level set method for structural topology optimization[J]. *Computer methods in applied mechanics and engineering*, 2003, 192(1): 227-246.
- [9] Allaire G, Jouve F, Toader A M. Structural optimization using sensitivity analysis and a level-set method[J]. *Journal of computational physics*, 2004, 194(1): 363-393.
- [10] Guo Z Z, Chen Y Z. Study on topological optimization design of constrained damping plate based on evolutionary structural optimization [J]. *Journal of Machine Design*, 2006, 23(10): 3-6.
- [11] Ling Z, Ronglu X, Yi W, et al. Topology optimization of constrained layer damping on plates using Method of Moving Asymptote (MMA) approach[J]. *Shock and Vibration*, 2011, 18(1-2): 221-244.
- [12] Sigmund O, Petersson J. Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima[J]. *Structural optimization*, 1998, 16(1): 68-75.

## Free Surface Flow Simulation Using Moving-Grid Finite-Volume Method

\*Sadanori Ishihara<sup>1</sup>, †Kenichi Matsuno<sup>2</sup>, †Masashi Yamakawa<sup>2</sup>, †Takeshi Inotomo<sup>1</sup>,  
and †Shinichi Asao<sup>3</sup>

<sup>1</sup>Graduate School of Science and Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan

<sup>2</sup>Department of Mechanical System Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto, Japan.

<sup>3</sup>Department of Mechanical Engineering, College of Industrial Technology,  
1-27-1, Nishikoya, Amagasaki, Hyogo, Japan.

\*Presenting author: d4821001@edu.kit.ac.jp

†Corresponding author: d4821001@edu.kit.ac.jp

### Abstract

In this paper, an interface-tracking method combined with the Moving-Grid Finite-Volume method is presented for simulating free surface flows. In the interface-tracking method, the calculation grid is moved and deformed according to the movement of the free surface. For tracking free surface, surface height equation for the free surface shape was solved. We applied this method to some flow cases with free surface. Numerical results show that the present new flows simulation method using Moving-Grid Finite-Volume method is very accurate and have a promising feature for free surface flows.

**Keywords:** Free Surface Flows, Computational Fluid Dynamics, Moving-Grid, Finite Volume Method

### Introduction

Free surface flows are important from a point of view of engineering. For example, sloshing flow in tanks, mixing in vessels, jet from nozzle and injection molding are free surface flows.

Many numerical methods are presented for free surface flows [Scardovelli and Zaleski (1999)].

These method can be classified to two approaches, interface-capturing method and interface-tracking method.

In the interface-capturing method, fixed grid is used. For free surface capturing, particle movement or some function is solved. MAC method [Harlow and Welch (1965)], VOF method [Hirt and Nichols (1981)] and level set method [Sussman et al. (1994)] are used in this approach. These method can solve bubble flow and breaking waves. However, these method often have interface smearing.

On the other hand, in the interface-tracking method, moving grid is used. For free surface tracking, computational grids are moved and deformed according to movement of free surface. ALE [Okamoto and Kawahara (1990); Lo and Young (2004); Ushijima (1998)] and Finite Volume method [Apsley and Hu (2003); Muzafferija and Peric (1997)] are used in this approach. This approach is very simple and can track free surface with sharp interface. However, computational grids are usually large deformed.

Free surface flow can be interpreted as a moving boundary problem. For moving boundary problems, Moving-Grid Finite-Volume Method was suggested [Mihara and (1999)]. This method can solve flow with moving grids with satisfying physical and geometrical conservation laws. The method has been applied to various flows [Mihara et al. (1999); Watanabe and Matsuno (2009); Matsuno (2010)]. However, these applications have been limited to single phase flows.

The purpose of this paper is to extend the Moving-Grid Finite-Volume Method to free surface flows. The main advantages of this method is simple treatment with free surface and satisfying physical and geometrical conservation laws.

In some test cases with free surface, comparison with analytical solution or experimental data are presented.

## Governing Equations

The governing equations are the continuity equation and the nondimensionalized incompressible Navier-Stokes equations. These equations are written as follows:

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0, \quad (1)$$

and

$$\frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{E}}{\partial x} + \frac{\partial \mathbf{F}}{\partial y} + \frac{\partial \mathbf{G}}{\partial z} = \mathbf{H}, \quad (2)$$

where  $x, y$  and  $z$  are coordinates,  $t$  is time.  $u, v$  and  $w$  are velocities in  $x, y$  and  $z$  directions, respectively.  $\mathbf{q}$  is the velocity vector,  $\mathbf{q} = [u, v, w]^T$ .  $\mathbf{E}, \mathbf{F}$  and  $\mathbf{G}$  are flux vectors in  $x, y$  and  $z$  directions, respectively.  $\mathbf{H}$  is the body force term including gravity.

Flux vectors are written as follows:

$$\mathbf{E} = \hat{\mathbf{E}} - \mathbf{E}_v + \mathbf{E}_p, \quad \mathbf{F} = \hat{\mathbf{F}} - \mathbf{F}_v + \mathbf{F}_p, \quad \mathbf{G} = \hat{\mathbf{G}} - \mathbf{G}_v + \mathbf{G}_p, \quad (3)$$

where  $\hat{\mathbf{E}}, \hat{\mathbf{F}}$  and  $\hat{\mathbf{G}}$  are the advection flux vectors,  $\mathbf{E}_v, \mathbf{F}_v$  and  $\mathbf{G}_v$  are the viscous flux vectors, and  $\mathbf{E}_p, \mathbf{F}_p$  and  $\mathbf{G}_p$  are pressure flux vectors in  $x, y$  and  $z$  directions, respectively. The elements of flux vectors and body force term are:

$$\mathbf{E} = \begin{bmatrix} u^2 \\ uv \\ uw \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} vu \\ v^2 \\ vw \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} wu \\ wv \\ w^2 \end{bmatrix}, \quad \mathbf{E}_v = \frac{1}{\text{Re}} \begin{bmatrix} u_x \\ v_x \\ w_x \end{bmatrix}, \quad \mathbf{F}_v = \frac{1}{\text{Re}} \begin{bmatrix} u_y \\ v_y \\ w_y \end{bmatrix}, \quad \mathbf{G}_v = \frac{1}{\text{Re}} \begin{bmatrix} u_z \\ v_z \\ w_z \end{bmatrix},$$

$$\mathbf{E}_p = \begin{bmatrix} p \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{F}_p = \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix}, \quad \mathbf{G}_p = \begin{bmatrix} 0 \\ 0 \\ p \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} X \\ 0 \\ -\frac{1}{\text{Fr}^2} \end{bmatrix}, \quad (4)$$

where  $p$  is the pressure,  $\text{Re}$  is the Reynolds number and  $\text{Fr}$  is the Froude number.

The subscripts  $x, y$  and  $z$  indicate derivatives with respect to  $x, y$  and  $z$  respectively.  $X$  is the body force in  $x$  direction. Reynolds number and Froude number are:

$$\text{Re} = \frac{\bar{U}_0 \bar{L}_0}{\bar{\nu}}, \quad \text{Fr} = \frac{\bar{U}_0}{\sqrt{\bar{g} \bar{L}_0}}, \quad (5)$$

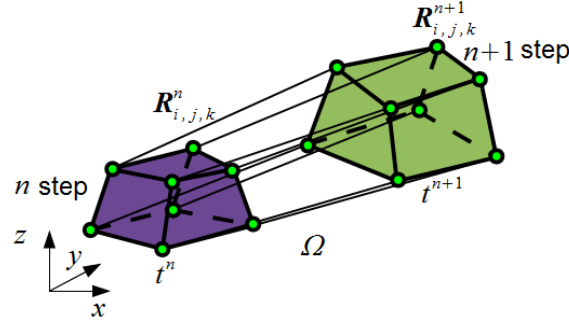
where  $\bar{L}_0$  is the characteristic length,  $\bar{U}_0$  is the characteristic velocity,  $\bar{\nu}$  is the kinematic viscosity and  $\bar{g}$  is the gravitational acceleration. Over bar shows the dimensional quality.

## Discretization method and numerical method

### Discretization method

In discretization for these equations, we use Moving-Grid Finite-Volume Method. This discretization method is based on cell-centered Finite-Volume Method in space-time unified

domain. In three-dimensional case, four-dimensional polyhedron in the  $(x, y, z, t)$  control volume is used.



**Figure 1. Schematic drawing of control volume**

Fig. 1 shows schematic drawing of structured control volume in  $(x, y, z, t)$  unified domain.  $\mathbf{R}$  is grid position vector,  $\mathbf{R} = [x, y, z]^T$ , where superscript  $n$  shows time step and subscript  $i, j, k$  show structured grid point indexes. The purple region is  $n$  time step computational cell, green region is  $n + 1$  step computational cell. The control volume is four-dimensional polyhedron  $\Omega$ . Eq. (2) is integrated with the control volume  $\Omega$  as

$$\int_{\Omega} \left( \frac{\partial \mathbf{q}}{\partial t} + \frac{\partial \mathbf{E}}{\partial x} + \frac{\partial \mathbf{F}}{\partial y} + \frac{\partial \mathbf{G}}{\partial z} \right) d\Omega = V_{\Omega} \mathbf{H}, \quad (6)$$

where  $V_{\Omega}$  is a four-dimensional volume ( $V_{\Omega} = \int_{\Omega} d\Omega$ ). Eq. (6) is divergence form in  $(x, y, z, t)$  dimension. By using Gaussian divergence theorem, Eq. (6) is written as,

$$\int_{\Omega} \left[ \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial t} \right) \cdot (\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{q}) \right] d\Omega = \oint_{\partial\Omega} [(\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{q}) \cdot \mathbf{n}] d(\partial\Omega) = V_{\Omega} \mathbf{H}, \quad (7)$$

where  $\mathbf{n}$  is the vector normal to control volume surface.  $\partial\Omega$  is the surface of the control volume. The  $n_x, n_y, n_z$  and  $n_t$  are components of  $\mathbf{n}$  in  $x, y, z$  and  $t$  directions, respectively.

From Eq. (7), we can write as,

$$\sum_{l=1}^8 (q n_l + \Phi - \Psi + \Pi) = V_{\Omega} \mathbf{H}, \quad (8)$$

where,

$$\Phi = \hat{\mathbf{E}} n_x + \hat{\mathbf{F}} n_y + \hat{\mathbf{G}} n_z, \quad \Psi = \frac{1}{\text{Re}} (\mathbf{E}_v n_x + \mathbf{F}_v n_y + \mathbf{G}_v n_z), \quad \Pi = \mathbf{E}_p n_x + \mathbf{F}_p n_y + \mathbf{G}_p n_z. \quad (9)$$

The subscript  $l$  in Eq. (8) denotes the surface of the control volume in four dimension  $(x, y, z, t)$ .

The  $l = 7$  surface normal vector is the computational cell at  $n$  time step,  $l = 8$  surface normal vector is the computational cell at  $n + 1$  time step. For example, Fig. 2 shows the control volume surface at  $l = 2$ .

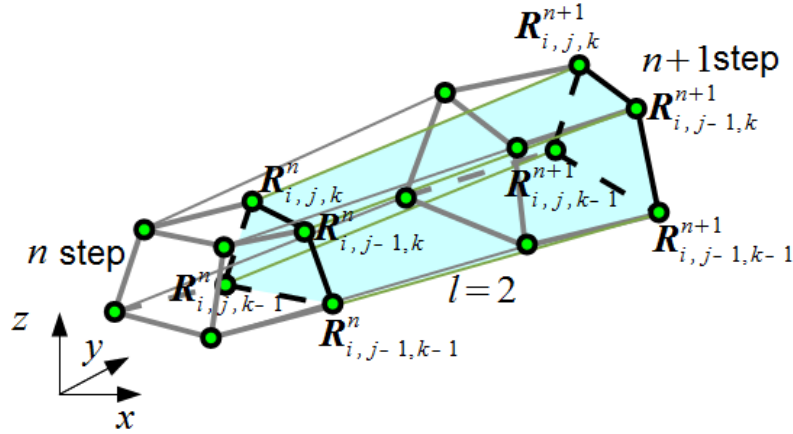


Figure 2. Schematic drawing of control volume surface

Control volume surfaces at  $l = 7$ ,  $l = 8$  are perpendicular to the  $t$  axis. Eq. (8) becomes as follows:

$$\mathbf{q}^{n+1}(n_t)_8 + \mathbf{q}^n(n_t)_7 + \sum_{l=1}^6 (\mathbf{q}^{n+1/2} n_t + \Phi^{n+1/2} - \Psi^{n+1/2} + \Pi^{n+1/2}) = V_\Omega \mathbf{H}. \quad (10)$$

This Eq. (10) is discretized equation.

#### Numerical method

To solve Eq. (10), we use SMAC method [Amsden and Harlow (1970)]. Intermediate velocity is solved iteratively using LU-SGS method [Yoon and Jameson (1988)]. The inviscid term  $\Phi$  and moving grid term  $\mathbf{q}n_t$  are evaluated using QUICK method [Leonard (1979)]. The viscous term  $\Psi$  and pressure gradient term  $\Pi$  are evaluated using central difference scheme. The Poisson equation about pressure correction is solved iteratively using Bi-CGSTAB [van der Vorst (1992)].

#### Interface-Tracking method

In present method, a surface height equation [Apsley and Hu (2003)] is solved for free surface height. Fig. 3 shows free surface shape.

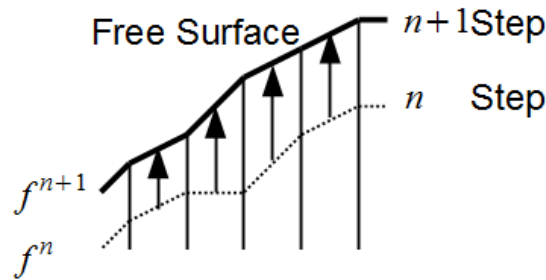


Figure 3. Free surface height

In Fig. 3,  $f = f(x, y, t)$  is the surface height function. The surface height equation is as follows:

$$\frac{\partial f}{\partial t} + u \frac{\partial f}{\partial x} + v \frac{\partial f}{\partial y} = w. \quad (11)$$

Eq. (11) can be discretized as follows:

$$f_{i,j}^{n+1} = f_{i,j}^n - \Delta t (u f_x + v f_y - w)^n, \quad (12)$$



where  $\Delta t$  is time step size, superscript  $n$  shows time step and subscript  $i, j, k$  shows grid point indexes. The  $uf_x$  and  $vf_y$  in Eq. (12) are evaluated using 1st order upwind differencing scheme. Once the free surface height is solved, computational grids moved and deformed according to the movement of the free surface. In the present study, the surface tension is neglected. The pressure on the free surface is fixed by  $p = 0$ .

## Numerical Results

### Sloshing flow case

To check the present method, sloshing flow case is solved. A comparison was made with the experimental result [Okamoto and Kawahara (1990)]. The geometry of the domain is shown in Fig. 4.

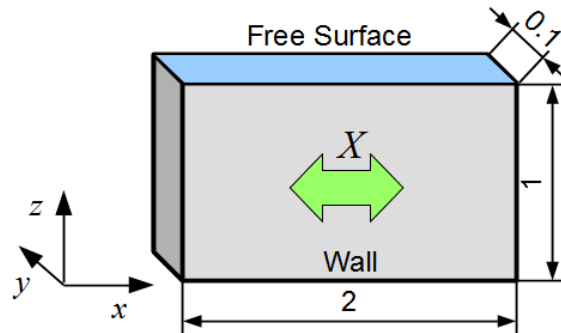


Figure 4. Geometry of sloshing flow case

The size of the initial domain is  $2 \times 0.1 \times 1$ . The calculation domain is nondimensionalized by initial free surface height.

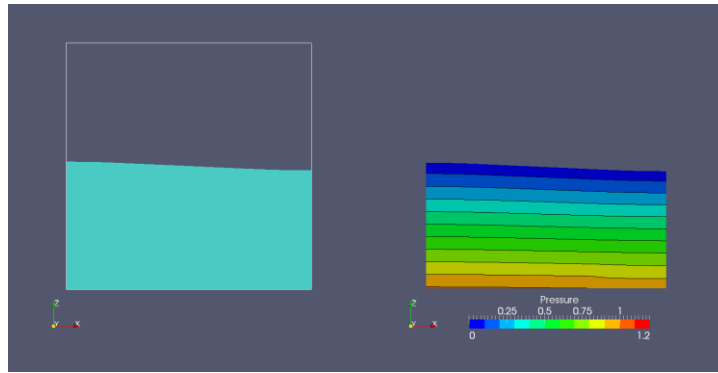
In our case, fluid flow with oscillating body force is solved. The body force is as follows:

$$X = -A\varpi^2 \sin \varpi t, \quad (13)$$

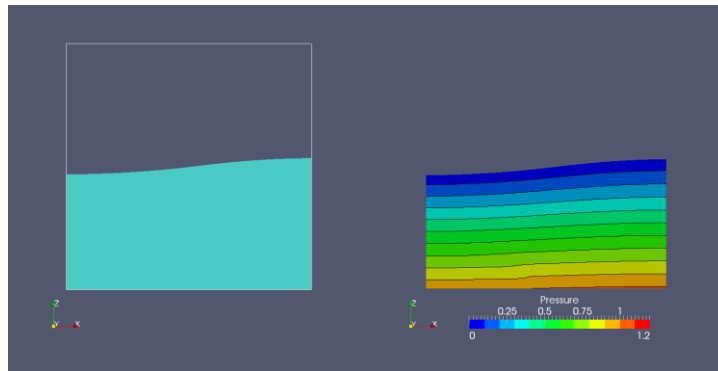
where  $A = 0.00186$  and  $\varpi = 1.20$  are the amplitude and frequency of the oscillation, respectively. The numerical grid used for calculation had  $61 \times 11 \times 51$  grid points. The time step size is 0.001. The Reynolds number is  $1.1 \times 10^6$  and the Froude number is 1.

The initial condition of the velocity is given by  $u = v = w = 0$ . The initial condition of the pressure is given by  $p = 0$ . The boundary conditions is as follows. In the wall boundary, the velocity is slip condition and the pressure is Neumann boundary condition. In the free surface, the velocity is 0th extrapolated and the pressure is fixed by  $p = 0$ .

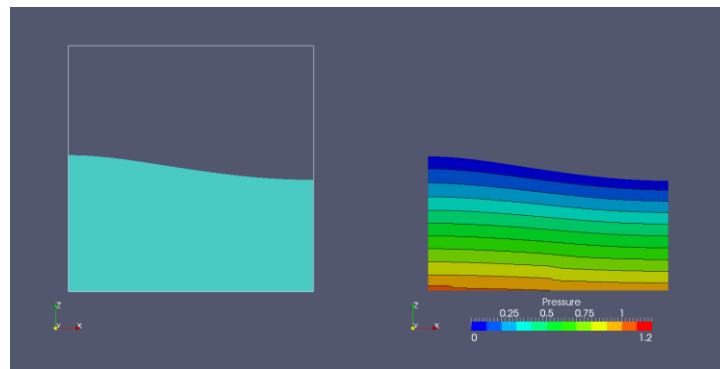
Fig. 5 shows numerical simulation results. In these figures, the left column show the free surface shape and the right column show pressure distributions at  $t=2.6, 5.2, 7.9$  and  $10.5$ . From the free surface shape, nonlinear surface movement is appeared.



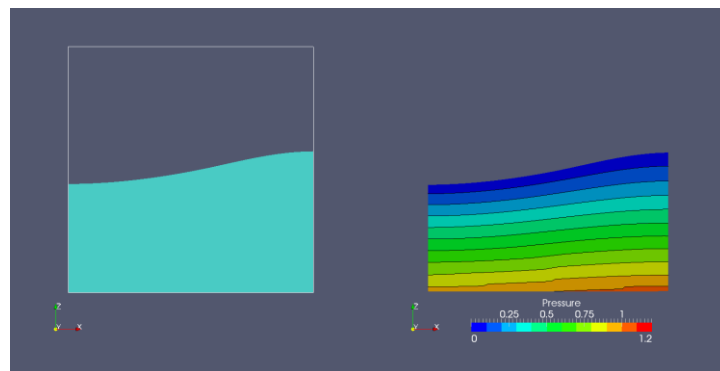
(a)  $t=2.6$



(b)  $t=5.2$

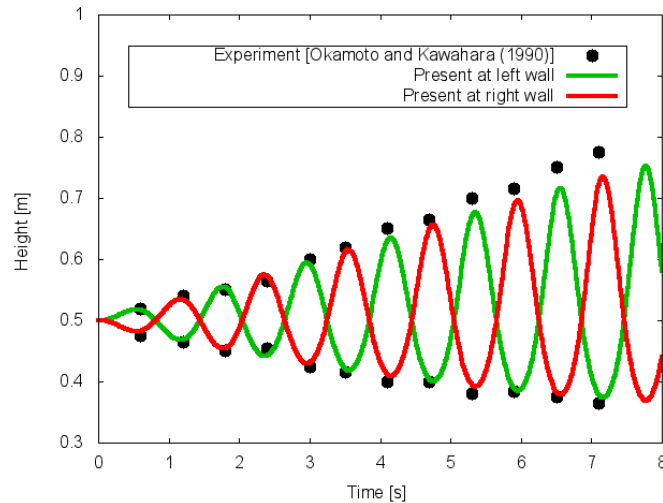


(c)  $t=7.9$



(d)  $t=10.5$

Figure 5. Results of sloshing flow case (left : surface shape, right : pressure distributions)

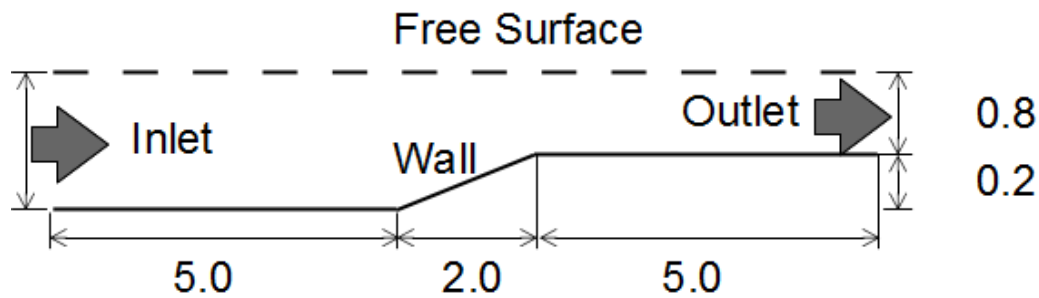


**Figure 6. Time history of the free surface elevation at the left wall and the right wall**

Fig. 6 shows the time history of the free surface elevation at the left wall and right wall. The initial free surface height is 0.5 m. Points are experimental results [Okamoto and Kawahara (1990)], green lines is present result at the left wall and red line is present result at the right wall. As shown in Fig. 6, surface height at the left wall and the right wall increase alternately, and present results agree reasonably with experimental one. From these results, the present method can apply sloshing analysis.

*Ramp flow case*

In inviscid case, flow over a ramp [Apsley and Hu (2003); Muzaferija and Peric (1997)] is solved. This case is basic test case with free surface flow. The geometry of the domain is shown in Fig. 7. Four Froude number condition cases were solved: subcritical flow at  $Fr=0.3, 0.32$  and supercritical flow at  $Fr=1.92, 2$ . These conditions are same as references  $Fr=0.3, 2.0$  [Apsley and Hu (2003)] and  $Fr=0.32, 1.92$  [Muzaferija and Peric (1997)].



**Figure 7. Geometry of ramp flow case**

The calculation domain is nondimensionalized by initial free surface height. The numerical grid used for calculation had  $61 \times 11 \times 31$  grid points. The time step size is 0.005. In this case, inviscid flow is assumed. The Froude number is 0.3, 0.32, 1.92 and 2.0.

The initial condition of velocity is given by  $u = 1, v = w = 0$ . The initial condition of pressure is given by hydrostatic pressure. The boundary conditions as follows. In the inlet boundary, the velocity is fixed by  $u = 1, v = w = 0$ , and the pressure is Neumann boundary condition. In the outlet boundary, the velocity is 0th extrapolated and the pressure is 0th extrapolated. In the bottom wall boundary, the velocity is slip condition and the pressure is Neumann boundary condition. In the

front and back boundary, the velocity is slip condition and the pressure is Neumann boundary condition. In the free surface, the velocity is 0th extrapolated and the pressure is fixed by  $p = 0$ . Fig .8 shows the free surface shape of  $Fr=0.3$  case at  $t=200$ . Fig. 9 shows the free surface shape of  $Fr=2.0$  case at  $t=200$ .



**Figure 8. Surface shape of ramp flow case ( $Fr=0.3$ )**



**Figure 9. Surface shape of ramp flow case ( $Fr=2.0$ )**

As shown in Fig. 8, the free surface height at the outlet boundary decrease from the inlet boundary. On the other hand, as shown in Fig. 9, the free surface height at the outlet boundary increase from the inlet boundary. These results are caused by subcritical ( $Fr=0.3$ ) or supercritical ( $Fr=2.0$ ) conditions.

Table 1 shows the free surface height from the bottom wall at the outlet boundary. In Table 1, 1-d theory shows analytical results using 1-d theory [Apsley and Hu (2003)]. Present shows present results. Error shows our results error from 1-d theory. As shown in Table 1, our results are agree with reference solutions and 1-d theory solutions.

**Table 1. Free surface height of ramp flow case**

	Froude number			
	0.3	0.32	1.92	2
1-d theory	0.7689	0.7635	1.0897	1.0776
Present	0.7949	0.7940	1.0910	1.0794
Error [%]	3.39	3.98	0.12	0.17
Apsley and Hu (2003)	0.7687	—	—	1.0792
Muzaferija and Peric (1997)	—	0.7752	1.0992	—

**Conclusions**

In this paper, new flow simulation method with free surface is presented. This method is based on the Moving-Grid Finite-Volume Method and coupled with the interface-tracking method. We applied present new method to some flow cases. From the comparison with experimental or numerical data, present method using the Moving-Grid Finite-Volume method is very accurate and have a promising feature for free surface flows.

## Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers 24560192, 25420120.

## References

- Amsden, A. A. and Harlow, F. H. (1970) A simplified MAC technique for incompressible fluid flow calculations, *Journal of Computational Physics* **6-2**, 322–325.
- Apsley, D. and Hu, W. (2003) CFD simulation of two- and three-dimensional free-surface flow, *International Journal for Numerical Methods in Fluids* **42**, 465–491.
- Harlow, F. H. and Welch, J. E. (1965) Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface, *The Physics of Fluids*, **8-12**, 2182–2189.
- Hirt, C. W. and Nichols, B. D. (1981) Volume of fluid (VOF) method for the dynamics of free boundaries, *Journal of Computational Physics* **39-1**, 201–205.
- Scardovelli, R. and Zaleski, S. (1999) Direct numerical simulation of free surface and interfacial flow, *Annual Review of Fluid Mechanics* **31**, 561–603.
- Sussman, M., Smereca, P. and Osher, S. (1994) A level set approach for computing solutions for incompressible two-phase flow, *Journal of Computational Physics* **114-1**, 146–159.
- Leonard, B. P. (1979) A stable and accurate convective modeling procedure based on quadratic interpolation, *Computer Methods in Applied Mechanics and Engineering* **19**, 59–98.
- Lo, D. C. and Young, D. L. (2004) Arbitrary Lagrangian-Eulerian finite element analysis of free surface flow using a velocity-vorticity formulation, *Journal of Computational Physics* **195-1**, 175–201.
- Matsuno, K. (2010) Development and applications of a moving grid finite volume method, in developments and applications in engineering computational technology, Chapter 5, ed. Topping, B. H. V. et al., *Saxe-Coburg Publications*, 103–129.
- Mihara, K., Matsuno, K. and Satofuka, N. (1999) An iterative finite-volume scheme on a moving grid (1st report, the fundamental formulation and validation), *Transactions of the Japan Society of Mechanical Engineers Series B* **65-637**, 2945–2953 (in Japanese).
- Muzafferija, S. and Peric, M. (1997) Computation of free-surface flows using the finite-volume method and moving grids, *Numerical Heat Transfer Part B* **32**, 360–384.
- Okamoto, T. and Kawahara, T. (1990) Two-dimensional sloshing analysis by Lagrangian finite element method, *International Journal for Numerical Methods in Fluids* **11**, 453–477.
- Ushijima, S. (1998) Three-dimensional arbitrary Lagrangian-Eulerian numerical prediction method for non-linear free surface oscillation, *International Journal for Numerical Method in Fluids* **26**, 605–623.
- van Der Vorst, H. (1992) Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear Systems, *SIAM Journal on Scientific Computing* **13-2**, 631–644.
- Watanabe, K. and Matsuno, K. (2009) Moving computational domain method and its application to flow around a high-speed car passing through a hairpin curve, *Journal of Computational Science and Technology* **3**, 449–459.
- Yoon, S. and Jameson, Y. (1988) Lower-upper symmetric-Gauss-Seidel methods for the Euler and Navier-Stokes equations, *AIAA Journal* **26**, 1025–1026.

## An effective improved algorithm for Finite Particle Method

†Yang Yang Fei Xu Lu Wang Xiao-ting Li

School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China

†Corresponding author: npuyang@nwpu.edu.cn

**Abstract:** The low accuracy near the boundary or the interface in SPH method has been paid extensive attention. Finite Particle Method (FPM) is a significant improvement to the traditional SPH method, which can greatly improve the computational accuracy for boundary particles. However, there are still some inherent defects for FPM, such as the long computing time and the potential numerical instability. By conducting matrix decomposition and structural analysis on the basic equations of FPM, an improved FPM method (IFPM) is proposed, which can not only maintain the  $C^1$  consistence of FPM for boundary particles, but also keep the invertibility of the coefficient matrix in traditional FPM and greatly reduce the computing time. Finally, some numerical results show that IFPM is really an effective improvement for traditional FPM.

**Key words:** SPH; Finite Particle Method; matrix decomposition; stability;  $C^1$  consistence

### 1. Introduction

Traditional Smoothed Particle Hydrodynamics method (SPH) is the most important meshfree particle methods [Lucy. (1997); Gingold and Monaghan. (1997); Liu and Liu. (2003)], which has been widely applied in Fluid Dynamics [Yang et al. (2014) ;Feng et al. (2013)], Continuum Elasticity [Liu et al. (2011)], Solid Mechanics [Libersky et al. (1993)] and so on. However, the low accuracy near the boundary or the interface is also the most remarkable defect for SPH by now. Some numerical method has been proposed to improve this problem, such as the Corrective Smoothed Particle Method (CSPM) [Chen and Beraun. (2000)] and discontinuous SPH [Liu et al. (2003); Xu et al. (2013)], but all of them could not avoid the error accumulation.

In 2005, M B Liu, G R Liu and G M Zhang, R C Batra proposed a new-type SPH method based on Taylor series expansion respectively[Liu et al. (2005); Zhang and Batra. (2004)], named Finite Particle Method (FPM). Compared with SPH, FPM has the advantages of free selection on the basis function, high accuracy near the boundary, and it is also not sensitive to the smooth length and the irregular distribution of particles. In addition, FPM could get the function value and derivative value simultaneously, which avoids the error propagation when the low-order derivative is used in the calculation of the high-order derivative in SPH.

However, there are still two disadvantages for FPM, which are the long computation time and computational instability. The former is caused by the large amount of calculation on solving the linear equations for each particle in the

computational domain, and the latter is because the invertibility of the coefficient matrix in the linear equations cannot be well satisfied all the time. Therefore, in this paper, an improved algorithm for FPM is proposed, which is abbreviated to IFPM and is proven to cover the shortage in FPM effectively.

## 2. Original FPM

In 1D case, considering a Taylor series expansion at  $x_i$  up to the first-order derivative,

$$f(x) = f(x_i) + (x - x_i)f'_x(x_i) \quad (1)$$

Multiplying both sides of (1) with the basis function  $\varphi_1(x)$  and  $\varphi_2(x)$  respectively, and integrating in the computational domain,

$$\begin{aligned} \int f(x)\varphi_1(x)dx &= f(x_i)\int \varphi_1(x)dx + f'_x(x_i)\int (x - x_i)\varphi_1(x)dx \\ \int f(x)\varphi_2(x)dx &= f(x_i)\int \varphi_2(x)dx + f'_x(x_i)\int (x - x_i)\varphi_2(x)dx \end{aligned} \quad (2)$$

Expression (2) could be seemed as linear equations, and expressed as the following matrix form,

$$\begin{bmatrix} \int \varphi_1(x)dx & \int (x - x_i)\varphi_1(x)dx \\ \int \varphi_2(x)dx & \int (x - x_i)\varphi_2(x)dx \end{bmatrix} \begin{bmatrix} f(x_i) \\ f'_x(x_i) \end{bmatrix} = \begin{bmatrix} \int f(x)\varphi_1(x)dx \\ \int f(x)\varphi_2(x)dx \end{bmatrix} \quad (3)$$

Further, the particle form of the above equations (3) can be obtained as follows,

$$\begin{bmatrix} \sum_{j=1}^N \varphi_1(x_j)\Delta d_j & \sum_{j=1}^N (x_j - x_i)\varphi_1(x_j)\Delta d_j \\ \sum_{j=1}^N \varphi_2(x_j)\Delta d_j & \sum_{j=1}^N (x_j - x_i)\varphi_2(x_j)\Delta d_j \end{bmatrix} \begin{bmatrix} f(x_i) \\ f'_x(x_i) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^N f(x_j)\varphi_1(x_j)\Delta d_j \\ \sum_{j=1}^N f(x_j)\varphi_2(x_j)\Delta d_j \end{bmatrix} \quad (4)$$

where  $\Delta d_j$  is the particle size. Equation (4) is the basic expression of FPM in 1D case.

Similarly, the basic expression of FPM in 2D case could be derived as follows,

$$\begin{bmatrix} \sum_{j=1}^N \varphi_1(x_j, y_j)\Delta S_j & \sum_{j=1}^N (x_j - x_i)\varphi_1(x_j, y_j)\Delta S_j & \sum_{j=1}^N (y_j - y_i)\varphi_1(x_j, y_j)\Delta S_j \\ \sum_{j=1}^N \varphi_2(x_j, y_j)\Delta S_j & \sum_{j=1}^N (x_j - x_i)\varphi_2(x_j, y_j)\Delta S_j & \sum_{j=1}^N (y_j - y_i)\varphi_2(x_j, y_j)\Delta S_j \\ \sum_{j=1}^N \varphi_3(x_j, y_j)\Delta S_j & \sum_{j=1}^N (x_j - x_i)\varphi_3(x_j, y_j)\Delta S_j & \sum_{j=1}^N (y_j - y_i)\varphi_3(x_j, y_j)\Delta S_j \end{bmatrix} \begin{bmatrix} f(x_i, y_i) \\ f'_x(x_i, y_i) \\ f'_y(x_i, y_i) \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^N f(x_j, y_j)\varphi_1(x_j, y_j)\Delta S_j \\ \sum_{j=1}^N f(x_j, y_j)\varphi_2(x_j, y_j)\Delta S_j \\ \sum_{j=1}^N f(x_j, y_j)\varphi_3(x_j, y_j)\Delta S_j \end{bmatrix} \quad (5)$$

where  $\Delta S_j$  is the particle size.

### 3. IFPM

#### 3.1 1D case

In 1D case, IFPM could be derived based on the matrix decomposition on the coefficient matrix and the constant term in the basic FPM equation (4).

First, the coefficient matrix decomposition:

$$A = \begin{bmatrix} \sum_{j=1}^N \varphi_1(x_j) \Delta d_j & \sum_{j=1}^N (x_j - x_i) \varphi_1(x_j) \Delta d_j \\ \sum_{j=1}^N \varphi_2(x_j) \Delta d_j & \sum_{j=1}^N (x_j - x_i) \varphi_2(x_j) \Delta d_j \end{bmatrix} = \begin{bmatrix} \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_N) \\ \varphi_2(x_1) & \varphi_2(x_2) & \cdots & \varphi_2(x_N) \end{bmatrix} \begin{bmatrix} \Delta d_1 & & & \\ & \Delta d_2 & & \\ & & \ddots & \\ & & & \Delta d_N \end{bmatrix} \begin{bmatrix} 1 & x_1 - x_i \\ 1 & x_2 - x_i \\ \vdots & \vdots \\ 1 & x_N - x_i \end{bmatrix} \triangleq K \cdot D \cdot C \quad (6)$$

Second, the constant term decomposition:

$$B = \begin{bmatrix} \sum_{j=1}^N f(x_j) \varphi_1(x_j) \Delta d_j \\ \sum_{j=1}^N f(x_j) \varphi_2(x_j) \Delta d_j \end{bmatrix} = \begin{bmatrix} \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_N) \\ \varphi_2(x_1) & \varphi_2(x_2) & \cdots & \varphi_2(x_N) \end{bmatrix} \begin{bmatrix} \Delta d_1 & & & \\ & \Delta d_2 & & \\ & & \ddots & \\ & & & \Delta d_N \end{bmatrix} \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_N) \end{bmatrix} \triangleq K \cdot D \cdot F \quad (7)$$

Based on (6) and (7), the basic FPM equation (4) could be expressed as follows,

$$KDCf = KDF \quad (8)$$

where  $f = [f(x_1) \ f(x_2) \ \cdots \ f(x_N)]^T$ ,  $N$  is the number of the particles in the support of  $x_i$ .

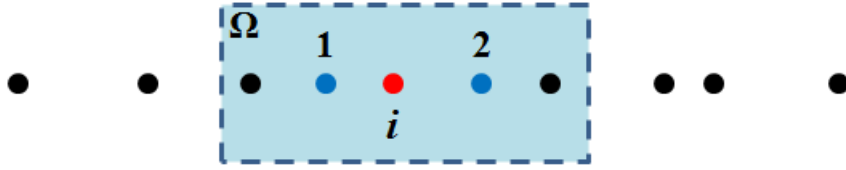
Solving equations (8) is equivalent to solve the following equations,

$$KD(Cf - F) = 0 \quad (9)$$

Since FPM is free on the selection of the basis function, we could just consider the case of  $rank(K) = 2$ , i.e.  $K$  is a row full-rank matrix. Specially, as shown in Figure 1, if just two nearest particles to  $x_i$  in its support (i.e.  $N = 2$ ) are chosen and introduced into the approximate calculation equations (9) of  $x_i$ , the matrix  $K$ ,  $C$  and  $F$  have the following reduced forms,

$$K = \begin{bmatrix} \varphi_1(x_1) & \varphi_1(x_2) \\ \varphi_2(x_1) & \varphi_2(x_2) \end{bmatrix}, \quad D = \begin{bmatrix} \Delta d_1 & 0 \\ 0 & \Delta d_2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & x_1 - x_i \\ 1 & x_2 - x_i \end{bmatrix}, \quad F = \begin{bmatrix} f(x_1) \\ f(x_2) \end{bmatrix} \quad (10)$$





**Figure 1. The schema on the selection of computational particles in 1D-IFPM**

Here, the matrices  $K$  and  $D$  are invertible, and thus the equations (8) is equivalent to the following simplified equations (11),

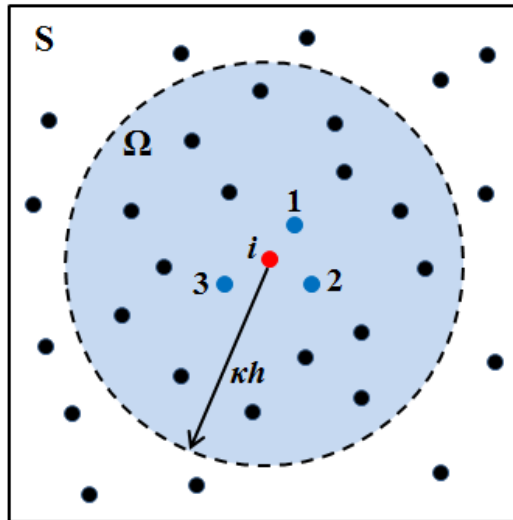
$$Cf - F = 0 \tag{11}$$

Equations (11) are the basic equations of proposed 1D-IFPM in this paper. It can be clearly found that the coefficient matrix  $C$  of equations (11) is always invertible, and thus 1D-IFPM is always stable.

*3.2 2D case*

Similar to section 3.1, 2D-IFPM equation could be derived after the decomposition and deformation to equation (5). The difference is that, as shown in Figure 2, three nearest particles to  $x_i$  in its support (i.e.  $N = 3$ ) should be chosen and introduced into the approximate calculation equations in order to keep that  $K$  is a invertible matrix. The basic equation of 2D-IFPM are shown as follows (12),

$$\begin{bmatrix} 1 & x_1 - x_i & y_1 - y_i \\ 1 & x_2 - x_i & y_2 - y_i \\ 1 & x_3 - x_i & y_3 - y_i \end{bmatrix} \begin{bmatrix} f(x_i, y_i) \\ f_x(x_i, y_i) \\ f_y(x_i, y_i) \end{bmatrix} = \begin{bmatrix} f(x_1, y_1) \\ f(x_2, y_2) \\ f(x_3, y_3) \end{bmatrix} \tag{12}$$



**Figure 2. The schema on the selection of computational particles in 2D-IFPM**

It can be also clearly found that the coefficient matrix of equations (12) is always invertible, and thus 2D-IFPM is always stable.

Meanwhile, it must be mentioned that the uniform distribution of particles is actually a special case of random distribution, thus the above IFPM method is also suitable for uniformly distributed particles.

#### 4. Numerical tests and analysis

##### 4.1 Particle approximation accuracy

Generally speaking, if one method could reproduce a  $k$ -order polynomial, this method could be called with  $C^k$  consistence. In order to verify the accuracy of IFPM, the following examples are tested. For 1D case, considering the function  $f(x)=1$  and  $f(x)=x$  in  $[0,10]$  respectively, 11 particles are randomly distributed and their coordinates are shown as follows. The smooth length  $h=1$ , and the cubic B-spline function and its derivative function are chosen as the kernel function for SPH and the basic functions for FPM [Monaghan. (1992)].

**(R)**: 0.3770, 2.3160, 3.9550, 4.8890, 6.2410, 6.7910, 7.9620, 8.8520, 9.1330, 9.8800

The numerical results are shown in the Table 1~2, where *Error* is defined as the difference between the reproduced value and the exact value, i.e.  $Error = \hat{f} - f$ . It can be seen that both IFPM and FPM have distinct accuracy compared with SPH method. They could reproduce both the interior particles and the particles near the boundary accurately, which means they could be called with  $C^1$  consistence.

**Table 1. Reproduced results for  $f(x) = 1$**

Exact $f$	Reproduced results and error					
	SPH		FPM		IFPM	
	$\hat{f}$	<i>Error</i>	$\hat{f}$	<i>Error</i> / $\times 10^{-15}$	$\hat{f}$	<i>Error</i> / $\times 10^{-15}$
1	0.8977	-0.1023	1.0000	0	1.0000	-0.2220
1	1.2028	0.2028	1.0000	0	1.0000	0
1	1.1022	0.1022	1.0000	0	1.0000	0
1	1.0648	0.0648	1.0000	-0.1110	1.0000	0
1	1.0745	0.0745	1.0000	-0.2220	1.0000	0
1	1.0971	0.0971	1.0000	0	1.0000	0
1	0.9539	-0.0461	1.0000	0	1.0000	0
1	1.0076	0.0076	1.0000	0	1.0000	0
1	0.9476	-0.0524	1.0000	-0.1110	1.0000	0
1	0.5817	-0.4183	1.0000	0	1.0000	0
1	0.8977	-0.1023	1.0000	0	1.0000	-0.2220

**Table 2. Reproduced results for  $f(x) = x$**

Exact $f$	Reproduced results and error					
	SPH		FPM		IFPM	
	$\hat{f}$	Error	$\hat{f}$	Error/ $\times 10^{-14}$	$\hat{f}$	Error/ $\times 10^{-15}$
0.3770	0.3386	-0.0384	0.3770	-0.0056	0.3770	-0.2220
2.3160	2.8021	0.4861	2.3160	0	2.3160	0
3.9550	4.5517	0.5967	3.9550	0.0888	3.9550	0
4.8890	5.0218	0.1328	4.8890	-0.0888	4.8890	0
6.2410	6.8541	0.6131	6.2410	-0.0888	6.2410	0
6.7910	7.3309	0.5399	6.7910	0	6.7910	0
7.9620	7.6692	-0.2928	7.9620	0	7.9620	0
8.8520	8.8754	0.0234	8.8520	0	8.8520	0
9.1330	8.5581	-0.5749	9.1330	-0.1776	9.1330	0
9.8800	5.5330	-4.3470	9.8800	0	9.8800	0
0.3770	0.3386	-0.0384	0.3770	-0.0056	0.3770	-0.2220

For 2D case, considering the function  $f(x, y) = 1$  and  $f(x, y) = x + y$  in  $[0, 100] \times [0, 100]$  respectively, 5664 particles are randomly distributed in this area. The numerical results are shown in the Table 3, where *MSE* represents the Mean Square Error. It can be seen that IFPM could also keep  $C^l$  consistence in 2D case.

**Table 3. Reproduced results in 2D case**

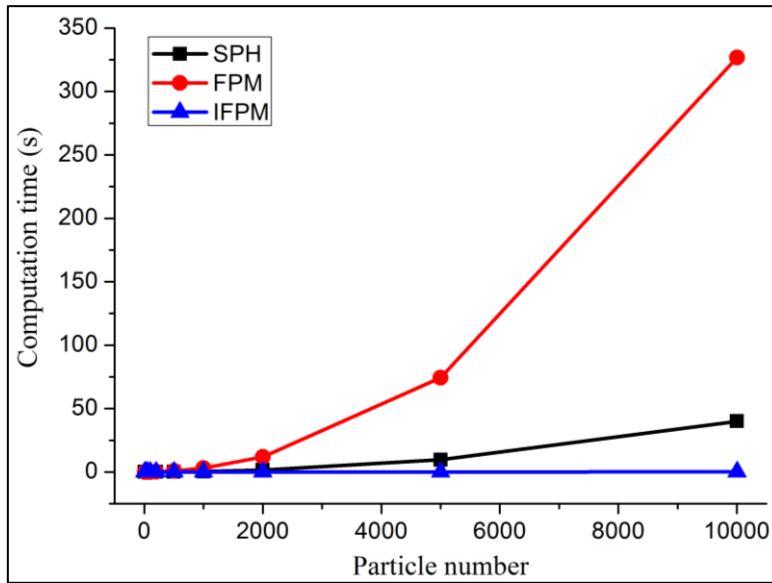
Function type	<i>MSE</i>		
	SPH	FPM	IFPM
$f(x, y) = 1$	0.0357	$1.5834 \times 10^{-32}$	$7.4650 \times 10^{-31}$
$f(x, y) = x + y$	427.7285	$3.8521 \times 10^{-28}$	$2.1792 \times 10^{-27}$

#### 4.2 Computation time analysis

In order to compare the computing time among SPH, FPM and IFPM, we consider the function  $f(x) = x$  in  $[0, 100]$ . The computing time of the three methods with increasing of the number of particles in the computational domain from 11 to 10001 are shown in Figure 3.

It is shown that with increasing of the particle number, the computing time of FPM increases rapidly, and SPH method increases steadily. While the computing time of IFPM proposed in this paper has no obvious change. When the number of particles increases from 11 to 10001, the computing time just increases from 0.0002s to

0.1176s. Therefore, IFPM method can greatly reduce the computing time compared with FPM.



**Figure 3. Comparison on the computing time among the three methods**

#### 4.3 Calculation of high-order functions

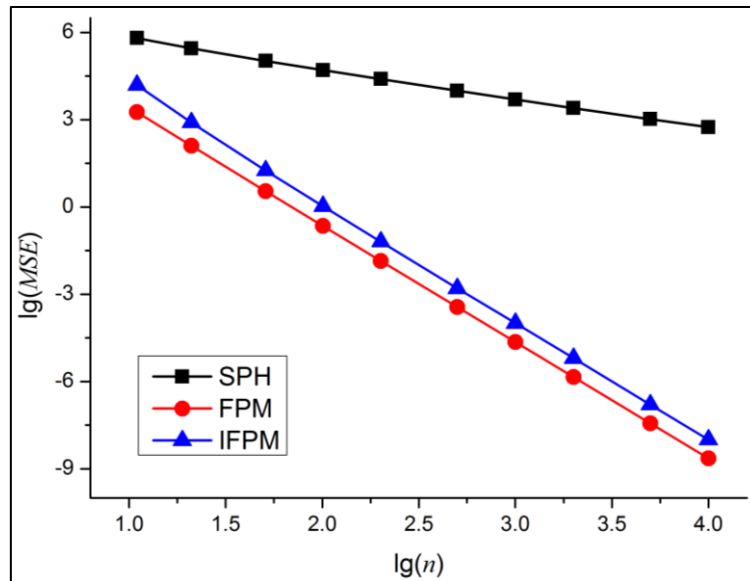
Table 4 shows the *MSE* among SPH, FPM and IFPM for the function  $f(x)=1$ ,  $f(x)=x$  and  $f(x)=x^2$  in  $[0, 100]$  with the number of particles 11 and 10001. It can be found that the accuracy of IFPM clearly decreases when the function times is bigger than 1.

**Table 4. *MSE* of reproduced results by using different methods**

Particle number	Function	SPH	FPM	IFPM
11	$f(x)=1$	0.0089	$2.2411 \times 10^{-33}$	0
	$f(x)=x$	54.6845	$9.7532 \times 10^{-29}$	$9.9513 \times 10^{-33}$
	$f(x)=x^2$	$6.3848 \times 10^5$	$1.8257 \times 10^3$	$1.5455 \times 10^4$
10001	$f(x)=1$	$1.3038 \times 10^{-5}$	$4.9299 \times 10^{-36}$	0
	$f(x)=x$	0.0598	$1.3449 \times 10^{-28}$	$3.7050 \times 10^{-30}$
	$f(x)=x^2$	554.8744	$2.2677 \times 10^{-9}$	$1.0068 \times 10^{-8}$

However, it is also shown in Table 4 that the accuracy of IFPM could be improved by increasing the total number of particles in the computational domain. Figure 4 shows the *MSE* of reproduced results for  $f(x)=x^2$  with increasing of the number of particles by using IFPM, where the data in Figure 4 is the denary logarithm of

original results. It is shown that the  $MSE$  will keep linear decrease with the increased particle number, which means the accuracy of IFPM is controllable.



**Figure 4.**  $MSE$  of reproduced results with the number of particles

#### 4. Conclusions

Based on the numerical tests above, it can be found that the proposed IFPM method could not only keep the high accuracy of FPM in both the interior area and the boundary area for the constant function and the linear function, but also modify the deficiencies of the long computing time and computational instability in the traditional FPM. For the high-order functions, the accuracy of IFPM is controllable, which could be improved by increasing the number of particles. Therefore, IFPM is an effective improvement for traditional FPM.

#### Acknowledgements

This paper is supported by the National Natural Science Foundation of China (Grant No. 11272266) and the graduate starting seed fund of Northwestern Polytechnical University (Grant No. Z2014083).

#### References

- Lucy L B. (1977) Numerical approach to testing the fission hypothesis, *The Astronomical Journal* **82**, 1013-1024.
- 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.
- Liu, G. R. and Liu, M. B. (2003) *Smoothed Particle Hydrodynamics — A Meshfree Particle Method*, World Scientific, Singapore.
- Yang X F, Peng S L, Liu M B. (2014) A new kernel function for SPH with applications to free surface

- flows, *Applied Mathematical Modelling* **38**, 3822-3833.
- Feng D L, Liu M B, Li H Q, et al. (2013) Smoothed particle hydrodynamics modeling of linear shaped charge with jet formation and penetration effects, *Computers & Fluids* **86**, 77-85.
- Liu M B, Zong Z, Chang J Z. (2011) Developments and applications of Smoothed Particle Hydrodynamics, *Advances in Mechanic* **41**, 219-236. (in Chinese)
- Libersky L D, Petschek A G, Carney T C, et al. (1993) High strain Lagrangian hydrodynamics - a three dimensional SPH code for dynamic material response, *Journal of Computational Physics* **109**, 67-75.
- Chen J K, Beraun J E. (2000) A generalized smoothed particle hydrodynamics method for nonlinear dynamic problem, *Computer Methods in Applied Mechanics and Engineering* **190**, 225-239.
- Liu M B, Liu G R, Lam K Y. (2003) A one-dimensional meshfree particle formulation for simulating shock waves, *Shock Waves* **13**, 201-211.
- Xu Fei, Zhao Yan, Yan Rui, et al. (2013) Multidimensional discontinuous SPH method and its application to metal penetration analysis. *International Journal for Numerical Methods in Engineering* **93**, 1125-1146.
- Liu M B, Xie W P, Liu G R. (2005) Modeling incompressible flows using a finite particle method, *Applied Mathematical Modelling* **29**, 1252-1270.
- Zhang G M, Batra R C. (2004) Modified smoothed particle hydrodynamics method and its application to transient problems, *Computational Mechanics* **34**, 137-146.
- Monaghan J J. (1992) Smoothed particle hydrodynamics, *Annual Review of Astronomy and Astrophysics* **30**, 543-574.

# Multiscale modelling of multiple-cracking fracture behaviour of engineered cementitious composites

Ting Huang<sup>1</sup>, †Y.X. Zhang<sup>1</sup>, and \*Chunhui Yang<sup>2</sup>

<sup>1</sup>School of Engineering and Information Technology, University of New South Wales, Canberra,  
Australian Defence Force Academy, ACT 2600, Australia

<sup>2</sup>School of Computing, Engineering and Mathematics, University of Western Sydney, Penrith, NSW  
2751, Australia

\*Presenting author: R.Yang@uws.edu.au

†Corresponding author: Y.Zhang@adfa.edu.au

## Abstract

A multiscale numerical approach is presented in this paper to model the tensile behaviour of engineered cementitious composites. A unit cell model with a random realisation of the matrix cracking strength distribution and fibre volume fraction distribution is employed. Cracks are adaptively introduced within cracked elements by means of extended finite element method along with a cohesive law which defines the relationship between the stress and the crack opening displacement. The developed numerical method is used to predict the tensile properties of an ECC, and good agreement between the numerical results and experimental results from direct tension tests validate the developed numerical method.

**Keywords:** Engineered cementitious composites (ECC), Tensile behaviour, Multiple cracking, Multiscale numerical modelling, Unit cell model

## Introduction

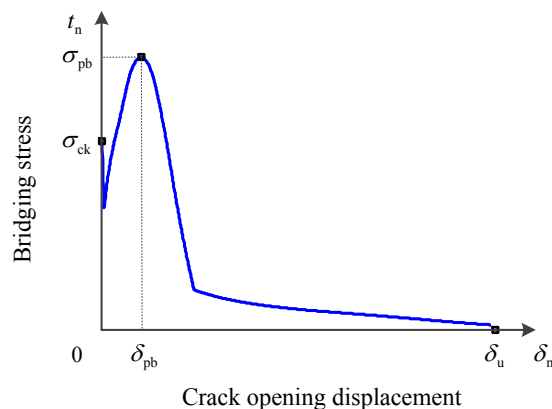
High performance fibre reinforced cementitious composites (HPFRCC) are a special family of fibre-reinforced cement-based composites, and they distinguish from the conventional quasi-brittle fibre reinforced concrete (FRC) by macroscopic pseudo strain-hardening behaviour in tension with multiple cracking developed up to relatively high strain levels. The associated enhanced ductility, energy absorption capacity and toughness compared with FRC are expected to significantly enhance the serviceability and sustainability of cementitious materials. Engineered cementitious composites (ECC) are a unique member of HPFRCC, featuring extraordinary tensile ductility with a low to moderate fibre content. For example, ECC materials reinforced with polyethylene fibre or the polyvinyl alcohol (PVA) fibre with a fibre volume fraction no greater than 2 % exhibited a tensile strain capacity of 3 – 6 % [Li (1998)].

The tensile behaviour of ECC has been most often characterized by results of direct tension or flexure tests. Although some numerical models have been developed over the past few years [Spagnoli (2009); Kabele (2010)] to model the tensile strain-hardening behaviour of ECC, these models could only qualitatively assess the tensile strain capacity of an ECC in comparison to that of conventional FRC. In this paper a multiscale numerical approach based on the unit cell (UC) model is developed to predict the tensile behaviour of ECC. The bridging behaviour of individual fibres is regarded as the composite characteristics at microscale. The formation and opening of one single crack is regarded as a representation of the material behaviour at mesoscale, and it is derived by superimposing the bridging responses of all fibres

involved in the crack-bridging action. The effect of crack bridging provided by matrix and fibres is considered for the individual crack through a cohesive law that relates the bridging stress to the crack opening displacement. Finally, the overall microscopic tensile stress-strain relationship for ECC is determined from the joint response of sequentially emerged individual cracks and the contribution of the uncrack matrix. The UC should be able to accommodate a random number of cracks under the tensile loading. Various micromechanics crack bridging model are available to derive the cohesive law based on the material constituent parameters and fibre/matrix interface parameters. As a result, the presented multiscale approach has the great advantage to link the material parameters at microscale with composite properties at macroscale. The multiscale modelling framework for HPFRCC was first proposed by Kabele [Kabele (2007)] and an analytical method was employed to derive the overall tensile stress-strain relationship of the proposed model. However, in this work the overall tensile stress-strain relationship of the UC model is obtained numerically by means of the extended finite element method (XFEM).

### A cohesive law

It is assumed that cracks can occur at any material point during the loading where the maximum principle stress attains the local matrix cracking strength  $\sigma_{ck}$ . After crack occurs, the crack response is governed by the cohesive law utilizing the cohesive crack concept, which corresponds to a characteristic material response at the mesoscale. In this study, the contributions to the crack bridging effect from both matrix and fibres are taken into account, and therefore, the cohesive law is the superposition of the crack bridging law of plain matrix and the crack bridging law due to fibres [Wang (2005)]. Especially, the fibre bridging law is computed by summing up the force contributions of all fibres bridging the crack. A representative cohesive law of the ECC is shown in Fig. 1 schematically. It should be noted that the cohesive law of ECC possesses a hardening branch prior to the softening, i. e. the ultimate bridging strength  $\sigma_{pb}$  exceeds the matrix cracking strength such that more cracks can be triggered, and this essentially differentiates ECC from conventional quasi-brittle FRCs.



**Figure 1. The cohesive law of ECC**

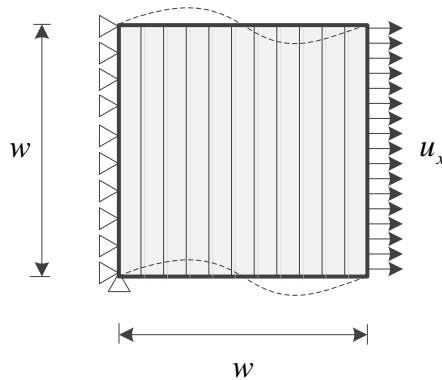
In Fig. 1,  $\delta_{pb}$  is crack opening displacement corresponding to the peak bridging stress and  $\delta_u$  is the ultimate cracking opening displacement.



When fibres on one of the cracks have exhausted their load carrying capacity and became unable to carry increasing load, softening starts on this plane. Consequently, only this crack continues to open while others undergo unloading, which triggers fracture localization. For cracks that undergo unloading, it is assumed that the protruding fibres elastically contract [Kabele (2007)].

**A unit cell (UC) model**

The most common UC approach assumes that the microstructure of composite tends to spatially repeat itself at macroscale, i. e. the whole macroscopic specimen consists of periodically arranged unit cells. Therefore the UC approach is based on describing the microstructure of material in a unit cell, which can statistically represent the macroscopic material point, and periodic boundary conditions are imposed on the boundaries of the UC to capture the effect of the surrounding medium. A UC model as shown in Fig. 2 under uniaxial tension loading is adopted. The UC is meshed by a one-dimensional array of quadrilateral strips along the loading direction, and the matrix strength as well as the fibre volume fraction of each element is randomly generated based on a probability distribution function calibrated from experimental observations. Element size is determined by referring to the minimum spacing  $x_d$  of the cracks, which can be theoretically estimated for the randomly distributed short fibre composites based on the fibre parameters and interface parameters [Wu and Li (1992)]. Consequently, only one crack can occur in an element in one simulation. A crack initiates when the axial stress reaches the matrix strength at an element, and it is introduced directly within the element through the entire middle cross-section perpendicular to the loading direction by means of the XFEM. While the behaviour of uncracked matrix is considered to be linear-elastic, the crack behaves according to the non-linear cohesive law.



**Figure 2. A unit cell model (UC) with boundary conditions for uniaxial tension and the periodic boundary conditions on the top and bottom boundaries (dash lines)**

The macroscopic material properties are then obtained as the volume average of their microscopic counterparts at the UC level as given by [Nguyen et al. (2011)]:

$$\sigma_{xx}^M = \frac{1}{w} f_{R_x}^m \tag{1}$$

$$\varepsilon_{xx}^M = \frac{1}{w} u_x \quad (2)$$

where  $\sigma_{xx}^M$  and  $\varepsilon_{xx}^M$  are macroscopic stress and strain in the  $x$  direction;  $f_{R,x}^m$  stands for the sum of internal  $x$ -forces at all nodes along the right boundary of the UC on which an prescribed displacement  $u_x$  has been imposed.

### Numerical validation and discussion

In this research, the experimental results obtained from the selected mix design named as Mix-3 in literature [Wang (2005)] is used to validate the proposed numerical model. In the test, the interface parameters, the macroscopic tensile stress-strain relationship as well as the matrix flaw size distribution were measured. The direct tension tests were conducted on specimens with a gage length of 180 mm. Here, a UC of a dimension nearly two-thirds of this characteristic length, which can be discretized into 60 elements, is used. The computational macroscopic material properties including the ultimate tensile strength and tensile strain capacity, as well as the experimental results are summarized in Table 1.

**Table 1. Summary of computational and experimental tensile properties**

	Ultimate tensile strength (MPa)	Tensile strain capacity (%)
Prediction	5.04	1.79
Experiment [Wang (2005)]	4.84	1.86

While the measured ultimate tensile strength is 4.84 MPa and the tensile strain capacity is 1.86 %, the computational tensile properties of the ECC are 5.04 MPa and 1.79 % respectively, which are in reasonably good agreement. This demonstrates the effectiveness of the proposed numerical multiscale approach in predicting the tensile properties of ECC.

### Conclusions

In this study, a multiscale numerical approach is developed for modelling of the multiple-cracking fracture behaviour of ECC under direct tension. The macroscopic material properties including the ultimate tensile strength and tensile strain capacity are obtained based on the homogenized response of a UC under uniaxial tension proposed in the work. The UC is able to accommodate a number of cracks accompanying the strain hardening, which is one of the distinct features of ECC. These cracks are adaptively introduced within cracked elements by means of the XFEM, with the effect of crack bridging provided by matrix and fibres implemented through a cohesive law.

The approach was verified by comparison of the predicted tensile properties of an ECC with experimental results obtained from the uniaxial tension tests.

## References

- Kabele, P. (2007) Multiscale framework for modelling of fracture in high performance fibre reinforced cementitious composites, *Engineering fracture mechanics* **74**(1), 194-209.
- Kabele, P. (2010) Stochastic finite element modelling of multiple cracking in fibre reinforced cementitious composites, *Fracture and Damage of Advanced Fibre-reinforced Cement-based Materials*, 155-163.
- Kouznetsova, V. G. (2002) Computational homogenization for the multi-scale analysis of multi-phase materials, PhD thesis, Technische Universiteit Eindhoven.
- Li, V. C. (1998) Engineered cementitious composites - tailored composites through micromechanical modelling, *Fibre Reinforced Concrete: Present and the Future*. N. Banthia, A. Bentur and A. Mufti, Canadian Society for Civil Engineering, Montreal: 64-97.
- Nguyen, V. P., O. Lloberas-Valls, M. Stroeven and L. J. Sluys (2011) Homogenization-based multiscale crack modelling: from micro-diffusive damage to macro-cracks, *Computer Methods in Applied Mechanics and Engineering* **200**(9), 1220-1236.
- Spagnoli, A. (2009) A micromechanical lattice model to describe the fracture behaviour of engineered cementitious composites, *Computational Materials Science* **46**(1), 7-14.
- Wang, S. (2005) Micromechanics based matrix design for engineered cementitious composites, PhD Thesis, University of Michigan.
- Wu, H. C. and V. C. Li (1992) Snubbing and bundling effects on multiple crack spacing of discontinuous random fibre-reinforced brittle matrix composites, *Journal of the American Ceramic Society* **75**(12), 3487-3489.

## Numerical study on the thermodynamic characteristics in a Twin Swirl pulverized coal Combustor

Yinli LIU, \*†Hao TANG, Yongfen WANG

College of Energy and Power Engineering,  
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

\* Presenting author: hao.tang@nuaa.edu.cn

†Corresponding author: hao.tang@nuaa.edu.cn

### Abstract

Numerical simulations were conducted to study the combustion characteristics in a coaxial Twin Swirl Combustor (TSC) burning pulverized coal. Results under different stoichiometric ratios show that volatiles release in between two recirculation zones close to the secondary air inlet, and char burns mostly in the near-wall region due to the centrifugal force brought in by two swirling airflows. Intensive near-wall burning of pulverized coal favors a relatively high wall temperature, contributing to continuous molten slag discharge. An evenly distributed temperature profile with an average of about 1500 K in the chamber is obtained, which is beneficial for low NO<sub>x</sub> emissions. The case with stoichiometric ratios of 0.9 yielded lower NO<sub>x</sub> emission rates near exit while remains high volatile conversion rate and char burnout rate, as compared with the case with stoichiometric ratios of 0.8.

**Keywords:** Twin swirl combustor, Computational Fluid Dynamics (CFD), thermal, NO<sub>x</sub>

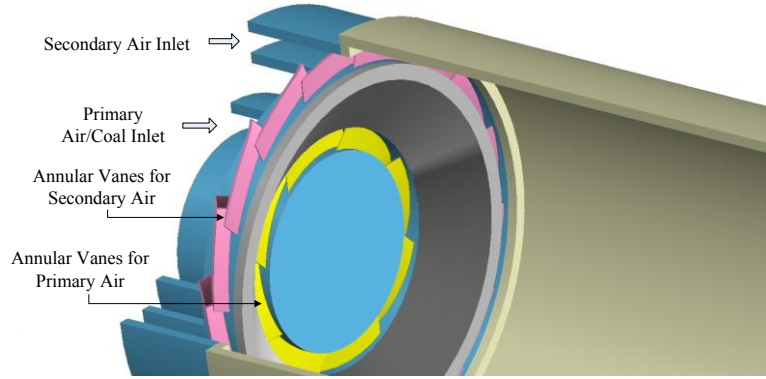
### Introduction

Confined swirling flows are widely used in most the industrial instruments, i.e. internal combustion engines and industrial burners. A strong swirling airflow in the chamber will cause negative-pressure effects, generating internal recirculation zones (IRZ). The existence of recirculation is beneficial for both premix and non-premixed combustion. In pulverized coal combustion, it can help increase the gas recirculation for flame stabilization, and prolonging the travelling time of the coal particles, which is beneficial for reaching high-level burnout rate.

In this paper, a Twin Swirling Combustor (TSC) burning pulverized coal has been proposed. A schematic plot for the inlet structures of TSC is shown in Figure 1. With two swirling airflows, coal particles are burned near the chamber wall in intensive combustion rates. Under the centrifugal force, burnout particles are captured by the chamber wall in the form of fusion slag. With the operation pressure and/or the gravity effect, the ash is removed in the form of molten slag and discharged from the bottom of the combustor to a water quenched slag hopper, where it forms crystal pellets.

CFD simulation was conducted on the Twin Swirl pulverized coal Combustor (TSC) to evaluate the combustion performance inside the chamber in a slagging combustion condition. Results under different stoichiometric ratios were obtained in a fuel-rich condition for restraining of NO<sub>x</sub> generation. C1 and C2 were given to each case, with the stoichiometric ratio ( $\alpha$ ) of 0.9 and 0.8, respectively. Comparative analyses were

made on the combustion performance as well as the pollutant emission rates. A schematic plot for the inlet structures of TSC is shown in Fig. 1. The non-swirling inner primary and outer secondary airflows were turned into two coaxial swirling airflows with the same swirl direction after they flow through the fixed annular vanes. Detailed geometries for TSC can be referred to [Liu and Tang (2014)]. Inlet conditions for case C1 and C2 are described in Table 1. Coal proximate analyses and ultimate analyses results are illustrated in Table 2.



**Figure 1. Schematic plot of inlet structures of TSC**

**Table 1 Operation conditions for different cases**

	Stoichiometric ratio $\alpha$	Air flow rate (kg/s)	Coal feed rate (kg/s)	Temperature (K)	
				Primary air	Secondary air
C1	0.9	5.559e-2	7.676e-3	293	673
C2	0.8	5.559e-2	8.529e-3	293	673

**Table 2 Properties of the pulverized coal**

Proximate analysis (%)				Ultimate analysis (%)				
FC	Volatile	Ash	Moisture	C	H	O	N	S
36.9	45.5	12.9	4.7	77.58	6.57	14.71	1.12	0.02

### Numerical models and mathematical methods

The thermal simulation of the pulverized coal combustion in TSC was performed using a finite volume method. Coal combustion is modeled as a diluted two-phase (solid-gas) reacting flow using an Eulerian-Lagrangian approach. For the gas phase, the governing equations of mass, momentum, species and energy are written in conservative form:

$$\frac{\partial(\rho\Phi)}{\partial t} + \frac{\partial(\rho u_j \Phi)}{\partial x_j} = \frac{\partial}{\partial x_j} \left( \Gamma_\Phi \frac{\partial \Phi}{\partial x_j} \right) + S_\Phi \quad (1)$$

With  $\Phi$ ,  $t$ ,  $u$ ,  $\Gamma_\Phi$  and  $S_\Phi$  denoting Favre-averaged variables, time, velocity diffusion coefficient, and source term, respectively [Muller et al. (2010)].

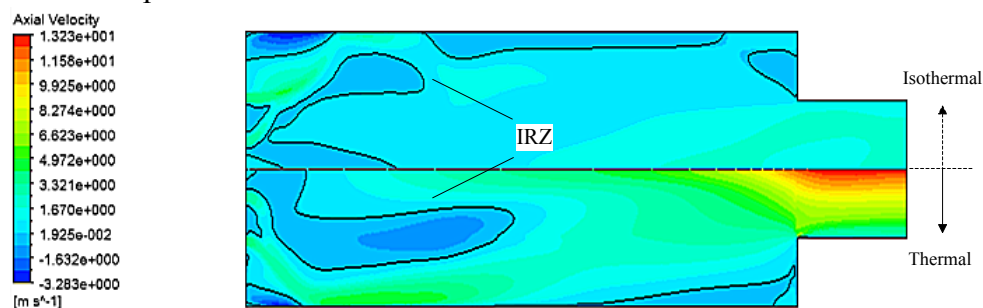
The RNG  $k$ - $\varepsilon$  model [Yakhot and Orszag (1986)] was adopted to simulate the turbulent flows in the chamber. The devolatilisation process of the coal particles was

simulated using the two-competing-rate model [Kobayashi et al. (1976)]. The turbulence-chemistry interaction was modeled using the finite-rate/eddy-dissipation model with the  $\beta$ -Probability Density Function (PDF) methods [Smoot and Smith (1985)]. The Discrete Ordinates (DO) model [Raithby and Chui (1990)] was used to calculate the radiation heat transfer. For the prediction of  $\text{NO}_x$  emission, both the thermal and fuel  $\text{NO}_x$  were calculated, while the prompt  $\text{NO}_x$  was ignored [Zhou et al. (2014)]. The formation of the thermal  $\text{NO}_x$  is modeled by the extended Zeldovich mechanism [Zeldovich (1946)]. Fuel  $\text{NO}_x$  is generated when nitrogen originally bound in the coal particles combines with excess oxygen.

Mesh independence test had been implemented before simulations started. A total number of about 100,000 quadrilateral mesh cells were chosen for all the calculations. The isothermal cases were first simulated with converged results. Then coal particles were added to couple with the continuous phase calculation. Convergence criteria were set to five orders of magnitude reduction and at least 20,000 iterations were carried out to ensure the convergence.

## Results and discussion

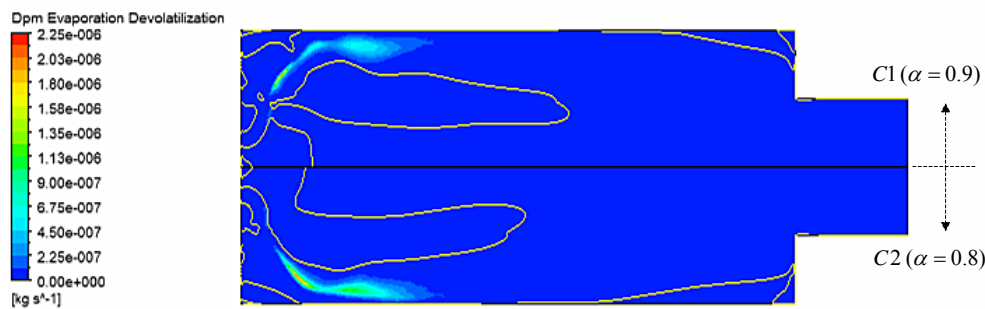
The contours of the axial velocity are shown in Figure 2. Only half of the calculation domain is shown due to axisymmetric flow characteristics. The upper half depicts the axial velocity under isothermal condition, while the lower half describes the axial velocity of C2, which is under thermal condition. The recirculation zones are highlighted with black curves, in which the axial velocity is negative. The IRZ moves more downstream towards the axis direction under thermal condition, and the area of IRZ is larger compared with the isothermal one. There is another recirculation zone near the combustor wall. When combustion process is taken into consideration, this recirculation zone becomes smaller and move downstream. The discrepancy in the shape and location of the IRZ can be due to two aspects. Firstly, the chemical reactions and heat generated during coal combustion have huge impacts on the flow field inside the chamber. Also, when coal particles are added into the chamber, the interaction between the air and coal particles can also affect the aerodynamic field of the continuous phase.



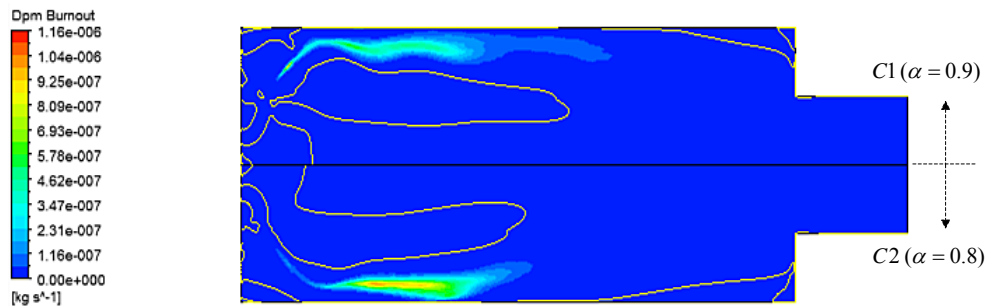
**Figure 2. Contours of axial velocity (upper half, isothermal results; lower half, thermal results)**

Figure 3 compares the particle evaporation/devolatilization rate and particle burnout rate for C1 and C2 are shown in Figure 4. In each figure, the upper half is the contour of C1 ( $\alpha = 0.9$ ) and the lower half is that of C2 ( $\alpha = 0.8$ ). The light yellow outlines in Figure 3 and Figure 4 depict the recirculation zones where the axial velocity is negative. Though the length of the IRZ of C1 is slightly larger than that of C2, the other recirculation zones are almost with same shape and located in same place. As

seen in Figure 3, the place where volatiles release from coal particles is located just in between the two recirculation zones near the secondary air inlet. Moreover, the contour of particle burnout rate shown in Figure 4 indicate that char also burns out in between two recirculation zones, though the char burnout district extends further downstream along the axial direction than that of the particle evaporation/devolatilization zone. As is shown in Figure 3, the contours for particle evaporation/devolatilization rates of C1 and C2 differ little from each other. For C2, a higher evaporation/devolatilization rate is obtained. Similarly, higher burnout intensity is fulfilled in C2, while the length of the char burnout area of C1 is larger than that of C2. The char burnout zones for both cases are located in the vicinity of wall regions, due the centrifugal force aroused by the swirling airflows. Large evaporation/devolatilization rate and char burnout rate near the chamber wall indicates an intensive combustion in the near-wall region, which is beneficial for particle deposition and melting slag discharge.

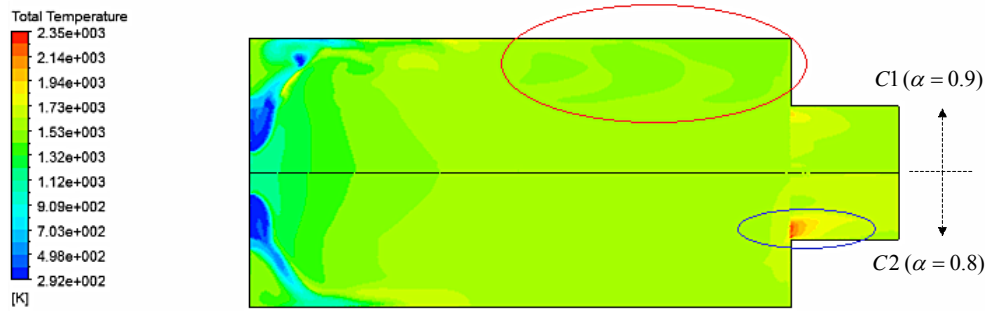


**Figure 3. Contours of particle evaporation/devolatilization rate (upper half, C1; lower half, C2)**



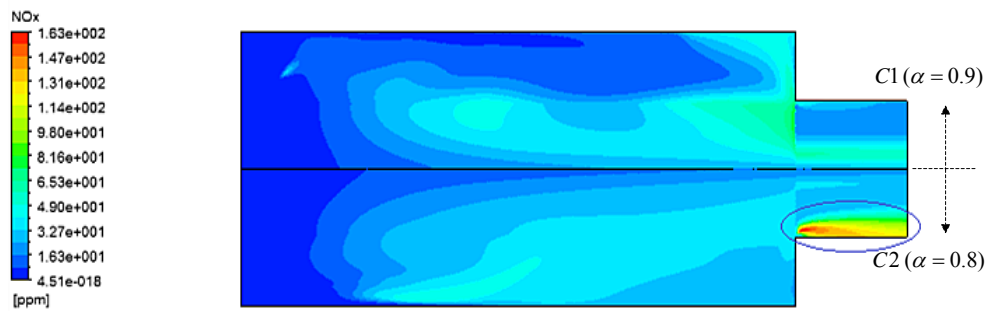
**Figure 4. Contours of particle burnout rate (upper half, C1; lower half, C2)**

The temperature contours in the TSC chamber of C1 and C2 are depicted in Figure 5. A relatively uniform temperature distribution is obtained, which is beneficial for low thermal NO<sub>x</sub> emission. In most part of the chamber for both cases, the temperature can reach up to 1600 – 1700 K. Within this temperature range, slagging combustion can be realized. The temperature profiles in C1 and C2 are similar in most of the parts in the chamber. This is because the same inlet conditions for the air flow, which results in the similar aerodynamic field and combustion performance shown in Figure 2 to 4. For C1, there is a near-exit zone with a relatively lower temperature near the wall (marked with a red ellipse in Figure 5). Moreover, in the corner near the contraction part of C2, there is an area with a temperature peak of 2300K (marked with a blue ellipse in Figure 5), which indicating a large generating rate of thermal NO<sub>x</sub> in this place.



**Figure 5. Contours of axial velocity (upper half, C1; lower half, C2)**

The distribution of  $\text{NO}_x$  in the TSC chamber is shown in Figure 6. Here only thermal and fuel  $\text{NO}_x$  was calculated while prompt  $\text{NO}_x$  was ignored, considering the characteristics of the pulverized coal combustion. As is shown in Figure 6,  $\text{NO}_x$  is mainly generated in the near-axis regions downstream and the upper corner near the contraction part in C1; while for C2,  $\text{NO}_x$  is more evenly distributed along the radial direction. There is a  $\text{NO}_x$  concentration peak in the contraction part near the exit of C2 (highlighted in a blue ellipse in Figure 6). This is contributed by high thermal  $\text{NO}_x$  generation rate in this area, in accordance with the temperature peak shown in Figure 5.



**Figure 6. Contours of axial velocity (upper half, C1; lower half, C2)**

**Table 3 Summary of simulation results**

	Particle Average Residence Time (s)	Average Volatile Conversion Rate (%)	Average Char Burnout Rate (%)
C1	1.00	99.61	93.52
C2	0.93	99.93	92.96
	Average Temperature in TSC Chamber (K)	Average $\text{NO}_x$ Concentration Rate in TSC Chamber (ppm)	Average $\text{NO}_x$ concentration rate @ exit (ppm)
C1	1472.75	21.47	25.89
C2	1516.54	23.21	71.46

Table 3 shows the combustion performance and the  $\text{NO}_x$  generation rates under a fuel-rich condition of C1 and C2. The average particle residence time during combustion is around 1 s, which is favorable for highly efficient coal combustion. In both cases, the average volatile conversion rate is larger than 99.5%, and C1 yields a merely higher average char burnout rate of 93.52% to 92.96%. The average



temperature in the TSC chamber of C2 is about 45 K higher than that of C1. This is because with the same amount of combustion air and almost the same combustion efficiency, more burned coal means more heat is generated. Low NO<sub>x</sub> combustion is realized for both cases, and C1 acquire a relatively lower NO<sub>x</sub> generation rate. The evenly distributed temperature profile in the chamber, along with the reducing atmosphere greatly inhibits the NO<sub>x</sub> generating rate during the pulverized coal combustion. The average NO<sub>x</sub> concentration rate near the exit for C1 is only 25.89 ppm, compared with 71.45 ppm for C2.

## Conclusions

In this paper, the thermodynamics properties of a Twin Swirl Combustor (TSC) burning pulverized coal were studied. Simulation results show that the pulverized coal is burnt efficiently, with an average volatile conversion rate of larger than 99.5% and an average char burnout rate of 93%. Meanwhile, the NO<sub>x</sub> generation rate is reduced in a low level in the chamber. C1 with the stoichiometric ratio of 0.9 yielded lower NO<sub>x</sub> emission rates near combustor exit while remained high volatile conversion rate and char burnout rate, as compared with C2 with the stoichiometric ratio of 0.8.

Two swirling airflows in the chamber form strong swirling and recirculation zones. The particle residence time in the chamber is prolonged for around 1 s, which is beneficial for the highly efficient coal combustion. Most of the coal particles are burned in the vicinity of chamber wall. The increased wall temperature is favorable for ash deposition and molten slag discharge. An evenly distributed temperature profile as well as the fuel rich environment in the TSC chamber can further inhibit NO<sub>x</sub> generation during pulverized coal slagging combustion.

A detailed measurement of the thermal properties in the TSC is essential to further compare and evaluate the performance of the TSC and is expected to be conducted soon.

## Acknowledgement

This work is supported by Funding of Jiangsu Innovation Program for Graduate Education (CXLX11\_0193).

## References

- Liu, Y. L. and Tang H. (2014) Numerical study on the interaction mechanism between swirl and reverse flow rate in a twin swirl combustor, *Advance Materials Research* **960-961**, 341-348.
- Müller, M., Lemp, O., Leiser, S., Schnell, U., Grathwohl, S., Maier, J., and Mönckert, P. (2010). Advanced modeling of pulverized coal combustion under oxy-fuel conditions. In *The 35th International Technical Conference on Clean Coal and Fuel Systems. Clearwater, FL*.
- Yakhot, V. and Orszag S. A. (1986) Renormalization group analysis of turbulence. I. Basic theory, *Journal of Scientific Computing* **1**, 1-51.
- Kobayashi, H., Howard, J. B., and Sarofim, A. F. (1976) Coal devolatilization at high temperatures. In *Symposium (international) on combustion, Cambridge, Massachusetts*.
- Smoot, L. D., & Smith, P. J. (1985). *Coal combustion and gasification*, Plenum Press, New York, USA.
- Raithby, G. D. and Chui, E. H. (1990). A finite-volume method for predicting a radiant heat transfer in enclosures with participating media. *Journal of Heat Transfer*, **112**(2), 415-423.
- Zhou, H., Yang, Y., Liu, H. Z. and Hang, Q. J. (2014) Numerical simulation of the combustion characteristics of a low NO<sub>x</sub> swirl burner: Influence of the primary air pipe, *Fuel* **130**, 168-176.
- Zeldovich, Y. B. (1946) The oxidation of nitrogen in combustion and explosions. *Acta Physicochim. URSS*, **21**(4), 577-628.

## The inadequacy of elastic properties from tensile tests for Lamb wave analysis

W.H. Ong<sup>1</sup>, \*W.K. Chiu<sup>1</sup>, N. Rajic<sup>2</sup> and C. Rosalie<sup>2</sup>

<sup>1</sup>Department of Mechanical & Aerospace Engineering, Monash University, Clayton, Victoria, Australia

<sup>2</sup>Defence Science & Technology Organisation, Aerospace Division, Airframe Diagnostic Systems Group, Fishermans Bend, Victoria 3207, Australia

\*Presenting and corresponding author: wing.kong.chiu@monash.edu

### Abstract

Lamb wave analyses have been conducted on metallic and composite structures. The dispersion characteristics for these plate structures are evaluated using the Rayleigh-Lamb equations and an assumption that the material properties are known. In most studies, the elastic moduli used are determined using standard tensile tests. These properties have been used to adequately model the dispersive behavior of Lamb wave in metallic plates. It has been shown that the use of elastic properties from tensile tests is able to model the dispersive Lamb wave behavior. However, this may not be the case for plates made from woven fibre reinforced plastics. The effects of material properties on the propagation of Lamb wave in fibre reinforced composite plates are significant. In most of these investigations, the elastic properties used are obtained from tensile tests. It will be shown that the use of such properties will lead to significant inaccuracy in the determination of the dispersion characteristics of a composite plate.

**Keywords:** Lamb wave, elastic properties of composite materials

### Introduction

A modelling capability for propagating Lamb waves has always been desirable to gain understanding, reduce costly experimentation and perform analyses not possible experimentally. A key factor to accurate modelling of Lamb wave propagation is correctly defining the properties of the propagating medium. Much of the previous literature was focused on aluminium plates. In all these studies, the elastic properties used in structural analysis are also used in modelling Lamb wave propagation. This assumption was valid because such aluminium is a homogenous and isotropic material. Its elastic property does not exhibit significant strain rate effects.

Recently the use of advanced composite structures, in particular carbon fibre reinforced polymers (CFRP) have become commonplace and as a result accurate modelling in composite plates is now desirable. Unlike aluminum, the elastic properties of the fabricated CFRP plates are affected by the processing conditions and fibre compaction. The advances in the computational and experimental structural analyses have led to the use CFRP component in significant load bearing structures. Copious studies have reported good agreement between these analyses that suggest the validity of the elastic properties used.

Naik (1994) reported that stiffnesses of woven composite panels are governed by weave parameters such as weave architecture, yarn sizes, yarn spacing and yarn crimp [1]. The concept of the repeating unit cell (RUC) was used to help define the elastic properties of woven composites [2]. However, in the work by West [3], they showed that the axial stiffness of woven composite is not dependent on the crimp angle.

Lamb waves are low amplitude, high frequency elastic waves travelling on a medium. The length-scale of Lamb wave in the propagating direction is large compared with the microstructures of metallic plates. However, in this case of woven composite plates, the length-scale of the propagating Lamb wave may be comparable to that of its RUC. To this end, this paper will present a set of results that will highlight the apparent shortcoming in using elastic properties that are determined for structural analyses for analyzing Lamb wave propagation in a woven composite plate.

In this study, a set of results obtained with an aluminium plate will first be presented. The computational analyses of the Lamb wave propagation were conducted using published elastic

properties commonly used in structural analyses. The aim of this part of the work is to validate the computational capability and the experimental scanning laser vibrometry equipment. A quasi-isotropic test plate was fabricated from woven composite material. A series of tensile tests was performed to determine the elastic modulus of the test plate. These results were compared with theoretically derived elastic moduli using the manufacturer has published material properties. These properties were subsequently used to model the propagating Lamb wave on this woven composite plate. These computational results were compared with the experimental test data. This concerted work highlights the need for a new approach to establish a set of elastic properties of woven composite plates.

### Laser vibrometry

The dispersion curves of the test plates were acquired using a laser vibrometry facility described in detail in [4, 5]. The PZT bonded to the plate is made to actuate Lamb waves by a 50V peak-peak drive signal from a Krohn-Hite model 7602 amplifier. The out-of-plane displacements from the Lamb waves are detected on the plate surface by a Polytec OFV 505 laser vibrometer. The positioning of the laser vibrometer relative to the plate is controlled by a stepper motor driven high resolution X-Y table set to 5000 steps per mm.

### Metallic plate

An aluminium plate measuring 300x300x4mm had a Ferroperm Pz27 PZT Ø10mm x 1mm disc placed 75mm from the bottom edge as shown in Figure 1. Also shown is “line 1” which will be used to record the Lamb waves propagating across the plate by LV. Line 1 measured 210mm long and spans from the top edge of the plate to the edge of the PZT disc. During the laser vibrometry scan 256 points were collected along this line. A frequency sweep was conducted at each point using Hanning windowed sinusoids with centre frequencies ranging from 100 kHz to 1 MHz in 100 kHz intervals. At 200 kHz the window resulted in a 5 cycle sinusoid, the window was kept constant in time such that the energy at each frequency would be similar.

The data along line 1 was then transformed using a 2D FFT at each frequency and summated to reveal the full spectrum of Lamb modes shown in Figure 2. The theoretical Lamb wave modes are overlaid on the 2D FFT contour plot in the form of dispersion curves and demonstrate a good agreement. These curves were predicted using DISPERSE with the material properties summarized in Table 1. The properties are from the material library supplied by DISPERSE and the values common across engineering applications (such as FEA software) and literature. The excellent agreement between the theoretical and experimental dispersion curves show:

- a. The use of elastic properties commonly used in structural analyses of aluminium structures can also be used in modelling Lamb wave propagation in aluminium plates.
- b. The veracity and efficacy of the scanning laser vibrometry test facilities.

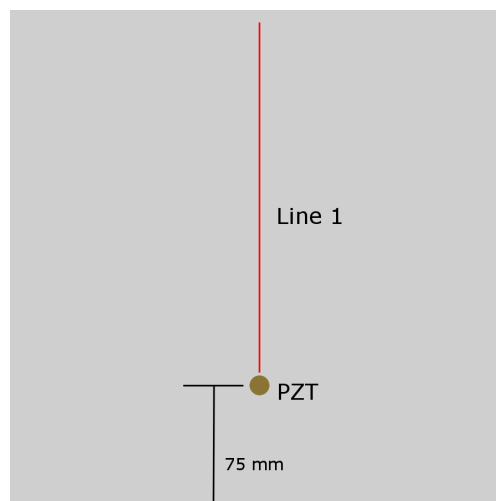


Figure 1 Aluminium plate with PZT and scanned line.

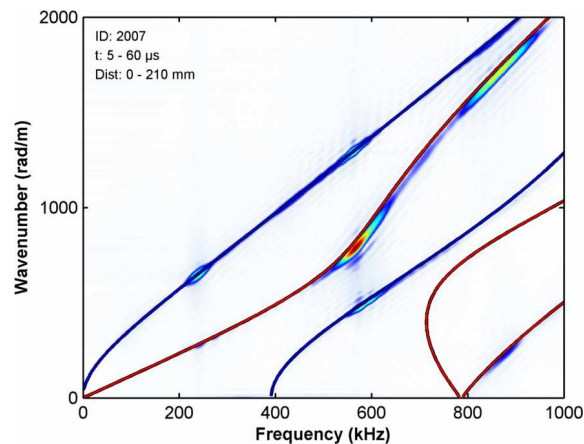


Figure 2 2D FFT on aluminium plate showing agreement with mathematical modelling.

Table 1 Aluminium material properties

Density (kg/m <sup>3</sup> )	2700
Poisson's ratio	0.3375
Young's modulus (GPa)	70.758

### Composite Plate

A composite specimen was fabricated using Hexcel M18 prepreg with woven carbon G939 fabric. The quasi-isotropic test plate used is fabricated from 16 bi-axial plies in the layup  $[0/90, +/-45, +/-45, 0/90, 0/90, +/-45, +/-45, 0/90]_s$ . Upon curing, the plate was trimmed to a planar dimension of 400mm x 400mm. A Ferroperm Pz27 Ø10mm x 1mm PZT disc was placed at the centre of the plate as shown in Figure 3. The dimensions and the mass of the test plate were measured and the material density is calculated to be 1434 kg/m<sup>3</sup>.

A series of uniaxial test specimens were cut from the fabricated composite plate as shown in Figure 4a to measure the stiffness in the 0-degree direction. A unidirectional strain gauge was bonded at the centre of each of the coupon aligned with the tensile direction. An Instron tensile machine equipped with a 2kN load cell was used as shown in Figure 4b. Stress across the strain gauge was calculated by dividing the load by the cross sectional area of the test specimen. The stress-strain curves from the 4 test specimens are shown in Figure 5. The effective moduli of the test specimens were calculated from these results are shown in Table 2. The average modulus was found to be 43.9 GPa.

Using the in-plane modulus provided by Hexcel and estimates for the remaining values (Table 3), the effective modulus of the test plate was calculated using classical laminate theory. This method returned an effective modulus of 46.2 GPa which agreed well with the experimental value of 43.9 GPa, which corresponds to a 5% variation.

The material properties in Table 3 were used to calculate the dispersion curves in DISPERSE. Figure 6 shows the comparison between the experimentally measured dispersion curve compared with those calculated using DISPERSE. The measured wavenumbers of majority of all modes are higher than the theoretical values. Curiously, there is a region where there is apparent agreement between the theoretical and the experimental dispersion characteristics (see Figure 6). The deviation of the experimental and theoretical dispersion curves are greater than 5%.

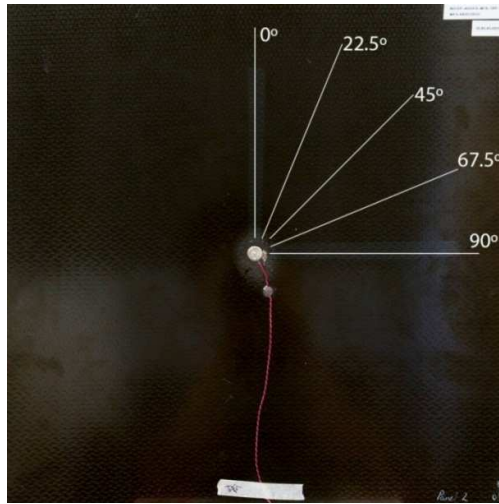
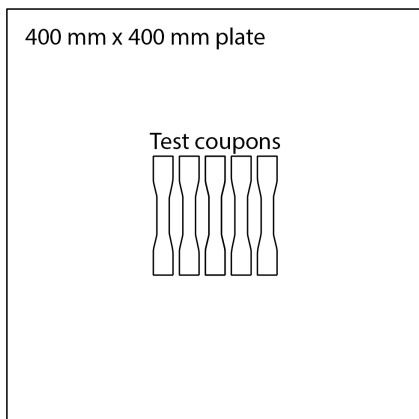


Figure 3 Photograph of specimen showing scan lines.



a)

b)

Figure 4 a) Test coupons cut from plate. b) Test coupon in tension loading.

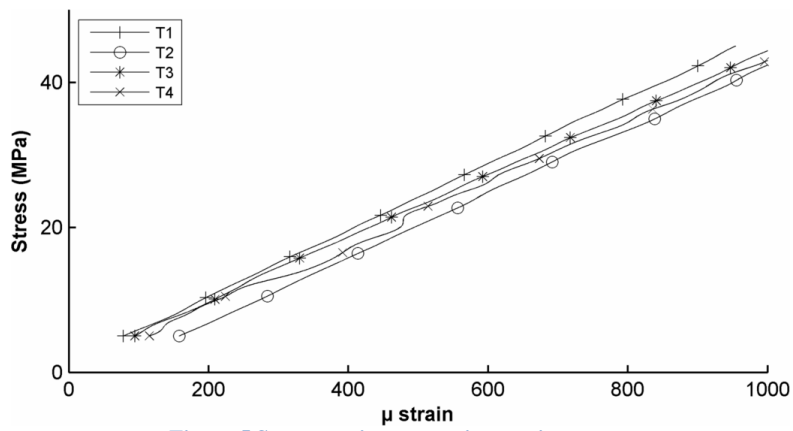


Figure 5 Stress-strain curve using strain gauges.

Table 2 Effective tension moduli.

Specimen	Laminate theory using initial properties	T1	T2	T3	T4	T1-4 Average
Effective modulus (GPa)	46.2	45.7	44.3	43.0	42.7	43.9

Table 3 Published and estimated material properties

$E_{11}=E_{33}$ (GPa)	67
$E_{22}$ (GPa)	8.6
$G_{12}=G_{13}=G_{23}$ (GPa)	5
$\nu_{12}=\nu_{13}=\nu_{32}$	0.35
Ply thickness (mm)	0.269

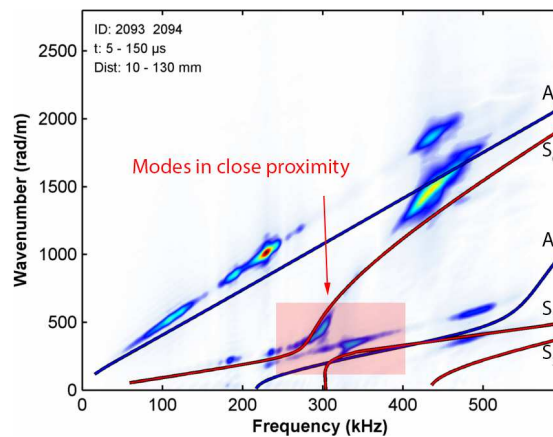


Figure 6 Preliminary reference 2D FFT showing modes in close proximity.

## Discussions

The results presented above can be summarized as follows:

- Excellent agreement between the theoretical and experimental dispersion curves can be obtained when a metallic plate. The efficacy of the laser vibrometry test facility is verified.
- The use of elastic properties as defined by tensile tests is sufficient for modelling the dispersive characteristics of Lamb wave.
- The elastic modulus of a composite specimen determined from a tensile test agreed well with the theoretical modulus calculated using manufacturer's data.

This discussion establishes a possible explanation for the disagreement between the measured and theoretical dispersion curves obtained for the composite panel. According to Naik [1] elastic stiffnesses of woven composite panels are governed by weave parameters such as weave architecture, yarn sizes, yarn spacing and yarn crimp. The concept of the repeating unit cell (RUC) was used to help define the elastic properties of woven composites [2]. The definition of the RUC shows that the need to consider the length-scales when defining the elastic property of woven composite. Figure 7 shows the weave pattern of the test plate used in this paper. The width of the

tow is approximately 1.9 mm. Figure 7 outlines the concept of the RUC for the test plate used in this paper. The important length-scales associated with this RUC are shown in Figure 7.

The measured wavelength of the Lamb wave modes ranges from 3 mm to 25 mm. Whilst these are small compared with the thickness of the plates, it is evident that they are comparable to the length-scale of the RUC as defined in Figure 7. Therefore, it is expected that the perturbations associated with the propagating Lamb wave mode will have length scales that are comparable to the RUC. In this respect, one can expect the yarn crimp angle to have an effect on the elastic property of the woven composite. To this end, one can expect a reduction in the effective elastic modulus that will bring about a reduction in the group velocity of the Lamb wave modes. This is consistent with the results shown in Figure 6. However, the region where modes are in close proximity (see Figure 6) show that merely altering the elastic moduli by trial-and-error is not sufficient. A systematic approach based on the dispersive characteristics of the propagating Lamb wave mode including the higher order modes is required to establish a set of elastic properties that will allow for an accurate representation of the dispersion curves of woven composites.

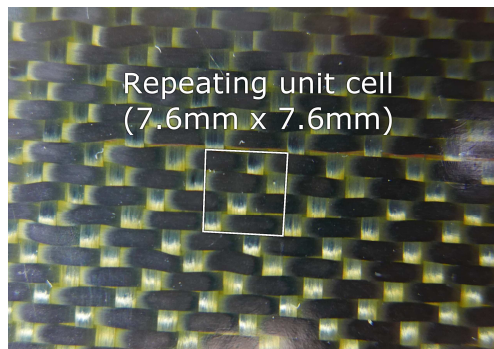


Figure 7 Photograph of CFRP test specimen showing satin weave.

## Conclusions

The efficacy of the non-contact laser vibrometry for studying the propagation of Lamb wave propagation in plate-like structures was first demonstrated. In the experiments where an aluminium plate was used, there was an excellent agreement between the theoretical and experimental dispersion curves. The elastic constants for aluminium used in the theoretical analyses were obtained from standard tensile tests.

Similar investigations were repeated on a composite plate incorporating a satin weave carbon fibre fabric. Firstly, the results from a series of tensile test confirmed that the in-plane elastic constant of were consistent with manufacturer's data. The elastic properties were then used to determine the dispersion curves of Lamb waves propagating in the test specimen. The agreement between the theoretical and experimentally measured Lamb modes was poor.

An explanation for this disagreement is attributed to the length-scale of the RUC of the woven fabric of the CFRP. It was found that the wavelength of the propagation Lamb wave is comparable with the length-scale of the RUC. In this respect, one expects the elastic constant to be less than that determined from standard tensile tests which is consistent with the higher wave number measured. These results show a need for an appropriate test method for determining the elastic constants that is governed by the dispersion relationship of the propagating Lamb waves.

### **Acknowledgements**

This project is funded by the US Navy Office of Naval Research – Global (N62909-14-1-N248). The financial support provided is gratefully acknowledged.

### **References:**

1. Naik, R.A., *Analysis of woven and braided fabric reinforced composites*, 1994, NASA: Hampton, Virginia. CR-194930
2. Dixit, A., H.S. Mali, and R.K. Misra, *Unit cell model of woven fabric textile composite for multiscale analysis*. International Tribology Conference Malaysia 2013, 2013. **68**: p. 352-358.
3. West, A.C. and D.O. Adams, *Axial yarn crimping effects in braided composite materials*. Journal of Composite Materials, 1999. **33**(5): p. 402-419.
4. Ong, W.H. and W.K. Chiu, *Enhancement of Lamb wave-based in situ structural health monitoring through guided local geometry changes*. Structural Health Monitoring, 2013. **12**(4): p. 339-358.
5. Ong, W.H. and W.K. Chiu, *Numerical modelling of scattered Lamb waves through varied damage size in challenging geometry*. Structural Health Monitoring, 2013. **12**(3): p. 278-295.



# Prediction on the onset of global failure of irregular honeycombs under compression

†**Youming Chen, Raj Das, and Mark Battley**

Centre for Advanced composite materials, Department of Mechanical Engineering, University of Auckland, Auckland,  
New Zealand.

†Corresponding author: cyou659@aucklanduni.ac.nz

## Abstract

In the present study, two irregular honeycombs are manufactured by a 3D printer and tested under compression. Experimental observation shows that the first fracture of cell walls is significantly critical, which is indeed the onset of global failure of honeycombs. To predict the onset of global failure of honeycombs, failure criteria based on stress or strain at integration point level need to be used, which is difficult to realize due to stress singularity at cell wall joints. To circumvent this issue, a mesh level, 40 elements within each cell wall, is chosen in FE analysis so that numerical results at cell wall joints are close to engineering solutions. Failure criteria based on von Mises stress, equivalent plastic strain, tensile plastic strain, surface tensile stress and bending moment are then employed to predict the first fracture of cell walls. It is found that failure criteria based on von Mises stress, equivalent plastic strain, tensile plastic strain yield the same predictions which are within 9% of experiment value for strength and within 24% for strain at failure. The cell walls which are predicted to be the most likely to fracture first in the honeycombs agree well with these cell walls that rupture first in experiments.

**Keywords:** Irregular honeycombs, fracture and failure, FE modelling

## Introduction

As manufacturing techniques advance, cellular solids have seen a wide range of applications in automotive, aerospace, aircraft, marine, construction and packaging industries due to their advantage in weight-saving, impact-absorbing, thermal-insulating, noise-abating and so forth. Naturally, the mechanics of cellular solids have become a field of interest for researchers over the past few decades (a comprehensive review can be found in [Gibson and Ashby (1997)]).

For the safe and efficient use of cellular solids, characterization of mechanical properties of cellular solids is vitally important and has been carried out extensively [Deshpande and Fleck (2000); Deshpande and Fleck (2001); Motz and Pippin (2001); Kabir, Saha et al. (2006); Daniel and Cho (2011); Battley, Clark et al. (2013)]. With tests under different loading conditions, the constitutive relations of cellular solids can be established [Deshpande and Fleck (2000); Deshpande and Fleck (2001)], which, however, are generally valid for a specific cellular solid. Generalization of constitutive relations of cellular solids can be found in [Gibson and Ashby (1997)]. To develop tools of predicting the properties of cellular solids, cellular solid modelling have been progressed from simple single cell models [Gibson and Ashby (1982)] to complicated models of random geometry [Ribeiro-Ayeh (2005); Chen, Das et al. (2014)]. Through modelling, it is recognized that the global response of cellular solids is strongly related to their relative density, microstructure and the properties of base materials [Gibson, Ashby et al. (1982); Gibson and Ashby (1982); Ribeiro-Ayeh (2005); Chen, Das et al. (2014)].

The majority of modelling work has been undertaken to study the elastic response of cellular solids [Zhu, Knott et al. (1997); Grenestedt and Tanaka (1998); Simone and Gibson (1998); Simone and Gibson (1998); Grenestedt and Bassinet (2000); Roberts and Garboczi (2001); Gan, Chen et al. (2005); Ribeiro-Ayeh (2005); Li, Gao et al. (2006); Chen, Das et al. (2014)]. In terms of instability analysis of cellular solids which involves geometrical and material nonlinearity, limited work has been conducted. Papka and Kyriakides [Papka and Kyriakides (1994); Papka and Kyriakides (1998)] numerically reproduced the crushing process of aluminum hexagonal honeycombs with integration of geometric imperfections. Zhu etc. [Zhu, Thorpe et al. (2006)] studied the effect of cell irregularity on the high strain compression of Voroni honeycombs using Abaqus Riks method.

In their models cell wall material was assumed to be elastic, therefore the typical stress plateau of honeycombs under compression did not occur and the nonlinearity of response of Voroni honeycombs under high strain arises solely from large deformation of cell walls. Jounaid and Sab. [Jounaid and Sab (2009)] pointed out global buckling (a great number of cell walls buckle simultaneously) disappear in irregular honeycombs, and thus elastic buckling analysis is not able to predict the instability of irregular honeycombs. Kyriakides etc. [Jang, Kyriakides et al. (2010); Gaitanaros, Kyriakides et al. (2012)] simulated the crushing of an open-cell aluminum foam using LS-DYNA. The typical three stages: initial elastic, stress plateau and densification regimes were reproduced. It is noteworthy that the spread of localized band of crushed cells agrees well with experimental observation made using X-ray tomography. Nammi ect. [Nammi, Myler et al. (2010)] modelled the compressive response of a closed-cell aluminum foam using a repeating unit cell constructed from the tetrakaidecahedra structure, and remarkable stress fluctuation in stress plateau regime were predicted. Daxner and Bische [Robert (2005); Daxner, Bitsche et al. (2006); Daxner (2010)] predicted the yield surface of Kelvein and Weaire-Phelan closed-cell foams using macroscopic plastic effective strain and energy dissipated as criteria to define the onset of global failure of foams. Mills [Mills (2010)] investigated the yield surface of polyethylene and polystyrene closed-cell foams using the Kelvin foam model, with foam failure defined by the moment when the yield zone (equivalent plastic strain larger than 0.01) crosses the foam structure.

In above mentioned simulations, cell wall materials are assumed to be elastic-plastic (except [Zhu, Thorpe et al. (2006); Jounaid and Sab (2009)], without consideration of material fracture. This is valid for cellular solids made from ductile material such as aluminum subjected to compression, because no cell wall fracture was experimentally observed during compressive tests. However, for these cellular solids made from brittle or quasi-brittle materials, cell walls fracture at some stage, leading to the failure of cellular solids. Additionally, when cellular solids are subjected to tension or shear, they fail by cell walls tearing apart. Therefore, material failure criteria need to be incorporated into models to predict the failure of cellular solids. To apply material failure criteria, accurate stress or strain (Von Miss stress or equivalent plastic strain) at integration point should be used. However, due to stress singularity at the corners of cells (cell wall joints), stress and strain are actually mesh dependent. Hence, appropriate mesh size should be determined and adopted.

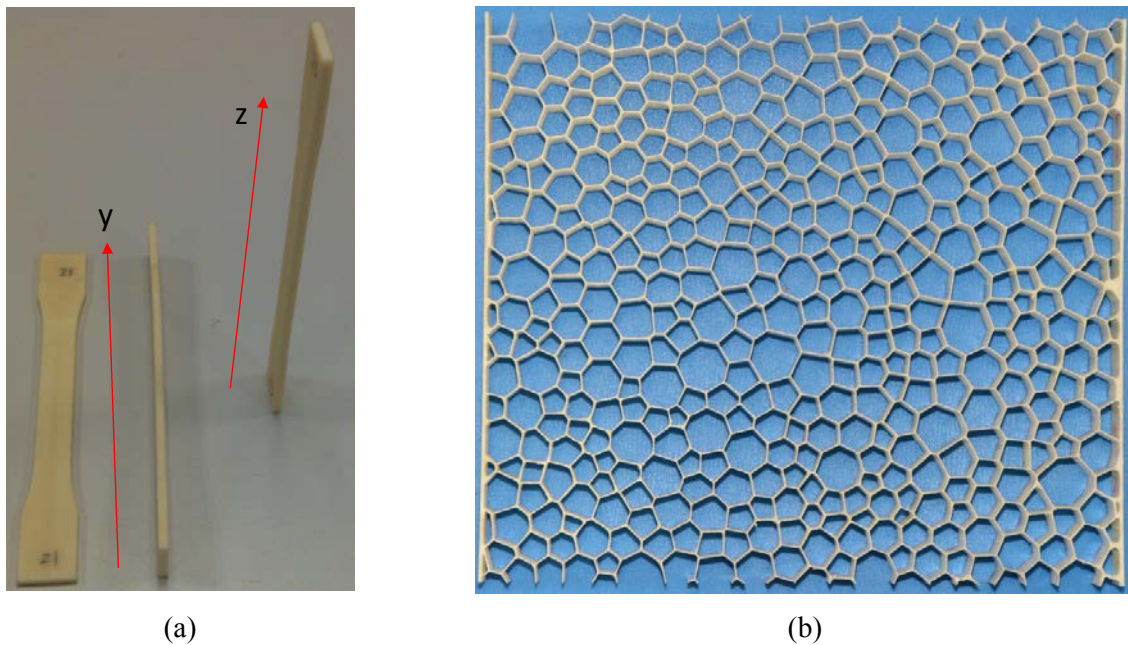
Due to three dimensionality, it is difficult to observe the deformation and failure pattern at cell wall level in foams, and thus is challenging to validate foam models. Honeycombs, as a member of the family of cellular solids, are often employed to study the mechanics of cellular solids due to their geometric simplicity [Papka and Kyriakides (1994); Papka and Kyriakides (1998)]. The deformation of each cell wall in honeycombs can be readily captured by camera and used to validate numerical models [Papka and Kyriakides (1994)]. Nevertheless, in most of cellular solids, variations in microstructures are ubiquitous, for instance, irregular cell shape, non-uniform cell size and wall thickness. Therefore, irregular honeycombs, such as honeycombs with random wall thickness and cell size, are more geometrically representative of realist cellular solids. It is a challenging task to manufacture irregular honeycombs in conventional manner, but with 3D printing it become easier.

This paper aims to predict the onset of global failure of irregular honeycombs under compression which fail by cell wall fracture. To accomplish this, two irregular honeycombs are manufactured by a 3D printer and tested under compression firstly. During the tests, progressive cell wall fracture is captured by a camera positioned squarely toward the specimens. The mechanical properties of base material of the honeycombs are measured with printed dogbone specimens and subsequently incorporated into finite element (FE) models. Static analyses are conducted in Abaqus for these two honeycombs. Then different failure criteria are utilized to predict the onset of global failure of the honeycombs, and finally these predictions are compared.

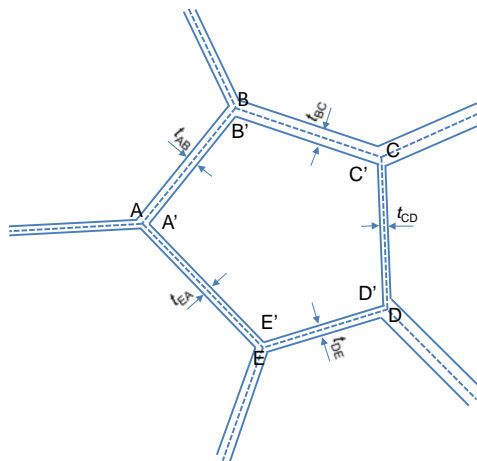
## Experiments

In the present study specimens were manufactured by a 3D printer (Projet HD 3500 Plus), which has a net build volume of  $203 \times 178 \times 152$  mm. The build material VisiJet M3 X is utilized. The highest printer resolution of  $16 \mu\text{m}$  layers ( $750 \times 750 \times 1600$  DPI (xyz)) was chosen. As 3D printers build parts by stacking a layer of materials on top of the previous layer, the properties of printed materials may be anisotropic and different from that of raw materials, and thus the properties of the printed materials still need to be measured. To this end, dogbone specimens of type I were printed in three orthogonal orientations (see Figure 1a) and tensile tests were conducted with them.

Two irregular honeycombs of uniform wall thickness but having cell size (diameter of incircle of cells) following prescribed log-normal distributions (see Figure 1b) were printed here. To print these specimens, the geometry need be created first, which is done in the software Abaqus using python scripts. Firstly, a precursor honeycomb without wall thickness (dashed line in Figure 2) is created, which is indeed a 2D Laguerre tessellation. The details of the generation of 2D Laguerre tessellations are given in [Chen, Das et al. (2014)]. Once the precursor honeycomb is generated, the points, lines, and cells are labelled and stored in a hierarchical order, so that the topologies can be readily known, i.e., which lines construct a cell and which points form a line. Thereafter, cell wall thickening is performed for all the cells one by one. To illustrate the thickening process, we take a cell ABCDE (see Figure 2) as an example. First we determine the inner normal of each edge of the cell ABCDE. As the thickness of each wall is given, five linear equations which represent lines coincidental with inner edges ( $A'B'$ ,  $B'C'$ ,  $C'D'$ ,  $D'E'$  and  $E'A'$ ) can be developed. Solving a set of two linear equations which represent two adjoining inner edges can yield the coordinates of an inner vertex, such as  $A'$ . Following the same process by rotation, the coordinates of the other four inner vertices ( $B'$ ,  $C'$ ,  $D'$ ,  $E'$ ) are calculated. Each cell is analysed through this process, and then all the edges of cell walls are determined. With all the topological data obtained by this process, a sketch is drawn and extruded to 3D geometry in Abaqus. Two thin plates are added at two edges (see Figure 1b) to protect cell walls at edges from localized damage due to contact with compression platens.



**Figure 1: (a) printing orientations of dogbone specimens; (b) printed regular honeycomb**



**Figure 2: Schematics of cell wall thickening**

As specimen size is limited by the maximum building volume of the 3D printer, the more the cells in the specimens are, the smaller the cell size is and the thinner the cell walls are, which make it difficult for the camera to capture deformation at cell wall level and for the 3D printer to print. Here the average cell size is chosen to be 4.46 mm, with a standard deviation of 1 mm, and cell wall thickness is set to be 0.3 mm. The specimens are 170 mm wide, 190 mm tall and 25 mm thick. Compressive tests were subsequently carried out with a synchronized camera positioned squarely toward the specimens to capture deformation in honeycombs.

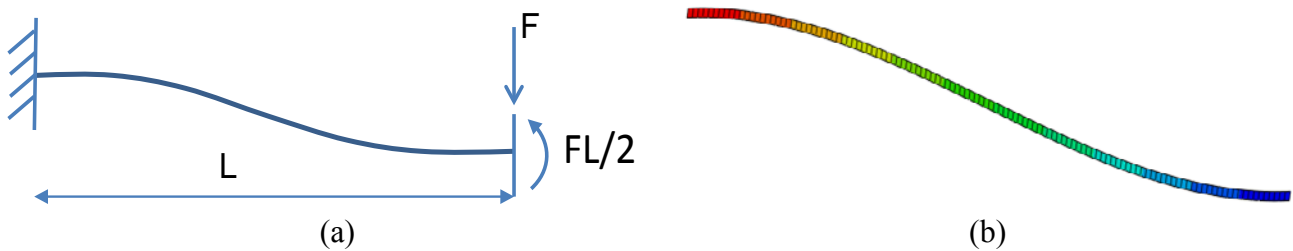
### Computational models

In modelling of compression of ductile honeycombs, it is not necessary to accurately calculate stress and strain at integration point level since the predicted global response is generally not sensitive to them. Nevertheless, it is critical for brittle honeycombs, because stress and strain at integration point level determine whether or not cell walls rupture and thus affect the prediction of global response of honeycombs. However, it is impossible to calculate the exact stress and strain at cell wall joints due to stress singularity. Engineering solutions are more effective and practical in predicting beam failure than exact solutions. For example, Davidge [Davidge (1979)] pointed out that brittle beams rupture when maximum surface tensile stress ( $\sigma_{\max}$ ) is 1.1 times the ultimate tensile strength of beam material, maximum surface tensile stress is calculated as

$$\sigma_{\max} = \frac{6M_{\max}}{bt^2} \quad (1)$$

where  $M_{\max}$  is maximum bending moment, and  $b$  and  $t$  are the width and thickness of the beam, respectively. In a compressed honeycomb, cell walls are basically under the conditions similar to the beam shown in Figure 3a. If we adjust mesh size in FE analysis to make numerical results (surface stress and bending moment) match engineering solutions, then the numerical results at integration points can be utilized to determine cell walls fracture. To find out the mesh size, a mesh sensitivity study is conducted for the beam shown in Figure 3a. The beam is modeled using shell elements (S4R), with one element along width direction, as shown in Figure 3b. Beam material is assumed to be elastic. One end of the beam is fixed and the other is constrained from rotating and pushed downward a third of length of the beam. Three beam length-to-thickness ratios, 10, 20 and 30, are considered. The maximum surface stress and bending moment are read at the integration points. The engineering solutions for surface stress and bending moment are calculated by Eq. (1) and Eq. (2).

$$M_{\max} = \frac{FL}{2} \quad (2)$$



**Figure 3: (a) beam with one end fixed and one end constrained from rotating subjected to a load at the free end; (b) FE model of the beam**

The honeycombs are modelled by shell element (S4R) as well. The geometry of shell models is extruded from the precursor sketches mentioned in the experiments section. Instead of full model, the compression of honeycombs is seen as plain strain and thus only a slice of honeycomb (0.5 mm thick) is modelled, with one element along thickness direction. All elements are constrained moving along honeycomb thickness direction. The tensile properties measured with printed dogbone are incorporated into models. Since the simulations is up to the point when cell wall start to break and no cell wall contact occurs before that, static analyses with large deformation considered are conducted.

## Results and discussion

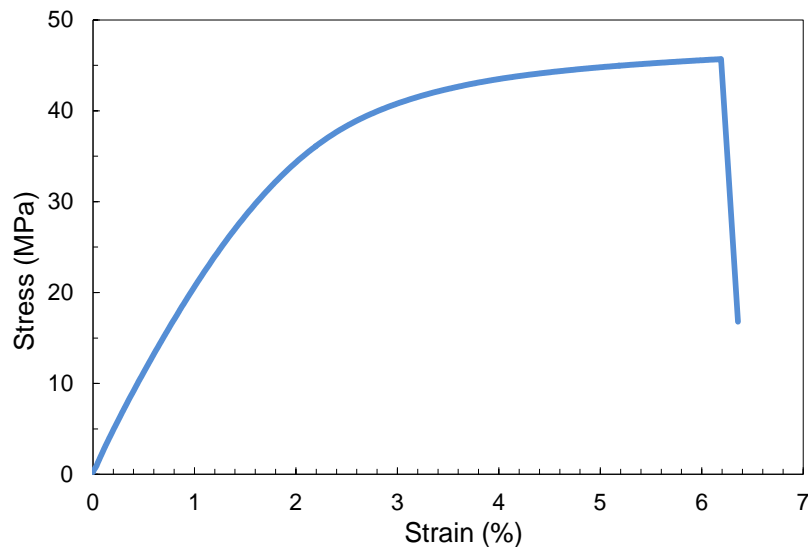
In this section, the tensile properties measured with dogbone specimens are presented firstly. Then the compressive response of the honeycombs is described. The results of mesh study are subsequently illustrated. Finally, the predictions on the onset of honeycomb failure with different criteria are developed and discussed.

### Experimental results

Table 1 lists the measured modulus, ultimate tensile strength and elongation at break of material printed in different orientations. It can be seen that the printed specimens give Young's modulus reasonably closed to that of raw material. The ultimate stresses of the specimens printed standing up along y direction and lying along y direction are slightly smaller than that of raw material, while the specimens printed along z direction exhibit ultimate stress 35.5% lower than that of raw material, which is attributable to the way of 3D printer making part by stacking materials. As honeycombs are built up along thickness direction, the properties measured with specimens printed standing up along Y direction, as shown in Figure 4, is adopted in numerical models.

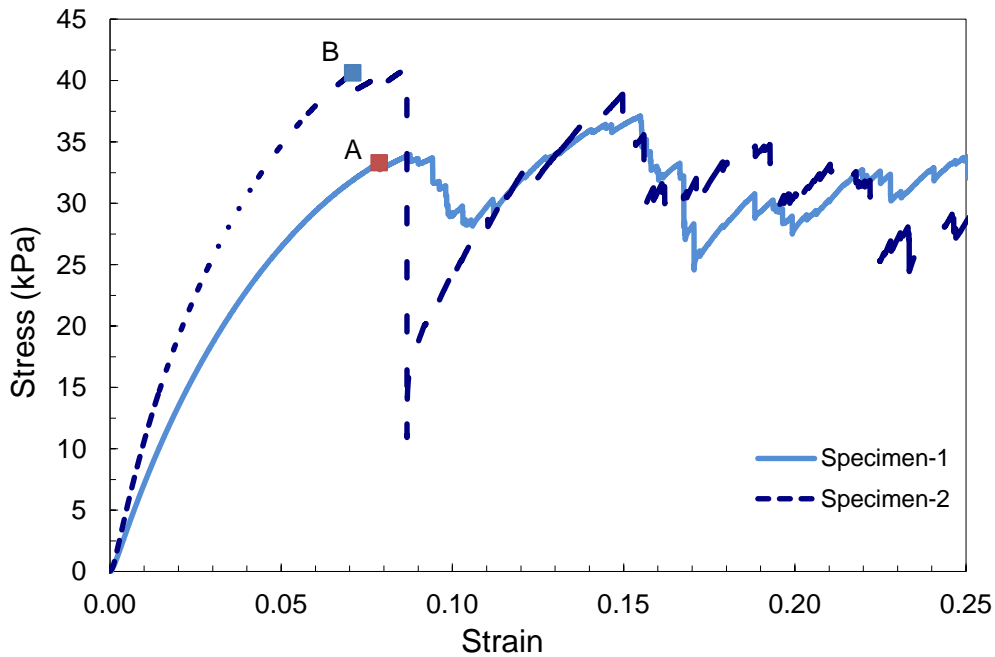
**Table 1. Measured properties of printed dogbone specimens**

	Modulus (Chord 0.05%-0.25%) (GPa)	Ultimate tensile strength (MPa)	Elongation at break (standard) (%)
Lying along Y direction	$2.05 \pm 0.06$	$44.3 \pm 0.36$	$3.98 \pm 0.07$
Standing up along Y direction	$2.34 \pm 0.03$	$45.9 \pm 0.36$	$5.71 \pm 0.42$
Z direction	$2.27 \pm 0.03$	$31.6 \pm 2.57$	$1.98 \pm 0.30$
Raw material	2.17	49.0	8.3

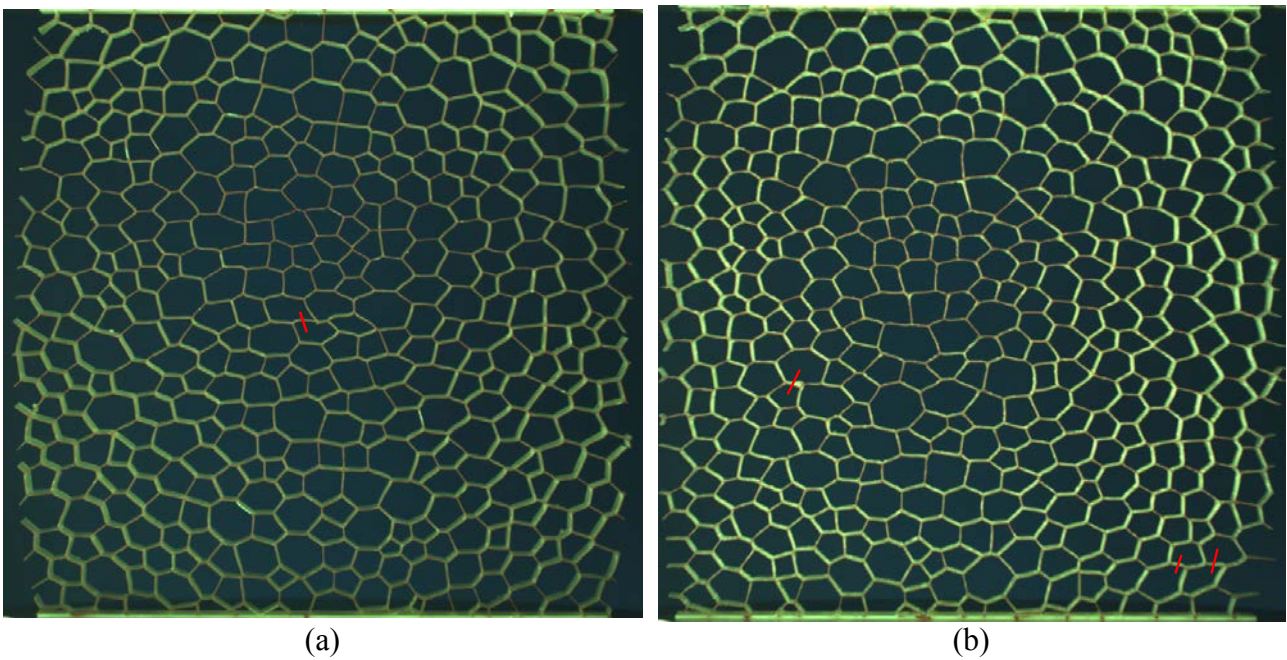


**Figure 4: Stress-strain curve of the dogbone specimens printed standing up along Y direction**

During the compression, it was observed that cell walls fracture progressively. Figure 5 shows the compressive stress-strain of the two irregular honeycombs. Each drop in stress in Figure 5 corresponds to one or multiple cell wall fracture. The first fracture of cell wall occurs at a strain of 0.078 (marked as point A in Figure 5) for specimen-1 and at a strain of 0.071 (marked as point A in Figure 5) for specimen-2, giving failure stresses 40.6 kPa and 33.3 kPa, respectively. It is noteworthy that the first fracture of cell walls are important as it marks the onset of global failure of honeycombs. The configurations of specimen-1 and specimen-2 corresponding to point A and B are shown in Figure 6, and the cell walls that rupture at the moment are marked by a line.



**Figure 5: Compressive stress-strain curves of the honeycombs**

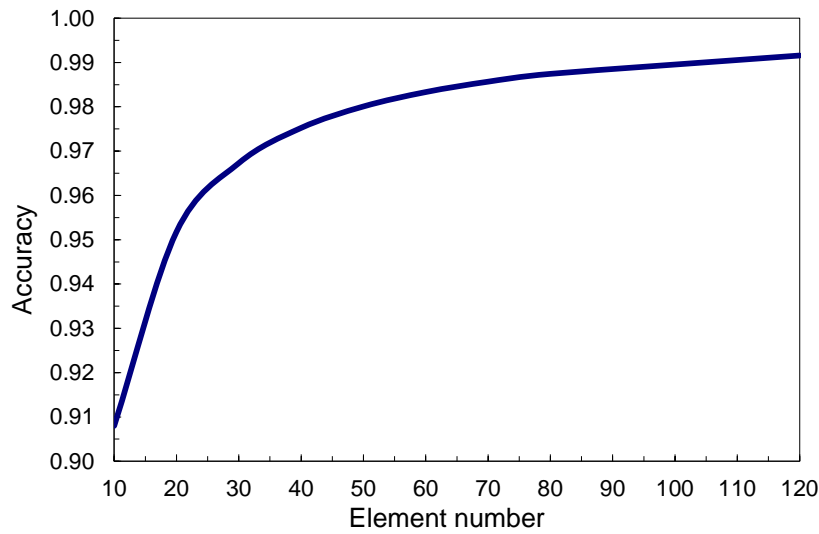


**Figure 6: Configurations of specimens: (a) specimen-1;(b) specimens 2 at the point when the first fracture of cell walls occurs**

*Numerical results*

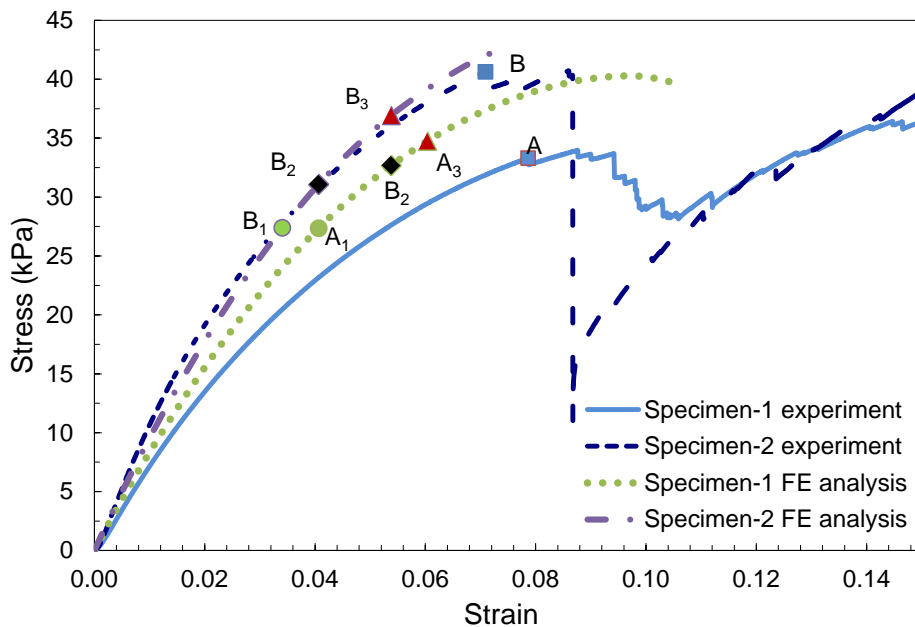
As to the mesh sensitivity study with the beam shown in Figure 3a, the maximum surface tensile stress and bending moment at the fixed end obtained by FE analysis are compared against engineering solutions here. Figure 7 shows the variation of ratio of numerical results to engineering solutions with element number in the beam. The ratio increases with increasing element number. With 40 elements, numerical results are 97.5% of engineering solutions. Additionally, it is found that the variation of this ratio with element number is insensitive to beam configuration (thickness-

length ratio). Considering computational cost and accuracy, the mesh level of 40 elements within each cell wall is chosen in the subsequent honeycomb simulations.



**Figure 7: Variation of accuracy of stress prediction at the ends of the beam with element number**

Figure 8 shows the comparison of compressive stress-strain curves of honeycombs between experiments and FE analysis. It can be seen that for specimen-2, the predicted stress-strain curve agrees well with experimental counterpart before the first fracture of cell wall (point B), while the predicted stress for specimen-1 is a little larger than experimental value, which could be attributable to imperfections induced during manufacturing.



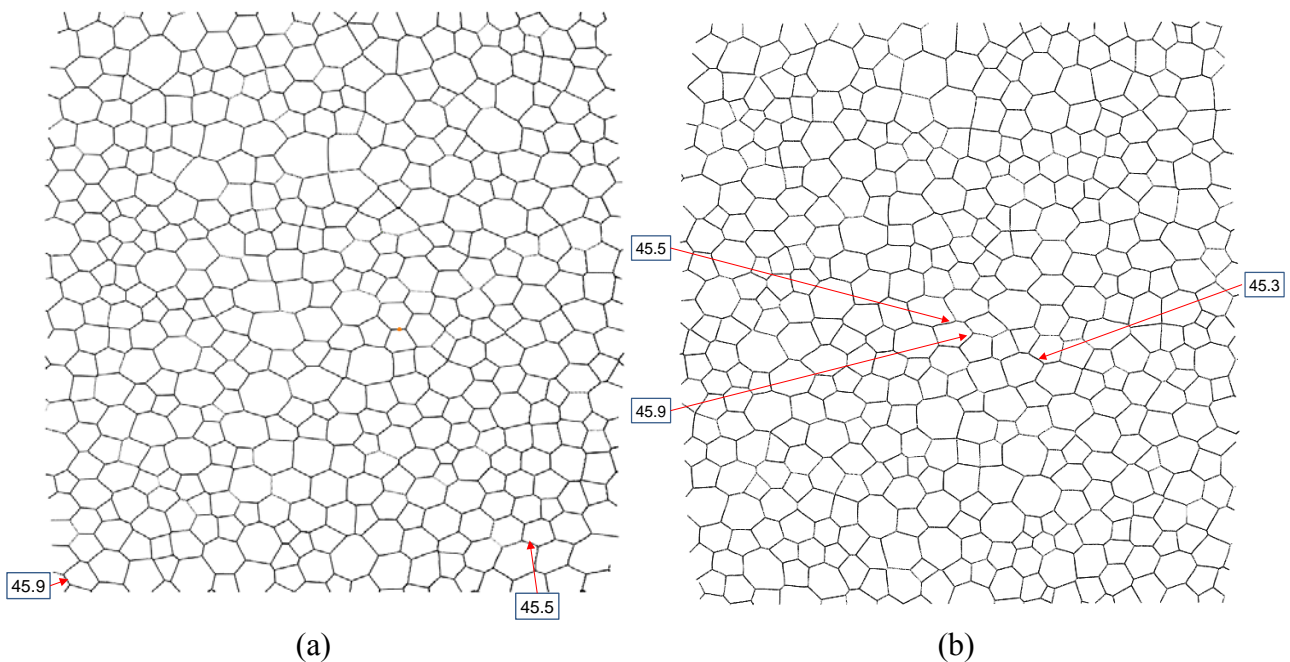
**Figure 8: Comparison of stress-strain curve between experiments and FE analysis**

To predict the first fracture of cell walls, criteria based on tensile stress, bending moment, von Mises stress, tensile plastic strain and equivalent plastic strain are used. When tensile stress, bending moment, von Mises stress, tensile plastic strain and equivalent plastic strain at an integration point of a cell wall reach specified values, the cell wall fracture and the honeycomb fails. In order to establish failure criteria, the data from tensile tests with printed dogbone are

employed there, namely, ultimate stress (45.9 kPa) and plastic strain (0.04) at breaking point. All the criteria are listed in Table 2. For bending moment criterion, as mentioned in the computational model section, 1.1 times the ultimate tensile stress are used, and bending moment here is actually bending moment force per unit width. With these criteria, The onset of failure of honeycombs are predicted and marked in Figure 8, with points A<sub>1</sub> and B<sub>1</sub> by bending moment criterion, points A<sub>2</sub> and B<sub>2</sub> by tensile stress criterion and points A<sub>3</sub> and B<sub>3</sub> by von Mises stress, equivalent plastic strain and tensile plastic strain criteria. The corresponding strength and strain at failure of the honeycombs are listed in Table 2. It is noticeable that criteria based on von Mises stress, equivalent plastic strain and tensile plastic strain yield the same results, which are the closest to experimental results, within 9% for strength and within 24% for strain at failure. Figure 9 shows all the cell walls which have maximum von Mises stress larger than 45 at the point of honeycomb failure. These cell walls are the most vulnerable and have the greatest potential to fracture first in the honeycombs, which agree well with the cell walls that fracture first in experiments.

**Table 2. Strength and strain at failure of the honeycombs predicted by different criteria**

Variables	Criteria	Predicted strength (kpa)		Predicted strain at failure	
		Specimen-1	Specimen-2	Specimen-1	Specimen-2
Tensile stress ( $\sigma_{\max}$ ) (kPa)	$\sigma_{\max}=45.9$	32.7	31.1	0.054	0.041
Bending moment ( $M_{\max}$ ) (N)	$M_{\max}=45.9 \times 1.1 \times 0.3^2 / 6$	27.4	27.4	0.041	0.034
von Mises stress ( $\sigma_{\text{mises}}$ ) (kPa)	$\sigma_{\text{mises}}=45.9$	34.8	36.9	0.060	0.054
Tensile plastic strain ( $\epsilon_p$ )	$\epsilon_p=0.4$	34.8	36.9	0.060	0.054
Equivalent plastic strain ( $\epsilon_{pe}$ )	$\epsilon_{pe}=0.4$	34.8	36.9	0.060	0.054
Experimental results		33.3	40.6	0.078	0.071



**Figure 9: Configurations of (a) specimen-1; (b) specimen-2 at failure point in FE analysis and cell walls which may fracture first**



## Conclusions

This paper concerns with predicting the onset of failure of irregular honeycombs. Due to stress singularity, the exact stress and strain at cell wall joints cannot be calculated. To circumvent this issue, a mesh level with 40 elements within each cell wall is adopted in FE models so that the calculated stress and strain at cell wall joints are close to engineering solutions. Then tensile stress, plastic strain, bending moment, von Mises stress and equivalent plastic strain at integration point level are used to predict the global failure of honeycombs. It is found that criteria based on von Mises stress, equivalent plastic strain and tensile plastic strain give the same predictions, which are the closest to experimental values, within 9% for strength and within 24% for strain at failure. The cell walls which are predicted to have the greatest potential to fracture first in the honeycombs agree well with these cell walls that rupture first in experiments. In summary, with a mesh level of 40 elements in each cell wall and von Mises stress, equivalent plastic strain and tensile plastic strain criteria, the onset of global failure of brittle honeycombs can be reasonably predicted. However, the deviation of predicted strain at failure from experiment value is not negligible. More specimens will be printed and tested in the future to further confirm the modelling.

## References

- Battley, M. A., A. M. Clark, et al. (2013). "Shear strength of sandwich core materials subjected to loading rates relevant to water slamming." Journal of Reinforced Plastics and Composites: 0731684413509424.
- Chen, Y., R. Das, et al. (2014). "Effects of cell size and cell wall thickness variations on the stiffness of closed-cell foams." International Journal of Solids and Structures.
- Daniel, I. and J.-M. Cho (2011). "Characterization of anisotropic polymeric foam under static and dynamic loading." Experimental mechanics **51**(8): 1395-1403.
- Davidge, R. W. (1979). Mechanical behaviour of ceramics, CUP Archive.
- Daxner, T. (2010). Finite element modeling of cellular materials. Cellular and Porous Materials in Structures and Processes, Springer: 47-106.
- Daxner, T., R. D. Bitsche, et al. (2006). "Space-filling polyhedra as mechanical models for solidified dry foams." Materials Transactions **47**(9): 2213.
- Deshpande, V. and N. Fleck (2001). "Multi-axial yield behaviour of polymer foams." Acta Materialia **49**(10): 1859-1866.
- Deshpande, V. S. and N. A. Fleck (2000). "High strain rate compressive behaviour of aluminium alloy foams." International Journal of Impact Engineering **24**(3): 277-298.
- Deshpande, V. S. and N. A. Fleck (2000). "Isotropic constitutive models for metallic foams." Journal of the Mechanics and Physics of Solids **48**(6-7): 1253-1283.
- Gaitanaros, S., S. Kyriakides, et al. (2012). "On the crushing response of random open-cell foams." International Journal of Solids and Structures **49**(19-20): 2733-2743.
- Gan, Y. X., C. Chen, et al. (2005). "Three-dimensional modeling of the mechanical property of linearly elastic open cell foams." International Journal of Solids and Structures **42**(26): 6628-6642.
- Gibson, L., M. Ashby, et al. (1982). "The mechanics of two-dimensional cellular materials." Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **382**(1782): 25-42.
- Gibson, L. J. and M. Ashby (1982). "The mechanics of three-dimensional cellular materials." Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **382**(1782): 43-59.
- Gibson, L. J. and M. F. Ashby (1997). Cellular solids : structure and properties. Cambridge ; New York, Cambridge University Press.
- Grenestedt, J. L. and F. Bassinet (2000). "Influence of cell wall thickness variations on elastic stiffness of closed-cell cellular solids." International Journal of Mechanical Sciences **42**(7): 1327-1338.
- Grenestedt, J. L. and K. Tanaka (1998). "Influence of cell shape variations on elastic stiffness of closed cell cellular solids." Scripta Materialia **40**(1): 71-77.
- Jang, W. Y., S. Kyriakides, et al. (2010). "On the compressive strength of open-cell metal foams with Kelvin and random cell structures." International Journal of Solids and Structures **47**(21): 2872-2883.
- Jouneid, F. and K. Sab (2009). Elastic Buckling of 2-D Random Honeycombs: Does a Representative Volume Element Exist? IUTAM Symposium on Mechanical Properties of Cellular Materials. H. Zhao and N. A. Fleck, Springer Netherlands. **12**: 77-86.
- Kabir, M. E., M. C. Saha, et al. (2006). "Tensile and fracture behavior of polymer foams." Materials Science and Engineering: A **429**(1-2): 225-235.
- Li, K., X. L. Gao, et al. (2006). "Effects of cell shape and strut cross-sectional area variations on the elastic properties of three-dimensional open-cell foams." Journal of the Mechanics and Physics of Solids **54**(4): 783-806.

- Mills, N. (2010). "Deformation mechanisms and the yield surface of low-density, closed-cell polymer foams." Journal of Materials Science **45**(21): 5831-5843.
- Motz, C. and R. Pippan (2001). "Deformation behaviour of closed-cell aluminium foams in tension." Acta Materialia **49**(13): 2463-2470.
- Nammi, S., P. Myler, et al. (2010). "Finite element analysis of closed-cell aluminium foam under quasi-static loading." Materials & Design **31**(2): 712-722.
- Papka, S. D. and S. Kyriakides (1994). "In-plane compressive response and crushing of honeycomb." Journal of the Mechanics and Physics of Solids **42**(10): 1499-1532.
- Papka, S. D. and S. Kyriakides (1998). "Experiments and full-scale numerical simulations of in-plane crushing of a honeycomb." Acta Materialia **46**(8): 2765-2776.
- Ribeiro-Ayeh, S. (2005). Finite element modelling of the mechanics of solid foam materials, Karlstad University.
- Robert, B. (2005). Space-filling polyhedra as mechanical models for solidified dry foams. PhD, Vienna University of Technology.
- Roberts, A. P. and E. J. Garboczi (2001). "Elastic moduli of model random three-dimensional closed-cell cellular solids." Acta Materialia **49**(2): 189-197.
- Simone, A. and L. Gibson (1998). "The effects of cell face curvature and corrugations on the stiffness and strength of metallic foams." Acta Materialia **46**(11): 3929-3935.
- Simone, A. and L. Gibson (1998). "Effects of solid distribution on the stiffness and strength of metallic foams." Acta Materialia **46**(6): 2139-2150.
- Zhu, H., J. Knott, et al. (1997). "Analysis of the elastic properties of open-cell foams with tetrakaidecahedral cells." Journal of the Mechanics and Physics of Solids **45**(3): 319-343.
- Zhu, H. X., S. M. Thorpe, et al. (2006). "The effect of cell irregularity on the high strain compression of 2D Voronoi honeycombs." International Journal of Solids and Structures **43**(5): 1061-1078.

## Development of an anatomically-based SPH model for cranial ballistic injury

†E.E. Kwon<sup>1</sup>, M.R. Singh<sup>1</sup>, R.D. Vallabh<sup>1</sup>, \*R. Das<sup>1</sup>, J.W. Fernandez<sup>2,3</sup> and M.C. Taylor<sup>4</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Auckland, New Zealand

<sup>2</sup>Auckland Bioengineering Institute (ABI), University of Auckland, New Zealand

<sup>3</sup>Department of Engineering Science, The University of Auckland, Auckland, New Zealand

<sup>4</sup>The Institute of Environmental Science and Research (ESR), Christchurch Science Centre, New Zealand

\*Presenting author: r.das@auckland.ac.nz

†Corresponding author: eryl.kwon@auckland.ac.nz

### Abstract

Forensic investigation has the primary challenge of assessing cause from limited evidence. To inform the investigation process, computational modelling can assess: i) the potential ballistic pathways, by analysing entry wound and blood spatter patterns; and ii) the influence of target material effects and cranial geometry. The retrograde ejection of blood and tissue following projectile impact from the entry wound is called ‘backspatter’ and can aid in informing the investigator about the proximity of the shooter, with the potential to differentiate between suicide and homicide. However, the ‘backspatter’ phenomenon is not well understood. This study presents (i) the development of an anatomically-based model of cranial ballistic injury using the Smoothed Particle Hydrodynamics (SPH) method; (ii) simulation of the tail splashing and temporary cavitation mechanisms by utilizing a range of scalp and bone simulants and comparison with experiment; (iii) evaluation of cranial stress and strain and energy dissipation; and (iv) evaluation of the effects of bullet characteristics on the creation of the entry wound by parametric analysis.

**Keywords:** Smooth Particle Hydrodynamics, SPH, Ballistic Simulation, Backspatter

### Introduction

Backspatter refers to ejection of biological tissue from ballistic wounds, opposite to the line of fire (Stephens and Allen 1983, Karger 2008). It is widely accepted that backspatter occurs, particularly in close-range contact shots to the head (Stephens and Allen 1983, Yen, Thali et al. 2003). The stain patterns resulting from backspatter are of critical importance in a crime scene because of the direction against the line of fire (Grosse Perdekamp, Vennemann et al. 2005), providing a potential connection between the victim and the shooter. However, the literature on backspatter is limited compared to other areas of blood spatter research, providing a less solid foundation and a less thorough understanding of the mechanisms (Radford 2009).

The three main mechanisms contributing to backspatter reported in the literature are (i) subcutaneous gas pockets; (ii) temporary cavities; and (iii) tail splashing (Karger 2008). Subcutaneous gas pockets are temporary spaces between skin and bone created due to muzzle gas from contact or near-contact shots (Karger and Brinkmann 1997). Temporary cavities form when the passage of a projectile through near-water density organs, such as the musculature or the brain, creates pressure waves to radially expand the bullet trail temporarily (Karger 2008). In the case of distant shots, the temporary cavity in brain is believed to be a major contributor (Foote 2012). Tail splashing results from a backward streaming of fluid and fragments along the sides of the bullet in the retrograde direction upon projectile impact (Amato, Billy et al. 1974).

Due to the extreme speed of a projectile, the ballistic event happens very quickly, making study of the backspatter event difficult, with little prior work collecting detailed physical evidences to allow thorough validation. Additionally, the backspatter event is the result of various combinations of mechanisms and variables, making it difficult to isolate each mechanism to gauge their contribution. To reduce the complexity issues, this study focuses only on distant shots, so the subcutaneous gas pocket mechanism can be eliminated from consideration. The focus of this research is solely on the tail splash and temporary cavitation.

In order to collect the necessary detailed physical evidences, it is essential to form proper models, as human samples are not viable due to ethical issues. Previous literature has shown that the problem of finding a target for experimentation that best simulates the properties of the human cranium has been a significant constraint with much controversy. The use of animal models, involving calves or swine has been established well in the literature (Karger, Nusse et al. 1996, Karger, Nusse et al. 1997, Karger, Nusse et al. 2002, Radford 2009). The animal models used commonly have ethical issues, different geometry to that of a human cranium, and individual biodiversity which makes them unreliable as an experimental model (Yen, Thali et al. 2003). To provide an alternative to the animal models, physical models using synthetic materials have been developing, with increased geometric details and complexity in their construction (Stephens and Allen 1983, Thali, Kneubuehl et al. 2002, Radford 2009, Foote 2012, Carr, Lindstrom et al. 2014, Kwon 2014). Physical models have no ethical issues and the experimental results agree well with the reported human case studies (Thali, Kneubuehl et al. 2002, Carr, Lindstrom et al. 2014). However, the high cost of each sample and the manufacturing difficulties are major hurdles to using physical models to study backspatter.

Computational models are increasingly being used as an alternative to complicated, costly, and potentially ethically challenging experiments. In ballistic experiments, the experimental parameters, such as speed of the bullet or the sample material properties, are not easily controlled. However, the computational model allows fast and easy customisation of the experimental parameters. This gives the computational model a clear advantage for forensic crime scene reconstructions, where the bullet calibre or the impact speed is not known, or the victim had significantly different biological material properties due to aging or disease. For example, if the victim was suffering from osteoporosis (a disease effecting bone density, causing weak and brittle bones), the current studies in animal or physical models would have to be re-done to compensate for this factor. The computational model has unmatched analytic advantage as the each layer of the model can be visualised separately without damaging the other layers. For animal or physical models, the skin layer must be damaged in order to analyse the bone layer defect size and crack propagation.

There are two computational methods through which study of backspatter could be made possible: finite element method (FEM) and Smoothed Particle Hydrodynamics (SPH). FEM offers the advantage of being able to model structures with intricate shapes and indirectly quantify their complex behaviour at a point (Raul, Deck et al. 2007). However, this method has a significant limitation in that at high speeds and deformations, mesh integrity is lost. In contrast, SPH, originally developed for astrophysics (Gingold and Monaghan 1977), is a conservative, mesh-less method that can handle complicated, highly-deformable geometry, large void areas, and dynamic ballistic behaviour (Stellingwerf and Wingate 1993). The SPH method has been utilized in several applications, including explosive fragmentation of metal casings involving intense shock and high rate plastic deformation (Kong, Wu et al. 2013) and the high speed impact of a metal sphere on a thin metallic plate with a range of materials and velocities (Kalameh, Karamali et al. 2012).

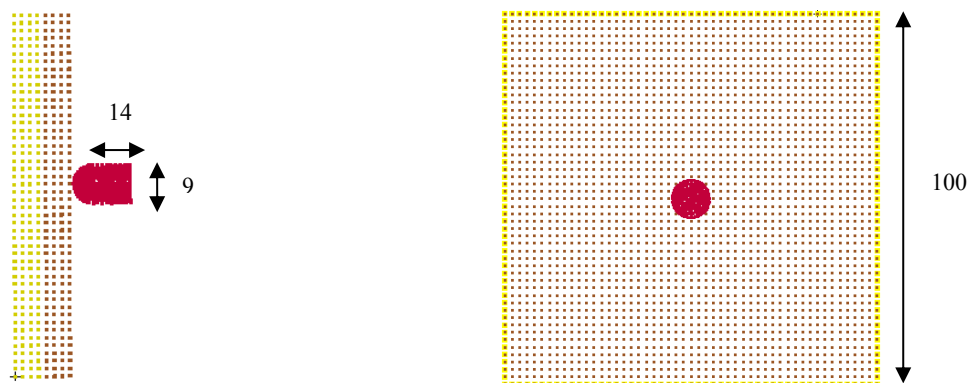
In this study we present both a simple flat plate SPH model as well as an anatomically accurate SPH model of the male human head. The less-computationally-expensive simple geometry is used to test various material constitutive models and different simulants, as well as ballistic characteristics. The anatomical geometry was used for the final models, which represent a more detailed and realistic ballistic simulation. Both geometries consist of two layers: the scalp and the skull. The scalp comprises of the skin and the connective tissues overlying the top of the head (Harris, Yoganandan et al. 1993). The skin is a highly non-linear, anisotropic, viscoelastic and nearly incompressible material (Crichton, Donose et al. 2011, Crichton, Chen et al. 2012). The skull is the firm bone that encases the brain. On projectile impact, the skull is expected to show radial fractures originating from the point of impact and concentric fractures around the bullet entry and exit sites (Viel, Gehl et al. 2009). Common characteristics often observed but not always present, include bevelling and keyhole defects (Quatrehomme and Iscan 1997). Another major layer of the human cranium, the brain, has been omitted in this research to concentrate the investigation on the contribution of scalp and skull to backspatter generation.

In this paper, we present findings of the SPH-based simple and anatomically-accurate model of cranial ballistic injury, including: (i) simulation of the tail splashing and temporary cavitation mechanisms, utilizing a range of scalp and bone simulants and comparison with experiment; (ii) evaluation of cranial stress and strain and energy dissipation; and (iii) evaluation of the effects of bullet calibre and speed on the creation of the entry wound by parametric analysis.

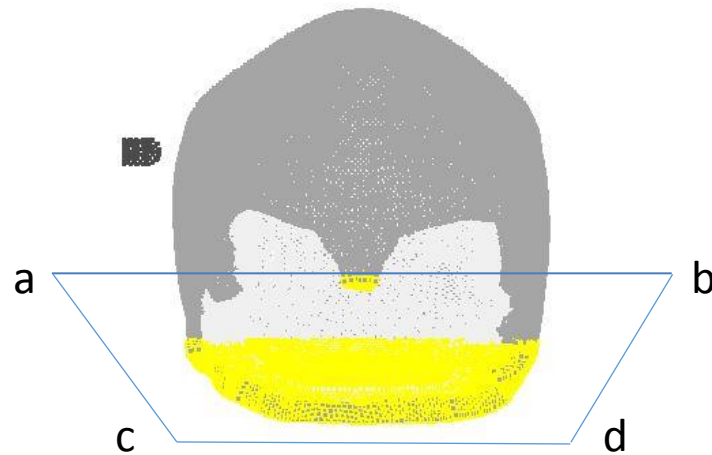
## Methods

The simple plate geometry (Figure 1) is composed of two flat plates of 100 x 100 x 5 mm, each representing the scalp layer and the skull layer of a human cranium. The minimum SPH particle size for this geometry was 0.7 mm, which contained ~ 580,000 SPH particles.

A representative scalp and cranium geometry was adapted from the Physiome Project repository of a 38 year old male from the Visible Human male data set (Spitzer, Ackerman et al. 1996) and developed at the Auckland Bioengineering Institute. The dimensions of the scalp layer were acquired from MRI scan results taken from images from the Centre for Advanced MRI (Auckland University). The SPH particle size was set to 2.5 mm, which contained ~141,000 SPH particles.



**Figure 1. Simple geometry with a 9 mm bullet, Scalp layer is shown as red, Skull layer is shown as yellow, the pinned boundary is shown as yellow edge on the right**



**Figure 2. Anatomical geometry with a 9 mm bullet, scalp is shown in grey, skull is shown in white, the pinned boundary below the plane abcd is highlighted with yellow**

Both geometries had equivalent physical models, with matching geometry and simulant materials. The experimental result was recorded using high speed photography, at a frame rate of 30,000 frames per second. The purpose of the physical models was to be the validation tool for the computational models.

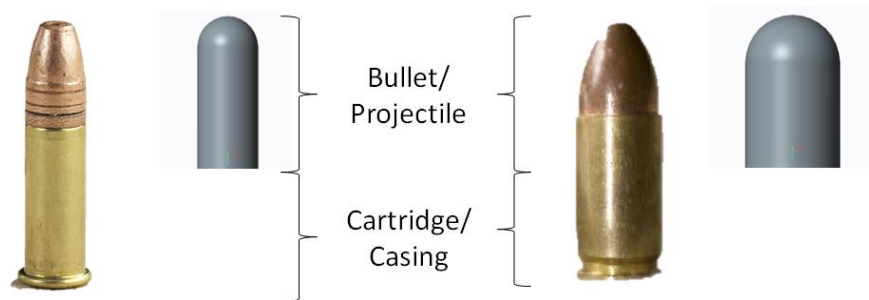
A total of 16 scalp-skull simulant combinations, consisting of four scalp simulant and four skull simulant materials, were simulated using the simple geometry model. For the scalp simulants, the material properties of the ballistic gelatine, Room Temperature Vulcanizing (RTV) silicone\_1, RTV silicone\_2 and natural rubber were modelled using either viscoelastic or hyperelastic failure. The material constants were obtained from either the Mooney-Rivlin or the Yeoh models (Wang, Deng et al. 2004, Korochkina, Jewell et al. 2008). Both models use strain energy potential to characterise rubber. For the skull simulants, the material properties of the Medium-Density Fiberboard (MDF), Particle board, Polyurethane\_1 and Polyurethane\_2 were modelled using isotropic-elastic failure. No material constants were required for the elastic failure.

The anatomical geometry model was simulated using the best performing simulants selected based on the simple geometry model simulation results. The scalp layer was modelled using RTV silicone\_2 and the skull layer was modelled using Polyurethane\_2.

The simple geometry model was also used for the parametric studies of the effect of bullet calibre and the velocity on ballistic response of the target. 2 different types of bullets were used, a 9 mm Luger and a .22 Long Rifle (LR). The 9 mm is a 115grain, Full Metal Jacket, and has solid lead core with a copper gilding material coating. The .22 LR had less wounding potential and penetrating potential compared to the 9 mm Luger. Both the 9 mm and .22 bullets were modelled as a cylinder with hemisphere tip, with material properties of copper and density of lead to represent the solid lead core. The bullet was modelled using Johnson and Cook high velocity impact copper model from ANSYS material library (Johnson and Cook 1985).

The 9 mm was simulated with 300 m/s and 370 m/s impact velocity, each representing the lower limit and the upper limit of the projectile speed respectively. In ballistic experimentation, the physical models were shot from a distance over 1 m. This isolated the tail splashing and temporary cavitation as backscatter generation mechanisms by eliminating the subcutaneous gas pockets that would have formed if the shot distance were to be closer. As seen in Figures 1 and 2, the bullet has been positioned at the centre of the simple geometry model and on the right temple of the

anatomical geometry model. The anatomical position was chosen based on the frequency of the cranial ballistic wounding site.



**Figure 3. Physical and computational bullet geometry comparison between 0.22 (left) and 9 mm (right) – only the projectile part is modelled in computational models**

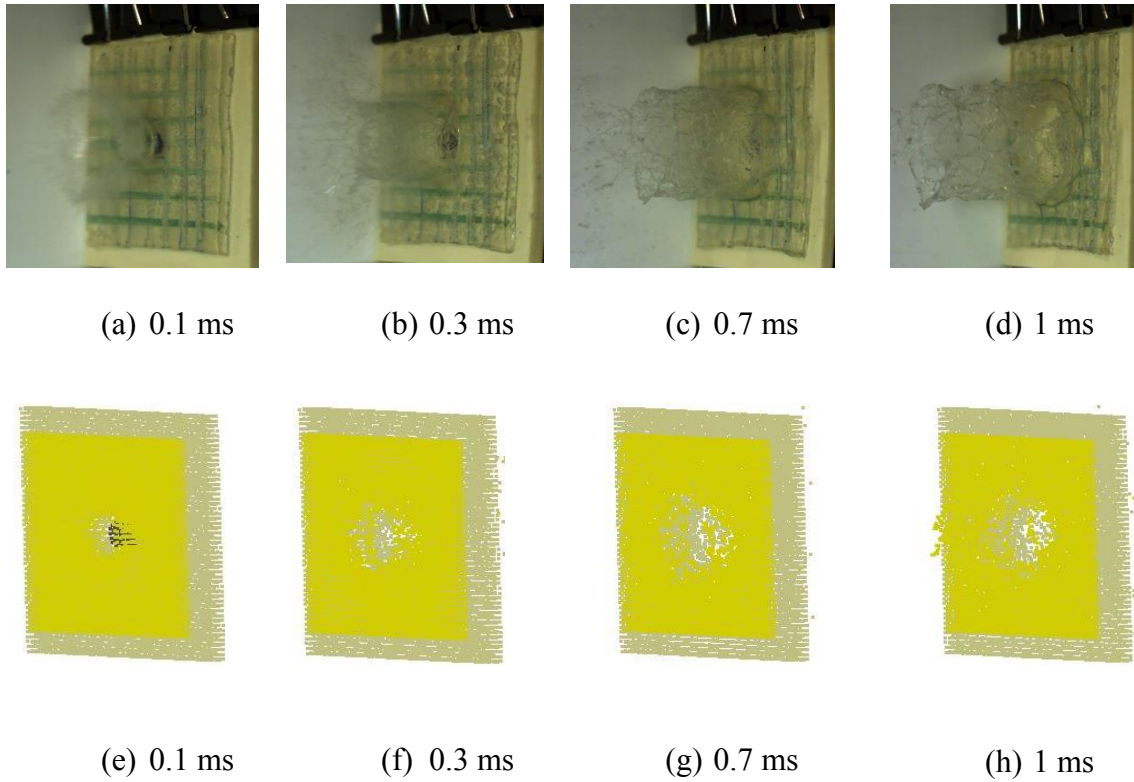
## Results and Discussion

### *Simple geometry*

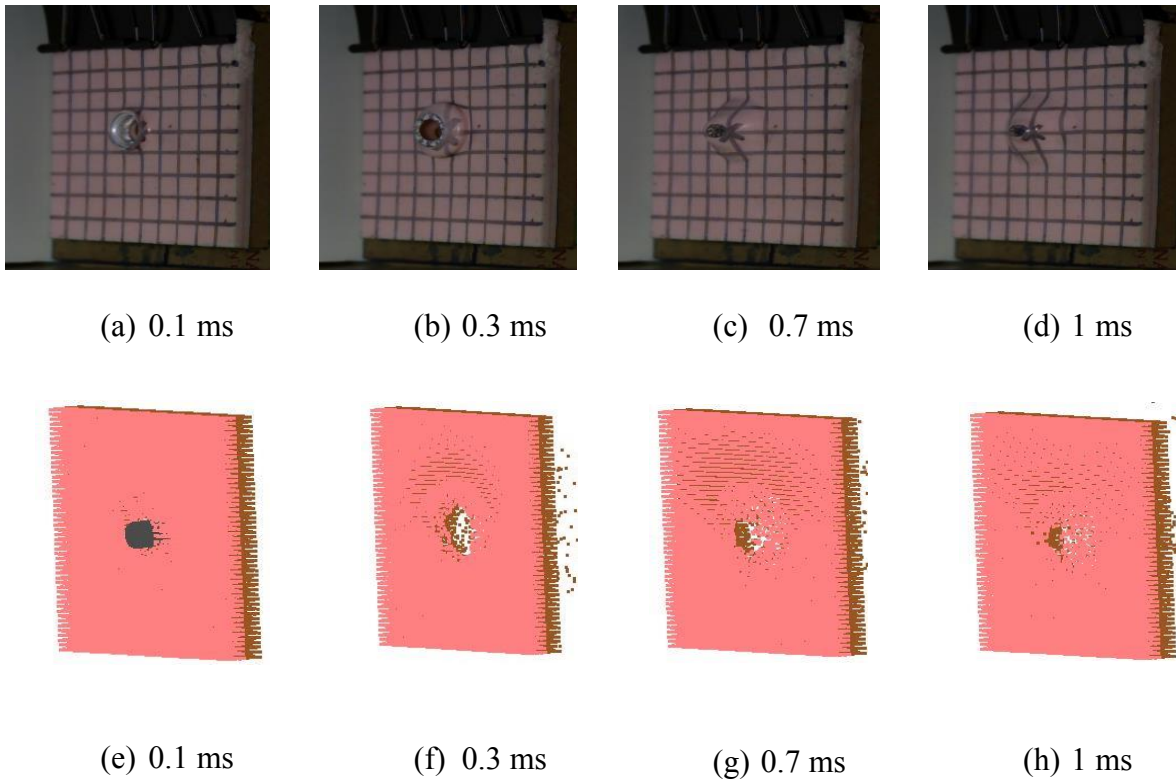
The experimental and simulation results were analysed to determine the most suitable scalp and skull simulant combination to represent human cranial ballistic wounding. The simulation results were validated by comparison to the experimental results. The comparison was based on both qualitative and quantitative observations made over dynamic deformation behaviour and static defect dimensions and characteristics. Two of the most typical dynamic ballistic deformation results from the testing and simulation of the physical and computational models are compared chronologically in Figures 4 and 5.

The first main point of analysis was the ability to replicate the tail splash and temporary cavitation backscatter generation mechanisms. The gelatine skin simulant was better at demonstrating the tail splashing mechanism than any other scalp simulants used. The tail splash from the Gelatin-Polyurethane\_1 model is illustrated, at 0.1 ms post-impact, in Figure 4 (a) and (e). On the other hand, the Silicone\_1-Polyurethane\_2 model illustrates the subcutaneous temporary cavity mechanism very well. The computational model successfully simulated the retrograde elastic bulging of the scalp simulant as well as the reduction of the bullet entry hole size. The choice of skull simulants affected the magnitude of the ballistic response, producing more backscatter for the Polyurethane simulants compared to the MDF or Particle Board. Also, there was minimal skull simulant backscatter observed, which concurs with the literature findings (Burnett 1991, Coupland, Rothschild et al. 2011).

It is expected that the tough integument of the human scalp will prevent the crushed tissues to backscatter as freely as the gelatine simulant. Therefore, the Silicone\_2 was accepted as the best scalp simulant. The human skulls also show bevelling and minor crack formation at the bullet entrance, and the lack of ability to produce such characteristics eliminated both the MDF and Particle Board as a viable skull simulant. The Polyurethane\_2 was chosen as the best skull simulant based on the bone defect diameter and morphology.



**Figure 4. Comparison of experimental (a-d) and simulation (e-h) results of a 9 mm projectile impact on Gelatin-Polyurethane\_1 model**

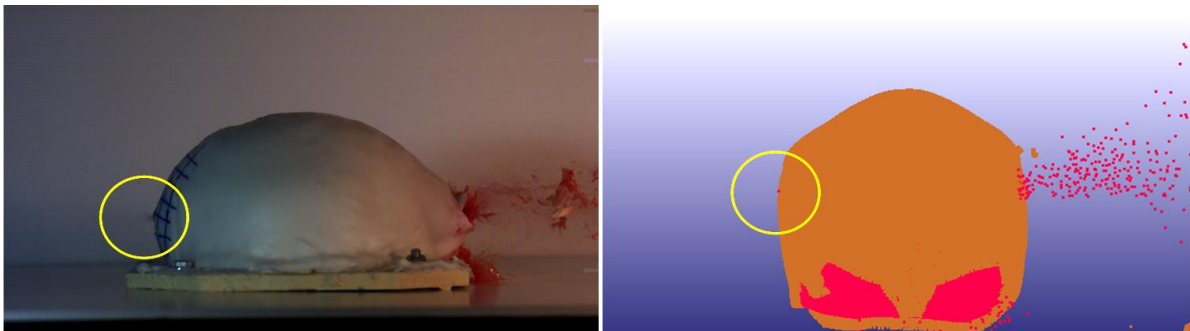


**Figure 5. Comparison of experimental (a-d) and simulation (e-h) results of a 9 mm projectile impact on Silicone\_1-Polyurethane\_2 model**



### *Anatomical geometry*

The anatomical geometry model simulated using the Silicone\_2-Polyurethane\_2 simulant combination produced an extended subcutaneous temporary cavity, due to surface curvature and increased kinetic energy absorption. This increased magnitude of temporary cavity dynamics illustrated the existence of a delayed backspatter due to the temporary cavity mechanism. Crucially, the tail splash mechanism produced backspatter at the early impact stage, and the anatomical geometry model produced a chronologically separated second ejection of the backspatter. This delayed backspatter was observed around 2.6 ms and coincided with the collapse of the temporary cavity. Therefore, the pressure created by the collapse of the temporary cavity mechanism is responsible for the retrograde ejection of the fragments inside the cavity of this backspatter via path of least resistance (Figure 6).

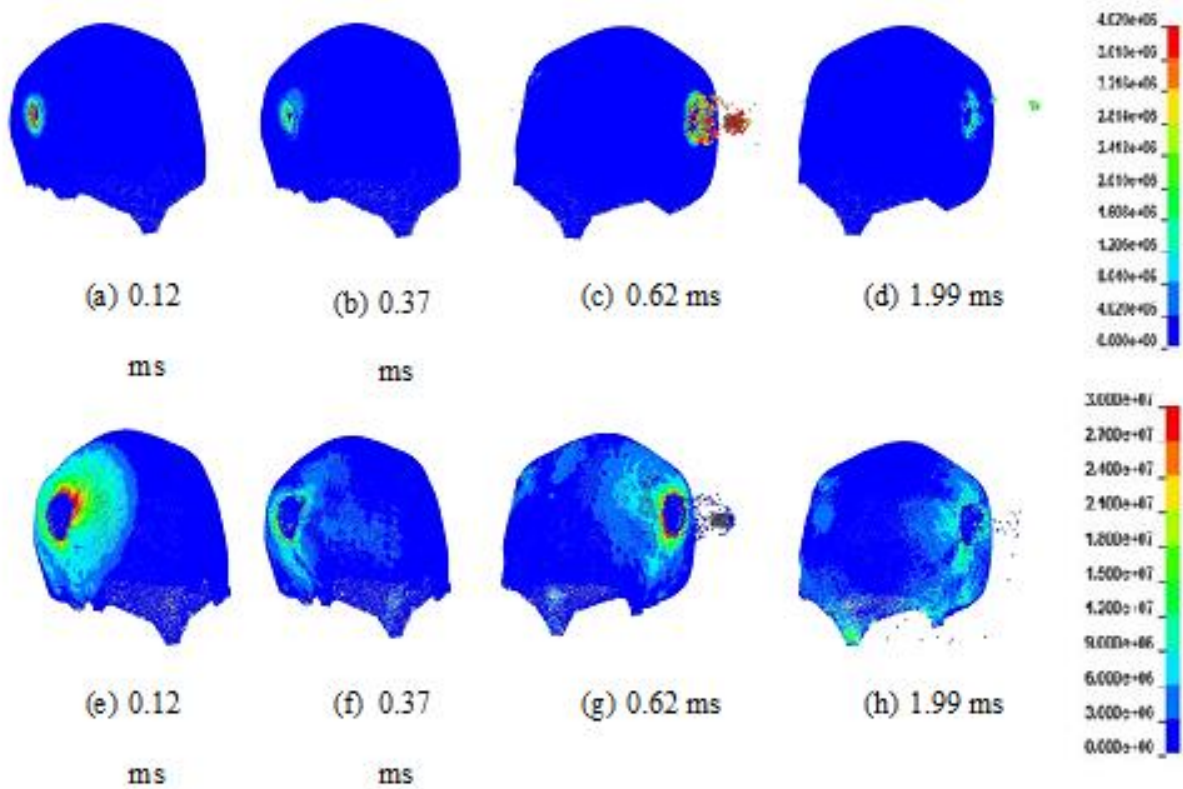


**Figure 6. Comparison of experimental (left) and simulation (right) results of a 9 mm projectile impact on Silicone\_2-Polyurethane\_2 anatomical geometry model, at 2.6 ms post-impact. The delayed backspatter is highlighted with yellow circles.**

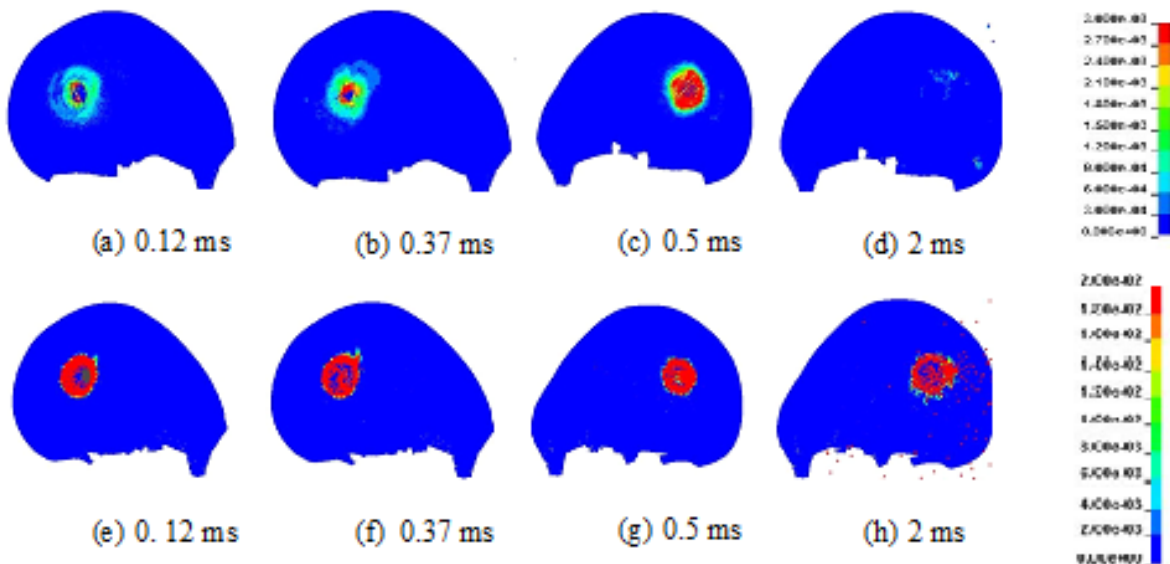
### *Computational model unique analysis*

One of the main advantages of the simulation is that it allows detailed analysis of the impact event without the use of complicated actuators and associated error. In experimentation, it is almost impossible to isolate the stress and strain developed in each layer of the cranial model. For the aforementioned practical reasons, there has been no analysis on the stress patterns developed in an animal model nor a physical model reported in a literature previously.

However, the computational modelling has ability to calculate the stress and strain of the target without additional processing. The anatomical geometry model simulation result was analysed for the Von Mises Stress (Figure 7) and Plastic strain (Figure 8). These are typical engineering measurements used to analyse stressed and deformation developed in a structure. The plastic strain was useful to monitor the area of the deformation while the Von Mises stress was used to visualise stress distribution in the bone layer.



**Figure 7. Evolution of Von Mises stresses on the scalp (a-d) and the skull (e-h) of the anatomical geometry model**



**Figure 8. Evolution of mean effective strain on the scalp (a-d) and the skull (e-h) of the anatomical geometry model**

For the stress analysis, the Silicone\_2 scalp simulant stretches and fails at the impact point immediately upon the projectile impact. This is due to the high stresses generated around the entry site as shown in Figure 7 (a). A maximum stress of 4.02 MPa is generated at the bullet entry site, which is greater than the tensile strength of the material, leading to extension and failure. This is in accordance with the observations made by Jussila et al (Jussila, Leppaniemi et al. 2005), who assert that ‘an impacting bullet makes the skin to stretch, partially crush and finally rupture, allowing the bullet to enter underlying tissue’. The impact stresses are localized and most of the surrounding areas of the head are not affected, as indicated by the blue regions. At a time of 0.3 ms, Figure 7 (b), the entrance cavity oscillates and closes up, causing build-up of high pressures. Following the exit of the bullet shown in Figures 7 (c) and 7 (d), the hyperelastic nature of the material leads to minor oscillations as it tries to recover to its original state and close up the cavity.

On bullet impact, high stresses in the order of 30 MPa are produced in the Polyurethane\_2 skull simulant. Since the stresses are greater than the ultimate tensile strength of the resin it fails at the impact point in a brittle manner. The stresses then radiate outwards as shown in Figure 7 (e) and 7 (f). The magnitude of the radiating stresses (4-10 MPa) is well below the failure stress of the material. High stresses produced at the exit site initiate crack propagation along the cranium as indicated by the faint red lines in Figures 7 (g) and 7 (h). The entry site presents a clean ‘punched out’ external appearance while the exit wound is much larger and irregular in shape. This is the desired result supported by other sources in literatures (Quatrehomme and Iscan 1998, Quatrehomme and Iscan 1999).

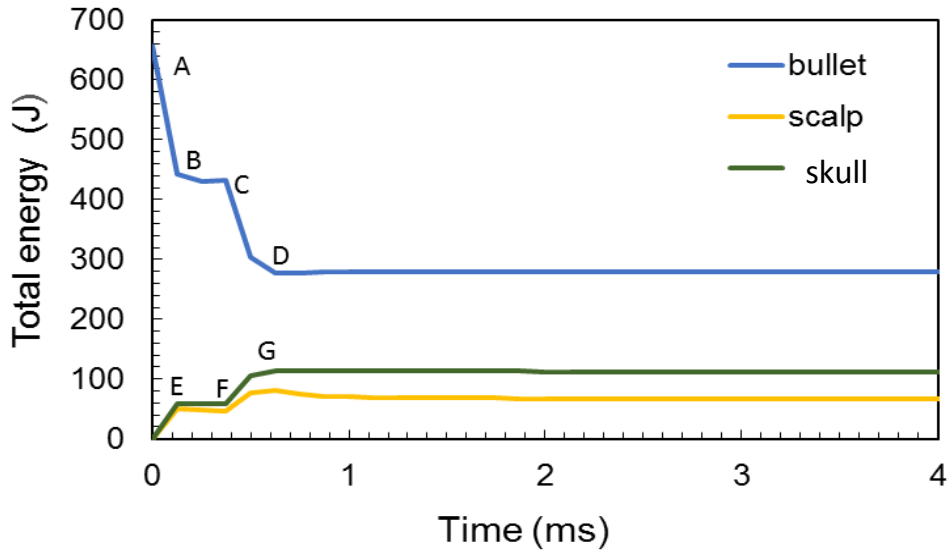
For the strain analysis, the Silicone\_2 scalp simulant strains and starts tearing as shown in Figure 8 (a). A similar tearing of the skin from a contact shot has been recorded in literature (Faller-Marquardt, Bohnert et al. 2004). Even after failure, the Silicone\_2 tries to close up the entrance wound causing the surrounding material to strain as indicated by the light blue particles in Figure 8 (b). The elastic deformation of the silicone rubber on bullet exit can be clearly perceived by the red regions in Figure 8 (c). Over time, the oscillations die away and the strain diminishes as shown in Figure 8 (d).

When the bullet enters the Polyurethane\_2 skull layer of the model, strains up to 2% are observed at the impact point leading to fracture. In addition, the high strains caused around bullet entry and exit sites in Figures 8 (f) and 8 (h) are much larger than the bullet diameter. An implication of this observation is that the polyurethane resin demonstrates some amount of plastic deformation and is therefore not perfectly brittle.

The energy graph of each part and layers involved (Figure 9) provides a very valuable insight into the ballistic impact event. It shows how the original kinetic energy of the bullet is converted to various other forms of energy to result in the ballistic response of the target as a whole. Initially, the particles that shape the scalp and cranium layers are at rest (Point A and Origin). On impact (energy change to point B and E), it is assumed that the bullet’s kinetic energy changes into: bullet deformation energy, damage energy, heat generation, and impact energy (Komuński, Kubiak et al. 2009). Since the bullet used in experimentation has a full metal jacket it is assumed not to lose energy due to bullet deformation. Energy lost in the form of heat is thought to be minimal in our analysis.

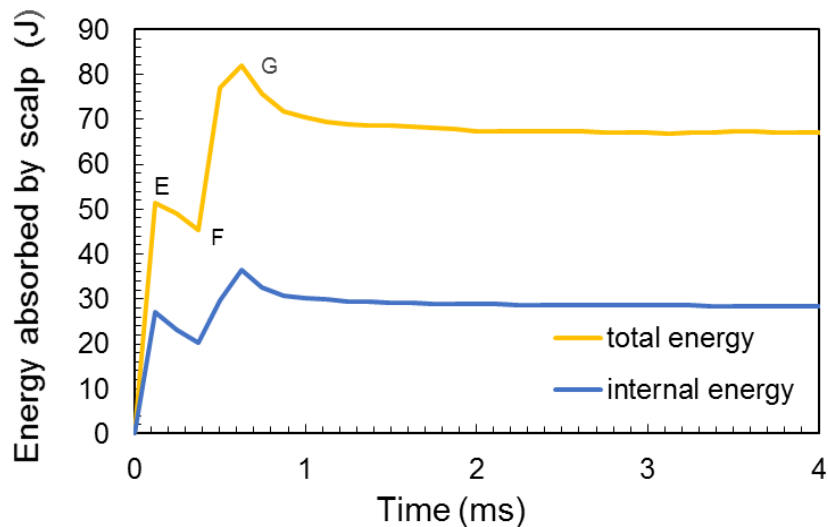
The scalp and cranium layers each absorb 50-60J of energy as the bullet passes through (Point E). The rest of the energy (117J, point A to point B in Figure 9) is used to fracture and damage the material. Between points E and F, the total energy of the cranial layer remains fairly constant, but

the scalp loses a fraction of its energy. A similar effect is observed between points F and G as the scalp layer gains 50J of energy and then loses a small percentage of it.



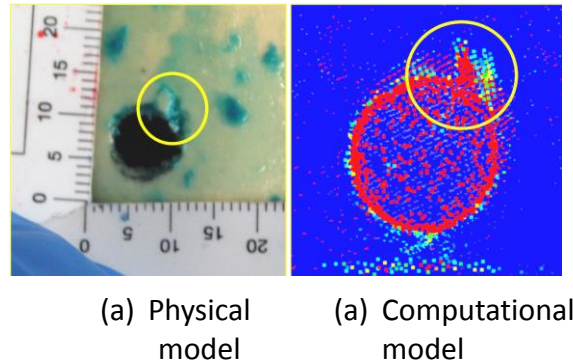
**Figure 9. The total energy of the bullet, scalp and skull over time**

This behaviour can be best explained by considering the total energy absorbed by the hyperelastic scalp over time in relation to its internal energy, as shown in Figure 10. Due to its elastic nature, a fraction (5J) of the total internal energy (25J) at 0.12ms is utilized in deforming the material as it tries to recover to its original state (Point E and F). Thus, the scalp loses 20% (5J/25J) of its internal energy in the form of strain energy. Once the scalp layer partially recovers to a position of equilibrium, it stops oscillating and the total energy of the system remains constant over time. The cranium is represented by a hard, brittle material so it sustains its internal energy after bullet exit (Point G and beyond) without further deformation. In summary, a total of 380J of energy is lost by the projectile, while the scalp and cranium layers gain 66J (17.4% of projectile energy) and 112J (29.5%) respectively. The scalp layer loses 20% of its internal energy due to straining of the material.



### Figure 10. Comparison of the scalp layer internal energy to its total energy

The ballistic parametric study showed that the larger calibre bullet produces the larger defect size, which concurs with the experimental observations. The simulation produced similar defect characteristics for the bone layer, producing bevelling and keyhole defect of similar dimensions (Figure 11).



**Figure 11. Comparison between the bone layer keyhole defect of (a) physical and (b) computational (anatomical geometry) model**

#### *Bullet characteristic parametric study*

The calibre and speed of the projectile used was varied in the anatomical geometry model to investigate the dependence of bullet characteristics on the backspatter generation and the stress/strain developed in the model.

The .22 LR produced entrance defect smaller than the 9 mm. Also the stress generated by the .22 LR was more localized on the impact site. More significantly, the .22 LR impact did not produce any backspatter from both the physical and computational models.

The bullet speed change does not have a physical experimental result to compare to. The simulational observation was: i) higher impact speed of 370 m/s resulted in larger area of stress development with greater impact stresses for both the .22 LR and the 9 mm; ii) The percentage energy absorbed from the projectile was almost independent of the impact velocity. This signifies more energy is absorbed into the target as the projectile speed increases. In an experimental study by Clemenson et al. (Clemenson, Falconer et al. 1973), the maximum pressure created in the head varied by approximately the square of the projectile velocity.

### Conclusions

This study presents both a simplified and anatomically-accurate computational models of human cranium for ballistic backspatter research. The computational model successfully incorporated human anatomical geometry into the scalp and the skull layers. The simple geometry was used to increase simulational efficiency in simulant evaluation.

The tail splash and temporary cavitation mechanism has been witnessed and confirmed as a major backspatter cause/mechanism when the subcutaneous gas pocket mechanism was eliminated. When combined with the unique ability of the computational models to provide the stress, strain and energy graphs, a detailed chronological description of the temporary cavitation mechanism was made, which was not reported in literatures before.

The effect of the bullet calibre or speed change resulted in desired variation in the simulational results.

Overall, the computational model has these benefits

- No ethical issue
- Cost-effective
- Ease of experimentation
- Higher control of experimental variables
- Ease of customisation for the use of forensic case studies and crime reconstructions
- Unmatched analytical advantage, providing otherwise unattainable values such as stress, strain and energy of individual parts.
- Non-invasive analysis

## References

- Amato, J. J., L. J. Billy, N. S. Lawson and N. M. Rich (1974). "High Velocity Missile Injury: An Experimental Study of the Retentive Forces of Tissue." *The American Journal of Surgery* **127**: 454-459.
- Burnett, B. R. (1991). "Detection of Bone and Bone-Plus-Bullet Particles in Backspatter from Close-Range Shots to Heads." *Journal of Forensic Science* **36**(6): 1745-1752.
- Carr, D., A.-C. Lindstrom, A. Jareborg, S. Champion, N. Waddell, D. Miller, M. Teagle, I. Horsfall and J. Kieser (2014). "Development of a skull/brain model for military wound ballistics studies." *International Journal of Legal Medicine*: 1-6.
- Clemedson, C.-J., B. Falconer, L. Frankenberg, A. Jönsson and M. D. J. Wennerstrand (1973). "Head injuries caused by small-calibre, high velocity bullets." *Zeitschrift für Rechtsmedizin* **73**(2): 103-114.
- Coupland, R. M., M. A. Rothschild, M. J. Thali, B. P. Kneubuehl, R. M. Coupland, M. A. Rothschild and M. J. Thali (2011). *Introduction to Wound Ballistics*. B. P. Kneubuehl, Springer Berlin Heidelberg: 1-2.
- Crichton, M. L., X. Chen, H. Huang and M. A. F. Kendall (2012). "Elastic modulus and viscoelastic properties of full thickness skin characterised at micro scales." *Biomaterials*.
- Crichton, M. L., B. C. Donose, X. Chen, A. P. Raphael, H. Huang and M. A. F. Kendall (2011). "The viscoelastic, hyperelastic and scale dependent behaviour of freshly excised individual skin layers." *Biomaterials* **32**(20): 4670-4681.
- Faller-Marquardt, M., M. Bohnert and S. Pollak (2004). "Detachment of the periosteum and soot staining of its underside in contact shots to the cerebral cranium." *Int J Legal Med* **118**(6): 343-347.
- Foote, N. R. (2012). *The Role of the Temporary Cavity in Cranial Backspatter*. Master of Science in Forensic Science, The University of Auckland.
- Gingold, R. A. and J. J. Monaghan (1977). "Smoothed particle hydrodynamics-theory and application to non-spherical stars." *Monthly notices of the royal astronomical society* **181**: 375-389.
- Grosse Perdekamp, M., B. Vennemann, D. Mattern, A. Serr and S. Pollak (2005). "Tissue defect at the gunshot entrance wound: what happens to the skin?" *Int J Legal Med* **119**(4): 217-222.
- Harris, G. F., N. Yoganandan, D. Schmaltz, J. Reinartz, F. Pintar and A. Sances Jr (1993). "A biomechanical impact test system for head and facial injury assessment and model development." *Journal of Biomedical Engineering* **15**(1): 67-73.
- Johnson, G. R. and W. H. Cook (1985). "Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures." *Engineering Fracture Mechanics* **21**(1): 31-48.
- Jussila, J., A. Leppaniemi, M. Paronen and E. Kulomaki (2005). "Ballistic skin simulant." *Forensic Sci Int* **150**(1): 63-71.
- Kalameh, H. A., A. Karamali, C. Anitescu and T. Rabczuk (2012). "High velocity impact of metal sphere on thin metallic plate using smooth particle hydrodynamics (SPH) method." *Frontiers of Structural and Civil Engineering* **6**(2): 101-110.
- Karger, B. (2008). *Forensic Ballistics*. *Forensic Pathology Reviews*. M. Tsokos. Totowa, NJ, Humana Press. **5**: 139-172.
- Karger, B. and B. Brinkmann (1997). "Multiple gunshot suicides: potential for physical activity and medico-legal aspects." *Int J. Legal Med.* **10**: 188-192.
- Karger, B., R. Nüsse and T. Bajanowski (2002). "Backspatter on the Firearm and Hand in Experimental Close-Range Gunshots to the Head." *The American Journal of Forensic Medicine and Pathology* **23**(3): 211-213.
- Karger, B., R. Nüsse, G. Schroeder, S. Wustenbecker and B. Brinkmann (1996). "Backspatter from Experimental Close-Range Shots to the Head I-Microbackspatter." *International Journal of Legal Medicine* **109**: 66-74.
- Karger, B., R. Nüsse, H. D. Troger and B. Brinkmann (1997). "Backspatter from Experimental Close-Range Shots to the Head II-Microbackspatter and the Morphology of Bloodstains." *International Journal of Legal Medicine* **110**: 27-30.
- Komuński, P., T. Kubiak, M. Łandwijit and R. Romek (2009). "Energy transmission from bullet impact onto head or neck through structures of the protective ballistic helmet-tests and evaluation." *Techniczne Wyroby Włókiennicze* **17**: 18-23.
- Kong, X., W. Wu, J. Li, F. Liu, P. Chen and Y. Li (2013). "A numerical investigation on explosive fragmentation of metal casing using Smoothed Particle Hydrodynamic method." *Materials & Design* **51**: 729-741.
- Korochkina, T. V., E. H. Jewell, T. C. Claypole and D. T. Gethin (2008). "Experimental and numerical investigation into nonlinear deformation of silicone rubber pads during ink transfer process." *Polymer Testing* **27**(6): 778-791.

- Kwon, E. E. (2014). Development of Physical and Numerical Models to Study Cranial Backspatter. Master of Engineering, University of Auckland.
- Quatrehomme, G. and M. Y. Iscan (1997). "Bevelling in exit gunshot wounds in bones." Forensic Sci Int **89**(1-2): 93-101.
- Quatrehomme, G. and M. Y. Iscan (1998). "Gunshot wounds to the skull: comparison of entries and exits." Forensic Sci Int **94**(1-2): 141-146.
- Quatrehomme, G. and M. Y. Iscan (1999). "Characteristics of gunshot wounds in the skull." J Forensic Sci **44**(3): 568-576.
- Radford, G. E. (2009). Modelling Cranial Gunshot Wounds and Backspatter. Master of Science, University of Otago.
- Raul, J., C. Deck, F. Meyer, A. Geraut, R. Willinger and B. Ludes (2007). "A finite element model investigation of gunshot injury." International Journal of Legal Medicine **121**: 143-146.
- Spitzer, V., M. J. Ackerman, A. L. Scherzinger and D. Whitlock (1996). "The visible human male: a technical report." Journal of the American Medical Informatics Association **3**(2): 118-130.
- Stellingwerf, R. and C. Wingate (1993). "Impact modeling with smooth particle hydrodynamics." International Journal of Impact Engineering **14**(1): 707-718.
- Stephens, B. G. and T. B. Allen (1983). "Back Spatter of Blood from Gunshot Wounds - Observations and Experimental Simulation." Journal of Forensic Sciences **28**(2): 437-439.
- Thali, M. J., B. P. Kneubuehl and R. Dirnhofer (2002). "A "Skin-skull-brain" Model for the Biomechanical Reconstruction of Blunt Forces to the Human Head." Forensic Science International **125**: 195-200.
- Viel, G., A. Gehl and J. P. Spherhake (2009). "Intersecting fractures of the skull and gunshot wounds. Case report and literature review." Forensic science, medicine, and pathology **5**(1): 22-27.
- Wang, W., T. Deng and S.-g. ZHAO (2004). "Determination for Material Constants of Rubber Mooney-Rivlin Model [J]." Special Purpose Rubber Products **4**: 003.
- Yen, K., M. J. Thali, B. P. Kneubuehl, O. Peschel, U. Zollinger and R. Dirnhofer (2003). "Blood-Spatter Patterns: Hands Hold Clues for the Forensic Reconstruction of the Sequence of Events." American Journal of Forensic Medicine and Pathology **24**(2): 132-140.

## **Cementless total hip arthroplasty with and without screw fixation: a randomized study of 78 hips**

**Zhi-qi Zhu, Peng Li, Ming Hou, De-gui Wang, Xue-yang Lin**

Department of Orthopaedic Surgery, Longgang District People's Hospital of Shenzhen, Shenzhen China, 518172

### **Abstract**

[Objective] This study is to evaluate the clinical and radiographic results of cementless total hip arthroplasty with and without screw fixation. [Methods] Cementless total hip arthroplasty was performed from August 2009 to August 2011. 76 patients (78 hips) were randomized screw fixation group (group A, n = 40) and non-screw fixation group (group B, n = 38). Consecutive radiographs were compared to evaluate acetabular condition at 3 months, 12 months and 24 months after surgery and the last follow-up, respectively. Harris Hip Scores (HHSs) were determined before surgery and at the most recent follow-up. The Kaplan-Meier survivorship analysis was used to estimate the survival rate of the prosthesis in the two groups. [Results] 39 hips in group A and 36 hips in group B were available for complete clinical and radiographic analysis while the lost-to-follow-up patients in group A and group B were 1 and 2, respectively. The mean follow-up period was 4.2 years (range, 2.3~5.5 years). In group A, radiolucent lines were present around the cup in 4 hips (10.3%), 3 hips (7.7%) and 1 hip (2.6%) at 3 months, 12 months and 24 months, respectively. By contrast, in group B, radiolucent lines appeared around the cup in 1 hip (2.8%), 2 hips (5.6%) and 2 hips (5.6%) at 3 months, 12 months and 24 months, respectively. Osteolysis and migration were observed in 1 hip in group A and 2 hips in group B. The Harris scores were  $95 \pm 2.1$  points in group A and  $93 \pm 5.5$  points in group B, respectively. The mean surgery time and operative blood loss were less in group A than those in group B. The year mean polyethylene liner wear was 0.08 mm in group A and 0.1 mm in group B, respectively. Kaplan-Meier survivorship at 4.2 years was 95% in group A and 93% in group B with radiographic loosening as the end point. [Conclusion] Additional screw fixation in principle is not necessary in press-fit cups.

**Key words:** screw, arthroplasty, hip, cementless

### **Introduction**

Applications of cementless prosthesis in primary total hip arthroplasty replacement have been reported to obtain excellent results. However, it still remains controversial that whether screw fixation is needed when using cementless press-fit acetabular hip prosthesis. To address this problem, we followed up 76 patients (78 hips) who underwent total hip arthroplasty with cementless hemispheric and non-press-fit acetabular prosthesis in our department from August 2008 to December 2011, to investigate the clinical effect of the acetabular screw in the fixation of acetabular



prosthesis and to explore whether acetabular screw fixation is necessary in cementless total hip arthroplasty if the bone is under good condition.

## **Clinical data and methods**

### *Patients data*

From August 2008 to December 2011, 76 patients (78 hips) that need to receive total hip arthroplasty replacement but with good condition of acetabular bone were randomly divided into two groups. Group A, non-screw fixation group, is composed of 40 hips, including 21 male and 19 female with a mean age of 56 years old (range, 46-77 years) and average BMI of 21 (range, 15-32). The reasons for primary total hip arthroplasty in group A included aseptic necrosis of femoral head in 20 cases, femoral femoral neck fracture in 16 cases, ankylosing spondylitis or joint stiffness in 2 cases and osteoarthritis in 2 cases. Group B, screw fixation group, comprises of 38 patients, including 20 male and 18 female with a mean age of 54 years old (range, 21-77 years) and average BMI 21 (range, 15-32). The reasons for primary total hip arthroplasty in group B included aseptic necrosis of femoral head in 19 cases, femoral neck fracture in 19 cases and osteoarthritis in 3 cases. The age and weight of patients in group A is not significantly different from those in group B respectively ( $p>0.05$ ).

### *Surgery approaches*

All patients received the surgery by the same group of surgeons in our department. The operation was performed using a direct lateral transgluteal approach and with hip revealed in a conventional way. The operation materials included cementless and press-fit metal cup, polyethylene liner and metal femoral head in 28 mm diameter. The patients in rroup A were directly implanted with the press-fit acetabular components and the surgery took 75 min in average with 248 ml of mean intraoperative bleeding. The patients in group B were implanted with the acetabular and followed by fixation in the top with 2-3 screws. The operation took 89 min in average with 291 ml of mean introperative bleeding. The drainage tubes were removed after surgery and the patients was allowed to direct full weight bearing and do the functional exercises with weight-bearing ambulation. The average period of hospitalization for the patients was 18 days.

### *Follow-up after surgery*

The patients were followed up at 1 month, 6 months, 24 months and last follow-up, respectively after operation. The Harris hip score (HHS) was assessed, the standardised anteroposterior and lateral view radiographs were taken and the radiological changes were evaluated. A vertical line against the teardrop joining line was drawn from the hip center, which would meet the teardrop joining line at an intersection. The distance between a line joining the inferior aspect of the teardrop and the acetabular component was measured as the horizontal migration of the acetabular prosthesis. The acetablular component migration greater than  $4^{\circ}$  in rotation or greater than 4 mm in width was considered to be unstable. Criteria for acetabular

loosening included continuous radiolucencies around the cup in zones 1 to 3 with respect to Delee-Charnley categories and the incidence frequency of radiolucent lines was recorded as well. The heterotopic ossification was graded according to Brooker classification and the polyethylene wear was measured according to Dorr method [Dorr and Wan (1996)] and the correction was performed based on the femoral head diameter (28 mm).

### *Statistical analysis*

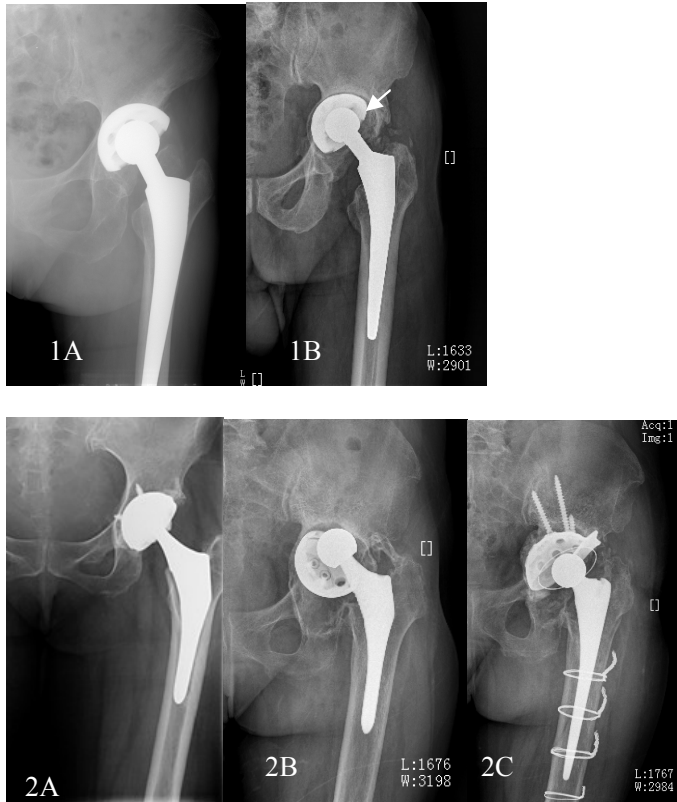
Statistical analysis was performed with the SPSS version 13.0 software. Since this was a comparative study, the results were analyzed with the paired student's t-test or chi-square test and a set point of  $p < 0.05$  was considered as statistically significant. We use Kaplan-Meier survival analysis to determine the survivorship with component loosening or revision for any reason as the endpoint. The 95% confidence intervals (95% CI) were presented and the survivorship of the prosthesis from the two groups was compared.

## **Results**

### *Clinical evaluation*

Overall, 39 hips in group A and 36 hips in group B were available for complete clinical and radiographic analysis. The average follow-up period was 4.2 years (range, 2.3-5.5 years) after surgery. The lost-to follow-up patients in group A and group B were 1 and 2, respectively. In group A the average Harris hip score increased from  $38 \pm 8.1$  before surgery to  $95 \pm 2.1$  at the last follow-up  $93 \pm 5.5$  while in group B the mean Harris hip score was improved from  $36 \pm 8.6$  preoperatively to  $95 \pm 2.1$  at the recently follow-up. It could be easily seen that the Harris hip scores of the both two groups were significantly increased after surgery ( $p < 0.01$ ).

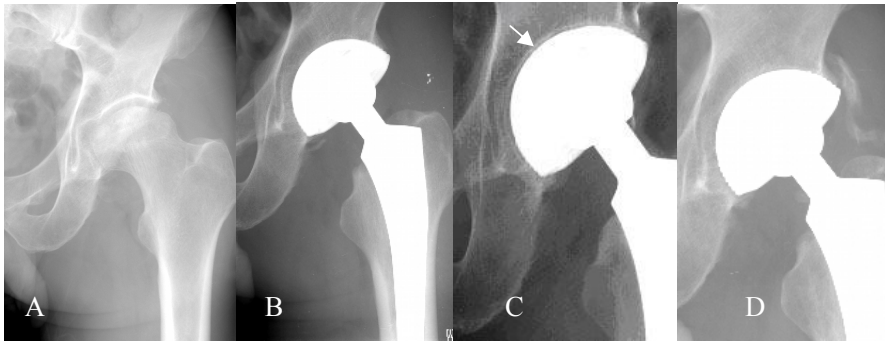
Complications associated with total hip replacement surgery were observed. In group A, one patient was found to have hip dislocation on the surgery side 20 days after surgery but the patient did not have dislocation anymore after closure. Additionally, one patient (2.6%) was reported to have osteolysis between 2 and 3 acetabular and cup aseptic loosening. But the patient only suffered from mild pain and refused revision (Figure 1). In group B, 2 patients (5.6%) had to undergo revision surgery due to loosening accompanying with moderate to severe pain. To be more specific, one of the patients suffered from septic loosening and acetabular osteolysis 2.8 years after surgery but the patient recovered well after revision and no infection occurred anymore. Overall, no patient was found to have symptomatic deep vein thrombosis or nerve damage.



**Figure 1-2 X-ray examples of hips needed revision: 1A)** A female patient with 67 years old, who suffered from fracture of the femoral neck and received cementless total hip replacement without acetabular screw fixation. **1B)** The patient was found to have acetabular osteolysis and cup migration at 4-year follow-up but she only got mild pain and refused revision. **2A)** Another female patient with 55 years old, who underwent cementless total hip replacement with screw fixation due to femoral necrosis. **2B)** The patient was found to get septic loosening and acetabular osteolysis with apparent cup migration 2.8 years after surgery. **2C)** The Harris hip score was 89 one year after revision was performed using acetabular reinforcement ring.

#### *Radiological evaluation*

In group A, radiolucent lines were present around the cup in 4 hips (10.3%), 3 hips (7.7%) and 1 hip (2.6%) at 3 months, 12 months and 24 months, respectively. At 4 years after operation, one patient had to undergo revision due to acetabular osteolysis and aseptic loosening with moderate pain. In addition, 10 hips (26%) developed heterotopic ossification and the average year polyethylene wear was 0.08 mm (range, 0-0.25 mm). By contrast, in group B, radiolucent lines appeared around the cup in 1 hip (2.8%), 2 hips (5.6%) and 2 hips (5.6%) at 3 months, 12 months and 24 months, respectively. Among them, 2 hips had to undergo revision resulting from osteolysis and migration and 11 hips shown heterotopic ossification. The average year polyethylene wear was 0.10 mm (range, 0.03-0.28 mm). The X-ray examples representative of the typical cases were shown in Figure 3.



**Figure 3. X-ray examples of typical cases:** **A)** A female patient with 61 years old, who suffered from aseptic necrosis of femoral head; **B)** The patient underwent total hip replacement without screw fixation; **C)** Radiolucent line of 1 mm in width was present in acetabulum at 6-month follow-up; **D)** The radiolucent line disappeared and heterotopic ossification graded at level 3 appeared at 2-year follow-up

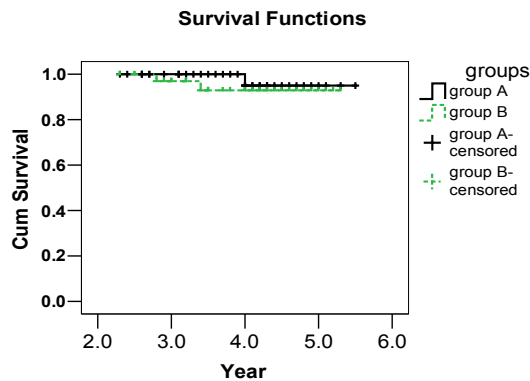
#### *Statistical data*

The clinical data of group A and B were analyzed using the Student's paired t-test or chi-square test. It could be easily seen from the table one that the operation time and blood loss during operation in non-screw fixation group were less than those in screw fixation group. Furthermore, more radiolucent lines could be observed in group A at 3-month follow-up after surgery.

**Table 1. The clinical data and statistical analysis of patients in group A and B**

<b>Evaluation</b>	<b>Group A(n=39)</b>	<b>Group B(n=6)</b>	<b><i>p</i>*</b>
Harris hip scores, mean	95	93	>0.05
Operation time in min, mean	75	89	<0.05*
Bleeding in ml, mean	248	291	<0.05*
Heterotopic ossification frequency	26%	31%	>0.05
Annual polyethylene wear in mm, mean	0.08	0.10	>0.05
Occurrence of radiolucencies (3 <sup>rd</sup> .m.)	10.3%	2.8%	<0.05*
Occurrence of radiolucencies (12 <sup>rd</sup> .m.)	7.7%	5.6%	>0.05
Occurrence of radiolucencies (24 <sup>rd</sup> .m.)	2.6%	5.6%	>0.05

The Kaplan-Meier survival analysis demonstrated that the survivorship of group A and B were 0.95 (0.91-0.99) and 0.93 (0.88-0.98), respectively. Apparently, there was no significant difference between the groups in survivorship.



**Figure 4. The Kaplan-Meier survival curve of group A and B**

### Discussion

Cementless acetabular prosthesis design and materials have greatly improved over the past decades. The improvement in smoothness of acetabular inner surface, the locking mechanism of polyethylene liner and the particle coating on cup surface are able to promote bone ingrowth, which significantly increase the survivorship of cementless acetabular components and enable the application of the primary total hip arthroplasty replacement to achieve excellent results. However, it still remains controversial when the use of screw fixation is essential when using a hemispherical press-fit acetabular prosthesis.

Supporters of the use of screws fixation believed that screw in the press-fit acetabular component could enhance the initial stability and osseointegration and prevent acetabular prosthesis migration in the long-term [Heller et al. (2013); Hsu and Lin (2010); Roth et al. (2006)]. In the vitro studies, researchers have found that the screw can increase the primary stability of acetabular cup by 26%. When the bone quality was not in an excellent condition, bone defect would have an impact on cup fixation. By contrast, the use of screws could significantly improve the initial cup stability, especially in bone with central defects [Heller et al. (2013)]. Hsu et al. [Hsu and Lin (2010)] demonstrated via their in vitro study that well-fitted screws would enhance the initial cup stability and the stability would enhance with the increase in the screws number. However, the screws would locally reduce the micromotion at the cup-bone interface. As a result, it was not advisable to place the screws close to one another. In addition, the screws should be placed peripherally and separately to expand the stable region, avoiding the micromotion at the cup-bone interface. Zilkens et al [Zilkens et al. (2011)] reported that 107 patients who underwent acetabular replacement combined with screw fixation were found to achieve good results at 2.6-year follow-up. They believed that screws fixation would not increase the occurrence of osteolysis and radiolucencies after acetabular replacement.

Opponents of the use of additional screws have argued that the stability of a press-fit acetabular prosthesis depends on surgical techniques and excellent press-fit technique rather than the use of screws fixation. The recent press-fit technique will provide adequate initial fixation so that adjunctive screw fixation is not indicated and

insufficient to prevent late migration. Furthermore, the in vitro results can only simulate the immediate state after surgery but do not represent the outcomes after bone ingrowth into the components. Won et al. [Won et al. (1995)] found that adding the screws to the component would reduce the micromotion at the superior cup but could increase it on the inferior side sometimes. Therefore, the author believed that screw fixation was unnecessary to improve the initial cup stability. Additionally, from the biomechanics point of view, it is not necessary to use the non-weight-bearing acetabular screws for fixation at the ischial or pubic bone. Garavaglia et al [Garavaglia et al. (2011)] reported that 335 patients underwent total hip arthroplasty without screws fixation. No hip had to be revised due to aseptic loosening or osteolysis during the follow-up up to 10 years and the ten-year survival rate of the prosthesis was about 98.8%. Udomkiat et al. [Udomkiat et al. (2002)] reported that 110 cases primary total hip replacements were performed with a porous-coated socket that was implanted using a press-fit technique. The aseptic loosening rate was only 0.9% at 10-year follow-up. The press-fit component could improve osseointegration to the coated without the loss of cup elasticity modulus. Furthermore, it could transmit the biomechanical force from the cup to the joint to reduce stress-shielding and osteolysis [Morscher et al. (2002)]. Röhrli et al [Röhrli et al. (2006)] reported that 50 patients were operated with cups using additional screw fixation. And 28 hips were observed to display osteolytic lesions with varying degrees mainly relate to screws. The author suggested that the joint pressure transmitted by the cup screws and the access of polyethylene wear debris into the cup-bound interface through the screw holes would result in component osteolysis, which would increase the risk of aseptic loosening.

In this study, we found that the cup radiolucent lines appeared more in non-screw fixation group at 3 months after surgery. But most of the radiolucent lines disappeared at 2-year follow-up. By contrast, in screw fixation group, the cup radiolucencies initially were similar to those found after the component reached to a stable stage. However, overall, there was no significant difference in the survival rate regardless screw fixation. Similar results were reported by other research groups. Pakvis et al. [Pakvis et al. (2012)] also reported that the cup radiolucent lines were present more within 2 months of healing period. But after the component was stable no significant difference could be observed from the two groups with or without the use of screw in two months after the healing period. Iorio et al. [Iorio et al. (2010)] carried out an investigation on 775 cases up to 10-year follow-up. He found that screw fixation did not have a favorable or adverse effect on the clinical outcome of the radiolucencies. However, several researches suggested that more radiological changes around the cup occurred in screw fixation group after bone healing period. Roth et al. [Roth et al. (2006)] divided 220 patients into two groups based on the use of screws or not. He found that radiolucent lines shown more in the non-screw fixation group at 5 months after surgery. But they disappeared and no cup migration occurred at 25-month follow-up. On contrary, more radiolucent lines were observed in the screw fixation group during 5-25 months after surgery. He suggested that the causes for more radiolucent lines were found in the screw fixation group could be explained by the

fact that the joint pressure transmitted by the cup screws and the access of polyethylene wear debris into the cup-bound interface through the screw holes. Therefore, he came to the conclusion that an additional screw fixation in principle was not necessary in press-fit cups. In this study, more radiolucent lines were found in the non-screw group at the early postoperative, which might be ascribed to the fact that the cup and bony acetabulum did not reach osseointegration at this stage. However, after the healing period, no significant difference in radiolucencies was found between the non-screw fixation and screw fixation group. In addition, the radiolucencies and osteolysis were inclined to increase in the screw group.

This study suggested that the acetabular screws could improve the initial stability of the cup prior to healing period. But after healing period, it made no significant difference to the component stability regardless of the use of screw fixation. However, it should be noted that the non-screw fixation procedure would reduce the operative time, blood loss and operation fees. Additionally, it could avoid injury on blood vessels and nerves caused by screws fixation and it also made it easier to perform possible revision. Therefore, the screwless fixation procedure might favor patients when the patient bone condition allowed doing it. Nevertheless, screws fixation should be carried out to improve the component stability if the following symptoms appeared, including osteomalacia after osteolysis, intraoperative acetabular fractures, acetabular dysplasia and acetabular bone defects on the edges in which good press-fit and initial stability could not be satisfied.

The authors are aware that this investigation has some limitations, such as an inadequate cases and a more subjective evaluation which would probably generate bring data bias. In recent years, some authors have advocated to use objective tests, such as using RSA (Radiostereometric Analysis) and EBBA (Einzel Bild Roentgen Analyse) to measure the component migration or use of computer-aided quantitative measurement software to analyze the radiological images [Garavaglia et al. (2011); Röhrli et al. (2006; Zilkens et al. (2011)]. However, the present study still utilizes the conventional analysis strategy. Furthermore, another limitation of this study is a short follow-up period. So a long-term follow-up should be performed.

In summary, this study suggested that when performing the total hip arthroplasty replacement, the inherent stability of the cup could be achieved via press-fit technique. In addition, the use of screw fixation did not improve the stability of the cup when the acetabular bone was in good quality condition. Thus, we arrive to the conclusion that screws are not necessary for the total hip arthroplasty replacement.

## Reference

- Dorr, L. D. and Wan, Z. (1996) Comparative results of a distal modular sleeve, circumferential coating, and stiffness relief using the Anatomic Porous Replacement II, *The Journal of Arthroplasty* **11**, 419-428.
- Garavaglia, G, Lübbecke, A., Barea, C., Roussos, C., Peter, R. and Hoffmeyer, P. (2011) Ten-year results with the Morscher press-fit cup: an uncemented, non-modular, porous-coated cup inserted

- without screws, *International Orthopaedics* **35**, 957-963.
- Heller, S., Brosh, T., Kosashvili, Y., Velkes, S., Burg, A. and Dudkiewicz, I. (2013) Locking versus standard screw fixation for acetabular cups: is there a difference?, *Archives of Orthopaedic and Trauma Surgery* **133**, 701-705.
- Hsu, J.-T. and Lin, D.-J. (2010) Effects of Screw Eccentricity on the Initial Stability of the Acetabular Cup in Artificial Foam Bone of Different Qualities, *Artificial Organs* **34**, E10-E16.
- Iorio, R., Puskas, B., Healy, W. L., Tilzey, J. F., Specht, L. M. and Thompson, M. S. (2010) Cementless Acetabular Fixation With and Without Screws: Analysis of Stability and Migration, *The Journal of Arthroplasty* **25**, 309-313.
- Morscher, E. W., Widmer, K. H., Bereiter, H., Elke, R. and Schenk, R. (2002) [Cementless socket fixation based on the "press-fit" concept in total hip joint arthroplasty], *Acta Chir Orthop Traumatol Cech* **69**, 8-15.
- Pakvis, D., Luites, J., Hellemond, G. v. and Spruit, M. (2012) A cementless, elastic press-fit socket with and without screws, *Acta Orthopaedica* **83**, 481-487.
- Röhrli, S. M., Nivbrant, B., Snorrason, F., Kärrholm, J. and Nilsson, K. G. (2006) Porous-coated cups fixed with screws: A 12-year clinical and radiostereometric follow-up study of 50 hips, *Acta Orthopaedica* **77**, 393-401.
- Roth, A., Winzer, T., Sander, K., Anders, J. O. and Venbrocks, R. A. (2006) Press fit fixation of cementless cups: how much stability do we need indeed?, *Archives of Orthopaedic and Trauma Surgery* **126**, 77-81.
- Udomkiat, P., Dorr, L. D. and Wan, Z. (2002). Cementless Hemispheric Porous-Coated Sockets Implanted with Press-Fit Technique without Screws: Average Ten-Year Follow-up, Vol 84.
- Won, C. H., Hearn, T. C. and Tile, M. (1995) Micromotion of cementless hemispherical acetabular components. Does press-fit need adjunctive screw fixation?, *Journal of Bone & Joint Surgery, British Volume* **77-B**, 484-489.
- Zilkens, C., Djalali, S., Bittersohl, B., Kalicke, T., Kraft, C. N., Krauspe, R. and Jager, M. (2011) Migration pattern of cementless press fit cups in the presence of stabilizing screws in total hip arthroplasty, *Eur J Med Res* **16**, 127-132.



## Global sensitivity analysis for structural models by sparse grid integration

†Zhou Changcong, Zhang Feng, and Wang Wenxuan

School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China

†Corresponding author: zccput@163.com

### Abstract

Global sensitivity indices based on variance can effectively apportion the output uncertainty to the inputs. How to efficiently and accurately perform the global sensitivity analysis is of great concern for researchers. In this work, the employment of sparse grid integration to the estimate of global sensitivity indices is discussed. The new method can be used for sensitivity analysis of the structural models involving independent variables or correlated variables, and can further decompose the variance contribution in the correlation cases. Advantage of the sparse grid integration in estimating integrals is well inherited by the new method, to ensure the accuracy while keeping the computational burden controllable. Numerical and engineering examples have been studied to test the applicability of the proposed method.

**Keywords:** Sensitivity analysis; variance; uncertainty; sparse grid integration

### 1. Introduction

Global sensitivity indices are playing an important role in identifying and representing uncertainties in engineering, and many researchers have proposed their own indices, as well as corresponding computing techniques [Borgonovo (2007); Sobol (2001)]. Among these different indices, the variance based ones have attracted increasing interests as they are able to capture the influence of the full range of variation of each input factor, and reflect the interaction effects among input factors. Variance based sensitivity analysis has been acknowledged as a versatile and effective tool in the uncertainty analysis.

In the real world, input factors of a model are often correlated to each other, and sometimes the correlation may have significant impacts on the sensitivity results [Borgonovo and Tarantola (2008); Kucherenko et al. (2012); Mara (2009)]. Xu and Gertner (2008) pointed out that the contribution of uncertainty to the output by an individual input factor should be decomposed into two parts: the uncorrelated part, which means this part is completely immune from the other input factors and is produced by this input factor “individually and independently”, and the correlated part, which means this part is produced by the correlation of this input variable with the others. Mara and Tarantola (2012) proposed a set of variance-based sensitivity indices to perform sensitivity analysis of models with correlated inputs, which can distinguish between the mutual dependent contribution and the independent contribution of an input to the model response variance.

An important task in the sensitivity analysis is to improve the computational efficiency, especially in engineering cases where the models involved usually take a long processing time [Saltelli et al. (2010)]. It is found that the variance based sensitivity indices can be viewed as nested expressions of expectation operator and variance operator. Enlightened by this feature, in this work a new method is

developed for global sensitivity analysis by using the sparse grid integration (SGI) technique. The sparse grid technique, which is based on one-dimensional formulae and then extended to higher dimensions, has been extensively utilized for highly dimensional multivariate integration, as well as interpolation [Barthelmann et al. (2000); Gerstner and Griebel (2003); Gerstner and Griebel (1998); Smolyak (1963)]. To a certain extent, this technique avoids the “curse of dimension” of conventional integration algorithms, which means the computational cost grows exponentially with the dimension of the problem. In the SGI technique, multivariate quadrature formulae are constructed with combinations of tensor products of suitable one-dimensional formulae, thus the function evaluations needed and the numerical accuracy become independent of the dimension of the problem up to logarithmic factors. Existing literature has reported the high efficiency and accuracy of sparse grid applications.

## 2. Review on variance based sensitivity indices

### 2.1 Sensitivity analysis of model output with independent inputs

Let  $Y = g(\mathbf{X})$  be the performance function of the model under investigation, with  $Y$  the output,  $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$  the vector of independent input variables, where  $X_i$  is the  $i$ th variable. Sobol (2001) proposed that the performance function can always be decomposed into summands of different dimensions, that is

$$g(\mathbf{X}) = g_0 + \sum_{i=1}^n g_i(X_i) + \sum_{i<j}^n g_{i,j}(X_i, X_j) + \dots + g_{1,2,\dots,n}(X_1, X_2, \dots, X_n) \quad (1)$$

Sensitivity analysis based on variance is to quantify the contribution of an individual input variable to the output variance, and Sobol proposed the variance decomposition equation based on Eq.(1),

$$V(Y) = \sum_{i=1}^n V_i + \sum_{i=1, j>i}^n V_{ij} + \dots + V_{1,2,\dots,n} \quad (2)$$

$V_i$  is the first order variance contribution of  $X_i$ , and can be formulated as

$$V_i = V_{X_i}(E_{X_{-i}}(Y | X_i)) \quad (3)$$

where  $\mathbf{X}_{-i}$  denotes the vector of all input variables except  $X_i$ , i.e.  $\mathbf{X}_{-i} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)^T$ .  $V_{ij}$  and higher order variance terms in Eq.(2) denote the contribution to the output variance of variable interaction brought by the form of the performance function. When only the first order variance contribution is considered, the variance decomposition can be reformulated as

$$V(Y) = \sum_{i=1}^n V_i \quad (4)$$

The first order variance contribution  $V_i$  is also referred to as the main effect of  $X_i$  on the output variance, and it measures the first order effect of  $X_i$  on the output, ignoring the interactions between  $X_i$  and the other variables. When taking the interactions into consideration, the total contribution of  $X_i$  is measured by  $E_{X_{-i}}(V_{X_i}(Y | \mathbf{X}_{-i}))$ . According to the known identity:

$$V_{X_{-i}}(E_{X_i}(Y | \mathbf{X}_{-i})) + E_{X_{-i}}(V_{X_i}(Y | \mathbf{X}_{-i})) = V(Y) \quad (5)$$

$V_{X_{-i}}(E_{X_i}(Y | \mathbf{X}_{-i}))$  can be seen as the first order effect of  $\mathbf{X}_{-i}$ , thus  $V(Y)$  minus  $V_{X_{-i}}(E_{X_i}(Y | \mathbf{X}_{-i}))$  should give the contribution of all terms in the variance

decomposition which includes  $X_i$ .

To normalize the variance contribution, the main effect index is defined as [Saltelli et al. (2010)]

$$S_i^M = \frac{V_{X_i}(E_{X_{-i}}(Y | X_i))}{V(Y)} \quad (6)$$

and the total effect index is defined as

$$S_i^T = \frac{E_{X_{-i}}(V_{X_i}(Y | X_{-i}))}{V(Y)} \quad (7)$$

## 2.2 Sensitivity analysis of model output with correlated inputs

The variance decomposition in Eq.(2) is proposed under the assumption of input independence, which thus means it may not hold when the input variables are correlated. For correlated variables, the main effect index and total effect index still reflect the variance contribution of variables, yet the difference from the independence case lies in that both indices are now composed by two parts. Take  $V_i$  for example, according to Xu and Gertner (2008), it can be divided into two parts, i.e. the variance contribution of the uncorrelated part of  $X_i$ , denoted by  $V_i^U$ , and the variance contribution of the correlated part with the other variables, denoted by  $V_i^C$ . It should be pointed out that the variance contribution of the correlated part of  $X_i$  is way different from the interaction contribution of  $X_i$  to the output variance. These two are absolutely different concepts, as the former comes from the correlation among the input variables, and the latter is produced by variable interactions in the performance function.

Thus, the main effect index,  $S_i^M$ , can be decomposed in the following way,

$$S_i^M = S_i^{MC} + S_i^{MU} \quad (8)$$

where  $S_i^{MC}$  denotes the correlated main contribution of  $X_i$  to the output variance, and  $S_i^{MU}$  denotes the uncorrelated main contribution. Similarly, the total effect index,  $S_i^T$ , can be decomposed as

$$S_i^T = S_i^{TC} + S_i^{TU} \quad (9)$$

where  $S_i^{TC}$  denotes the correlated total contribution of  $X_i$  to the output variance, and  $S_i^{TU}$  denotes the uncorrelated total contribution.

In fact, in the existing literature measuring the contributions of inputs to the output variance for cases involving correlated inputs is still a tricky issue. Researchers have proposed different variance based sensitivity indices based on different considerations, and it is difficult to judge which one is better. Agreement is highly needed to give an exact and unambiguous definition of the ANOVA for correlated inputs just as the one provided by Sobol decomposition when the inputs are independent. In this work, the sensitivity indices talked about are related to those in the work of Mara and Tarantola (2012).

## 3. Variance based sensitivity analysis with SGI

The SGI technique has been proven to be an effective tool in the uncertainty analysis, and it will be employed to perform the variance based sensitivity in this section. It

should be pointed out that normally the SGI procedure is carried out in the independent space, which means it cannot be directly used when input variables are correlated. In this section, the procedure using SGI for sensitivity analysis of models with independent variables is first introduced. Afterwards, the use of SGI for variance based sensitivity analysis with correlated variables is discussed, which is partly based on the work of Mara and Tarantola (2012).

### 3.1 Algorithm based on SGI for sensitivity analysis of independent variables

For the performance function  $Y = g(\mathbf{X})$  with independent variables, the expectation and variance of the output can be estimated by SGI according to the following formulae,

$$E(Y) = \int g(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \approx \sum_{j=1}^p w_j g(\mathbf{x}_j) \quad (10)$$

$$V(Y) = \int (g(\mathbf{x}) - E(Y))^2 f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} \approx \sum_{j=1}^p w_j (g(\mathbf{x}_j) - E(Y))^2 \quad (11)$$

where the  $n$ -dimension quadrature point  $\mathbf{x}_j$  and the associated weight  $w_j$  ( $j=1,2,\dots,p$ ) are obtained by the Smolyak algorithm [Gerstner and Griebel (2003); Gerstner and Griebel (1998)].

With the output variance obtained,  $V_{X_i}(E_{X_{-i}}(Y|X_i))$  and  $E_{X_{-i}}(V_{X_i}(Y|X_{-i}))$ , both of which can be seen as nested expressions of expectation operator and variance operator, need to be estimated to get the indices  $S_i^M$  and  $S_i^T$ .

Consider  $V_{X_i}(E_{X_{-i}}(Y|X_i))$  first. Keep  $X_i$  fixed, and treat  $X_{-i}$  as variables, then  $Y$  can be seen as the function of  $X_{-i}$ . The inner expectation can be estimated by SGI as follows,

$$E_{X_{-i}}(Y|X_i) = \int g(\mathbf{x}_{-i}, x_i) f_{X_{-i}}(\mathbf{x}_{-i}) d\mathbf{x}_{-i} \approx \sum_{j=1}^m w_j g(\mathbf{x}_{-i}^j, x_i) \quad (12)$$

where  $\mathbf{x}_{-i}^j$  denotes the  $j$ -th value of  $X_{-i}$ . Apparently  $E_{X_{-i}}(Y|X_i)$  can be seen as a univariate function of  $X_i$ , and is redefined as

$$\varphi(X_i) = E_{X_{-i}}(Y|X_i) \quad (13)$$

Thus,

$$V_{X_i}(E_{X_{-i}}(Y|X_i)) = V(\varphi(X_i)) \quad (14)$$

The variance of the univariate function  $\varphi(X_i)$  can be estimated by SGI as follows,

$$V(\varphi(X_i)) = \int (\varphi(x_i) - E(\varphi(x_i)))^2 f_{X_i}(x_i) dx_i \approx \sum_{j=1}^{\tilde{m}} w_j (g(x_i^j) - E(\varphi(X_i)))^2 \quad (15)$$

where

$$E(\varphi(X_i)) = \int \varphi(x_i) f_{X_i}(x_i) dx_i \approx \sum_{j=1}^{\tilde{m}} w_j g(x_i^j) \quad (16)$$

Finally, the main effect index can be obtained as

$$S_i^M = \frac{V_{X_i}(E_{X_{-i}}(Y|X_i))}{V(Y)} = \frac{V(\varphi(X_i))}{V(Y)} \quad (17)$$

The total effect index  $S_i^T$  can be obtained in the similar manner.

### 3.2 Algorithm based on SGI for sensitivity analysis of correlated variables

When the input variables are correlated, the algorithm based on SGI in Section 3.1 is

no longer applicable. In this section, the SGI technique is extended to the case of correlated variables to perform the variance based sensitivity analysis. In this work, the correlation matrix, which is a symmetric matrix composed of Pearson correlation coefficients, is adopted to measure the correlation between random variables.

### 3.2.1 Estimate of the response variance

Eq.(10) and Eq.(11) in Section 3.1 can be used to estimate the expectation and variance of the model output only when the variables are independent. Thus, for correlated variables there has to be a transformation from correlation space to independence space. In this work, we only consider the problems with normal distributions, i.e.  $X_i \sim N(\mu_{X_i}, \sigma_{X_i}^2)$ .

The covariance matrix of  $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$  is denoted as

$$\mathbf{C}_X = \begin{pmatrix} \sigma_{X_1}^2 & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & \sigma_{X_2}^2 & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & \sigma_{X_n}^2 \end{pmatrix} \quad (18)$$

where  $Cov(X_i, X_j) = Cov(X_j, X_i) = \rho_{ij} \sigma_{X_i} \sigma_{X_j}$ , as  $\rho_{ij}$  is the Pearson correlation coefficient of  $X_i$  and  $X_j$ . The joint PDF of  $\mathbf{X}$  is

$$f(\mathbf{X}) = (2\pi)^{-\frac{n}{2}} |\mathbf{C}_X|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{X} - \boldsymbol{\mu}_X)^T \mathbf{C}_X^{-1} (\mathbf{X} - \boldsymbol{\mu}_X)\right\} \quad (19)$$

where  $\boldsymbol{\mu}_X = (\mu_{X_1}, \mu_{X_2}, \dots, \mu_{X_n})^T$  is the vector of input expectations,  $|\mathbf{C}_X|$  and  $\mathbf{C}_X^{-1}$  are the determinant value and reverse matrix of  $\mathbf{C}_X$  respectively.

An orthogonal matrix  $\mathbf{A}$  exists which would introduce random variable vector  $\mathbf{U} = (U_1, U_2, \dots, U_n)^T$  by the following formula [Shi et al. (2009)]:

$$f(\mathbf{A}\mathbf{U} + \boldsymbol{\mu}_X) = (2\pi)^{-\frac{n}{2}} (\lambda_1 \lambda_2 \cdots \lambda_n)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^n \frac{U_i^2}{\lambda_i}\right) \quad (20)$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  is the latent roots of  $\mathbf{C}_X$ .

Thus the  $n$ -dimension correlated input variables  $\mathbf{X} = (X_1, X_2, \dots, X_n)^T$  can be transformed to independent normal variables  $\mathbf{U} = (U_1, U_2, \dots, U_n)^T$  by

$$\mathbf{U} = \mathbf{A}^T (\mathbf{X} - \boldsymbol{\mu}_X) \quad (21)$$

where the column vectors of  $\mathbf{A}$  are the latent root vectors of  $\mathbf{C}_X$ , and the probability distribution of the independent normal variables can be obtained as  $U_i \sim N(0, \lambda_i)$ .

By Eq.(21), we further get

$$\mathbf{X} = \mathbf{A}\mathbf{U} + \boldsymbol{\mu}_X \quad (22)$$

Substitute Eq.(22) into  $Y = g(\mathbf{X})$ , the performance function in the independence space can be obtained as follows,

$$Y = g(\mathbf{X}) = \phi(\mathbf{U}) \quad (23)$$

where  $\phi(\bullet)$  denotes the mapping relationship between  $\mathbf{U}$  and the output. Because distribution parameters of  $\mathbf{U}$  in the dependence space have been obtained, the expectation and variance of the model output can be readily estimated by Eq.(10) and

Eq.(11).

### 3.2.2 Orthogonalization of correlated variables

For two normal variables,  $X_i$  and  $X_j$ ,  $E(X_i|X_j)$  can be used to define the marginal relationship between them. The value of  $E(X_i|X_j)$  measures how much  $X_i$  is correlated to  $X_j$ , and equals  $E(X_i)$  when the two variables are independent. In the same manner,  $E(X_i|X_j, X_k)$  defines the correlation of  $X_i$  to  $X_j$  and  $X_k$ . Mara and Tarantola declared that for non-normal random variables, higher conditional moments are needed to characterize the correlation between the variables. Issues concerning correlated non-normal variables would become much complicated in most cases. In this work, the scope of research is confined to problems only involving normally distributed variables.

For a vector of correlated normal variables,  $\mathbf{X}=(X_1, X_2, \dots, X_n)^T$ , the following relationship holds,

$$f(X_1, X_2, \dots, X_n) = f(X_1)f(X_2|X_1)\cdots f(X_n|X_1, X_2, \dots, X_{n-1}) \quad (24)$$

where  $f(X_1, X_2, \dots, X_n)$  is the joint PDF,  $f(X_1)$  is the marginal PDF of  $X_1$ ,  $f(X_2|X_1)$  is the marginal PDF of  $X_2$  conditioned on  $X_1$ ,  $f(X_n|X_1, X_2, \dots, X_{n-1})$  is the marginal PDF of  $X_n$  conditioned on  $(X_1, X_2, \dots, X_{n-1})^T$ . As we have talked about,  $E(X_i|X_j)$  can quantify the marginal relationship between two correlated normal variables, thus the following equations hold,

$$\begin{aligned} \bar{X}_1 &= X_1 \\ \bar{X}_2 &= X_{2-1} = X_2 - E(X_2|X_1) \\ \bar{X}_3 &= X_{3-12} = X_3 - E(X_3|X_1, X_2) \\ &\dots \\ \bar{X}_n &= X_{n-12\dots(n-1)} = X_n - E(X_n|X_1, X_2, \dots, X_{n-1}) \end{aligned} \quad (25)$$

Mara and Tarantola (2012) pointed out that the above transformation is one of Rosenblatt's for normal variables. A new set of variables  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^T$  can be generated by the above transformation. In fact, these new variables are obtained by subtracting the correlation part from the original correlated variables, thus the new variables are independent from each other.

Besides, the orthogonalization from correlated variables to independent ones is not unique. This can be seen from the transformation in Eq.(25), which clearly depends on the ordering of variables. If we reorder the original input variables as  $\mathbf{X}=(X_2, \dots, X_n, X_1)$ , which is valid except the corresponding correlation matrix needs to be modified accordingly, and apply the orthogonalization, a new set of independent variables will be obtained and denoted as  $(\bar{X}_2, \dots, \bar{X}_n, \bar{X}_1)$ . It should be reminded that  $(\bar{X}_2, \dots, \bar{X}_n, \bar{X}_1)$  is different from  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)$  initially obtained by orthogonalizing  $(X_1, X_2, \dots, X_n)$ . In fact, through changing the ordering of the original correlated variables in sequence, a total of  $n!$  sets of independent variables can be generated by the orthogonalization. However, among all these sets, only a total of  $n$  sets generated by cycling the orthogonalization are used in this work, i.e.  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)$ ,  $(\bar{X}_2, \dots, \bar{X}_n, \bar{X}_1)$ ,  $\dots$ ,  $(\bar{X}_n, \dots, \bar{X}_{n-2}, \bar{X}_{n-1})$ .

### 3.2.3 Interpretations on sensitivity indices of the newly independent variables

It is necessary to find out the relationship between the sensitivity indices of the independent variables and those of correlated ones. Let us focus on the main effect index first, which is denoted as

$$\begin{aligned}
 \bar{S}_1^M &= V[E(Y | \bar{X}_1)] / V(Y) \\
 \bar{S}_2^M &= V[E(Y | \bar{X}_2)] / V(Y) \\
 \bar{S}_3^M &= V[E(Y | \bar{X}_3)] / V(Y) \\
 &\dots \\
 \bar{S}_n^M &= V[E(Y | \bar{X}_n)] / V(Y)
 \end{aligned} \tag{26}$$

Now consider the explanation of the sensitivity indices. Because  $\bar{X}_1 = X_1$ ,  $\bar{S}_1^M$  is the full marginal contribution of  $X_1$  to the response variance, which means  $\bar{S}_1^M = S_1^M$ .  $\bar{S}_2^M$  is the marginal contribution of  $X_2$  to the response variance without its correlative contribution with  $X_1$ , as  $\bar{X}_2$  is uncorrelated with  $X_1$ . In the same manner,  $\bar{S}_3^M$  is the marginal contribution of  $X_3$  to the response variance without its correlative contribution with  $(X_1, X_2)^T$ . For the last new variable  $\bar{X}_n$ ,  $\bar{S}_n^M$  is the uncorrelated marginal contribution to the response variance, which means  $\bar{S}_n^M = S_n^{MU}$ .  $\bar{X}_1$  keeps all the information concerning  $X_1$  including its correlated part with the other variables while  $\bar{X}_n$  only keeps the independent part of  $X_n$  excluding all of its correlated part.

Clearly, by the variance based sensitivity analysis on  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^T$ , the main effect index of  $X_1$ , i.e.  $S_1^M$ , as well as the uncorrelated main effect index of  $X_n$ , i.e.  $S_n^{MU}$ , can be obtained. Remember that a total of  $n$  sets of independent variables can be generated by orthogonalizing the original correlated variables in cycle. In the same way of obtaining  $S_1^M$  and  $S_n^{MU}$ , the indices  $S_2^M$  and  $S_1^{MU}$  can be obtained by the sensitivity analysis on  $(\bar{X}_2, \dots, \bar{X}_n, \bar{X}_1)$ . Change the order of the original correlated variables, and perform orthogonalization to get the corresponding independent variables, which is further analyzed to get the sensitivity indices, the full and uncorrelated contributions of each correlated variable can be obtained. Similarly, the total effect index  $S_i^T$  of the correlated inputs can be thus estimated and decomposed.

### 3.2.4 Computational issues

In this section, the computational issues involved in the procedure of sensitivity analysis are addressed. Still, we take the analysis on  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^T$  as an example. To analyze the sensitivity indices on  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^T$ , firstly the statistical characteristics have to be known.

By Eq.(25) we know the mean value and standard deviation of  $\bar{X}_1$  are equal to those of  $X_1$ . For the  $i$ th ( $i > 1$ ) variable, the following equation holds according to Eq.(25),

$$\bar{X}_i = X_i - E(X_i | X_1, X_2, \dots, X_{i-1}) \tag{27}$$

Only the first  $i$  variables are involved in Eq.(27), i.e.  $\mathbf{X}_{1-i} = (X_1, X_2, \dots, X_i)^T$ , of which the mean vector is  $\boldsymbol{\mu}_{\mathbf{X}_{1-i}} = (\mu_{X_1}, \mu_{X_2}, \dots, \mu_{X_i})^T$ , and the corresponding covariance matrix can be taken from Eq.(18) as

$$\mathbf{C}_{X_{1:i}} = \begin{pmatrix} \sigma_{X_1}^2 & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_i) \\ \text{Cov}(X_2, X_1) & \sigma_{X_2}^2 & \cdots & \text{Cov}(X_2, X_i) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_i, X_1) & \text{Cov}(X_i, X_2) & \cdots & \sigma_{X_i}^2 \end{pmatrix} \quad (28)$$

Now consider the conditional mean value in Eq.(27),  $E(X_i | X_1, X_2, \dots, X_{i-1})$ . The mean vector and covariance matrix for  $X_{1:i}$  can be rewritten as  $\boldsymbol{\mu}_{X_{1:i}} = \begin{bmatrix} \mu_{X_i} \\ \boldsymbol{\mu}_{X_{1:(i-1)}} \end{bmatrix}$  and

$$\mathbf{C}_{X_{1:i}} = \begin{bmatrix} \mathbf{C}_{X_i X_i} & \mathbf{C}_{X_i X_{1:(i-1)}} \\ \mathbf{C}_{X_{1:(i-1)} X_i} & \mathbf{C}_{X_{1:(i-1)} X_{1:(i-1)}} \end{bmatrix}, \text{ where}$$

$$\mathbf{C}_{X_i X_i} = \sigma_{X_i}^2 \quad (29)$$

$$\mathbf{C}_{X_{1:(i-1)} X_i}^T = \mathbf{C}_{X_i X_{1:(i-1)}} = [\text{Cov}(X_i, X_1) \quad \text{Cov}(X_i, X_2) \quad \cdots \quad \text{Cov}(X_i, X_{i-1})] \quad (30)$$

$$\mathbf{C}_{X_{1:(i-1)} X_{1:(i-1)}} = \begin{pmatrix} \sigma_{X_1}^2 & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_{i-1}) \\ \text{Cov}(X_2, X_1) & \sigma_{X_2}^2 & \cdots & \text{Cov}(X_2, X_{i-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_{i-1}, X_1) & \text{Cov}(X_{i-1}, X_2) & \cdots & \sigma_{X_{i-1}}^2 \end{pmatrix} \quad (31)$$

The conditional distribution of  $X_i$  conditioned on  $(X_1, X_2, \dots, X_{i-1})^T$  is still a normal distribution, and its mean value can be calculated in the following way [Shi et al. (2009)],

$$E(X_i | X_1, X_2, \dots, X_{i-1}) = \mu_{X_i} + \mathbf{C}_{X_i X_{1:(i-1)}} \mathbf{C}_{X_{1:(i-1)} X_{1:(i-1)}}^{-1} (\mathbf{X}_{1:(i-1)} - \boldsymbol{\mu}_{X_{1:(i-1)}}) \quad (32)$$

From the above equation, it can be seen that  $E(X_i | X_1, X_2, \dots, X_{i-1})$  is in fact an expression containing the first  $(i-1)$  variables. Substitute Eq.(32) into Eq.(27), we can get an expression containing the first  $i$  variables. In other words,  $\bar{X}_i$  can be viewed as a function of the first  $i$  correlated variables, of which the expression can be explicitly obtained. In the same way in Section 3.2.1, the mean value and standard deviation of  $\bar{X}_i$  can be conveniently obtained using the SGI technique, which are denoted as  $\mu_{\bar{X}_i}$  and  $\sigma_{\bar{X}_i}$ , respectively. After obtaining the mean and standard deviation of the independent variables, the sensitivity analysis still cannot be performed immediately, as the original performance function is a mapping of the output response with the original correlated variables, not with the new independent ones. Thus the performance function has to be rebuilt to describe the relationship between the output response and the new independent variables.

From Eq.(25), the following relationship holds,

$$\begin{aligned} X_1 &= \bar{X}_1 \\ X_2 &= \bar{X}_2 + E(X_2 | X_1) \\ X_3 &= \bar{X}_3 + E(X_3 | X_1, X_2) \\ &\cdots \\ X_n &= \bar{X}_n + E(X_n | X_1, X_2, \dots, X_{n-1}) \end{aligned} \quad (33)$$



As discussed above,  $E(X_2 | X_1)$  is an expression only containing  $X_1$ . Now substitute  $X_1 = \bar{X}_1$  into  $E(X_2 | X_1)$ , then  $X_2$  can be transformed into an expression only containing  $\bar{X}_1$  and  $\bar{X}_2$ . Similarly, substitute the expression of  $X_1$  and  $X_2$  into  $E(X_3 | X_1, X_2)$ , then  $X_3$  can be transformed into an expression only containing  $\bar{X}_1$ ,  $\bar{X}_2$  and  $\bar{X}_3$ . Repeat this process, the expression for  $X_n$  can be finally obtained, which should only involve  $(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n)^T$ . This process can be denoted as

$$\begin{aligned} X_1 &= \beta_1(\bar{X}_1) \\ X_2 &= \beta_2(\bar{X}_2) \\ X_3 &= \beta_2(\bar{X}_1, \bar{X}_2, \bar{X}_3) \\ &\dots \\ X_n &= \beta_n(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) \end{aligned} \quad (34)$$

in which  $\beta_i$  denotes a mapping relationship.

Substitute Eq.(34) into the original performance function, a new performance function between the output response and the new independent variables can be obtained, which is denoted as

$$Y = \bar{g}(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) \quad (35)$$

According to the previous discussions, variance based sensitivity analysis can be easily performed for the performance function in Eq.(35). The sensitivity indices thus obtained are then used to interpret the main effect index, total effect index, as well as the decompositions, for the original correlated variables according to Section 3.2.3.

#### 4. An automobile front axle

In the automobile engineering, the front axle is an important component that bears heavy loads. Due to the rigid requirements for its strength, stiffness and fatigue life, mechanical property of the front axle must be strictly tested before the mass production [Lu et al. (2012)]. Variance can act as an important index to measure the robustness of the front axle, considering the uncertainty existing in the structure. The I-beam structure is widely used in the design of front axle due to its high bending strength and light weight. Consider the I-beam structure shown in Figure 1. The maximum normal stress and shear stress are  $\sigma = M / W_x$  and  $\tau = T / W_\rho$  respectively, where  $M$  is the bending moment,  $T$  is the torque,  $W_x$  and  $W_\rho$  are the sectional factor and polar sectional factor given as

$$W_x = \frac{a(h-2t)^3}{6h} + \frac{b}{6h} [h^3 - (h-2t)^3] \quad (36)$$

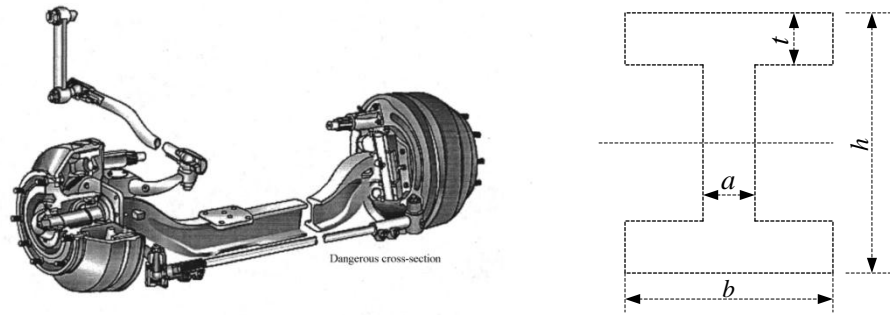
$$W_\rho = 0.8bt^2 + 0.4[a^3(h-2t)/t] \quad (37)$$

Consider the static strength of the front axle, the performance function can be thus established as

$$g = \sigma_s - \sqrt{\sigma^2 + 3\tau^2} \quad (38)$$

where  $\sigma_s$  is the yielding stress, and  $\sigma_s = 460$  MPa according to the material property. In the real engineering, uncertainty is unavoidable in the manufacture process, and randomness exists in the external loads. In this example, the geometry parameters of the I-beam, i.e.  $a$ ,  $b$ ,  $t$ ,  $m$ , and the loads  $M$  and  $T$  are taken as random variables. The

probability distribution information is given in Table 1.



**Figure 1. Sketch of the automobile front axle**

**Table 1. Distribution information of the inputs for the I-beam**

Input (unit)	Distribution type	Mean	Standard deviation
$a$ (mm)	Normal	12	0.06
$b$ (mm)	Normal	65	0.325
$t$ (mm)	Normal	14	0.07
$h$ (mm)	Normal	85	0.425
$M$ (N·mm)	Normal	$3.5 \times 10^6$	$1.75 \times 10^4$
$T$ (N·mm)	Normal	$3.1 \times 10^6$	$1.55 \times 10^4$

By the above illustration, we may get the impression that the nonlinearity of the performance function is high, thus the interaction effect on the output variance should be noticeable. However, conclusions might be different if we perform the quantitative sensitivity analysis. First, consider the contribution of the inputs to the output variance under the assumption that the inputs are independent from each other. The results obtained by the proposed method and MCS are reported in Table 2.

When the inputs are independent, a total of 451 runs of the performance function are needed by the proposed method, and comparison with MCS shows the accuracy of the proposed method is acceptable. Another important feature in the sensitivity results is that, for each input, the main effect index is very close to the total effect index. It indicates the effect of interactions between inputs on the output variance is negligible, despite of the nonlinearity of the performance function. The independent inputs can be ranked as  $\{t, T, b, a, h, M\}$  in the descending order according to their contributions to the output variance.

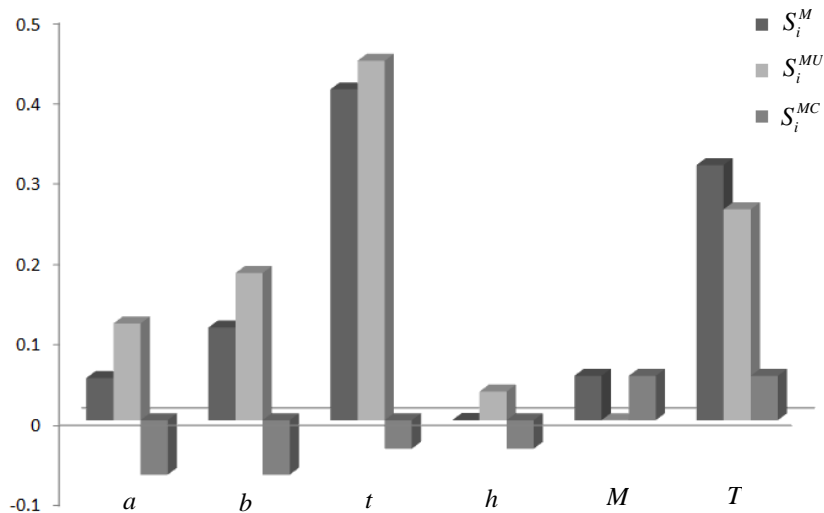
**Table 2 Sensitivity indices of the independent inputs for the I-beam**

		$a$	$b$	$t$	$h$	$M$	$T$
$S_i^M$	Proposed method	0.112	0.171	0.418	0.033	0.0001	0.265
	MCS	0.115	0.172	0.410	0.032	0.0004	0.262
$S_i^T$	Proposed method	0.113	0.171	0.418	0.033	0.0001	0.265
	MCS	0.117	0.171	0.417	0.033	0.0001	0.269

As a matter of fact, in the real engineering it is inappropriate to assume the independence among the inputs. In most cases, one input is probably correlated to another, and such correlation may have notable effect on the output performance. Now consider the sensitivity analysis on the I-beam structure under the assumption of

input correlation. Assume  $\rho_{ab} = -0.3$ ,  $\rho_{th} = -0.3$ , and  $\rho_{MT} = 0.4$ . Since interactions between the inputs have little effect on the output variance, thus only the main effect index is estimated and decomposed by the proposed method, and shown in Figure 2.

The sensitivity analysis results in Figure 2 can provide helpful interpretations of the model. Still, the input  $t$  and  $T$  contribute the most to the output variance, as is the case when the inputs are independent. Thus, these two inputs should be carefully controlled if we want to reduce the output variation, especially the former one. Besides,  $M$  becomes more important than  $h$  after the correlation is introduced. It can also be noticed that the contributions of the correlated part by the first four inputs are negative, which is caused by the negative correlation coefficient. In this example, the inputs can be seen as independent pairs of dependent variables, e.g. the input  $a$  is only correlated to  $b$  and independent from the rest. As a result, the correlated contributions of two correlated inputs are equal, e.g.  $S_a^{MC} = S_b^{MC}$ . From Figure 2 it can be clearly seen that correlated part of the inputs plays an important role in contributing to the output variance, sometimes even more significant than the uncorrelated part. With the proposed method, more information has been explored, which can be referred to by analysts to improve the model performance.



**Figure 2 Sensitivity indices of the correlated inputs for the I-beam**

## 5. Conclusions

In this work, application of the SGI technique to the global sensitivity analysis is discussed. Sensitivity analysis under the case of independent variables is much different from that of correlated variables, as in the latter case both the computation and interpretation are more complicated. When the variables are independent, both the main effect index and total effect index can be estimated by the two-stage use of the SGI technique. The whole procedure is considerably clear and simple. When it comes to the correlated variables, necessary steps need to be taken before the sensitivity results can be obtained. The whole procedure can be generalized as follows: (1) estimate the output variance with the SGI technique, (2) orthogonalize the correlated variables to independent variables, (3) estimate the statistical information of the independent variables by SGI, as well as the mapping between these new variables and the output, and (4) perform the sensitivity analysis with the SGI technique for the new performance function and independent variables. The sensitivity results are then used to interpret the contributions of the original correlated

variables to the output variance. The proposed method inherits the merits of the SGI technique, and can estimate the integrals involved in the sensitivity analysis with acceptable accuracy, while keeping the computational burden under control. Applications to the examples have shown that the proposed method can be seen as a viable choice for sensitivity analysis of engineering models.

### Acknowledgement

This work is supported by the development fund for important program of NWPU (20136102120032).

### References

- Barthelmann, V., Novak, E. and Ritter, K. (2000) High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics* **12**: 273–288.
- Borgonovo, E. (2007) A new uncertainty importance measure. *Reliability Engineering and System Safety* **92**: 771–784.
- Borgonovo, E. and Tarantola, S. (2008) Moment independent and variance based sensitivity analysis with correlations: An application to the stability of a chemical reactor. *International Journal of Chemical Kinetics* **40**: 687–398.
- Gerstner, T. and Griebel, M. (2003) Dimension-adaptive tensor-product quadrature. *Computing* **71**: 65–78.
- Gerstner, T. and Griebel, M. (1998) Numerical integration using sparse grids. *Number Algorithms* **18**: 209–232.
- Kucherenko, S., Tarantola, S. and Annoni, P. (2012) Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications* **183**: 937–946.
- Lu, H., Zhang, Y.M., Zhao, C.L. and Zhu, L.S. (2012) Reliability sensitivity estimation of mechanical components with multiple failure modes. *Journal of Mechanical Engineering* **48**(2): 63–67 (in Chinese).
- Mara, T.A. (2009) Extension of the RBD-fast method to the computation of global sensitivity indices. *Reliability Engineering and System Safety* **94**: 1274–1281.
- Mara, T.A. and Tarantola, S. (2012) Variance-Based Sensitivity Indices for Models with Dependent Inputs. *Reliability Engineering and System Safety* **107**: 115–121.
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010) Variance Based Sensitivity Analysis of Model Output. Design and estimator for the total sensitivity index. *Computer Physics Communications* **181**: 259–270.
- Shi, Y., Xu, W., Qin, C. and Xu Y. (2009) *Mathematical statistics*, Beijing: Science Press.
- Smolyak, S.A. (1963) Quadrature and interpolation formulas for tensor products of certain classes of functions. *Soviet Mathematics Doklady* **4**: 240–243.
- Sobol, I.M. (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* **55**: 271–280.
- Xu, C. and Gertner, G.Z. (2008) Uncertainty and sensitivity analysis for models with correlated parameters, *Reliability Engineering and System Safety* **93**: 1563–1573.

## **Preliminary work on the potential of extending structural health monitoring concepts for healing assessment**

†\*W.K. Chiu<sup>1</sup>, W.H. Ong<sup>1</sup>, M. Russ<sup>2</sup> and Z. Chiu<sup>3</sup>

<sup>1</sup>Department of Mechanical & Aerospace Engineering, Monash University, Clayton, Victoria, Australia

<sup>2</sup>The Alfred Hospital, Commercial Road, Prahran, Victoria, Australia

<sup>3</sup>Faculty of Medicine, Nursing and Health Sciences, Monash University, Clayton, Victoria, Australia

\*Presenting author: wing.kong.chiu@monash.edu

†Corresponding author: wing.kong.chiu@monash.edu

### **Abstract**

External fixations are used to treat skeletal fractures. Patients are required to be immobilized for up to 12 weeks to allow the fracture to heal. This paper presents some preliminary finite element findings to highlight the potential of attaching sensing elements on the fixation to monitor the progression of healing of the fracture. This work paves the way for the extension of structural health monitoring concepts for orthopaedic devices.

**Keywords:** Structural health monitoring, bone healing assessment, bone fracture

### **Introduction**

Musculoskeletal injuries are often associated with traumas. One of the associated effects of these injuries is bone fracture which is a complete or incomplete breakage of a bone, as a result of excessive force or trauma to the site. Following a fracture, there are a variety of treatment choices, most commonly internal and external fixation, should operative methods be chosen. In general, intermedullary nailing, a type of internal fixation, is chosen, as it has been associated with decreased complications such as pin tract sepsis and joint stiffness (Wu, 2006). However, as noted by Ingari and Powell (2007), temporary external fixation has been indicated in situations where temporary fixation is necessary and risks of infection high, such as in combat. In these circumstances after a short period of time it is possible to convert to internal fixation, but in these cases risks of infection are high, and often the temporary fixation is modified to be permanent. The objectives of orthopaedic injury management are to prevent infection, promote fracture healing and restore function.

The healing process of fractures is a complicated procedure, on both a macroscopic and microscopic level. Due to the precise balance required between anabolic and catabolic phases, delayed unions, mal-unions and non-unions are common, occurring in 5-10% of all long bone fractures (Griffin et al, 2011). Hernigou et al (2005) define a delayed union as a fracture site continuing to sustain clinical and radiological signs of fracture outside of the expected healing time, or the absence of signs of progressive repair between the 3rd and 6th month of repair following a fracture. Mal-unions are defined as a pathological union of a fracture, usually involving shortening and rotational or angular deformity (Wu, 2006), while non-unions are defined as a lack of union within the expected healing time (Griffin et al, 2011). Although there is a lack of a standardised definition of these pathologies, they are clinically significant and have an impact on quality of life.

In the above studies, it is evident that an ability of assess the degree of union of the fixated fracture is of fundamental importance. It is needed to (1) determine if a re-operation is required, and (2) to assess the effectiveness of the external fixation as a definitive treatment. The most common method of determining state of union, following clinical data collected, is a plain radiograph or a CT scan. Unfortunately, both have significant drawbacks to the patient, primarily exposure to radiation, and both are fundamentally inconclusive in determining state of union. Progresses in Engineering research has led to a significant advancement in Structural Health Monitoring (SHM) which have been shown to offer the prospect of a quantum gain in performance and efficiency for the structural integrity management of expensive assets such as aircraft and infrastructure. The key enabling technologies for this revolution include primarily the rapid and continuing advances that have been

made in the past three decades in the development of miniaturised sensors, actuators, and of various multifunctional materials and structural concepts Srinivasan, A.V. and D.M. McFarland (2001).

For the purpose of this paper, a fixated femur will be used to outline this potential. A saw-bone femur fixated with a Hoffmann II external fixator was used. This paper will present a set of results that will establish the fundamentals required to underpin the ability to integrate sensing into an external fixator for union and healing assessment. It will be shown that the modal response of a fixated femur is sensitive to the state of union of the fractured region.

**Finite element analysis**

The aims of this part of the work are to:

1. Determine the potential of locating sensing devices on the external fixation to determine the state of healing of the fractured femur.
2. Qualitatively identify suitable sensor placement locations on the fixation and the appropriate frequency bandwidth that will facilitate the monitoring of the healing of the fracture femur.

This will be conducted with a series of finite element analyses of the fixated femur. The geometry of the fixated femur shown in Figure 1 was digitally scanned using structured light 3D scanning technology. A finite element model was created from the scanned geometry as shown in Figure . Due to its complex geometry, it was meshed using tetrahedral elements set to a spacing of 2mm. This resulted in 345,966 elements and 79,793 nodes as shown in Figure 2. The fracture was simulated with a 3mm thick slice in the middle of the femur (shown in Figure 3) and was assigned a Young’s modulus 1% of the flawless bone. The properties used for the fixation and the saw-bone are shown in Table 1. The unlabeled fixation components in Figure were assigned titanium material properties. A constant 1% viscous damping was used in the simulation. In this respect, the finite element analyses reported in this section is use for qualitative purposes.

Table 1 Material properties used in simulations (Rudman et al, 2006).

Material	Density (kg/m <sup>3</sup> )	Young’s Modulus (GPa)	Poisson’s Ratio
Cortical bone	1500	17	0.3
Stainless steel pin	7817	198	0.272
Composite rod	1500	134	0.3
Titanium	4430	121	0.34



Figure 1: Saw-bone femur fixated with Hoffmann II external fixation.

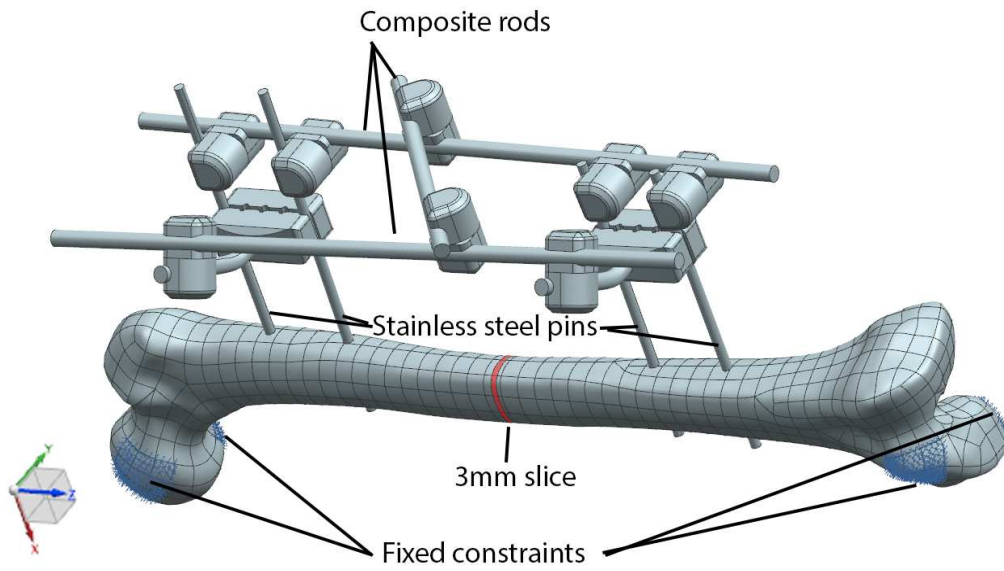


Figure 2: 3D model of fixated saw-bone specimen.

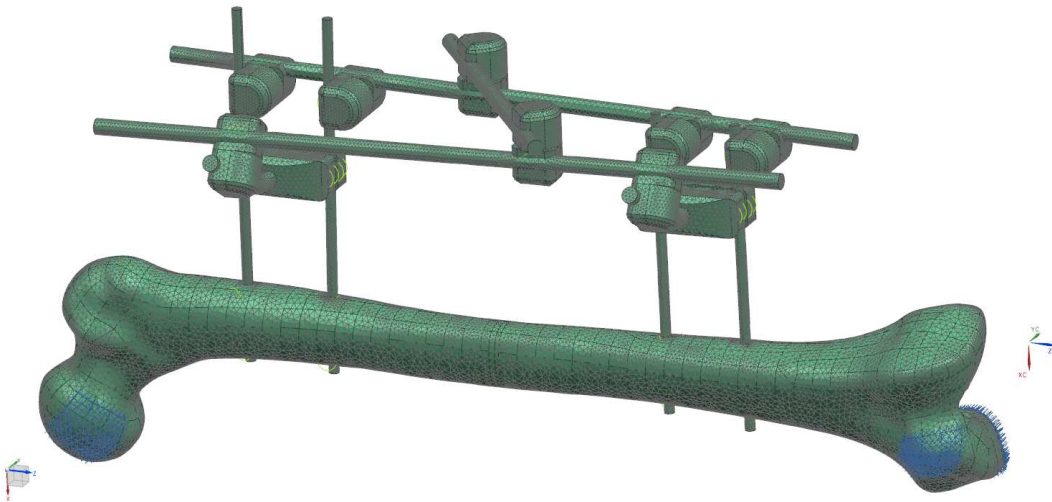


Figure 3: 3D model illustrating mesh density.

The fixed constraints applied to the fixated femur are shown in Figure 3. The dynamic response of the model was solved using NX Nastran. The mode shapes and natural frequencies of the fixated femur were first analysed. To facilitate the assessment of the healing and union of the fracture femur, it is important that the desired mode must include significant deformation of the femur.

**Identification of sensing location**

Figures 4(a) and (b) show two typical mode shapes that were calculated. They were selected to highlight the significance of using the appropriate modes for monitoring of the union and healing of the fracture femur. It was found that the lower modes are likely to be insensitive to the presence of the fracture because it is dominated by the deformation of the fixation (see Figure 4a). However, the higher mode (e.g. the 6<sup>th</sup> mode) shown in Figure 4b shows the deformation of the femur. Therefore, it is expected that the higher order resonant behaviour of the femur is likely to be affected by the changing material properties of the femur in the fractured region that is representative of the state of healing and/or union.

The 6<sup>th</sup> mode shape shown in Figure 4(b) also includes significant deformation of the legs of the fixation. It is inferred from these results that the sensors can suitably be located on the legs of the fixation to determine the state of healing. This is essential as they are easily accessible and are external to the body.

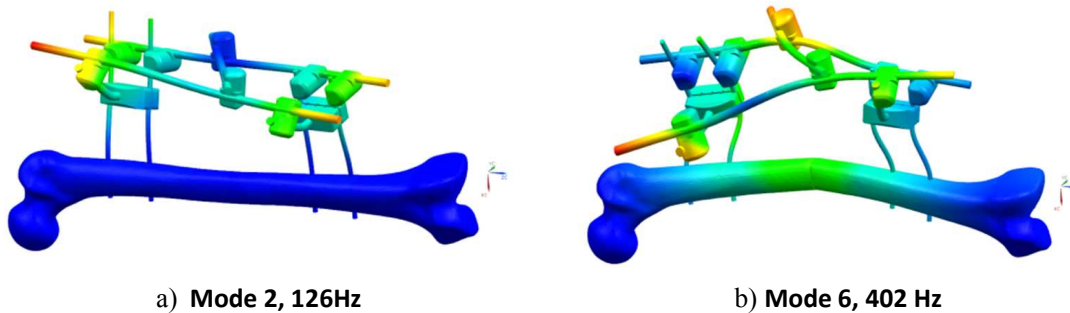


Figure 4: Mode shapes.

**Forced analyses**

To confirm this, a series of forced dynamic response of the fixated femur were conducted with a forced input is applied to the fixation as shown in Figure 5. A unit force is applied over a frequency bandwidth of 1000 Hz. The 4 sensing locations are selected and the transfer functions between the surface stress on these locations and the force input were calculated. Sensing locations 1 and 2 and 3 and 4 were located on the opposite sides of the fracture location. The calculations were conducted with the following femur configurations: (a) no fracture; (b) the fractured femur simulated with the properties at the fractured location reduced to 1% of its original.

The forced responses of the fixated femur with and without the fractured region are shown in Figures 6 and 7. These results shows that the effects of the fractured femur on the transfer function are only evident in the higher modes above 300 Hz. This is because significant deformation of the femur was present only in the higher modes. The results also showed that the sensing location shown in Figure 5 is potentially sensitive to the state of healing and union of the fractured femur.

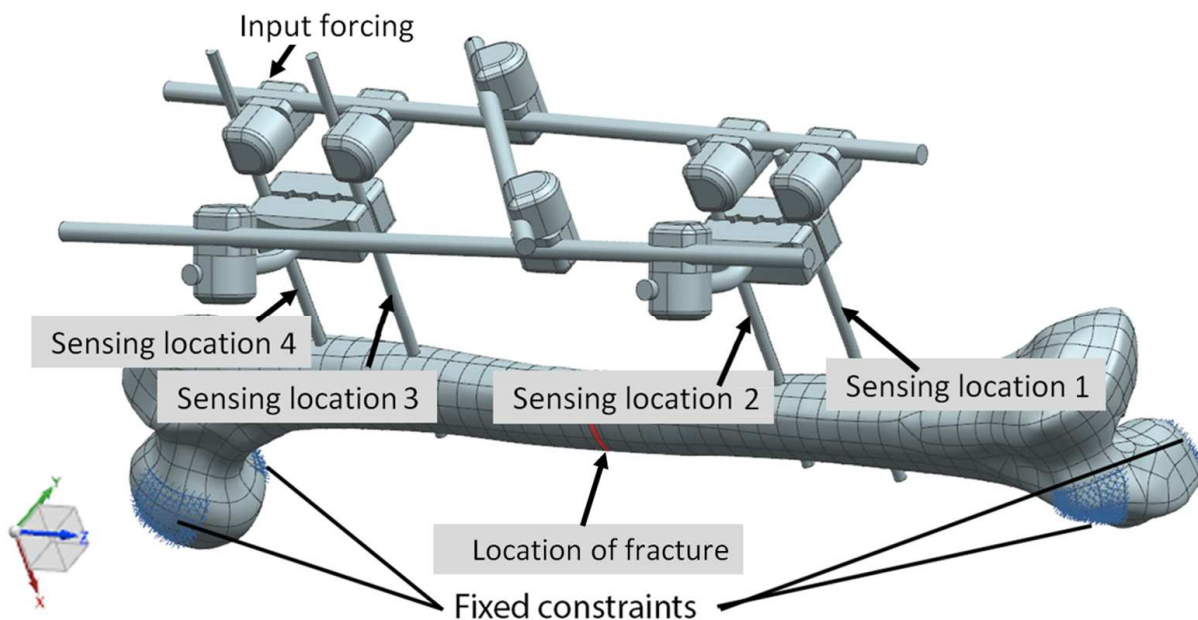


Figure 5: Input and output locations for transfer function calculation.



The transfer function obtained for sensing locations 1 and 2 are shown in Figure 6(a) and 6(b). The results obtained from these two sensing locations are similar both in magnitude and in phase. The natural frequencies for the modes above 250 Hz increases when the fracture is totally healed and united. Figure 7(a) and 7(b) shows the transfer function obtained for sensing locations 3 and 4. The healing of the femur can be indicated by the appearance of natural modes at about 300 Hz and 480 Hz.

This set of results clearly demonstrates the significance of locating the sensing element in the proximity of the drive point. It also highlights the potential of integrating structural health monitoring concepts into the orthopaedic devices for the monitoring of healing and union of a fractured femur. An important feature is the possibility of locating sensors external to the human body for healing assessment.

### Conclusions

The work presented in this paper described the potential of integrating structural health monitoring concepts into fixators to determine the state of union of a fractured long bone. The work presented showed how the inclusion of actuation and sensing protocol can be established to assess the state of union of the externally fixated saw-bone femur. This preliminary study highlights the potential of locating sensors external to the human body in assessing the state of healing of a fixated fractured femur.

### Acknowledgements

This project is funded by the US Navy Office of Naval Research – Global (N62909-14-1-N248). The financial support provided is gratefully acknowledged.

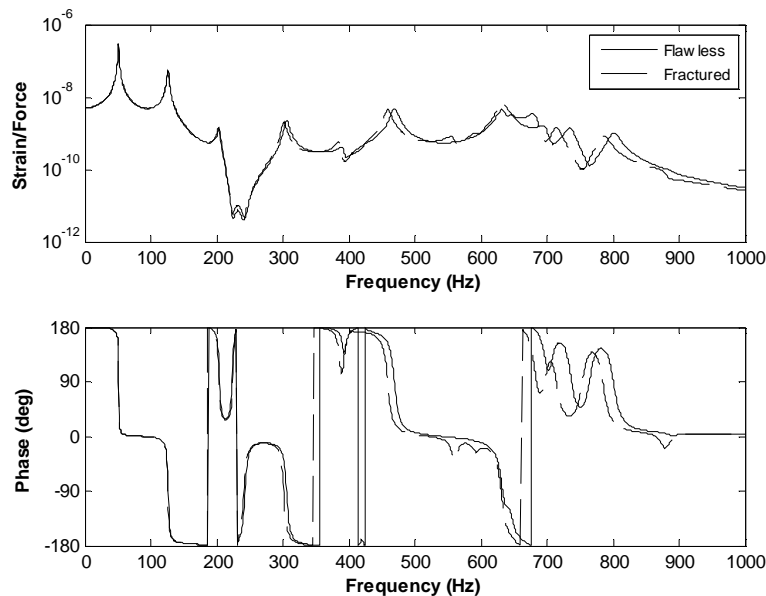


Figure 6(a): Transfer function for a fractured and flawless bone on pin 1.

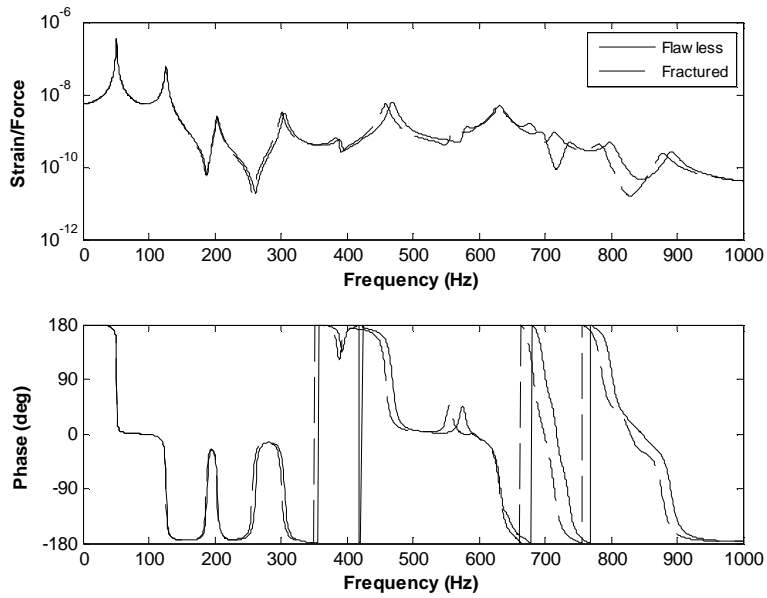


Figure 6(b): Transfer function for a fractured and flawless bone on pin 2.

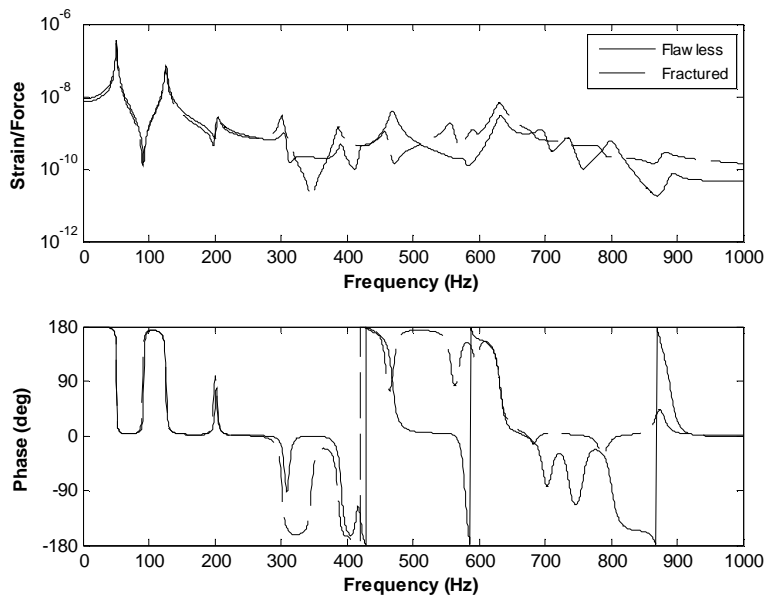


Figure 7(a): Transfer function for a fractured and flawless bone on pin 3.

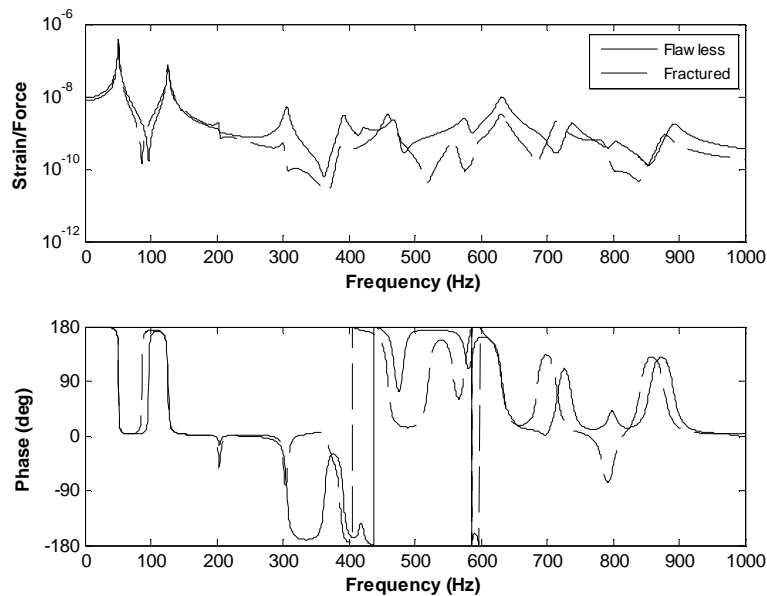


Figure 7(b): Transfer function for a fractured and flawless bone on pin 4.

**References:**

- Hernigou P, Poignard A, Beaujean. Rouard H. (2005) Percutaneous autologous bone-marrow grafting for non-union. *J Bone Joint Surg (Am)* Vol 87, pp 1430-1437.
- Ingari, J.V., Powell, E. (2007) Civilian and Detainee Orthopaedic Surgical Care at an Air Force Theatre Hospital. *Technique in Hand and Upper Extremity Surgery* Vol 11 Iss 2, pp 130-134.
- Rudman KE, Aspden RM, Meakin JR (2006). Compression or tension? The stress distribution in the proximal femur. *BioMedical Engineering OnLine*, 5:12 doi:10.1186/1475-925X-5- 12.
- Srinivasan, A.V. and D.M. McFarland, (2001) *Smart Structures - Analysis and Design*. Cambridge University Press.
- Wu, C.C., (2006) Treatment of Long-Bone Fractures, Malunions and Nonunions: Experience at Chang Gung Memorial Hospital, Taoyuan, Taiwan. *Chang Gung Medical Journal* Vol 29 pp 347-357

## Finite element analysis of different chipbreaker types for turning tool process

Dyi-Cheng Chen<sup>1,\*</sup>, Ci-Syong You<sup>1</sup>, Chia-An Tu<sup>1</sup>

<sup>1</sup>Department of Industrial Education and Technology,  
National Changhua University of Education, Changhua 500, Taiwan

\*Corresponding author: dcchen@cc.ncue.edu.tw

### Abstract

This paper used rigid-plastic finite element DEFORM<sup>TM</sup> 3D software, studying on 6061 aluminum alloy cutting plastic deformation behavior. The main design including a variety of different types of turning tools chipbreakers that under the same conditions, chipbreakers including without chipbreaker design, square chipbreaker design and an elliptical chipbreaker design, in order to analyze the effective stress after cutting, the effective strain after cutting, tool temperature, tool wear analysis in different chipbreaker types. The results of simulation analysis hoped to confirm the applicability of finite element method for cutting 6061 aluminum alloy.

**Keywords:** Turning tool, Chipbreakers design, 6061 aluminum alloy.

### Introduction

Cutting is an important method for forming a metal material, in order to improve processing efficiency and product surface quality, together with co industry stringent requirements. Must examine cutting inherent principle, since the metal cutting technology is quite complex and involves many factors. It used analytical methods are difficult to accurately describe the chip formation process, and funding for purely experimental study required a lot of investment and time. The finite element analysis techniques aided design as an effective tool to the workpiece material properties such as temperature, strain and strain rate functions for the interaction between the various machining parameters.

Includes regarding the mechanical processing research: [Toru et al. (2014)] research the cutting tool hot-working, transfers heat from heated up cutting tool, the work piece carries on the simulation analysis, the cutting experiment showed that the cutting tool heated up the pattern to improve the cutting process capability. [Stepan et al. (2014)] confirmed for calculating the stability of fixed milling, surveys in the working conditions carries on the confirmation the general numerical algorithm, its can forecast that the cylindrical milling cutter geometric form carries on the milling process the stability condition. [Gandjar et al. (2014)] described five axis milling, during using an analysis method to define the CWE half fine mill ring-like and plane face cutter in the carving part. [Totis et al. (2014)] research cutting tool shape optimization. It has carried on the optimization to the design of module through the use finite element analysis method, the dynamic mathematical model of common sensor having. Through conducting modal analysis and cutting test. The experimental result has proven the new equipment outstanding characteristics. [Jaroslava and Zdenk et al. (2014)] studied the cutting blade and cutter life of cutting blade radius, the process steam turbine outer covering division.

### Finite element analysis

In metal shaping process, because comes under influence of the plastic deformation, creates the production of flaw mostly is caused for the material ductile fracture by processing. The brittle failure of metallic material among crystallizing cleavage surfaces, the atom unifies to destroy for the focus micro phenomenon mutually, namely breaks out the destruction, and before the material destroys. The strain capacity does not have minimum even the plastic deformation, and formation

destruction of non-early warning. The DEFORM™ software according to the design system and processing process, because of the analysis pattern that the two-dimensional or three-dimensional flow distorts, can simulate the metallic material in the die forming, after ductile fracture value and distortion, temperature and plastic flow speed, stress and strain of distributed situation.

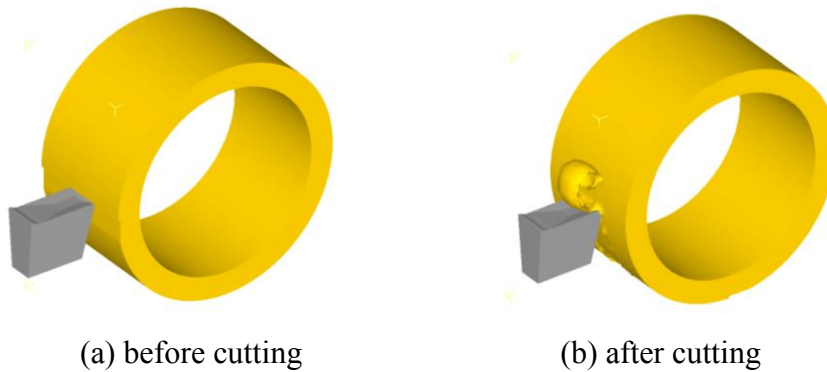
**Simulation and parameter setting**

Table 1 illustrated the different tools chipbreakers cutting simulation parameters of 6061 aluminum alloy, including fixed parameters include: speed 200 rpm, tool temperature 200 °C, friction factor 0.7.

**Table 1. Different tools chipbreakers cutting simulation parameters of 6061aluminum alloy**

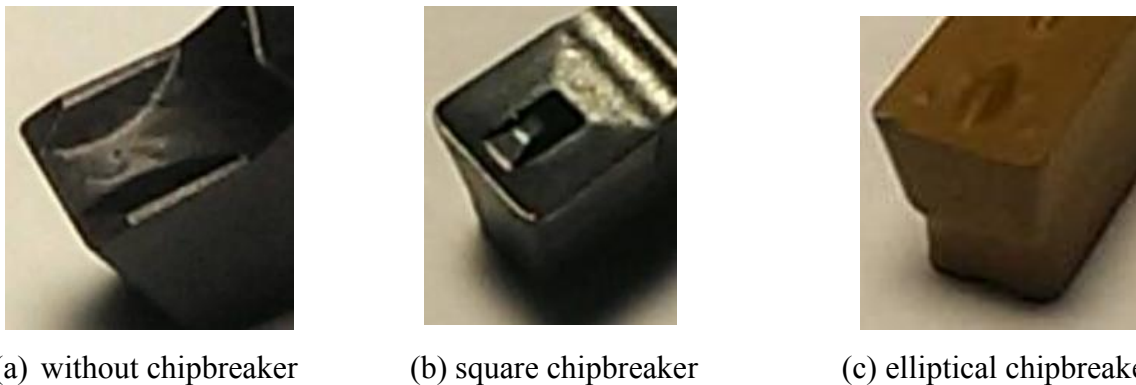
turning tools chipbreakers types	Rotational speed (rpm)	Tool Temperature (°C)	friction factor
without chipbreaker			
square chipbreaker	200	20	0.7
elliptical chipbreaker			

Figure 1 (a) and (b) show simulation diagram for an 6061 aluminum alloy chipbreaker turning cutting before forming and after forming .



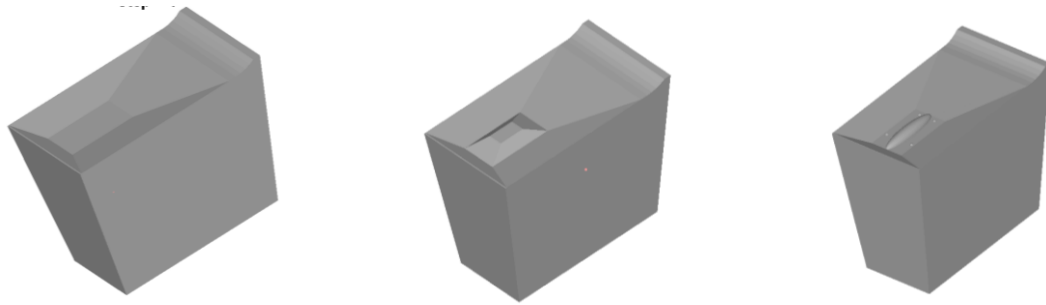
**Figure 1. 6061 aluminum alloy chipbreaker turning cutting**

Figure 2 shows 6061 aluminum alloy turning chipbreaker entity diagram, which (a) without chipbreaker, (b) square chipbreaker and (c) elliptical chipbreaker.



**Figure 2. Entity diagram of 6061 aluminum alloy chipbreaker turning cutting**

Figure 3 shows simulation diagram of 6061 aluminum alloy turning chipbreaker, which (a) without chipbreaker, (b) square chipbreaker and (c) elliptical chipbreaker.

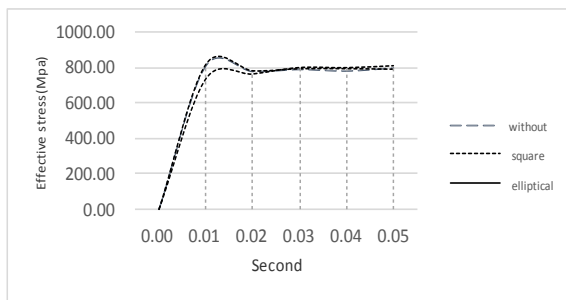


(a) without chipbreaker      (b) square chipbreaker      (c) elliptical chipbreaker

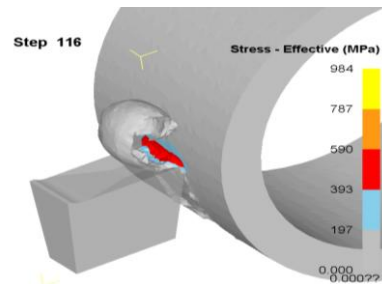
**Figure 3. Simulation diagram of 6061 aluminum alloy 6061 chipbreaker turning cutting**

**Results and discussion**

Figure 4 shows effective stress diagram for different chipbreaker, the maximum value at 0.1 second for an elliptical chipbreaker. Moreover stress producing is 841MPa, square chipbreakers stress arising is smallest value 790MPa. Stress arising from 0.02-0.05 seconds is no difference. Figure 5 shows effective stress diagram of without chipbreaker tool. It can be seen that the maximum effective stress generated in the tool and billet contact.

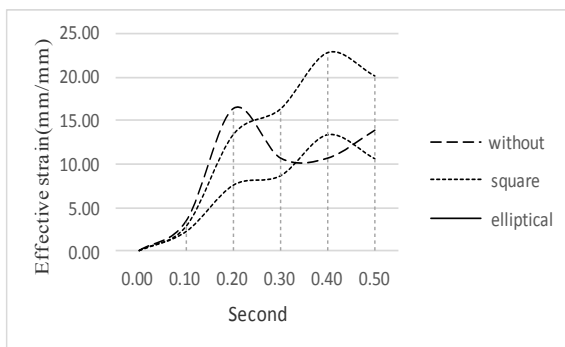


**Figure 4. Different chipbreaker effective stress diagram**

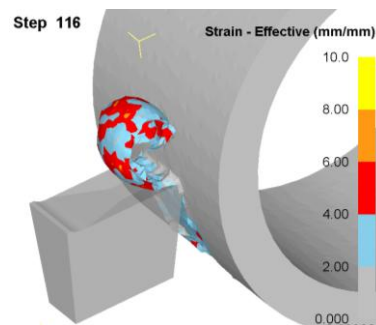


**Figure 5. Effective stress of without chipbreaker**

Figure 6 shows effective strain diagram for different chipbreaker. The without chipbreaker produce the greatest strain 16(mm / mm) at 0.2 seconds. The square chipbreakers is arising smallest effective strain 7.5(mm/mm). Strain The elliptical chipbreakers is arising largest effective strain at 0.4 seconds. Figure 7 shows strain- effective diagram for without chipbreaker. It can be seen that the maximum effective strain generated in the chip curler.



**Figure 6. Different chipbreaker effective strain diagram**



**Figure 7. Effective strain of without chipbreaker**

Figure 8 shows different chipbreaker temperature diagram, the square chipbreaker produced maximum value at 0-0.5 seconds, the temperature is 141°C. Figure 9 shows without chipbreaker

temperature diagram. It can be seen that the maximum temperature generated in the tool and billet contact.

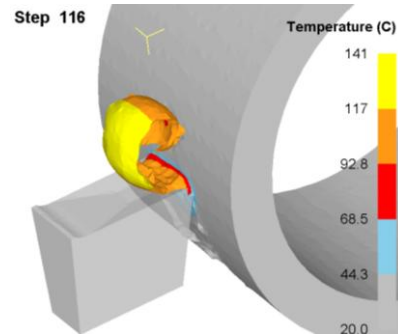
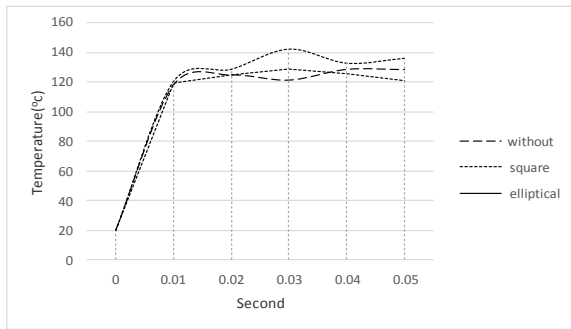
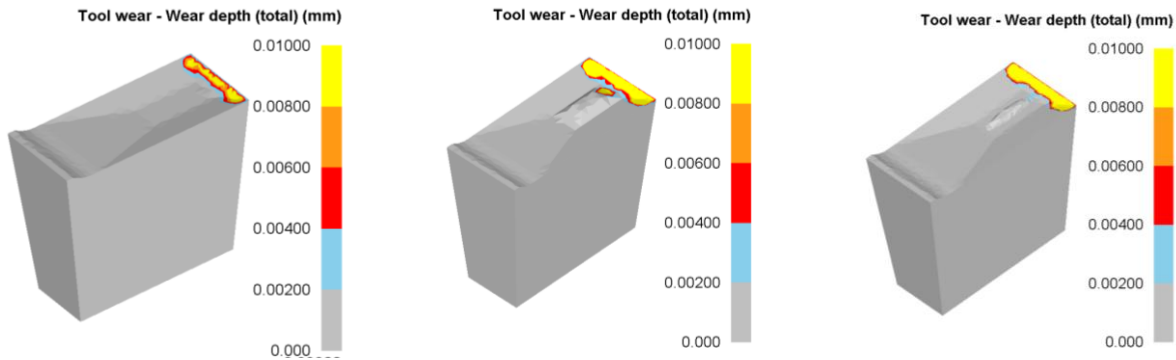


Figure 8. The temperature of different chipbreaker

Figure 9. The temperature of without chipbreaker

Figure 10 shows wear simulation diagram for different chipbreakers, (a) the wear is homogeneous at the tip for without chipbreakers, (b) square type chipbreaker generate maximum wear in the chipbreaker area and (c) elliptic chipbreaker induced homogeneous wear at the tip area.



(a) without chipbreaker

(b) square chipbreaker

(c) elliptical chipbreaker

Figure 10. Wear simulation diagram of different chipbreakers

## Conclusions

This article in view of the aluminum alloy 6061 materials, has built the finite element model of description cutting process for without chipbreaker design, square chipbreaker design and an elliptical chipbreaker design. The results have shown that: (1) square chipbreakers stress arising is smallest value 790MPa; (2) the maximum effective strain generated in the chip curler; and (3) square type chipbreaker generate maximum wear in the chipbreaker area.

## References

- Gandjar, K. and Hendriko, H., Emmanuel, D. (2014) An analytical method for obtaining cutter workpiece engagement during a semi-finish in five-axis milling, *Computer-Aided Design* **55**,81-93.
- Jaroslava, F. and Zdenk, J. (2014) Influence of the cutting edge radius and the cutting edge preparation on tool life and cutting forces at inserts with wiper geometry, *Procedia Engineering* **69**,565-573.
- Stepan, G. and Munoa, J., Insuperger, T., Surico, M. (2014) Cylindrical milling tools: Comparative real case study for process stability, *CIRP Annals - Manufacturing Technology* **63**,385-388.
- Totis, G. Adams, O., Sortino, M., Veselovac, D., Klocke, F. (2014) Development of an innovative plate dynamometer for advanced milling and drilling applications, *Measurement* **49**,164-181.
- Toru, K. and Kanako, H., Mamoru, M. (2014) Efficient and precise cutting of zirconia ceramics using heated cutting tool, *CIRP Annals - Manufacturing Technology* **63**,105-108.

## Study of the characteristic of droplet transfer in laser-MIG hybrid welding based on the phase matching control of laser pulse and arc waveform

\*G. Song <sup>‡</sup>, J.Wang <sup>‡</sup>, and †L.M.Liu <sup>‡</sup>

<sup>‡</sup>School of Materials Science and Engineering, Dalian University of Technology, China.

\*Presenting author: songgang@dlut.edu.cn

†Corresponding author: liulm@dlut.edu.cn

### Abstract

This paper puts forward an experimental program on accurate phase matching control between laser pulse and MIG arc waveform, and achieves continuous trigger control of the laser pulse. With the relationship between laser pulse, MIG arc and droplet, the effect of laser pulse on droplet transfer is studied by controlling laser pulse to act on the peak value and basic value of the MIG arc waveform respectively during laser-MIG hybrid welding process on aluminum alloy 6061. High-speed camera is used to acquire the droplet transfer. Experimental results show that the form of droplet transfer hasn't changed with the laser pulse acting. However, laser pulse can change the necking form of droplet and the speed for droplet transfer by changing the electromagnetic force. The necking form of droplet changes from symmetrical necking to asymmetrical necking when the laser pulse acts on the peak value, however there is no significant change for the necking form when the laser pulse acts on the basic value. The speed of droplet transfer is faster than that without laser, and it improves with the increase of the laser power. The speed of droplet transfer with laser acting on the basic value is faster than that with laser acting on the peak value. Welding appearance with laser pulse acting on the basic value is better than that with laser pulse acting on the peak value. It is thought that this study can technically support the aluminum alloy welding with high speed and low thermal damage at small welding currents.

**Keywords:** Phase matching, Hybrid welding, Drop transfer, Welding appearance

### Introduction

As an advanced method, laser-MIG hybrid welding has many advantages in increasing welding speed and improving drop transfer and so on. Liu S et al studied the droplet transfer mode and forming process in 5kW CO<sub>2</sub> laser-MAG hybrid welding, it is found that arc characteristic, droplet transfer mode and final weld bead geometry are strongly affected by the distance between laser and arc, the droplet transfer mode is changed from globular transfer to projected transfer with the increasing DLA[Liu et al. (2012)]. Lei Z et al studied characteristics of droplet transfer in 3kW CO<sub>2</sub> laser-MIG hybrid welding, it was found that the droplet transfer mode is changed from short circuiting transfer to projected transfer due to interaction between CO<sub>2</sub> laser and MIG arc in CO<sub>2</sub> laser-MIG hybrid welding process, and the frequency of droplet can be improved by adjusting the parameters[Lei et al. (2004)].

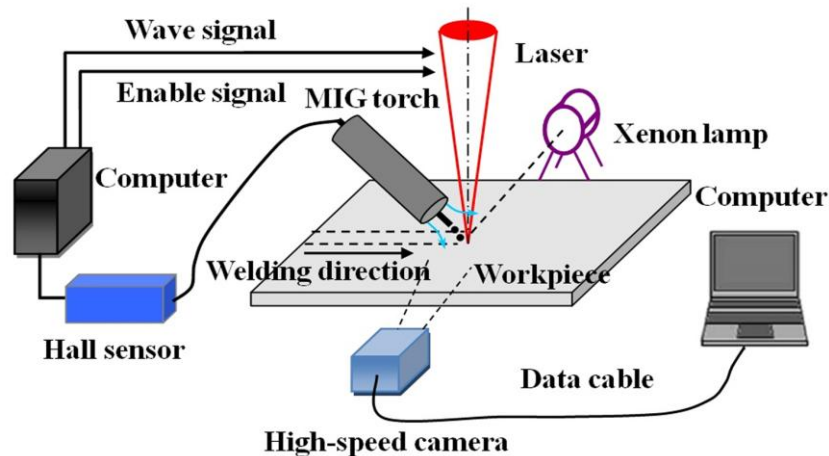
The above researches mainly focus on the influence of high-power laser on the arc and droplet transfer. Our team discovered the low power pulsed laser enhancing arc discharge phenomenon. It found that the arc plasma can be stabilized by the laser pulse, and the enhancing effect of laser on arc plasma leads to the characteristic variation of the arc plasma, including the arc composition, the arc behavior, the arc electron temperature and density, which results in a very stable arc discharge in high speed welding conditions [Liu and Chen (2013a;2013b;2011)]



Considering the low power pulsed laser enhancing arc discharge phenomenon, this study mainly focuses on the improvement of the laser pulse to the droplet transfer based on the matching control of laser pulse and arc waveform. Using high-speed camera, the characteristics of droplet transfer after employing this system is studied. By comparing the variation in the characteristics of droplet transfer with laser acting on the peak value and basic value of the MIG arc waveform respectively, the interactions between laser and droplet transfer are analyzed.

### Experimental setup

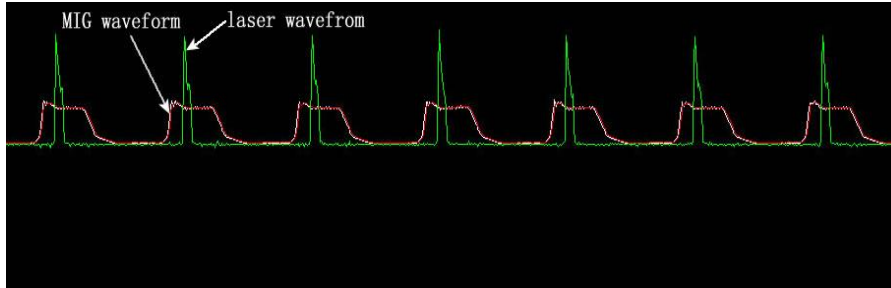
As is shown schematically in Fig. 1, the heat source is composed of a pulsed Nd: YAG laser (LWS-1000) hybrid and a direct current pulsed MIG arc. The laser, with a wave-length of  $1.064\mu\text{m}$ , is focused by a lens with a focal distance of 150 mm into a spot measuring about 0.5 mm on the surface of the workpiece. The MIG arc acts behind the laser along the welding direction with an adjustable DLA, and the angle of the metal electrode axis to workpiece is  $45^\circ$ . Argon with a purity of 99.99% is used as the shielding gas, and the flow rate through the MIG nozzle is  $15\text{L}\cdot\text{min}^{-1}$ .



**Figure 1. Schematic diagram of experimental devices**

A xenon lamp is positioned towards the laser acting point and perpendicularly to the welding direction to avoid the influence of arc to droplet transfer, and on the opposite side a high-speed camera is placed to acquire the droplet transfer. A suitable optical filter should be attached to the camera lens to ensure suitable light intensity. The acquiring frequency of the camera is set to be 1000 frame/s and the interval time between every two successive photos is 1ms. By a cable, the data acquired is transferred to a computer that can translate the data into visible images through corresponding software.

In order to acquire the accurate matching control of laser pulse and arc waveform, the matching control system is developed. Hall sensor is used to detect the arc wave in welding process. As shown in Fig. 1, the data signal of MIG arc is transmitted to computer and acquired by NI PCI-6221, the DAQ system process the signal and capture the feature points of arc by the computer, thus accurately analyze the waveform field of MIG arc, and realize a continuous trigger control of laser pulse on peak value and basic value of MIG arc respectively by outputting the wave signal and enable signal to YAG laser directly. The function of this system is as shown in Fig 2. The laser pulse distribute in the pulsed MIG arc waveform uniformly.

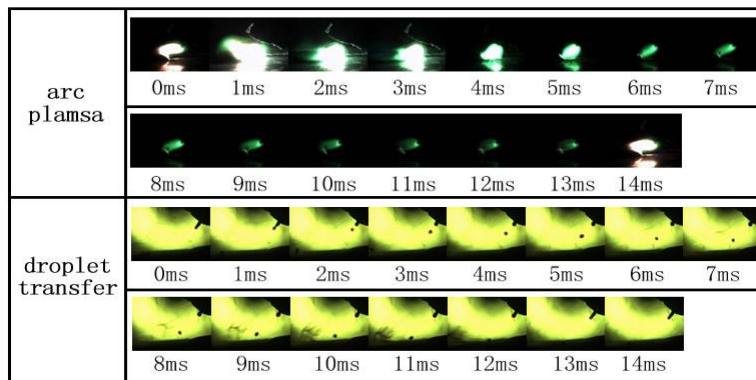


**Figure 2. Schematic diagram of function of the matching control system**

Experiments are performed at atmospheric pressure and room temperature. A 6061 aluminum alloy plate with a thickness of 2 mm is applied as the workpiece. Oxide film and grease on the plates are cleaned before welding. In this welding system, workpiece travels in a direction, while the MIG torch, the high-speed camera and the laser beam maintain fixed. Therefore, the information of droplet transfer can be acquired continuously and steadily.

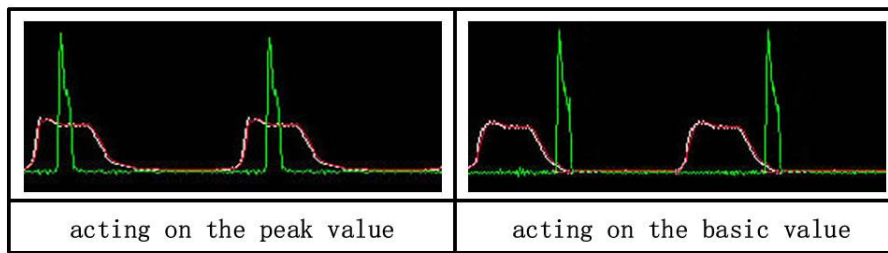
**Experiment results**

In the single DC pulsed MIG welding, there are different modes of droplet transfer (such as one droplet per pulse and one droplet every two pulses). It's the mode of drop transfer that determines the pattern of pulse scatter distribution, so we should confirm the mode of drop transfer in this experimental condition. In MIG welding process, the metal electrode is the anode during the DC MIG arc, in which the electrons are accelerated to impact and heat the welding wire. Thereinto, the peak value is used to melt the welding wire, and the basic value is used to maintain the burning of the arc. As is shown in fig.3, the cycle of arc plasma is about 14ms (the peak value is about 3ms and the basic value is about 1ms), and the cycle of drop transfer is the same, so the mode of drop transfer is one drop per pulse in this experimental condition.



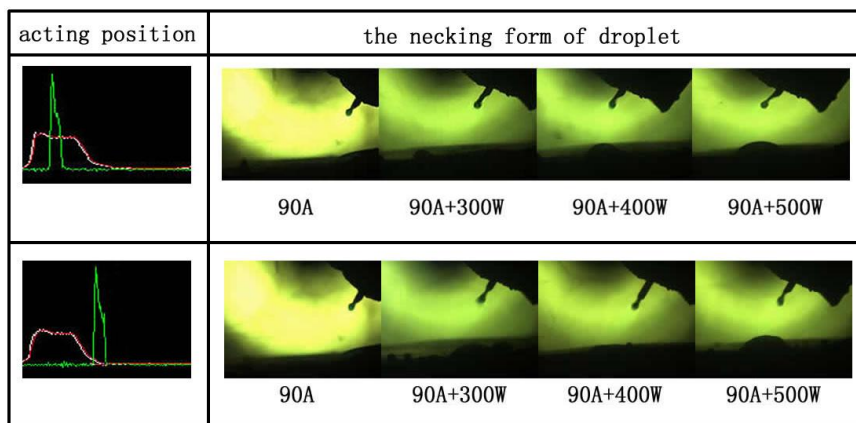
**Figure 3. The relationship of the MIG arc and drop transfer**

Therefore, the pattern of laser pulse distribution is determined to be the format as shown in the fig.4. The laser pulse acts on the peak value and basic value respectively. That is to say, the pulse scatter act on the period when the droplet forming and entering into molten pool respectively.



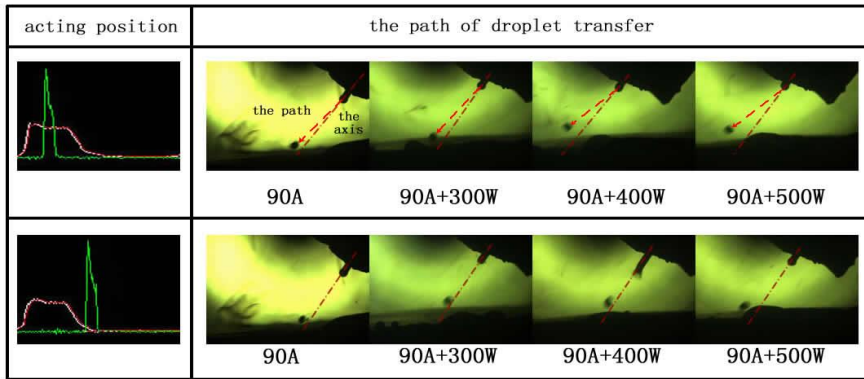
**Figure 4. Schematic diagram of the MIG arc waveform position of laser pulse acts on**

Fig.5 shows the necking mode with different laser pulse distribution modes between pulsed laser and DC pulse MIG arc. It is found that the necking mode under different laser pulse distribution modes has different variation rules: when the laser pulse act on the peak value, the necking of droplet deviates from the axis more easily with the increase of laser power, that is to say the necking form is changed from symmetrical necking to asymmetrical necking; to the contrary, there was no significant deviation in the necking of droplet during the laser pulse act on the basic value.



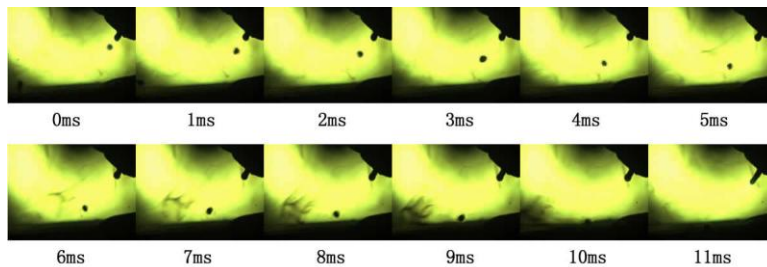
**Figure 5. The diagram of the necking form of droplet**

Due to the difference of laser pulse acting on the peak value and basic value, the path of droplet shares certain regularity during the droplet entering into molten pool: There is no deviation between the path of droplet transfer and the metal electrode axis without laser pulse, and the angle between the path and axis increases with the increasing of the laser power when the laser pulse acting on the peak value. During the laser pulse acting on the basic value, there is no significant deviation between the path of droplet and the metal electrode axis, just as is shown in Fig 6. It indicates that there are some differences of the interactions between laser pulse and arc plasma under different laser pulse distribution modes.



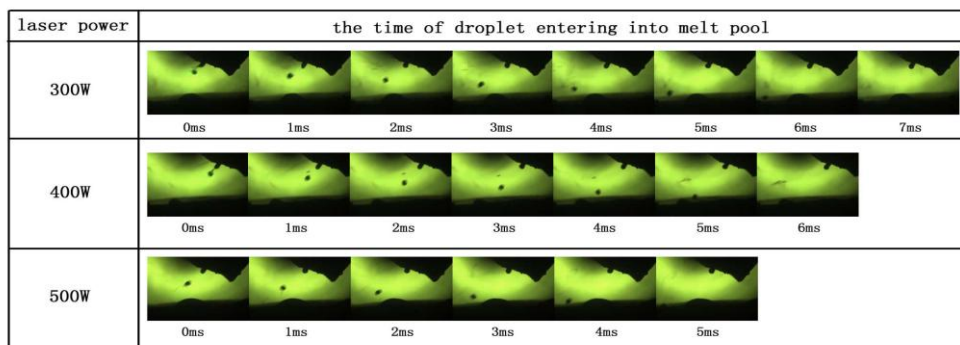
**Figure 6. The diagram of the path of droplet transfer**

In the droplet transfer cycle shown in Fig.7, the time for the droplet entering into melt pool is about 11ms in the single MIG welding, and it almost occupies the whole time of the basic value (the basic value is about 11ms).



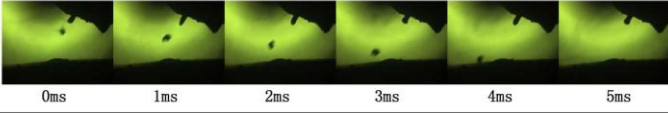
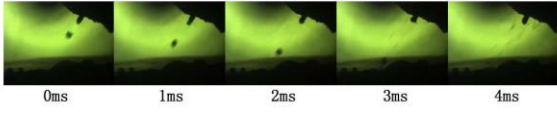
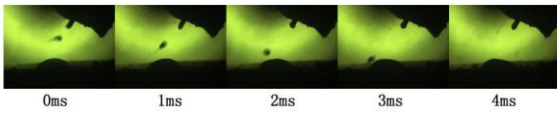
**Figure 7. The time of MIG droplet entering into melt pool**

Figure 8 is a typical metal transfer cycle from the comparative experiment between different pulse scatter distribution modes under different laser power. As can be seen, during the laser pulse acting on the peak value, the time for the droplet entering into melt pool is reduced with the increase of the laser power. For example, the time with 400W power is about 6ms, and it decreases about 4ms than that in the single MIG welding, and decreases about 1ms than that with 300W power.



**Figure 8. The time of droplet entering into melt pool during laser pulse acting on peak value**




As is shown in Fig 9, during the laser pulse acting on the basic value, there are the same phenomena compared with the laser acting on the peak value: the time of droplet entering into melt pool is reduced with the increase of the laser power. The only difference is that the time of droplet entering into melt pool is shorter when laser pulse acting on the basic value. For example, the time with 400W power is about 4ms, and it decreases about 1ms than that with 300W power.

laser power	the time of droplet entering into melt pool
300W	 0ms 1ms 2ms 3ms 4ms 5ms
400W	 0ms 1ms 2ms 3ms 4ms
500W	 0ms 1ms 2ms 3ms 4ms

**Figure 9. The time of droplet entering into melt pool during laser pulse acting on basic value**

Therefore, the time of droplet entering into melt pool with laser pulse is shorter than that without laser pulse, and the time with laser acting on the basic value is shorter than that with laser acting on the peak value.

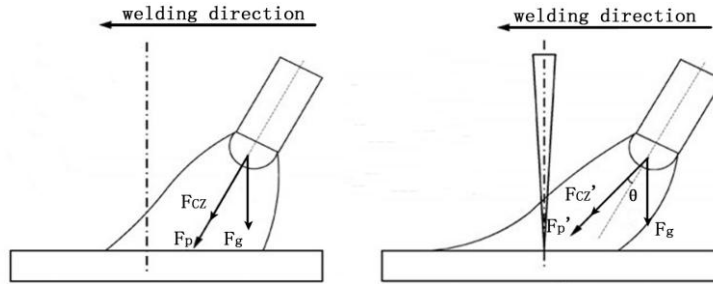
In order to ensure the impact on the formation of the laser pulse distribution pattern, the welding experiment is performed on aluminum plate with the laser of 500W and speed of 1500mm/min. Under the same welding conditions, the laser pulse is controlled to act on different positions of MIG arc wave form. Fig. 10 shows the hybrid weld bead appearances under the welding parameters with different laser pulse distribution modes. It can be seen that the weld appearance with laser pulse acting on the basic value is better than that with laser pulse acting on the peak value. When the laser pulse acts on the basic value, the laser-induced plasma can stabilize the small basic value easily, and the droplet is more advantageous to enter the melt pool because of the shorter time.

weld appearance	 without laser
	 laser acting on the basic value
	 laser acting on the peak value

**Figure 10. The appearances with different laser pulse distribution modes**

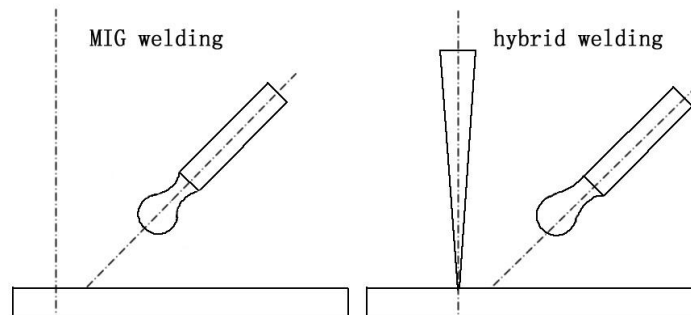
**Discussion**

From above results, it can be seen that, in the process of laser-MIG hybrid welding, the laser pulse cause the above-mentioned change, and the force on droplet is the primary cause of this change, so we should analyze the force on the droplet caused by the laser pulse.  $F_g$  is the gravity force,  $F_p$  is the plasma flow force,  $F_{CZ}$  is the electromagnetic pinch force. The plasma flow force is caused by the electromagnetic pressure difference, which is one part of the electromagnetic pinch force. Because the size of droplet has not changed, so compared with the single welding, the  $F_g$  also has no change in the laser-MIG hybrid welding.



**Figure 11. Schematic diagram of force of droplet of MIG welding and laser-MIG hybrid welding**

When the laser pulse acts on the MIG arc, there is the laser-induced plasma generating on the surface of the materials. The cation and free electron in the laser-induced plasma attract the arc to discharge in this area, and the boot of arc will be compressed [13, 15]. Therefore, the hybrid arc is changed because of the influence of laser pulse. The higher the laser power is, the stronger the electromagnetic attraction gets. That change also makes the direction of the electromagnetic pinch force and plasma flow force originally along with the axis of the wire offset at an angle  $\theta^\circ$ , just as is shown in Fig 11. Therefore, the extent and direction of the electromagnetic pinch force both change under the function of laser pulse. Therefore, the necking mode of droplet transfer changes regularly from symmetrical necking to asymmetrical necking by the influence of changed electromagnetic pinch force and plasma flow force when the laser pulse acts on the peak value of the MIG arc waveform, just as is shown in Fig 12. The droplet deviates from the wire axis, and with the laser power increasing, the larger the attraction it has, the bigger the angle of deviation is. When the laser pulse acts on the basic value of the MIG waveform, that is the droplet enter into the melt pool, the effect of laser pulse on the necking of droplet is not obvious for the reason that the droplet has already fallen off, then it also has little effect on the path of droplet entering into melt pool.



**Figure 12. Schematic diagram of necking form of droplet for MIG welding and laser-MIG hybrid welding**

When the laser pulse acts on the MIG arc, the root of MIG arc can be compressed at the place being acted by laser due to the effect of laser pulse<sup>[15,16]</sup>, which results in the increasing of the electromagnetic pressure difference between the electrode and the surface of workpiece. The electromagnetic pressure difference causes the change of the plasma flow force. Therefore, the extent of plasma flow force is also increased.

The time of droplet entering into melt pool is mainly influenced by the speed of the droplet. The speed of the droplet is mainly influenced by the plasma flow force after the droplet falls off. The plasma flow force increases due to the action of the laser pulse, therefore, it is the increasing of plasma flow force that gives rise to the shorter time of the droplet entering into the melt pool. The laser power is increasing, the faster speed it has, and the faster the speed of droplet has, then the shorter of the time that the droplet entering into the melt pool is. The time with laser acting on the basic value is shorter than that with laser acting on the peak value, this is because that the influence of the plasma flow force to the speed of droplet transfer with laser pulse acting on the basic value is stronger than that with laser pulse acting on the peak value.

### Conclusions

The influence of the laser pulse to the droplet in laser-MIG hybrid welding on 6061 aluminum alloys is studied in this paper, and the following conclusions can be drawn: Laser pulse can change the necking form of droplet and the speed for droplet transfer by changing the electromagnetic force. The necking form of droplet deviates from the axis more easily with the increase of laser power during the laser pulse acting on the peak value, and because of the influence of the necking of droplet, the angle between the path for the droplet entering into melt pool and wire axis increases with the increasing of the laser power, but there is no significant change for the path during the laser pulse acting on the basic value. The time of droplet entering into melt pool has a few characteristics under the action of laser pulse. The time of droplet entering into melt pool is shorter than that without laser. The time with laser power of 300W, 400W and 500W is 7ms, 6ms and 5ms during the laser pulse acting on the peak value. And the time with laser power of 300W, 400W and 500W is 5ms, 4ms and 4ms during the laser pulse acting on the basic value. The weld appearance with laser pulse acting on the basic value is better than that with laser pulse acting on the peak value.

### Acknowledgements

The authors gratefully acknowledge the sponsorship Supported by Program for Important National Science & Technology Specific Projects of China (No.2012ZX04008-031).

### References

- Liu, F., Zhang, W.Y. and Xue, C (2012) Influence of Area Energy for Welding Seam and Droplet Transfer on Hybrid Laser-arc Welding, *Journal of Mechanical Engineering* **48**, 84-90.
- Lei, Z.L., Chen, Y.B., Li, L.Q and Wu, L (2004) Characteristics of droplet transfer in CO<sub>2</sub> laser-MIG hybrid welding with projected mode, *APPLIED LASER* **24**, 361-364.
- Liu, L. and Chen, M (2013) Effect of laser pulse on recovery delay of arc plasma based on ion migration behavior in the pulsed laser-arc hybrid welding process, *Optics and Lasers in Engineering* **51**, 96-103.
- Liu, L., Chen, M. and Li, C (2013) Effect of electric arc on laser keyhole behavior based on direct observation during low power pulsed laser-arc hybrid welding process, *Optics and Lasers in Engineering* **51**, 1153-1106.
- Chen, M. and Liu, L.M. (2011) Study on attraction of laser to arc plasma in laser-TIG hybrid welding on magnesium alloy. *IEEE Transaction of Plasma Science* **39**, 1104-1109.

# Efficient Global Optimization Applied to Multi-Objective Design Optimization of Lift Creating Cylinder Using Plasma Actuators

\*†Masahiro Kanazaki<sup>1</sup>, Takashi Matsuno<sup>2</sup>, Kengo Maeda<sup>2</sup>, and Mitsuhiro Kawazoe<sup>2</sup>

<sup>1</sup>Graduate School of System Design, Tokyo Metropolitan University Japan

<sup>2</sup>Graduate School of Engineering, Tottori University, Japan

\*Presenting author: kana@tmu.ac.jp

†Corresponding author: <sup>1</sup>kana@tmu.ac.jp

## Abstract

A Kriging based genetic algorithm (GA) was employed to optimize the parameters of the operating conditions of plasma actuators (PAs). In this study, the multi-objective problem around a circular cylinder was considered. The objective functions are the lift maximization and the drag minimization. Two PAs were installed on the upper and the lower side of the cylinder. This problem was similar to the airfoil design, because the circular has potential to work as airfoil due to the control of flow circulation by the PAs with four design parameters. The aerodynamic performance was assessed by wind tunnel testing to overcome the disadvantages of time-consuming numerical simulations. The developed optimization system explores the optimum waveform of parameters for AC voltage by changing the waveform automatically. Based on these results, optimum designs and global design information were obtained while drastically reducing the number of experiments required compared to a full factorial experiment. An analysis of variance and a scatter plot matrix were introduced for design knowledge discovery. According to the discovered design knowledge, it was found that duty ratios for two PAs are an important parameter to create lift while reducing drag.

**Keywords:** Plasma Actuator, Genetic Algorithm, Efficient Global Optimization, Experimental Evaluation

## Introduction

Plasma actuators (PAs, shown in Fig. 1) are flow control devices that utilize atmospheric pressure discharge [Matsuno et al. (2008; 2009)] they have gained attention in recent years, because their advantages of being fully electronically driven with no moving parts and having a simple structure and a fast response are potentially ideal for application to subsonic flow control. Such active flow control devices have potential to control of the circulation around arbitrary objects and produce the lift-creating object even if it is not airfoil geometry.

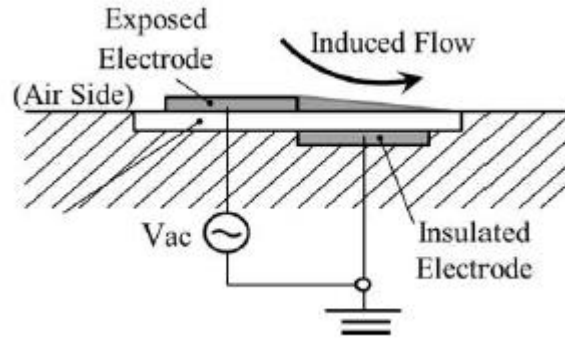
In this study, the design problem is defined as the optimization of lift creation and drag reduction via flow circulation controlled by the PAs. A circular cylinder model is used as a model and two PAs are installed. Thus, the objective functions considered in this paper are the maximizing lift and the minimizing drag around the circular cylinder. A multi-objective genetic algorithm (MOGA)-based efficient design technique was employed with wind tunnel testing to efficiently find the optimum designs. Through the design case, the applicability of the present wind tunnel testing to the multi-objective/ multi-parameter design problem was also investigated.

Design problems are often solved by GAs based on numerical simulation, such as computational fluid dynamics (CFD) [Kanazaki, et al. (2004), Kanazaki and Jeong (2007)]. However, there are several difficulties with solving the flow field around PAs. First, the accuracy of existing simulation methods is still insufficient. Second, the computational cost is very high for design techniques such as GAs. Several days are needed to acquire the results for each case, whereas the actual flow physics finishes in a few seconds.

In MOGA based efficient design technique, Kriging surrogate model was applied to represent the input/output relationship in the experimental data to reduce the experimental cost. This optimization technique, which is called efficient global optimization (EGO) [Donald, et al. (1998); Jeong, et al.



(2005); Kanazaki and Jeong (2007); Kanazaki, et al. (2008)], enables the optimization of global parameters in a small number of experiments while simultaneously obtaining information on the design space. The EGO based on Kriging surrogate model can find efficiently near-global optimum. In this study, Kriging surrogate model based GA performs optimization during a wind tunnel experiment in real time. The design system is automated developing the interface between the optimization and the wind tunnel testing.



**Figure 1. Schematic of plasma actuator.**

### Overview of Active Flow Control by Means of Plasma Actuator

In this research, a PA consisting of an exposed electrode and insulated electrode was used. A nonconductor was placed between the two electrodes, and AC voltage was applied. Fig. 1 shows the setup; this type of PA is called a single dielectric barrier discharge (SDBA) PA. The flow around the PA can be controlled by changing the number and location of PAs and the waveform of the AC voltage. Thus, the optimal technique for solving the design problem has to handle many parameters to acquire the best flow control.

Generic home-style AC voltage has a waveform with a constant frequency. However, several studies have reported that pulse width modulation (PWM) is effective for flow control of PAs. PWM is a drive system that turns the AC voltage on or off, as shown in Fig. 2. The frequency of on/off is defined as the "modulation frequency" ( $f_{\text{mod}}$ ) and is expressed by following equation:

$$f_{\text{mod}} = \frac{1}{T_1} \quad [\text{Hz}] \quad (1)$$

where  $T_1$  is the time of one cycle and  $T_2$  is the time the AC voltage is on. The ratio of  $T_2$  to  $T_1$  is defined as the duty ratio, which is an important parameter for PWM. The duty ratio ( $D_{\text{cycle}}$ ) is expressed by the following equation:

$$D_{\text{cycle}} = 100 \frac{T_2}{T_1} \quad [\%] \quad (2)$$

### Design Method: Efficient Global Optimization

#### *Efficient Global Optimization (EGO)*

The optimization procedure (Fig. 4) for PA design consists of the following steps. First,  $N$  design samples are selected by Latin hypercube sampling (LHS) [Donald, et al. (1998); Jeong, et al. (2005); Kanazaki and Jeong (2007); Kanazaki, et al. (2008)], which is a space filling method, and then assessed for the construction of an initial Kriging surrogate model. Second, an additional design sample is added, and the design accuracy is improved by constructing a Kriging model based on all  $N + 1$  samples. Note that the additional sample is selected by using expected improvement (EI) maximization [Donald, et al. (1998); Jeong, et al. (2005); Kanazaki and Jeong (2007); Kanazaki, et al. (2008)]. GA is applied to solve this maximization problem. This process is iterated

until the improvement of the objective functions becomes negligible. Through the design procedure proposed in this paper, all samples are evaluated by the wind tunnel testing.

The Kriging models express the value  $y(x_i)$  at the unknown design point  $x_i$  as

$$y(x_i) = \mu + \epsilon(x_i) \quad (i = 1, 2, \dots, m) \quad (3)$$

where  $m$  is the number of design variables,  $\mu$  is a constant of the global model, and  $\epsilon(x_i)$  represents a local deviation from the global model. The correlation between  $\epsilon(x_i)$  and  $\epsilon(x_j)$  is strongly related to the distance between the corresponding points,  $x_i$  and  $x_j$ . In the Kriging models, the local deviation at an unknown point  $x$  is expressed using stochastic processes. Specifically, a number of design points are calculated as sample points and then interpolated using a Gaussian random function as the correlation function to estimate the trend of the stochastic process.

Once the models are constructed, the optimum point can be explored using an arbitrary optimizer. However, it is possible to miss the global optimum design, because the approximate model includes uncertainty. Therefore, this study introduced EI values as the criterion. This study solves the lift maximization problem, then EI for maximization problem can be calculated as follows:

$$E[I(x)] = (f_{\max} - \hat{y}) \Phi \left( \frac{f_{\max} - \hat{y}}{s} \right) + s\phi \left( \frac{f_{\max} - \hat{y}}{s} \right) \quad (4)$$

EI for maximization problem can be calculated as follows:

$$E[I(x)] = (\hat{y} - f_{\min}) \Phi \left( \frac{\hat{y} - f_{\min}}{s} \right) + s\phi \left( \frac{\hat{y} - f_{\min}}{s} \right) \quad (5)$$

where  $f_{\max}$  and  $f_{\min}$  are the maximum and the minimum values among sample points, respectively.  $s$  is root mean square error (RMSE) and  $\hat{y}$  is the value predicted by Eq. 3 at an unknown point  $x$ .  $\Phi$  and  $\phi$  are the standard distribution and normal density, respectively. EI considers the predicted function value and its uncertainty, simultaneously. Therefore, by selecting the point where EI takes the maximum value, as the additional sample point, robust exploration of the global optimum and improvement of the model can be achieved simultaneously as shown in Fig. 4 because this point has a somewhat large probability to become the global optimum. In this study, the maximization of EI is carried out using GA.

GAs (Fig. 5(a)) was first proposed by Holland in the early 1970s [Holland, et al. (1975)] and are based on the evolution of living organisms with regard to adaptation to the environment and the passing on of genetic information to the next generation. GAs can find a global optimum because they do not use function gradients, which often lead to an exact local optimum. Thus, GA is a robust and effective method that can handle highly nonlinear optimization problems involving nondifferentiable objective functions. Owing to this advantage, GAs were applied to this experimental system. The GA used in this study [Kanazaki, et al. (2004); Donald, et al. (1998); Jeong, et al. (2005); Kanazaki and Jeong (2007); Kanazaki, et al. (2008)] utilizes a real-coded representation, the blended crossover (BLX- $\alpha$ ), and the uniform mutation. The selection probability of individuals for the crossover and mutation is expressed as follows:

$$prob = c(1 - c)rank1 - 1.0 \quad (6)$$

where rank is the value of fitness ranking among the population. In BLX- $\alpha$ , children are generated in a range defined by the two parents as shown in Fig. 5(b). The range is often extended equally on both sides as determined by the parameter  $\alpha$ .

*Analysis of Variance (ANOVA)*

In this study, Kriging model based ANOVA [Donald, et al. (1998); Jeong, et al. (2005); Kanazaki and Jeong (2007)] is employed to investigate the effect of the design variables to objective functions. Variance of a surrogate model can be calculated as,

$$\mu_i(x_i) \equiv \int \cdots \int \hat{y}(x_1, \cdots, x_n) dx_1, \cdots, dx_{i-1}, dx_{i+1}, \cdots, dx_n - \mu \tag{7}$$

where the total mean  $\mu$  is calculated as

$$\mu \equiv \int \cdots \int \hat{y}(x_1, \cdots, x_n) dx_1, \cdots, dx_n \tag{8}$$

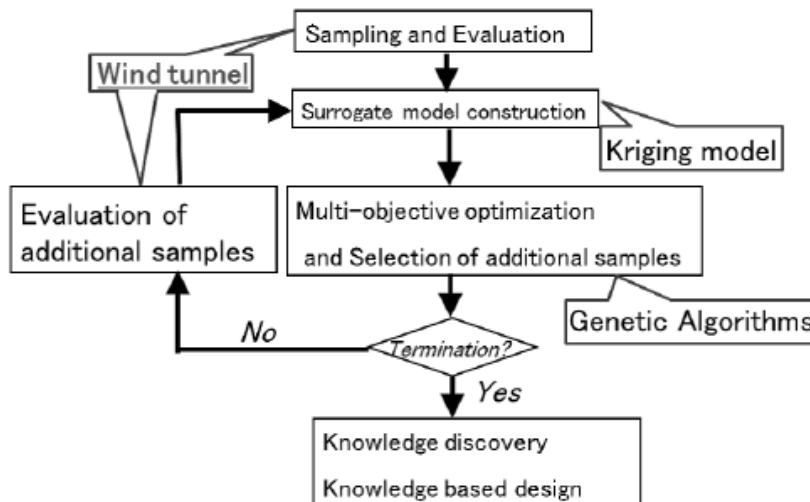
The proportion of the variance attributed to the design variable  $x_i$  to the total variance of the model can be expressed as:

$$p \equiv \frac{\int [\mu_i(x_i)]^2 dx}{\int \cdots \int [\hat{y}(x_1 \cdots x_n) - \mu]^2 dx_1 \cdots dx_n} \tag{9}$$

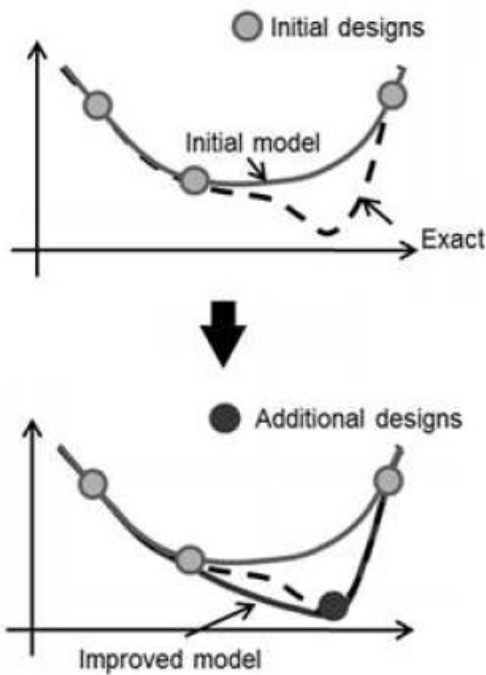
The value obtained by Eq. (9) indicates the sensitivity of an objective function to the variance of a design variable.

*Scatter Plot Matrix (SPM)*

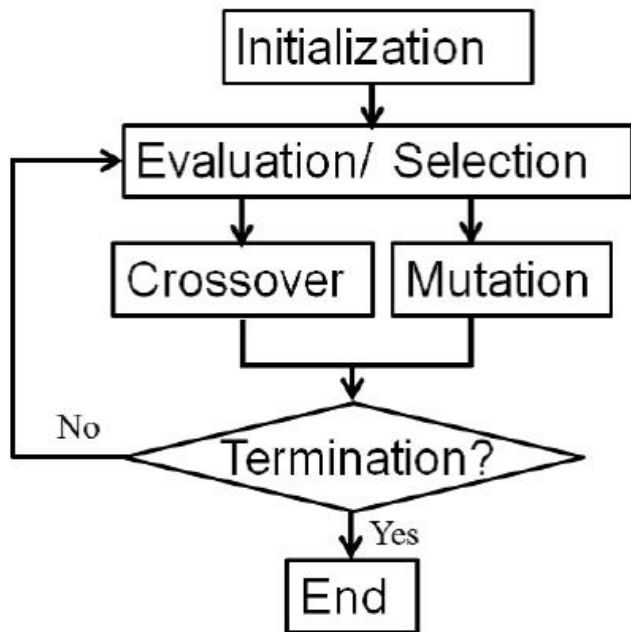
The solution and the design space of the multivariable design problem obtained by EGO are observed by the SPM which is one of the data mining, because the Kriging model cannot be visualized directly when the design problem has over four attribute values. SPM arranges two-dimensional scatter plots like a matrix among the objective functions and the design variables and facilitates the investigation of the design problem investigation. Each of the rows and columns is assigned attribute values such as design variables, objective functions, and constraint values. The diagonal elements show mutual same plots. Therefore, it can be said that the SPM shows scatter plots on the upper triangular part of the matrix and the correlation coefficients on the lower triangular part as additional information. modeFrontier™ ver. 4. 4. 2 is employed in this study.



**Figure 3. Optimization procedure based on wind tunnel evaluation**



**Figure 4. Improvement of the global model by expected improvement (EI) maximization**



**Figure 5. Schematic illustration of genetic algorithm**

## Experimental Setup

### *Wind Tunnel and Model*

The wind tunnel experiments were conducted in the subsonic closed-return wind tunnel of the Aerodynamics Laboratory at Tottori University. The wind tunnel has a closed test section with a  $0.70 \text{ m} \times 1.0 \text{ m}$  cross-section and  $2.0 \text{ m}$  length (Fig. 6). A two-dimensional circular model ( $105 \text{ mm}$  in diameter) was used as shown in Fig. 7. Model was placed on a flat plate and mounted to a support connected to a six-component external balance for measurement of the aerodynamic forces and moments. The output of the balance was amplified and acquired with a data acquisition board (National Instruments PXI-8106). The output signal contains noise from the atmospheric discharge of the plasma actuators. To eliminate this effect, the clean portion of the signal, during which the discharge did not appear, was extracted and used as a “clean” portion of the data.

### *PA and Its Power Supply*

In this study, two PAs were installed on the surface of the model. PA#1 and PA#2 were installed with mount angles of  $\theta_1 = 85.0^\circ$  and  $\theta_2 = -85.0^\circ$ , respectively, as shown in Fig. 7. The reference waveform of a high-voltage AC input was amplified by a solid-state high-power amplifier; the input power was increased up to  $400.0 \text{ W}$  with amplitude of  $70.0 \text{ Vpp}$ . A high-voltage transformer was used to achieve an AC input with amplitude of up to  $30 \text{ kV}$  at a frequency of  $5.0\text{-}15.0 \text{ kHz}$ . The voltage and current of the AC input were monitored by an oscilloscope along with the reference waveform.

### *Integration of Experiment System*

Figure 8 shows the schematic illustration of the developed system. EGO is executed in the workstation and receives the experimental data via LabVIEW® from the balance in the wind tunnel. The condition of the AC voltage can be automatically set during the optimization process based on

balance measurements.

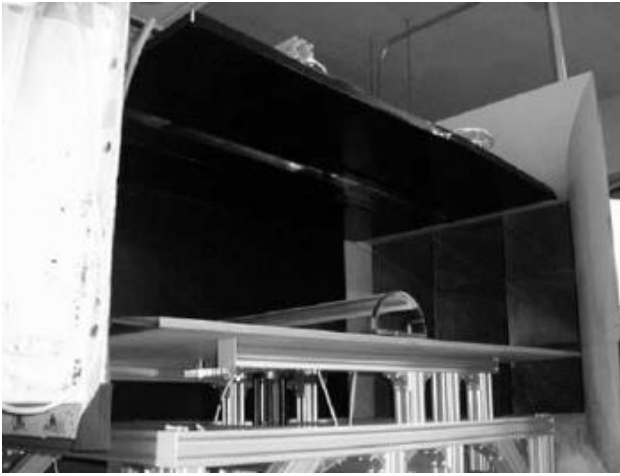


Figure 6. Test section of the wind tunnel

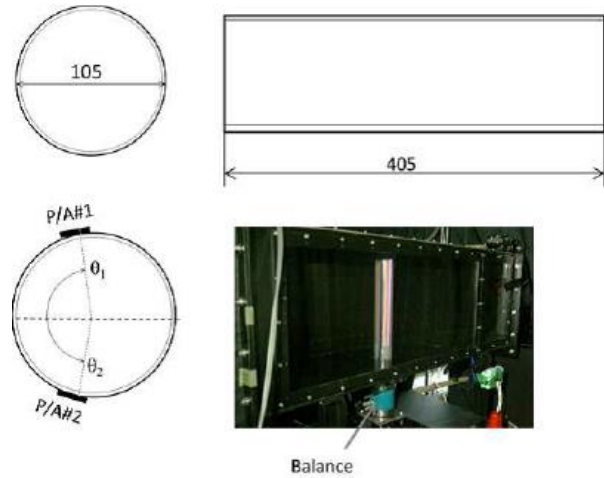


Figure 7. Circular cylinder model and the location of plasma actuators

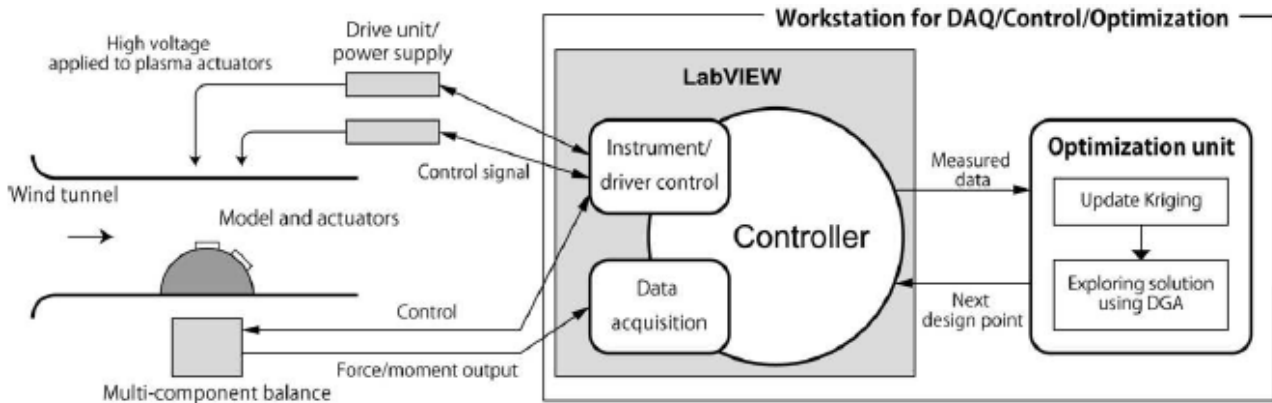


Figure 8. Schematic diagram of the integrated optimization system

## Formulation

In this study, multi-objective/ multi-parameter design problem which has four design variables was considered and the lift creation and drag reduction effect due to circulation control by PAs was investigated. The objective function was maximization of the lift coefficient ( $C_l$ ) and the minimization of the drag coefficient ( $C_d$ ) around the circular cylinder model. This design problem can be expressed as follows:

$$\begin{cases} \text{Maximize } C_l \\ \text{Minimize } C_d \end{cases} \quad (10)$$

The flow velocity was set to 10.0 m/s. The design problem expressed in Eq. (10) was solved by changing four parameters ( $f_{\text{mod}}$ ,  $D_{\text{cycle1}}$ ,  $D_{\text{cycle2}}$ ,  $\phi$ ) related to the AC voltage waveform. In this case, two PAs are applied different  $D_{\text{cycle}}$ ;  $D_{\text{cycle1}}$  and  $D_{\text{cycle2}}$ , for each design and the difference between  $D_{\text{cycle1}}$  and  $D_{\text{cycle2}}$  is decided by a phase difference  $\phi$ . The design space is defined as follows:

$$\begin{cases} 30.0 \leq f_{\text{mod}} \leq 200.0 \text{ [Hz]} \\ 0.0 \leq D_{\text{cycle1}} \leq 50.0 \text{ [%]} \\ 0.0 \leq D_{\text{cycle2}} \leq 50.0 \text{ [%]} \\ -90.0 \leq \phi \leq 90.0 \text{ [deg.]} \end{cases} \quad (11)$$

$\phi$  is the phase difference between PA#1 and PA#2. Consequently, the time lag can be expressed as  $\phi / f_{\text{mod}}$ .

## Results

### *Design Exploration Result*

In this section, the design problem expressed by Eq. (10) is discussed. To construct the initial Kriging model, 15 samples were obtained by LHS. To acquire additional samples, the island GA was executed with the following specifications: BLX-0.5 ( $\alpha = 0.5$ ), four subpopulations, 16 individuals for each subpopulations (64 individuals generated in total) and 64 generations. The EGO process will be stopped after ten or more additional samples show better function value than that of initial samples.

After the objective function was converted, seven additional samples were obtained, for a total of 22 sample designs. Figure 9 shows the history of  $C_1$  values for the sampling process. According to the history, the objective function converged well with a small number of samples.

### *Design Knowledge by Analysis of Variance*

Figure 10(a) shows the main effects and the two-way interaction of the design variables for objective function for  $C_1$ . According to Fig. 10(a),  $f_{\text{mod}}$  and  $D_{\text{cycle2}}$ , which defines the driving condition of PA on the lower side of the cylinder, has a predominant influence on  $C_1$ . In addition, two-way interaction of  $f_{\text{mod}} - D_{\text{cycle2}}$  is also effective to  $C_1$ . These results suggest that the circulation which creates aerodynamic lift around the model is decided by duty ratio PA on the lower side.

Figure 10(b) the main effects and the two-way interaction of the design variables for objective function for  $C_d$ . According to Fig. 10(a),  $f_{\text{mod}}$  which defines the driving condition of each PA on the cylinder has a predominant influence on  $C_d$ . It is reasonable result because higher  $f_{\text{mod}}$  create higher volume force which can reduce the flow separation. As this result, the drag is affected by  $f_{\text{mod}}$ .

### *Visualization of Design Problem by SPM*

Figure 11 shows the visualization results obtained by SPM, which shows the scatter plot for all parameter combinations. Plots colored by red represent designs which achieve higher aerodynamic performance. According to Fig. 11, higher  $f_{\text{mod}}$  and  $D_{\text{cycle2}}$  are always required for higher  $C_1$  and lower  $C_d$ . In addition,  $C_d$  and  $f_{\text{mod}}$  shows the high correlation (-0.882.) This result suggests that the lower  $C_d$  can be carried out with the higher  $f_{\text{mod}}$ .

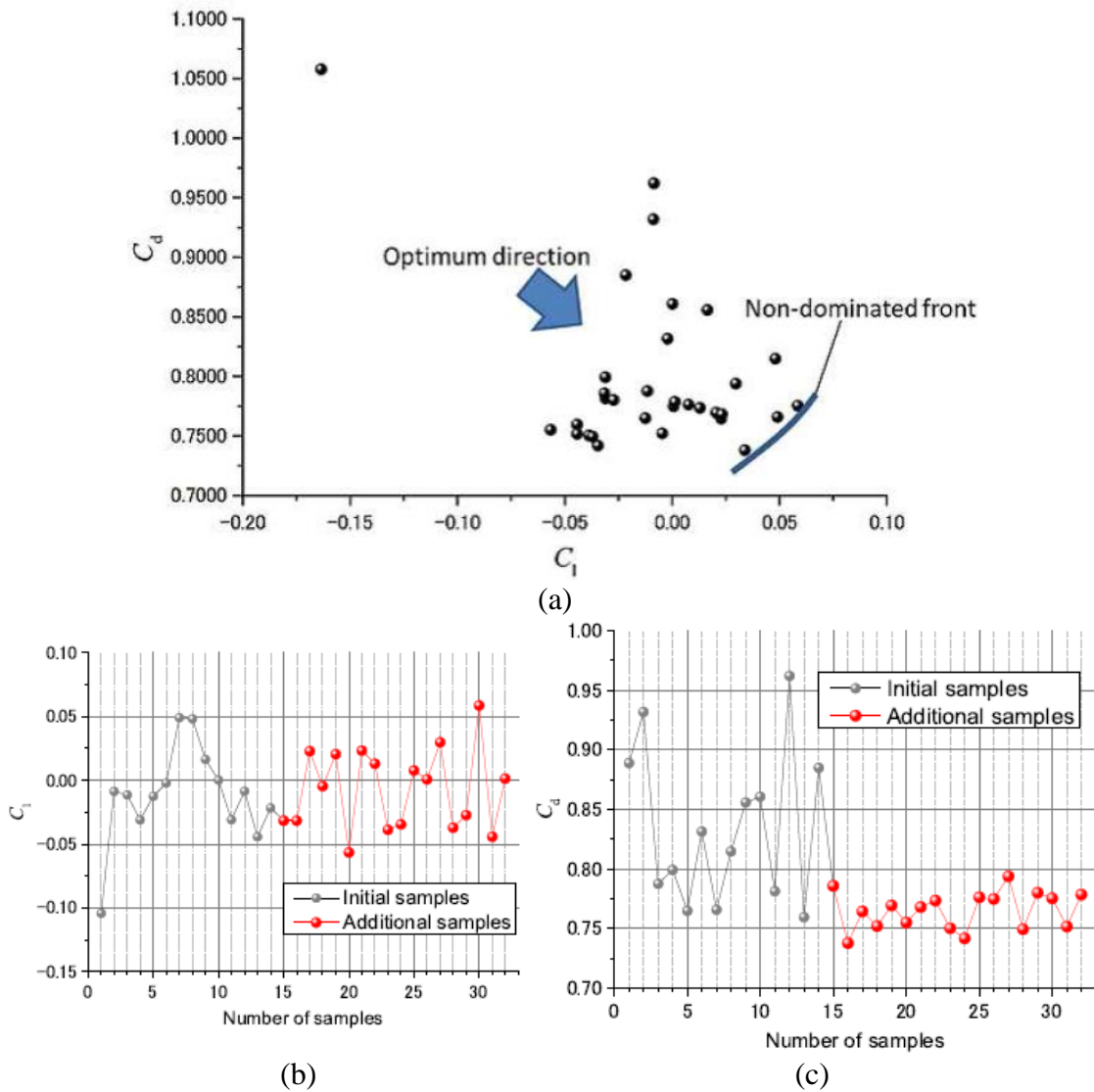


Figure 9. Solutions of the design problems (a)Design result, (b)Progression of objective function with sample number for the  $C_d$  and (c)Progression of objective function with sample number for the  $C_d$

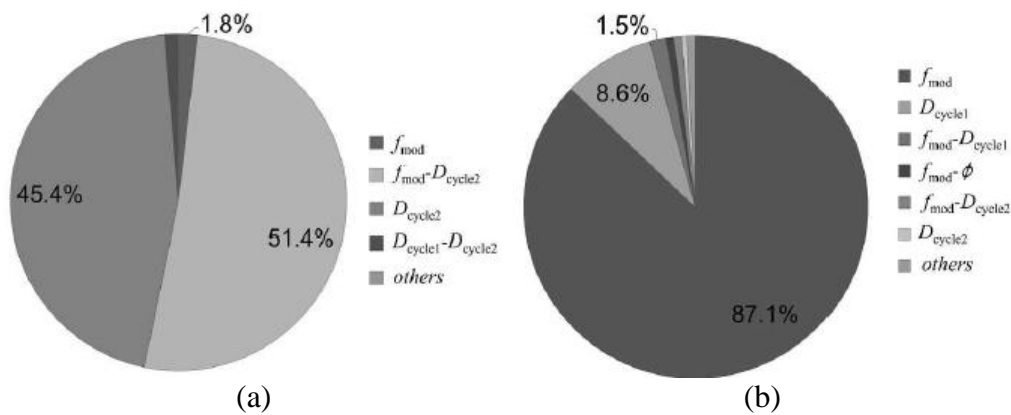


Figure 10. ANOVA results. (a)Effect of design variables of the design variables for  $C_1$  and (b)Effect of design variables of the design variables for  $C_d$

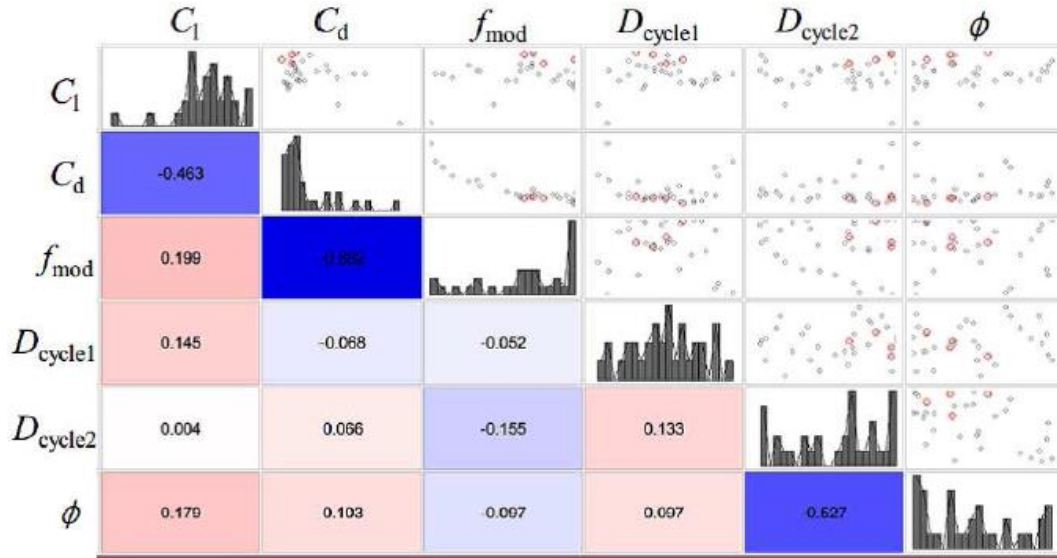


Figure 11. Visualization of the design problem using SPM

## Conclusions

Aerodynamic control performance of plasma actuators was optimized using wind tunnel test-based EGO. In this study, the lift-creating cylinder using plasma actuators was considered. This problem was that the circulation around a circular cylinder model was controlled to maximize the lift around the model. In addition, this study also considered the drag minimization around the cylinder, thus the design problem was formulated as the multi-objective problem. The optimization technique is firstly integrated in the operating system of the wind tunnel experiment to enable automation of the data-acquisition/ optimization process. Using the developed system, multi-objective design problem (lift maximization/ drag minimization) was solved. After several additional samples are obtained, the analysis of variance and the scatterplot matrix is employed for the knowledge discovery. Using these techniques, it is found that duty ratio and modulation frequency for the plasma actuators installed on the lower surface have the dominant effect for this problem. It is also found that the higher modulation frequency is required for the plasma actuator to minimize the drag.

## References

- Matsuno, T., Kawazoe, H., and Corke, T. C. (2008) Forebody vortex control on high performance aircraft using PWM-controlled plasma actuators, *Proc. of the 26th International Council of Aeronautical Sciences, ICAS 2008-3.3.4*, Alaska, USA , [CD-ROM].
- Matsuno, T., Kawazoe, H., and Nelson, R. C. (2009) Aerodynamic control of high performance aircraft using pulsed plasma actuators, *Proc. of the 47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, AIAA paper 2009-0697, Florida, USA , [CD-ROM].
- Kanazaki, M., Obayashi, S., and Nakahashi, K., (2004) Exhaust manifold design with tapered pipes using divided range MOGA, *Engineering Optimization, Taylor & Francis*, 36(2), 149-164.
- Donald, R. J., Matthias, S., and W. William. J. (1998) Efficient global optimization of expensive black-box function, *Journal of Global Optimization* 13, 455-492.
- Jeong, S. Murayama, M., and Yamamoto, K. (2005) Efficient optimization design method using Kriging model, *Journal of Aircraft*, 42(2), 413-420.
- Kanazaki, M., and Jeong, S. (2007) High-lift airfoil design using kriging based moga and data mining, *Korea Society for Aeronautical and Space Sciences International Journal* 8(2), 28-36.
- Kanazaki, M., Yokokawa, Y., Murayama, M., Ito, T., Jeong, S., and Yamamoto, K., (2008) Nacelle chine installation based on wind tunnel test using efficient design exploration, *Transaction of Japan Society and Space Science*, 51(173), 146-150.
- McKay, M. D., Beckman, R. J., and Conover, W. J., (1979) A Comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, 21(2), 239-24.
- Holland, J.H, (1975) *Adaptation in natural and artificial systems*, University of Michigan Press Ann Arbor.



# The Failure Behavior of Composite Honeycomb Sandwich Structure with Stringer Reinforcement and Interfacial Debonding

†\*Ruixiang Bai<sup>1</sup>, S.H. Bao<sup>1</sup>, X.Y. Liu<sup>2</sup>, and X.H. Liang<sup>1</sup>

<sup>1</sup>State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of technology, China.

<sup>2</sup>School of Aeronautics and Astronautics, Shanghai Jiaotong University, Shanghai 200240, China.

\*Presenting author: bairx@dlut.edu.cn

†Corresponding author: bairx@dlut.edu.cn

## Abstract

As an efficient lightweight structure, composite honeycomb sandwich panel has been widely used in many industries. The composite honeycomb sandwich structure with stringer reinforcement is a new type of sandwich structure. This paper investigated the damage and failure behavior of composite honeycomb sandwich structure with stringer reinforcement under in-plane compression condition. Some critical damage modes and failure behavior of composite sandwich structure with stringer reinforcement were revealed. Three different kinds of debonding damage of interface between sheet and core were considered, the failure modes as well as the whole failure process were obtained by numerical simulation.

**Keywords:** Sandwich structure, Stringer reinforcement, Composite, Interfacial Debonding, Failure

## Introduction

Advanced sandwich structure is usually a large thickness of honeycomb core bonded with composite sheets. With larger in-plane stiffness and strength, the composite material faceplate is mainly used to bear the axial load, bending moment and shearing action, while the honeycomb core, subject to bending and shear load, is mainly used to maintain the stability the relative position of sheets and transfer lateral load. With the advantages of high specific stiffness and specific strength, the structure can get high flexural stiffness and compressive yield strength under the condition of low specific gravity<sup>[1]</sup>.

The faceplate and core of advanced sandwich structure are anisotropic, which is a very important characteristic. Through the reasonable design of the composite faceplate or rational choose of the core structure, optimization sandwich structure can be designed and manufactured to meet the specific needs of various engineering applications<sup>[2]</sup>. The composite honeycomb sandwich structure with stringer reinforcement is a new type of sandwich structure, whose purpose is to further balance improve the axial and bending specific stiffness and specific strength of the structure, at the same time increase the reliability of the structure.

Due to the characteristics of manufacturing technology and the intrinsic properties of the materials, the debonding defect is easy to occur in the interface between the core and the sheets during service life<sup>[3]</sup>. As a result, the strength under static load will be decreased. Moreover, the failure mode of the sandwich structure will be more complicated, and the defects will seriously affect the accuracy of strength prediction. For the composite honeycomb sandwich structure with stringer reinforcement, the effect of stiffener on the failure modes of sandwich structure is worth studying.

## The equivalent of the material parameters and numerical model

There are two main simulation methods for the sandwich structure<sup>[4]</sup>. For hierarchical model, each single layer of the structure is considered respectively, and the constraints according to continuity for each interface also should be given appropriately to meet the requirements of stresses generality for adjacent layers. For the equivalent single-layer model, the sheet and core are replaced by a single-layer with certain thickness. The unified expression of displacement field is given along whole thickness direction by using the deformation theory of plate and shell. For the hierarchical

model, it has a large number of independent variables, while for the equivalent single-layer model, as the independent variables are less, it is commonly used in finite element method. To simplify the analysis, the equivalent single-layer model is adopted. The selected aluminum honeycomb core of the sandwich structure is transformed to homogeneous orthotropic material in finite element modeling. There are a variety of equivalent ways for the elastic parameters of honeycomb core<sup>[5]</sup>. The equivalent elastic parameters of hexagonal honeycomb core proposed by Zhao Jin-Sen<sup>[6]</sup> are adopted in this paper to derivated formula and calculate the equivalent material parameters of the simplified model. The equivalent formulas are as follows:

$$\left. \begin{aligned} E_{cx} = E_{cy} &= \frac{4}{\sqrt{3}} E_s \left( 1 - 3 \frac{t^2}{l^2} \right) \frac{t^3}{l^3} \\ E_{cz} &= \frac{2}{\sqrt{3}} E_s \frac{t}{l} \\ G_{cxy} &= \frac{\sqrt{3}}{3} E_s \left( 1 - \frac{t^2}{l^2} \right) \frac{t^3}{l^3} \\ G_{cxz} = G_{cyz} &= \frac{\sqrt{3}}{3} \left( \frac{t}{l} \right) G_s \\ \nu_{12} &= 1 - 4 \frac{t^2}{l^2} \end{aligned} \right\} \quad (1)$$

Where  $E_s$  and  $G_s$  are elastic parameters of the honeycomb core,  $l$  and  $t$  are wall length and wall thickness of a unit cell of the honeycomb core.

The equivalent properties of honeycomb core are given in Table 1 below.

**Tab.1 The equivalent properties of honeycomb core**

Elastic Parameter (Gpa)					Poisson's ratio	
$E_1$	$E_2$	$E_3$	$G_{12}$	$G_{13}$	$G_{23}$	$\nu_{12}$
0.31	0.31	1003	0.078	189	189	0.99

The traditional composite honeycomb sandwich structure consists of two composite sheets, adhesive layer and aluminum honeycomb core. For the stringer reinforced sandwich structure discussed in this paper, two buried aluminum stiffeners are contained. The adhesive layer is simulated by cohesive element in finite element analysis.

The overall dimensions of the two kinds of sandwich structure are uniform, the length is 90 mm, the width is 50 mm, and the total thickness of 15 mm, among which, the thickness of the honeycomb core is 12 mm, the thickness of adhesive layer is 0.1mm, and both of the thickness of the upper and lower sheets are 1.4 mm. The components size meets the requirements of ASTM C364-99 standard. the upper and lower faceplates are composite laminates for the two kind of sandwich structure, whose length and width directions are defined as  $x$  and  $y$  axis, respectively. The composite laminates has a total of 10 layers, the thickness of each layer is 0.14 mm, and the stacking sequence is  $[0/0/45/-45/90]_s$ . The material parameters of the composite laminates are shown in table 2. In addition, the two buried aluminium stiffeners are 90mm in length, 4mm in width, and 12mm in height. The distance from the two stiffeners to the center line of the sandwich structure is 12 mm. The material parameters of the aluminium stiffeners are shown in table 3.

**Tab.2 Properties of T300/QY8911**

Elastic Parameter (Gpa)				
$E_1$	$E_2 = E_3$	$G_{12} = G_{13}$	$G_{23}$	$\nu_{12}$
126	10.7	4.47	3.57	0.33
Strength Parameter (Mpa)				
$X_T$	$X_C$	$X_T$	$Y_C$	$S$
1548	1226	55.5	218	89.9

**Tab.3 Properties of the aluminum stringer**

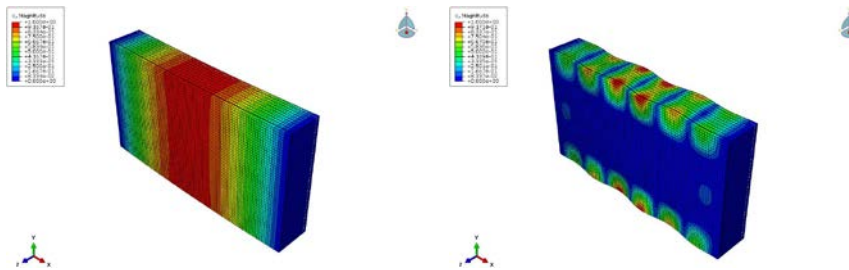
Properties (units)	Young's modulus (GPa)	Poisson's ratio $\nu_{12}$
Aluminum stringer	69.5	0.33

**The failure analysis of the sandwich structure**

*The linear buckling analysis*

Lanczos vector method is adopted to analyze and compare the linear buckling deformation characteristics of composite honeycomb sandwich structure without reinforcement and with stringer reinforcement, respectively. The main buckling modes of two kinds of sandwich structure are calculated, which are shown in figure 1.

In figure 1, we can see that the mainly buckling modes of the two kinds of sandwich structure are different under in-plane compression condition. global buckling instability mainly occurs to the sandwich structure without reinforcement, while partial buckling mainly occurs to the sandwich structure with stringer near the free boundary on both sides. Due to the existence of the stiffener, the buckling deformation of the honeycomb core is inhibited, and the overall stiffness of the structure is enhanced effectively. What's more, the buckling load of the sandwich structure with stringer is 525.71 KN, which is much higher than structure without reinforcement buckling load of 121.28 KN. Therefore, the stringer Reinforcement significantly improves the buckling bearing capacity of the composite honeycomb sandwich structure.

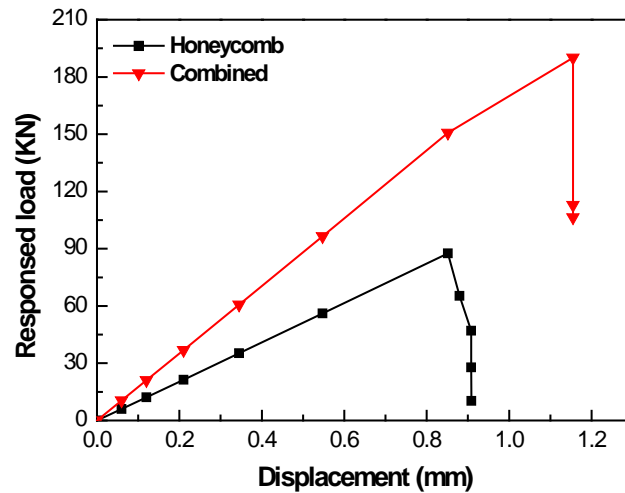


**Fig.1 The first buckling mode of the composite honeycomb sandwich structure with and without stringer reinforcement**

*The nonlinear failure analysis*

Figure 2 shows the load-displacement response of two different composite honeycomb sandwich structures under in-plane compression condition by nonlinear buckling analysis. Through observation, we know that the overall axial stiffness of the structure changed little, and axial compression stiffness approximate to linear under in-plane compression condition. After reaching limit loading points, failure damage occurs to both of the two kinds of sandwich structure, and the

continue carrying capacity loses quickly. Through comparison, we know that the limit load of the sandwich structure with stringer reinforcement is 190.03 KN, which is much higher than that of the structure without reinforcement as 87.52 KN. Therefore, the stringer Reinforcement effectively improves ultimate bearing capacity of the composite honeycomb sandwich structure. In addition, the specific strength of sandwich structure with stringer reinforcement is 1.05 times bigger than without reinforcement, which further evidences that composite honeycomb sandwich structure with stringer reinforcement has excellent structural performance.



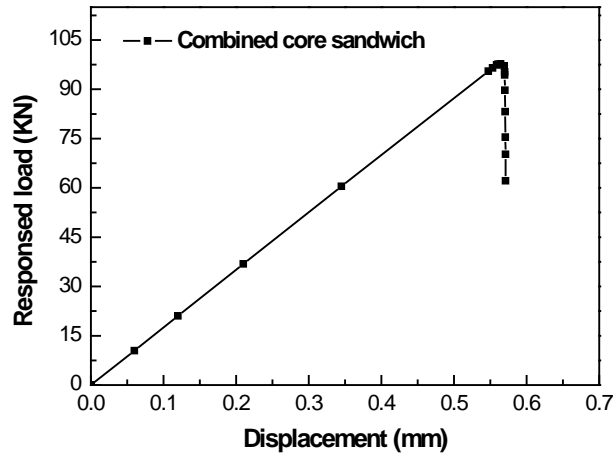
**Fig.2 The load-displacement response of two different composite honeycomb sandwich structures**

Comparing the results of nonlinear failure analysis and the linear buckling analysis respectively, we know that the linear buckling load is greater than the limit load for both of the two kinds of sandwich structure. Accordingly, the overall stability of composite honeycomb sandwich structures under in-plane compression condition is high, and the stiffness of the structure is further enhanced through stringer reinforcement. Therefore, the buckling failure is not the main failure modes of the structure, strength and damage are the main factors dominate the failure modes of sandwich plate in general.

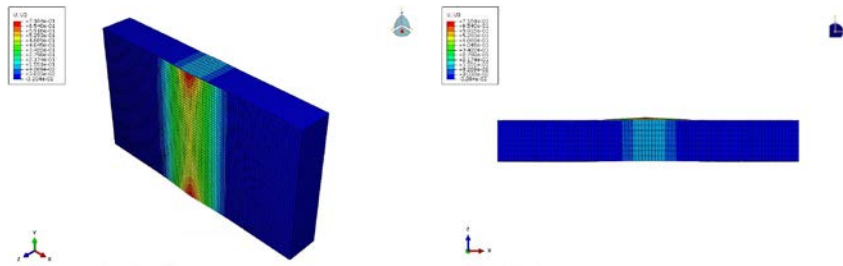
#### *The failure analysis of the sandwich structure with through interfacial debonding*

Considering a through-the-width sheet/core interfacial debonding in middle area of the reinforced composite honeycomb sandwich structure, and the length of debonding is 30mm. Figure 3 shows the load-displacement response of reinforced sandwich structure with a through-the-width interfacial debonding by nonlinear analysis. Analysis shows that, the relationship between load and axial displacement keeps linear, and will lose load carrying capacity quickly when reaches the limit load. The limit load of the sandwich structure with stringer reinforcement and through interfacial debonding is 97.72KN, which is much lower than that of the perfect reinforced sandwich structure as 190.03 KN. Therefore, the through-the-width interfacial debonding reduces ultimate bearing capacity of the sandwich structure.

Figure 4 shows the out-plane displacement of the sandwich structure with stringer reinforcement under the limit load. The result shows that local buckling occurred in the debonding area, and lead to the final broken. Also, because the stiffener improves the overall stiffness, local buckling only occurs at the debonded sheet near both sides of free boundary.

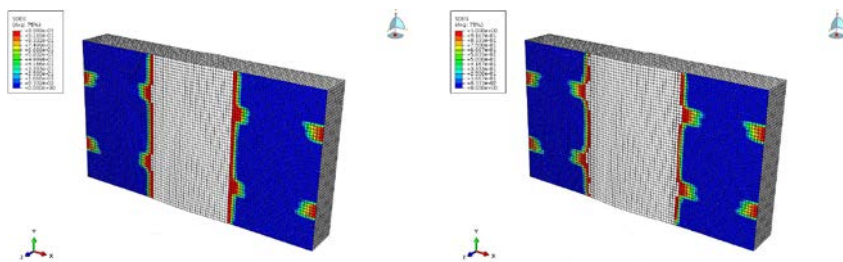


**Fig.3 The load-displacement response of sandwich structure with stringer reinforcement and through-the-width interfacial debonding**



**Fig.4 The out-plane displacement of the sandwich structure with stringer reinforcement under the limit load**

Figure 5 shows a symmetrical through-the-width interfacial debonding propagation behavior located at both side of the core. Under compression load, partial buckling occurs in the upper and lower sheets in the zone of debonding. With the increase of compression load, the lower sheet in debonding area contacts the core quickly thus inhibits the failure and propagation of the adhesive layer. At the same time, the upper sheet in debonding area bulges outward, free buckling occurs. Because the stiffness of the stiffener is higher than the honeycomb core, the debonding propagation starts at the interface between sheet and stiffener. With the increase of the compression loading, the displacement of bulging outward increasing gradually and the debonding propagation gradually extends to the interface of sheet/core near the initial debonding propagation of sheet/stiffener.



**Fig.5 The propagation of symmetrical through-the-width interfacial debonding in sandwich structure with stringer reinforcement**

## Conclusions

Reinforced by stringer reinforcement, the overall stiffness of the composite Honeycomb Sandwich Structure with Stringer Reinforcement is enhanced effectively, the buckling and ultimate bearing capacity are improved. Under in-plane compression condition, the buckling failure is not the main failure modes of the structure, while strength conditions are main factors controlling the sandwich structure damage in general. Due to the existence of interfacial debonding, local buckling will occur in the debonding area, and lead to the final broken. With the increase of the compression loading, the displacement of bulging outward increasing gradually and the debonding propagation gradually extends to the interface of sheet/core near the initial debonding propagation of sheet/stiffener.

## Acknowledgments

This work was supported by National Basic Research Program of China(No.2014CB046506), and the Project of State Key Laboratory of Structural Analysis for Industrial Equipment (No. S14207)

## References

- Fan, Q. X. (1998) Honeycomb Sandwich Composite, *Journal of Beijing institute of light industry*, 2, 77-81. (In Chinese)
- Liu, S. L., Guo, W. T. and Ma, Y. P. (2014) Mechanical Properties of Web-core Reinforced Sandwich Structures, *Development and application of materials*, 5, 62-67. (In Chinese)
- Zhang, L. H., Han, Y., Wo, X., Y., Xia, Y., F. and Tan, F. (2006) Analysis on Process Defects of Honey Comb Sandwich, *Spacecraft Recovery & Remote Sensing*, 27, 1, 57-61 (In Chinese)
- H., A. (2000). An alternative determination of transverse shear stiffnesses for sandwich and laminated plates. *International Journal of Solids & Structures*, 37, 25, 3503-3520(18).
- Fu, M., H., Xu, O., T. and Chen, Y. (2015) An Overview of Equivalent Parameters Honeycomb Cores, *Material View*, 5, 024. (In Chinese)
- Zhao, J.,S. (2006) Research On Equivalent Models Of The Mechanical Function For Aluminium Honeycomb Sandwich Panel, *Nanjing University of Aeronautics and Astronautics The Graduate School* (In Chinese)
- Bai, R. X. and Chen, H. R. (2002) Investigation Of Nonlinear Buckling Behavior Of Composite Sandwich Plates With Interfacial Delamin Ation And Matrix Micro-Cracks, *Journal of composite materials*, 19, 2, 80-84. (In Chinese)

## Seismic earth pressure on a multi-story underground structure

†\*Z.Y. Chen<sup>1,2</sup>, W. Chen<sup>3</sup>, and H. Fan<sup>3</sup>

<sup>1</sup> Associate professor, Department of Geotechnical Engineering, Tongji University, Shanghai, China.

<sup>2</sup> State Key Laboratory of Disaster Reduction in Civil Engineering, Shanghai, China.

<sup>3</sup> Master candidate, Department of Geotechnical Engineering, Tongji University, Shanghai, China.

\*Presenting author: zhiyichen@tongji.edu.cn

†Corresponding author: zhiyichen@tongji.edu.cn

### Abstract

For a multi-story underground structure, distribution and amplitude of seismic earth pressure along its side wall depth during an earthquake are critical for seismic design and safety evaluation. In this paper, a series of 1-g shaking table tests were conducted on a four-story subway station firstly. Experimental results showed that the distribution of maximum lateral dynamic earth pressure appeared an “S” shape, which was distinguished from that of a single-story underground structure. In the latter case, it is generally in a form of a triangle distribution. Furthermore, parametric study was carried out through nonlinear dynamic time history analyses using the general purpose finite element code ABAQUS. Attention was paid on influences of types of soils, structural stiffness, and vertical earthquake component on the distribution of seismic earth pressure. Numerical results showed that structures surrounded by sand suffered larger dynamic earth pressure than those surrounded by clay. Peak dynamic earth pressure of a flexible structure was a little smaller than that of a rigid one. And vertical earthquake component excited lateral dynamic earth pressure in some degree.

**Keywords:** Seismic earth pressure, Multi-story underground structure, Shaking table test, Parametric study, Stiffness

### Introduction

Nowadays, with the rapid development of economy and society modern underground transportation, represented by the subway, has become popular. In China under the current plan, 36 cities were approved to build rail transit, and it is planning to reach 6000 km of rail transit by 2020 [Sun (2013)]. At the same time, the world has experienced a high incidence of earthquakes. The M 9.1 earthquake in Sumatra in 2004, the M 8.0 earthquake in Wenchuan, China in 2008, and the M 9.0 earthquake in the northeast of Japan in 2011 all caused a great loss of human life and property. Obviously, strong earthquakes heavily threaten underground subways those are building or built. Typical lateral design of foundation systems and retaining structures often relies on static earth pressure theories and tends to neglect seismic effects frequently due to the lack of understanding thereof and the shortcoming of experimental data. While this assumption can be accurate for foundation systems exposed to small levels of seismic shaking, stronger accelerations due to larger magnitude earthquakes can cause significant damage to the foundations and superstructures [Luu (2013)].

Hence, theoretical, numerical and experimental studies of seismic earth pressure on underground structures had been conducted in recent years. Wang et al. [2010] proposed a new method to calculate the seismic earth pressure of shallow buried underground structures by combining Xie theory and M-O formula. Ostadan [2005] conducted a series study on seismic soil structure interaction of building walls resting

on firm foundation materials and proposed a simplified method for predicting the maximum seismic pressures. Gazetas et al. [2004] carried out a numerical study of dynamic stresses imposed on a variety of retaining systems under short-duration and impulsive base excitation. Psarropoulos et al. [2005] developed a general finite-element method that specifically focused on the distribution of dynamic earth pressures on rigid and flexible walls. Shaking table test and centrifuge test are popular to study the dynamic soil structure interaction. Yang et al. [2003] conducted shaking table tests on a double-story subway station and found that the dynamic lateral earth pressures were large at the middle part and small at the top and bottom. Madabhushi and Zeng [2007] conducted centrifuge tests to investigate the seismic response of a cantilever retaining wall under earthquake loading and pointed that the effect of an earthquake was more severe on a cantilever retaining with saturated backfill than that with dry backfill.

However, amounts of studies are almost aimed at retaining structures. There are still some differences between retaining structures and underground subway structures, such as design structural parameters, boundary conditions, and stress conditions. Hence it is not appropriate to simply apply the computational methods for retaining structures to subway structures. Furthermore, most researches aimed at traditional subway structure, which was single story or double story. Owing to new requirements on functions of underground space, modern subway station developing towards having a multilevel and complex structural form. With the increase of layers, lateral stiffness of stations decreases significantly, which changes the distribution and amplitude of seismic earth pressures.

In this study, a series of 1-g shaking table tests were conducted on a four-story subway station firstly. Due to the lacking of resources and time, many influence factors cannot be considered in tests. Thus, the experimental results are used to confirm the ability of the numerical technique to simulate the dynamic earth pressure. Then study was carried out through nonlinear dynamic time history analyses using the general purpose finite element code ABAQUS. Influences of types of soils, structural stiffness, and horizontal and vertical earthquake components on the distribution of seismic earth pressure were studied.

### **Shaking table tests**

#### *Experimental setup*

The shaking table test was carried out using the MTS Company shaking table facility at the State Key Laboratory for Disaster Reduction in Civil Engineering, Tongji University. The table can be input with three-dimensional and six-degree-of-freedom motions. The dimensions of the table are 4 m × 4 m. The working frequency ranges from 0.1 to 50 Hz. The shaking table vibrates with two maximum horizontal direction accelerations of 1.2 g and 0.8 g, and a maximum acceleration of 0.7 g vertically. A flexible container was used in the test. The cylindrical soil container was 3000 mm in diameter. Figure 1 shows the shaking table and the soil container.

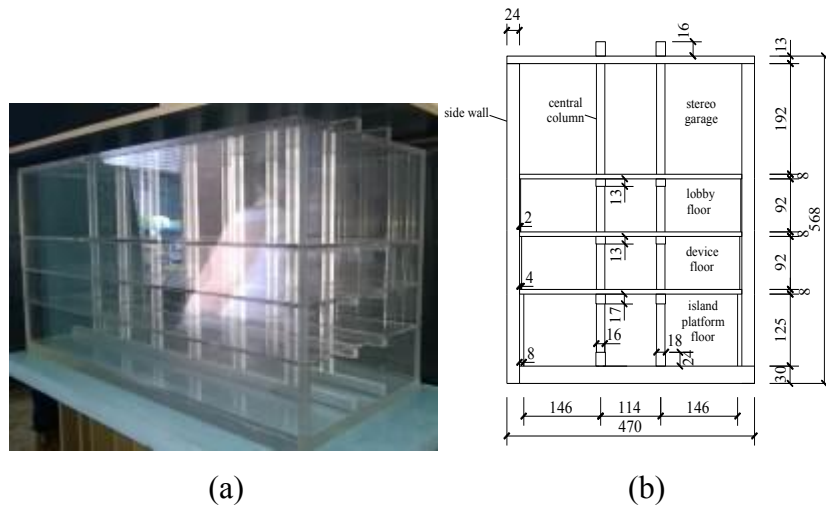




**Figure 1. Shaking table and soil container**

*Scale factor design and materials*

The prototype design of the model structure is a modern subway station with height of 28.3 m. The station was designed originally to be a six-story island platform station, and then because of the need for parking, the first to third floors underground were merged into one layer to function as a stereo garage. The second floor is the lobby floor, the third is a floor that houses equipment, and the fourth is an island platform. The total length of the station is 155 m, and the width varies from 23.6 to 28.35 m.



**Figure 2. (a) The model structure and (b) its dimensions**

On account of the differences in dimensions between a modern subway station and typical station, the scale factor design should be based on the size and bearing capacity of the shaking table, size of the soil container, boundary effect, and convenience of model manufacturing. The geometric scale factor is set to 1:50. Figure 2 shows the model structure and its dimensions. Scale factors among the physical quantities can be deduced using the Buckingham  $\pi$  law:

$$\begin{cases} S_{\sigma} = S_E \\ S_t = S_E^{-\frac{1}{2}} S_l S_{\rho}^{\frac{1}{2}} \\ S_v = S_E^{-\frac{1}{2}} S_{\rho}^{\frac{1}{2}} \\ S_a = S_E S_l S_{\rho}^1 \end{cases}, \quad (1)$$

where  $S_\sigma, S_E, S_t, S_l, S_v, S_a$  denote the stress scale factor, elastic modulus scale factor, time scale factor, geometric scale factor, velocity scale factor, and acceleration scale factor, respectively. Table 1 shows the scale factors of the model structure.

**Table 1. Scale factors of the model structure**

Type	Physical quantity	Scale factor
Material properties	Stress	0.106
	Strain	1.000
	Elastic modulus	0.106
	Poisson's ratio	1.000
	Equivalent density	1.765
Geometry properties	Length	0.020
	Linear displacement	0.020
	Angular displacement	1.000
Loading	Area	$4.00 \times 10^{-4}$
	Force	$4.24 \times 10^{-5}$
	Linear load	$2.12 \times 10^{-3}$
	Area load	0.106
	Moment	$8.48 \times 10^{-7}$
Dynamic properties	Mass	$1.41 \times 10^{-5}$
	Stiffness	$2.12 \times 10^{-3}$
	Duration	$8.16 \times 10^{-2}$
	Frequency	12.253
	Velocity	0.245
	Acceleration	3.003

Organic glass was chosen as the material of the model structure owing to its good homogeneity, high strength and low elastic modulus, providing flexibility to the design of the scale factor. This material is also suited to accurate manufacturing. The elastic modulus of three specimens were 3.60, 3.21, and 3.19 MPa, respectively. The average value was 3.33 MPa.

The synthetic model soil was a mixture of sand and sawdust. According to trial tests, adding sawdust to sand can reduce both the density and dynamic shear modulus, which complies with similitude requirements. Employing the Buckingham  $\pi$  law, the scale factors of geometry, density, shear modulus, and inertial acceleration were selected as essential parameters and adjusted to satisfy

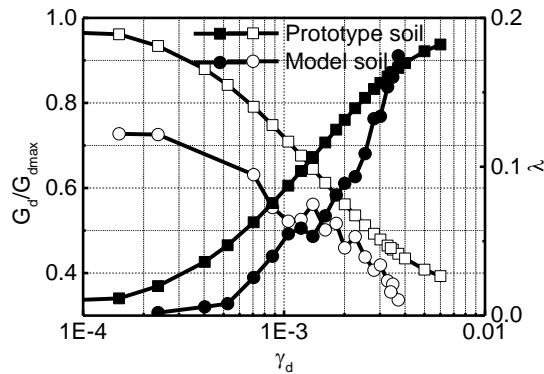
$$S_G / (S_l \cdot S_\rho) = S_a \quad (3)$$

where  $S_G, S_l, S_\rho, S_a$  denote the shear modulus ratio, geometry ratio, density ratio, and inertial acceleration ratio, respectively. The scale factors of soil are presented in Table 2.

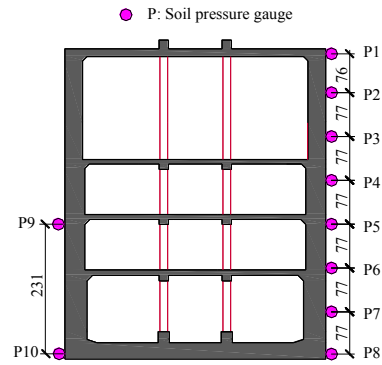
**Table 2. Scale factors of the model soil**

Physical quantity	Symbols	Scale factor
Shear modulus	$S_G$	0.020
Length	$S_L$	0.020
Density	$S_\rho$	0.333
Acceleration	$S_a$	3.003

To satisfy the scale factor equation and taking the maximum dynamic shear modulus,  $G_d / G_{d \max} - \gamma_d$  curve, and  $\lambda - \gamma_d$  curve into consideration, the most appropriate mass ratio of sawdust to sand was 1:2.5, where  $G_d, G_{d \max}, \lambda, \gamma_d$  denote the dynamic shear modulus, maximum dynamic shear modulus, damping ratio, and dynamic shear strain respectively. The density of the mixture was  $0.7 \text{ kg/m}^3$ , the density scale factor was 0.39, the confining pressure ratio was 0.02, and the modulus obtained in the test was 1.81 MPa. The  $G_d / G_{d \max} - \gamma_d$  curve and  $\lambda - \gamma_d$  curve obtained in a dynamic tri-axial test are presented in Figure 3.



**Figure 3. Dynamic properties of the model soil**



**Figure 4. Layouts of soil pressure meters**

#### *Layout of sensors*

The layouts of ten soil pressure gauges attached at the side wall of the model structure are shown in Figure 4. Soil pressure gauges P1–P8 were arranged to explore the distribution of the dynamic earth pressure and P9 and P10 were used to check the dynamic earth pressure.

#### *Test schema*

For the purpose of investigating the dynamic earth pressure under different intensities and types of ground motions, the three ground motions were scaled to two levels, 0.2 g and 0.6 g. Table 3 gives test cases.

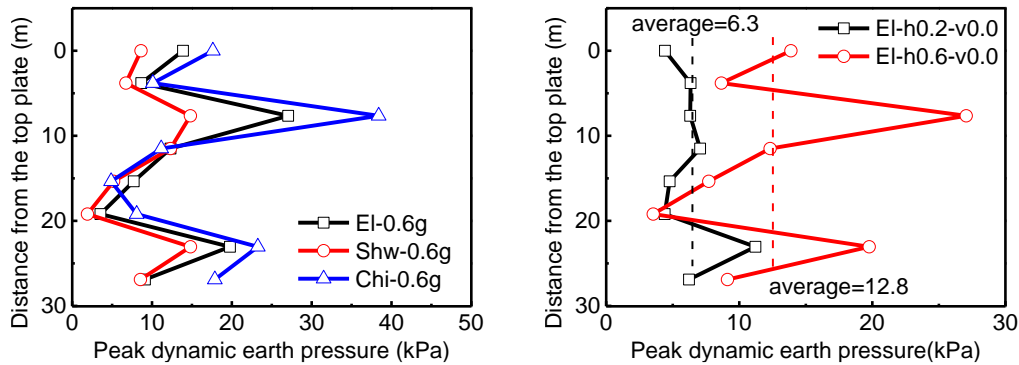
**Table 3. Test program**

Case	Test case	Ground motion	Horizontal peak acceleration (g)
1	El-0.2g	El Centro	0.2
2	El-0.6g	El Centro	0.6
3	Chi-0.6g	ChiChi	0.6
4	Shw-0.6g	Shanghai wave	0.6

#### *Results from shaking table tests*

For convenience, data obtained from tests were converted from model to prototype according to scale factors listed in Table 1. Figure 5(a) shows the peak dynamic lateral earth pressure along the side wall with different types of ground motions. It can be seen that, the amplitude of dynamic earth pressure closely related to ground

motions. Dynamic earth pressures amplitude under ChiChi ground motions were larger than the other. Specifically, the average peak dynamic earth pressure under ChiChi motion was 1.8 times the value under Shanghai wave. It is due to the pulse-like effect [Chen et al. (2015)]. But the distribution patterns were similar. Under three ground motions, the distribution of peak lateral dynamic earth pressure appeared an “S” shape, which was distinguished from that of a single-story underground structure. In the latter case, it is generally in a form of a triangle distribution. This may be because of the large height of the structure.



**Figure 5. Peak dynamic lateral earth pressure (a) with different types of ground motions; (b) with different peak horizontal accelerations**

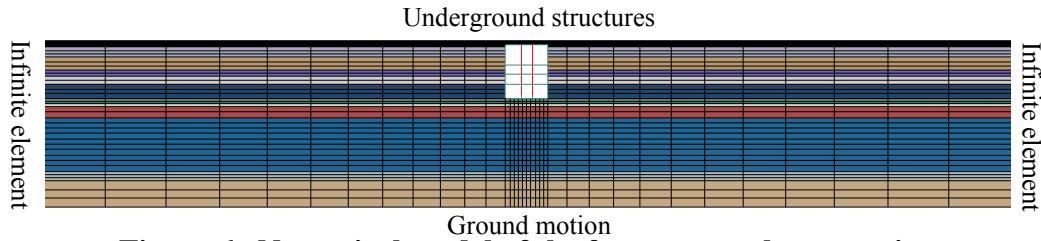
Figure 5(b) shows the influence of peak horizontal acceleration on the dynamic earth pressures. It is noted that the dash lines in the figure indicated the average values of pressures. It is found that with the increase of the magnitude of the earthquake, dynamic earth pressure increased remarkably. The average value of Case 3 was 2 times the value of Case 1. The influence of peak horizontal acceleration on pressure was different in each measurement point. But whether under a small earthquake or a large one, the distribution both appeared an “S” shape.

### Nonlinear dynamic time history analysis model

Due to the lacking of time and resources, it is difficult to identify many influences factors. For instance, the relative stiffness of soils to structures is an important factor in the soil-structure interaction, but during the test process it is hard to change overall soils or structures without disturbing the soils. Hence, experimental results were used to confirm the ability of the numerical technique to simulate the dynamic earth pressure. And numerical simulation can be used to study influences of the structural stiffness and types of surrounding soils on the distribution of seismic earth pressure.

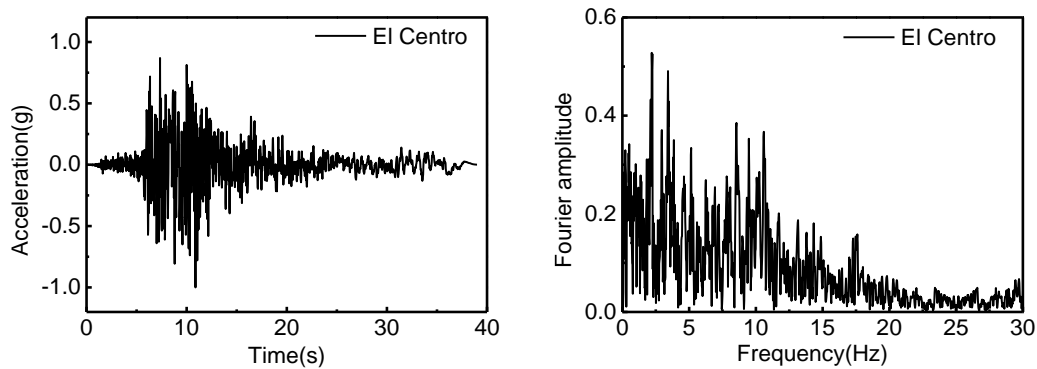
As shown in Figure 6, the plane strain model was chosen with dimensions of 1000 m long and 85 m high. In order to decrease the influence caused by seismic reflection, infinite element was imposed to the lateral boundaries. The horizontal and vertical displacements were fixed at the bottom. The structure was assumed to exhibit an elastic behavior throughout the entire analysis. Thus concrete in the structure was modeled as a linear elastic material, with unit weight 25kN/m<sup>3</sup>, Poisson’s ratio 0.15 and Young’s modulus 24GPa. Actual spacing of the column was taken into consideration with the reduced stiffness. The Mohr-Coulomb model was adopted to simulate soils. Soils were divided into 14 layers and soil properties were obtained from geotechnical investigations. 4-nodes plane strain element (CPE4R) and

quadrilateral plane strain infinite element (CINPE4) were adopted for soil, and beam element (B21) for structure. The interface between structure and ground was modeled as a frictional surface whose contact was assumed to follow the Coulomb friction law. A coefficient of friction equal to 0.4 was assumed which corresponds to a friction angle of 22 degrees.



**Figure 6. Numerical model of the four-story subway station**

El Centro record was used as the input ground motion, which was also the motion used in shaking table test. Figure 7 shows the motion and its Fourier spectrum.

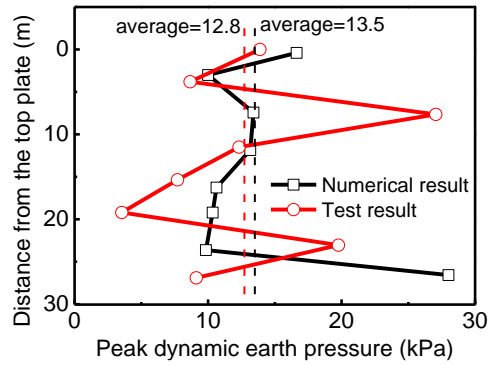


**Figure 7. Acceleration time history and Fourier spectrum of El Centro record**

## Parametric study

### *Comparison between numerical and test results*

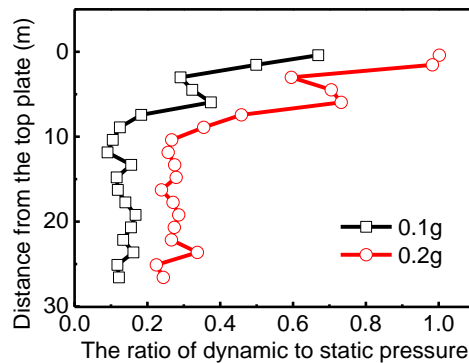
To verify the reliability and accuracy of the numerical model. A nonlinear dynamic time history analysis model was established in consistent with the dimensions of shaking table tests using the above-mentioned method. Figure 8 shows the comparison between numerical and test results in Case 3 (El-h0.6-v0.0). It can be seen that the distribution of dynamic earth pressure obtained from numerical results was also close to an “S” shape. The distributions were different in some degree. This may be due to the constitutive model of the model soil. The Mohr-Coulomb model is more appropriate for prototype soil rather than the synthetic model soil made of sand and sawdust. It is noted that the average dynamic pressure of numerical and test results were close. The error was 5.5%, respectively. Hence, the numerical model is able to investigate the amplitude of pressure in parametric studies.



**Figure 8. Comparison between numerical and test results in Case 3**

*The ratio of dynamic earth pressure to static pressure*

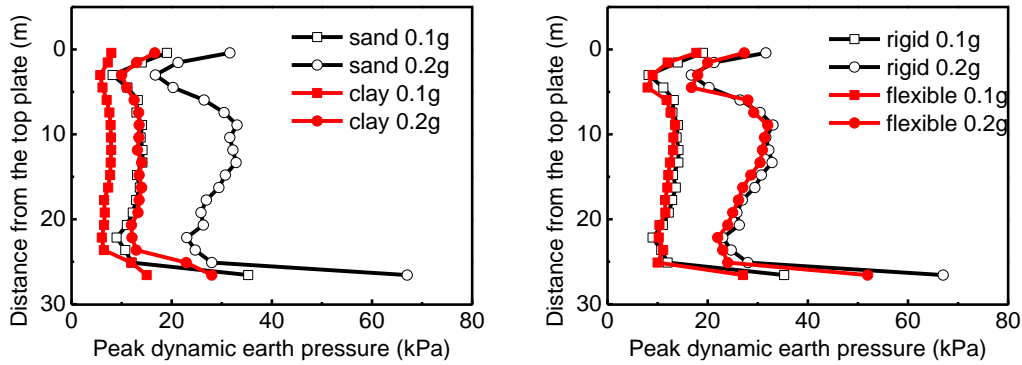
Figure 9 shows the ratio of dynamic to static pressure under two levels of the earthquake. Firstly, it can be seen that with the increase of peak acceleration, dynamic earth pressure increased notably. Then the ratio at the top of the structure was much larger than that at the bottom. It is because the static earth pressure at top was much smaller than that at the bottom. It is worth noting that at the lower part of the structure, the degree of increase was close. Specifically, ratios under two levels were about 0.15 and 0.27, respectively. It means that when the peak horizontal acceleration was 0.2 g, the lateral earth pressure increased nearly thirty percent, which needed attention in practice.



**Figure 9. The ratio of dynamic earth pressure to static earth pressure**

*Soil types*

In order to study the influence of soil type, the second layer to the eighth layer of prototype soils were combined as one homogeneous soil layer. And this layer was defined as sand layer and clay layer in two models, respectively. Frictional angle and cohesion of clay were 16 degrees and 17 kPa, and values of sand were 35 degrees and 0. Figure 10(a) shows the influence of soil type under two levels of the earthquake. It is seen that when the earthquake was small, distributions of dynamic earth pressure were close to linear one in most parts of the structure; when the earthquake was large, distributions in the sand tend to be the “S” shape. Structures surrounded by sand suffered larger dynamic earth pressure than that by clay. Specifically, when peak horizontal acceleration was 0.2 g in clay, dynamic earth pressure was very close to that in the sand when peak acceleration was 0.1 g.



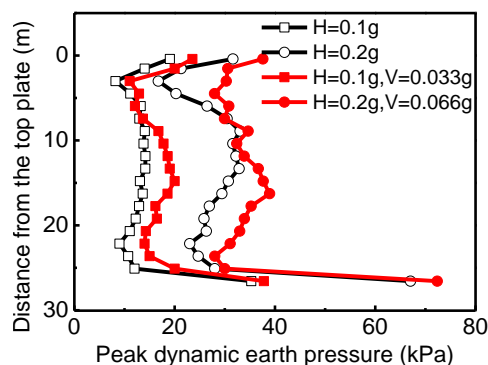
**Figure 10. Influences of (a) soil type and (b) stiffness of structure**

#### *Stiffness of the structure*

The diaphragm wall is 1.2 m wide. Dimensions of the inside wall vary from 0.4 m to 0.8 m (from top to bottom). The original structure with diaphragm wall was defined as a rigid structure while the structure without diaphragm wall was regarded as a flexible one. The lateral stiffness of the wall consist of the diaphragm wall and the inside wall on the bottom story was about 16 times the stiffness of the wall of the flexible one. Figure 10 shows the difference between the rigid structure and the flexible structure. It is seen that peak dynamic earth pressure of the flexible one was a little smaller than that of the rigid one whether under a small or large earthquake. But the difference was small. Properties of soils had more influences on dynamic earth pressure than stiffness of the structure.

#### *Vertical earthquake*

To study influence of vertical earthquake component on the dynamic earth pressure, peak vertical acceleration was scaled to be 2/3 of peak horizontal acceleration. Figure 11 presents the peak dynamic earth pressure under two levels of the earthquake with and without vertical component. Specifically, when peak horizontal acceleration was 0.1 g, average values of peak dynamic earth pressure with and without vertical component were 13.93 kPa and 17.77 kPa, respectively. Values were 29.39 kPa and 34.99 kPa when peak horizontal acceleration was 0.2 g. Hence, vertical earthquake component excited lateral dynamic earth pressure.



**Figure 11. Influences of vertical earthquake component**

### **Conclusions**

In this paper, to study distributions and amplitudes of dynamic earth pressure along the side wall of an underground subway station, a series of 1-g shaking table tests

were conducted on a four-story subway station firstly. Experimental results were used to confirm the reliability of the numerical technique. Then parametric study was carried out through nonlinear dynamic time history analyses using the general purpose finite element code ABAQUS. The following conclusions are drawn from the results of the study.

(1) The distribution of peak lateral dynamic earth pressure appeared an “S” shape, which was distinguished from that of a single-story underground structure. Different type of ground motions had influences on amplitudes of pressure, but distribution shapes were similar. And with the increase of the magnitude of the earthquake, dynamic earth pressure increased remarkably

(2) At the lower part of the structure, ratios of dynamic to static earth pressure were similar. Specifically, when the peak horizontal acceleration was 0.2 g, lateral earth pressure on the lower part of the structure increased nearly thirty percent, which needed attention in practice. And vertical earthquake component excited lateral dynamic earth pressure.

(3) Properties of soils had more influences on dynamic earth pressure than stiffness of the structure. Structures surrounded by sand suffered larger dynamic earth pressure than that by clay. Peak dynamic earth pressure of the flexible structure was a little smaller than that of the rigid one whether under a small or large earthquake, but the difference was small.

### Acknowledgements

This research was supported by the State Key Laboratory of Disaster Reduction in Civil Engineering (SLDRCE14-B-11), National Natural Science Foundation of China (Grant No. 51278524), and Innovation Program of Shanghai Municipal Education Commission (14ZZ034). All support is gratefully acknowledged.

### Reference

- Chen, Z. Y., Chen, W., Li, Y. Y., and Yuan, Y. (2015) Seismic performance of central columns for a multi-story subway station under vertical earthquake excitations, *Engineering structures* (submitted)
- Gazetas, G., Psarropoulos, P., Anastasopoulos, I., and Gerolymos, N. (2004) Seismic behaviour of flexible retaining systems subjected to short-duration moderately strong excitation, *Soil Dynamics and Earthquake Engineering* **24**, 537–550.
- Luu, A. L. (2013) Seismic earth pressures measured during a shake table experiment of underground structures. University of California, Irvine.
- Madabhushi, S., and Zeng, X. (2007) Simulating seismic response of cantilever retaining walls, *Journal of geotechnical and geoenvironmental engineering* **133**, 539–549.
- Ostadan, F. (2005) Seismic soil pressure for building walls: An updated approach, *Soil Dynamics and Earthquake Engineering* **25**, 785–793.
- Psarropoulos, P., Klonaris, G., and Gazetas, G. (2005) Seismic earth pressures on rigid and flexible retaining walls, *Soil Dynamics and Earthquake Engineering* **25**, 795–809.
- Sun C. F. (2013) Approving of numbers of urban railway projects which reach four trillion by 2020, *Sichuan Cement* **07**, 72–73 (In Chinese)
- Wang, W. P., Tao, L. J., Zhang, B., Li, W. B., and Wei, Y. J. (2010) A New Method for Seismic Active Pressure Calculation of the Underground Structure, *Journal of Disaster Prevention and Mitigation Engineering* **30**, 620–623.
- Yang, L. D., Yang, C., and Ji, Q. Q. (2003) Shaking table test and numerical calculation on subway station structures in soft soil, *Journal of Tongji University* **31**, 1135–1140.



## Stress analysis of functionally graded plates with different gradient direction

†\*M. Amirpour<sup>1</sup>, R. Das<sup>1</sup>, and E.I. Saavedra Flores<sup>2</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Auckland, Auckland 1010, New Zealand

<sup>2</sup>Departamento de Ingeniería en Obras Civiles, Universidad de Santiago de Chile, Av. Ecuador 3659, Santiago, Chile

\*Presenting author: maedeh.amirpournolla@auckland.ac.nz

†Corresponding author: maedeh.amirpournolla@auckland.ac.nz

### Abstract

The purpose of this paper is to investigate the general deformation pattern and stress field of a thin rectangular FGM plate for different property gradient directions, i.e. perpendicular and parallel to the loading direction, both analytically and numerically. The relevant governing equations of elasticity are solved with static analysis with power law distribution of volume fraction of constituents, and the analytical solutions for the displacements and stresses are derived. The resultant solutions are verified against numerical solutions obtained using the finite element method (FEM). The finite element (FE) solution is obtained using solid elements with spatially graded property distribution (at different gauss points), which is implemented by a user material subroutine (UMAT) in the ABAQUS FE software. The obtained results demonstrate that the direction of material property gradient and the nature of its variation have significant effects on the mechanical behavior of FGM plates. Moreover, the comparison between the exact solution and numerical simulation shows the efficiency of graded solid elements in modelling of thin FGM plate. **Keywords: FGM, stress field, power-law distribution, finite element method, elasticity, graded solid elements.**

### Introduction

Functionally Graded Materials (FGMs) are advanced engineered materials whereby material composition and properties vary spatially in macroscopic length scales, which are created by specialized manufacturing processes. The main advantage of FGMs is the elimination of stress concentration and discontinuity in the interface due to the monotonous variation of volume fraction of the constituents. There are various mathematical models to describe distribution of volume fraction of constituents, i.e. power-law and exponential law. Some researchers used exponential function for defining material property variation (Guo & Noda, 2014; Z. Wang, Guo, & Zhang, 2013). Another form of mathematical model using a power-law distribution has been widely used in a number of studies, especially for the mechanical engineering field (Cheng & Batra, 2000; Navazi, Haddadpour, & Rasekh, 2006; Sun & Luo, 2011). There are a significant growth in literatures corresponding to FGMs in the mechanics of FGMs (Chi & Chung, 2006; Thai & Choi, 2013; C. Wang & Xu, 2014), manufacture process (El-Desouky, Kassegneb, Moonb, McKittrickc, & Morsib, 2013; Kieback, Neubrand, & Riedel, 2003), crack growth and damage (Bocciarelli, Bolzon, & Maier, 2008; Eghtesad, Shafiei, & Mahzoon, 2012; Torshizian & Kargarnovin, 2014) in last 10 years.

One of the wide applications of FGM is in plate structures as thermal barriers. Therefore, understanding the mechanical behavior of an FGM plate is vital for

effective design of structures employing FGMs to meet desired and safety criteria. An exact solution for exponentially graded FGM plates with simply-supported boundary condition under a surface load was expanded by Pan (Pan, 2003). Further, Chi (Chi & Chung, 2006) derived an analytical formulation for three types of distribution function namely power-law, sigmoid and exponential based on classical plate theory (CPT) for rectangle simply supported FGM plate with transverse loading. They evaluated the exact solution with numerical simulation by MARC FE program with 16 layers of different material properties in thickness direction. Recently, (Akbarzadeh, Hosseini zad, Eslami, & Sadighi, 2010) used the first-order shear deformation theory (FSDT) and the third-order shear deformation theory (TSDT) for FGM plate with the power-law distribution of the volume fraction. They obtained the natural frequencies and dynamic responses of the FGM plate analytically. All the mentioned studies have been carried out with regards to variation of material constituents through the thickness. However, no studies are known on distribution of volume fraction through the length of the plate.

During manufacturing process of FGMs, the reliability requirements for the product should be considered to meet desired or application-specific performance criteria. One approach to produce FGMs is use of additive manufacturing (3D printing), which can control local composition and microstructure. Furthermore, the gradient distribution and its relationship with the loading direction will affect the macro stiffness and mechanical behavior.

This paper presents the behavior of the thin rectangular FGM plate with CPT assumptions under transverse loading. The close forms solutions based on Fourier series expression for power-law distributions in two different gradient directions (the thickness direction and length direction) are obtained. Moreover, the analytical solutions are proved by the numerical simulation of the finite element method using ABAQUS. In numerical modelling, the graded solid elements are implemented in the user material subroutine (UMAT). These results will possibly enable us to understand the behavior of new materials with controlled macro properties.

## Analytical Solutions

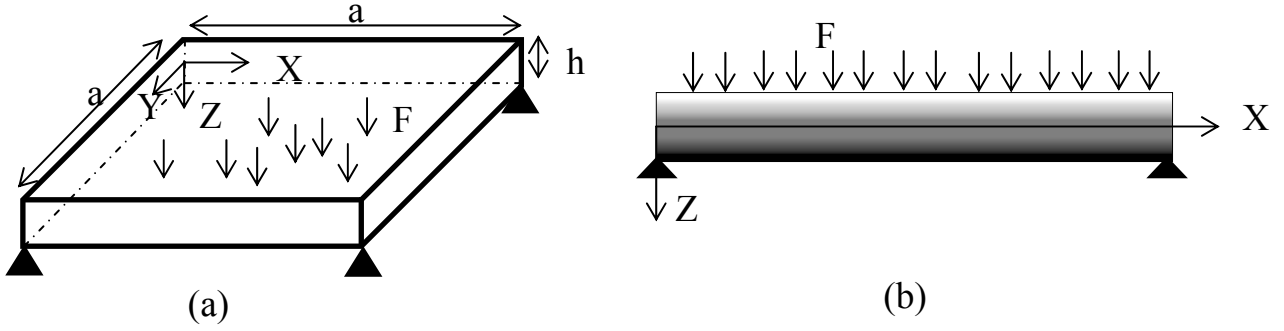
### State I- Gradient distribution in thickness direction

Let us consider an elastic rectangular plate. As shown in Fig. 1a, coordinates  $x$  and  $y$  define the plane of the plate, whereas the  $z$ -axis originated at the middle surface of the plate is in the thickness direction. The Poisson ratio is assumed to be constant. However, the Young's moduli in thickness direction vary with Eq. (1) and the elastic modulus can be determined by the rule of mixture:

$$V_i(z) = \left(\frac{1}{2} + \frac{z}{h}\right)^n \quad (1)$$

$$E(z) = V_i(z)E_i + (1 - V_i(z))E_m = \left(\frac{1}{2} + \frac{z}{h}\right)^n (E_i - E_m) + E_m \quad (2)$$

Where  $n$  is the power-law index,  $h$  is the thickness of the plate,  $E_i$  and  $E_m$  are the Young's moduli of inclusion and matrix, respectively.



**Figure 1. a) The Geometry of FGM Plate. b) The distribution of volume fraction in Z**

**direction :**  $V_i(z) = \left(\frac{1}{2} + \frac{z}{h}\right)^n$

It is assumed that the deformations and the stresses of the thin FGM plate are based upon classical plate theory. So, the strain and stress fields are:

$$\varepsilon_x = \varepsilon_{x0} - z \frac{\partial^2 w}{\partial x^2} \quad (3a)$$

$$\varepsilon_y = \varepsilon_{y0} - z \frac{\partial^2 w}{\partial y^2} \quad (3b)$$

$$\gamma_{xy} = \gamma_{xy0} - 2z \frac{\partial^2 w}{\partial x \partial y} \quad (3c)$$

$$\sigma_x = \frac{E(z)}{1-\nu^2} \left[ \varepsilon_{x0} + \nu \varepsilon_{y0} - z \frac{\partial^2 w}{\partial x^2} - \nu z \frac{\partial^2 w}{\partial y^2} \right] \quad (3d)$$

$$\sigma_y = \frac{E(z)}{1-\nu^2} \left[ \varepsilon_{y0} + \nu \varepsilon_{x0} - z \frac{\partial^2 w}{\partial y^2} - \nu z \frac{\partial^2 w}{\partial x^2} \right] \quad (3e)$$

$$\tau_{xy} = \frac{E(z)}{1-\nu^2} \frac{1-\nu}{2} \left[ \gamma_{xy0} - 2z \frac{\partial^2 w}{\partial x \partial y} \right] \quad (3f)$$

In this work, the thickness of thin plate is assumed to be in the range  $1/20 \sim 1/100$  of its length. So, the transverse shear deformations should be eliminated.

The matrix formation of axial forces and momentum are:

$$\begin{pmatrix} N_x \\ N_y \\ N_{xy} \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{12} & A_{11} & 0 \\ 0 & 0 & A_{66} \end{bmatrix} \begin{pmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{pmatrix} + \begin{bmatrix} B_{11} & B_{12} & 0 \\ B_{12} & B_{11} & 0 \\ 0 & 0 & B_{66} \end{bmatrix} \begin{pmatrix} -\frac{\partial^2 w}{\partial x^2} \\ -\frac{\partial^2 w}{\partial y^2} \\ -2\frac{\partial^2 w}{\partial x \partial y} \end{pmatrix} \quad (4a)$$

$$\begin{pmatrix} M_x \\ M_y \\ M_{xy} \end{pmatrix} = \begin{bmatrix} B_{11} & B_{12} & 0 \\ B_{12} & B_{11} & 0 \\ 0 & 0 & B_{66} \end{bmatrix} \begin{pmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{pmatrix} + \begin{bmatrix} C_{11} & C_{12} & 0 \\ C_{12} & C_{11} & 0 \\ 0 & 0 & C_{66} \end{bmatrix} \begin{pmatrix} -\frac{\partial^2 w}{\partial x^2} \\ -\frac{\partial^2 w}{\partial y^2} \\ -2\frac{\partial^2 w}{\partial x \partial y} \end{pmatrix} \quad (4b)$$

The coefficients of above equations are depending on the material properties of FGM plate. So, for the above rectangular FGM plate we have:

$$A_{11} = \int_{-h/2}^{h/2} \frac{E(z)}{1-\nu^2} dz = \int_{-h/2}^{h/2} \frac{\left(\frac{1}{2} + \frac{z}{h}\right)^n (E_i - E_m) + E_m}{1-\nu^2} dz = \frac{h}{1-\nu^2} \frac{(nE_m + E_i)}{n+1}$$

$$A_{12} = \nu A_{11} = \frac{\nu h}{1-\nu^2} \frac{(nE_m + E_i)}{n+1} \quad (5a)$$

$$A_{66} = \frac{1-\nu}{2} A_{11} = \frac{1-\nu}{2} \frac{h}{1-\nu^2} \frac{(nE_m + E_i)}{n+1}$$

$$B_{11} = \int_{-h/2}^{h/2} \frac{zE(z)}{1-\nu^2} dz = \frac{1}{1-\nu^2} \int_{-h/2}^{h/2} z \left(\frac{1}{2} + \frac{z}{h}\right)^n (E_i - E_m) + zE_m dz = \frac{h^2}{1-\nu^2} \frac{n(E_i - E_m)}{2(n+1)(n+2)}$$

$$B_{12} = \nu B_{11} = \frac{\nu h^2}{1-\nu^2} \frac{n(E_i - E_m)}{2(n+1)(n+2)} \quad (5b)$$

$$B_{66} = \frac{1-\nu}{2} B_{11} = \frac{1-\nu}{2} \frac{h^2}{1-\nu^2} \frac{n(E_i - E_m)}{2(n+1)(n+2)}$$

$$C_{11} = \int_{-h/2}^{h/2} \frac{z^2 E(z)}{1-\nu^2} dz = \frac{1}{1-\nu^2} \int_{-h/2}^{h/2} z^2 \left(\frac{1}{2} + \frac{z}{h}\right)^n (E_i - E_m) + E_m dz =$$

$$\frac{h^3}{12(1-\nu^2)} \left( \frac{3(n^2 + n + 2)(E_i - E_m)}{(n+1)(n+2)(n+3)} + E_m \right) \quad (5c)$$

$$C_{12} = \nu C_{11} = \frac{\nu h^3}{12(1-\nu^2)} \left( \frac{3(n^2 + n + 2)(E_i - E_m)}{(n+1)(n+2)(n+3)} + E_m \right)$$

$$C_{66} = \frac{1-\nu}{2} C_{11} = \frac{1-\nu}{2} \frac{h^3}{12(1-\nu^2)} \left( \frac{3(n^2 + n + 2)(E_i - E_m)}{(n+1)(n+2)(n+3)} + E_m \right)$$

With definition of the uniform distributed transverse loading by Fourier series, the equilibrium equation can be written as:

$$\frac{\partial^2 M_x}{\partial x^2} + \frac{\partial^2 M_y}{\partial y^2} + 2 \frac{\partial^2 M_{xy}}{\partial x \partial y} = -F = -\sum \sum F_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \quad (6)$$

$$F_{mn} = \begin{cases} \frac{16F}{\pi^2 mn} & n, m = 1, 3, 5, \dots \\ 0 & n, m = 2, 4, 6, \dots \end{cases}$$

With definition of stress function  $\Phi(x, y)$  and  $N_x = \frac{\partial^2 \Phi}{\partial y^2}$ ,  $N_y = \frac{\partial^2 \Phi}{\partial x^2}$ ,  $N_{xy} = \frac{\partial^2 \Phi}{\partial x \partial y}$ , the strains at the middle surface are expressed in terms of the stress function and the deflection:

$$\begin{pmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{pmatrix} = \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{12} & D_{11} & 0 \\ 0 & 0 & D_{66} \end{bmatrix} \begin{pmatrix} \frac{\partial^2 \Phi}{\partial y^2} \\ \frac{\partial^2 \Phi}{\partial x^2} \\ \frac{\partial^2 \Phi}{\partial x \partial y} \end{pmatrix} + \begin{bmatrix} E_{11} & E_{12} & 0 \\ E_{12} & E_{11} & 0 \\ 0 & 0 & E_{66} \end{bmatrix} \begin{pmatrix} -\frac{\partial^2 w}{\partial x^2} \\ -\frac{\partial^2 w}{\partial y^2} \\ -2\frac{\partial^2 w}{\partial x \partial y} \end{pmatrix} \quad (7a)$$

$$\begin{pmatrix} M_x \\ M_y \\ M_{xy} \end{pmatrix} = - \begin{bmatrix} E_{11} & E_{12} & 0 \\ E_{12} & E_{11} & 0 \\ 0 & 0 & E_{66} \end{bmatrix} \begin{pmatrix} \varepsilon_{x0} \\ \varepsilon_{y0} \\ \gamma_{xy0} \end{pmatrix} + \begin{bmatrix} F_{11} & F_{12} & 0 \\ F_{12} & F_{11} & 0 \\ 0 & 0 & F_{66} \end{bmatrix} \begin{pmatrix} -\frac{\partial^2 w}{\partial x^2} \\ -\frac{\partial^2 w}{\partial y^2} \\ -2\frac{\partial^2 w}{\partial x \partial y} \end{pmatrix} \quad (7b)$$

where:

$$D_{11} = \frac{A_{11}}{A_{11}^2 - A_{12}^2} = \frac{1}{1 - \nu^2 A_{11}} = \frac{n+1}{h(nE_m + E_i)}$$

$$D_{12} = \frac{-A_{12}}{A_{11}^2 - A_{12}^2} = \frac{-\nu A_{11}}{A_{11}^2 - \nu^2 A_{11}^2} = \frac{-\nu(n+1)}{h(nE_m + E_i)} \quad (8a)$$

$$D_{66} = \frac{-1}{A_{66}} = -\frac{2}{1-\nu} \frac{1-\nu^2}{h} \frac{(n+1)}{(nE_m + E_i)}$$

$$E_{11} = \frac{A_{12}B_{12} - A_{11}B_{11}}{A_{11}^2 - A_{12}^2} = \frac{\nu^2 A_{11}B_{11} - A_{11}B_{11}}{A_{11}^2 - \nu^2 A_{11}^2} = -\frac{hn(E_i - E_m)}{2(n+2)(nE_m + E_i)}$$

$$E_{12} = \frac{A_{12}B_{11} - A_{11}B_{12}}{A_{11}^2 - A_{12}^2} = 0 \quad (8b)$$

$$E_{66} = -\frac{B_{66}}{A_{66}} = -\frac{B_{11}}{A_{11}} = -\frac{hn(E_i - E_m)}{2(n+2)(nE_m + E_i)}$$

$$F_{11} = B_{11}E_{11} + B_{12}E_{12} + C_{11} = \frac{h^2}{1-\nu^2} \left( \frac{n(E_i - E_m)}{2(n+1)(n+2)} \right) \left( -\frac{hn(E_i - E_m)}{2(n+2)(nE_m + E_i)} \right) +$$

$$\frac{h^3}{12(1-\nu^2)} \left( \frac{3(n^2 + n + 2)(E_i - E_m)}{(n+1)(n+2)(n+3)} + E_m \right) \quad (8c)$$

$$F_{12} = \nu F_{11}$$

$$F_{66} = C_{66} + B_{66}E_{66}$$

In order to find the relation between stress function and deflection, we substitute the Eqs. (7a) and (7b) in Eq. (6):

$$E_{12} \frac{\partial^4 \Phi}{\partial x^4} + E_{12} \frac{\partial^4 \Phi}{\partial y^4} + 2(E_{11} - E_{66}) \frac{\partial^4 \Phi}{\partial x^2 \partial y^2} + F_{11} \frac{\partial^4 w}{\partial x^4} + F_{11} \frac{\partial^4 w}{\partial y^4} + 2(F_{12} + 2F_{66}) \frac{\partial^4 w}{\partial x^2 \partial y^2} =$$

$$\sum \sum F_{mn} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \quad (9)$$

A compatibility equation is then used to provide another governing equation. By using Eq. (7a), the compatibility equation

$$\frac{\partial^2 \varepsilon_x}{\partial y^2} + \frac{\partial^2 \varepsilon_y}{\partial x^2} = \frac{\partial^2 \gamma_{xy}}{\partial y \partial x} \quad (10)$$

Can be expressed as:

$$D_{11} \frac{\partial^4 \Phi}{\partial x^4} + D_{11} \frac{\partial^4 \Phi}{\partial y^4} + 2(D_{12} - D_{66}) \frac{\partial^4 \Phi}{\partial x^2 \partial y^2} - E_{12} \frac{\partial^4 w}{\partial x^4} - E_{12} \frac{\partial^4 w}{\partial y^4} - 2(E_{11} - E_{66}) \frac{\partial^4 w}{\partial x^2 \partial y^2} = 0 \quad (11)$$

The stress function and deflection can be obtained by simultaneous solution of Eqs. (9) and (11). With definition of Fourier series for  $\Phi(x, y)$  and  $w$  alongside

simultaneous solution of Eqs. (9) and (11), the stress function and deflection can be obtained.

$$\begin{aligned}
 w(x, y) &= \sum \sum w_{mn} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \\
 \Phi(x, y) &= \sum \sum \Phi_{mn} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a}
 \end{aligned} \tag{12}$$

where

$$w_{mn} = \left( \frac{1}{F_{11} \left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \right) F_{mn}$$

$$\Phi_{mn} = 0$$

So, the stress and strain fields for the rectangular FGM plate with the material gradation in thickness direction are:

$$\varepsilon_x = \frac{E_{11} + z}{F_{11}} \sum \sum \frac{F_{mn} \left( \frac{m\pi}{a} \right)^2}{\left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \tag{13a}$$

$$\varepsilon_y = \frac{E_{11} + z}{F_{11}} \sum \sum \frac{F_{mn} \left( \frac{n\pi}{a} \right)^2}{\left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \tag{13b}$$

$$\gamma_{xy} = \frac{-2(E_{11} + z)}{F_{11}} \sum \sum \frac{F_{mn} \left( \frac{n\pi}{a} \right) \left( \frac{m\pi}{a} \right)}{\left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \cos \frac{m\pi x}{a} \cos \frac{m\pi y}{a} \tag{13c}$$

$$\sigma_x = \frac{E(z)}{1-\nu^2} \frac{E_{11} + z}{F_{11}} \sum \sum F_{mn} \frac{\left( \frac{m\pi}{a} \right)^2 + \nu \left( \frac{n\pi}{a} \right)^2}{\left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \tag{13d}$$

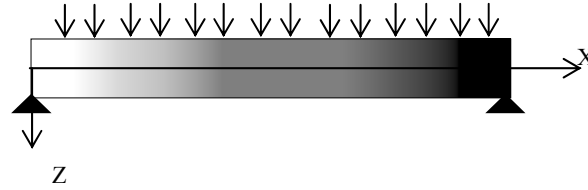
$$\sigma_y = \frac{E(z)}{1-\nu^2} \frac{E_{11} + z}{F_{11}} \sum \sum F_{mn} \frac{\left( \frac{n\pi}{a} \right)^2 + \nu \left( \frac{m\pi}{a} \right)^2}{\left( \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right)^2} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \tag{13e}$$

$$\tau_{xy} = \frac{-E(z)}{1+\nu} \frac{E_{11} + z}{F_{11}} \sum \sum F_{mn} \frac{\left(\frac{m\pi}{a}\right)\left(\frac{n\pi}{a}\right)}{\left(\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{a}\right)^2\right)^2} \cos \frac{m\pi x}{a} \cos \frac{m\pi y}{a} \quad (13f)$$

### State II- Gradient distribution in length direction

Consider the rectangular FGM plate-Fig. 2- similar to state 1 while the Young's moduli vary in length direction based on Eq. (14).

$$E(x) = \left(\frac{x}{a}\right)^n (E_i - E_m) + E_m \quad (14)$$



**Figure 2. The distribution of volume fraction X direction :  $V_i(x) = \left(\frac{x}{a}\right)^n$**

For this case, the coefficients become:

$$A_{11} = \int_{-h/2}^{h/2} \frac{E(x)}{1-\nu^2} dz = \int_{-h/2}^{h/2} \frac{\left(\frac{x}{a}\right)^n (E_i - E_m) + E_m}{1-\nu^2} dz = \frac{h}{1-\nu^2} E(x)$$

$$A_{12} = \nu A_{11} = \frac{\nu h}{1-\nu^2} E(x) \quad (15a)$$

$$A_{66} = \frac{1-\nu}{2} A_{11} = \frac{1-\nu}{2} \frac{h}{1-\nu^2} E(x)$$

$$B_{11} = \int_{-h/2}^{h/2} \frac{zE(x)}{1-\nu^2} dz = \frac{E(x)}{1-\nu^2} \int_{-h/2}^{h/2} z dz = 0$$

$$B_{12} = \nu B_{11} = 0$$

$$B_{66} = \frac{1-\nu}{2} B_{11} = 0 \quad (15b)$$

$$C_{11} = \int_{-h/2}^{h/2} \frac{z^2 E(x)}{1-\nu^2} dz = \frac{E(x)}{1-\nu^2} \int_{-h/2}^{h/2} z^2 dz = \frac{h^3}{12(1-\nu^2)} E(x)$$

$$C_{12} = \nu C_{11} = \frac{\nu h^3}{12(1-\nu^2)} E(x) \quad (15c)$$

$$C_{66} = \frac{1-\nu}{2} C_{11} = \frac{1-\nu}{2} \frac{h^3}{12(1-\nu^2)} E(x)$$

It can be seen that the axial forces and bending momentum are uncoupled. This phenomenon is different from the state 1. The equilibrium Eq. (6) becomes:

$$C_{11} \left[ \frac{\partial^4 w}{\partial x^4} + \frac{\partial^4 w}{\partial y^4} + 2 \frac{\partial^4 w}{\partial x^2 \partial y^2} \right] + \frac{\partial C_{11}}{\partial x} \left[ 2 \frac{\partial^3 w}{\partial x \partial y^2} + 2 \frac{\partial^3 w}{\partial x^3} \right] + \frac{\partial^2 C_{11}}{\partial x^2} \left[ \frac{\partial^2 w}{\partial x^2} + \nu \frac{\partial^2 w}{\partial y^2} \right] = F \quad (16)$$

In order to satisfy the equilibrium equation and boundary conditions, the deflection function  $w$  of FGM plate should be the form of:

$$w(x, y) = \sum \sum w_{mn} \sin \frac{m\pi x}{a} \sin \frac{m\pi y}{a} \quad (17)$$

By substituting in the equilibrium equation we can find  $w_{mn}$  :

$$\left\{ C_{11} \left[ \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right]^2 + \frac{\partial^2 C_{11}}{\partial x^2} \left[ \left( \frac{m\pi}{a} \right)^2 + \nu \left( \frac{n\pi}{a} \right)^2 \right] \right\} w_{mn} = F_0$$

$$w_{mn} = \frac{F_0}{C_{11} \left[ \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right]^2 + \frac{\partial^2 C_{11}}{\partial x^2} \left[ \left( \frac{m\pi}{a} \right)^2 + \nu \left( \frac{n\pi}{a} \right)^2 \right]}$$

So, the stress and strain become:

$$\varepsilon_x = -z \frac{\partial^2 w}{\partial x^2} \quad (18a)$$

$$\varepsilon_y = -z \frac{\partial^2 w}{\partial y^2} = z \sum \sum \frac{\left( \frac{n\pi}{a} \right)^2 F_0}{C_{11} \left[ \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right]^2 + \frac{\partial^2 C_{11}}{\partial x^2} \left[ \left( \frac{m\pi}{a} \right)^2 + \nu \left( \frac{n\pi}{a} \right)^2 \right]} \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \quad (18b)$$

$$\gamma_{xy} = -2z \sum \sum \frac{\left( \frac{m\pi}{a} \right) \left( \frac{n\pi}{a} \right) F_0}{C_{11} \left[ \left( \frac{m\pi}{a} \right)^2 + \left( \frac{n\pi}{a} \right)^2 \right]^2 + \frac{\partial^2 C_{11}}{\partial x^2} \left[ \left( \frac{m\pi}{a} \right)^2 + \nu \left( \frac{n\pi}{a} \right)^2 \right]} \cos \frac{m\pi x}{a} \cos \frac{n\pi y}{a} \quad (18c)$$

where

$$\begin{aligned} \frac{\partial^2 w}{\partial x^2} &= \sum \sum \left( \lambda^{(2)} - \lambda \left( \frac{m\pi}{a} \right)^2 \right) \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} + 2\lambda^{(1)} \left( \frac{m\pi}{a} \right) \cos \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \\ \lambda &= \frac{16F_0}{\pi^2} \cdot \frac{1}{C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2)} \\ \lambda^{(1)} &= -\frac{16F_0}{\pi^2} \cdot \frac{\frac{\partial C_{11}}{\partial x} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^3 C_{11}}{\partial x^3} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2)}{\left( C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)^2} \\ \lambda^{(2)} &= -\frac{16F_0}{\pi^2} \cdot \left\{ \frac{\left( \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^4 C_{11}}{\partial x^4} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right) \cdot \left( C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)^2}{\left( C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)^4} \right. \\ &\quad \left. - \frac{2 \left( \frac{\partial C_{11}}{\partial x} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^3 C_{11}}{\partial x^3} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)^2 \cdot \left( C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)}{\left( C_{11} m n \frac{\pi^4}{a^4} (m^2 + n^2)^2 + \frac{\partial^2 C_{11}}{\partial x^2} m n \frac{\pi^2}{a^2} (m^2 + \nu n^2) \right)^4} \right\} \quad (19) \end{aligned}$$

$$\sigma_x = \frac{E(x)}{1-\nu^2} [\varepsilon_x + \nu \varepsilon_y] \quad (20a)$$

$$\sigma_y = \frac{E(x)}{1-\nu^2} [\varepsilon_y + \nu \varepsilon_x] \quad (20b)$$

$$\tau_{xy} = \frac{E(x)}{1-\nu^2} \frac{1-\nu}{2} \gamma_{xy} \quad (20c)$$



**Finite Element Models**

In the numerical simulation, a square FGM plate with  $a=50$  cm,  $h=2$  cm and material properties of metal and ceramic constituents-Table 1- is considered.

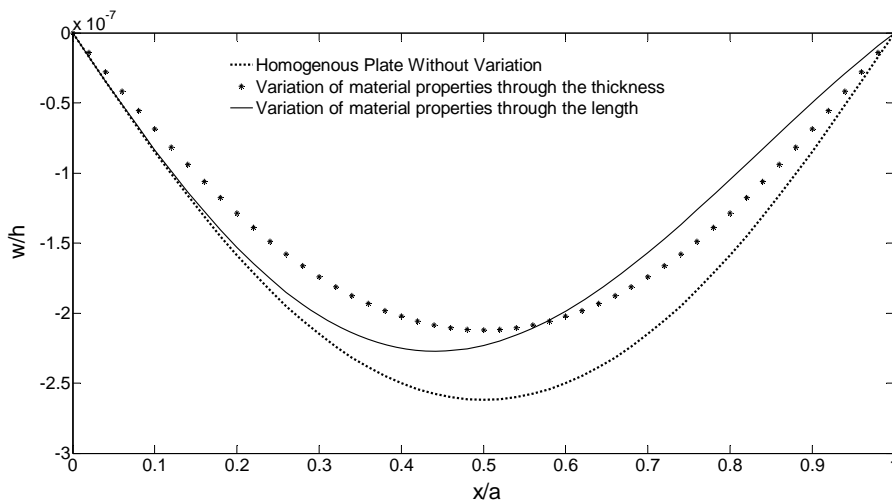
Table 1. material properties

Metal (Ti-6Al-4V)	Ceramic ( $ZrO_2$ )
$E_m = 66.2GPa$	$E_i = 117.0GPa$
$\nu = 0.33$	$\nu = 0.33$

By applying the  $n=3$ ,  $F_0 = 1N/cm^2$  and Fourier series coefficient ( $m= n =20$ ), the theoretical results can be evaluated with numerical modelling. Graded solid elements are implemented by means of direct sampling properties at the Gauss points of the elements. The user material subroutine (UMAT) for modelling of FGM plate with graded elements is provided in ABAUQS.

**Results**

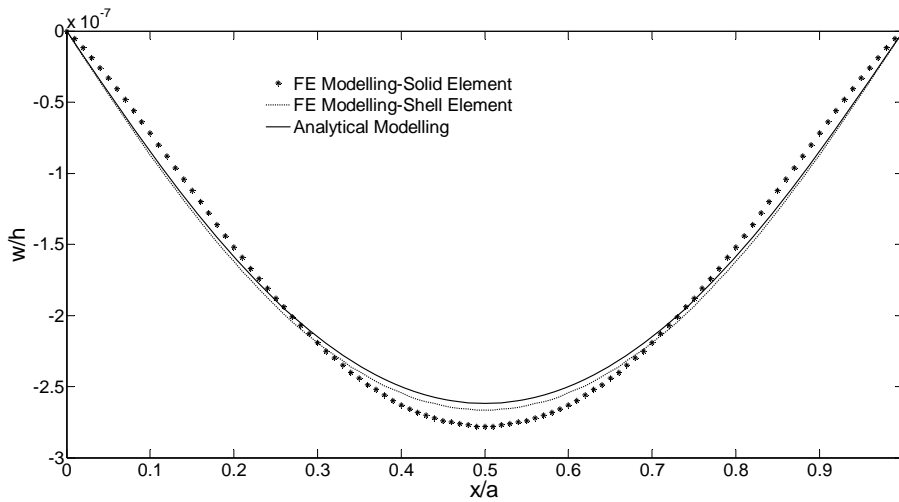
Fig.3 compares the deflection of the plate with different gradient variation (homogenous plate without variation and variation in Z and X direction, respectively). It can be seen that by applying the higher stiffness material (ceramic) as an inclusion in the metal matrix, the deflection of plate decrease significantly. Moreover, the FGM plate with variation of material constituents in thickness direction has the least maximum deflection with the symmetry parabolic shape. While the maximum deflection for FGM plate with X variation does not occur in the middle of the length. It happens in 0.4 times length.



**Figure 3. The deflection of the plate**

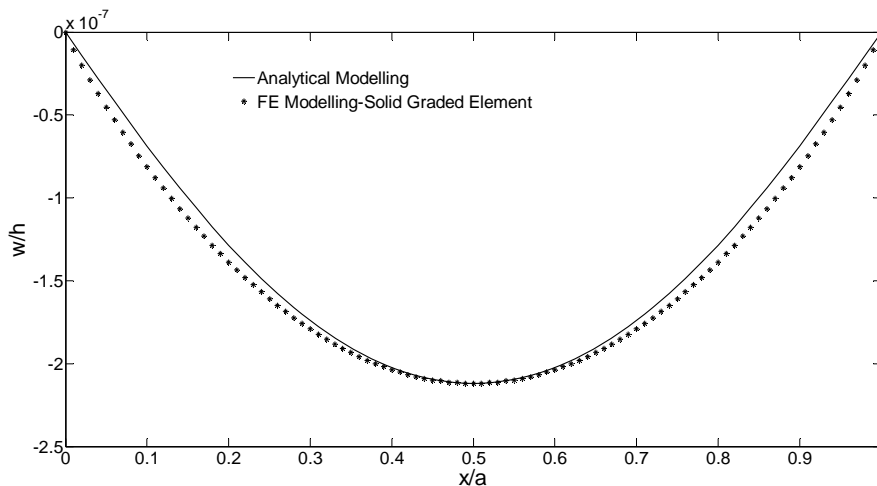
Fig. 4 shows the theoretical and FE (both Shell and Solid elements) results for a square homogenous plate without variation. As revealed in Fig. 4, a good agreement is obtained in using solid elements with mesh size 0.25 with shell elements. So, the

solid elements can be used for modelling of a thin FGM plate instead of shell elements.

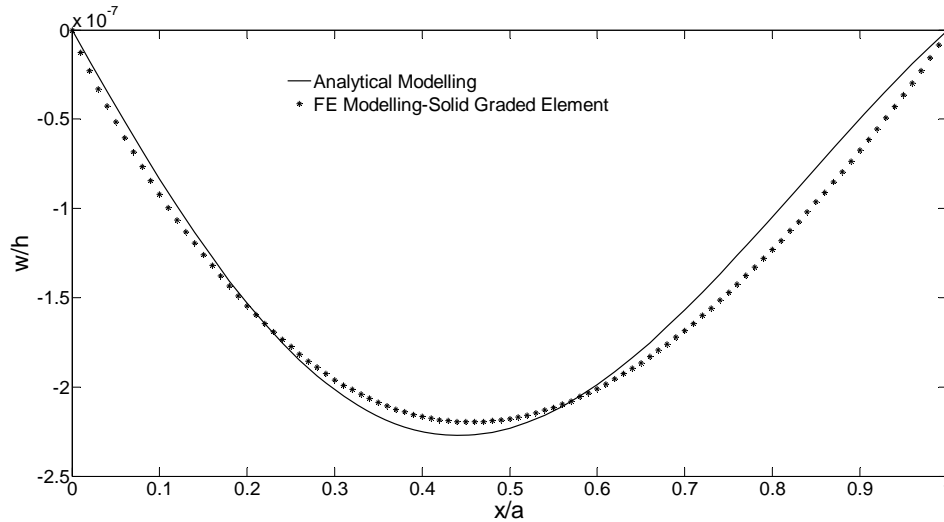


**Figure 4. The deflection of the homogenous plate**

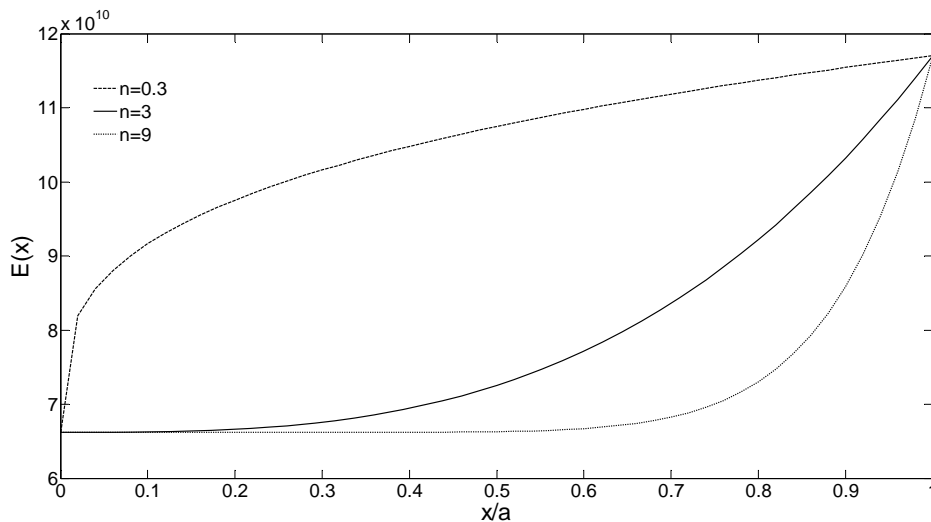
Figs. 5 and 6 represent the deflection of FGM plate for Z and X gradient variation analytically and numerically, respectively. The analytical and numerical results agree very well in Z and X variation with the maximum error less than 10%. As can be seen from the comparison of two graphs in Fig. 6, the changes between analytical and numerical solution after critical point-maximum deflection- is much pronounced compared to the ones before that. It could be caused by power-law index (n) effect. It can be observed from Fig. 7 that the jump of E occurs at the critical point for  $n=3$  – approximately 0.4 times of plate length-. Indeed, before this point there is a little bit variation in Young modulus and after that the considerable growth of E is obtained. Moreover, as n is raised the jump of E occurs in near the end of the plate ( $x=a$ ).



**Figure 5. The deflection of the FGM plate with variation of material properties through the thickness**

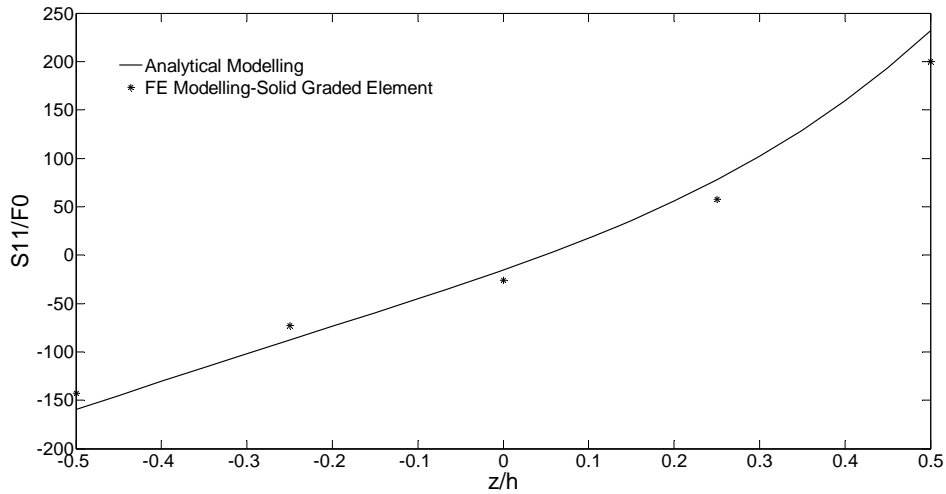


**Figure 6. The deflection of the FGM plate with variation of material properties through the length**

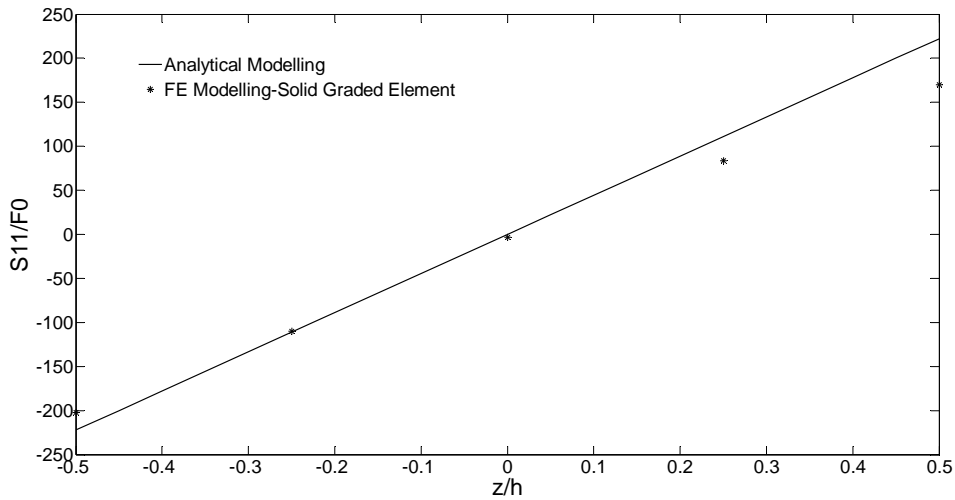


**Figure 7. The effect of power-law index in E(x)**

The variation of the stress in x direction at the center of the FGM plate along the thickness for state I and II is depicted in Figs. 8 and 9. The stress for FGM plate with Z gradient variation is a function of Z of order 4. This phenomenon coincides with the analytical formulation in Eq. (13d), in which the stresses are proportional to  $z.E(z)$ . While for X gradient variation, the linear function of Z is presented (based on Eq. (20a)). The maximum tensile and compressive stress in the center of the FGM plate is at the bottom and top edge, respectively. In addition, the good agreement is obtained between analytical and numerical results with the maximum error less than 15%.



**Figure 8. The stress  $\sigma_x$  at the center of FGM plate for variation of material properties through the thickness**



**Figure 9. The stress  $\sigma_x$  at the center of FGM plate for variation of material properties through the length**

**Conclusions**

In this study, the theoretical formulation of the simply- supported thin square FGM plate with power-law distribution of volume fraction through the thickness (state I) and length (state II) under transverse loading is derived. The analytical results are also validated by finite element analysis. The FE solution is obtained using solid elements with spatially graded property distribution, which is implemented by a user material subroutine (UMAT) in ABAQUS. The results lead to the following conclusions:

1. Strain at middle surface and deflection for state I are coupled while this phenomenon is different from State II. There is extra coefficient in strain equations in state II that is due to gradient variation in X direction and differentiation of deflection with respect to X.

2. Under the assumption of thin plate the theoretical results agree very well with those of FE simulation with solid graded element.
3. There is not a linearly proportion between stress and thickness direction in state I. because the Young's moduli of this state are functions of  $z$  of order  $n$ -power law index-, so the stress indicates a function of  $z$  of order  $n+1$ .
4. The maximum deflection occurs in the middle plate for state I while it happens before that in state II-at nearby 0.4 times of length-.

The approach outlined in this paper would be beneficial for the ideal FGM plate which variation of volume fraction of constituents is controlled and obey a specific function of distribution. This aim could be achieved with aid of 3D printing that can control local composition and microstructure.

## References

- Akbarzadeh, A. H., Hosseini zad, S. K., Eslami, M. R., & Sadighi, M. (2010). Mechanical behaviour of functionally graded plates under static and dynamic loading. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 1(-1), 1-8.
- Bocciarelli, M., Bolzon, G., & Maier, G. (2008). A constitutive model of metal-ceramic functionally graded material behavior: Formulation and parameter identification. *Computational Materials Science*, 43(1), 16-26.
- Cheng, Z. Q., & Batra, R. C. (2000). Exact Correspondence between Eigenvalues of Membranes and Functionally Graded Simply Supported Polygonal Plates. *Journal of Sound and Vibration*, 229(4), 879-895.
- Chi, S.-H., & Chung, Y.-L. (2006). Mechanical behavior of functionally graded material plates under transverse load—Part I: Analysis. *International Journal of Solids and Structures*, 43(13), 3657-3674.
- Eghesad, A., Shaffei, A. R., & Mahzoon, M. (2012). Study of dynamic behavior of ceramic-metal FGM under high velocity impact conditions using CSPM method. *Applied Mathematical Modelling*, 36(6), 2724-2738.
- El-Desouky, A., Kassegneb, S. K., Moonb, K. S., McKittrick, J., & Morsib, K. (2013). Rapid processing & characterization of micro-scale functionally graded porous materials. *Journal of Materials Processing Technology*, 1251– 1257.
- Guo, L., & Noda, N. (2014). Investigation Methods for Thermal Shock Crack Problems of Functionally Graded Materials—Part I: Analytical Method. *Journal of Thermal Stresses*, 37(3), 292-324.
- Kieback, B., Neubrand, A., & Riedel, H. (2003). Processing techniques for functionally graded materials. *Materials Science and Engineering*, A362, 81–105.
- Navazi, H. M., Haddadpour, H., & Rasekh, M. (2006). An analytical solution for nonlinear cylindrical bending of functionally graded plates. *Thin-Walled Structures*, 44(11), 1129-1137.
- Pan, E. (2003). Exact Solution for Functionally Graded Anisotropic Elastic Composite Laminates. *Journal of Composite Materials*, 37(21), 1903-1920.
- Sun, D., & Luo, S.-N. (2011). Wave propagation and transient response of functionally graded material circular plates under a point impact load. *Composites Part B: Engineering*, 42(4), 657-665.
- Thai, H.-T., & Choi, D.-H. (2013). Analytical solutions of refined plate theory for bending, buckling and vibration analyses of thick plates. *Applied Mathematical Modelling*, 37(18-19), 8310-8323.
- Torshizian, M. R., & Kargarnovin, M. H. (2014). The mixed-mode fracture mechanics analysis of an embedded arbitrary oriented crack in a two-dimensional functionally graded material plate. *Archive of Applied Mechanics*, 84(5), 625-637.
- Wang, C., & Xu, X. (2014). A new hybrid element analysis for exponentially varying FGM based on an asymptotic stress field. *Acta Mechanica Sinica*, 27(3), 315-330.
- Wang, Z., Guo, L., & Zhang, L. (2013). A general modelling method for functionally graded materials with an arbitrarily oriented crack. *Philosophical Magazine*, 94(8), 764-791.

## Stochastic analysis of a radial-inflow turbine in the presence of parametric uncertainties

\*A. Zou<sup>1</sup>, †E. Sauret<sup>1</sup>, J.-C. Chassaing<sup>2</sup>, S. C. Saha<sup>1</sup>, and YT Gu<sup>1</sup>

<sup>1</sup>School of Chemistry, Physics & Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4001, Australia.

<sup>2</sup>Sorbonne Université, UPMC Uni Paris 06, CNRS, UMR7190, d'Alembert Institute, F-75005, Paris, France.

\*Presenting author: a.zou@qut.edu.au

†Corresponding author: emilie.sauret@qut.edu.au

### Abstract

This paper presents an uncertainty quantification study of the performance analysis of the high pressure ratio single stage radial-inflow turbine used in the Sundstrand Power Systems T-100 Multi-purpose Small Power Unit. A deterministic 3D volume-averaged Computational Fluid Dynamics (CFD) solver is coupled with a non-statistical generalized Polynomial Chaos (gPC) representation based on a pseudo-spectral projection method. One of the advantages of this approach is that it does not require any modification of the CFD code for the propagation of random disturbances in the aerodynamic and geometric fields. The stochastic results highlight the importance of the blade thickness and trailing edge tip radius on the total-to-static efficiency of the turbine compared to the angular velocity and trailing edge tip length. From a theoretical point of view, the use of the gPC representation on an arbitrary grid also allows the investigation of the sensitivity of the blade thickness profiles on the turbine efficiency. The gPC approach is also applied to coupled random parameters. The results show that the most influential coupled random variables are trailing edge tip radius coupled with the angular velocity.

**Keywords:** Uncertainty Quantification, gPC, CFD, Radial Turbine

### Introduction

In order to develop robust turbines' design, it is critical to consider uncertainties in the optimization process. Turbine design is nowadays mainly achieved using Computational Fluid Dynamics (CFD) technique. However, the design is made at the optimal operating conditions and any likely variations in operations can dramatically alter the efficiency of the turbine [Odabae et al. (2014); Sauret and Gu (2014)]. The concern is even greater while working with high-density fluids in low-temperature Organic Rankine Cycles (ORC) which have the potential to extract more energy from the high-density fluids. Due to the complexity of the geometry and computational cost associated with these geometries and fluids, no advanced uncertainty quantification (UQ) has been proposed so far for these ORC turbines and very little work has been done on the uncertainty quantification in turbomachinery in general. Relatively recently, Gopinathrao et al. [Gopinathrao et al. (2009)] and Loeven and Bijl [Loeven and Bijl (2010)] successfully performed non-intrusive Polynomial Chaos and Probabilistic Collocation respectively on a transonic axial compressor but no study has been found on radial turbines.

UQ is a mathematical approach employed to determine the likely certain outcomes in an uncertain system. Any engineering system is subject to uncertainties, which can come from the random variation of geometric parameters and operating conditions for example. These uncertainties cannot be removed from the system and are called "aleatory" uncertainties [Faragher (2004)]. In addition, the numerical representation of this system also introduces uncertainties through the mathematical models and boundary conditions used. These "epistemic" uncertainties [Faragher (2004)], however, can be reduced as they are due to modelling errors. In the numerical simulations, all these parameters are constants, which make it impossible to isolate the influence of these parameters without using uncertainty quantification methods.

So far, different approaches have been developed in order to quantify these uncertainties. Monte Carlo (MC) technique is a typical approach employed to perform probabilistic analysis. However, it is suffering from expensive computational cost and poor convergence rates, especially for complex geometries [Sankaran and Marsden (2011)]. To overcome such issue, other approaches have been developed, such as so-called Polynomial Chaos (PC) method based

on the spectral stochastic finite elements approach [Spanos and Ghanem (1989); Ghanem and Spanos (1991)], generalized Polynomial Chaos (gPC) as extended by Xiu [Xiu et al. (2002)], stochastic collocation method [Mathelin and Hussaini (2003)], and Multi-Element generalized Polynomial Chaos (ME-gPC) method [Wan and Karniadakis (2005); Kewlani and Iagnemma (2009)]. Non-intrusive methods are also becoming more popular as they don't require the modification of the deterministic solver. A comparison between intrusive and non-intrusive methods is presented by Onorato et al. [Onorato et al. (2010)] while non-intrusive approaches are detailed in [Loeven et al. (2007)].

In this work, a generalized Polynomial Chaos (gPC) method is coupled to the deterministic CFD solver and applied to the performance analysis of a radial-inflow turbine. A uniform distribution of the random parameters associated with Legendre polynomials is chosen. Due to the curse of dimensionality the uncertain parameters are investigated separately with high-order spectral projections while the combined effect of the parameters is initially investigated using low-order polynomials. The impact of the variable input parameters are evaluated on the total-to-static efficiency of the radial turbine. The stochastic space of each random variable is correlated to the range of uncertainty of the physical input parameters. The sensitivity to the uncertain parameters and their potential coupled effects on the stochastic turbine efficiency are discussed in details.

### Generalized Polynomial Chaos Method

In this study, the generalized Polynomial Chaos (gPC) framework proposed by Spanos and Ghanem [Spanos and Ghanem (1989); Ghanem and Spanos (1991)] is used. The gPC representation of a random process  $u$  is defined as:

$$u(x, \Theta) = \sum_{\alpha \in \mathbb{N}^N} \hat{u}_\alpha(x) \phi_\alpha(\Theta) \quad (1)$$

where  $\Theta = \{\Theta_j(\omega)\}_{j=1}^N$ ,  $N \in \mathbb{N}^N$ , is a  $\mathbb{R}^N$  valued random array on a probability space  $(\Omega, \mathcal{A}, \mathcal{P})$  with probability distribution  $P_\Theta(d\theta)$  and  $d\theta$  is the Lebesgue measure.  $\phi_\alpha(\Theta)$  is the multivariate orthogonal polynomials, with total degree not greater than  $P$ . They are built as tensor products of orthogonal polynomials along each random dimension with respect to the probability measure  $P_\Theta(d\theta)$ . The modal coefficients in Eq. (1) are determined by:

$$\hat{u}_\alpha(x) = E\{u(x, \Theta) \phi_\alpha(\Theta)\} / E\{\phi_\alpha^2(\Theta)\} \quad \text{for } \alpha \in \mathbb{N}^N \quad (2)$$

where  $E$  represents the expectation. The order  $P$  of the polynomial basis is chosen based on accuracy requirements.

The modal coefficients can be re-written as:

$$\hat{u}_j(x) = \frac{1}{E\{\phi_\alpha^2(\Theta)\}} \sum_{i=0}^{N_q} w_i u(x, \Theta) \phi_j(\Theta) \quad (3)$$

where the weights  $w_i$  and nodes  $\Theta$  of the Gauss-Legendre quadrature are determined by solving an eigenvalue problem based on the Golub-Welsch algorithm.  $N_q = (N_d + 1)^{N_d} - 1$  is the number of cubature points, with  $N_d$ , the number of random parameters.

### Interpolated gPC

As demonstrated by Sauret et al. [Sauret et al. (2014)] interpolated gPC can provide useful approximations of the gPC approach. The method uses the existing deterministic solutions as an arbitrary grid on which preferably high-order interpolations are performed to carry out the stochastic projection. This is of particular interest for the blade thickness profiles evaluation for which re-creating the profiles for each quadrature point is extremely time consuming. Thus, this approach is used here despite the reduced accuracy as a preliminary estimation of the sensitivity of the turbine performance to the blade thickness profiles.

### Statistical Post-Processing of the gPC Method

Once the modal coefficients are determined using Eq. (1), the statistical properties of the random parameters are obtained thanks to the orthogonality of the polynomial basis. The mean  $\mu$  and the variance  $\sigma^2$  are thus obtained by:

$$\mu = \hat{u}_0 \quad (4)$$

$$\sigma^2 = \sum_{j=1}^M \hat{u}_j^2(x) E\{\phi_\alpha^2\} \quad (5)$$

Then the standard deviation,  $\sigma = \sqrt{\sigma^2}$  and the coefficient of variation,  $CoV = \sigma/\mu$  are obtained from Eqs. (3) and (4).

### gPC-CFD coupling

The gPC method used in this study is non-intrusive and thus doesn't require any modification of the deterministic solver. The gPC method is implemented in Matlab and automatically coupled with the CFD solver *ANSYS-CFX* using Python scripting. The CFD results are then sent back into Matlab for the statistical post-processing.

### Radial-Inflow Turbine

The radial-inflow turbine used in this work has been developed by Sundstrand and experimentally tested by Jones [Jones (1996)]. This geometry has become an open benchmark after the work of Sauret [Sauret (2012)] who reconstructed the geometry and provided initial CFD results.

The test case at nominal conditions is a 120 kW, 5.7 pressure ratio turbine used in the Sundstrand Power Systems T-100 Multi-purpose Small Power Unit. However, only the rig conditions have been experimentally tested which have a lower rotational speed and lower inlet pressure but the same pressure ratio. The rig conditions are used in this study for validation and application of the gPC method. The geometry is presented below in Figure 1 and the full details are presented in Sauret [Sauret (2012)]. The turbine has 19 stator blades and 16 rotor blades.



Figure 1. Rotor and Stator geometry of the radial-inflow turbine



### Deterministic volume-averaged CFD solver

Three-dimensional geometry and mesh of one blade passage including stator, rotor and part of diffuser are reproduced in *ANSYS* turbomachinery package. Reynolds-Averaged Navier-Stokes equations are solved in this simulation using *ANSYS-CFX version 15*. The  $k-\omega$  SST turbulence model is used for the simulations and high resolution schemes are used for both the advection and turbulence as recommended by [Louda et al. (2013)].

#### Boundary Conditions

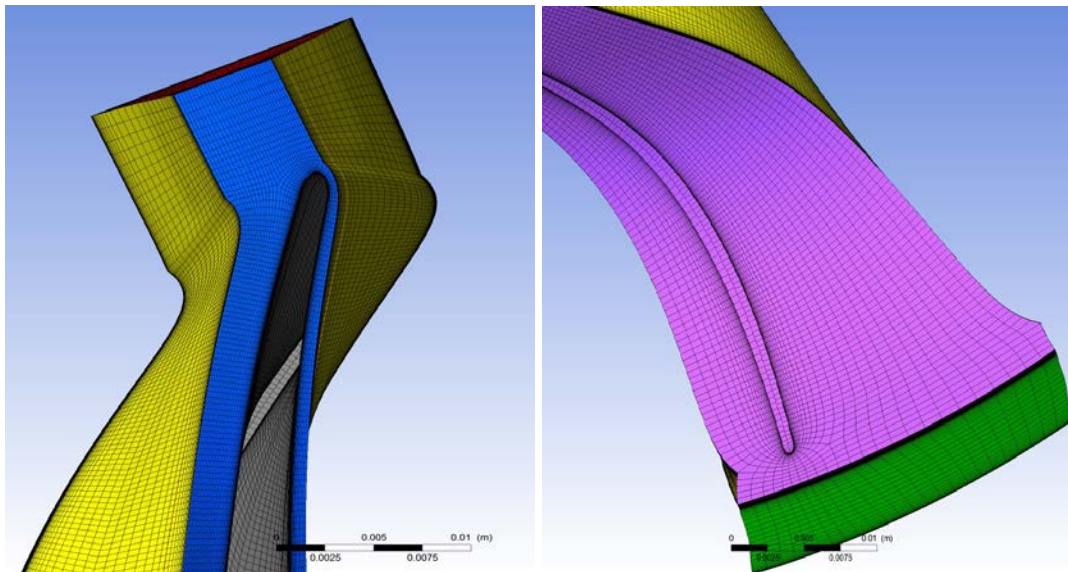
For the rig condition, the temperature of inlet of stator is  $477.6\text{ K}$  and the outlet static pressure is  $72.4\text{ kPa}$ . The mass flow rate at the inlet of the stator is  $\dot{Q}_m = 0.0173\text{ kg/s}$ . The working fluid is air, considered as ideal gas. The rotational speed is  $71700\text{ RPM}$ . [Sauret (2012)].

#### Mesh

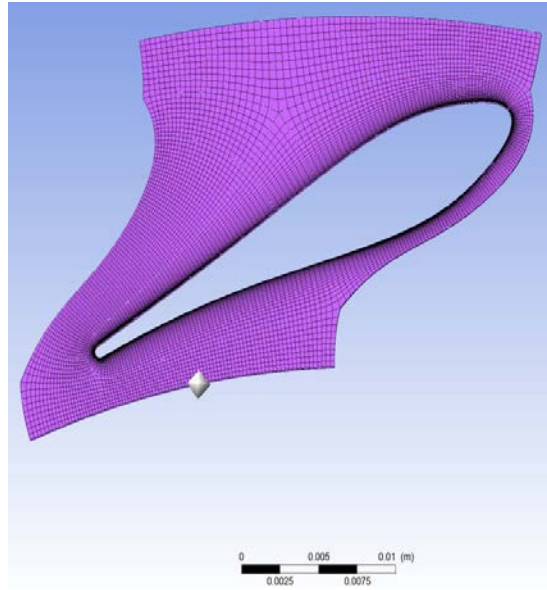
The mesh is generated using *ANSYS-TurboGrid* for the flow passage for both rotor and stator. The non-dimensional grid spacing at the wall  $y_w^+$  ranges from 20 to 140, which is the recommended range as the log-law wall function is valid for  $y_w^+$  values above 15 and under 100 for machine Reynolds number of  $1 \times 10^5$  where the transition affects the boundary layer formation and skin friction and up to 500 for Reynolds number of  $2 \times 10^6$  when the boundary layer is mainly turbulent throughout [Manual (2000)]. The boundary layer refinement control is  $4 \times 10^6$  with Near Wall Element Size Specification to reach the  $y_w^+$  (non-dimensional wall element size) requirement for the  $k-\omega$  SST turbulence model.

After a grid refinement study, the total mesh number is 712,082 including stator, rotor and part of diffuser. The grid quality was checked using indicators such as orthogonality of the cells and aspect ratios. The converged mesh is presented in Figure 2.

All of the computations were performed until full convergences of the flow variables were achieved. The residuals were dropped down below  $10^{-6}$ .



**Figure 2. Three-dimensional view of the O–H grid at the rotor blade at the hub and shroud.**

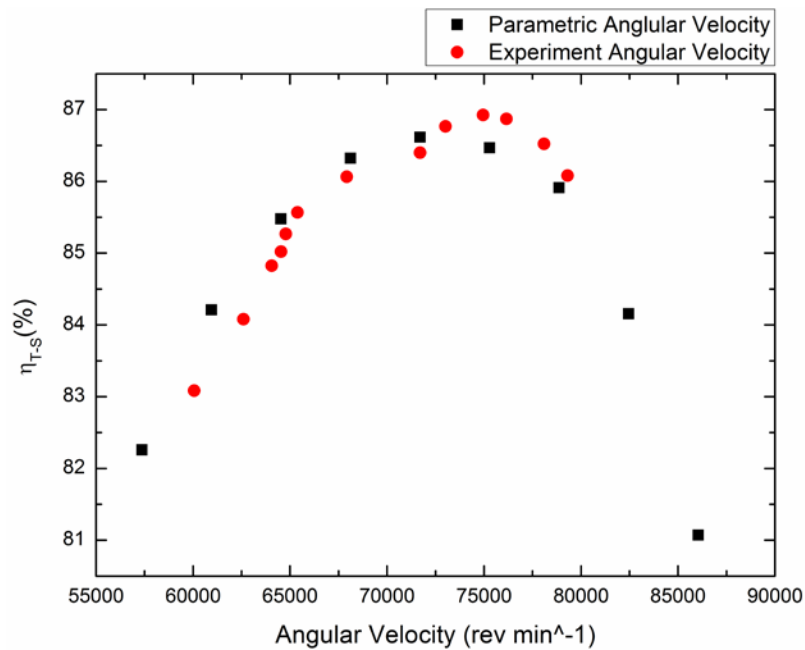


**Figure 3. Two-dimensional view of the O–H grid at the stator blade.**

## Results

### Validation

From Figure 4, three-dimensional CFD total-to-static efficiency is compared against the experimental data for the rig conditions. The results are in really good agreement with the experiments with a maximum difference less than 1%.



**Figure 4. Variation of total-to-static efficiency with rotational speed.**

*Parametric Study*

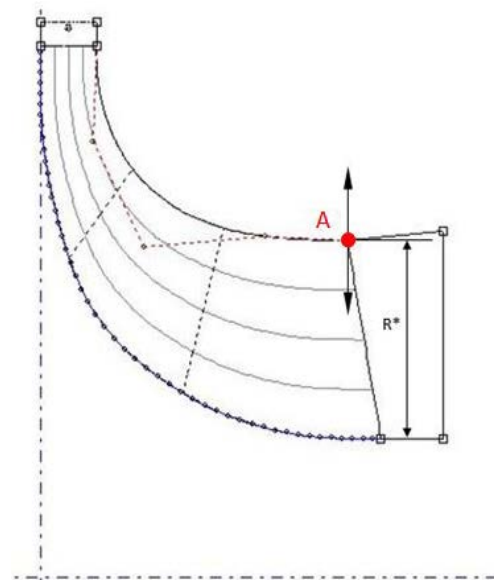
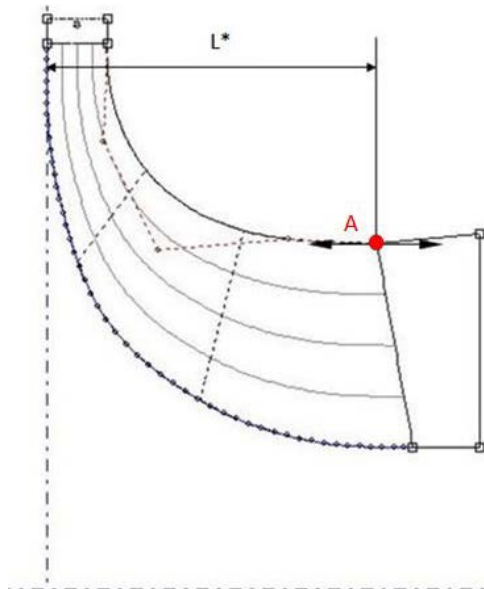
Four main parameters, angular velocity, TE (Trailing Edge) Tip Length, TE Tip Radius and blade thickness profiles (Table 1) have been initially investigated in order to identify a non-linear response surface on which the gPC method will be valuable to apply.

**Table 1. Characteristics of the studied uncertain parameters**

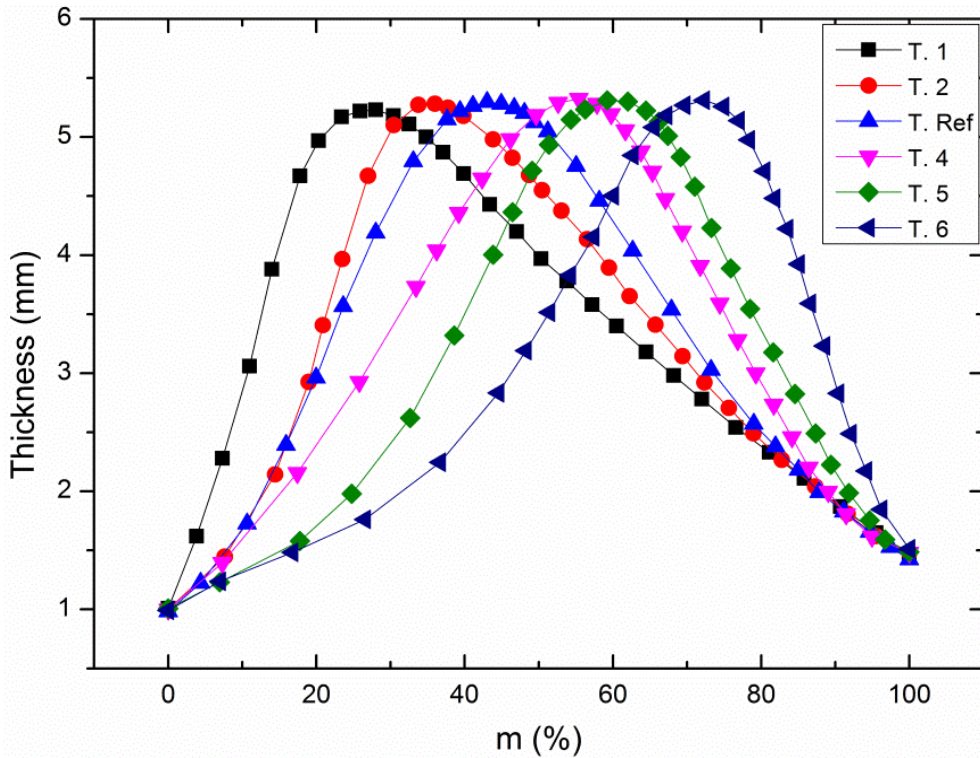
Uncertain Parameter	$\mu$	Support
Angular Velocity $\omega$ (rev.min <sup>-1</sup> )	71700	[57360, 86040]
TE Tip Length L (mm)	35.0012	[33.1, 42 ]
TE Tip Radius R (mm)	36.83	[31.1, 37]
Blade Thickness peak position along the meridional length (%)	41	[21, 71]

The TE Tip Length and TE Tip Radius are defined in Figures 5 and 6. The red point “A” in Figures 5 and 6 is the geometry changing point, corresponding to the TE position at the shroud. The arrows’ direction is the geometry changing direction. When “A” point is moving in horizontal direction, the TE tip length will vary. It is important to note that when “A” point is moving in the vertical direction (TE Tip Radius), the blade height will be modified and the tip clearance will be kept at the initial value.

In Figure 7, six different rotor blade thickness profiles have been manually established for the parametric study. The maximum value of the blade thickness is kept constant while its location is moved along the tip length, thus modifying the profile curve shape.

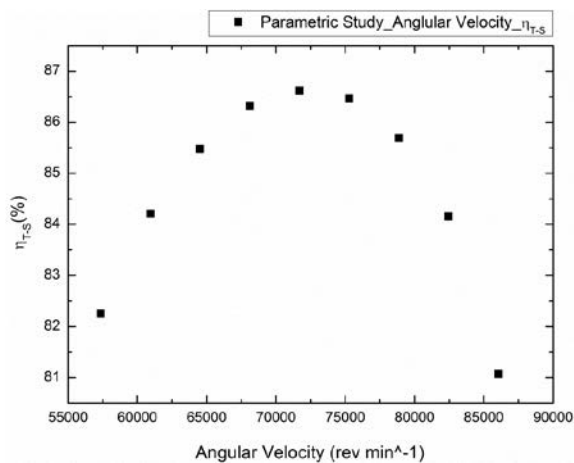


**Figure 5. TE Length geometric study.      Figure 6. TE Tip Radius geometric study.**

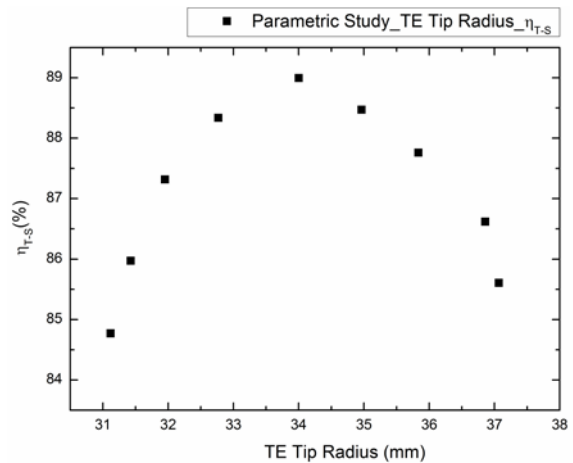


**Figure 7. Blade thickness profile geometric study.**

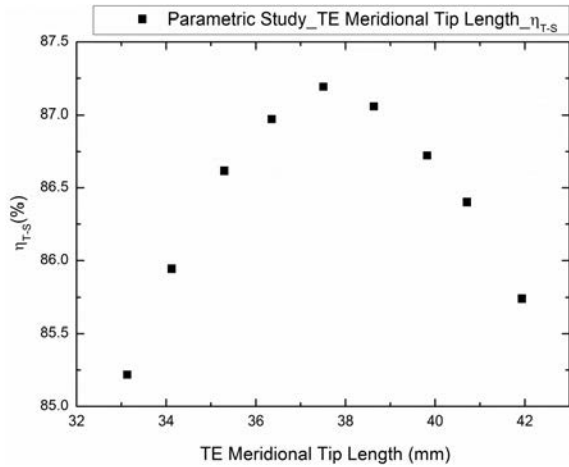
As shown in Figures 8-11, angular velocity, TE tip radius, TE tip length and blade thickness have non-linear response surfaces. One can also note that the maximum efficiency is obtained at values of the TE tip radius, TE tip length and blade thickness different from the initial Jones' geometry, indicating that optimization of this turbine can be achieved. Angular velocity, TE tip radius, TE tip length and blade thickness are thus used as random inputs for the application of the gPC method.



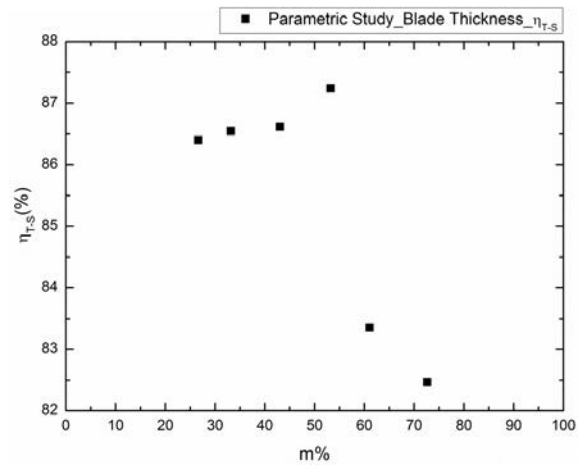
**Figure 8. Evolution of the total-to-static efficiency with the angular velocity.**



**Figure 9. Evolution of the total-to-static efficiency with TE tip radius.**



**Figure 10. Evolution of the total-to-static efficiency with the TE tip length.**



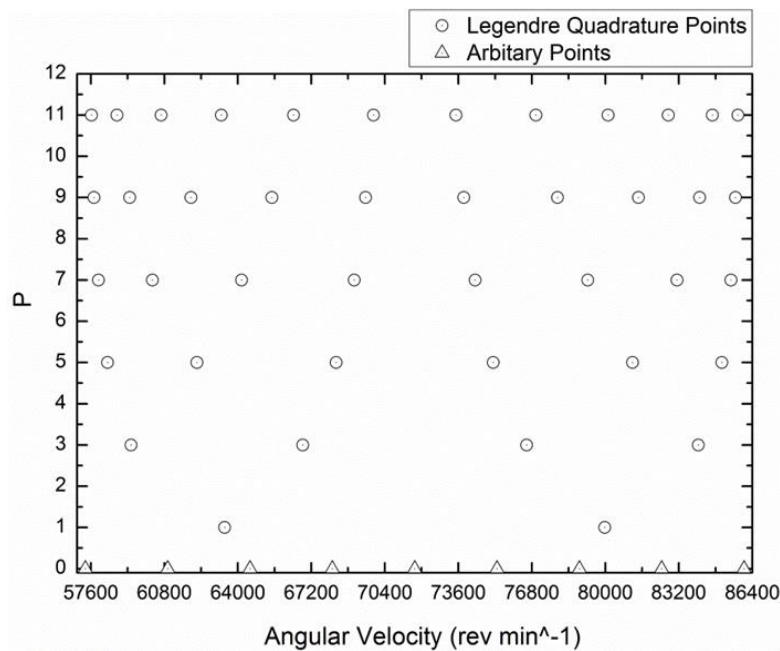
**Figure 11. Evolution of the total-to-static efficiency with the blade thickness profiles.**

### Uncertainty Quantification

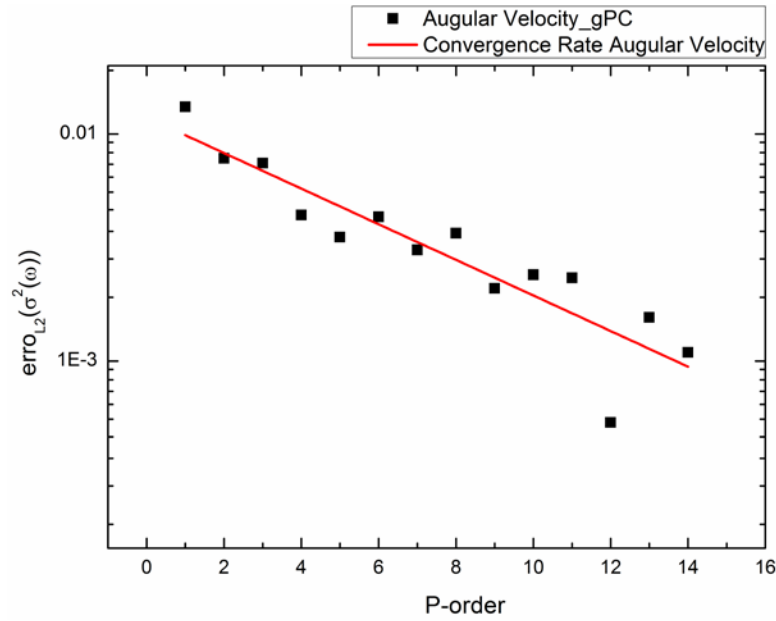
The mean and support for the 4 random parameters (angular velocity, TE tip radius, TE tip length and blade thickness) are summarized in Table 1.

#### Convergence Study

Figure 12 shows the CFD points and the gPC legendre quadrature points for  $P$  ranging is 1, 3, 5, 7, 9, 11 respectively, when angular velocity is the random variable.



**Figure 12. Legendre quadrature points and arbitrary support points for Angular Velocity for P=1, 3, 5, 7, 9, 11**



**Figure 13. Convergence rates of the variance of the angular velocity in respect to the total-to-static efficiency.**

In Figures 13, symbols represent simulations while lines are the corresponding linearly fitted decay rates. It can be seen that the error line trend decreases when the P-order is increasing, showing good convergence rates. Similar trends are observed for all uncertain variables.

#### Statistical results

In Table 2,  $\omega$ ,  $L$  and  $R$  refer to the angular velocity, TE tip length and TE tip radius respectively.

Regarding each uncertain parameter, the mean value  $\mu$ , the standard deviation  $\sigma$ , variance  $\sigma_v^2$  and the coefficient of variation  $\text{CoV} = \sigma/\mu$  of the total to static efficiency with the gPC at  $P = 11$  are presented in Table 2. It can be seen that the blade thickness profile has the most influential effect on the turbine total-to-static efficiency closely followed by the TE tip radius  $R$  while  $L$  doesn't appear to be a critical geometric parameter in regards to the efficiency. The gPC method was also applied for coupled uncertain parameters with a lower polynomial order  $P=5$  in order to minimize the computational cost. When parameters are coupled, the most influential coupled random variables on the total-to-static efficiency are  $R-\omega$ .

**Table 2. Mean, standard deviation and CoV of the total-to-static efficiency for each individual uncertain parameter for  $P=11$  and coupled parameters for  $P=5$ .**

Variable	gPC 1D (P=11)				gPC 2D (P=5)		
	$\omega$	$L$	$R$	Blade Thickness	$R-\omega$	$L-\omega$	$R-L$
$\mu$	85.09	86.72	85.65	85.5	83.27	85.34	81.68
$\sigma \times 10^{-3}$	13.409	1.720	16.611	17.400	25.360	14.971	13.102
$\sigma_v^2 \times 10^{-3}$	0.180	0.003	0.276	0.303	0.643	0.224	0.172
$\text{CoV} \times 10^{-3}$	15.759	1.983	19.393	20.400	30.454	17.543	16.040

#### Conclusion

In this paper, a deterministic 3D CFD solver is coupled with gPC approach and successfully applied to investigate a complete 3D high-pressure ratio radial-inflow turbine. The uncertainty quantification has been applied to the performance analysis of radial turbine for the propagation

of various aerodynamic and geometric uncertainties. The convergence rate for each uncertain parameter has been investigated, showing that the stochastic spectral projection decreases dramatically with the increase of polynomial order. The initial deterministic study highlighted the non-linear response of the total-to-static turbine efficiency in regard to the variations of the angular velocity, TE tip radius, TE tip length and blade thickness. From the preliminary study, for the CoV of the total to static efficiency, the most influential uncertainty is the blade thickness closely followed by the TE tip radius. When the gPC approach is applied to coupled random parameters, the most influential coupled random variables are the trailing edge tip radius with the angular velocity. In future work, other parameters of radial turbine and more dimensional gPC will be investigated, such as maximum blade thickness. Then stochastic collocation method will be applied for the uncertainty quantification analysis of the radial turbine.

## Acknowledgement

Dr. E. Sauret would like to thank the Australian Research Council for their financial support (DE130101183).

## References

- Faragher, John. 2004. *Probabilistic Methods for the Quantification of Uncertainty and Error in Computational Fluid Dynamic Simulations*: DTIC Document.
- Ghanem, Roger G and Pol D Spanos. 1991. *Stochastic finite elements: a spectral approach*. Vol. 387974563: Springer.
- Gopinathrao, N. P., D. GBagshaw, C. Mabilat and S. Alizadeh. 2009. "Non-Deterministic CFD Simulation of Transonic Compressor Rotor." In *ASME Turbo Expo 2009: Power and Land, Sea and Air, GT2009, Orlando, Florida, USA, June 8-12, 2009*, edited.
- Jones, AC. 1996. "Design and test of a small, high pressure ratio radial turbine." *Journal of turbomachinery* 118 (2): 362-370.
- Kewlani, Gaurav and Karl Iagnemma. 2009. "A multi-element generalized polynomial chaos approach to analysis of mobile robot dynamics under uncertainty." In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, edited, 1177-1182: IEEE.
- Loeven, G. J. A. and H. Bijl. 2010. "The application of the probabilistic collocation method to a transonic axial flow compressor." In *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, Orlando, FL; United States, 12-15 April, 2010*, edited.
- Loeven, G. J. A., J. A. S. Witteveen and H. Bijl. 2007. "Probabilistic Collocation: An Efficient Non-Intrusive Approach for Arbitrarily Distributed Parametric Uncertainties." In *45th AIAA Aerospace Science Meeting and Exhibit AIAA-2007, Reno, Nevada, USA, 8-11 January, 2007*, edited.
- Louda, Petr, Petr Sváček, Jaroslav Fořt, Jiří Fürst, Jan Halama and Karel Kozel. 2013. "Numerical simulation of turbine cascade flow with blade-fluid heat exchange." *Applied Mathematics and Computation* 219 (13): 7206-7214.
- Manual, ANSYS User's. 2000. "Ansys Inc". In *Canonsburg, PA*.
- Mathelin, Lionel and M Yousuff Hussaini. 2003. *A stochastic collocation algorithm for uncertainty analysis*. Hampton, Virginia: National Aeronautics and Space Administration, Langley Research Center. Langley Research Center.
- Odabae, M, Mohsen Modir Shanechi and K Hooman. 2014. "CFD Simulation and FE Analysis of a High Pressure Ratio Radial Inflow Turbine." In *19AFMC: 19th Australasian Fluid Mechanics Conference*, edited, 1-4: Australasian Fluid Mechanics Society.
- Onorato, G., G. J. A. Loeven, G. Ghorbaniasl, H. Bijl and C. Lacor. 2010. "Comparison of Intrusive and Non-Intrusive Polynomial Chaos Methods for CFD Applications in Aeronautics." In *V European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2010, Lisbon, Portugal, 14-17 June, 2010*, edited by J. C. F. Pereira and A. Sequeira (Eds).
- Sankaran, S. and A. L. Marsden. 2011. "A stochastic collocation method for uncertainty quantification and propagation in cardiovascular simulations." *J Biomech Eng* 133 (3): 031001-1-031001-12. doi: 10.1115/1.4003259.

- Sauret, Emilie. 2012. "Open design of high pressure ratio radial-inflow turbine for academic validation." In *ASME 2012 International Mechanical Engineering Congress and Exposition*, edited, 3183-3197: American Society of Mechanical Engineers.
- Sauret, Emilie and Yuantong Gu. 2014. "Three-dimensional off-design numerical analysis of an organic Rankine cycle radial-inflow turbine." *Applied Energy* 135: 202-211.
- Sauret, Emilie, Rodney Persky, Jean-Camille Chassaing and Didier Lucor. 2014. "Uncertainty quantification applied to the performance analysis of a conical diffuser." Paper presented at the 19th Australasian Fluid Mechanics Conference (AFMC), RMIT University, Melbourne, 8-11 December.
- Spanos, Pol D and Roger Ghanem. 1989. "Stochastic finite element expansion for random media." *Journal of engineering mechanics* 115 (5): 1035-1053.
- Wan, Xiaoliang and George Em Karniadakis. 2005. "An adaptive multi-element generalized polynomial chaos method for stochastic differential equations." *Journal of Computational Physics* 209 (2): 617-642.
- Xiu, Dongbin, Didier Lucor, C-H Su and George Em Karniadakis. 2002. "Stochastic modeling of flow-structure interactions using generalized polynomial chaos." *Journal of Fluids Engineering* 124 (1): 51-59.



## AERODYNAMIC ANALYSIS OF THE AIRFOIL OF A VAWT BY USING 2D CFD MODELLING

\*Luis F. Garcia<sup>1</sup>, †Julian E. Jaramillo<sup>2</sup> and Jorge L. Chacón<sup>3</sup>

<sup>1</sup>Department of Mechanical Engineering, Universidad Industrial de Santander, UIS.

<sup>2</sup>Department of Mechanical Engineering, Universidad Industrial de Santander, UIS.

<sup>3</sup>Department of Mechanical Engineering, Universidad Industrial de Santander, UIS

\*Presenting author: ingarcia1703@gmail.com

†Corresponding author: jejarami@uis.edu.co

### Abstract

Colombia is a country where the benefits of wind power industry are barely used because of the geography in some areas does not allow the implementation of onshore horizontal axis wind turbines. Furthermore, exist rural areas without access to the electrical grid. Therefore, there is currently a deficit of energy supply in some towns. This research took place in one of those areas (i.e. Chicamocha Canyon-Santander) where the answer to the energy supply problems could be the use of vertical axis wind turbines, which can be used for turbulent flows. Hence, one task of this research is the analysis of the wind resources in the Chicamocha Canyon in order to implement the wind energy. The wind turbines must be designed in such a way that the blades take good advantage of the wind resources in the area of interest. Consequently, in the current research the analysis of two different airfoils (i.e. NACA0018 and DÜ 06-W-200) through a 2D CFD simulation is carried out by means of a free-software (OpenFOAM). Predicted results using the “Spalart-Allmaras” turbulence model are similar to the wind tunnel data published in the literature. Moreover, global parameters such as dimensionless lift and drag coefficients were calculated. Finally, this research encourages VAWT studies under wind turbulent flows in order to achieve the best use of natural resources in Colombia.

**Keywords:** aerodynamic, airfoil, blades, wind turbine, turbulence modelling, CFD.

### Introduction

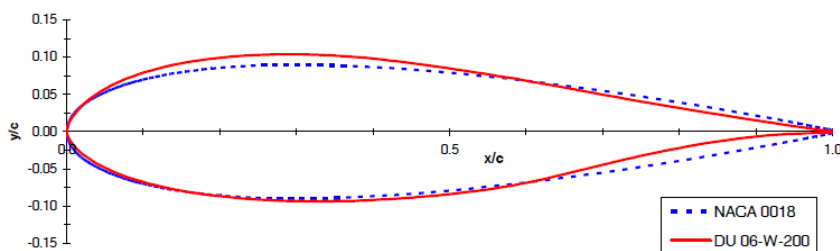
According to UPME\*, the energy demand in Colombia in the industrial and residential sectors is going to duplicate in the next 40 years. Furthermore, 75% of the energy demand is supplied by hydroelectric power that has a negative impact on the environment [UPME (2014)]. Nowadays, there is only one wind farm that produces 19.5 MW in the country [UPME (2014)]. Nevertheless, inside Colombia wind direction change constantly due to its topography and there is the need of developing wind power solutions capable to use this variant resource in order to fulfill Colombian energy demand. The 52% of Colombian territory is not connected to the local grid and according to the government [Olaya (2012)], more than 50% of 2012 resources were invested in maintenance, installation and adequacy of outdoor lighting for rural and urban locations in Santander. This high outgoings and the importance of preserve local environment stimulates alternative energy solutions that reduce expenses of local electricity. This research is the first one in analyzing wind power density at the “Cañón del Chicamocha”. It also determines the feasibility installation of Vertical Axis Wind Turbines (VAWT) at the place.

\* The Mining and Energy Planning Unit of Colombia, (Unidad de Planeación Minero-Energética).

The performance of a VAWT relies principally on its airfoil, which obtain lift or drag forces necessary to produce high efficient torque at its edge. Airfoil design and selection is an important task that depends on three main parts: wind flow conditions, airfoil shape and modelling.

Currently, Darrieus (\*) VAWT (based on lift aerodynamic force) uses the commercial NACA0018 (\*\*) airfoil. In a previous research [Claessens (2006)], a new airfoil for these turbines is developed. He presents the DU06W200 airfoil, which improves the performance of the NACA0018. The research made experiments and modelling of the airfoil based on Blade Element Momentum (BEM) theory. After that, [Castelli et al. (2012)] compared the airfoils DU06W200 and NACA0021. He evaluated their energy performance and aerodynamic forces that interact between the three wind turbine blades. The analysis was done with the commercial CFD software “Fluent 6.3.26” (\*\*\*) at wind speed of 9 [m/s] (much higher than the found at the “Cañón del Chicamocha”) under three different turbulence models:  $k-\omega$  SST,  $k-\epsilon$  Realizable and Spalart-Allmaras. [Chandrala et al. (2013)] analyzed the NACA0018 airfoil for horizontal wind turbines at wind speed of 32 [m/s]. He used the commercial software “ANSYS CFX” with the standard “k-E” turbulence model. Finally, [Boutillier (2011)] developed an experimental investigation of transition over the NACA0018 airfoil at a Reynolds number of  $1 \times 10^5$ . He focused the work specifically at the shear layer.

This research determines experimentally the feasibility installation of VAWT at “Cañón del Chicamocha”. Furthermore, the research is centered in the analysis of the airfoils DU06W200 and NACA0018 under the wind flow conditions at “Cañón del Chicamocha”. The study uses CFD through the free software “OpenFOAM” (\*\*\*\*) and the one equation turbulence RANS model developed by Spalart-Allmaras [NASA]. The difference between the airfoils can be seen at figure 1.



**Figure 1. Comparison between airfoils NACA0018 and DU06W200**  
[Claessens (2006)]

(\*) Darrieus VAWT’s consists of a number of curved airfoil blades mounted on a vertical rotating shaft or framework. The curvature of the blades allows the blade to be stressed only in tension at high rotating speeds [[http://en.wikipedia.org/wiki/Darrieus wind turbine/](http://en.wikipedia.org/wiki/Darrieus_wind_turbine/)].

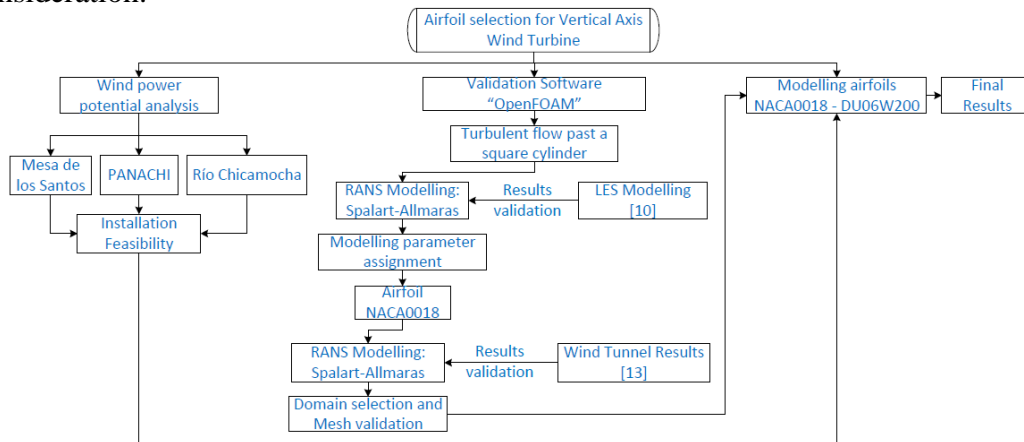
(\*\*) The NACA airfoils are airfoil shapes for aircraft wings developed by the National Advisory Committee for Aeronautics (NACA). The shape of the NACA airfoils is described using a series of digits following the word "NACA". The parameters in the numerical code can be entered into equations to precisely generate the cross-section of the airfoil and calculate its properties. [[http://en.wikipedia.org/wiki/NACA airfoil/](http://en.wikipedia.org/wiki/NACA_airfoil/)].

(\*\*\*) ANSYS Fluent is a commercial software that contains the broad physical modeling capabilities needed to model flow, turbulence, heat transfer for industrial applications. [<http://www.ansys.com/Products/Simulation+Technology/Fluid+Dynamics/Fluid+Dynamics+Products/ANSYS+Fluent/>].

(\*\*\*\*) OpenFOAM is a free, open source CFD software package developed by OpenCFD Ltd at ESI Group and distributed by the OpenFOAM Foundation. It has a large user base across most areas of engineering and science, from both commercial and academic organizations [<http://www.openfoam.com/>].

## 2. Methodology

Figure 2 shows the research scheme applied in this work. As it can be seen, it has two main components: an experimental, with the purpose of finding out wind potential, and the other main component, aimed to establish the best airfoil for the site under consideration.



**Figure 2. Research scheme.**

### 2.1. Installation feasibility of VAWT's at "Cañón del Chicamocha"

Wind flows properties at "Chicamocha Canyon" were collected for three years, since 2009, by "Parque Nacional del Chicamocha-PANACHI" [Panachi (2012)] three times per day at three different sites: "La Mesa de los Santos", Chicamocha River and "PANACHI". By using the collected data, the feasibility installation of VAWT was analyzed. Furthermore, wind flow conditions were established in order to realize a reliable aerodynamic research of the airfoils NACA0018 and DU06W200.

Wind flow direction varies according to the topography of the selected location and VAWT does not need to be pointed towards the wind to be effective, therefore, they can be used on the sites with highly variable wind direction. Moreover, its structural and esthetic principles have improved power in turbulent flows [Castelli et al. (2012)] [Manwell et al. (2010)].

Starting from conservation of mass and the continuity equation principle, the wind potential was analyzed. The air flow as a function from air density ( $\rho$ ) and its velocity ( $U$ ) that pass through the swept area [ $A$ ] of the turbine, can be expressed as:

$$\frac{dm}{dt} = \rho * A * U \quad (1)$$

Wind power potential can be expressed as kinetic energy in delta time:

$$\frac{P}{A} = \frac{1}{2} * \rho * U^3 \quad (2)$$

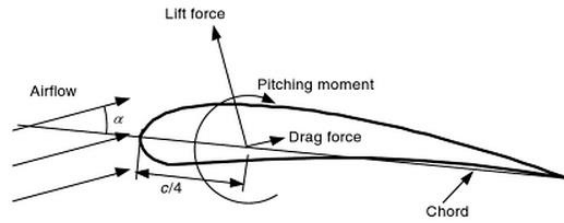
This research uses [Manwell et al. (2010)] criteria that establishes how significant wind power potential is at a selected location:

**Table 1. Wind power potential criteria [Manwell et al. (2010)].**

$P/A < 100 \text{ W/m}^2$	Poor
$P/A \approx 400 \text{ W/m}^2$	Good
$P/A > 700 \text{ W/m}^2$	Excellent

## 2.2. Aerodynamic study

Wind flow incidence over the airfoil generates force distribution over its surface, where result decomposes into lift, drag and pitching momentum as figure 3 shows.



**Figure 3. Forces and Momentum over an airfoil,  $\alpha$  angle of attack  $c$ , chord length. [Manwell et al. (2010)]**

These parameters are define dimensionless through its coefficients as:

$$\text{Lift Coefficient:} \quad C_l = \frac{L/l}{\frac{1}{2}\rho U^2 c} \quad (3)$$

$$\text{Drag Coefficient:} \quad C_d = \frac{D/l}{\frac{1}{2}\rho U^2 c} \quad (4)$$

$$\text{Pitching Momentum:} \quad C_m = \frac{M}{\frac{1}{2}\rho U^2 A c} \quad (5)$$

Dimensionless parameters analyze performance of airfoils under wind flow conditions of the location and through “Dynamic Similitude” concept. The mentioned parameters are function of attack angle, Mach and Reynolds number.

## 2.3. CFD airfoil studies

Fluid Mechanics governing equations in turbulent flow do not allow general analytic solution. Its study is approached from testing, dimensionless reasoning or mathematical simplified analysis. CFD is a new analysis technique for solving these equations and its final task is the resolution of mathematical equations that express turbulent fluid flow laws through software usage [Fernández (2012)].

Nowadays, different researches are improving airfoils performance of wind turbines through wind tunnel tests and theoretical studies. Nevertheless, these efforts are time consuming and need high technology laboratories (costly). Wind turbines CFD modelling provides affordable solutions to analyze aerodynamic issues of the airfoil and offers unlimited information over the domain since each cell modelling equals a tiny sensor for measuring each flow variable.

Governing equations in turbulent flow can be solved directly (DNS) or indirectly (INS). DNS solves each temporal and spatial fluctuation, meaning that high Reynolds numbers ( $10^5$ ) require high-density meshing and short temporal steps for solving each flow vortex [Fernández (2012)], what demands high computational costs. Hence, indirect equations solving is the feasible road for problem solving. That solution has two options that employ a turbulence model for closing the equations system when a temporal average (RANS) or a spatial filtering (LES) is applied. The research at hand uses RANS and “Spalart-Allmaras” one equation model to close the equations system (Equations 6-12). This model is selected due to it was developed to analyze turbulent flows around aerodynamic surfaces. In brief, analyzed cases were make under the

following conditions: incompressible, steady and turbulent flow, therefore solved equations are stationary.

Conservation of Mass:

$$\frac{\partial U_i}{\partial x_i} = 0 \quad (6)$$

Conservation of Momentum:

$$\rho \frac{\partial U_i}{\partial t} + \rho U_j \frac{\partial U_i}{\partial x_j} = -\frac{\partial P}{\partial x_i} + \frac{\partial}{\partial x_j} (2\mu S_{ji} - \overline{\rho u'_i u'_j}) \quad (7)$$

Mean strain-rate tensor:

$$s_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \quad (8)$$

Reynolds stress tensor:

$$\tau_{ij} = -\overline{\rho u'_i u'_j} \approx \mu_t \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) - \frac{2}{3} \rho k S_{ij} \quad (9)$$

Kinetic turbulent energy:

$$k = \frac{1}{2} (\overline{\rho u'^2} + \overline{\rho v'^2} + \overline{\rho w'^2}) \quad (10)$$

Turbulence eddy viscosity:

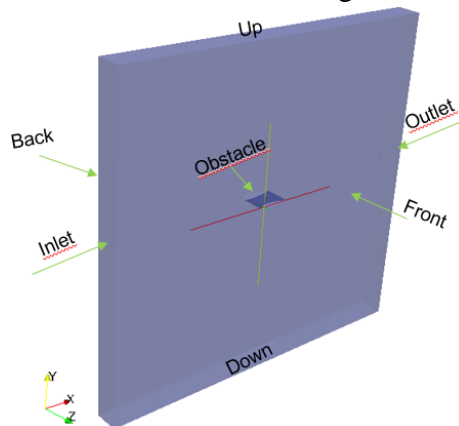
$$\mu_t = \rho \hat{v} f_{v1} \quad (11)$$

Spalart-Allmaras Model:

$$\begin{aligned} \frac{\partial \hat{v}}{\partial t} + u_j \frac{\partial \hat{v}}{\partial x_j} = & c_{b1} (1 - f_{t2}) \hat{S} \hat{v} - \left[ c_{w1} f_w - \frac{c_{b1}}{k^2} f_{t2} \right] \left( \frac{\hat{v}}{d} \right)^2 \\ & + \frac{1}{\sigma} \left[ \frac{\partial}{\partial x_j} \left( (v + \hat{v}) \frac{\partial \hat{v}}{\partial x_j} \right) + c_{b2} \frac{\partial \hat{v}}{\partial x_i} \frac{\partial \hat{v}}{\partial x_i} \right] \end{aligned} \quad (12)$$

#### 2.4. OpenFOAM Modelling description

According to the properties of the flow cases, velocity-pressure coupling was solved through SIMPLE (Semi Implicit Method for Pressure linked Equations) algorithm. The research analyze fluid flow over an airfoil. The domain splits into patches where boundary conditions are established as is shown in figure 4.



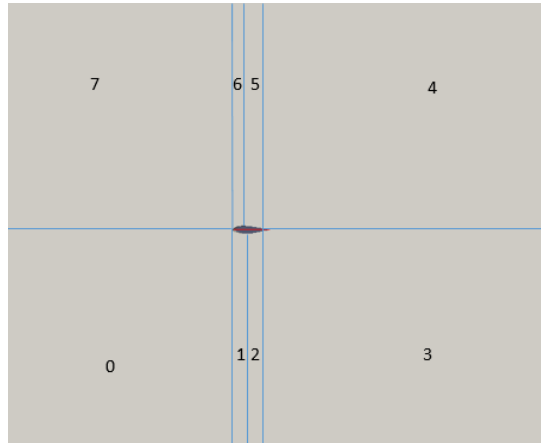
**Figure 4. Patches distribution over the domain.**

The domain distribution was the same for every modelling developed. There were only variations of the geometry at the “obstacle” patch and incidence flow according to the problem studied.

Previously modelling the airfoils under wind conditions from selected place, i.e. “Cañón del Chicamocha”, several tests were made and compared with published results in order to verify the selected mathematical model and ensure a good implementation of the software.

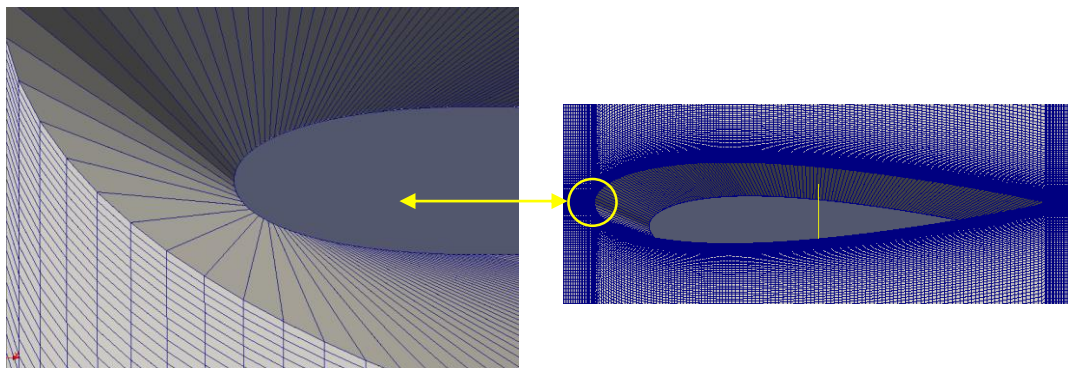
*2.5. Domain definition of the Airfoil*

The appropriate domain definition ensure a develop flow downstream the airfoil and assures independence of the imposed boundary conditions. Therefore, three different domains were studied for the NACA0018 airfoil, which were implemented by [Cao (2011)], [Joukowski (2013)] and [Claessens (2006)], named hereafter as 1, 2 and 3, respectively. Domains 1 and 2 represents boundaries as function of the chord length, while the domain 3 obeys to spatial ratio of wind tunnel. Each modelling was make with an 8-block domain as is shown in figure 5.



**Figure 5. Domain shaped blocks in OpenFOAM.**

Simulations were performed under Reynolds number of  $3 * 10^5$ , chord length of 60[cm], attack angle of  $0^\circ$  . Furthermore, different meshes were applied in order to study its influence. The main task is find the domain that under the same flow conditions presents less computational cost, more accurate results and  $y^+$  acceptable values. Figure 6 shows one of the analyzed grids in detail.



**Figure 6. Mesh detail around NACA0018 airfoil.**

After simulations were done, the three domains gave us these conclusions:

- Domain 1, showed convergence issues by using a high density mesh that decrease  $y +$  values. Furthermore, computational cost was high. These facts made harder the study of parameters variation.
- Domain 2, allows the developing of the flow downstream at the airfoil, and its shorter dimensions reduced computational costs. Therefore, the meshing density and the  $y +$  analysis were easier to run.
- Domain 3 gave a good flow resolution and its computational cost was less than domain 1. Nevertheless, wind flow direction could not be greater than  $5^\circ$  by the dimensions of the domain height.

Consequently, the Domain 2 was chosen to perform the simulation of the airfoils. Its geometry is formed by 5 chord lengths distance from the leading edge to the top face, 5.5 chord lengths to the bottom face, 5 chord lengths to the upwind face (inlet) and 6 chord lengths to the downwind face (outlet). The boundary conditions used at the patches shown in figure 4 are shown in table 2.

**Table 2. Boundary conditions at the Patches.**

Inlet	Patch
Outlet	Patch
FrontandBack	Empty
obstacle	Wall
upanddown	patch

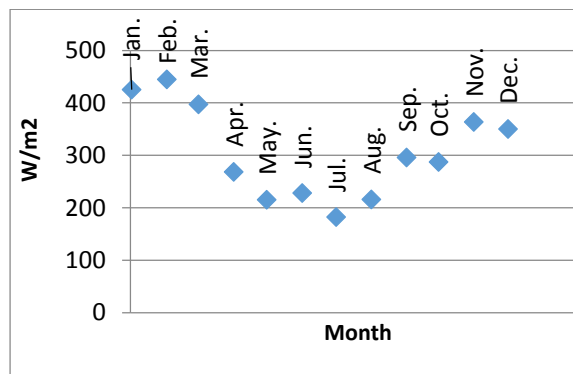
### 3. Discussion and analysis of results

#### 3.1. Wind power density analysis

Wind speed data collection from “Cañón del Chicamocha” was made with cup anemometers installed in three main parts of the location, i.e. “Mesa de los Santos”, Chicamocha river and “PANACHI”, above cableway supports with height of 20 [m] each.

##### 3.1.1. “Mesa de los Santos” wind speed measuring

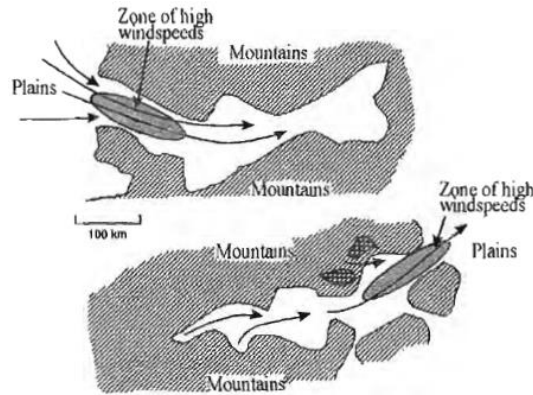
Average wind speed velocity has values from 5 to 7 [m/s] per year, which gives a maximum wind power density of  $450 [W/m^2]$  on February and minimum of  $180 [W/m^2]$  on July. Minimum values were find between May and August as figure 7 shows.



**Figure 7. “Mesa de los Santos” wind power density per month.**

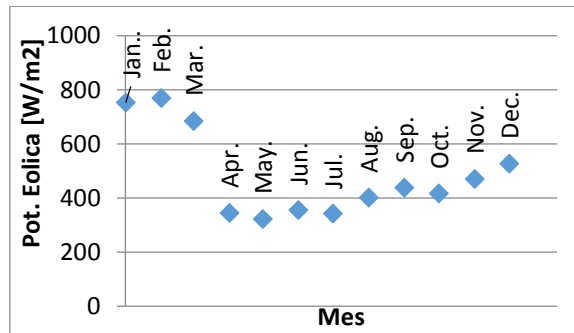
3.1.2. *Chicamocha river wind speed measuring*

At this location, wind flow accelerates due to mountains that surrounds the river, which acts as a diffusor redirecting the wind to a smaller section as figure 8 shows.



**Figure 8. Wind speed increase due to canalization of prevailing winds from the mountain [Manwell et al. (2010)].**

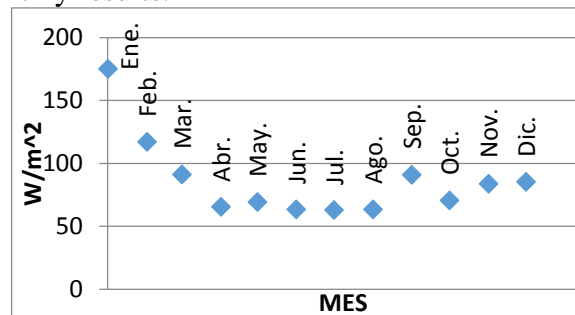
This phenomena effect is confirmed by wind speeds observed at the location, i.e. up to 8.75[m/s]. Therefore, wind power density values were up to 770[W/m<sup>2</sup>]. Furthermore, wind speed variations were higher, as figure 9 shows.



**Figure 9. Wind power density at Chicamocha River per month.**

3.1.3. *“PANACHI” measurement*

Lastly, wind flow behavior was analyzed in PANACHI, where maximum wind speed value was find in January, i.e. 5.65 [m/s], with a wind energy density of 180[W/m<sup>2</sup>]. Figure 10 shows monthly results.



**Figure 10. Average wind power density at PANACHI per month.**



Table 3 summarize the annual average wind speed and wind power density of the three locations. It is shown that the greatest and feasible place for VAWT locations is at Chicamocha River.

**Table 3. Wind power potential at “Cañón del Chicamocha”**

Place	Annual average wind speed [m/s]	Standard Deviation	Annual average wind power density [ $W/m^2$ ]
Mesa de los santos	5.974	0.736	306.188
Chicamocha River	6.928	1.084	485.115
Parque Nacional del Chicamocha (PANACHI)	4.325	0.536	86.643

### 3.2. Validation and verification

Obtained numerical results are validated by the lift and drag coefficients comparison against wind tunnel data from [Claessens (2006)]. Three different angles of attack of  $0^\circ$ ,  $10^\circ$  and  $20^\circ$  were simulated for a chord length of  $c=0.25$  [m], and Reynolds number of  $3 \times 10^5$ . Results are shown in tables 4, 5 and 6, respectively.

**Table 4.  $y^+$  results for different angles of attack.**

$\alpha$	$y^+_{min}$	$y^+_{max}$	$y^+_{mean}$
$0^\circ$	1.614	15.979	10.775
$10^\circ$	0.636	16.436	9.58
$20^\circ$	0.922	18.456	8.222

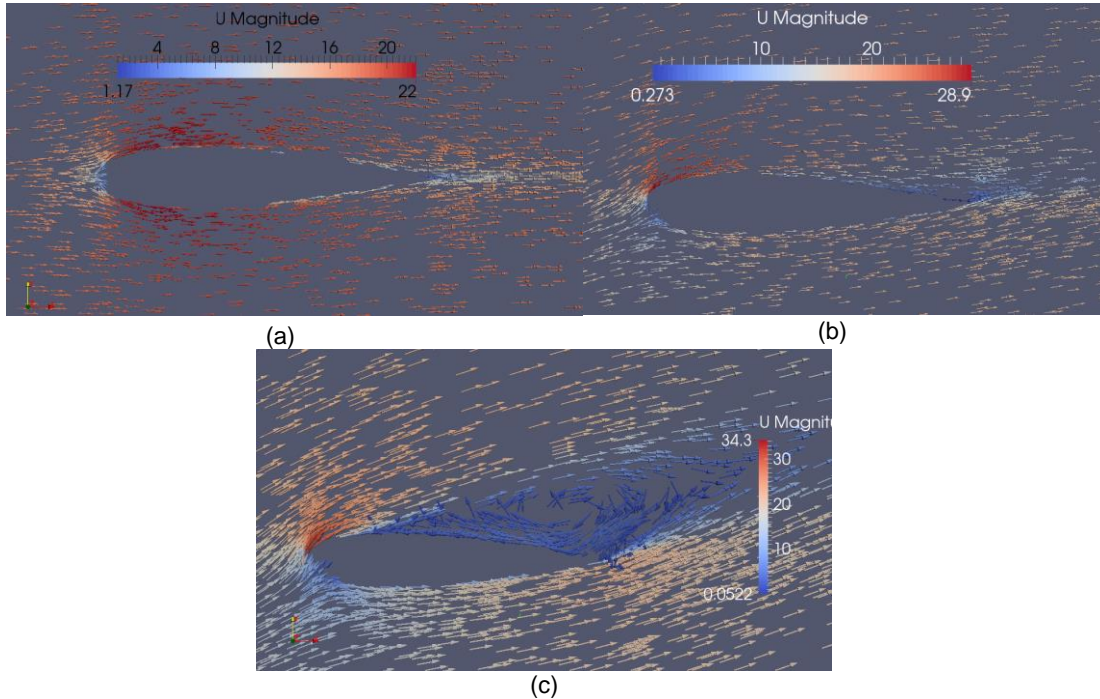
**Table 5. Lift coefficients of the airfoil NACA0018 at different angles of attack.**

$\alpha$	$C_{l_{simulacion}}$	$C_{l_{túnel}}$	%Error
$0^\circ$	0.0204	0.0193	5.7
$10^\circ$	0.664	0.803	17.31
$20^\circ$	0.769	0.615	25

**Table 6. Drag coefficients of the airfoil NACA0018 at different angles of attack.**

$\alpha$	$C_{d_{simulacion}}$	$C_{d_{túnel}}$	%Error
$0^\circ$	0.0379	0.0324	16.97
$10^\circ$	0.0646	0.059	9.49
$20^\circ$	0.206	0.243	15.22

According to these results, the accuracy of the implemented turbulence model is accepted. Its maximum variation was of 17% and minimum of 6% in comparison to the wind tunnel tests. The best performance of the airfoil was found at 10°, where the lift and drag coefficients ratio is the greatest, 10.3 approximately. Wind flow behavior of the airfoil NACA0018 is shown at figure 11.



**Figure 11. Wind velocity vectors over the airfoil at different angles of attack: (a) 0°, (b) 10°, and (c) 20°**

As figure 11 shows, during an angle of attack of 0° a greater acceleration of the flow is produced around the airfoil. Whereas angles of attack lower than 20° the vortex generation is not perceivable at the trailing edge.

### 3.3. Analysis of the airfoils NACA0018 and DU06W200

#### 3.3.1. Airfoil performance analysis for Reynolds numbers between $2 \times 10^5$ and $3.4 \times 10^5$

In order to analyze the finite volume method accuracy under Spalart-Allmaras turbulence model, simulations were carried out for the Reynolds number presented by [Claessens (2006)]. Its modelling conditions are:

- ✓ Angle of attack of 10°
- ✓ Chord length of 0.25[m].
- ✓ Fluid properties as follows:

	<b>Re <math>2 \times 10^5</math></b>	<b>Re <math>3 \times 10^5</math></b>	<b>Re <math>3.4 \times 10^5</math></b>
<b>Wind Speed[m/s]</b>	7.5	18	20
$\tilde{\nu}$	0.0107	0.0276	0.0306
$\nu$	0.0107	0.0276	0.0306

A drag polar (\*) was used in order to analyze the ratio between the drag and lift coefficients. Its values are shown at table 7.

**Table 7. Lift and drag coefficients from the airfoils NACA0018 and DU06W200 under different Reynolds numbers.**

Reynolds Number	NACA0018		DU06W200	
	Cd	Cl	Cd	Cl
200000	0.08	0.707	0.085	0.876
300000	0.065	0.664	0.075	0.884
340000	0.07	0.687	0.077	0.926

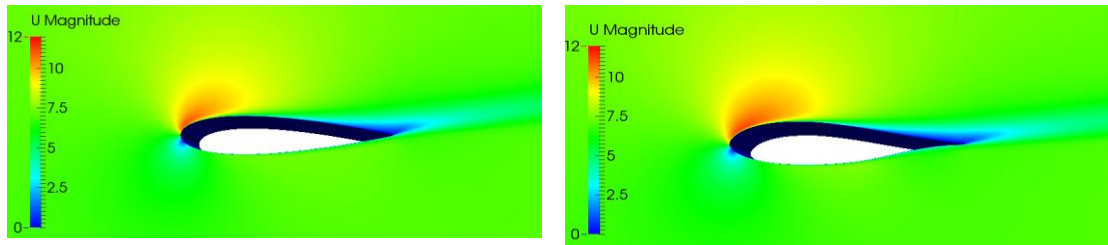
As it is shown, for the same Reynolds number the lift coefficients of the DU06W200 airfoil overcome in 23.3% the ones from NACA0018 airfoil. Furthermore, as Reynolds number increases the lift coefficient increases as well. Finally, as lift coefficients from DU06W200 airfoil are greater than NACA ones, [Claessens (2006)] results are confirmed.

3.3.2. Airfoils modelling under “Cañón del Chicamocha” wind flow

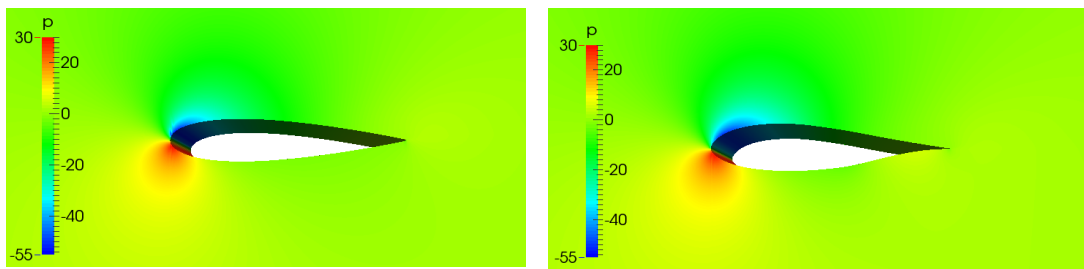
Parameters and conditions:

- ✓ Wind speed at the entrance is the highest found at the selected location.
- ✓ Steady state regime.
- ✓ Angle of attack of 10°.
- ✓  $c=0.25[m]$  and  $Re = 1.19 \times 10^5$ .

Predicted results are shown in figures 12 and 13.



**Figure 12. Wind speed average magnitude: (a) NACA0018 and (b) DU06W200**



**Figure 13. Pressure distribution at airfoils (a) NACA0018 and (b) DU06W200.**

(\*) The Drag Polar is the relationship between the lift and the drag coefficient.

As it is shown in figure 12, the wind speed at the leading edge of the airfoil DU06W200 is greater than that for the NACA0018. This effect shows the design optimization of the airfoil developed by [Claessens (2006)]. Therefore, the cambered airfoil DU06W200 generates a high-pressure peak followed by a sharp fall of its values, as figure 13 shows. This phenomenon generates turbulent flow in advance quickly since the boundary layer cannot follow this pressure increase [Claessens (2006)].

On the other hand, simulation results show that NACA0018 separation bubbles at the leading edge is greater than at DU06W200 airfoil. The increase of the laminar separation bubbles leads to develop a longer wake, which is directly linked to noise production from the blades [Claessens (2006)].

The effects of these airfoils characteristics are represented by the lift and drag coefficients, which are presented in table 8.

**Table 8. Lift and drag coefficients of the airfoils NACA0018 and DU06W200 under “Cañón del Chicamocha” wind flow**

AIRFOIL	Cl	Cd
NACA0018	0.707	0.0801
DU06W200	0.876	0.0853

The results show a 20% increase of the lift coefficient generate by DU06W200 airfoil, at cost of a 6% increase at its drag coefficient. Thus, the airfoil DU06W200 has a better performance than the airfoil NACA0018 under feasible wind flow conditions of the location, i.e. Chicamocha River.

**Conclusions**

- ✓ The implementation of Vertical Axis Wind Turbines is feasible at Chicamocha river, where average wind speed is 7 [m/s] and average wind power density is 485.115[W/m<sup>2</sup>].
- ✓ An airfoil domain with 5 chord lengths distance from the leading edge to the top face, 5.5 chord lengths to the bottom face, 5 chord lengths to the upwind face (inlet) and 6 chord lengths to the downwind face (outlet), allows a proper development of the wake using the Spalart-Allmaras turbulence model.
- ✓ Contrasted to the airfoil NACA0018, the new airfoil DU06W200 presented this results:
  - It has a lift coefficient increase of 20% with the same drag loses for NACA0018 at Chicamocha River wind speed.
  - There is a less wind flow recirculation at the trailing edge, which confirmed noise reduction of the DU06W200 airfoil.
  - Lift coefficients for the DU06W200 airfoil at Reynolds numbers between 2x10<sup>5</sup> and 3.4x10<sup>4</sup>, are 23% greater than the NACA0018 ones.
  - Therefore, the most appropriate airfoil for blades of a vertical axis wind turbines under wind flows of Chicamocha’s canyon is the DU06W200.

## Acknowledgments

At first, to the “Universidad Industrial de Santander” by supporting financially the research through the 1358 project from the “VIE”. Furthermore, to “Parque Nacional del Chicamocha (PANACHI)” represented by Maria Eva Bermúdez, whom allows the use of wind speed flow data from the location of interest.

## References

- UPME. Ponencia “Colombia: tendencias a largo plazo del sector de energía en Colombia”. Cartagena. 2014. p. 26.
- OLAYA, Leon. Informe de gestión 2012. Secretaria de Infraestructura, alcaldía de Bucaramanga. 2012. p. 47.
- CLAESSENS, Maarten. The design and testing of airfoils for application in small vertical axis wind turbine. Master of Science thesis. Aerospace engineering. The Netherlands-Delft: Delft University of Technology (TU Delft), 2006. 137 p.
- CASTELLI, Marco; SIMIONI, Giulia and BENINI, Ernesto. Numerical analysis of the influence or airfoil asymmetry on VAWT performance. World academy of Science, Engineering and Technology. 61, 2012.
- CHANDRALA, Monir; CHOUBEY, Abhishek; GUPTA, Bharat. CFD Analysis of horizontal axis wind turbine blade for optimum value of power. En: International journal of energy and environment, 2013, Volume 4, Issue 5,. p. 825-834.
- BOUTILIER, Michael. Experimental investigation of transition over a NACA0018 airfoil at a low Reynolds number. Master thesis of applied science in mechanical engineering. Canadá, Ontario: University of Waterloo. 2011, 237 p.
- PANACHI. Datos de la velocidad del viento en el Parque Nacional del Chicamocha, La punta de la mesa de los santos y el río Chicamocha: 2012.
- MANWELL, James; MCGOWAN, Jon and ROGERS, Anthony, Wind Energy Explained [Energía eólica explicada], Great Britain-Chippenham-Wiltshire: Wiley, 2010. 689 p.
- FERNÁNDEZ, Jesus. Técnicas numéricas en ingeniería de fluidos: Introducción a la dinámica computacional de fluidos (CFD) por medio del método de volúmenes finitos, Barcelona: Reverté, 2012. 386 p.
- CAO, Han. Aerodynamic analysis of small horizontal axis wind turbine blades by using 2D and 3D CFD modelling. Master Thesis: Preston, Inglaterra. University of the central Lancashire. 2011, p. 34.
- JOUKOWSKI, Airfoil at different Viscosities, 2013.
- NASA: <http://turbmodels.larc.nasa.gov/spalart.html>

## Structural Mechanics Optimization of the AISHa Ion Source

†\*F. Noto<sup>1</sup>, L. Andò<sup>1</sup>, L. Celona<sup>1</sup>, F. Chines<sup>1</sup>, G. Ciavola<sup>1,3</sup>, G. Costa<sup>1</sup>, S. Gammino<sup>1</sup>, O. Leonardi<sup>1</sup>, S. Marletta<sup>1</sup>, D. Nicolosi<sup>1</sup>, G. Torrasi<sup>1,2</sup>

<sup>1</sup> INFN - Laboratori Nazionali del Sud, Via S. Sofia 62, 95125 Catania, Italy

<sup>2</sup>Università Mediterranea di Reggio Calabria, Dipartimento di Ingegneria dell'Informazione, delle Infrastrutture e dell'Energia Sostenibile, Via Graziella, 89100 Reggio Calabria, Italy;

<sup>3</sup> CNAO Foundation, Strada Campeggi, 53, Pavia PV, Italia

\*Presenting author: francesco.noto@lns.infn.it

†Corresponding author: [francesco.noto@lns.infn.it](mailto:francesco.noto@lns.infn.it)

### Introduction

Different facilities for hadron therapy have been built or designed in the recent past and Italy is present in the field either with synchrotron-based and with cyclotron-based facilities. For both types of accelerators the availability of high brightness multiply charged ion beams is essential and R&D efforts in this subject are increasing. In order to cope the stringent requests of a hospital environment at INFN-LNS a new ion source called AISHa is under realization exploiting all the knowledge acquired in last decades in the ion source design and realization.

Therefore, the AISHa source has been designed by keeping in mind the typical requirements of hospital-based facilities, where the minimization of the mean time between failures (MTBF) is a key point together with the maintenance operations, which should be fast and easy. Some critical parts of the body source, in particular the plasma chamber and the hexapole containment chamber, have been studied and optimized with FEM softwares.

In the paper, the entire structure will be described, with a particular attention to thermo-mechanical simulations of the plasma chamber and the electrostatic and structural simulation of the hexapole containment chamber.

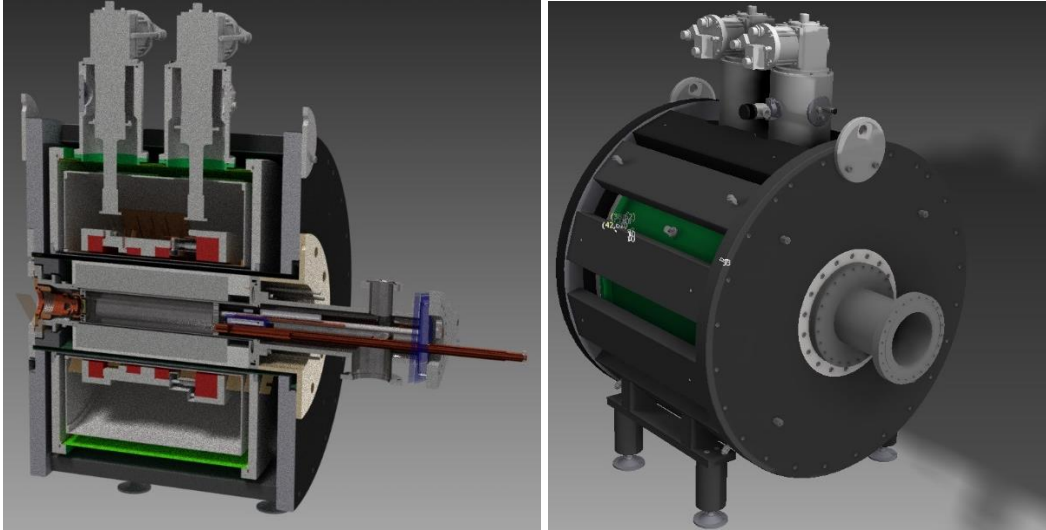
**Keywords:** Computation, Hadron Therapy, Ion Source, FEM, CFD

### Introduction

The INFN-LNS has a broad experience in the design, construction and commissioning of ECR and microwave ion sources. In particular, since the '90s two sources for highly charged ion beam production, named SERSE and CAESAR, are operational to provide the heavy ions to be accelerated by the LNS K-800 Superconducting Cyclotron [1,2]The AISHa ion source has been designed by taking into account the typical requirements of hospital-based facilities, where the minimization of the mean time between failures (MTBF) is a key point together with the maintenance operations which should be fast and easy. [1,2]Therefore, a so called 3rd generation ECR ion source is not suitable, being quite complex for unskilled operators.

The new AISHa source is designed to be an intermediate step between the 2nd generation ECRIS (unable to provide the requested current and/or brightness) and the 3rd generation ECRIS [1] (too complex and expensive). It is intended to be a multipurpose device, operating at 18 GHz, in order to achieve higher plasma densities. It should provide enough versatility for future needs of the hadron therapy, including the ability to run at larger microwave power to produce different species and highly charged ion beams. At the same time, the electrical power to be installed for its operation will be kept below 50 kW. This demand implies also the simplification of all ancillary systems including an oven for metallic ion beams, which permits the

production of new beams for hadron therapy and for other applications.[2,3] The AISHa source (figure 1) is funded within the framework of the program of Sicilian Government named PO FESR 2007-2013 and a pool of Sicilian SME is associated with INFN for this project. The source is potentially interesting for any hadrontherapy center using heavy ions.



**Figure 1: Layout of the AISHa source**

### **Mechanical development**

In order to minimize the maintenance operations, the development of this new source involved some mechanical and structural improvements with respect to similar devices. In particular hereinafter the optimization of the hexapole containment chamber and of the source plasma chamber are described.

### **Use of COMSOL Multiphysics**

The Finite Element Method approximates a Partial Differential Equations problem with a discretization of the original problem based on a mesh, which is a partition of the geometry into small units of simple shape called mesh elements. The PDE method looks for a solution in the form of a piecewise polynomial function, each mesh element defining the domain for one “piece” of it. Such a piecewise polynomial function will be expressed as a linear combination of a finite set of predefined basis functions. Let us consider for example a 2-dimensional problem with a single dependent variable  $p(x,y)$ . The solving method implemented in this code is based on a mesh with quadratic triangular elements. The expression “quadratic elements” refers to the fact that on each mesh element the sought piecewise polynomial function  $p^*(x,y)$  is at most a quadratic polynomial. In this case, the solution is expressed as:

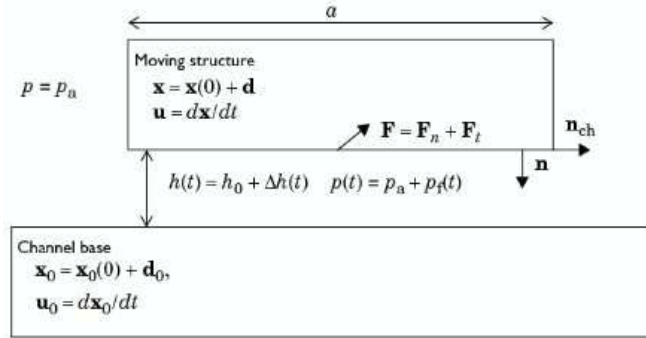
$$p(x, y) \cong p^*(x) = \sum_{i=1}^n p_i \varphi_i(x, y)$$

where  $i$  refers to a node of the mesh,  $p_i$  are the degrees of freedom,  $\varphi_i(x,y)$  are the basis functions and  $n$  is the total number of nodes, under the assumption that each triangle of the mesh possesses six nodes: three corner nodes and three mid-side nodes [4]. A basis function  $\varphi_i(x,y)$  has here the restriction to be a polynomial of degree at most 2 such that its value is 1 at node  $i$  and 0 at all other nodes [5]. The degree of freedom  $p_i$  is thus the value of  $p^*(x,y)$  at node  $i$ . The definition of the basis function associated to each node

of the mesh can be derived using for example a general method introduced by Silvester in 1969 [6].

### COMSOL's Thin-Film Flow Model for optimitation plasma chamber

All of COMSOL's single-phase fluid flow interfaces are based on the three fluid dynamics conservation equations known as the Navier-Stokes equations [4], concerning the conservation principles of mass, momentum and energy (Figure 2):



**Figure 2. Schematic diagram of the situation to which the thin-Film Flow Model applies**

The model that has been used in our simulations is the Thin-Film Flow Model [7] and belongs to the Computational Fluid Dynamics (CFD) module. The Thin-Film Flow Model can be used to model a thin channel fluid located between two moving structures, as schematized in Figure 2. The upper structure is referred to as the moving structure and the lower one as the channel base. Initially, both structures are surrounded by gas with a constant pressure  $p_a$  and the fluid can freely move into and out of the gap. Due to the movements of the structures, an additional and usually time-dependent pressure  $p_f$  appears in the gas inside the gap, which produces a normal force  $F_n$  on the structures. Also a viscous drag force  $F_t$  is created which resists the tangential movement of the structure. In the Thin-Film Flow Model, it is however assumed that:

- The film thickness  $h$  remains always very small with respect to the dimensions of the solid structures.
- The channel curvature is small.

Also the following assumptions are made:

- The inertial effects in the fluid are negligible compared to the viscous effects, thus the flow is laminar.
- The pressure  $p = p_a + p_f$  is constant over the film thickness  $h$ .
- The velocity profile over the film thickness is parabolic.
- The fluid is isothermal.

Given these assumptions, solving the fluid flow problem with the Navier-Stokes equations reduces to solving the following equation, called the Reynolds equation:

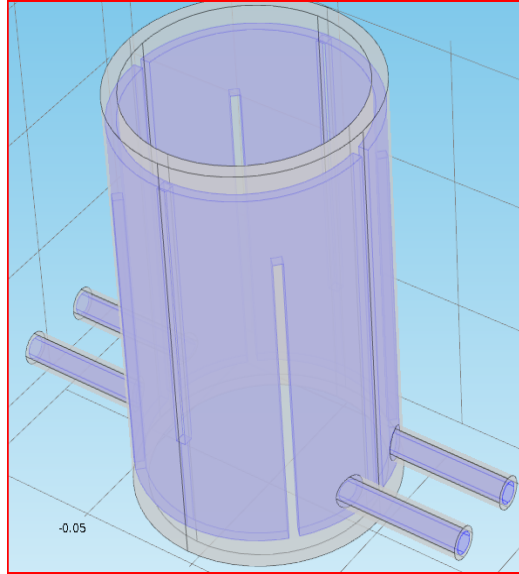
$$\frac{\partial(\rho h)}{\partial t} + \vec{\nabla}_{tg} \cdot (\rho h \vec{U}) - \rho(\vec{\nabla}_{tg} \Delta h_m \cdot u_m - \vec{\nabla}_{tg} \Delta h_b \cdot u_b) = 0$$

where  $\rho$  is the density,  $h = h_0 + \Delta h_m + \Delta h_b$  is the film thickness,  $t$  is the time,  $\vec{\nabla}_{tg}$  is a gradient computed only with the tangential derivatives along the channel boundaries,  $U$  is the mean film velocity,  $\Delta h_m$  and  $u_m$  are the normal displacement and the tangential velocity of the so-called "moving structure", respectively, and  $\Delta h_b$  and  $u_b$  are the normal displacement and the tangential velocity of the "channel base", respectively. The mean film velocity  $U$  is actually a function of the pressure  $p$ , the dynamic viscosity  $\mu$ , the film thickness  $h$ , the tangential velocities  $u_m$  and  $u_b$  of the solid structures and the relative flow rate function  $Q_{ch}$  that accounts for possible rarefied gas effects:



$$\vec{U} = \frac{\vec{\nabla}_{tg} p}{12\mu} h^2 Q_{ch} + \frac{u_m + u_b}{2}$$

Permanent water flow in the plasma chamber is required to provide the expected low temperature. [9] Therefore, the goal of our study was to optimize the design of the groove. This optimization of the particular was done starting from a model of the chamber used in other sources. The model was designed considering four cylinders:  $\phi = 92$  mm (chamber inner diameter),  $\phi = 94$  mm (water flow internal diameter),  $\phi = 102$  mm (water flow outside diameter) and  $\phi = 104$  mm (chamber external diameter), each divided in two half cylinders. For each of the two half-cylinders an input, an output and three septa were designed. The domains considered were two, one for water and one for the metal (AISI 316L and aluminum 3003-H18) (Figure 3).



**Figure 3. The first model, in gray the metal domain and in purple the water domain.**

For the inlet an initial velocity of the water has been determined by requiring a flow rate of 3 liters / minute ( $0.00005 \text{ m}^3/\text{s}$ ) and calculating the section of the input (the input to the diameter is 4 mm so the area is  $0.0000502 \text{ m}^2$ ). From the relation:

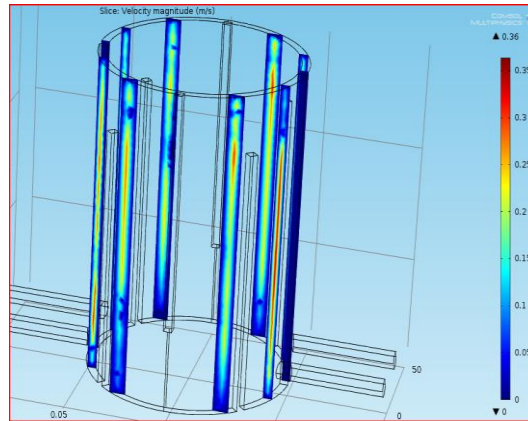
$$v = \frac{Q}{A} \quad (1)$$

by substituting the calculated values we obtain that the initial speed has to be 1 m/s. The initial water temperature was set at  $20^\circ \text{C}$ .

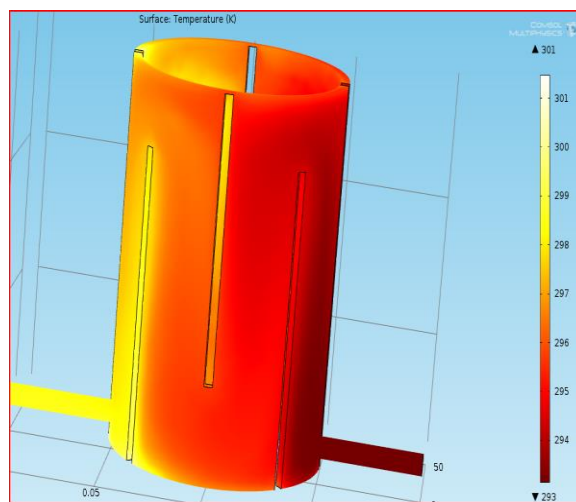
Concerning the mesh, for both domains we have chosen a tetrahedral mesh. In the domain of the water the choice was to use a mesh more dense than in the domain of the metal. The results obtained in the case of aluminum 3003-H18 show that the temperature reaches the maximum value of about  $300^\circ \text{K}$  at the output of water, that is, after the water has traveled the half cylinder, it is heated and has lost in part its cooling capacity. Nevertheless, this temperature is satisfactory, as we must have a maximum surface temperature of the lower chamber of  $50^\circ \text{C}$  in order not to damage the magnets that are in contact with it. The flow obtained is substantially laminar. Calculating the Reynolds number with the expression:

$$\text{Re} = \frac{v * D}{\mu} \quad (2)$$

and assigning the values:  $v=0.35$  m/s (the highest in the flow),  $D = 0.004$  mm and assuming the value  $1.01 * 10^{-6} \text{ m}^2/\text{s}$  for the kinematic viscosity  $\mu$  of the water, we get a Reynolds number equal to 1386 so we are in the field of laminar flow (Re less than 2300 is laminar flow) (Figures 4 and 5).

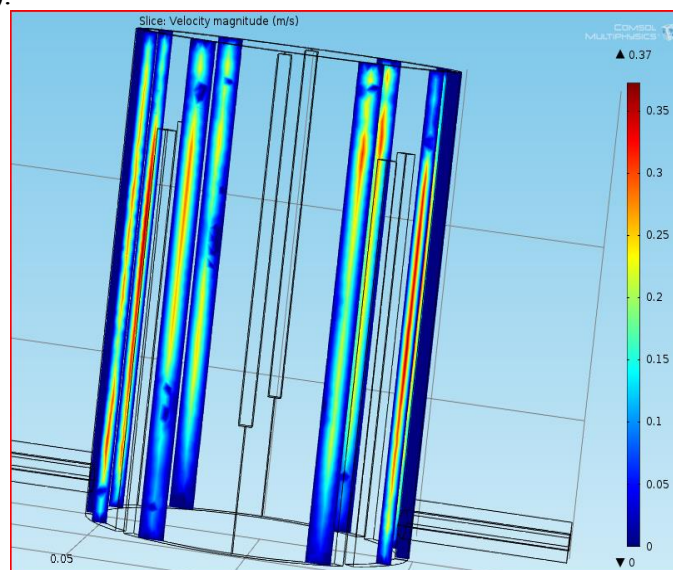


**Figure 4. Velocity speed in the aluminum case.**

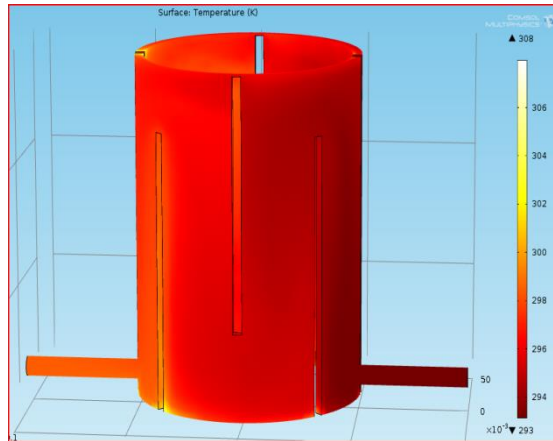


**Figure 5. Temperature in the aluminum case.**

Subsequently the simulation has been developed using the AISI 316L steel as metal material. The results obtained for the velocity of the water are similar to the case of aluminum, while the maximum temperature reached is lower than in the previous case (Figure 6 and 7).



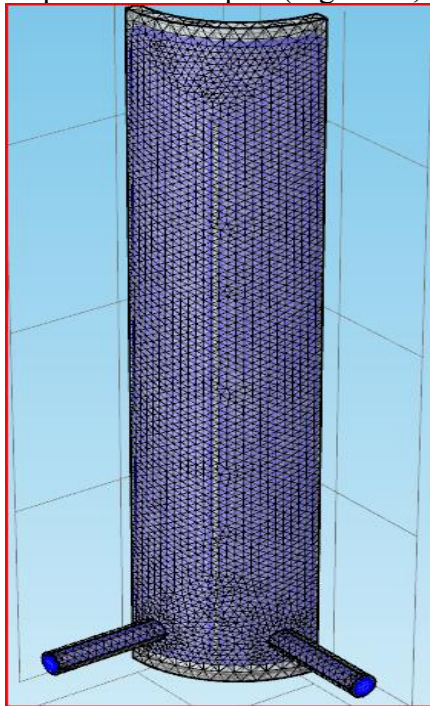
**Figure 6. Velocity speed in the aluminum case.**



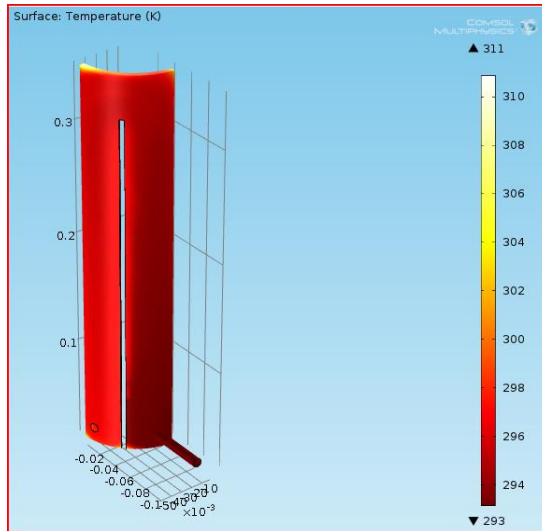
**Figure 7. Velocity speed in the aluminum case.**

This temperature decrease is due to the characteristics of the two materials: in fact the thermal conductivity of steel (14.6 W/m\*K), is lower than that of aluminum (155 W/m\*K); it is a measure of the ability of a material to transmit heat (i.e. the lower the value of  $k$ , the more insulating is the material).

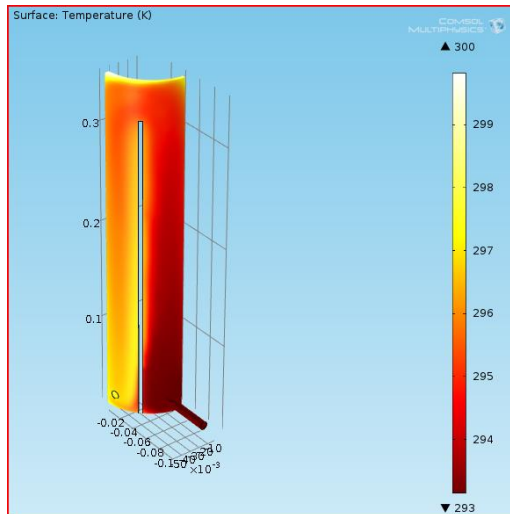
The first step has been developed using the COMSOL code, considering a length of the chamber of 155 mm. Then the simulation was performed for the actual size of the chamber, that is 655 mm, but as the plasma formation takes place within 360 mm it has been decided to simplify the simulation, and then the calculations, considering a total length of 360 mm. Another simplification was to divide the plasma chamber into 4 equal sectors each with an input and an output (Figures 8, 9, 10).



**Figure 8. The mesh in the new model**

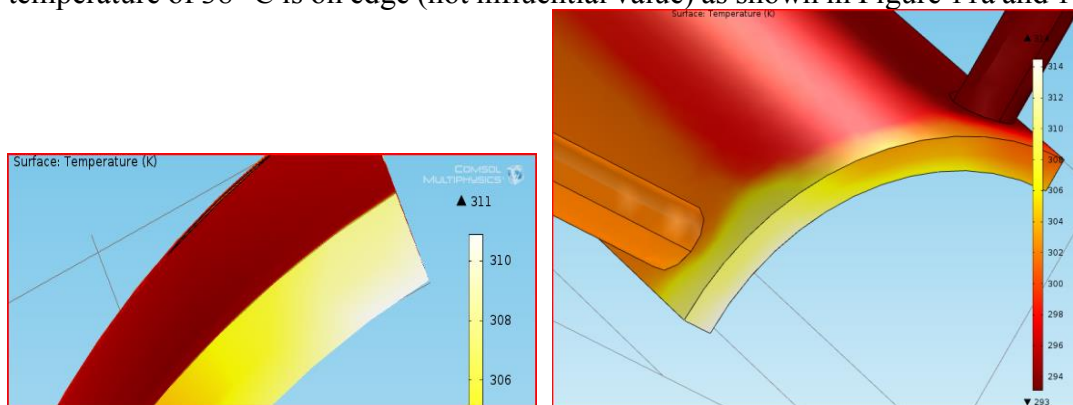


**Figure 9. The temperature values in the AISI 316L case.**



**Figure 10. The temperature values in the aluminum 3003-H18 case**

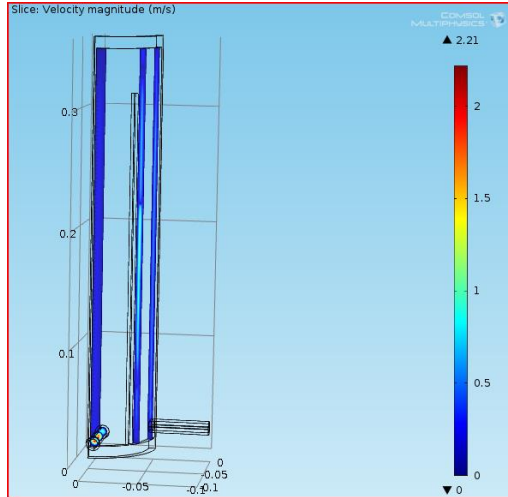
An interesting detail is that the  $D_t$  between a point near the entrance and a point near the water outlet is about  $4^\circ\text{C}$  and the average temperature is  $24^\circ\text{C}$ . The maximum temperature of  $38^\circ\text{C}$  is on edge (not influential value) as shown in Figure 11a and 11b.



**Figure 11a and 11b. The temperature edge.**

As in step 1, we have developed the same simulations in the case of another metal material, aluminum 3003-H18.

The results obtained for the water velocity are quite similar to each other (Figure 12).



**Figure 12. The water velocity values in the aluminum 3003-H18 case**

In the plasma chamber the speed values are on average equal to 0.5 m/s and therefore we are in laminar regime as the Reynolds number is about 2000. Pressure drop was calculated using the expressions of both the distributed and concentrated losses. The former is:

$$\Delta p = \rho * \xi * \frac{L}{D} * \frac{v^2}{2} \quad (3)$$

where  $\rho$  is the density of water,  $\xi$  is the friction factor calculated by means of the Moody (using the Reynolds number and the roughness of the material),  $L$  is the length of the chamber,  $D$  the diameter of the duct of water (assumed in this case a size of the water domain equal to 4 mm) and  $v$  the velocity of water (assumed an average speed of 0.5 m/s from the simulation): the calculation leads to a  $D_p$  of about 0.01 bar.

For concentrated losses the expression is:

$$\Delta p = \beta * \rho * \frac{v^2}{2} \quad (4)$$

where  $\rho$  is the water density,  $\beta$  is the coefficient of friction equal to 0.5 for the 90-degree elbow and  $v$  is the velocity of the water in different cases. The calculation leads to a  $D_p$  of about 0.01 bar to the entrance ( $v = 2$  m/s);  $D_p$  of about 0.002 bar for the output ( $v = 0.9$  m/s) and  $D_p$  of about 0.0006 bar in the elbows due to the septum ( $v = 0.5$  m/s): the sum of all the losses yields a loss of load of 0.023 bar.

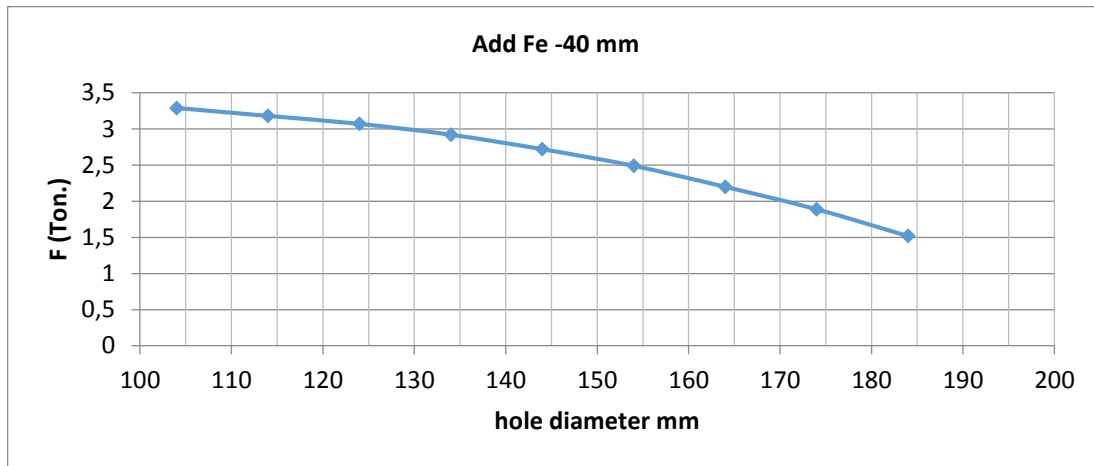
### **Minimization of the stray magnetic field in the extraction area.**

The AISHa magnetic field is provided by a set of four superconducting coils able to create a mirror field as shown in fig. 13 with respectively 2.7 T and 1.7 axial field at injection and extraction. The forces between the coils have been decreased to 30 tons for safe operation at maximum field, in this operating conditions different calculations have been performed to minimize the field in the extraction area to avoid unwanted effects on beam optics and to minimize sparks in the extraction gap.



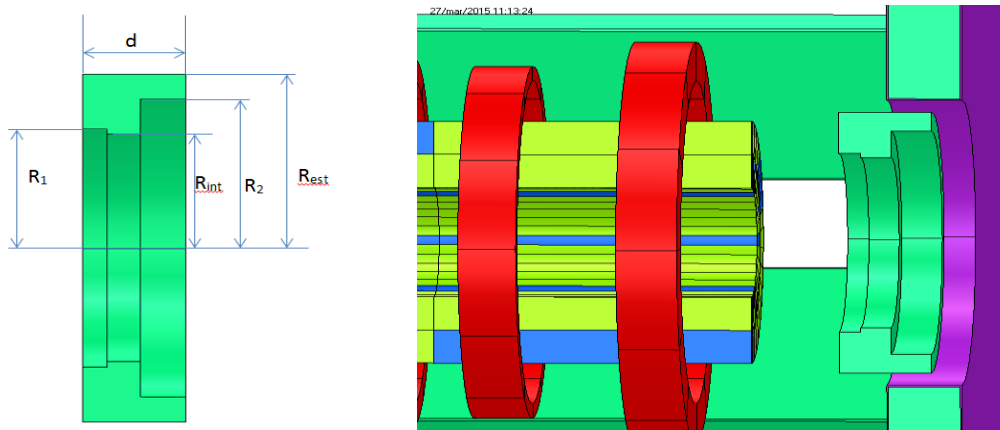
Iron Add -40 mm	Force (ton.)
$\phi_{int} = 104$ mm	3.29
$\phi_{int} = 114$ mm	3.18
$\phi_{int} = 124$ mm	3.07
$\phi_{int} = 134$ mm	2.92
$\phi_{int} = 144$ mm	2.72
$\phi_{int} = 154$ mm	2.49
$\phi_{int} = 164$ mm	2.20
$\phi_{int} = 174$ mm	1.89
$\phi_{int} = 184$ mm	1.52

**Table 2. Force variation function of the diameter variation.**



**Figure 15. Force variation function of the diameter variation.**

This configuration, albeit even if mechanically valid, was discarded due to static problems, in fact there discharges may occur due to the iron shape, therefore it was decided to change their profile (figure 16) and see the table 3.



**Figure 16a and 16b. New profile and iron parte in loco**

	d	R <sub>int</sub>	R <sub>1</sub>	R <sub>2</sub>	R <sub>est</sub>
Fe1	64	81	85	106	124
Fe2	74	81	85	106	124
Fe3	84	81	85	106	124
Fe4	94	81	85	106	124

**Table 3. Size variations of particular "added iron"**

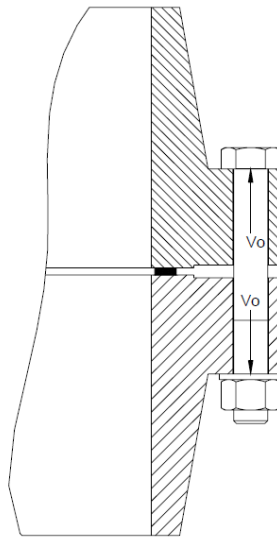
In the framework of pressure vessels, the threaded connections are used to connect the cover to the body of the containers, or to connect two pipe sections etc. However, the speech fact also applies to other cases. On the two the endpoints of the pipe to be connected are welded two flanges between which is interposed a more yieldable gasket material which serves to ensure the sealing.

At first the screw are tightened to ensure sealing; then the container is pressurized. To effect of the pressure screws are stretched further, while the gasket is download. However, a certain load must stay on it to prevent leakage.

In the phase of pre-tensioning the bolts are stretched by a total force  $W_1$  and correspondingly the seal is compressed by a force  $-W_1$ .

In the step of pressurizing the internal pressure  $p$  causes the onset of the force  $W_2 = \frac{\pi G^2 p}{4}$  where  $G$  is the mean diameter of gasket; this causes a shift  $\delta_2$  upwards cover compared to the coat. Correspondingly, screws stretch of  $\delta_2$  and gasket it increases its thickness  $\delta_2$ .

$K_b$  are the force that causes a lengthening unit of the screws (stiffness of the screws) -  $k_g$  the force which causes a decrease of unitary gasket thickness (stiffness of the gasket).



**Figure 17.** Flange connection between mantle and cover

To create the screws axial load  $W_1$  calculated in the previous section (ie a load of  $W_1/N$  on each screw) must tighten the nuts with the proper tightening torque (popularly called tightening torque).

To calculate this we observe that the torque screw-nut screw is dynamically equivalent to a inclined plane. When the nut rotates in the the screwing element is brought from the position 1 to position 2 for example. To make this shift must be a force  $dT$  perpendicular to the axis screw.

For the balance, and considering that all elements are in the same condition

$$T = \frac{W_1}{N} \tan(\alpha + \varphi)$$

Where the  $\alpha$  helix angle of the thread for a screw with pitch  $p$  and  $\varphi$  friction angle.

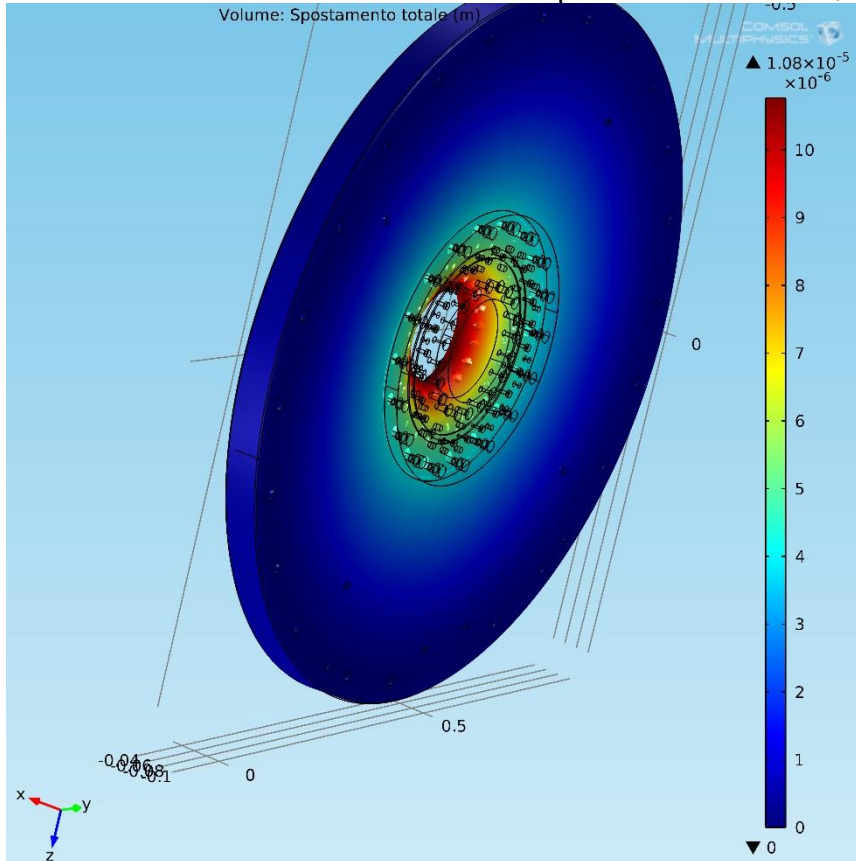
So

$$M_t = \frac{D_m}{2} \frac{W_1}{N} \tan \varphi'$$

where  $\varphi'$  angle of friction between the nut and the bearing surface and  $D_m$  is the mean diameter of the nut (average width between key and nominal diameter). The tightening torque is given by the sum of these two partial moments of which the second is



completely lost, while the first remains stored in screw as torque. Applying these calculations to our system, or the closing flange between the yoke and the plasma chamber and considering our values namely a diameter of 250 mm we obtain that the minimum number of screws is 12 and with a pressure of 30000 N/cm it is obtained that the minimum diameter of the life to be applied is a M14 with a minimum length of 35 mm. It was subsequently also set a finite element simulation to check the tightness of the flange with the assumptions made and the results obtained show that the theoretical calculations made were correct and the maximum displacement is  $1.08 \times 10^{-5}$  m (figure).



**Figure 18.** Connection flange FEM analysis

## Conclusions

The optimization of the AISHa plasma chamber design and all mechanical parts have been pursued with FEM codes.[8] The parts are currently under construction and the source will be assembled in next September at INFN-LNS.

## References

- [1] L. Celona, G. Ciavola, S. Gammino, L. Andò, D. Mascali - DESIGN OF THE AISHA ION SOURCE FOR HADRON THERAPY FACILITIES - Proceedings of ECRIS2012, Sydney, Australia, ISBN ISBN 978-3-95450-123-654
- [2] S. Gammino et al., Rev. Sci. Instrum. 70, 9 (1999), 3577
- [3]. S. Gammino, G. Ciavola, Rev. Sci. Instrum. 71, 2 (2000), 631
- [4] Felippa C.A., Introduction to Finite Element Methods, lecture notes, Department of aerospace engineering sciences of the University of Colorado, Boulder, 2004.
- [5] COMSOL Multiphysics User's Guide v4.1, COMSOL A B, 2010.
- [6] Lewis R.W., Nithiarasu P. & Seetharamu K.N., Fundamentals of the Finite Element Method for Heat and Fluid Flow, New York, John Wiley & Sons, 2004.

[7] CFD Module User's Guide v4.1, COMSOL AB, 2010.

[8] L. Celona et al. ECR IONS SOURCE DEVELOPMENT AT INFN-LNS – Proceedings of ECRIS2014

[9] F. Noto et al, Optimization of the gas flow in a GEM Chamber and development of the GEM foil Stretcher – Nuclear Technology & Radiation Protection, 2014, vol.29 suppl, pp s0-s0, DOI: 10.2298/NTRP140202?N

# Entropically Damped Artificial Compressibility for SPH

Prabhu Ramachandran \*

Department of Aerospace Engineering, IIT Bombay  
Powai, Mumbai 400076  
prabhu@aero.iitb.ac.in

Kunal Puri

Department of Aerospace Engineering, IIT Bombay  
Powai, Mumbai 400076  
kunal.r.puri@gmail.com

March 31, 2015

## Abstract

The Entropically Damped Artificial Compressibility (EDAC) method of Clausen [2013] is an interesting alternative to the traditional artificial compressibility technique. We apply the EDAC method in the context of Smoothed Particle Hydrodynamics. In this work we present preliminary results for this method. The method is easy to implement in a standard SPH framework. Our simulations are compared with those of the Transport Velocity Formulation of Adami and Hu [2013] as well as a standard SPH formulation where necessary.

The results appear very promising and seem comparable to those of the TVF method. Like the TVF, it does not produce spurious pressure oscillations. Unlike the TVF, which cannot be directly applied to free-surface problems, the new method has no such difficulty.

**Keywords:** SPH, Incompressible flow, Artificial compressibility, Entropically Damped Artificial Compressibility

## 1 Introduction

The Smoothed Particle Hydrodynamics method has been applied to a wide variety of problems including elastic dynamics, compressible fluid flow, viscous incompressible fluid flow, multi-phase problems etc. For incompressible flows, SPH implementations either resort to a weakly-compressible formulation using the artificial compressibility technique or use a pressure-based approach as seen in several incompressible SPH schemes. The difficulty with the incompressible projection based schemes is the need to solve an implicit problem which poses computational challenges when scaling up problems to larger sizes and involves increased complexity. The weakly-compressible formulation, while explicit, faces problems related to unphysical pressure oscillations. The pressure oscillations are reduced greatly when the Transport Velocity Formulation of Adami et al. [2013] (TVF) is used. The method introduces a background pressure that serves to reduce tensile instability and reduces pressure oscillations. In the words of the authors, the method produces “unprecedented accuracy and

---

\*Address all correspondence to this author.

stability”. The difficulty with the TVF formulation is that it cannot handle free-surface problems directly on account of the background pressure. The ability to naturally capture free-surface problems is a desired feature of the SPH method.

Recently, the Entropically Damped Entropically Damped Artificially Compressible method of Clausen [2013a,b] (EDAC) has been applied to finite-difference and finite-element schemes. The method employs a new approach and evolves the pressure in time. This evolution equation eliminates the need to use the stiff equation of state that is usually employed.

In the present work, we combine the Entropically Damped Artificially Compressible method with the SPH formulation and the boundary condition of Adami et al. [2012]. This results in a simple set of equations and produces results that are roughly comparable to that of the TVF. The method can also handle free-surfaces as it does not require a background pressure. The results for the dam break problem appear to be good.

We perform simulations for a few classic problems and show that the method produces stable and accurate results comparable to the best available SPH schemes albeit with much simpler equations and no additional corrections. To our knowledge this is the first time that this technique has been applied to the SPH.

## 2 The Numerical Method

As discussed in Clausen [2013a], the basic idea behind the EDAC method is to introduce an evolution equation for the pressure,  $p$ , instead of an equation of state. The evolution equations for momentum and pressure are written as,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \text{grad}(\mathbf{u}) = -\frac{1}{\rho} \nabla p + \text{div}(\sigma), \quad (1)$$

$$\frac{\partial p}{\partial t} + \mathbf{u} \cdot \text{grad}(p) = -\rho c_s^2 \text{div}(\mathbf{u}) + \nu \nabla^2 p \quad (2)$$

In the equations above,  $\mathbf{u}$  is the velocity of the fluid,  $p$  is the pressure  $\sigma$  represents the deviatoric part of the stress tensor,  $c_s$  is the speed of sound (which is set to be a multiple of the maximum speed of the fluid).  $\nu$  is the kinematic viscosity of the fluid. The density,  $\rho$  is held constant. We start with a zero pressure and the pressure evolves naturally from the equation (2) above.

## 3 Numerical Implementation

In order to simulate the equations above with the SPH, the following SPH discretizations are used. The density,  $\rho$  is held constant. The momentum equation is discretized using the standard SPH approach using Monaghan’s original formulation but without any artificial viscosity,

$$\frac{d\mathbf{u}_i}{dt} = - \sum_j m_j \left( \frac{p_j}{\rho_j^2} + \frac{p_i}{\rho_i^2} \right) \nabla W_{ij} \quad (3)$$

We add a real viscosity based on the Morris formulation. For equation (2) we use the following,

$$\frac{dp_i}{dt} = \sum_j m_j c_s^2 \mathbf{u}_{ij} \cdot \nabla W_{ij} + \frac{m_j}{(\rho_i + \rho_j)} \frac{4\nu p_{ij}}{(r_{ij}^2 + \eta h_{ij}^2)} \mathbf{r}_{ij} \cdot \nabla W_{ij}, \quad (4)$$

For the boundary conditions, we employ the formulation of Adami et al. [2012].

## 4 Results

Using the equations above, we simulate a few standard problems and present the results as compared with other SPH formulations.

### 4.1 Dam break

The first test case we consider is the traditional two-dimensional dam-break problem.

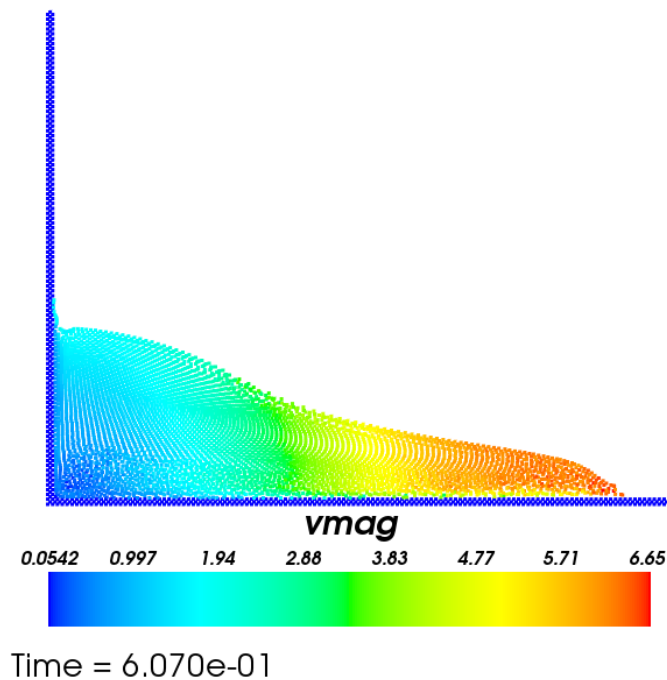


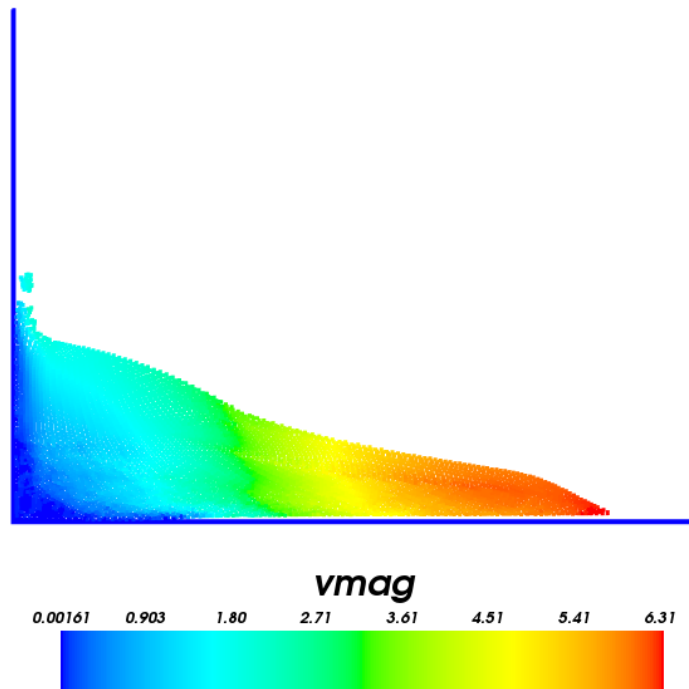
Figure 1: Two-dimensional dam-break problem solved with the EDAC-SPH method.

As can be seen, there is much less fluid sticking to the walls on the side and the results are comparable to those produced by the traditional SPH technique.

We have obtained good results for a variety of other problems like Couette flow, Poiseuille flow, Lid-Driven-Cavity and the Taylor Green vortex. We are in the process of comparing these results with other SPH schemes.

## 5 Conclusions

We have shown preliminary results of a new scheme that combines the EDAC method with SPH. The resulting formulation is very simple, does not display spurious pressure oscillations, does not require a background pressure and seems to produce good results for a few test problems. We will be performing more rigorous tests and will compare the new method carefully with results using the TVF and the standard SPH.



Time = 6.043e-01

Figure 2: Two-dimensional dam-break problem solved with the traditional SPH.

## References

- S. Adami, X.Y. Hu, and N.A. Adams. A generalized wall boundary condition for smoothed particle hydrodynamics. *Journal of Computational Physics*, 231(21):7057–7075, August 2012. doi: 10.1016/j.jcp.2012.05.005. URL <http://linkinghub.elsevier.com/retrieve/pii/S002199911200229X>.
- S. Adami, X.Y. Hu, and N.A. Adams. A transport-velocity formulation for smoothed particle hydrodynamics. *Journal of Computational Physics*, 241:292–307, May 2013. doi: 10.1016/j.jcp.2013.01.043. URL <http://linkinghub.elsevier.com/retrieve/pii/S002199911300096X>.
- Jonathan R Clausen. Entropically damped form of artificial compressibility for explicit simulation of incompressible flow. *Physical Review E*, 87(1):013309–1–013309–12, January 2013a. doi: 10.1103/PhysRevE.87.013309. URL <http://link.aps.org/doi/10.1103/PhysRevE.87.013309>.
- Jonathan R Clausen. Developing Highly Scalable Fluid Solvers for Enabling Multiphysics Simulation. Technical Report March, Sandia National Laboratories, 2013b. URL <http://prod.sandia.gov/techlib/access-control.cgi/2013/132608.pdf>.

## Numerical Investigation of Aerosol Particle Transport and Deposition in Realistic Lung Airway

Mohammad S. Islam<sup>1\*</sup>, Suvash C. Saha<sup>1†</sup>, Emilie Sauret<sup>1</sup>, Y.T. Gu<sup>1</sup>, Zoran D. Ristovski<sup>1</sup>

<sup>1</sup> School of Chemistry, Physics & Mechanical Engineering  
Queensland University of Technology  
2 George Street, GPO Box 2434, Brisbane QLD 4001, Australia.

\*Presenting author: m20.islam@qut.edu.au  
†Corresponding author: suvash.saha@qut.edu.au

### Abstract

Different human activities like combustion of fossil fuels, biomass burning, industrial and agricultural activities, emit a large amount of particulates into the atmosphere. As a consequence, the air we inhale contains significant amount of suspended particles, including organic and inorganic solids and liquids, as well as various microorganism, which are solely responsible for a number of pulmonary diseases. Developing a numerical model for transport and deposition of foreign particles in realistic lung geometry is very challenging due to the complex geometrical structure of the human lung. In this study, we have numerically investigated the airborne particle transport and its deposition in human lung surface. In order to obtain the appropriate results of particle transport and deposition in human lung, we have generated realistic lung geometry from the CT scan obtained from a local hospital. For a more accurate approach, we have also created a mucus layer inside the geometry, adjacent to the lung surface and added all apposite mucus layer properties to the wall surface. The Lagrangian particle tracking technique is employed by using ANSYS FLUENT solver to simulate the steady-state inspiratory flow. Various injection techniques have been introduced to release the foreign particles through the inlet of the geometry. In order to investigate the effects of particle size on deposition, numerical calculations are carried out for different sizes of particles ranging from  $1\ \mu\text{m}$  to  $10\ \mu\text{m}$ . The numerical results show that particle deposition pattern is completely dependent on its initial position and in case of realistic geometry; most of the particles are deposited on the rough wall surface of the lung geometry instead of carinal region.

**Keywords:** Particle Transport, Particle Deposition, Realistic Lung Airway, Mucus Layer, Pollutant Particles.

### 1.0 Introduction

As a result of industrialization and human activities, especially in the developing countries, air pollution has significantly increased over the last and current century. Inhalation and exhalation, either by nose or mouth, are an essential part of the human body mechanisms. During the breathing process, we inhale pollutant particulates mix with air that can cause fatal disease. In the case of inhalation, our nasal cavities and oral airways act as a filter. Most of the inhaled micron particles basically deposit in the nasal cavities and oral region during the inhalation process due to inertial impaction and strong turbulent dispersion [Feng and Kleinstreuer (2014)]. The

remaining particle passes through the trachea and deposit in the bronchi and alveoli [Tena and Clarà (2012)].

**Inhalation and exhalation processes have attracted considerable attention to the researchers in recent years.** Investigation the deposition pattern of inhaled particles in the human lung is very challenging due to the complex geometrical structure of the human lung [Kumar et al. (2009); Soni and Aliabadi (2013); Weibel (1963b)]. According to the Weibels's book, most likely in 1731, Rev. Stephen Hales first studied the elasticity of the air in human lung and proposed a dimension for the surface of the human lung [Weibel (1963a)]. The most comprehensive and simple human lung geometry was defined by Weibel's (1963). In case of modelling and simulation of particle transport and deposition in the human lung, Weibel's [Weibel (1963b)] lung model is still being used due to its geometric simplicity. Some geometrical developments of the idealized lung model have been conducted to overcome its complexity [Finlay (2001); Horsfield et al. (1971); Raabe et al. (1976a); Raabe et al. (1976b)]. The most commonly used model for foreign particle movement and detention in the human respiratory tract is that distinguished by a regular and irregular bifurcation pattern [Nowak et al. (2003); Zhang and Kleinstreuer (2004)]. As the Weibel's based human lung geometry model is not realistic. Therefore, in order to obtain the appropriate numerical results for particle transport and deposition, the CFD analysts now focus on realistic airway models [Ma and Lutchen (2006); Ma and Lutchen (2009); Wall and Rabczuk (2008)]. The anatomical based human airway models like Computerized Tomography (CT) scan or Magnetic Resonance Imaging (MRI) geometrical models are becoming more popular in the current literature.

Inhaled particle deposition in the human lung is mainly caused by inertial impaction, Brownian diffusion, gravitational sedimentation and interception [Choi and Kim (2007)]. The airborne particle deposition in human respiratory tract is principally governed by its shape [Hofmann et al. (2009); Kasper (1982)] and size [Hofmann (2011)]. Submicron-particles, which are less than 0.5  $\mu\text{m}$  are initially deposited in the human lung by Brownian diffusion [Hofmann (2011)], while larger particles are deposited by sedimentation and inertial impaction. Breathing pattern is also responsible for particle deposition in human airways. Due to the long residence time slow breathing patterns are more effective for sedimentation and Brownian diffusion, whereas impaction is favoured by fast breathing pattern [Hofmann (2011)]. [Zhang and Kleinstreuer (2002)] have investigated the air flow structure for nanoparticle in the upper airway. They have investigated the turbulent effect on the tracheobronchial area. Cebra and Summers (2004) have investigated the pressure and flow pattern for tracheal and central bronchial in anatomically realistic model and their investigation concluded that decreased pressure and increased shear stress in the stenosis part. Taherian et al. (2011) have performed the unsteady air flow simulation for a CT scan based realistic geometry and have shown the velocity distribution and variation of the vorticity in the lung. Farkas and Szöke (2013), first time created the mucus layer on their non-realistic geometry to simulate the particle deposition. However, they have considered a constant thickness of mucus for all generations which are not physiologically correct. From the above review, most of the above work has been conducted for non-realistic lung geometry. Moreover, none of them investigated the realistic injection properties.

In the present study, we have developed a model for realistic particle transport and deposition in human lungs by creating lung geometry directly from the CT scan



image obtained from an adult healthy patient. We have also applied realistic injection properties to release particles from the inlet and studied the effect of particles' sizes on the deposition.

## 2. Geometry Generation

In the present study, we have generated the realistic geometry of the human airway. There are several steps to generate the geometry from CT scan or MRI data. First, we have collected the CT image data from a radiologist from a local hospital and the data format is DiCom. The second step is to use AMIRA, the geometry generation commercial software to create the 3D geometry. Then, we have imported the 3D geometry into another software, GEOMAGIC for surface construction. Finally, SOLIDWORKS was used to create the mucus layer inside the geometry.

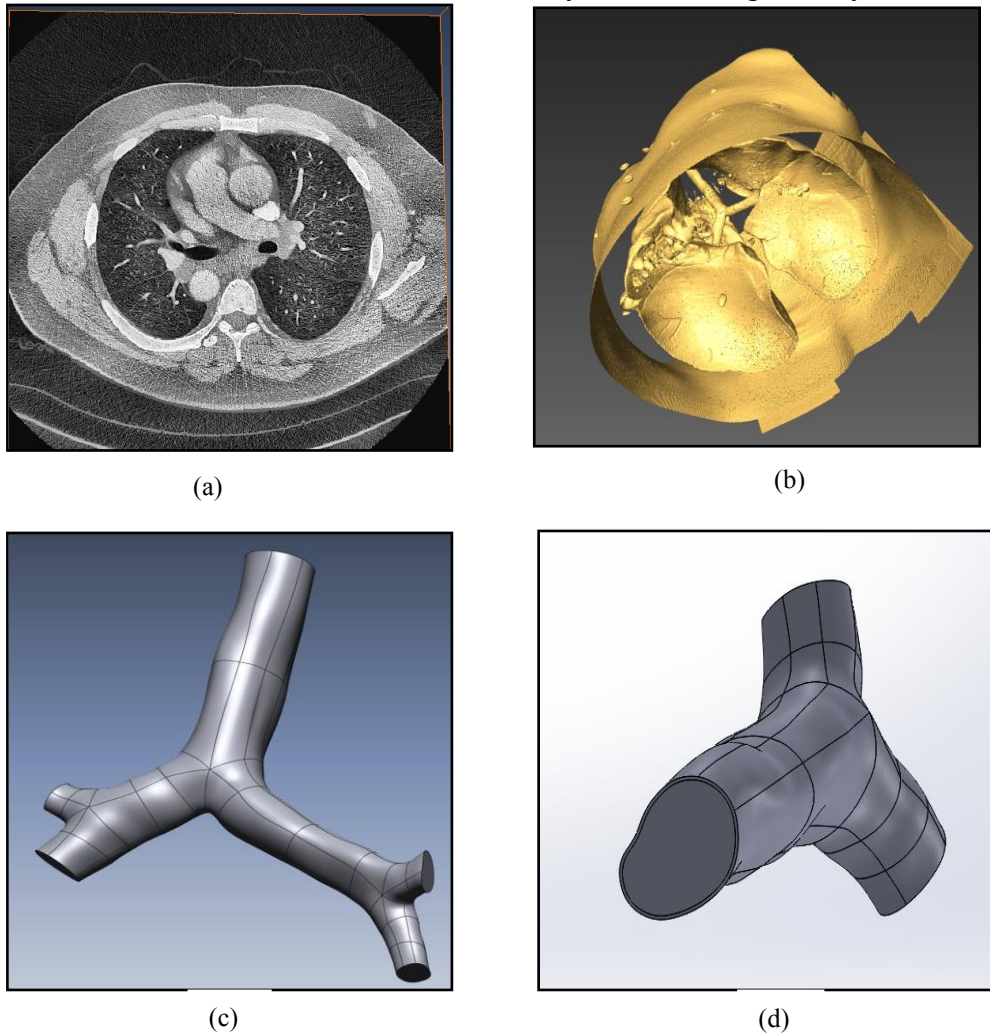


Fig. 1(a) CT scan data visualization, (b) Surface rendering and airway visualization, (c) 3D geometry of human lung, and (d) mucus layer of the geometry.

Fig. 1(a) represents the visualization of CT scan data and Fig. 1(b) represents the lung airway geometry after clearing the raw materials from the geometry. Fig. 1(c) represents the 3D human lung geometry and Fig. 1(d) represents the geometry with the mucus layer. As we know the mucus layer thickness varies from generation to generation, thus we have generated a multi-thickness mucus layer on our geometry.

## 2. Numerical Methods

The complete 3D geometry is imported in ANSYS 15.0 software for simulation purpose. The computational mesh was generated in ANSYS Workbench. A good quality unstructured mesh is generated for the geometry with some inflation layers to correctly capture the boundary layer. Dense mesh has been created on the bifurcation zone where higher velocity gradients are expected. The orthogonal quality of the mesh is calculated as 0.61 above the recommended 0.2 value [put ANSYS manual as reference]. We have generated an unstructured mesh with 1500000 nodes. The grid refinement test is also conducted to obtain the grid independency on numerical solution. The numerical solution of particle transport and deposition are carried out by Lagrangian particle tracking method. We have used the Lagrangian based Discrete Phase Model (DPM), where the air is treated as continuum phase and particle is treated as disperse phase. The continuum phase is governed by the Navier-Stokes equations and the disperse phase can exchange mass, momentum and energy with the continuum phase. In the DPM, the interaction with the continuous phase is considered. We have set the mucus layer properties with the density,  $998.2 \text{ kg/m}^3$  and the viscosity,  $0.89 \text{ kg/ms}$ . The pressure-velocity coupling and second order pressure spatial discretization have been used. The residual convergence criteria are set to  $10^{-06}$ . Different injection properties have been used to release the particle from the inlet surface. The boundary conditions used in FLUENT to solve the particle transport and deposition are; the velocity inlet and pressure outlet. The velocity magnitude at inlet is considered as  $1 \text{ m/s}$  and the zero pressure at pressure outlet. A no slip condition is set at the wall. The discrete phase model conditions are used as trap as the boundary condition type. The boundary condition trap means as soon as the particle will touch the wall, it will capture at the wall.

## 3. Results and Discussion

### 3.1 Simulation without Mucus Layer

In order to investigate the foreign particle transport and deposition, at first we have simulated the model without any mucus layer inside the geometry. The simulations are carried out for three different particles sizes,  $1 \mu\text{m}$ ,  $5 \mu\text{m}$  and  $10 \mu\text{m}$ . The effects of particle size on respiratory deposition have been obtained. According to the physics behind deposition, larger particle should deposit in the upper airway. Fig. 2(a) represents the particle deposition comparison for the three different particles' sizes. One thousand and sixteen particles are released in this simulation and it is observed that in the case of  $10 \mu\text{m}$  particles, 428 particles are deposited in the first generation and the rest of the particles have escaped from the simulations. Among the 428 deposited particles, 159 particles are deposited within the first 100 iterations, which represents that most of the particles are deposited in the upper portion of the generation. In case of  $5 \mu\text{m}$  particles, 358 particles are deposited among the 1016 particles, which is less than  $10 \mu\text{m}$  particles and the same scenario can be observed for the  $1 \mu\text{m}$  particles.

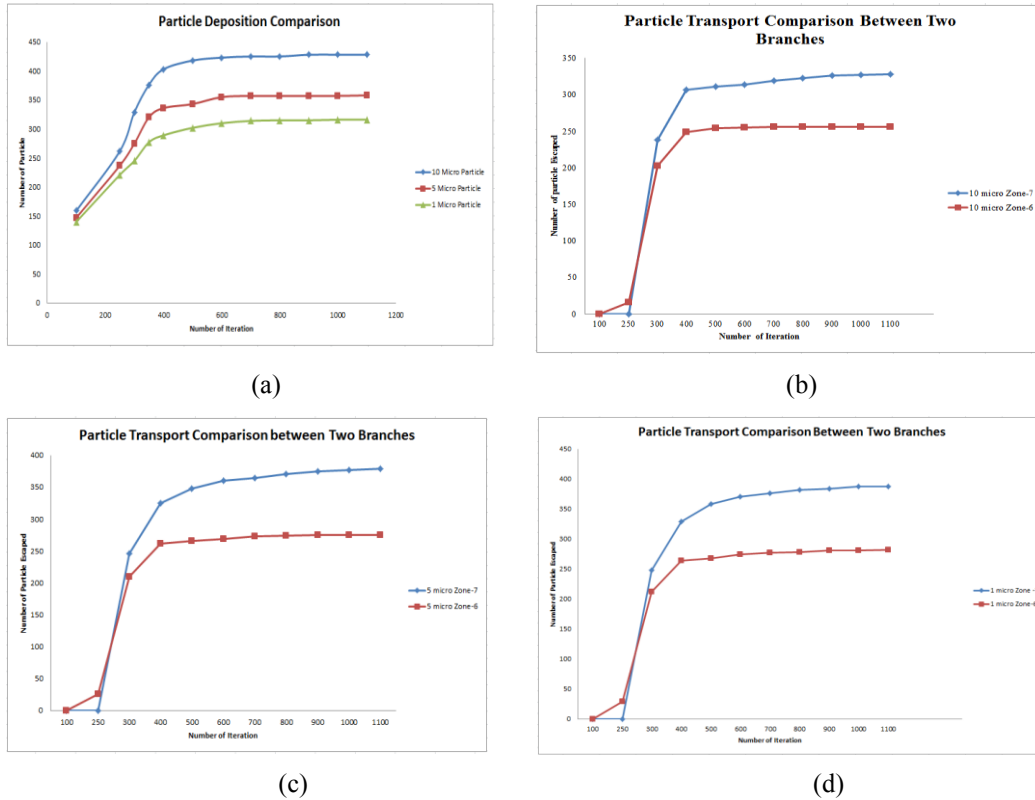


Fig. 2(a): Particle deposition comparison for different sizes of particles, particle transport comparison for 2(b)  $10 \mu m$ , 2(c)  $5 \mu m$  and 2(d)  $1 \mu m$ .

Fig. 2(b) represent the particle transport scenario for  $10 \mu m$  particles through the two daughter branches of the first generation. The diameter of the two daughter branches is different and the diameter of the zone 6 is much larger than zone 7. It will be almost 1.5 times larger than zone 7. An interesting scenario is that the branching pattern of the zone 7 is more vertical than zone 6. As the diameter of zone 6 is greater than zone 7, in general, most particles should escape through zone 6. However, realistically most of the particles have escaped through zone 7 despite its lower diameter. The reason for that is that zone 7 is much more vertical compared to zone 6. That is why, larger particles choose this way due to their larger size and inertia. Fig. 2(c) and Fig. 2(d) exhibit the same scenario for particle transport through the two daughter branches.

In order to get a clear idea about the inspiratory deposition pattern, the deposition scenarios for  $10 \mu m$  particles are visualized. Fig. 3(a) represents the particle deposition pattern for  $10 \mu m$  particles. However, for the realistic geometry, we observed that most of the particles are deposited in the wall surface instead of the carinal region. It should be noted that the particles have been released by using the inlet surface nodes. The geometrical structure of human lung is not symmetrical and due to its uneven surfaces, the geometrical shape is very complex. On the other hand, non-realistic geometry is symmetrical in shape and it does not have any curve, bends or wave on its wall surface. Moreover, in case of non-realistic geometry, the carinal angle is  $45^\circ$ . On the other hand, for the realistic geometry, the carinal angle is found to be more than  $90^\circ$ . As we know, large particles are deposited in the upper portion of the lung airway and inertial impaction is the main mechanism for large particle deposition. In general, all the particles we inhale from the atmosphere follow the air streamline. However, during their movement, when any curve, bends or uneven surface appears

in front of them, they can't follow the air stream due to its large inertia. As a result, those particles are captured by the wall surface.

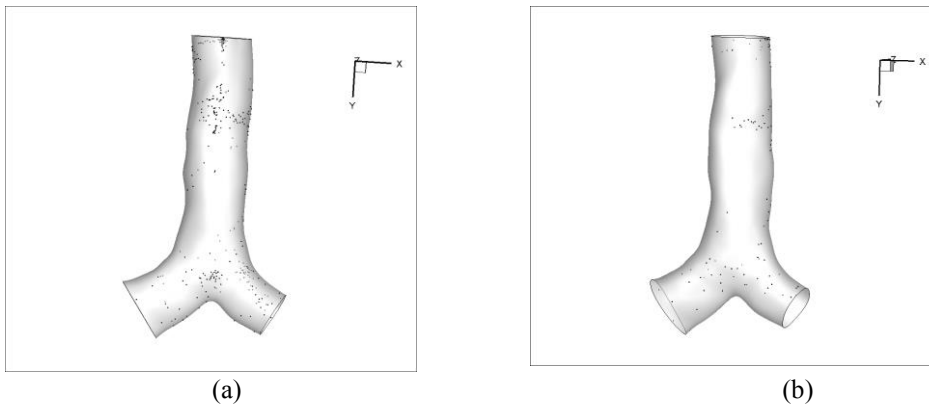


Fig. 3: Inspiratory deposition pattern of  $10 \mu m$  particles (a) Injection-I (b) Injection-II

Fig. 3 (a) shows that a large number of particles is deposited at the very upper portion of the trachea. Fig. 3 (b) depicts that less particles are deposited at the upper portion of the trachea. Two different injection properties have been used for these two cases and the deposition pattern is different for individual injection. The injection-I and injection-II contains 1016 and 800 particles respectively and the initial positions of the particles are different.

Fig. 4 (a) and Fig. 4 (b) represent the respiratory particle deposition pattern for  $5 \mu m$  particles for different injection. The deposition pattern shows that a large number of particles deposited at the two daughter bronchiole as well the parent bronchiole. We can clearly see that fewer particles are trapped in the right bronchiole compared to  $10 \mu m$  particle. If we compare the two daughter bronchioles, then of course larger particles are deposited at the right bronchiole i.e. zone 7. In case of realistic lung geometry, right branch of the geometry is more vertical than left one. That is why most of the particles have a tendency to go through the right bronchiole. As a result, a large number of particles are deposited at the right bronchiole.

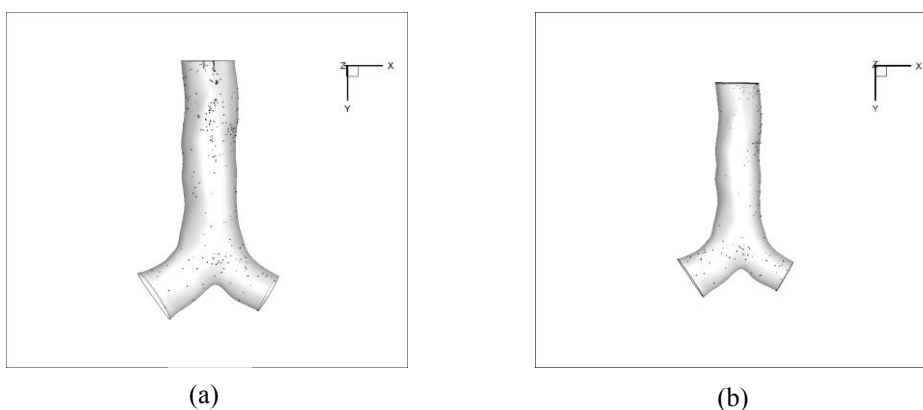


Fig. 4: Inspiratory deposition pattern of  $5 \mu m$  particles (a) Injection-I (b) Injection-II

The inspiratory particle deposition scenario for  $1 \mu m$  particles is shown in Fig. 5(a). From this figure we can clearly observe that very few numbers of particles are deposited at the left bronchiole. The deposition pattern for parent bronchiole is also different compare to larger particle deposition. It is interesting to mention that a good number of particles are deposited in the upper section of the parent bronchiole. Fig.

5(b) depicts the particle deposition pattern for different injection. This figure shows that fewer numbers of particles are deposited at the upper portion.

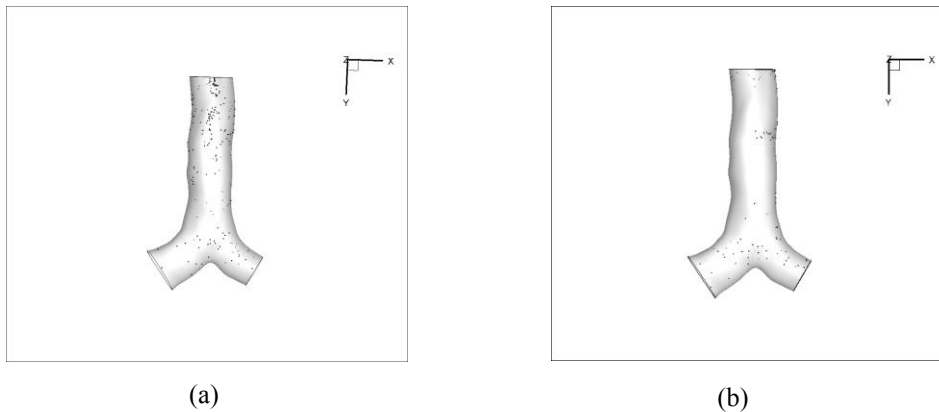
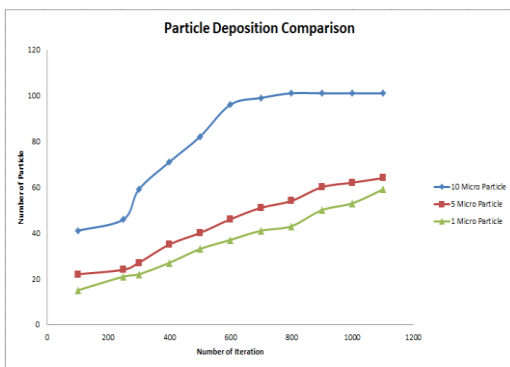


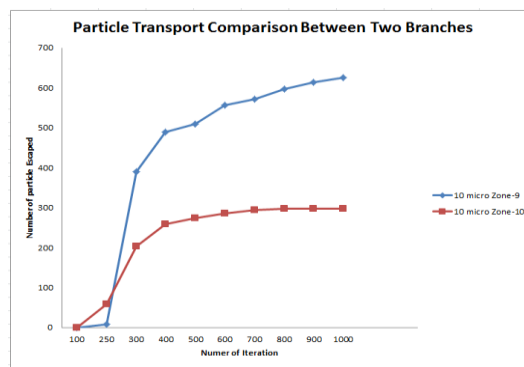
Fig. 5: Inspiratory deposition pattern of  $1\ \mu\text{m}$  particles (a) Injection-I (b) Injection-II.

### 3.2 Simulation with Mucus Layer

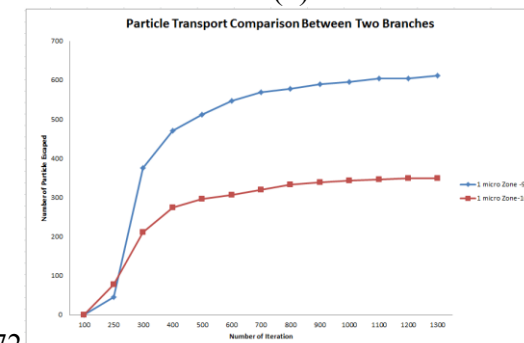
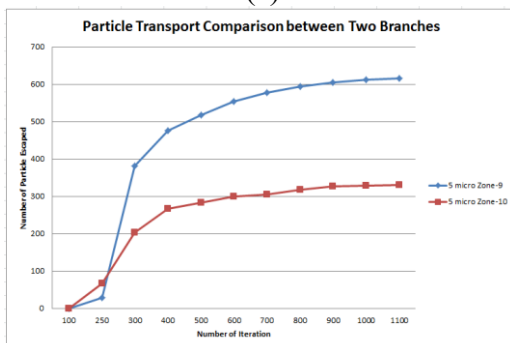
We have also investigated the particle transport and deposition for our lung geometry with mucus layer. The particle transport and deposition pattern have been investigated for  $10\ \mu\text{m}$ ,  $5\ \mu\text{m}$  and  $1\ \mu\text{m}$  particles. We have injected the particles by using the nodes of the inlet surface and the numbers of injected particles are 1054. All appropriate mucus layer properties have been added to simulate the particle transport and deposition. In order to show the deposition pattern, we have collected the trapped particle coordinates on the lung wall. After creating the mucus layer inside the geometry, the diameter of the inlet surface reduces. As a result, the deposition pattern shows a significant difference compared to the case without mucus layer. Fig. 6(a) represents the inspiratory particle deposition pattern for three different sizes of particles. Fig. 6(a) shows that there is a significant difference between the particle depositions for the three different sizes of particles. The graph for  $10\ \mu\text{m}$  particle shows that a large number of particles is deposited in the first generation compare to two other size particles.



(a)



(b)



(c)

(d)

Fig. 6(a): Particle deposition comparison for different size particle, particle transport comparison for 6(b)  $10 \mu m$ , 6(c)  $5 \mu m$  and 6(d)  $1 \mu m$  particle.

Fig. 6(b) represents the particle transport scenario for  $10 \mu m$  particles. From the figure, we can clearly see that a large number of particles are escaping through zone 7. As we explained earlier, two daughter bronchioles are not symmetrical and their diameters are also different. In this case, zone 7 is the smaller daughter bronchioles with small diameter compare to zone 6. In general, due to the large diameter of zone 6, most of the escaped particle should go through the zone 6. But, in this case, most of the particles are escaping through the zone 7 because this bifurcation is more vertical. Fig. 6(c) and Fig. 6(d) represents the particles transport scenario for  $5 \mu m$  and  $1 \mu m$  particle respectively.

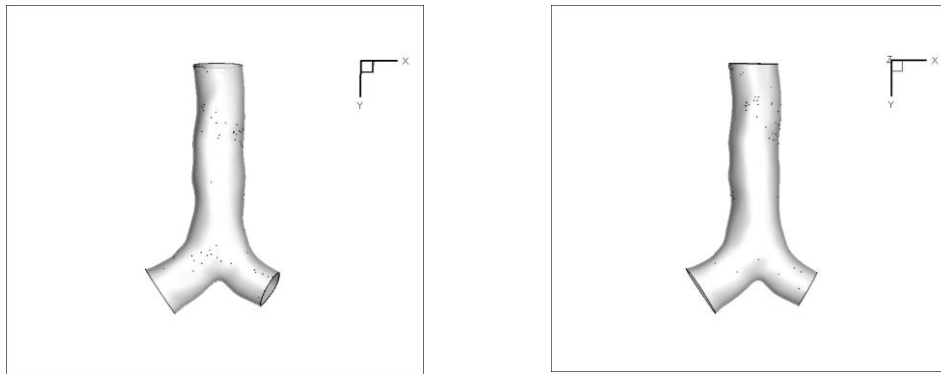


Fig. 7: Inspiratory deposition pattern of  $10 \mu m$  particles (a) Injection-I (b) Injection-II

Fig. 7(a) and Fig. 7(b) represent the respiratory particle deposition scenario for two different injection properties. From the above figure, we can clearly see that the deposition pattern is different for different injection. Fig. 7(a) shows that a good number of particles are deposited at the top of the carinal angle. On contrary, Fig. 7(b) shows that only couple of particles is deposited at the top of the carinal angle.

#### 4. Conclusion

In order to obtain the accurate numerical results of particle transport and deposition in human lung, it is really important to carry out numerical simulations for realistic lung geometry. To investigate the effects of the particle initial position on deposition, different injection techniques have been used and our results suggest that particle deposition is completely dependent on its initial position. The numerical simulations also suggest that the complex geometry of the airway plays an important role in the respiratory deposition. The numerical results also conclude that most of the particles are escaping through the right daughter bronchiole. In order to obtain the effects of mucus layer on deposition, we have carried out the simulation for two cases- with mucus layer and without mucus layer. We have drawn the deposition graph based on number of iterations. The particle deposition comparison graph shows a significant difference between these two cases. The deposition efficiency noticeably increases for  $10 \mu m$  particle in case of mucus layer. The following conclusion can be drawn from the above simulations:

- i) The particle deposition pattern is dependent on its initial injection position.

- ii) Respiratory geometry plays an important role in case of deposition and the deposition pattern is dependent on lung geometry.
- iii) Most of the larger particles have escaped through the right bronchiole in spite of its small diameter.
- iv) The deposition efficiency significantly increased for larger particle in case mucus layer.

## References

- Choi, J.-I. & Kim, C. S. 2007. Mathematical Analysis of Particle Deposition in Human Lungs: An Improved Single Path Transport Model. *Inhalation toxicology*, 19, 925-939.
- Feng, Y. & Kleinstreuer, C. 2014. Micron-Particle Transport, Interactions and Deposition in Triple Lung-Airway Bifurcations Using a Novel Modeling Approach. *Journal of Aerosol Science*, 71, 1-15.
- Finlay, W. H. 2001. *The Mechanics of Inhaled Pharmaceutical Aerosols: An Introduction*, Academic Press.
- Hofmann, W. 2011. Modelling Inhaled Particle Deposition in the Human Lung—a Review. *Journal of Aerosol Science*, 42, 693-724.
- Hofmann, W., Morawska, L., Winkler-Heil, R. & Moustafa, M. 2009. Deposition of Combustion Aerosols in the Human Respiratory Tract: Comparison of Theoretical Predictions with Experimental Data Considering Nonspherical Shape. *Inhalation toxicology*, 21, 1154-1164.
- Horsfield, K., Dart, G., Olson, D. E., Filley, G. F. & Cumming, G. 1971. Models of the Human Bronchial Tree. *Journal of Applied Physiology*, 31, 207-217.
- Kasper, G. 1982. Dynamics and Measurement of Smokes. I Size Characterization of Nonspherical Particles. *Aerosol Science and Technology*, 1, 187-199.
- Kumar, H., Tawhai, M. H., Hoffman, E. A. & Lin, C.-L. 2009. The Effects of Geometry on Airflow in the Acinar Region of the Human Lung. *Journal of biomechanics*, 42, 1635-1642.
- Ma, B. & Lutchen, K. R. 2006. An Anatomically Based Hybrid Computational Model of the Human Lung and Its Application to Low Frequency Oscillatory Mechanics. *Annals of biomedical engineering*, 34, 1691-1704.
- Ma, B. & Lutchen, K. R. 2009. Cfd Simulation of Aerosol Deposition in an Anatomically Based Human Large-Medium Airway Model. *Annals of biomedical engineering*, 37, 271-285.
- Nowak, N., Kakade, P. P. & Annapragada, A. V. 2003. Computational Fluid Dynamics Simulation of Airflow and Aerosol Deposition in Human Lungs. *Annals of biomedical engineering*, 31, 374-390.
- Raabe, O., Yeh, H., Schum, G. & Phalen, R. 1976a. Tracheobronchial Geometry: Human. *Dog, Rat, Hamster*.
- Raabe, O. G., Yeh, H.-C., Schum, G. & Phalen, R. F. 1976b. Tracheobronchial Geometry: Human, Dog, Rat, Hamster.
- Soni, B. & Aliabadi, S. 2013. Large-Scale Cfd Simulations of Airflow and Particle Deposition in Lung Airway. *Computers & Fluids*, 88, 804-812.
- Tena, A. & Clarà, P. 2012. Deposition of Inhaled Particles in the Lungs. *Archivos de Bronconeumología (English Edition)*, 48, 240-246.
- Wall, W. A. & Rabczuk, T. 2008. Fluid-Structure Interaction in Lower Airways of Ct-Based Lung Geometries. *International Journal for Numerical Methods in Fluids*, 57, 653-675.
- Weibel, E. R. 1963a. Chapter I - Introduction. In: WEIBEL, E. R. (ed.) *Morphometry of the Human Lung*. Academic Press.
- Weibel, E. R. 1963b. Morphometry of the Human Lung.
- Zhang, Z. & Kleinstreuer, C. 2004. Airflow Structures and Nano-Particle Deposition in a Human Upper Airway Model. *Journal of computational physics*, 198, 178-210.
- Zhang, Z. & Kleinstreuer, C. 2002. Transient Airflow Structures and Particle Transport in a Sequentially Branching Lung Airway Model. *Physics of Fluids (1994-present)*, 14, 862-880.

# The influence of the parameter $h$ and a new modified method of Homotopy analysis method for initial value problems

†Z. Wang<sup>1</sup>, \*L. Zou<sup>2,3</sup>, and Y.P. Qin<sup>1</sup>

<sup>1</sup>School of Mathematical Sciences, Dalian University of Technology, China.

<sup>2</sup>School of Naval Architecture, Dalian University of Technology, China.

<sup>3</sup>State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, China.

\*Presenting author: lizou@dlut.edu.cn

†Corresponding author: wangzhen@dlut.edu.cn

## Abstract

In this paper, we pay more attention to the embedding parameter  $h$ , which has an influence on the convergence region of solution series in Homotopy analysis method (HAM). We use some theorems to give the concrete influence and proof. Then we introduce a new modified method of HAM called the Piecewise homotopy analysis method (PHAM). Furthermore, examples such as NLS equation, Ricatti equation and Duffing equation are presented to illustrate the main results.

**Keywords:** Homotopy analysis method; the convergence region; Cauchy-Kowalevskaya theorem; Piecewise homotopy analysis method

## Introduction

In 1992, Liao [1] introduced a powerful method known as the Homotopy analysis method (HAM) to solve the nonlinear problems. Compared with perturbation techniques, on one hand, the validity of HAM is independent of whether or not there exist small parameters in the considered equation. And on the other hand, HAM has an embedding parameter to control the convergence region of the series, namely convergence-control parameter. That is to say, we can adjust and control the convergence region of the series by assigning the convergence-control parameter a proper value. In recent 20 years, this method has been successfully applied to solve many types of nonlinear problems [2-6]. In these papers, absolutely all of them note that the convergence region of solution series depend upon the convergence-control parameter [7]. For some special cases, Liao only shows some special equations' convergence region, such as Ricatti equation, Blasius equation [2] and so on. Liu [8] make contributions to giving the essence of the generalized Taylor theorem, which is the key of HAM. However, the influence that the closer the value of the convergence-control parameter is to zero, the larger the convergence region is to infinite usually known as one common things without giving a strict proof.

Inspired by Abdelrazec's [9] idear to prove the convergence of the Adomian decomposition method, we aim to the convergence-control parameter to show the detail influence and verify it on the NLS equation with cubic nonlinearity. On the basis of the convergence of HAM, we propose a new analysis method called Piecewise homotopy analysis method. To our surprise, this method has a lot of benefits than HAM under the application of some examples.

## Mathematical Formulation

Consider the abstract initial-value problem

$$\begin{cases} (1-q)L[\phi(t, q) - u_0(t)] = hqH(t)[L(\phi(t, q)) - N(\phi(t, q))], & t > 0, \\ u_0(0) = f \end{cases} \quad (1)$$

where the operator  $L$  is linear and  $N$  is nonlinear,  $N(u)$  is analytic near the initial data  $f$ . Without loss of generality, we take  $L = \frac{\partial}{\partial t}$ ,  $u_0(t) = f$ ,  $H(t) = -1$ , the part of  $t > 0$  in formula (1) could be transformed into



$$(1-q) \frac{\partial}{\partial t} \phi(t, q) + hq \left[ \frac{\partial}{\partial t} \phi(t, q) - N \left( \frac{\partial}{\partial t} \phi(t, q) \right) \right] = 0.$$

Simplify the above formulation, we have

$$\frac{\partial}{\partial t} \phi(t, q) = mN[\phi(t, q)] \tag{2}$$

with  $\phi(0, q) = f$ , where  $m = \frac{hq}{hq - q + 1}$ . Moreover, Eq. (2) could be reformulated as

$$\phi(t, q) = f + m \int_0^t N[\phi(s, q)] ds \tag{3}$$

with  $\phi(0, q) = f$ .

**Convergence Analysis**

**Lemma** (Cauchy’s estimate). Suppose that  $\phi(t)$  is differentiable in  $\{t \in \mathbb{R} : |t - t_0| < T\}$ , and to any  $\hat{t} \in (0, T)$ , there exist a  $M > 0$ , such that  $\phi(t) \leq M$  on  $C : |t - t_0| = \hat{t}$ , then

$$|\phi^{(k)}(t_0)| \leq \frac{Mk!}{\hat{t}^k}, \forall k \geq 0.$$

**Proof.** From the Cauchy integral formula, we obtain

$$\phi^{(k)}(t_0) = \frac{k!}{2\pi i} \int_C \frac{\phi(t)}{(t - t_0)^{k+1}} dt,$$

then

$$|\phi^{(k)}(t_0)| = \frac{k!}{2\pi} \left| \int_C \frac{\phi(t)}{(t - t_0)^{k+1}} dt \right| \leq \frac{k!}{2\pi} \frac{M}{\hat{t}^{k+1}} \cdot 2\pi \hat{t} = \frac{Mk!}{\hat{t}^k}$$

**Theorem 1** (Cauchy-Kowalevskaya). Suppose  $\phi(t)$  is the exact solution of Eq.(2),  $m < +\infty$ , then there exists a  $\tau > 0$  such that  $u : [0, \tau] \rightarrow \mathbb{R}$  is also an analytic real function.

**Proof.** As  $N(\phi)$  is analytic near  $f$ ,  $m < +\infty$ , so  $m^2 N(\phi)$  is also analysis near  $f$ . By Cauchy’s estimate, there exist  $a, b > 0$  such that

$$m^2 \frac{1}{k!} |\partial_\phi^k \phi(f)| \leq \frac{b}{a^k}, \forall k \geq 0, \tag{4}$$

where  $\partial_\phi^k N(\phi)$  make the sense of the Fréchet derivative, it means that  $\partial_\phi N(\phi) = N'(\phi)$ ,  $\partial_\phi^2 N(\phi) = N''(\phi)$ , ... The taylor series of  $N(\phi)$  at  $f$  converges when  $|\phi - f| < a$ . What’s more, if  $r = |\phi - f|$ , we have

$$m|N(\phi)| \leq m \sum_{k=0}^{\infty} \frac{1}{k!} |\partial_\phi^k N(f)| |\phi - f|^k \leq \frac{b}{m} \sum_{k=0}^{\infty} \frac{|\phi - f|^k}{a^k} = \frac{a}{a-r} \frac{b}{m} =: mg(r).$$

Consider the majorant function  $g(r)$ , it is obviously that

$$m \frac{1}{k!} |\partial_\phi^k N(f)| \leq m \frac{1}{k!} \partial_r^k g(0), \forall k \geq 0. \tag{5}$$

We get the initial-problem (2)’ majorant problem which is given by

$$\begin{cases} \dot{r}(t) = mg(r), & t > 0, \\ r(0) = 0 \end{cases} \quad (6)$$

where  $r > 0$ . The majorant problem (6) is a simple ODE with the exact solution

$$r(t) = a - \sqrt{a^2 - 2 \frac{ab}{m} t}.$$

It is clear that  $r(t)$  is analytic on  $(-\infty, \frac{a}{2b}m)$ . According to the comparison principle, if  $\phi(t)$  suit Eq. (3), then

$$|\phi(t) - f| \leq m \int_0^t |N(\phi(s))| ds \leq m \int_0^t |g(r(s))| ds = r(t) = \sum_{k=1}^{\infty} \frac{t^k}{k!} \partial_t^k r(0).$$

The Taylor series of the majorant problem's solution  $r(t)$  absolutely converges for all  $|t| < \frac{a}{2b}m$ . In

order to prove  $\phi(t)$  is also analytic in  $t \in [0, \tau)$ , where  $\tau = \frac{a}{2b}m = \frac{a}{2b} \frac{hq}{hq - q + 1}$ , we must to prove

$$|\partial_t^k \phi(0)| \leq \partial_t^k r(0), \quad k \geq 1.$$

Suppose that's right, then the Taylor series for  $\phi(t)$  has the majorant series. According to Weierstrass M-test, consequently, it converges. First we prove the bound above by computing the following formulas for  $k = 1, 2, 3$ ,

$$\partial_t \phi(t) = mN(\phi(t)),$$

$$\partial_t^2 \phi(t) = mN'(\phi(t))N(\phi(t)),$$

$$\partial_t^3 \phi(t) = mN''(\phi(t))N(\phi(t))N(\phi(t)) + mN'(\phi(t))N'(\phi(t))N(\phi(t)).$$

As a result,

$$|\partial_t \phi(0)| \leq m|N(\phi(0))| \leq mg(r(0)) = \partial_t r(0),$$

$$|\partial_t^2 \phi(0)| \leq m|N'(\phi(0))||N(\phi(0))| \leq mg'(r(0))g(r(0)) = \partial_t^2 r(0),$$

$$\begin{aligned} |\partial_t^3 \phi(0)| &\leq |\partial_t^3 \phi(0)|m|N''(\phi(0))||N(\phi(0))||N(\phi(0))| + m|N'(\phi(0))||N'(\phi(0))||N(\phi(0))| \\ &\leq mg''(r(0))g(r(0))g(r(0)) + mg'(r(0))g'(r(0))g(r(0)) = \partial_t^3 r(0), \end{aligned}$$

Generally speaking, for all  $k \geq 0$ ,

$$\phi^{(k+1)}(t) = mP_k(N(\phi(t))),$$

where  $P_k(N(\phi(t)))$  is a polynomial of  $N$ , and its Fréchet derivatives up to the  $k$ th-order with positive coefficients. In consequence, we have

$$|\partial_t^{(k+1)} \phi(0)| \leq m|P_k(N(\phi(0)))| \leq mP_k(|N(\phi(0))|) = \partial_t^{(k+1)} r(0),$$

So far, this bound have already concluded the proof.

**Theorem 2.** For the formula

$$\tau(h) = \frac{a}{2b} \frac{hq}{hq - q + 1}, \quad (7)$$

the closer the value of  $h$  is to zero, the larger the convergence region  $\tau(h)$  is to  $\infty$ .

**Proof.** As  $\tau(h) > 0$ , we have

$$h \in (-\infty, 1 - \frac{1}{q}) \cup (0, +\infty).$$

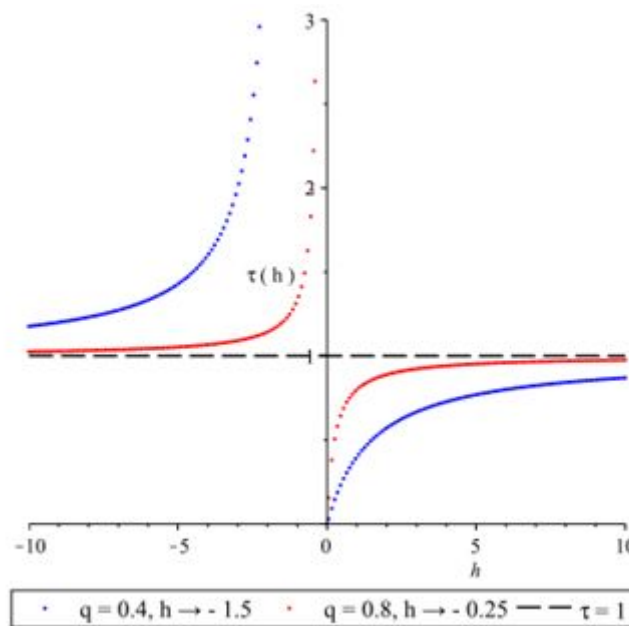
Eq. (7) equivalent to

$$\tau(h) = \frac{a}{2b} \frac{1}{1 - \frac{1}{h} (1 - \frac{1}{q})}, \tag{8}$$

From Eq. (8), we have the conclusion for the convergence region  $\tau$ .

- (i) the closer the value of  $h$  is to  $\infty$ , the closer the convergence region  $\tau$  is to  $\frac{a}{2b}$ .
- (ii) the closer the value of  $h$  is to  $1 - \frac{1}{q}$ , the larger the convergence region  $\tau$ . Especially, when  $q \rightarrow 1^-$ , that is,  $h \rightarrow 0^-$  and  $\tau \rightarrow +\infty$ .

The proof has been finished in (ii) and the conclusion can be showed in Figure 1.



**Figure 1.** Take different value of  $q = 0.4, 0.8$  with  $a = 2, b = 1$ , we find the closer the corresponding value of  $1 - \frac{1}{q}$  is to  $-1.5, -0.25$ , the closer  $\tau$  is to  $+\infty$ .

### HAM for NLS Equation

Some nonlinear equations are often presented as

$$\begin{cases} L(u) = N(u), & t > 0, \\ u(0) = f. \end{cases} \tag{9}$$

For example, consider the following continuous NLS equation

$$iu(t) = -\frac{1}{2}u_{xx} - |u|^2u, \quad t > 0, \tag{10}$$

where  $u(x, t)$  is an amplitude function with the property of  $|u|^2 = \bar{u} u$ , this equation usually called NLS equation with cubic nonlinearity. Obviously, the equation fits to the above abstract formulation (9) with

$$N(u) = \frac{1}{2} i \partial_x^2 u + i |u|^2 u .$$

Do the transformation  $u(x, t) = F(x) e^{it}$ , then use  $u(t)$  instead of  $F(x)$ . Moreover, take the initial data  $u(0) = 1, \dot{u}(0) = 0$ , we obtain

$$\begin{cases} \ddot{u}(t) = u - 2u^3, & t > 0, \\ \dot{u}(0) = 0, \\ u(0) = 1, \end{cases} \tag{11}$$

The exact solution of (11) is  $\text{sech}(t)$ . It is straightforward represent

$$u(t) = \sum_{k=0}^{\infty} a_k t^k, \tag{12}$$

under the set of base functions

$$\{t^k, k = 0, 2, 4, \dots\}.$$

With the aid of the Eq. (11) and under the rule of solution expression, we choose the initial approximation

$$u_0(t) = 1,$$

and the auxiliary linear operator

$$L = \frac{\partial^2}{\partial t^2},$$

and

$$N(\phi(t, q)) = \frac{\partial^2}{\partial t^2} \phi(t, q) - \phi(t, q) + 2\phi^3(t, q),$$

with  $\phi(0, q) = 1$ . According to the above conditions, we have the zeroth-order deformation equation

$$(1 - q)L[\phi(t, q) - u_0(t)] = hqH(t)N(\phi(t, q)) \tag{13}$$

From Eq. (13), we know  $\phi(t, 0) = u_0(t)$  and  $\phi(t, 1) = u(t)$  when  $q = 0$  and  $q = 1$ . With respect to the embedding parameter  $q$ , we define

$$u_k(t) = \frac{1}{k!} \frac{\partial^k}{\partial q^k} \phi(t, q),$$

and  $\phi(t, q)$  can be expanded in Taylor series

$$\phi(t, q) = \phi(0, q) + \sum_{k=0}^{\infty} u_k(t) q^k,$$

Define the following vector

$$\vec{u}_k(t) = \{u_0(t), u_1(t), \dots, u_k(t)\}.$$

Differentiating Eq. (13)  $k$  times with respect to  $q$ , then take  $q = 0$  and dividing by  $k!$ . At last, we get the  $k$ th-order deformation equation

$$L[u_k - \chi_k u_{k-1}] = hH(t)R_k[\vec{u}_{k-1}],$$

with the initial condition  $u_0(0) = 1, u_k(0) = 0 (k > 0)$ , where

$$R_k[\vec{u}_{k-1}] = \ddot{u}_{k-1} - u_{k-1} + 2 \sum_{i=0}^{k-1} \sum_{j=0}^{k-1-i} u_i u_j u_{k-1-i-j} - (1 - \chi_k)$$

and

$$\chi_k = \begin{cases} 0, & k \leq 1, \\ 1, & k > 1. \end{cases}$$

According to the both of the rule of solution expression and the coefficient ergodicity, the corresponding auxiliary function should be determined uniquely

$$H(t) = 1.$$

Then we successively have

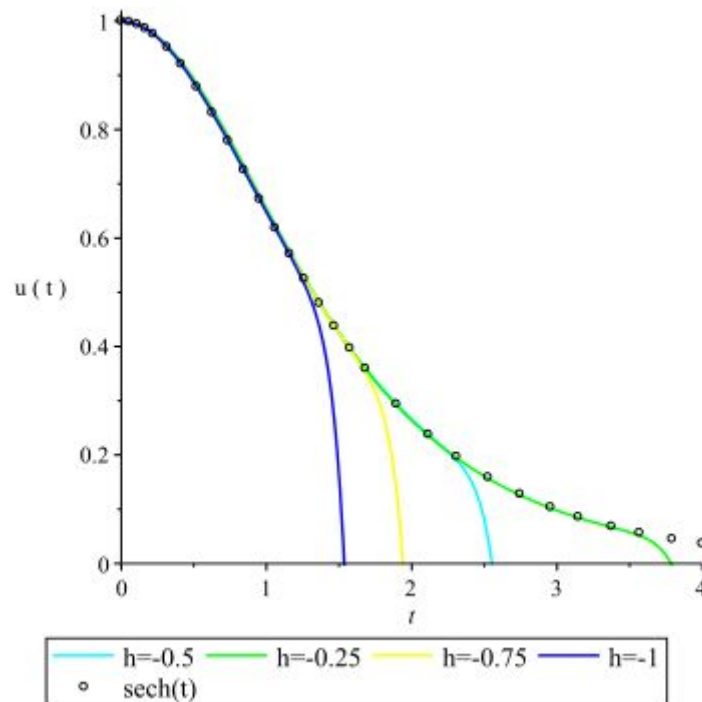
$$u_1(t) = \frac{1}{2}ht^2,$$

$$u_2(t) = \frac{1}{2}ht^2 + \frac{1}{2}h^2t^2 + \frac{5}{24}h^2t^4,$$

$$u_3(t) = \frac{1}{2}ht^2 + h^2t^2 + \frac{1}{2}h^3t^2 + \frac{5}{12}h^2t^4 + \frac{5}{12}h^3t^4 + \frac{61}{720}h^3t^6,$$

.....

According to the formula  $u(t) = \sum_{k=0}^{\infty} u_k(t)$ ,  $u(t)$  could be obtained and as is shown in Figure 2.



**Figure 2.** Take different value of  $h = -1, -0.75, -0.5, -0.25$ , we find the value of  $h$  is closer to 0, the corresponding graph is more fit with the graph of the exact solution.

### Piecewise Homotopy Analysis Method

Considered the graph of  $h = -1$  from Figure 2, we find the homotopy analysis solution are very closed to the exact solution in  $[0,1]$  but away from the exact solution in  $[1,4]$ . In generally, compare the approximate solution with the exact solution for any given value of  $h$ , we can choose a  $t_0$ , such that the homotopy analysis solution and the exact solution be overlap for any  $t \in [0, t_0]$  in an

negligible error, but be away for  $t \in [t_0, \infty]$ . Now, conclude the idea of Piecewise homotopy analysis method: First, we can get the homotopy analysis solution of one nonlinear equation with the initial guess solution  $u_0(t)$ , here we assume the approximate solution is  $a_0(t)$  with the starting point  $(0, u_0(t))$ , then we can obtain the graphs of the homotopy analysis solution  $a_0(t)$  and the exact solution. Second, we can choose a point  $t_0$  on the overlap section, the part  $t \in [0, t_0]$  is to be preserved because the homotopy analysis solution  $a_0(t)$  is absolutely equal to the exact for any  $t \in [0, t_0]$  in an negligible error. Third, we choose the point  $(t_0, a_0(t_0))$  as the starting point and use HAM to the nonlinear equation again to get a new analysis approximate solution  $a_1(t)$  and its graph. We can find a point  $t_1 (> t_0)$  on the overlap section between the new approximate  $a_1(t)$  and the exact solution's graphs, then keep the part  $t \in (t_0, t_1]$  for the same reason.  $t_2 (> t_1), t_3 (> t_2), t_4 (> t_3), \dots$  and  $a_2(t), a_3(t), a_4(t), \dots$  would be find after repeat the above step again and again, where  $\{t_k | k=0,1,2,\dots\}$  is not unique. At last, the Piecewise homotopy analysis solution  $a(t)$  could be obtained as the following form.

$$\begin{cases} a_0(t), & 0 \leq t \leq t_0 \\ a_1(t) = a_0(t - t_0) \Big|_{u_0=a_0(t_0)}, & t_0 \leq t \leq t_1 \\ a_2(t) = a_1(t - t_1) \Big|_{u_0=a_0(t_1)}, & t_1 \leq t \leq t_2 \\ \dots \\ a_k(t) = a_{k-1}(t - t_{k-1}) \Big|_{u_0=a_0(t_{k-1})}, & t_{k-1} \leq t \leq t_k \\ \dots \end{cases}$$

To calculate conveniently, we often choose the distance of  $t_k$  and  $t_{k+1}$  as a small and constant real number, and this number is called the tep size,  $d$ . That is,  $d = t_{k+1} - t_k, k = 0,1,2,\dots$ . The detail answer is that,  $t_0 = d, t_1 = 2d, \dots, t_k = (k + 1)d, \dots$ . In this case, the Piecewise homotopy analysis solution  $a(t)$  could be represented as

$$\begin{cases} a_0(t), & 0 \leq t \leq d \\ a_1(t) = a_0(t - d) \Big|_{u_0=a_0(d)}, & d \leq t \leq 2d \\ a_2(t) = a_1(t - 2d) \Big|_{u_0=a_0(2d)}, & 2d \leq t \leq 3d \\ \dots \\ a_k(t) = a_{k-1}(t - kd) \Big|_{u_0=a_0(kd)}, & kd \leq t \leq (k + 1)d \\ \dots \end{cases}$$

**Piecewise Homotopy Analysis Method for Ricatti Equation**

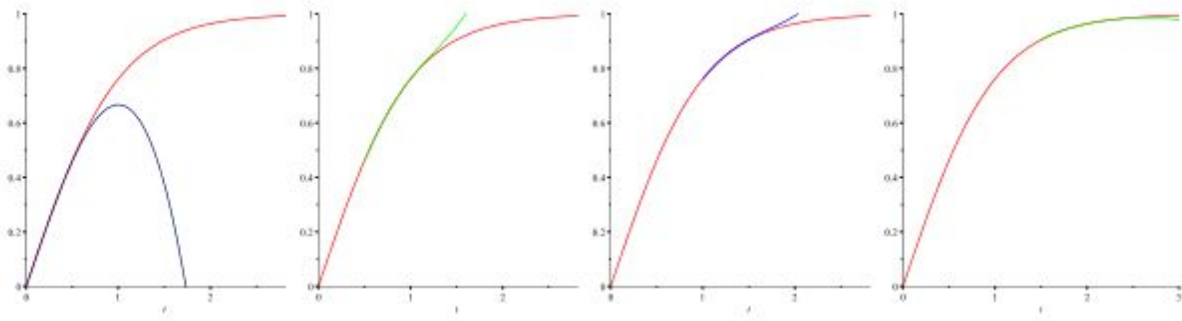
Consider the Ricatti equation,

$$\dot{u}(t) = 1 - u^2(t), \quad t > 0, \tag{14}$$

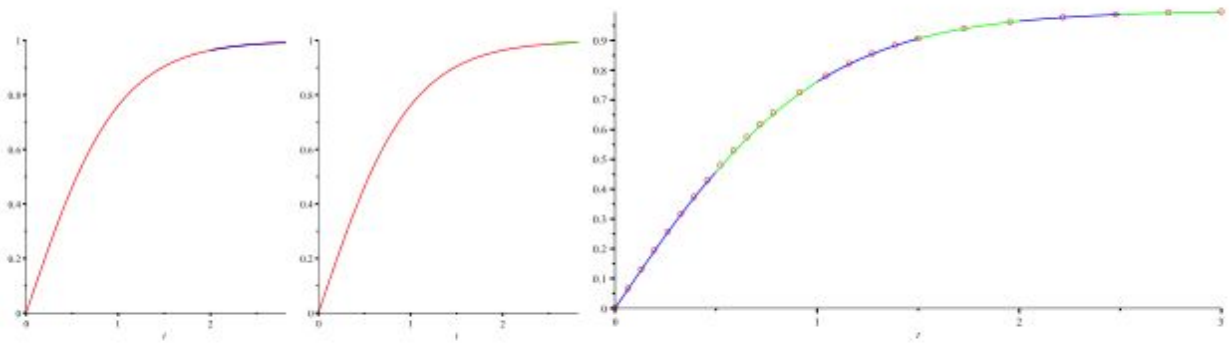
with initial condition  $u(0) = 0$ . The exact solution of (14) is  $u(t) = \tanh(t)$ .

For simplicity, we let  $h = -1$  and just take 4th-order homotopy analysis solution  $u(t) \approx t - \frac{1}{3}t^3$  after using the HAM to Eq.(14) with  $u_0 = t$ , because the graph of the approximate solution  $t - \frac{1}{3}t^3$  and

the exact solution  $\tanh(t)$  has already overlap in  $[0,0.5]$ , that is to say,  $[0,0.5]$  is a valid interval for Ricatti equation. As is showed on Figure 3(a).



(a) Blue:  $a_0, 0 \leq t \leq 3.0$  (b) Green:  $a_1, 0.5 < t \leq 3.0$  (c) Blue:  $a_2, 1.0 < t \leq 3.0$  (d) Green:  $a_3, 1.5 < t \leq 3.0$



(e) Blue:  $a_4, 2.0 < t \leq 3.0$  (f) Green:  $a_5, 2.5 < t \leq 3.0$  (g) solid:  $a(t), 0 \leq t \leq 3$

**Figure 3.** (a-f): the red line is the figure of  $\tanh(t), 0 \leq t \leq 3$ . (g): the point of circle is the exact solution  $\tanh(t)$ , the blue and green line is the 4-th order approximate solution  $a(t)$  on divided sections with  $t_0 = 0.5, t_1 = 1.0, t_2 = 1.5, t_3 = 2.0, t_4 = 2.5, t_5 = 3.0$ .

In generally, use the HAM, from Liao's book [2], we know the  $k$ th-order deformation equation ,

$$\dot{u}_k - \chi_k \dot{u}_{k-1} = hH(t) \left[ \dot{u}_{k-1} + \sum_{j=0}^{k-1} u_j u_{k-1-j} - (1 - \chi_k) \right]$$

we can get  $u(t)$ 's 4th-order homotopy analysis solution under the initial data  $u_0$ .

$$a_0(t) = u_0 + (1 - u_0^2)t + (u_0^3 - u_0)t^2 + \left(-u_0^4 + \frac{4}{3}u_0^2 - \frac{1}{3}\right)t^3 + \left(u_0^5 - \frac{5}{3}u_0^3 + \frac{2}{3}u_0\right)t^4 \quad (15)$$

with  $h = -1, H(t) = 1$  near  $t = 0$ . Note that  $u_0$  is an undetermined initial data. It can be easily find that  $a_0(t) = t - \frac{1}{3}t^3$  when  $u_0 = 0$ . Then we use the Piecewise homotopy analysis method.

Take  $t_0 = 0.5$ , we know the 4th-order Piecewise homotopy analysis solution is  $a_0(t)$  on  $[0,0.5]$ ; take  $t_1 = 1.0$ , the 4th-order Piecewise homotopy analysis solution is  $a_1(t) = a_0(t - 0.5)|_{u_0 = a_0(0.5)}$  on  $(0.5, 1]$ . In generally, take  $t_k = 0.5(k + 1)$ , we can get the 4th-order Piecewise homotopy analysis solution  $a_k(t)$  on  $(0.5k, 0.5(k + 1)]$ .

By the Piecewise homotopy analysis method,  $u(t)$ 's 4th-order Piecewise homotopy analysis solution  $a(t)$  is

$$\begin{cases} a_0(t), & 0 \leq t \leq 0.5 \\ a_1(t) = a_0(t - 0.5) \Big|_{u_0 = a_0(0.5)}, & 0.5 \leq t \leq 1.0 \\ a_2(t) = a_1(t - 1.0) \Big|_{u_0 = a_0(1.0)}, & 1.0 \leq t \leq 1.5 \\ \dots \\ a_k(t) = a_{k-1}(t - 0.5k) \Big|_{u_0 = a_0(0.5k)}, & 0.5k \leq t \leq 0.5(k+1) \\ \dots \end{cases}$$

The figure could be obtained in  $t \in [0, \infty]$  with the values of  $t_0 = 0.5, t_1 = 1.0, \dots, t_k = 0.5(k+1), \dots$  as showed on Figure 3(g).

**Table 1.** Comparisons of  $a(t)/\tanh(t)$ ,  $a'(t)/\tanh'(t)$  and  $a''(t)/\tanh''(t)$  with different values of  $t$ .

$t$	$a(t)$	$\tanh(t)$	$a'(t)$	$\tanh'(t)$	$a''(t)$	$\tanh''(t)$
0.0	0.000000	0.000000	1.000000	1.000000	0.000000	0.000000
0.5 <sup>+</sup>	0.458333	0.462117	0.789931	0.786448	-0.724103	-0.726862
0.5 <sup>-</sup>	0.458333	0.462117	0.750000	0.786448	-1.000000	-0.726862
1.0 <sup>+</sup>	0.760947	0.761594	0.420960	0.419974	-0.640656	-0.639700
1.0 <sup>-</sup>	0.760947	0.761594	0.437507	0.419974	-0.520278	-0.639700
1.5 <sup>+</sup>	0.906028	0.905148	0.179113	0.180707	-0.324564	-0.327133
1.5 <sup>-</sup>	0.906028	0.905148	0.192241	0.180707	-0.246151	-0.327133
2.0 <sup>+</sup>	0.964366	0.964028	0.069998	0.070651	-0.135008	-0.136219
2.0 <sup>-</sup>	0.964366	0.964028	0.069814	0.070651	-0.137663	-0.136219
2.5 <sup>+</sup>	0.986599	0.986614	0.026623	0.026592	-0.052532	-0.052472
2.5 <sup>-</sup>	0.986599	0.986614	0.024931	0.026592	-0.063039	-0.052472
3.0 <sup>+</sup>	0.994970	0.995055	0.010034	0.009866	-0.019968	-0.019634
3.0 <sup>-</sup>	0.994970	0.995055	0.009109	0.009866	-0.025581	-0.019634

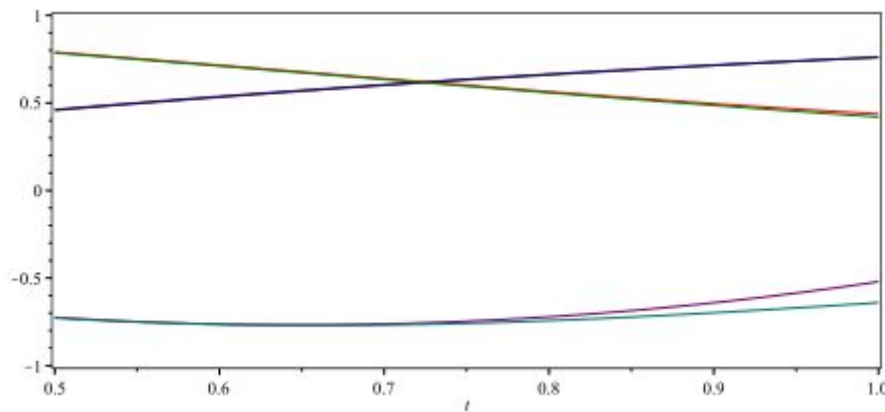
Error analysis: We know that the solution of HAM is analytic and the Piecewise homotopy analysis solution  $a(t)$ 's accurate is  $O(0.5^4)$  when  $u_0 = 0$ , or  $O(0.5^3)$  when  $u_0 \neq 0$ . The first-order derivative of the approximate solution  $a(t)$ 's accurate is  $O(0.5^3)$  when  $u_0 = 0$ , or  $O(0.5^2)$  when  $u_0 \neq 0$ .



$u_0 = 0$ . The second-order derivative of the approximate solution  $a(t)$ 's accurate is  $O(0.5^2)$  when  $u_0 \neq 0$ , or  $O(0.5^1)$  when  $u_0 = 0$ . All of those is showed on Figure 4, and the graphs of  $a(t), a'(t), a''(t), t \in (0.5, 1.0]$  are showed on Figure 4.

Table I shows that the error of  $a(t)$  is satisfied with the conjecture. For instance, when  $t = 0.5^-$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.003784(\leq 0.5^3)$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.036448(\leq 0.5^2)$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.273138(\leq 0.5^1)$ ; when  $t = 1.0^-$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.0006476(\leq 0.5^4)$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.017533(\leq 0.5^3)$ , the error between  $a(t)$  and  $\tanh(t)$  is  $0.119422(\leq 0.5^2)$ , and so do the others value of  $t$ . Moreover, the errors of  $a^{(n)}(t^+), n = 0, 1, 2$ . is smaller than the errors of  $a^{(n)}(t^-), n = 0, 1, 2$ .

Figure 4 shows the errors in  $(0.5, 1.0]$ . From the graph, we find that the larger value of  $t$ , the bigger error of every  $a(t), a'(t), a''(t)$ . And of course, to any  $t \in (0.5, 1.0]$ , the error of  $a(t)$  is smaller than the error of  $a'(t)$  and the error of  $a'(t)$  is smaller than the error of  $a''(t)$ .



**Figure 4.** From left top to bottom, the middle two lines is the exact solution and the approximate solution, whose precision is  $0.5^4$ ; the first two lines is the first-order derivative of the exact solution and the approximate solution, whose precision is  $0.5^3$ ; the last two lines is the second-order derivative of the exact solution and the approximate solution, whose precision is  $0.5^2$ .

**Piecewise Homotopy Analysis Method for Duffing Equation**

Let's consider the following Duffing equation as an example,

$$\ddot{u} + \frac{u^3}{1+u^2} = 0, \quad t > 0, \tag{16}$$

with initial conditions  $\dot{u}(0) = 0, u(0) = A$ . Clearly, Eq.(16) can be deformed into

$$\ddot{u} + \ddot{u}u^2 + u^3 = 0, \tag{17}$$

In order to use the Piecewise homotopy analysis method, we should apply the HAM to Eq.(17) in the first place.

The zeroth-order homotopy equation is

$$(1 - q)L[\phi(t, q) - u_0(t)] = hqH(t)N[\phi(t, q)] \tag{18}$$

with  $\phi(0, q) = 0, \phi(0, q) = A$ , where

$$N[\phi(t, q)] = \ddot{\phi}(t, q) + \ddot{\phi}(t, q)\phi^2(t, q) + \phi^3(t, q).$$

Take

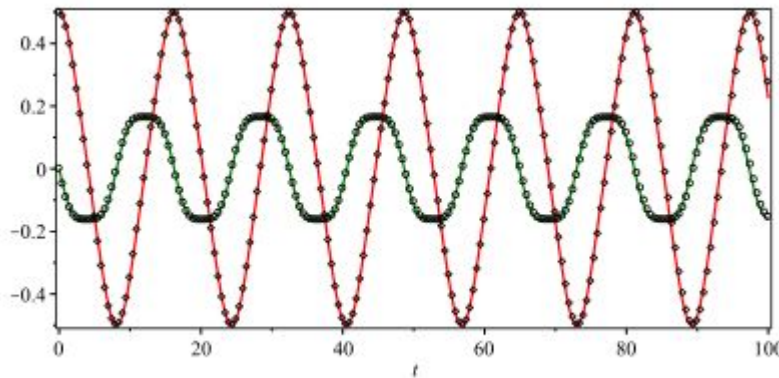
$$\phi(t, q) = u_0(t) + \sum_{k=0}^{\infty} u_k(t)q^k, L = \frac{\partial^2}{\partial t^2}, h = -1, H(t) = 1$$

into Eq. (18), then differentiating the Eq. (18) k times with respect to q and dividing by k! at last, we obtain the kth-order deformation equation

$$\ddot{u}_k - \chi_k \ddot{u}_{k-1} = -\ddot{u}_{k-1} - \sum_{j=0}^{k-1} \sum_{i=0}^j (\ddot{u}_i + u_i)u_{j-i}u_{k-1-j} + (1 - \chi_k).$$

u(t)'s mth-order homotopy analysis solution is

$$\tilde{a}_0(t) = \sum_{k=0}^m u_k(t).$$



**Figure 5.** The rhombus points are Eq.(16)'s 4th-order Piecewise homotopy analysis solution  $\tilde{a}_0(t)$ , the red line is the 4th-order homotopy analysis solution  $\tilde{a}_0(t)$ ; the cycle points are the first-order derivative of  $\tilde{a}_0(t)$ , the green line is the first-order derivative of  $\tilde{a}_0(t)$ . From the above equations, we know  $\tilde{a}_0(t)$  is only based on  $u_0(t)$  and  $m$ . Of course, we always want to take a big value of  $m$  to get a better  $\tilde{a}_0(t)$ . Due to the limitations of the paper, we get the 4th-order  $\tilde{a}_0(t)$  with initial data  $u_0 = a + bt$ .

4th-order  $\tilde{a}_0(t)$  homotopy analysis solution with the initial data  $u_0 = a + bt$  is

$$\begin{aligned} \tilde{a}_0(t) = & a + bt + \left(\frac{1}{2}a^5 - \frac{1}{2}a^3 - \frac{1}{2}a^7 + \frac{1}{2}a^9\right)t^2 + \left(\frac{3}{2}a^8b - \frac{1}{2}a^2b + \frac{5}{6}a^4b - \frac{7}{6}a^6b\right)t^3 \\ & + \dots \\ & + \left(\frac{211}{416000}b^7a^2 + \frac{271}{1108800}b^9\right)t^{15} + \frac{211}{3328000}b^8at^{16} + \left(\frac{211}{56576000}\right)b^9t^{17} \end{aligned}$$

The first initial data  $\dot{u}(0) = 0, u(0) = A$  can be obtained when  $a = 0, b = A$ , and in this situation,

$$\begin{aligned} \tilde{a}_0(t) = & At - \frac{1}{20}A^3t^5 + \frac{1}{42}A^5t^7 + \left(\frac{1}{480}A^5 - \frac{1}{72}A^7\right)t^9 + \left(-\frac{9}{3080}A^7 + \frac{1}{110}A^9\right)t^{11} \\ & + \left(\frac{11}{3360}A^9 - \frac{11}{124800}A^7\right)t^{13} + \frac{271}{108800}A^9t^{15} + \frac{211}{56576000}A^9t^{17} \end{aligned} \tag{19}$$

According to the above 4th-order homotopy analysis solution  $\tilde{a}_0(t)$ , we can use the Piecewise homotopy analysis method to solve Eq. (16). Here we also take  $t_0 = 0.5, t_1 = 1.0, \dots, t_k = 0.5(k + 1), \dots$

to get the different initial data of the corresponding intervals, then we can obtain the 4th-order Piecewise homotopy analysis solution  $\tilde{a}_0(t)$  of the corresponding intervals. As is showed in figure 5.

Figure 5 shows the comparison of Eq.(16)'s 4th-order Piecewise homotopy analysis solution  $\tilde{a}_0(t)$  and 4th-order homotopy analysis solution  $a_0(t)$  on  $[0,100]$ , and the comparison of the two solution's first-order derivative on  $[0,100]$  when  $A=0.5$ .

## Conclusions

In this paper, we first prove that the closer the value of  $h(< 0)$  is to zero, the larger the convergence region is to  $+\infty$  for initial-value problems. Based on the convergence, another analysis method called Piecewise homotopy analysis method is proposed to solve nonlinear equations.

Compared HAM with Piecewise homotopy analysis method through two examples, we find the latter has the following characteristics: First of all, Piecewise homotopy analysis method has all the characteristics that HAM has. Second, Piecewise homotopy analysis method could get a better solution with a relatively small value of  $m$  than HAM for large  $t$ , it over comes a shortcoming in HAM that the farther point's approximation effect is not as good as the point near the initial point, and for this reason the calculations are reduced to get the same accuracy of large  $t$ . Moreover, Piecewise homotopy analysis method not only has a higher accuracy than HAM, but its derivatives have an acceptable accuracy. In addition, Piecewise homotopy analysis method can solve equations with arbitrary initial data, which is depend on the step size. At last, a higher accuracy can get by shortening the tep size.

## References

- [1]. S.J. Liao. The proposed homotopy analysis technique for the solution of nonlinear problems. Ph.D. Thesis, 1992.
- [2]. S.J. Liao. Beyond perturbation: Introduction to the homotopy analysis method. Chapman and Hall/CRC Press, Boca Raton, 2003.
- [3]. Y. Tan and S. Abbasbandy. Homotopy analysis method for quadratic riccati differential equation. Communications in Nonlinear Science and Numerical Simulation, 13(3):539 – 546, 2008.
- [4]. H. Xu and J.. Cang. Analysis of a time fractional wave-like equation with the homotopy analysis method. Physics Letters A, 372(8):1250 – 1255, 2008.
- [5]. S.J. Liao. An analytic approximate approach for free oscillations of self-excited systems. International Journal of Non-Linear Mechanics, 39(2):271 – 280, 2004.
- [6]. S.J. Liao. Homotopy analysis of nonlinear progressive waves in deep water. Journal of engineering mathematics, 2003.
- [7]. S.j. Liao. A kind of approximate solution technique which does not depend upon small parameters ii. an application in fluid mechanics. International Journal of NonLinear Mechanics, 32(5):815 – 822, 1997.
- [8]. C. S. Liu. The essence of the generalized taylor theorem as the foundation of the homotopy analysis method. Communications in Nonlinear Science and Numerical Simulation, 16(3):1254 – 1262, 2011.
- [9]. A. Abdelrazec and D. Pelinovsky. Convergence of the adomian decomposition method for initial-value problems. Numerical Methods for Partial Differential Equations, 27(4):749–766, 2011.

## Test of the GEM Front Tracker for the SBS Spectrometer at JLab

F. Mammoliti<sup>1,2</sup>, V. Bellini<sup>1,2</sup>, M. Capogni<sup>3</sup>, E. Cisbani<sup>4,5</sup>, P. Musico<sup>6</sup>, F. Noto<sup>1,2</sup>, R. Perrino<sup>7</sup>, L.Re<sup>1,2</sup>, S. Spinali<sup>1,2</sup>, M.C. Sutera<sup>2</sup>

<sup>1</sup>*Dipartimento di Fisica ed Astronomia, Università di Catania, via Santa Sofia 64, I-95123 Catania, Italy*

<sup>2</sup>*INFN - Sezione di Catania, via Santa Sofia 64, I-95123 Catania, Italy*

<sup>3</sup>*Enea Istituto Nazionale di Metrologia delle Radiazioni Ionizzanti(INMRI-ENEA), C.R.Casaccia, C.P. 2400 I-00100 Roma, Italy*

<sup>4</sup>*INFN - Sezione di Roma La Sapienza, P.le Aldo Moro, 2 I-00185 Roma, Italy*

<sup>5</sup>*Istituto Superiore di Sanità, viale Regina Elena 299, I-00161 Roma, Italy*

<sup>6</sup>*INFN – Sezione di Genova, via Dodecaneso 33, I-16146 Genova, Italy*

<sup>7</sup>*INFN – Sezione di Lecce, via Arnesano, I-73100 Lecce, Italy*

Corresponding Author: Francesco Mammoliti, Dipartimento di Fisica e Astronomia Via Santa Sofia 64, 95123 Catania (Italy). Email: francesco.mammoliti@ct.infn.it, tel.: +390953785266.

A new Large-Acceptance Forward Angle Spectrometer (Super BigBite) is under development at JLab/Hall A to optimally exploit the exciting opportunities offered by the 12 GeV upgrade of the electron beam. The tracking of this new apparatus is based on the Gas Electron Multiplier (GEM) technology, which has been chosen to optimize cost/performance, position resolution and meet the high hits rate ( $>1$  MHz/cm<sup>2</sup>).

The first GEM detector were designed, built and tested, during different periods, at the DESY test beam facility in Hamburg, by using an electron beam with an energy from 2.0 to 6.0 GeV. In particular, different chambers, with a dimension of 40x50 cm<sup>2</sup>, were equipped with a new implementation of the APV25 readout chip. Measurements were performed at different impact points and angles between the electron beam and the plane of the GEM chambers.

In this report we present the technical characteristics of the detector and comment on the presently achieved performance.

**Keywords:** GEM Detector, Tracker.

### INTRODUCTION

The Jefferson Laboratory [1] (JLab) is one of the most important experimental facility providing a multi GeV, high intensity, longitudinally polarized, electron beam. The laboratory is undergoing a major upgrade of its CEBAF electron beam and experimental halls. In late 2013, CEBAF will deliver electron with energy up 12 GeV (twice the present limit) with excellent intensity (up to 100  $\mu$ A) and longitudinal polarization (up to 85%). In order to take advantage of the new scenario, the equipments of the 3 existing experimental Halls are under upgrading to optimally exploit the opportunities of the new beam. In particular members of Hall A collaboration are developing a new reconfigurable spectrometer, the Super BigBite (SBS [2], Fig. 1), featuring very forward angle (down to 7 degree), large momentum (2-10 GeV/c) and angular (64 mrad) acceptance, high rate capability (1 MHz/cm<sup>2</sup>) and very high luminosity environment (up to  $10^{39}/(s \cdot cm^2)$ ). The new spectrometer will consist, in its full configuration, of a dipole magnet with field integral up to 3 T·m (it will operate at about 2 T·m), a primary charged particle tracker (first tracker), 2 identical proton polarimeters (made of a Carbon analyzer and large tracker), and an hadron calorimeter. SBS will initially serve 4 experiments [3] dedicated to the study of the nucleon structure in terms of elastic

electromagnetic form factors at high 4-momentum transfer  $Q^2$  up to  $15 \text{ GeV}^2$  and of transverse momentum distributions of the quarks in the SIDIS (Semi Inclusive Deep Inelastic Scattering) region. The tracking systems of SBS will be mainly based on GEM chambers. In the next section the main features of the SBS tracker and of the GEM detector will be presented and finally the preliminary results of the test performed at DESY will be discussed.

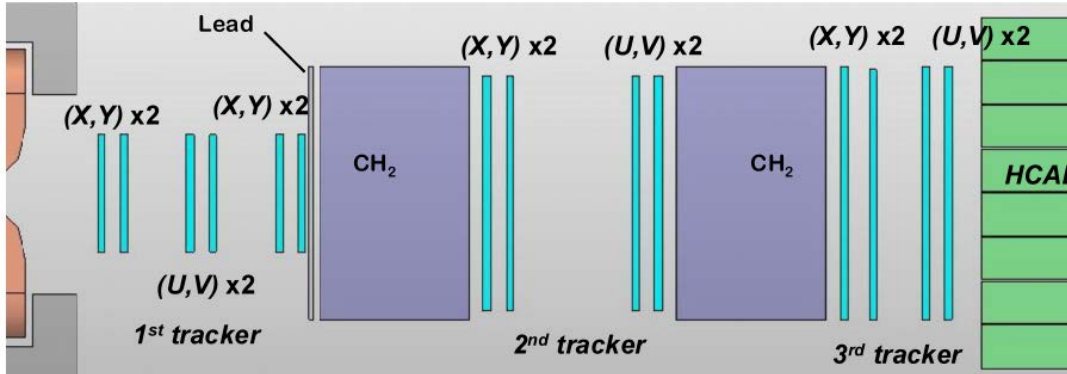
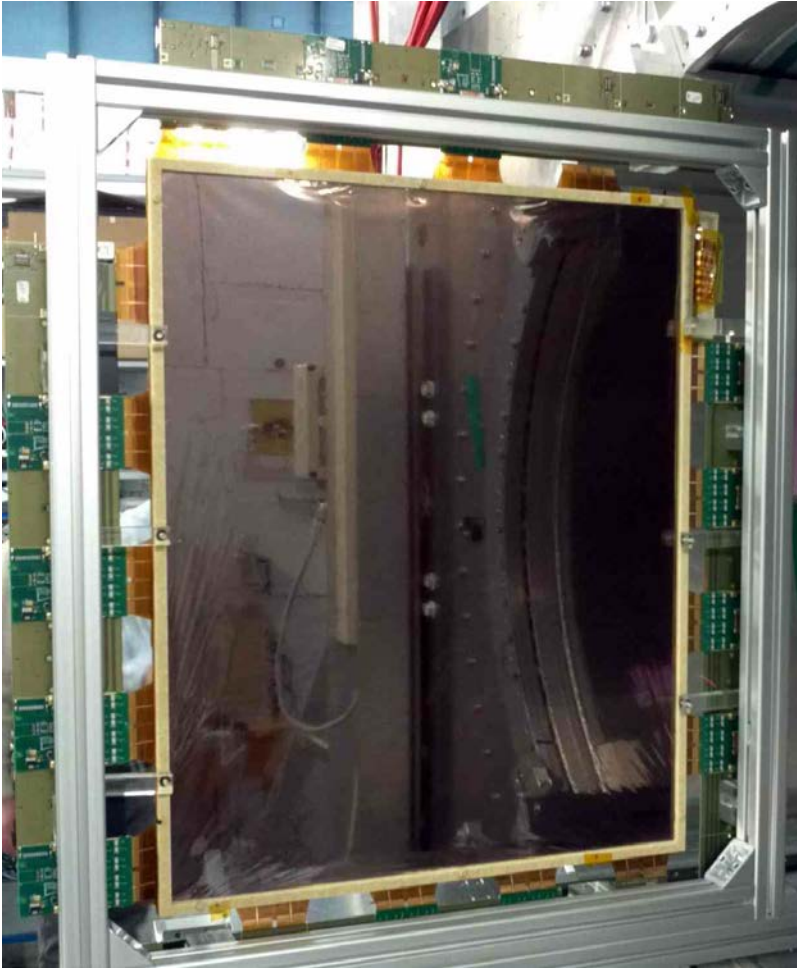


Figure 1. Schematic layout of the SBS Spectrometer.

## SBS TRACKER AND GEM DETECTOR

The SBS tracking system is made of three stations. The primary (front) tracker, placed just after the dipole momentum analyzing magnet, will consist of six large area ( $40 \times 150 \text{ cm}^2$ ) and high resolution ( $\sim 70 \mu\text{m}$ ) GEM chambers, for a total tracker length of about 50 cm. Each chamber is made by 3 adjacent GEM modules of  $40 \times 50 \text{ cm}^2$  active rectangular area, for a total of 18 modules. It is designed to be capable to track accurately particles emerging from the electron scattering in a large background of soft photons ( $\sim 0.5 \text{ MHz/cm}^2$ ) and MIPs ( $\sim 0.2 \text{ MHz/cm}^2$ ). The primary tracking will be reinforced by combination with two small ( $10 \times 20 \text{ mm}^2$ ) planes of silicon  $\mu$ strips placed in proximity of the target. The other stations are meant to track particles after a polarization analyzer wall and will require less accuracy. The primary tracker is under the responsibility of INFN groups. GEM technology [4] has been chosen to optimize cost/performance, position resolution and meet the high rate ( $> 1 \text{ MHz/cm}^2$ )[5]. The single module is made of 3 GEM foils and double layer x/y strips readout with  $400 \mu\text{m}$  strip pitch (figure 2). The 8 mm wide mechanical frame incorporates high voltage feeding protection resistors and gas inlet/outlet holes. The signals from each triple GEM module are read out in two coordinates through COMPASS-like [6] strip conductors planes. The front-end electronics [7] (FE) for the  $\sim 100\text{K}$  channels of the tracker is based on the APV25 [8] chip, successfully used in the LHC experiment CMS. The APV25 is a serial output analogue ASIC running at 40 MHz. The FE cards, each with 128 channels, are placed around the GEM module. Custom backplanes are used to distribute power and control to the FE cards and to collect the analogue outputs.



**Figure 2. Fully equipped GEM module.**

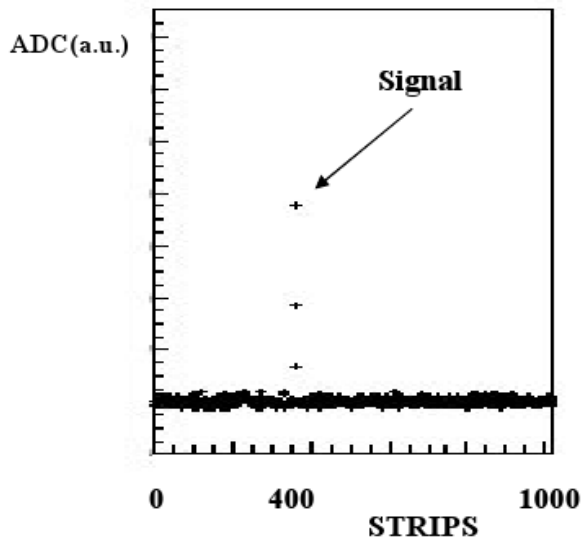
In Figure 2 a fully equipped  $40 \times 50 \text{ cm}^2$  GEM module prototype setup under test at DESY is shown. The module is equipped with the APV25 electronics and 18 front-end cards are located behind the 4 rectangular backplanes that sit along the 4 sides of the module. During the test, a gas mixture of Ar (70%) and  $\text{CO}_2$  (30%) has been used and HV has been powered by the first version of the HV-GEM system [9] providing 7 independent HV levels. Moreover, precise tracking has been performed by small silicon strip detectors located before the GEM.

The test has been performed in the T22 DESY Test electron/positron beam area [10]. The test beam is originated from the lepton synchrotron DESY II by converted bremsstrahlung on carbon fiber target. The energy of the beam varies between about 1 and 6 GeV/c with typical intensity of 1000 particle/( $\text{s} \cdot \text{cm}^2$ ) (divergence is about 2 mrad).

## DATA ANALYSIS AND DISCUSSION

In this section we present preliminary results of data analysis performed on about hundreds of beam runs obtained by using three GEM chamber prototypes with a dimension of two of  $40 \times 50 \text{ cm}^2$ . All chambers were readout by the APV electronics which were under development at the same time. During the test, different configuration have been used: energy of the electron beam (from 2.0 to 6.0 GeV), HV settings, angle between the beam and the plane of the chamber and position of the chamber with respect to the beam. Moreover, in order to have pedestals, without beam runs were

acquired.

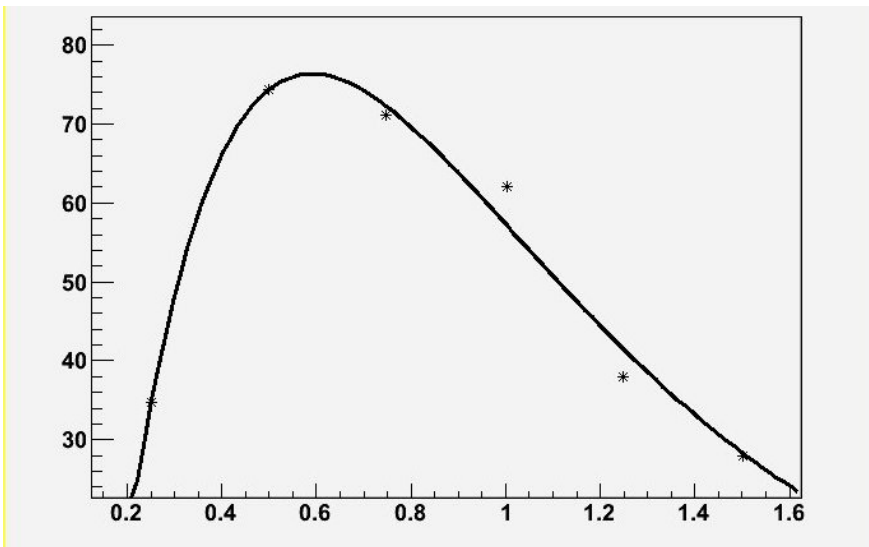


**Figure 3. Single electron signal in the x-direction.**

A single signal in the x-direction is shown in figure 3: it is clearly visible at about strip #400 and it is obtained after the pedestal subtraction. By using APV 25 chips, it is possible to register different parts of the signal (every 25 ns), event by event. The shape of the signal was fitted by using the formula:

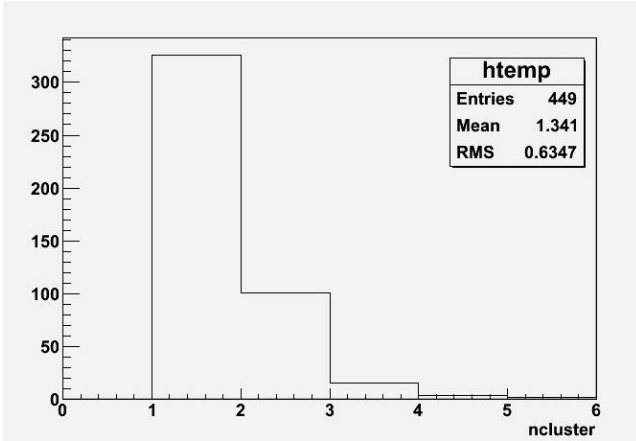
$$A \left( 1 - e^{-\frac{(t-t_0)}{\tau_1}} \right) \cdot e^{-\frac{(t-t_0)}{\tau_2}}$$

in which  $\tau_1$  and  $\tau_2$  are the slope and falling time of the signal, respectively,  $t_0$  is the stop time and A is the Amplitude (see figure 4).

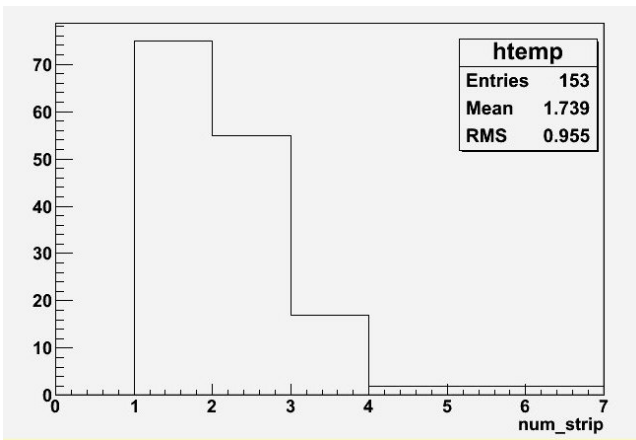


**Figure 4. Signal shape: time is in the x-direction and ADC in the y-direction, respectively.**

Adjacent firing strips are grouped in clusters and both number of clusters (figure 5) and number of strips (figure 6) of each cluster was evaluated. In both cases, distributions are consistent with the data from COMPASS GEM characterization [11].

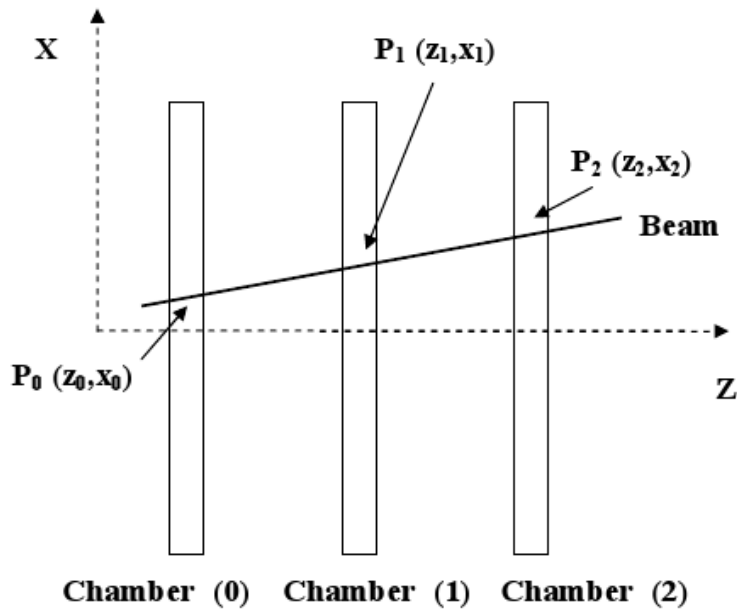


**Figure 5. Number of clusters.**



**Figure 6. Number of strips on each cluster.**





**Figure 7. Schematic layout of the 3 chambers.**

The schematic layout of the 3 chambers during the test is shown in figure 7, in which chamber #0, #1 and #2 have an area of  $40 \times 50 \text{ cm}^2$ . In order to select the single events of a run, we check if there is a cluster on each chamber in the x-direction. Each cluster provides the hit position  $P_n(z_n, x_n)$  and  $\sigma_n$ , where  $n$  is the index of the chamber. By using two points, for example  $P_0(z_0, x_0)$  and  $P_1(z_1, x_1)$ , a straight line  $X = a \cdot Z + b$  is reconstructed in the two-dimensional space ( $a$  and  $b$  are obtained by a linear fit). Finally, we consider  $P_2(z_2, x_2)$  and if  $|x_2 - az_2 - b| < \sigma_2$  than the signal of the 3 chambers belongs to the same particle otherwise it is rejected. As told before, we have used 3 chambers and an event is good if there is a cluster in all chambers otherwise it is rejected. Finally we define efficiency the ratio between the number of good events and the total number of events given by the trigger. In particular we have found an efficiency of about 90%.

## CONCLUSIONS

The main purpose of the test was to verify the overall functionality of the main solution adopted in the first GEM prototype under simplified beam conditions. Both GEM hardware and readout electronics were under early development and therefore final results on efficiency and chamber resolution were not very indicative. Anyway, the GEM chambers operated fairly stably during the test and the preliminary results show reasonable indications of the general validity of the adopted solutions for the distribution of the collected charge. The data analysis also pointed out some critical aspects to be further investigated.

## REFERENCES

- [1] Thomas Jefferson National Accelerator Facility, [www.jlab.org](http://www.jlab.org).
- [2] The SBS project web site: <http://hallaweb.jlab.org/12GeV/SuperBigBite/>
- [3] 12GeV HallA upgrade web site: <http://hallaweb.jlab.org/12GeV/>.
- [4] Sauli F., GEM: A new concept for electron amplification in gas detectors, *Nucl. Instr. Meth. A*

386 (1997), p. 531.

[5] Bellini V. et al., GEM tracker for high luminosity experiments at the JLab Hall A, proceedings of the MPGD 2012 conference, published on JINST.

[6] Ketzer B., *Nucl. Instr. Meth. A* 494 (2002), p.142.

[7] Musico p. et al., 2011 IEEE Nuclear Science Symposium Conference Record.

[8] French M.J. et al., Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker, *Nucl. Instrum. Meth. A* 466 (2002), p. 359.

[9] 12GeV HallA upgrade web site: <http://hallaweb.jlab.org/12GeV/>.

[10] DESY Test Beams web site: <http://adweb.desy.de/~testbeam/>

[11] Altunas C. et al., CERN-EP/2002-008, CERN, 21 January 2002.

# PROBABILISTIC STRUCTURAL ANALYSIS OF THE COMPOSITE CREW MODULE- SUBSTRUCTURING WITH HIGH RESOLUTION GRID

<sup>1</sup>Dr. Shantaram S. Pai, <sup>2</sup>Dr. Vinod K. Nagpal\*, and <sup>2</sup>Ian Miller

<sup>1</sup>NASA Glenn Research Center, Cleveland, Ohio, USA.

<sup>2</sup>N&R Engineering and Management Service Corporation, Parma Heights, Ohio, USA.

\*Corresponding author: vnagpal@nrengineering.com

## ABSTRACT

A probabilistic structural analysis of an experimental Composite Crew Module (CCM) has been performed using the most current Orion Crew Module design loads. The purpose of the analysis was to quantify the risk of failure for different factors of safety in response to uncertainties in the design variables.

The CCM was found to be very safe with ample margin of safety and a low probability of failure. As a further analysis activity, the loads were scaled up and the composite material's strength was scaled down to assess a bounding scenario; still the design was found to be safe.

This analysis shows that for given factor of safety and uncertainty in the design space, the probability of failure associated with the design can be significantly different. Further, designing structures based on factor of safety without quantifying the influence of uncertainties in the design variables can lead to structures with unknown risk.

**KEY WORDS:** Composite Analysis, Probability of Failure, Composite Crew Module, Space Vehicle, Composite Failure Theories Applications

## INTRODUCTION

NASA is evaluating the use of composite materials for building its next generation crew exploration vehicle (CEV) [Kirsch 2011]. In addition, NASA is interested in assessing the applicability of probabilistic methods for developing reliability based designs of components, structures and systems made from composite materials. For these two reasons, NASA initiated a task to probabilistically analyze the composite material that was developed for the new crew module. This paper describes a procedure for estimating the reliability of the CCM structure given various missions.

Most design practices use deterministic analysis combined with factors of safety (FOS) to provide a margin between the assumed peak load and a mean failure point. Unfortunately, this approach does not provide any information on the potential failure rate or reliability of the system. Factors of safety do not provide any information which relates the loading range relative to the structural allowable range. In reality, there is statistical scatter in both the loading variables and the structure's ability to resist those loads. Hence designs using deterministic analysis combined with FOS can lead to an overly conservative design with excess weight and cost or under design with unknown magnitude of risk.

Probabilistic analyses quantify the effect that uncertainties in the input variables have on the response variables. As a result, the probabilistic analyses provide quantitative information to enable designers to either calculate a realistic risk or the actual FOS in an existing design. Thus the designer can improve the design by either selecting a predetermined risk or adjusting the FOS to meet the predetermined risk. The authors [Nagpal et al 2010 and Pai 2008] have used probabilistic approaches in quantifying probability of

failure and performing sensitivity analyses of different structural problems and furthermore demonstrated its application to improving design practices.

The probability of failure, a measure of reliability, is critical for determining the structural integrity of the vehicle under the various flight and ground loads encountered during operation. Performing an accurate estimation for the probability of failure is more critical now than before because of recent changes in NASA's requirements. The objective is to reduce the loss-of-crew/loss-of-vehicle rate by a factor of 2 to 3. In addition, the cost/pound of payload is also to be reduced by the same factor. These requirements present a formidable challenge given the rigors and uncertainties of space flight coupled with possible future budget constraints.

The use of Polymer Matrix Composite (PMC) materials for structural components is steadily increasing particularly for aerospace applications [Harris et al, 2008]. The key advantages of using composite materials are that they are light weight and can be manufactured with desired properties [NASA SP-8108]. However, the fabrication process of a component made from composite materials is quite complex and uniquely developed specifically for a given structural component [NASA PD-ED-1217]. This complexity in the fabrication process is due to a large number of variables such as fiber volume ratio, ply orientation, density, type and strength of fibers and other variables.

Designing, components made from composite materials using probabilistic methods requires the characterization of uncertainties in the material properties, fabrication processes and loading variables. In addition, there are further uncertainties associated with the degradation of composite materials in a space environment.

Probabilistic methods using physics-of-failure approaches provide valuable quantitative information which is not only used to develop a risk-based design, but also provides the probability of failure for a design which has been developed using a factor of safety, Figure 1. This figure displays the relationship between this CCM model's factor of safety and the probability of failure. There is an inverse relationship between the probability of failure and the factor of safety. That being, demanding a larger factor of safety leads to a decrease in the probability of failure. For human rated systems NASA HQ requires that the probability of failure be no greater than  $1.0E-04$ . Due to the safe nature of this design, this study estimates that one could relax the safety of factor to just below 1.4 and still satisfy the risk requirement. Figure 1 assumes that the uncertainties in the design variables are known to some extent. Later it will be shown how to deal with the situation when the design variables have a large amount of variance.

Inherent uncertainties in material properties, loads and geometry complicate the design process [Andrzej, Nowak, and Collins, 2000] and affect safety, performance and cost. Traditionally, factors of safety (FOS) have been developed empirically based on years of engineering experience and are assumed to account for these uncertainties [Melchers 1999]. Any design based on the FOS lacks quantitative information on risk of failure, consequently it remains unknown whether the design is within the acceptable risk limits or not.

In the absence of expensive-to-acquire experimental data, a probabilistic approach provides extremely valuable quantitative information for generating cost effective risk-based design and additionally provides precise direction for further resource investments. Uncertainties in the design variables are defined through probability distributions which are used in conjunction with a probabilistic methodology to propagate these uncertainties up to a probabilistic system response. Figure 2 shows a family of curves that relate the probability of failure with an assumed uncertainty in the design variables. As the uncertainty in the design variables increase, for a given factor of safety, the probability of failure also increases. The probabilistic analysis was first conducted assuming that all the material properties had 5% uncertainty, for which a probability of failure is then produced. The full probabilistic analysis is again conducted for each: 10%, 15%, 18%, 20%, and 40% uncertainty in the design variables. The 5% and 10% cases are considered to represent realistic scatter in the material properties, due to fabrication issues, whereas the other cases represent an attempt to envelope the design space to find cost savings based on less stringent manufacturing quality controls. All six test cases were used to generate the curve named "NO Cov", where all the design variables were assumed to be independent random variables. The entire procedure

described above is repeated. This time assuming that the design variables are not fully independent but have 10% and 30% mutual correlation.

### **Technical Approach**

The following procedure was used to quantify the probability of failure. First to be discussed is the deterministic finite element model; its material properties and the loads that were analyzed in this study. Next, several probabilistic analyses of the finite element model is presented, each case assuming different levels of uncertainty in the design variables. Each probabilistic analysis generates a distribution for the stresses or strains in a region of interest to the designer. These stress/strain distributions are input into the reliability calculation, which is used to determine a probability of failure. The probability of failure is estimated by considering the area under a material level limit state equation such as  $(1 - \frac{\text{stress}}{\text{strength}})$ . This type of the limit state equation is more realistic when the stresses and the material properties' allowables are not statistically independent, as can be explicitly seen in the Hoffman's failure criterion.

### ***Finite element model***

The finite element model analyzed was provided by the NASA Langley Research Center, Figure 3. The ANSYS workbench translator was used to convert the model from a NASTRAN format. The translated model was then verified for its accuracy of translation by comparing results of the original model obtained using ANSYS and NASTRAN. This comparison showed that the values of the maximum stress from the two analyses were within 0.25%.

The model was of the CCM pressure vessel's outer structure and included the mass of the secondary structure. This ensured that the loads were accurate for the mass and center of gravity of the CCM with a payload. A preliminary global analysis was performed using the material properties provided by NASA. From the results of this analysis, regions of high stresses, also called critical regions, in the structure were identified. The finite element meshes in these regions were further refined to perform high resolution analysis

Composite shell elements with orthotropic material properties were used to model the structure. The properties indicating the thickness of each ply, fiber angle and materials in the refined mesh area are given in Table 1. The CCM wall is made of 10 composite material plies and one aluminum honeycomb layer ("Shapegrid" TradeMark). The honeycomb layer, numbered as 6, is sandwiched between layers of composite plies numbered 1-5 and 7-11, as listed in Table 1. The composite plies are made of two different materials; plies 3 and 9 are made of one type of material, labeled as 1, and plies 1, 2, 4, 5, 7, 8, 10 and 11 are made of another type of material, labeled as 2. The composite plies are designed using high strength carbon fiber, IM7 and Cytec's 977-2 epoxy. Properties for both types of materials 1 and 2 including their elastic properties and allowable strengths are listed in Table 5. From these properties, it is obvious that material 1 is considerably stiffer and stronger than material 2. The honeycomb material, labeled as material 6, is a filler material and it doesn't have appreciable stiffness and strength properties.

### ***Loading and boundary conditions***

The CCM is subjected to three different loading conditions during its flight and landing. The loadings were provided by the NASA Johnson Space Flight Center as adaptations of the Orion Crew Vehicle loadings. In all cases the CCM is subjected to 15.2 psi internal pressure. This analysis considers only the first two of the three loading conditions, see Table 2. Since the original FE model was designed for the launch abort and docking load conditions. For these two loading conditions, the analysis includes 5 sub-loading conditions for abort loads and 8 sub-loading conditions for docking loads. The sub-loadings represent different variations in the two loadings. The two loading conditions considered are:

1. High altitude abort loads with five (5) sub-load conditions.
2. On-orbit docking loads with eight (8) sub-load conditions.

Both the abort and docking load conditions include acceleration loads. Where  $A_x$ ,  $A_y$  and  $A_z$  are accelerations in x (along the height), y (perpendicular to the paper) and z (horizontal in the plane of the paper) directions.

### ***Deterministic analyses***

Deterministic structural analyses were performed to identify the most critical regions with high stresses in the crew module, see Table 3. The high stress region is in the cylinder near its interface with the cone shape structure near one of the door openings, Figure 4.

Of the thirteen different load cases analyzed only two loads, high altitude abort loads and on-orbit docking loads were found to be the most critical loads. The most critical regions were identified using these two loads. A high resolution analysis at the ply level was conducted to determine peak stresses and identify potential failures.

Overall, the high altitude abort sub-load conditions resulted in higher stresses than the on-orbit docking sub-loading conditions. Since all high altitude abort sub-loading conditions are very similar in loading, variations in stresses among these sub-loading conditions are insignificant. Out of all the thirteen sub-loading conditions, the maximum stresses occurred in ply 3 under the fourth case, which was a high altitude abort sub-loading. Therefore this loading case was selected to perform the subsequent probabilistic analyses. The peak stress and type of ply failure were used in performing probabilistic structural analyses for several different levels of assumed uncertainties in the design variables. Ply 3 is the stiffest of all plies found in the critical stress region. Figure 4 illustrates stress distribution in ply 3, for the fourth high altitude abort load, without external and internal connections with the aero-shell and payloads.

### ***Failure criteria***

The probability of failure was estimated from the area which is both under a given failure criteria's PDF and to the left of the failure criteria limit of 0.0. A distribution for each of the following three ply level failure criteria was constructed from the probabilistic analysis: (1) inter-ply delamination criterion, (DELFC), (2) Hoffman's failure criterion (HFC) and (3) modified distortion energy criterion (MDE) [Robert Aiello, 1989], see Appendix A.

The MDE and HFC criteria have been expressed in terms of actual and allowable stresses and the DELFC has been expressed by actual and allowable strains. Allowable strain for DELFC is the shear strain between two plies. Plies 2, 3 and 5 are identified to have the most probable failures based on a deterministic analysis and thus selected for the probabilistic analyses, Table 4. The three ply level failure criteria were used to define the limit state equation in a "stress-strength formulation".

All three failure criteria have been calculated independently to estimate a probability of failure in the critical regions based on uncertainties in ten input design variables; three material properties and strengths (limiting stress) for each of materials 1 and 2 and the pressure and acceleration loads. There are five limiting stresses ( $\sigma_{xt}$ ,  $\sigma_{xc}$ ,  $\sigma_{yt}$ ,  $\sigma_{yc}$  and  $\sigma_{xy}$ ) for both material 1 and material 2.  $\sigma_x$ ,  $\sigma_y$  and  $\sigma_{xy}$  are the allowable stresses in both the x and y directions respectively. Further, the subscript (t) represents the materials' allowable in tension and (c) is the material's allowable in compression. The allowable with respect to shear stresses is denoted by  $\sigma_{xy}$ .

The DELFC is based on strains;  $\epsilon_x$ , in the x direction, and  $\epsilon_y$ , in the y direction, and shear strain,  $\epsilon_{xy}$  and shear strength of the epoxy which holds the plies together. DELFC will occur only if the actual shear strain in epoxy is more than allowable shear strain. The allowable shear strain is calculated from allowable shear stress, ( $S_{xy} = 7,000$  psi), Elastic Modulus ( $E_x = 500,000$  psi) and Poisson's ratio ( $\nu = 0.41$ ) using the relations  $G_{xy} = E_x / 2(1+\nu)$  and  $\Phi(\text{Allowable}) = S_{xy}/G_{xy}$ . This allowable shear deformation is calculated as 0.0395.

Table 4 shows the deterministic results for all three failure criteria for each ply in the critical region. For the HFC and MDE criteria, plies 2 and 3 have the lowest of the Failure Criteria (FC) values for both first

and second cases. On a scale from zero to one the MDE for ply 2 is 0.663. For the DELFC criterion, plies 2 and 5 have the lowest FC values. Since plies 3 and 9 have a zero fiber angles and are sandwiched between plies which also have a zero fiber angle, they have no response for DELFC. Ply 2 is next to ply 1 which has fiber angle of 45, and ply 5 has fiber angle of 45, they show the lowest FC values for DELFC response. Again low FC is more likely to fail than the plies with higher FC values.

Note at this stage, Table 4 gives point values for the three failure criteria. In the next section the probabilistic analysis will fit distributions to these criteria to quantify the expected scatter in the failure of the plies.

### ***Probabilistic analyses***

All design variables including material and geometric properties, boundary conditions and applied loading variables were selected as input variables for the probabilistic analysis. The input variables' inherent uncertainties are represented by a standard deviation. The design variables' uncertainties are then propagated through the finite element model of the CCM. The resulting response function, such as elastic stress, is then fit to a distribution. The distribution of the uncertainties in the peak stresses given systematic perturbations to all the relevant design variables are then inputted into the reliability calculations. A distribution for each failure criteria was generated based on the stress distribution in the critical region.

The analyses began by introducing uncertainty into the relevant state variables. A probabilistic discretization of the design space is created by assigning a probability distribution function (PDF) to each design variable, see Table 5. On successive sample runs of the finite element model, one uses these PDFs to determine the numerical values that the design variables take on. PDF and cumulative distribution function (CDF) for the response function are then constructed. This is accomplished by fitting a response surface, via a least squares methodology, through both the perturbed basic variables and the subsequent structural response.

This technique permits the entire spectrum of system response for the structure under uncertain loading conditions to be discovered. To arrive at this information traditional deterministic solution techniques were used, only the basic variables were modified to permit uncertainty in the model's design variables [Miller 2006].

### ***Results of probabilistic analyses***

The deterministic stress analysis results together with mean strengths and uncertainties of input variables were used to perform probabilistic structural analyses using FPI (Fast Probability Integrator) module of the NESSUS code. FPI determines the PDF of the response variable, stress, based on the uncertainties of the input variables and subsequently conducts the reliability analyses using the limit state function for all the failure criteria.

For the sake of the probabilistic analysis the material's components of strength ( $\sigma_{xt}$ ,  $\sigma_{xc}$ ,  $\sigma_{yt}$ ,  $\sigma_{yc}$  and  $\sigma_{xy}$ ) were considered to be one variable for each constituent ply material; they were perturbed concomitantly. All three components of the acceleration vector also increase or decrease simultaneously by the same proportion.

For the first probabilistic analysis, Table 5, all input variables were assumed to have uncertainties with standard deviations of 5%; the NASA experts concurred with this level of uncertainties. These standard deviations are considered realistic and represent a high quality fabrication of composites. The results for this level show that the PDF of the failure criteria does not overlap a limit state value of zero, see Figures 5-10. This means that there is practically no risk of failure. Hence, the design is found to be completely safe with respect to all three criteria of failure.

In the second probabilistic analysis, Table 5, the uncertainties standard deviations are doubled to 10% for all variables, twice the magnitude of uncertainties of those in the first scenario. This case represents a lower quality fabrication of composites. The uncertainties in this scenario are exaggerated intentionally in order to demonstrate the reserved safety of the current design. The HFC failure criterion PDF for ply 2

approaches the limiting value but does not cross it; once again implying a completely safe design. Figures for this scenario are not shown because of the page limitations.

Based on several scores of stress analyses for both the 5% and 10% uncertainty scenarios, normalized CDF and PDF using all three failure criteria for all the plies have been generated, more than 30 plots were created. These plots show the CDFs and PDFs of the three failure criteria (HFC, MDE and DELFC) along with their limit states for ply 2 for 5% uncertainties in the input variables. The limiting value of 0.0 for all the failure criteria is also shown in the PDF plots only. However, only three of these plots are shown here in Figures 5-10. These Figures are representative of the other 30 results which are not shown here to avoid repetitiveness of discussion. The results for all cases are shown in an unpublished report [N&R Engineering 2011] submitted to NASA.

### ***Trade study for reliability and cost effectiveness***

Since the analyses for both scenarios yield safe designs, the next natural step is to conduct a trade study with the objective of finding a lower cost composite material. In the trade study uncertainties in the design were systematically increased in set bounding scenarios: 15%, 18%, 20%, and 40%. This was done to provide guidelines for key managers on how they might be able to reallocate their resources more cost effectively. The uncertainties in these analyses have been highly exaggerated to show their quantitative impact on reliability and cost of a design. These analyses illustrate how probabilistic methods can assist the designer in selecting a lower cost composite i.e. composite fabricated using lenient tolerances while maintaining the desired level of reliability. Figure 2, discussed earlier shows a summary of results from the first two scenarios and this trade study.

### ***Simulating a fabrication processes for composites***

It is important to quantify the impact that the quality of a fabrication process can have. The next example demonstrates that for a given reliability a designer can select the appropriate tolerance level and corresponding cost for the design.

In this scenario, consider the case of three composites with same mean strength but fabricated with different quality standards; high, medium and low. A high quality composite implies tight-tolerances fabrication process, poor quality composite implies lenient-tolerances fabrication process and medium quality composite implies composite fabrication process with tolerances in between the two.

A graphical comparison of the composites with the same mean values of and hypothetically large uncertainties in all the input variables is shown in Figure 11. In this figure, there are 3 PDFs with the same mean value, a normalized mean of 0.588. The one with tightest scatter, green color, represents the case of tight and most expensive fabrication tolerances, and the one with widest scatter, blue color, represents the case of most lenient and inexpensive of the three cases fabrication tolerances. The third PDF, purple color, represents the case of tolerances in between the other two.

The three tolerances cases are represented by three standard deviations (SD); SD of 0.207 represents the case of tightest tolerance and SD of 0.287 the most lenient tolerance and SD of 0.239 the in-between tolerance. The quantified risk is area under the PDF to the left of the red vertical line. For the tightest tolerance case the risk is 0.0225%, most lenient tolerance case is 2.204% and for the in-between tolerance case is 0.696%.

This figure shows that for the same factor of safety, the risk in the design is significantly different and it illustrates that designing based on factor of safety without quantifying the influence of uncertainties could lead to a design with unknown risk. The factor-of-safety for all three tolerance cases is the same, 2.427. It is obvious from this demonstration that the level of risk would have been unknown if the probabilistic method was not used to quantify it.

Figure 12 illustrates a graphical comparison of the three PDFs representing three composite fabricated using the same tolerances. This example demonstrates how probabilistic structural analysis is used for accurately selecting an appropriate material for design. The analysis was performed using the same



magnitudes of uncertainties but with different mean values for the input variables. All three PDFs overlap the failure-criteria limit of zero, thereby allowing one to find the lower bound on acceptable materials. Such a study provides the relationship between the probability of failure and a factor of safety. Thus an analyst has more control over the amount of risk involved with the choice of a given factor of safety.

Also, this scenario demonstrates how probabilistic methods can be used in making key decisions such as setting cost-saving fabrication tolerances in addition to providing a quantitative estimate of the risk for a given factor of safety.

**Summary**

A probabilistic structural analysis of an experimental Composite Crew Module (CCM) of the future spacecrafts has been performed using the most current Orion Crew Module design loads obtained from the NASA Johnson Space Center. The purpose was to quantify the probability of failure for different factors of safety in response to uncertainties in the design variables. The results of these reliability calculations provide a quantitative means for selecting appropriate composite materials. An additional purpose was to demonstrate how probabilistic methods provide cost savings by allowing a reliability manager the ability to select proper factors of safety for predetermined risk.

Structural analyses were performed for 13 different load cases and the most critical load condition and the corresponding critical regions of high stresses were identified. A high resolution analysis at ply level was conducted in that region to determine peak stresses and identify potential failures. The peak stress and type of failure were then used in performing the probabilistic structural analyses.

Based on probabilistic structural analysis, the CCM is found to be very safe with ample margin of safety and a low probability of failure. As a further analysis activity, the loads were scaled up and the composite material’s strength was scaled down to assess a bounding scenario. Still the design is found to be safe.

Probabilistic methods demonstrate how to select materials for the structure based on setting tolerances (cost) and factor of safety for predetermined risk. Without using probabilistic methods, the reliability of the design remains unknown; in contrast to using a Factor of Safety method alone.

Finally probabilistic methods provide one more important piece of information for the key managers when optimizing the use of the critical resources. This information consists of sensitivities of the input variables such as geometry, material properties, etc., on the response variables such as stress. For improving the quality of the end products, managers need to know which resources have high sensitivities, and thus are controlling the structures’ performance or safety.

Table 1: Section Properties and Materials

Numbers of Layers = 11				
Total thickness = 1.0754 In.				
Ply (Layer)	Ply thickness in.	Material ID	Fiber angle	Ply Location
1	0.0081	2	45	↓ Inside
2	0.0081	2	0	
3	0.0053	1	0	
4	0.0081	2	0	
5	0.0081	2	45	
6	1	6	0	
7	0.0081	2	45	
8	0.0081	2	0	
9	0.0053	1	0	
10	0.0081	2	0	
11	0.0081	2	45	

Table 2: Composite Crew Module Loading Conditions

Load Cases	Internal Pressure (psi)	Gravity Acceleration X (in/s <sup>2</sup> )	Vehicle Accelerations						
			X (in/s <sup>2</sup> )	Y (in/s <sup>2</sup> )	Z (in/s <sup>2</sup> )	Rx (rad/s <sup>2</sup> )	Ry (rad/s <sup>2</sup> )	Rz (rad/s <sup>2</sup> )	
High Altitude Abort	1	15.2	-386.04	5408.401	11.54460	-354.3054			
	2	15.2	-386.04	5413.442	228.0736	-138.1016			
	3	15.2	-386.04	5385.366	226.8812	-568.2328			
	4	15.2	-386.04	5386.233	-203.9546	-568.3412			
	5	15.2	-386.04	5414.363	-205.0886	-138.1558			
Water Landing				G's					
	6		-386.04	348.836	279.488	305.435	The water landing load cases are not considered in this analyses.		
	7		-386.04	222.436	-192.225	307.354			
			-386.04	-110.446	-112.088	-103.202			
On-Orbit Docking							0.08	0.137	0.041
	9	15.2		11.1952	3.8604	4.24644	0.50265	0.86080	0.25761
	10	15.2		11.1952	-3.8604	4.24644	0.50265	0.86080	0.25761
	11	15.2		11.1952	3.8604	-4.24644	0.50265	0.86080	0.25761
	12	15.2		-11.1952	-3.8604	-4.24644	0.50265	0.86080	0.25761
	13	15.2		-11.1952	3.8604	4.24644	0.50265	0.86080	0.25761
	14	15.2		-11.1952	3.8604	-4.24644	0.50265	0.86080	0.25761
	15	15.2		11.1952	-3.8604	-4.24644	0.50265	0.86080	0.25761
16	15.2		-11.1952	-3.8604	4.24644	0.50265	0.86080	0.25761	

Load Case	Ply 1	Ply2	Ply3	
<b>High Altitude Abort</b>	Maximum Principal Stress (psi)			
	1	28,848	32,478	55,355
	2	28,269	32,432	54,947
	3	30,218	33,109	55,542
	4	29,362	32,491	56,377
	5	27,409	31,811	55,751
<b>On-Orbit Docking</b>	9	21,449	26,272	44,254
	10	22,228	26,154	44,371
	11	21,206	26,227	44,024
	12	22,497	26,270	44,537
	13	22,140	26,433	44,650
	14	21,717	26,388	44,420
	15	21,805	26,109	44,141
	16	22,920	26,315	44,767
<b>Pressure Only</b>		25,079	419,767	

Table 3: Typical Stress Distribution 11 plies

Table 4: Failure Criteria, calculated stresses and strains at the ply level for mean design values (5% uncertainty case)

Material	Ply	HFC	MDE	DELFC	SX	SY	SHY	eX	eY	eXY
2	1	0.990	0.913	1.000	6895.17	12319.16	-2938.25	6.13E-04	1.15E-03	-3.50E-05
2	2	0.694	0.663	0.963	27155.71	-8183.74	-448.21	2.61E-03	-8.70E-04	-5.34E-04
1	3	0.886	0.772	1.000	5932.79	-127.83	-388.22	2.99E-03	-8.70E-04	-5.34E-04
2	4	0.701	0.670	1.000	26770.13	-8199.09	-448.33	2.57E-03	-8.70E-04	-5.34E-04
2	5	0.951	0.918	0.950	6450.67	11878.02	-2870.68	5.72E-04	1.11E-03	-3.42E-05
3	6	1.000	1.000	0.981	0.00	0.00	0.00	1.18E-03	-8.68E-04	-5.39E-04
2	7	0.878	0.955	0.985	-8587.95	-3146.32	-569.47	-8.14E-04	-2.66E-04	-6.78E-04
2	8	0.872	0.949	0.999	-2703.86	-9372.75	-457.93	-2.25E-04	-8.81E-04	-5.45E-04
1	9	0.975	0.995	1.000	-9366.80	-1303.79	-376.21	-2.43E-04	-8.81E-04	-5.45E-04
2	10	0.869	0.949	1.000	-3089.46	-9388.10	-458.05	-2.51E-04	-8.81E-04	-5.45E-04
2	11	0.869	0.951	0.989	-9132.45	-3587.47	-501.91	-8.55E-04	-3.09E-04	-5.98E-04

Table 5: Material Properties, Loads and Standard Deviations used in the Probabilistic study

Variables with (%) Uncertainties and Standard Deviation.								
All values are in (psi) except accelerations Ax, Ay, Az in lb/in^2								
	Material	Variable	Mean Value	Case 1		Case 2		Direction
				Std. Dev.	%	Std. Dev.	%	
Elastic Modulus	1	Ex	22,950,000	1,147,500	5%	2,295,000	10%	X
		Ey	1,360,000	68,000	5%	136,000	10%	Y
		Gxy	690,000	34,500	5%	69,000	10%	XY
	2	Ex	10,533,000	526,650	5%	1,053,300	10%	X
		Ey	10,533,000	526,650	5%	1,053,300	10%	Y
		Gxy	840,000	42,000	5%	84,000	10%	XY
Allowable Stress	1	Sxt	124,779.00	12,477.90	10%	24,955.80	20%	Tension
		Sxc	85,787.10	8,578.71	10%	17,157.42	20%	Compression
		Syt	100,000.00	10,000.00	10%	20,000.00	20%	Tension
		Syc	100,000.00	10,000.00	10%	20,000.00	20%	Compression
		Sxy	12,870.00	1,287.00	10%	2,574.00	20%	Shear
	2	Sxt	57,265.20	5,726.52	10%	11,453.04	20%	Tension
		Sxc	39,370.49	3,937.05	10%	7,874.10	20%	Compression
		Syt	57,265.20	5,726.52	10%	11,453.04	20%	Tension
		Syc	39,370.49	3,937.05	10%	7,874.10	20%	Compression
		Sxy	12,870.00	1,287.00	10%	2,574.00	20%	Shear
	Pressure	Pr	15.20	0.76	5%	0.76	5%	
	Accelerations	Ax	-5,000	-1,000	20%	-2,000	40%	
Ay		203.95	40.79	20%	81.58	40%		
Az		568.34	113.67	20%	227.34	40%		

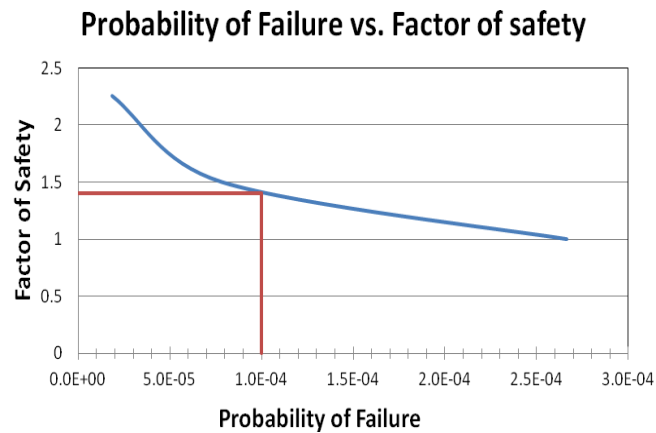


Figure 1: Probability of Failure vs. Factor of Safety

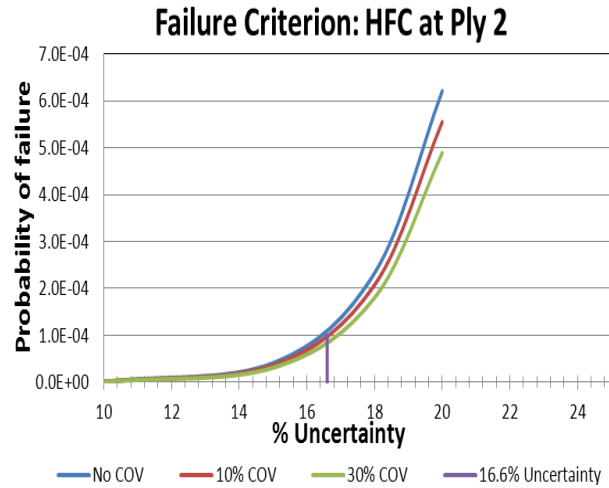


Figure 2: Probability of Failure as a function of the amount of uncertainty in the design space

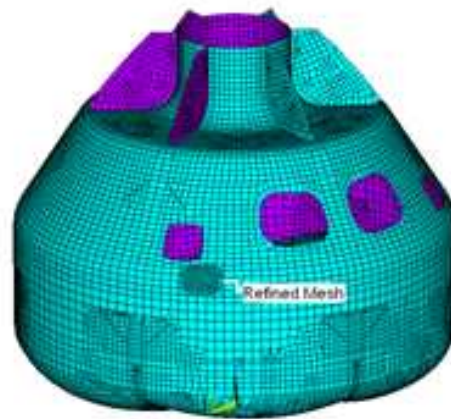


Figure 3: Refined Mesh for High Stress Location

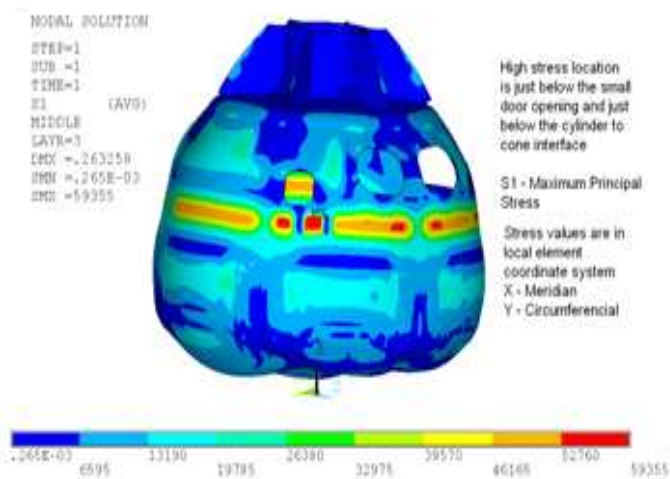


Figure 4: High Stress Plot for Maximum Principal Stress in Ply 3

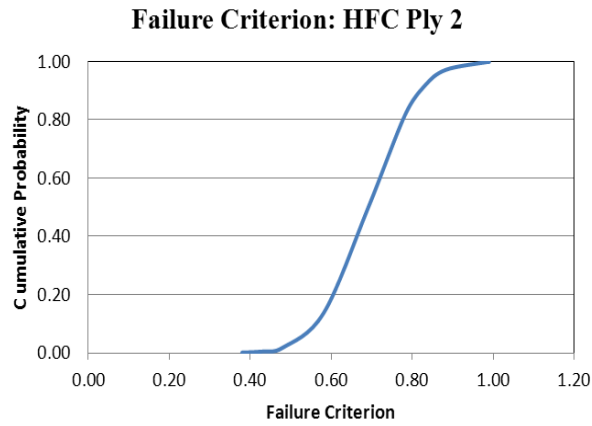


Figure 5: Cumulative Density Function (CDF) of Failure Criterion, HCF Ply 2

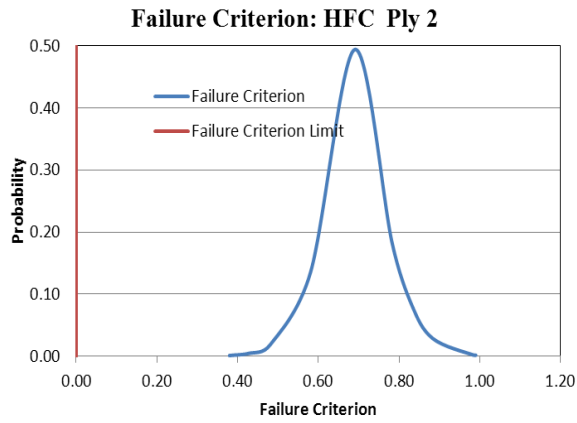


Figure 6: HCF Failure Criterion with Failure Criterion Limit for Ply 2 (5% uncertainties in the input variables)

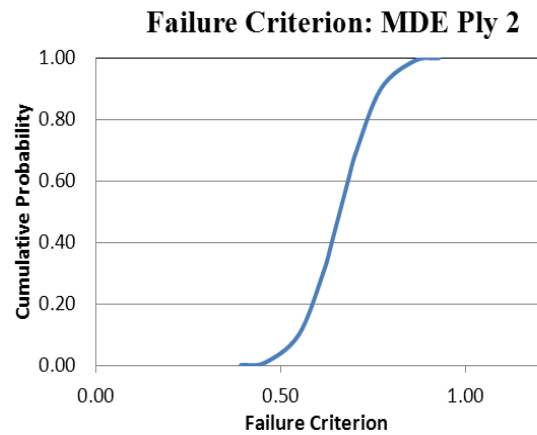


Figure 7: CDF of Failure Criterion, MDE Ply 2

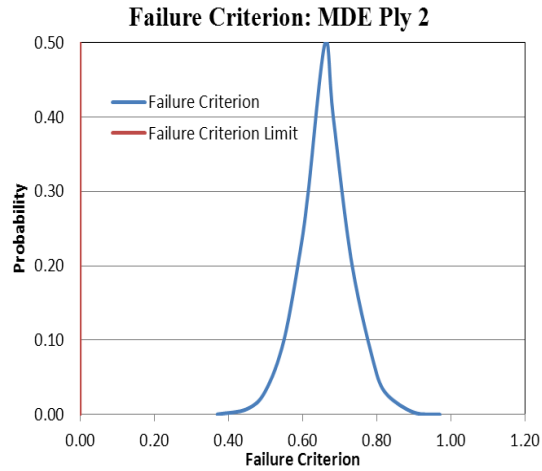


Figure 8: MDE Failure Criterion with Failure Criterion Limit for Ply 2 (5% uncertainties in the input variables)

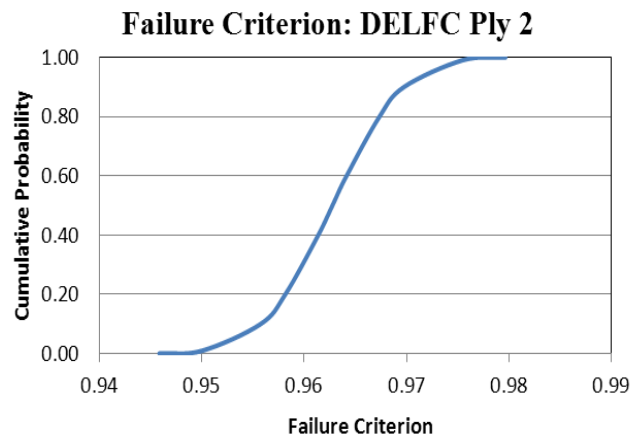


Figure 9: CDF of Failure Criterion: DELFC Ply 2

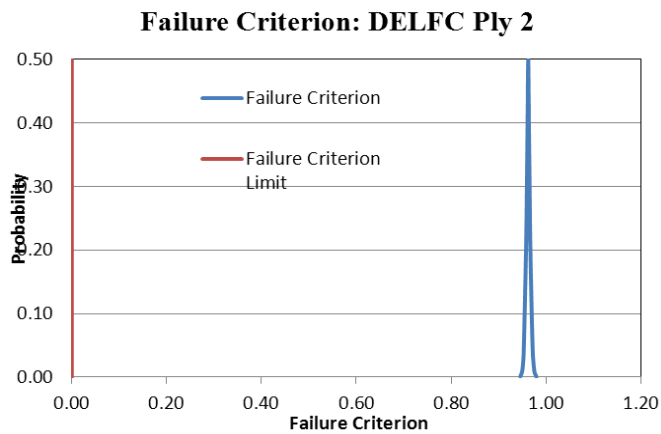


Figure 10: DELFC Failure Criterion with Failure Criterion Limit for Ply 2 (5% uncertainties in the input variables)

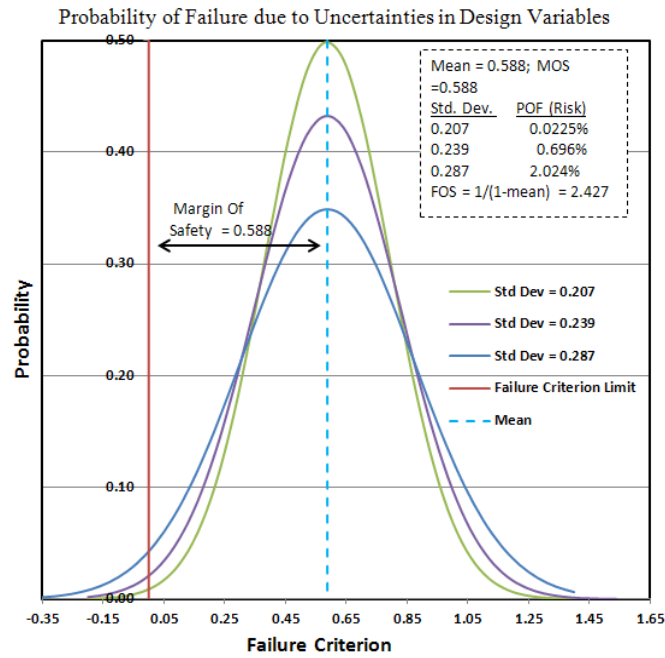


Figure 11: Probability of Failure due to Uncertainties in Design Variables. Factor of safety alone is not enough to properly assess risk.

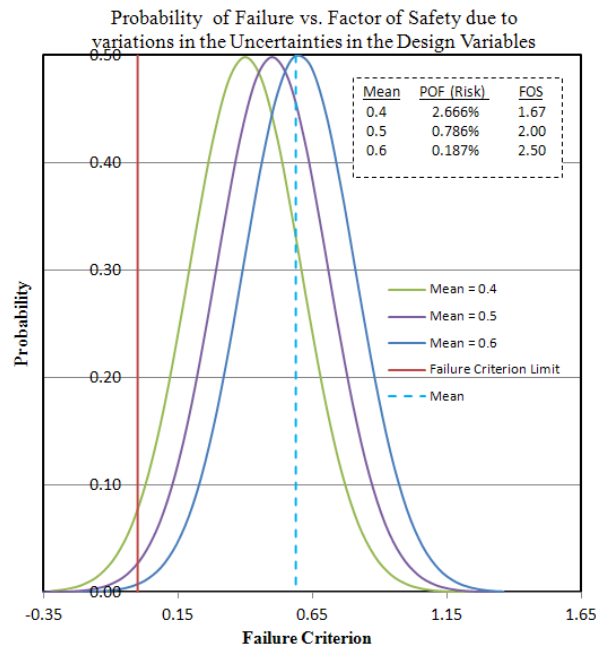


Figure 12: Probability of Failure (POF) vs. Factor of Safety (FOS) due to variation in the Uncertainties in the Design Variables

### ACKNOWLEDGMENTS

Funding for this work was provided by NASA Glenn Research Center.

### REFERENCES

Robert A. Aiello, "Composite Blade Structural Analyzer (COBSTRAN) User's Manual". NASA Technical Memorandum 101461. April 1989.

DEPARTMENT OF DEFENSE HANDBOOK. COMPOSITE MATERIALS HANDBOOK VOLUME 2. POLYMER MATRIX COMPOSITES MATERIALS PROPERTIES. MIL-HDBK-17-2F

Charles Harris, David Bowles, and Mark Shuart, "The NASA–Virginia Tech Composites Program Overview of the History and State-of-the-Art of Large Composite Structures in Aerospace Vehicles". Technical Presentation. 2008

Michael Kirsch, "Composite Crew Module Materials and Processes". NASA Engineering and Safety Center Technical Assessment Report. NESC-RP-06-019. 2011

Robert E. Melchers, "Structural Reliability Analysis and Prediction". John Wiley & Sons, 2nd edition, 1999.

I. Miller, "Probabilistic Finite Element Modeling of Aerospace Engine Components Incorporating Time-Dependent Inelastic Properties For Ceramic Matrix Composite (CMC) Materials". Master's thesis, University of Akron. 2006

Pai, Dr. Shantaram S. Hoge, Peter, Nagpal, Dr. Vinod K. "PROBABILISTIC ESTIMATION OF CRITICAL FLAW SIZES IN THE PRIMARY STRUCTURE WELDS OF THE ARES I-X LAUNCH VEHICLE" ASME IGTI CONFERENCE, BERLIN GERMANY , 2008, GT2008-50626

Pappu L.N. Murthy, Christos C. Chamis, and Subodh K. Mital. Computational Simulation of Continuous Fiber-Reinforced Ceramic Matrix Composites Behavior. NASA Technical Paper 3602 July 1996

Nagpal, Dr. Vinod K. and Strack, William, Development of System Uncertainty Analysis Tool PRODAF for Cost Effective and Quantified Risk Based Designs, *November 19-21, 2010 - ICCM 2010, Zhangjiajie, China*

N&R Engineering "PROBABILISTIC STRUCTURAL ANALYSIS OF THE COMPOSITE CREW MODULE -SUBSTRUCTURING WITH HIGH RESOLUTION GRID", Final Report, December 2011

NASA PREFERRED RELIABILITY PRACTICES. "STRUCTURAL LAMINATE COMPOSITES FOR SPACE APPLICATIONS. PRACTICE" NO. PD-ED-1217. Advanced Composite Structures. NASA SP-8108. 1974

Andrzej S. Nowak and Kevin R. Collins, "Reliability of Structures". McGraw-Hill, 2000.

Michael Stamatelatos et al, "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners" (2002). Chapter 14.

J. Townsend, C. Meyers, R. Ortega, J. Peck, M. Rheinfurth, and B. Weinstock, "Review of the Probabilistic Failure Analysis Methodology and Other Probabilistic Approaches for Application in Aerospace Structural Design". NASA technical paper 3434. 1993



## ANNEX A

### Failure Criteria equations and nomenclature

Delamination criterion:

$$FC = 1 - \left( \frac{|\Delta\phi|}{\Delta\phi_{del}} \right)$$

$$\Delta\phi_{del} = \frac{S_{12}}{G_{12}}$$

$$\Delta\phi = \frac{1}{2}(\varepsilon_{c22} - \varepsilon_{c11})(\sin 2\theta_i - \sin 2\theta_{i-1}) + \frac{1}{2}(\varepsilon_{c12})(\cos 2\theta_i - \cos 2\theta_{i-1})$$

Hoffman's criterion:

$$FC = 1 - \left[ \left( \frac{\sigma_{P11\alpha}}{S_{P11\alpha}} \right)^2 + \left( \frac{\sigma_{P11\beta}}{S_{P11\beta}} \right)^2 - K_{P12\alpha\beta} \frac{\sigma_{P11\alpha}}{|S_{P11\alpha}|} \frac{\sigma_{P22\beta}}{|S_{P22\beta}|} + \left( \frac{\sigma_{P12}}{S_{P12}} \right)^2 \right]$$

Modified distortion energy criterion:

$$FC = 1 - \left[ \left( \frac{\sigma_{P11\alpha}}{S_{P11\alpha}} \right)^2 + \left( \frac{\sigma_{P11\beta}}{S_{P11\beta}} \right)^2 - K_{P12\alpha\beta} \frac{\sigma_{P11\alpha}}{|S_{P11\alpha}|} \frac{\sigma_{P22\beta}}{|S_{P22\beta}|} + \left( \frac{\sigma_{P12}}{S_{P12}} \right)^2 \right]$$

$$K_{P12\alpha\beta} = K' \frac{(1 + 4\nu_{P12} - \nu_{P13})E_{P22} + (1 - \nu_{P23})E_{P11}}{\sqrt{E_{P11}E_{P22}(2 + \nu_{P12} + \nu_{P13})(2 + \nu_{P21} + \nu_{P23})}}$$

where  $K'$ , the correlation factor, is assumed to be 1.

$$\alpha = \begin{cases} T, & \sigma_{P11} \geq 0 \\ C, & \sigma_{P11} < 0 \end{cases}$$

$$\beta = \begin{cases} T, & \sigma_{P22} \geq 0 \\ C, & \sigma_{P22} < 0 \end{cases}$$

$$S_{P11\alpha} = \begin{cases} S_{P11T}, & \alpha = T \\ S_{P11C}, & \alpha = C \end{cases}$$

Symbol	Description
$\varepsilon_{11}$	Strain in the 11 direction
$\varepsilon_{22}$	Strain in 22 direction
$\varepsilon_{12}$	Shear strain in plane 12
$\theta_i$	Fiber angle in the $i$ th ply
$\theta_{i-1}$	Fiber angle in the $i - 1$ th ply
$S_{12}$	Shear strength in plane 12
$G_{12}$	Shear strain in plane 12
$\sigma_{P11}$	Calculated stress in the 11 direction of ply p
$\sigma_{P22}$	Calculated stress in the 22 direction of ply p
$\sigma_{P12}$	Calculated stress in the 12 plane of ply p
$S_{P11T}$	Tensile strength in the 11 direction of ply p
$S_{P11C}$	Compressive strength in the 11 direction of ply p
$S_{P22T}$	Tensile strength in the 22 direction of ply p
$S_{P22C}$	Compressive strength in the 22 direction of ply p
$S_{P12}$	Strength in the 12 plane of ply p
$\nu_{12}$	Poisson's ratio
$\nu_{23}$	Poisson's ratio
$T$	Tension
$C$	Compression

## **Development of a simplified computational model to study cranial backspatter using SPH**

†\*E.E. Kwon<sup>1</sup>, A. Malhotra<sup>1</sup>, R. Das<sup>1</sup>, J.W. Fernandez<sup>2,3</sup> and M.C. Taylor<sup>4</sup>

<sup>1</sup>Department of Mechanical Engineering, University of Auckland, New Zealand

<sup>2</sup>Auckland Bioengineering Institute (ABI), University of Auckland, New Zealand

<sup>3</sup>Department of Engineering Science, The University of Auckland, Auckland, New Zealand

<sup>4</sup>The Institute of Environmental Science and Research (ESR), Christchurch Science Centre, New Zealand

\*Presenting author: [eryn.kwon@auckland.ac.nz](mailto:eryn.kwon@auckland.ac.nz)

†Corresponding author: [eryn.kwon@auckland.ac.nz](mailto:eryn.kwon@auckland.ac.nz)

### **Abstract**

Investigation of crime requires rigorous testing and sound scientific understanding of the evidence to assist with the reconstruction of the criminal event. From the multitude of forensic specializations, bloodstain pattern analysis evidence from cranial gunshot wounding is of particular interest because of the high mortality rate resulting from head wounding compared to other body parts. Traditionally, animal models and physical models of the human anatomy have been used to study the mechanism and extent of ballistic spatter, including backspatter. Backspatter is a retrograde spattering of the target material from the entry wound. The reverse directionality of backspatter has specific evidential value, as it may establish a link between the victim and the shooter via transfer of biological matter. Backspatter evidence has also been used in courts to distinguish between a homicide and a suicide. Despite the importance of backspatter, the understanding of its mechanism has remained inadequate due to ethical issues, difference in anatomical geometry associated with various animal samples, or material property difference among biological and synthetic materials used in physical models. Hence there is a need to develop simulation tools that will use computational models of cranium geometry and configurations relatively similar to those of humans. Such computational models can act as alternatives to animal or physical models for the investigation of backspatter in a variety of situations.

In this study, a mesh-free method called Smoothed Particle Hydrodynamics (SPH) is used to develop a computational model to simulate high speed ballistic impacts. The complex geometry of the human cranium was reduced to a simplified box model equivalent to average female anatomical internal volumes. The inhomogeneous and anisotropic behaviours of the biological materials in a cranium (skin, skull and brain) were simplified to homogeneous and isotropic materials for each component. A physical equivalent model was manufactured and tested under the same ballistic conditions, for computational model validation. The computational model matched well with its physical equivalent experimentation in both material deformation characteristics and the timing of key events. This demonstrated the potential of the simulation models as a better alternative to animal and physical models. The simulation captured the temporary cavity development in the brain simulatant well, as well as showing realistic fragmentation, including backspatter. The temporary deformation of the skin entry wound was also a good match to the physical experimentation. The simulation also helped identify the most suitable material models to simulate ballistic impact of the brain simulatant. This work provides the basis for a more complex, anatomically accurate geometric cranium model to further develop reliable and robust simulation of cranial ballistic impact.

**Keywords:** ballistic impact, backspatter, forensic biomechanics, cranial injury, computational modelling, smoothed particle hydrodynamics, hypervelocity impact.

## Introduction

One of the most important aspects of crime investigation is to determine the causal event from the evidence found. With the growing number of incidents involving firearms every year, the importance of establishing links between ballistic evidence and its causal parameters has been increasing. This research focuses on retrograde spatter ejected from the entry wound called 'backspatter'. Due to its directionality, backspatter may provide an evidential link between the victim and the assailant (Karger 2008). It can also help distinguish the difference between a homicide from a suicide (Yen, Thali et al. 2003). Backspatter evidence has also been used in court to determine firing distance and positions of the persons involved (Kleiber, Stiller et al. 2001). The most valuable backspatter evidence is a cranial backspatter. This is because 1) the backspatter phenomenon is most pronounced when a near-water density organ such as brain is impacted (Karger 2001); and 2) the human cranium is the most fatal site on the human body for ballistic impact. While the human head represents only 9% of the body, 50% of combat death has been attributed to cranial impact (Michael E. Carey 1989).

There are four possible ways to form a more detailed understanding of backspatter: 1) human research; 2) animal models; 3) physical (man-made) models; or 4) computational models. For ethical reasons, direct backspatter experimentation on human subjects is not an option. Therefore, the use of reliable models is crucial for backspatter research. Such models must be accurate and reliable if valid inferences are to be drawn. Animal models used to date, such as swine or bovine samples (Burnett 1991, Karger, Nusse et al. 1996, Karger, Nusse et al. 1997, Radford 2009), have been limited by the differences in anatomical proportions, with the animal samples featuring a smaller cranial cavity and thicker bone compared to that of a human. As with human testing, animal models also carry significant ethical concerns. To counter the problem, physical models have been developed using synthetic materials and simplified or anatomically realistic geometries (Stephens and Allen 1983, Pex and Vaughan 1987, Radford 2009, Foote 2012, Kwon 2014). The physical models successfully eliminated the ethical issues and more recent models have resolved the anatomical geometry problem (Foote 2012, Carr, Lindstrom et al. 2014, Kwon 2014). However, physical models also have major shortcomings. The cost of the model manufacturing and experimentation is still high and there is limited control over experimental variables and a lack of structural complexity as compared to biological materials. These limitations have promoted research into computational models to simulate cranial ballistic impact and the associated backspatter generation. The ease of model modification, experimentation, and analysis, combined with the low cost, makes the computational model a worthy, if not critical, research topic.

There are two computational methods that have been used to study ballistic impacts: 1) Finite Element Method (FEM) and 2) Smoothed-Particle Hydrodynamics (SPH). FEM is a widely recognised computational method to simulate ballistic impact. The FEM method has been utilised to successfully simulate impacts on ballistic helmets (Aare and Kleiven 2007, Yang and Dai 2010), ballistic gelatine (Datoc 2010), human mandible (Tang, Tu et al. 2012), human brain (Raul, Deck et al. 2007) and human frontal bone (Pinnoji and Mahajan 2007). However, due to its Eulerian approach, FEM cannot handle fragmentation, large deformation and high speed impact very well. This makes FEM unsuitable for backspatter research where the focus is on the fragmented particles produced from a high speed impact. Therefore, this research focussed instead on SPH, using it to obtain approximate numerical solutions of the ballistic impact simulation on human cranial model. The SPH is a Lagrangian computational method, using equations of fluid dynamics, by replacing

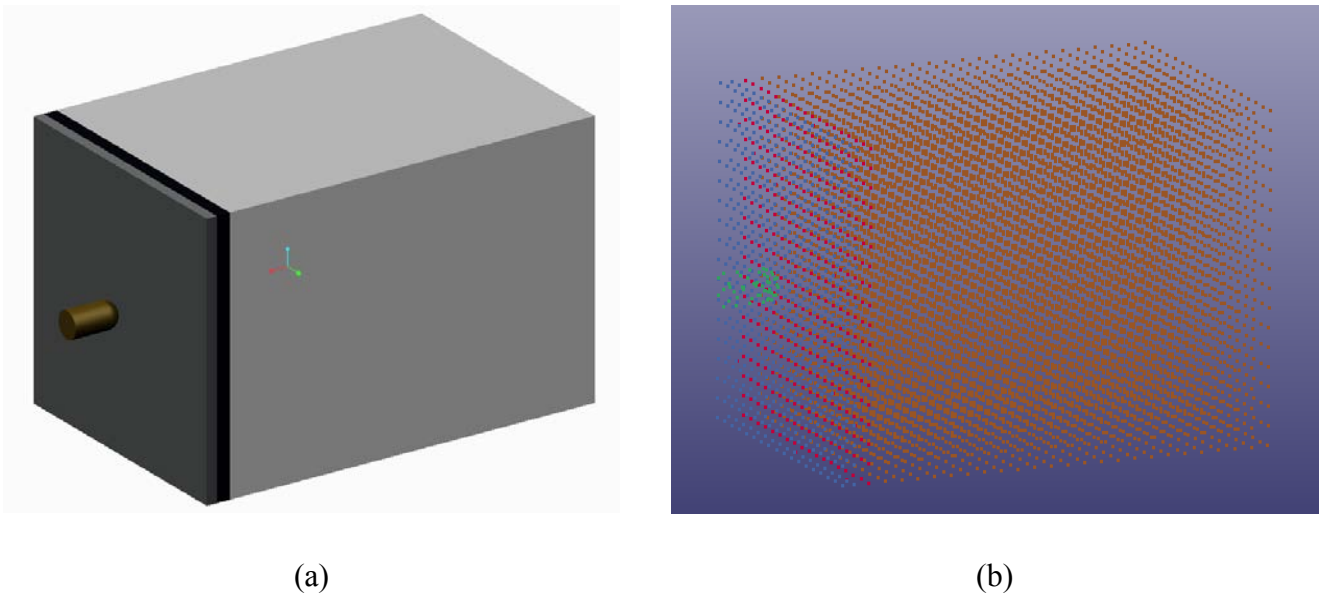
the body of fluid with a set of particles. SPH was originally invented to deal with problems in astrophysics; involving fluid masses moving arbitrarily in three dimensions in the absence of boundaries (Monaghan 2012). The mesh-free SPH technique simulates the gunshot wound appropriately and can handle large deformation and fragmentation (Monaghan 2012). SPH has been successfully used to model ballistic impact onto a variety of targets (Das, Collins et al. 2015).

In this research, the complex geometry of the human cranium was reduced to a simplified box model equivalent to average female anatomical internal volumes. The inhomogeneous and anisotropic behaviours of the biological materials in a cranium (skin, skull and brain) were simplified to homogeneous and isotropic materials for each component. As a validation method for the SPH computational model, a physical equivalent model of identical geometry and simulant materials has been manufactured and tested under the same ballistic conditions.

## Methods

### *Computational Model Development*

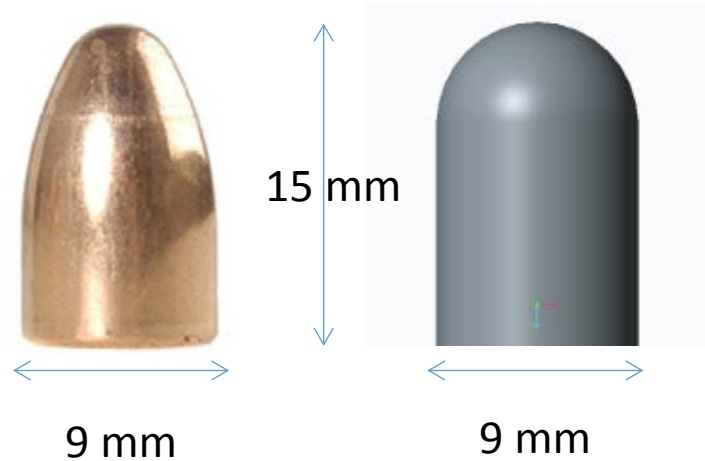
The scalp and the skull layers were represented by a 100 x 100 mm square. The scalp layer thickness was set to 3 mm, while the skull layer thickness was 5 mm. The thickness values were determined based on measurements of 114 MRI scan images acquired from the Centre of Advanced MRI, University of Auckland. The length of the brain layer was fixed to 140 mm so the internal volume of the brain layer was close to 1130 cm<sup>3</sup>, the average brain volume of a human female (Cosgrove, Mazure et al. 2007). The geometry of the model is shown in Figure 1 (a), while Figure 1 (b) shows the SPH model of the same geometry.



**Figure 1. (a) Geometry of the simplified computational model, complete with the bullet (brown), scalp (dark grey), skull (black) and the brain (light grey) layers (b) the SPH model of the same geometry**

Gravity and air resistance were assumed to be negligible for the purposes of the simulation. The model had a pinned boundary condition (no translation in x, y and z directions, while rotation is allowed) applied to the outside surfaces except the entry and exit side.

To match the experimental ballistic conditions, computational representation of a 9 mm Luger projectile was used to impact the computational model (Figure 2). The bullet was set to impact the model perpendicular to the scalp surface centre as shown in Figure 1 (a), with an initial speed of  $\sim 350\text{m/s}$ , which is the average bullet speed (MidwayUSA).



**Figure 2. Physical (left) and computational (right) bullet geometry comparison of the projectile part of the 9 mm Luger bullet**

The computational model used a hyperelastic failure model for the scalp and an isotropic-elastic model for the skull, based on a previous research (Kwon 2014, Das, Collins et al. 2015). The brain was modelled using either a viscoelastic model or the Plastic-Kinematic model to identify the most suitable material constitutive model to describe the ballistic deformation of a human brain.

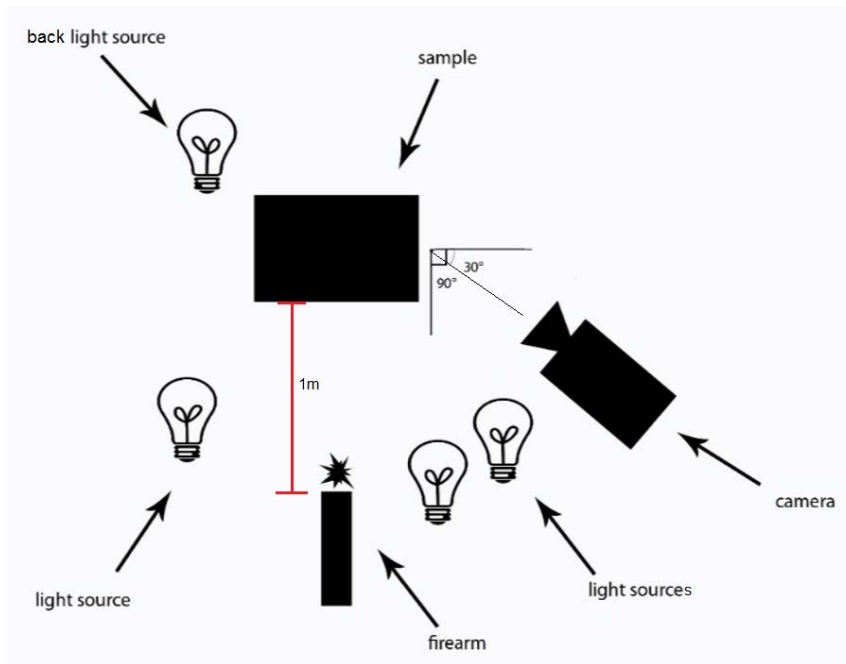
The material properties of the scalp and skull simulants were measured in the University of Auckland Centre of Advanced Composite Materials laboratories. The scalp simulant was a Room Temperature Vulcanising (RTV) Silicone, a polyurethane resin was used as the skull simulant, and the brain was modelled using 10% gelatine. The material properties of the 10% gelatine were obtained from the literature (Kelager 2006, Cohen, Cleary et al. 2009).

#### *Physical Equivalent Model*

Manufacturing of the physical model was carried out using a casting method. The skin layer was manufactured from curing RTV silicone with a cotton-cellulose sponge insert as the blood reservoir. The bone layer was made from a homogeneous cast of polyurethane. The brain layer was cast using a 10% gelatine solution. The quality of the gelatine was crucial in bullet trajectory visualization. To achieve the desired transparency, a few drops of cinnamon oil were added to the gelatine-water mixture.

The ballistic experimental setup for the physical model was carried out at the Firearms Testing Laboratory in the Mount Albert Science Centre of Environmental Science and Research. A sample was first secured on a holder. The target surface was set perpendicular to the bullet trajectory. A white surface was placed horizontally in front of the sample to capture any backspatter. A high speed camera, Photron Fastcam SA1, was positioned at an angle of approximately  $30^\circ$  to the surface. This oblique angle was used to capture both the scalp and the brain layer surface deformation at the same time. The frame rate was set to 30,000 frames per second to capture the details of the dynamic deformations and backspatter generation. The experimental setup schematic is shown in Figure 3.

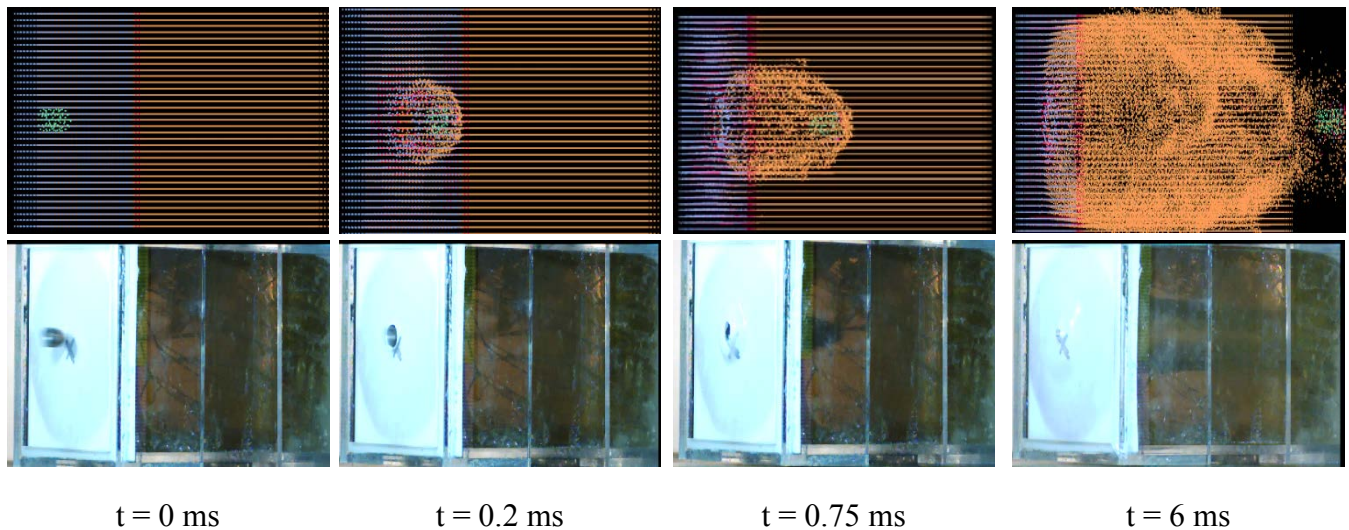
The samples were shot using a federal brand American Eagle 9 mm Luger calibre bullet fired from a Glock, Model 17 semi-automatic pistol. The bullet used was a 115 grain Full Metal Jacket (FMJ) projectile, which has a solid lead core with a copper material coating. The bullet was fired with a muzzle to target distance of approximately 1 m to minimize the effect of the muzzle gas on the ballistic response of the sample (Taylor, Laber et al. 2010). After each shot, the samples were photographed and archived with the collected backspatter for subsequent analysis.



**Figure 3. Experimental setup schematic**

## Results and Discussion

The high speed footage of the experimental results was visually compared to the simulation results. An image of the simulation result was extracted for each key stage of the ballistic deformation. The images were set to be of the same view angle and magnification, and had 50% transparency. The conditioned images were then overlayed onto the high speed footage for comparison. An excerpt of the key deformation stages for both physical and computational model is shown in Figure 4.



**Figure 4. Chronological comparison between the physical and computational model results**

The sequence of events observed from the physical model was highly comparable to that of simulation result, in both the spatial dimensions and the timing. For both the physical and the computational models: 1) at 3 ms it was observed that the maximum surface displacement and blow out diameter had been reached; 2) at 3.75 ms, the edges of the blow out are observed to be retracting, but the uniform cylindrical shape still remains; 3) at 6 ms the cavity of the blow out began to collapse while the brain temporary cavity is still expanding; 4) after the initial subcutaneous temporary cavitation, the scalp layer recovered its original position without any subsequent oscillatory motion.

Throughout the entire ballistic event, the temporary cavity in the computational model brain layer developed in similar size and speed to the physical equivalent as illustrated in Figure 4. The general tapered shape of the cavity is also replicated throughout the simulation result. Also, backscatter generation from both models was observed.

Of the two material constitutive models, the viscoelastic failure model was unsuccessful at simulating a realistic result. Under this simulation, the bullet stopped in the midst of the brain layer, losing its kinetic energy completely to the brain layer. This was different from the physical model experimental result. On the other hand, the Plastic-Kinematic model produced a simulation result that successfully goes through the computational model, mimicking the experimental result well.

## Conclusions

The computational model matched well with its physical equivalent experimentation in both material deformation characteristics and the timing of key events, demonstrating the potential of the simulation models as a better alternative to animal and physical models. The simulation successfully captured the temporary cavity development in the brain simulant, as well as showing realistic fragmentation, including backscatter. The temporary deformation of the skin entry wound was also a good match to the physical experimentation. The simulation also identified the most suitable material constitutive model to simulate ballistic impact of the brain simulant to be the Plastic-Kinematic mode.

As the first skin-skull-brain computational model of human cranium for ballistic backspatter research, this work provides the basis for a more complex anatomically accurate geometric cranium model to further develop reliable and robust simulation of cranial ballistic impact.

## References

- Aare, M. and S. Kleiven (2007). "Evaluation of head response to ballistic helmet impacts using finite element method." International Journal of Impact Engineering **34**: 596-608.
- Burnett, B. R. (1991). "Detection of Bone and Bone-Plus-Bullet Particles in Backspatter from Close-Range Shots to Heads." Journal of Forensic Science **36**(6): 1745-1752.
- Carr, D., A.-C. Lindstrom, A. Jareborg, S. Champion, N. Waddell, D. Miller, M. Teagle, I. Horsfall and J. Kieser (2014). "Development of a skull/brain model for military wound ballistics studies." International Journal of Legal Medicine: 1-6.
- Cohen, R. C., P. W. Cleary and B. Mason (2009). Simulations of human swimming using smoothed particle hydrodynamics. 7th International Conference on CFD in the Minerals and Process Industries, Commonwealth Scientific and Industrial Research Organisation.
- Cosgrove, K. P., C. M. Mazure and J. K. Staley (2007). "Evolving knowledge of sex differences in brain structure, function, and chemistry." Biological psychiatry **62**(8): 847-855.
- Das, R., A. Collins, A. Verma, J. Fernandez and M. Taylor (2015). "Evaluating Simulant Materials for Understanding Cranial Backspatter from a Ballistic Projectile." Journal of forensic sciences **60**(3): 627-637.
- Datoc, D. (2010). Finite Element Analysis and Modelling of a .38 Lead Round Nose Polytechnic State University. Master of Science in Biomedical Engineering, California Polytechnic State University.
- Foote, N. R. (2012). The Role of the Temporary Cavity in Cranial Backspatter. Master of Science in Forensic Science, The University of Auckland.
- Karger, B. (2001). "Forensic ballistic of gunshot wounds." Forensische ballistik von schussverletzungen **11**(3-4): 104-119.
- Karger, B. (2008). Forensic Ballistics. Forensic Pathology Reviews. M. Tsokos. Totowa, NJ, Humana Press. **5**: 139-172.
- Karger, B., R. Nusse, G. Schroeder, S. Wustenbecker and B. Brinkmann (1996). "Backspatter from Experimental Close-Range Shots to the Head I-Macrobackspatter." International Journal of Legal Medicine **109**: 66-74.
- Karger, B., R. Nusse, H. D. Troger and B. Brinkmann (1997). "Backspatter from Experimental Close-Range Shots to the Head II-Microbackspatter and the Morphology of Bloodstains." International Journal of Legal Medicine **110**: 27-30.
- Kelager, M. (2006). "Lagrangian fluid dynamics using smoothed particle hydrodynamics." University of Copenhagen, Denmark.
- Kleiber, M., D. Stiller and P. Wiegand (2001). "Assessment of shooting distance on the basis of bloodstain analysis and histological examinations." Forensic Science International **119**: 260-262.
- Kwon, E. E. (2014). Development of Physical and Numerical Models to Study Cranial Backspatter. Master of Engineering, University of Auckland.
- Michael E. Carey, M. D., Gurcharan S. Sama, Ph.D., J. Bryan Farrell, B.S., and Leo T. Happel, Ph.D. (1989). "Experimental missile wound to the brain." Journal of Neurosurgery **71**(5): 754-764.
- MidwayUSA. "Federal American Eagle Ammunition 9mm Luger 115 Grain Full Metal Jacket Box of 50." from <http://www.midwayusa.com/product/159241/federal-american-eagle-ammunition-9mm-luger-115-grain-full-metal-jacket-box-of-50>.
- Monaghan, J. J. (2012). "Smoothed Particle Hydrodynamics and Its Diverse Applications." Annual Review of Fluid Mechanics **44**(1): 323-346.
- Pex, J. O. and C. H. Vaughan (1987). "Observations of High Velocity Bloodspatter on Adjacent Objects." Journal of Forensic Sciences **32**(6): 1587-1594.
- Pinnoji, P. K. and P. Mahajan (2007). "Finite element modelling of helmeted head impact under frontal loading." Sadhana **32**(4): 445-458.
- Radford, G. E. (2009). Modelling Cranial Gunshot Wounds and Backspatter. Master of Science, University of Otago.
- Raul, J., C. Deck, F. Meyer, A. Geraut, R. Willinger and B. Ludes (2007). "A finite element model investigation of gunshot injury." International Journal of Legal Medicine **121**: 143-146.
- Stephens, B. G. and T. B. Allen (1983). "Back Spatter of Blood from Gunshot Wounds - Observations and Experimental Simulation." Journal of Forensic Sciences **28**(2): 437-439.
- Tang, Z., W. Tu, G. Zhang, Y. Chen, T. Lei and Y. Tan (2012). "Dynamic simulation and preliminary finite element analysis of gunshot wounds to the human mandible." Injury, Int J. Care Injured **43**: 660-665.
- Taylor, M. C., T. L. Laber, B. P. Epstein, D. S. Zamzow and D. P. Baldwin (2010). "The effect of firearm muzzle gases on the backspatter of blood." International Journal of Legal Medicine: 1-12.
- Yang, J. and J. Dai (2010). "Simulation-Based Assessment of Rear Effect to Ballistic Helmet Impact." Computer-Aided Design & Applications **7**(1): 59-73.
- Yen, K., M. J. Thali, B. P. Kneubuehl, O. Peschel, U. Zollinger and R. Dirnhofer (2003). "Blood-Spatter Patterns: Hands Hold Clues for the Forensic Reconstruction of the Sequence of Events." American Journal of Forensic Medicine and Pathology **24**(2): 132-140.



## Development of a fully coupled flow-geomechanics simulator for flow in saturated porous media

†\*Chao Zhang, Sadiq J. Zarrouk and Rosalind Archer

Department of Engineering Science  
University of Auckland, Private Bag 92019, Auckland, New Zealand

\*Presenting author: chao.zhang@auckland.ac.nz  
†Corresponding author: chao.zhang@auckland.ac.nz

### Abstract

In this paper, we propose a fully coupled flow-geomechanics simulator using the mixed finite element method. The mathematical model, including mass conservation of fluid, Darcy's law for velocity, and force equilibrium of solid skeleton, is derived in the framework of Biot's consolidation theory. Pore pressure, fluid velocity and solid displacement are chosen as primary variables. This has the advantage of satisfying element-wise mass conservation and describing the velocity in a continuous way, instead of as a derived value of pressure, as in traditional simulators. The mathematical model is then discretized in appropriate finite element spaces. Specifically, we use the constant discontinuous Galerkin space for pressure, the lowest order Brezzi-Douglas-Marini space for fluid velocity, and the linear Continuous Galerkin space for solid displacement. The system of equations is solved in a fully coupled manner, which ensures second order convergence and stability. Afterwards, the resulted model is validated using a wide range of benchmark problems, including the consolidation problems of Terzaghi, Mandel and Cryer. In all cases, our numerical results are in good agreement with the analytical solutions, which illustrates the effectiveness of our simulator, especially in capturing the Mandel-Cryer effect accurately.

**Keywords:** Mixed finite element, Poro-elasticity, Consolidation, Fully-coupled model, Porous media

### Introduction

In many applications involving porous media, it is of paramount importance to model the interaction between fluid flow and solid deformation in a tightly coupled manner to make accurate predications [Gambolati et al., 1991; Pao et al., 2001; Teatini et al., 2006; Yin et al., 2009]. This paper will focus on developing an alternative simulator to accurately describe this coupled process.

From the mathematics point of view, the mutual coupling between fluid flow and solid deformation leads to a complex initial-boundary-value problem. Over the past decades, a lot of researchers have studied this problem through different numerical methods, namely the finite difference method [Lee et al., 1998; Masson and Pride, 2007; Yanosik et al., 1979], the finite element method [Edwards, 2002; Lewis and Schrefler, 1987; Panneton and Atalla, 1997; Wheeler et al., 2014], and the finite volume method [Jenny et al., 2005; Mosharaf Dehkordi et al., 2014; Rozon et al., 1989]. Compared with the finite difference method and the finite volume method, not only could the finite element method handle well complex geometries, which is common in reservoir simulations, it is also good at multiple field problems [Fortin and Brezzi, 1991; Hesthaven and War-

burton, 2007].

In the literature, there are three coupling approaches (explicitly coupled, iteratively coupled and fully coupled) to study the fluid-solid interaction problem. In the explicitly coupled approach [Longuemare et al., 2002; Minkoff et al., 2003], the flow variables are assumed constant when solving the force equilibrium equation, and vice versa. This method requires relatively small time steps to ensure a physical solution. In the iteratively coupled scheme [Mikelić and Wheeler, 2013; Tran et al., 2004], each simulator solves its own governing equation and uses some correction terms to make sure that the equations are solved at the same time step. Since the coupled problem is split into a flow problem and a solid deformation problem, it results in two much smaller problems and saves computational time. Although an iterative method has the advantage of efficiency, the disadvantage is that it may introduce splitting errors. Especially in gas flow cases, splitting error may lead to unphysical solutions [Aarnes et al., 2007]. A third method is to solve the coupled flow and solid system simultaneously, which is referred to as the fully coupled approach [Khoshghalb and Khalili, 2010; Settari et al., 1998; White and Borja, 2011]. In this approach, each equation is discretized implicitly, which guarantees its robustness. However, the disadvantage is also obvious: it results in a larger system of equations and may require more computational time.

In this work, we present a fully coupled solver to account for the interaction between fluid flow and geomechanics by the finite element method. The mathematical model is proposed based on Biot's theory in poromechanics, and a subsequent numerical model is implemented by the mixed finite element techniques. Specifically, we use the lowest order Brezzi-Douglas-Marini function space (BDM1) and the constant Discontinuous Galerkin function space (DG0) for the fluid velocity and pressure field, respectively. This pair of function spaces ensures local mass conservation, which is relatively important in flow related problems. As to the solid displacement field, the standard linear Continuous Galerkin interpolant is used. The related system is then solved in a fully coupled manner, which ensures second order convergence.

In the following sections, we will show all the aspects in developing the flow-geomechanics simulator. Specifically, in Section 2 we briefly outline the mathematical equations used in this coupled problem. Section 3 discusses the weak form derivation, space and time discretization and finite element implementation. Section 4 examines the correctness of the developed solver by comparing the numerical results with some classic benchmark problems. Finally in Section 5, we draw conclusions.

## Mathematical Model

### *Governing equations*

As in classic continuum mechanics, the whole system is viewed as a number of overlapping continua representing the corresponding phase, namely solid skeleton and fluid. The representative elementary volume (REV) is large enough to preserve the physical properties, like porosity and permeability. On the other hand, the REV must be sufficiently small to be considered as a point in the macroscopic scale.

In the framework of Biot [1941] poromechanics, we make the following assumptions: (1) the pore water is incompressible, (2) the solid grains are incompressible, (3) small

strain theory is applicable, (4) the soil as a whole exhibit linear elastic deformation, and (5) the system is isothermal.

The flow of fluid through the porous medium is described by the mass conservation equation as:

$$\frac{\partial}{\partial t}(\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{u} = f_f, \quad (1)$$

where  $\mathbf{u}_s$  is solid skeleton displacement,  $\mathbf{u}$  is fluid velocity relative to solid skeleton, and  $f_f$  is fluid source term.

The relative velocity of fluid  $\mathbf{u}$  is usually governed by the Darcy's law as:

$$\mathbf{u} + \frac{\mathbf{k}}{\mu} \nabla p = 0, \quad (2)$$

where  $p$  is fluid pressure, and  $\mathbf{k}$  is permeability tensor, and  $\mu$  is dynamic viscosity. In simplified cases, isotropy can be assumed in porous media, and the corresponding permeability tensor is a diagonal one. By nature, most porous media systems are directionally dependent, and thus a full tensor permeability is usually more appropriate to accurately describe the flow path.

As we assume the system is quasi-static, the equation of force equilibrium can be expressed in terms of the total stress as:

$$\nabla \boldsymbol{\sigma} + \mathbf{f}_s = 0, \quad (3)$$

where  $\boldsymbol{\sigma}$  is total stress, and  $\mathbf{f}_s$  is body force, i.e. gravity. The relation between total stress  $\boldsymbol{\sigma}$ , effective stress  $\boldsymbol{\sigma}_e$ , and fluid pressure  $p$  is given as:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_e - \alpha p \mathbf{I}, \quad (4)$$

where  $\mathbf{I}$  is the identity matrix, and  $\alpha$  is the Biot coefficient. The sign convention, namely that positive stress is taken as tension and negative compression, is applied here.

The constitutive equation relating the effective stress to strain  $\boldsymbol{\varepsilon}$  reads:

$$\boldsymbol{\sigma} = 2G\boldsymbol{\varepsilon} + \lambda\varepsilon_v \mathbf{I}, \quad (5)$$

where  $\varepsilon_v$  is the volumetric strain,  $G$  and  $\lambda$  are Lamé constants.

By applying the small strain theory, strain is related to the solid skeleton displacement as:

$$\boldsymbol{\varepsilon} = \frac{1}{2}[\nabla \mathbf{u}_s + (\nabla \mathbf{u}_s)^T]. \quad (6)$$

#### *Initial and boundary conditions*

In order to complete the coupled equation system, we need to apply suitable initial and boundary conditions for both the fluid and solid part.

Initially, we assume that fluid pressure, fluid velocity, and solid displacements are known values as:

$$\begin{cases} p = p^0 \\ \mathbf{u} = \mathbf{u}^0 \\ \mathbf{u}_s = \mathbf{u}_s^0 \end{cases} \quad \text{in } \Omega, \quad (7)$$

at  $t = 0$ .

When  $t > 0$ , we consider the following boundary condition for the solid problem:

$$\begin{cases} \mathbf{u}_s = \mathbf{u}_s^b, & \text{on } \Gamma_{u_s} \\ \boldsymbol{\sigma} \mathbf{n} = \mathbf{g}, & \text{on } \Gamma_\sigma \end{cases}, \quad (8)$$

where  $\mathbf{n}$  denotes unit outward normal vector of the boundary, and  $\mathbf{g}$  is boundary force. As to the fluid flow part, the following boundary conditions are applied:

$$\begin{cases} p = p^b, & \text{on } \Gamma_p \\ \mathbf{u} = \mathbf{u}^b, & \text{on } \Gamma_u \end{cases}. \quad (9)$$

$\Gamma_p$ ,  $\Gamma_u$ ,  $\Gamma_{u_s}$  and  $\Gamma_\sigma$  are corresponding pressure, velocity, displacement and exterior stress boundaries.

Before concluding this section, we summarize the three governing equations described above:

$$\frac{\partial}{\partial t}(\nabla \cdot \mathbf{u}_s) + \nabla \cdot \mathbf{u} = f_f, \quad (10)$$

$$\mathbf{u} + \frac{\mathbf{k}}{\mu} \nabla p = 0, \quad (11)$$

$$\nabla \boldsymbol{\sigma} + \mathbf{f}_s = 0. \quad (12)$$

## Numerical Implementation

### *Weak form derivation*

We consider this problem posed on the physical domain  $\Omega$  with boundary  $\partial\Omega$ . Let triangulation  $\mathcal{T}_h = \{K\}$  be a partition of the domain  $\Omega$ , and  $K$  is a subset of  $\Omega$ . In order to perform finite element discretization, we need to introduce suitable function spaces for the test and trial functions. We set

$$L^2(\Omega) = \{p : \int_{\Omega} |p|^2 dx < +\infty\}, \quad (13)$$

$$H^1(\Omega) = \{u : u \in L^2(\Omega), Du \in L^2(\Omega)\}, \quad (14)$$

$$H(\text{div}, \Omega) = \{\mathbf{v} : \mathbf{v} \in L^2(\Omega)^d, \nabla \cdot \mathbf{v} \in L^2(\Omega)\}. \quad (15)$$

Fluid pressure must belong to  $L^2(\Omega)$ , fluid velocity belongs to  $H^1(\Omega)^d$  and solid displacement belongs to  $H(\text{div}, \Omega)$ , respectively.  $d$  denotes the number of space dimensions.

Multiplying Equation (10), Equation (11), Equation (12) by test functions  $q$ ,  $\mathbf{v}$  and  $\mathbf{v}_s$ , respectively, and integrating by parts on  $\Omega$ , we get the following residual formulation: Find  $p \in L^2(\Omega)$ ,  $\mathbf{u} \in H^1(\Omega)$  and  $\mathbf{u}_s \in H(\text{div}, \Omega)$  such that we have

$$\mathcal{R}_p = \int_{\Omega} \frac{\partial}{\partial t} (\nabla \mathbf{u}_s) q \, dx + \int_{\Omega} \nabla \cdot \mathbf{u} q \, dx - \int_{\Omega} f_f q \, dx = 0, \quad (16)$$

$$\mathcal{R}_u = \int_{\Omega} \mathbf{u} \cdot \mathbf{v} \, dx - \int_{\Omega} \frac{\mathbf{k}}{\mu} p \nabla \cdot \mathbf{v} \, dx + \int_{\Gamma} \frac{\mathbf{k}}{\mu} p \mathbf{v} \cdot \mathbf{n} \, ds = 0, \quad (17)$$

$$\mathcal{R}_{u_s} = \int_{\Omega} \boldsymbol{\sigma} \nabla \cdot \mathbf{v}_s \, dx - \int_{\Gamma} \mathbf{n} \cdot \boldsymbol{\sigma} \mathbf{v}_s \, ds - \int_{\Omega} \mathbf{f}_s \mathbf{v}_s \, dx = 0, \quad (18)$$

for all  $q \in L^2(\Omega)$ ,  $\mathbf{v} \in H^1(\Omega)$  and  $\mathbf{v}_s \in H(\text{div}, \Omega)$ . The symbol  $\mathcal{R}$  refers to residual.

### Space and time discretization

Fluid pressure  $p$ , fluid velocity  $\mathbf{u}$  and solid displacement  $\mathbf{u}_s$  are chosen as primary variables, as indicated earlier in this paper. In order to make the problem well-posed, the function spaces and the polynomial degrees of shape functions for the pressure and the velocity can not be chosen arbitrarily [Fortin and Brezzi, 1991]. Instead, a pair of finite element spaces that satisfy the in-sup condition is required. In this study, we choose the elementwise constant space (DG0) for pressure, the lowest order Brezzi-Douglas-Marini space (BDM1) for fluid velocity, and the linear Continuous Galerkin space (CG1) for solid displacement.

Compared with the standard finite element method, the advantages of this mixed form are: (1) it satisfies element-wise mass conservation, (2) the stability and convergence are ensured, (3) it avoids the numerical diffusion in standard finite element method, and (4) the velocity is described in a continuous way, instead of as a derived value of pressure.

We interpolate pressure, velocity and displacement in the discretized space as:

$$p_h = \sum_{i=1}^{n_{elem}} \phi_i P_i \quad (19)$$

$$\mathbf{u}_h = \sum_{i=1}^{n_{face}} \phi_i \mathbf{U}_i \quad (20)$$

$$\mathbf{u}_{sh} = \sum_{i=1}^{n_{node}} \phi_i \mathbf{U}_{s_i} \quad (21)$$

where  $p_h$ ,  $\mathbf{u}_h$  and  $\mathbf{u}_{sh}$  are discrete variables,  $\phi$  is shape function, and  $P$ ,  $\mathbf{U}$  and  $\mathbf{U}_s$  are values at corresponding elements, faces and nodes.  $n_{elem}$ ,  $n_{face}$  and  $n_{node}$  are respective numbers of elements, faces and nodes.

As for time discretization, we apply the conventional backward Euler scheme, which ensures first order accuracy:

$$\frac{\partial}{\partial t}(\nabla \mathbf{u}_s) = \frac{\nabla \mathbf{u}_s - \nabla \mathbf{u}_{s0}}{\Delta t} \quad (22)$$

where  $\Delta t$  is time step.

### *Coupling scheme*

As declared in the introduction section, the fully coupled scheme has the advantage of ensuring stability and avoids convergence issues. Therefore, the fully coupled approach is employed in our model. Specifically, all the three equations are solved simultaneously at each time step, and the implicit Newton-Raphson method is used to update the solution. At time step  $(n + 1)$ , the solution at  $(k + 1)$ th iteration is updated by adding the correction terms to the solution at  $k$ th iteration as:

$$\begin{bmatrix} p^{n+1} \\ \mathbf{u}^{n+1} \\ \mathbf{u}_s^{n+1} \end{bmatrix}^{k+1} = \begin{bmatrix} p^{n+1} \\ \mathbf{u}^{n+1} \\ \mathbf{u}_s^{n+1} \end{bmatrix}^k + \begin{bmatrix} \delta p^{n+1} \\ \delta \mathbf{u}^{n+1} \\ \delta \mathbf{u}_s^{n+1} \end{bmatrix}, \quad (23)$$

where the correction comes from solving the following linear system:

$$\begin{bmatrix} \frac{\partial \mathcal{R}_p}{\partial p} & \frac{\partial \mathcal{R}_p}{\partial \mathbf{u}} & \frac{\partial \mathcal{R}_p}{\partial \mathbf{u}_s} \\ \frac{\partial \mathcal{R}_u}{\partial p} & \frac{\partial \mathcal{R}_u}{\partial \mathbf{u}} & \frac{\partial \mathcal{R}_u}{\partial \mathbf{u}_s} \\ \frac{\partial \mathcal{R}_{u_s}}{\partial p} & \frac{\partial \mathcal{R}_{u_s}}{\partial \mathbf{u}} & \frac{\partial \mathcal{R}_{u_s}}{\partial \mathbf{u}_s} \end{bmatrix} \begin{bmatrix} \delta p^{n+1} \\ \delta \mathbf{u}^{n+1} \\ \delta \mathbf{u}_s^{n+1} \end{bmatrix} = \begin{bmatrix} \mathcal{R}_p \\ \mathcal{R}_u \\ \mathcal{R}_{u_s} \end{bmatrix}, \quad (24)$$

and the symbol  $\delta$  denotes correction.

### *Implementation in FEniCS*

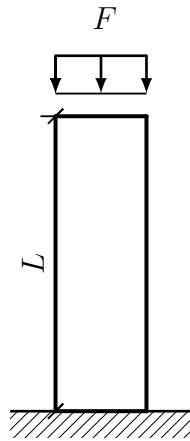
We create our code based on the recently developed open source project, FEniCS [Logg et al., 2012], which is a collection of libraries to facilitate the automated solution of partial differential equations using finite element method.

### **Numerical Tests**

In this section, we test the performance of the developed solver by solving a series of benchmark problems.

#### *Terzaghi's 1D problem*

The first example involves the consolidation problem from Terzaghi [1996], as illustrated in Figure 1. In this 1D problem, the top surface is assumed to be fully drained, while the bottom surface is impermeable. As for the solid part, the top surface is exposed to a constant vertical load and the bottom face is fixed. The length of the beam is taken as  $L = 15$  m, and other parameters are summarized in Table 1.



**Figure 1: Sketch of Terzaghi's 1D problem.**

**Table 1: Parameters in Terzaghi's problem**

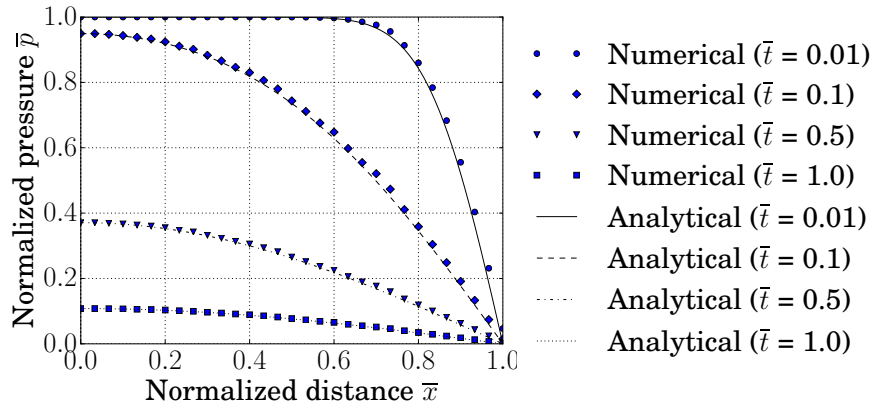
Parameter	Value
Young's modulus	100.0 MPa
Poisson's ratio	0.25
Biot coefficient	1.0
Overload	1000.0 Pa
Permeability	$1.0e^{-14} m^2$
Dynamic viscosity	$1.0e^{-5} m^2/sec$
Time step length	$1.0 e^{-3} sec$
Column length	15.0 m

The analytical solution is given by Verruijt [2013] through Laplace transformation:

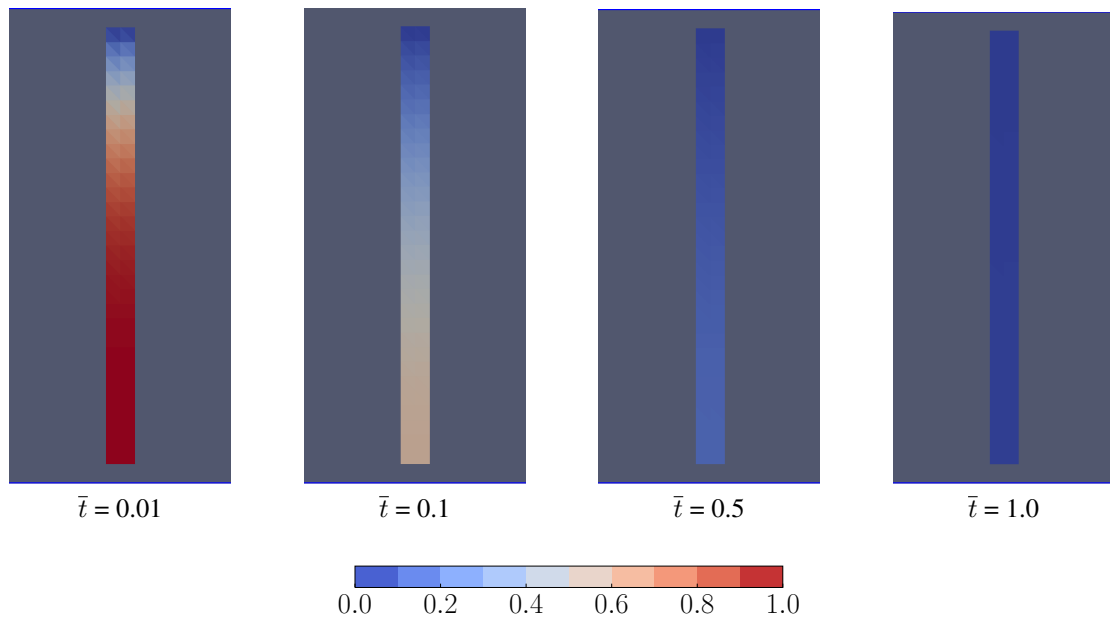
$$\bar{p} = \frac{4}{\pi} \sum_{k=1}^{\infty} \cos\left[(2k-1)\frac{\pi}{2}\bar{x}\right] \exp\left[-(2k-1)^2\frac{\pi^2}{4}\bar{t}\right], \quad (25)$$

where  $\bar{p}$  is normalized pressure  $\frac{p}{p_0}$ ,  $\bar{x}$  is normalized distance  $\frac{x}{L}$ ,  $\bar{t}$  is normalized time  $\frac{c_v t}{h^2}$ , and  $c_v$  is consolidation coefficient.

Our numerical results are shown in Figure 2 and Figure 3. It can be observed from Figure 2 that the agreement between our numerical result and the analytical solution appears excellent. At time  $t = 0$ , a constant load is applied to the top surface. This sudden increase in load will be initially suffered by the water, and thus the water pressure goes up everywhere in the beam. The pore pressure then gradually vanishes, since the top surface is fully drained. This dissipation process may take considerable time, depending on the permeability value.



**Figure 2: Comparison of the numerical and analytical solutions of Terzaghi's problem: excess pore pressure at different normalized time.**

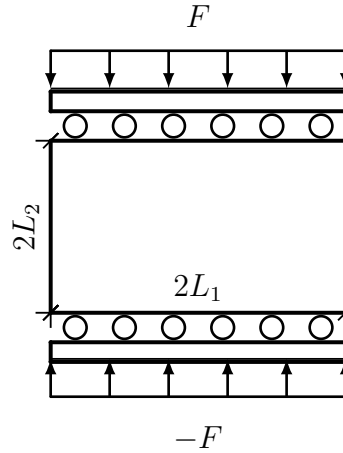


**Figure 3: Normalized pressure distribution of Terzaghi's problem at different normalized time**

### *Mandel's 2D problem*

In this example, we consider the 2D consolidation problem from Mandel [1953]. As illustrated in Figure 4, a rectangular soil sample is sandwiched between rigid frictionless plates at its top and bottom. The top and bottom surface are impermeable and the lateral surfaces are allowed to drain freely. The length and width of the sample are  $2L_1 = 2$  m and  $2L_2 = 2$  m, respectively. Other physical and computational parameters are given in Table 2.





**Figure 4: Sketch of Mandel's problem.**

**Table 2: Parameters in Mandel's problem**

Parameter	Value
Young's modulus	100.0 MPa
Poisson's ratio	0.25
Biot coefficient	1.0
Overload	1000.0 Pa
Permeability	$1.0e^{-14} m^2$
Dynamic viscosity	$1.0e^{-5} m^2/s$
Time step length	$1.0e^{-3} s$
Plate length	1.0 m
Plate width	1.0 m

The analytical solution is given by Abousleiman et al. [1996] and Coussy [2004] as:

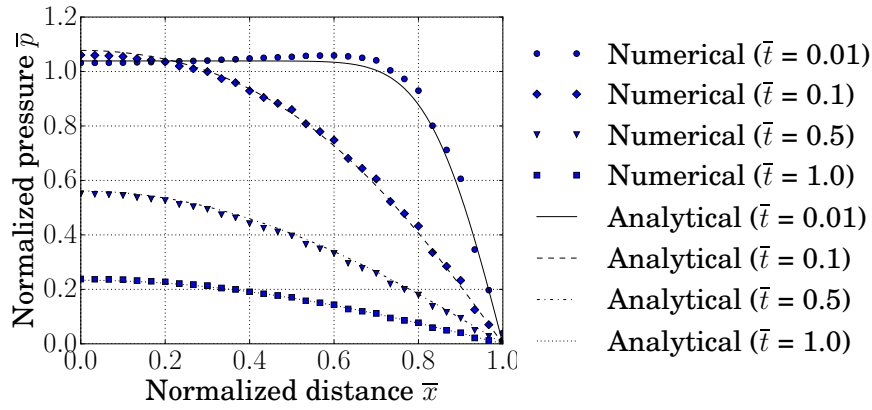
$$\bar{p} = 2 \sum_{k=1}^{\infty} \frac{\cos(\alpha_k \bar{x}) - \cos \alpha_k}{\alpha_k - \sin \alpha_k \cos \alpha_k} \exp(-\alpha_n^2 \bar{t}), \quad (26)$$

where  $\alpha_k$  is the solution of:

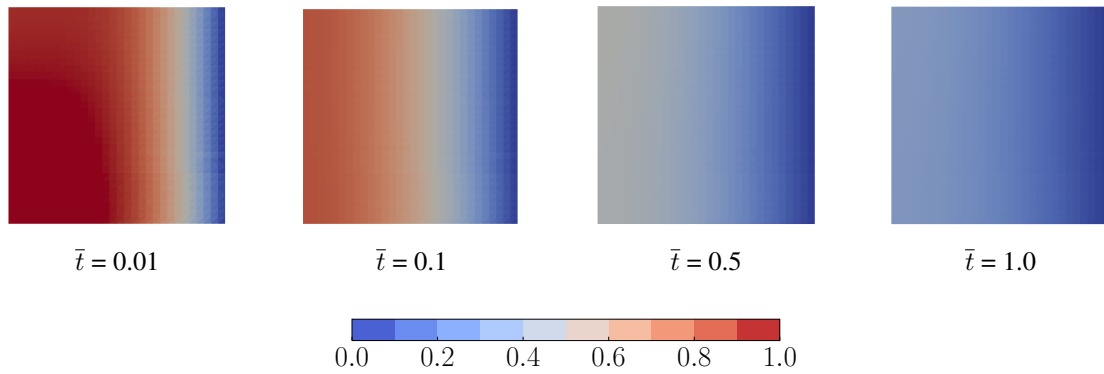
$$\frac{\tan \alpha_k}{\alpha_n} = \frac{1 - v}{v_n - v}, \quad (27)$$

and  $v_n$  and  $v$  stands for undrained and drained Poisson ration, respectively. For incompressible constituents,  $v_n$  is 0.5.

The numerical results in terms of normalized pressure versus normalized time are presented in Figure 5 and Figure 6. Similar to Terzaghi's 1D problem, the pressure jumps to some value and then gradually dissipates. It is important to note that the pressure at the centre of plate continues to increase after its initial creation by the Skempton effect [Skempton, 1954]. This is due to the fact that the generation of pore pressure is immediate, but the dissipation caused by the fluid flow is delayed by the small permeability and the flow path to escape from the lateral boundaries.



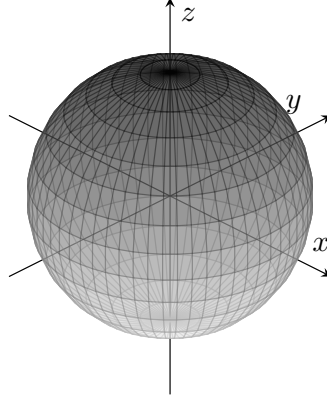
**Figure 5: Comparison of the numerical and analytical solutions of Mandel’s problem: excess pore pressure at different normalized time.**



**Figure 6: Normalized pressure distribution of Mandel’s problem at different normalized time**

*Cryer’s 3D problem*

Finally, we consider the sphere from Cryer [1963], a classic 3D consolidation problem. In this case, a 3D spherical soil sample, of radius  $a = 1.0\text{ m}$ , is loaded on its outer boundary by an instantaneous uniform confining pressure of magnitude  $F$ , see Figure 7. Table 3 summarizes all the useful parameters.



**Figure 7: Sketch of Cryer's problem.**

**Table 3: Parameters in Cryer's problem**

Parameter	Value
Young's modulus	10.0 MPa
Poisson's ratio	0.33
Biot coefficient	1.0
Overload	1000.0 Pa
Permeability	$4.95e^{-14} m^2$
Dynamic viscosity	$1.0e^{-5} m^2/s$
Time step length	$1.0e^{-2} s$
Sphere radius	1.0 m

We are interested in the pressure evolution at the centre of the sphere and the analytical solution is given by Verruijt [2013] as:

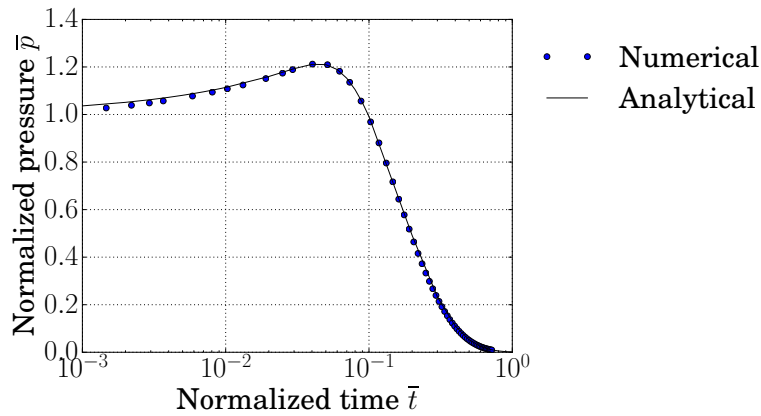
$$\bar{p} = \eta \sum_{k=1}^{\infty} \frac{\sin \alpha_k - \alpha_k}{\eta \alpha_k \cos \alpha_k / 2 + (\eta - 1) \sin \alpha_k} \exp(-\alpha_k^2 \bar{t}), \quad (28)$$

where  $\alpha_k$  is the positive roots of the equation:

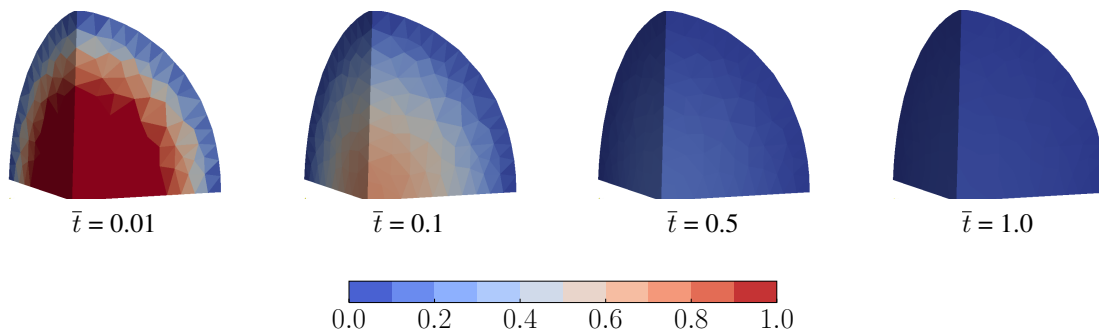
$$(1 - \eta \alpha_k^2) \tan \alpha_k = \alpha_k, \quad (29)$$

where  $\eta$  is defined as  $\frac{\lambda + 2G}{2G}$ .

We show the numerical results in Figure 8 and Figure 9. Initially when  $t = 0$ , the pressure at the centre jumps by an amount, and then continues to rise for a while before declining, as in the case of Mandel's problem. This non-monotonic pore pressure response, a rise in interior fluid pressure and the subsequent decaying to zero value, is referred to as the Mandel-Cryer effect. This is a distinctive phenomenon of Boit's consolidation, which is not observed by the traditional uncoupled theory [Terzaghi et al., 1943]. As shown in Figure 8, our model greatly captures the Mandel-Cryer effect.



**Figure 8: Comparison of the numerical and analytical solutions of Cryer's problem: excess pore pressure at the center.**



**Figure 9: Normalized pressure distribution of Cryer's problem at different normalized time**

## Conclusion

In this work, we have developed a fully coupled 3D model for flow-geomechanical simulation in porous media. By choosing fluid pressure, velocity and solid displacement as primary variables, the proposed mixed finite element formulation is able to ensure local mass conservation and express the velocity and displacement in a continuous way. The fully coupled approach is unconditionally stable, and eliminates the convergence issues encountered in explicit scheme and iterative coupled scheme. The performance of the resulting numerical model is validated according to Terzaghi's 1D problem, Mandel's 2D problem and Cryer's 3D problem. In all cases, our numerical results show good agreement with the analytical solutions.

## Acknowledgements

The authors wish to acknowledge the financial support from China Scholarship Council, and the contribution of the NeSI high-performance computing facilities and the eResearch centre at the University of Auckland. The first author would also like to express his gratitude to Dr. Jon Pearce for his invaluable help in developing the codes.

## REFERENCES

- Aarnes, Jørg E, Tore Gimse, and Knut-Andreas Lie. An introduction to the numerics of flow in porous media using matlab. In *Geometric modelling, numerical simulation, and optimization*, pages 265–306. Springer, 2007.
- Abousleiman, Y, AH-D Cheng, L Cui, E Detournay, and J-C Roegiers. Mandel's problem revisited. *Geotechnique*, 46(2):187–195, 1996.
- Biot, Maurice A. General theory of three-dimensional consolidation. *Journal of applied physics*, 12(2):155–164, 1941.
- Coussy, Olivier. *Poromechanics*. John Wiley & Sons, 2004.
- Cryer, CW. A comparison of the three-dimensional consolidation theories of biot and terzaghi. *The Quarterly Journal of Mechanics and Applied Mathematics*, 16(4):401–412, 1963.
- Edwards, Michael G. Unstructured, control-volume distributed, full-tensor finite-volume schemes with flow based grids. *Computational Geosciences*, 6(3-4):433–452, 2002.
- Fortin, Michel and Franco Brezzi. *Mixed and hybrid finite element methods*. New York: Springer-Verlag, 1991.
- Gambolati, Giuseppe, Giuseppe Ricceri, Werter Bertoni, Giovanni Brighenti, and Enzo Vuillermin. Mathematical simulation of the subsidence of ravenna. *Water Resources Research*, 27(11):2899–2918, 1991.
- Hesthaven, Jan S and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springer Science & Business Media, 2007.
- Jenny, P, SH Lee, and HA Tchelepi. Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. *Multiscale Modeling & Simulation*, 3(1):50–64, 2005.
- Khoshghalb, Arman and Nasser Khalili. A stable meshfree method for fully coupled flow-deformation analysis of saturated porous media. *Computers and Geotechnics*, 37(6):789–795, 2010.
- Lee, SH, LJ Durlofsky, MF Lough, WH Chen, and others. Finite difference simulation of geologically complex reservoirs with tensor permeabilities. *SPE Reservoir Evaluation and Engineering*, 1(6):567–574, 1998.
- Lewis, Roland Wynne and Bernard A Schrefler. *The finite element method in the deformation and consolidation of porous media*. 1987.
- Logg, Anders, Kent-Andre Mardal, and Garth Wells. *Automated solution of differential equations by the finite element method: The FEniCS book*, volume 84. Springer Science & Business Media, 2012.
- Longuemare, P, M Mainguy, P Lemonnier, A Onaisi, Ch Gérard, and N Koutsabeloulis. Geomechanics in reservoir simulation: overview of coupling methods and field case study. *Oil & Gas Science and Technology*, 57(5):471–483, 2002.
- Mandel, J. Consolidation des sols (étude mathématique)\*. *Geotechnique*, 3(7):287–299, 1953.
- Masson, Yder J and Steven R Pride. Poroelastic finite difference modeling of seismic attenuation and dispersion due to mesoscopic-scale heterogeneity. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 112(B3), 2007.

- Mikelić, Andro and Mary F Wheeler. Convergence of iterative coupling for coupled flow and geomechanics. *Computational Geosciences*, 17(3):455–461, 2013.
- Minkoff, Susan E, C Mike Stone, Steve Bryant, Malgorzata Peszynska, and Mary F Wheeler. Coupled fluid flow and geomechanical deformation modeling. *Journal of Petroleum Science and Engineering*, 38(1):37–56, 2003.
- Mosharaf Dehkordi, Mehdi, Mehrdad T. Manzari, H Ghafouri, and R Fatehi. A general finite volume based numerical algorithm for hydrocarbon reservoir simulation using blackoil model. *International Journal of Numerical Methods for Heat & Fluid Flow*, 24(8):1831–1863, 2014.
- Panneton, Raymond and Noureddine Atalla. An efficient finite element scheme for solving the three-dimensional poroelasticity problem in acoustics. *The Journal of the Acoustical Society of America*, 101(6):3287–3298, 1997.
- Pao, William KS, Roland W Lewis, and Ian Masters. A fully coupled hydro-thermo-poro-mechanical model for black oil reservoir simulation. *International journal for numerical and analytical methods in geomechanics*, 25(12):1229–1256, 2001.
- Rozon, Brad J and others. A generalized finite volume discretization method for reservoir simulation. In *SPE Symposium on Reservoir Simulation*. Society of Petroleum Engineers, 1989.
- Settari, Antonin, FM Mourits, and others. A coupled reservoir and geomechanical simulation system. *Spe Journal*, 3(03):219–226, 1998.
- Skempton, AW. The pore-pressure coefficients a and b. *Geotechnique*, 4(4):143–147, 1954.
- Teatini, P, M Ferronato, G Gambolati, and M Gonella. Groundwater pumping and land subsidence in the emilia-romagna coastland, italy: Modeling the past occurrence and the future trend. *Water Resources Research*, 42(1), 2006.
- Terzaghi, Karl. *Soil mechanics in engineering practice*. John Wiley & Sons, 1996.
- Terzaghi, Karl, Karl Terzaghi, Civil Engineer, Austria Czechoslovakia, Karl Terzaghi, Ingénieur Civil, Autriche Tchecoslovaquie, and Etats Unis. *Theoretical soil mechanics*, volume 18. Wiley New York, 1943.
- Tran, David, Antonin Settari, Long Nghiem, and others. New iterative coupling between a reservoir simulator and a geomechanics module. *SPE Journal*, 9(03):362–369, 2004.
- Verruijt, Arnold. Theory and problems of poroelasticity. *Delft University of Technology*, 2013.
- Wheeler, Mary, Guangri Xue, and Ivan Yotov. Coupling multipoint flux mixed finite element method with continuous galerkin methods for poroelasticity. *Computational Geosciences*, 18(1):57–75, 2014.
- White, Joshua A and Ronaldo I Borja. Block-preconditioned newton–krylov solvers for fully coupled flow and geomechanics. *Computational Geosciences*, 15(4):647–659, 2011.
- Yanosik, JL, TA McCracken, and others. A nine-point, finite difference reservoir simulator for realistic prediction of adverse mobility ratio displacements. *Soc. Pet. Eng. J*, 19(4):253–262, 1979.
- Yin, Shunde, Maurice B Dusseault, and Leo Rothenburg. Thermal reservoir modeling in petroleum geomechanics. *International journal for numerical and analytical methods in geomechanics*, 33(4):449–485, 2009.

## **Adaptive combined DE/FE algorithm for analyzing impact fracture problem**

**Wei Xu\* and Jiro Sakamoto**

School of Mechanical Engineering, College of Science and Engineering, Kanazawa University,  
Kanazawa, 920-1192, Japan

\*Corresponding and presenting author: [xwteamobj007@gmail.com](mailto:xwteamobj007@gmail.com)

### **Abstract**

This paper presents a description of a 3D adaptive combined DE/FE algorithm which can automatically convert the distorted finite elements into the spherical discrete elements during simulating the impact fracture of laminated glass. In this method a system is completely discretized into the finite elements at the initial moment without any discrete element existing until part of the finite elements becoming severely deformed. Subsequently each finite element, whose maximum tensile stress exceeds a user-specified conversion criterion, is converted into eight spherical discrete elements. At the same time the system is fragmented into two subdomains, the finite element (FE) and the discrete element (DE) subdomains. The impact fracture of a glass beam is simulated by the adaptive algorithm and the discrete element method, respectively. A satisfactory agreement of the simulation results is observed which validates the feasibility of such an adaptive algorithm; however, the computational efficiency of the adaptive algorithm is much higher than that of latter.

**Keywords:** Combined algorithm; Brittle fracture; Cohesive model; Discrete element method; Laminated glass

### **Introduction**

Laminated glass generally consists of two or more layers of glass sheets combined by the elastomeric interlayers of Polyvinyl Butyral (PVB) under heat treatment and has been widely used in automobile windshields, modern buildings, etc. thanks to its security and durability performance. The mechanical properties of laminated glass are more complicated than those of single glass due to the brittleness of glass, the hyper-elasticity of PVB interlayer and the coupled influence of both formers.

The fracture behavior and the flexural strength of laminated glass were not only influenced by the flexural stiffness of each layer but also by their interfacial bonding strength. A parametric study was carried out by Hidallana-Gamage et al. to investigate the influence of structural sealant joints on the mechanical properties of laminated glass panels under blast loading, the information in which might be employed to complement the guidance in the existing design standards [Hidallana-Gamage et al (2014)]. Foraboschi implemented the sacrificial ply design concept to design laminated glass in which the outer glass as a sacrificial ply was permitted to damage under small impact while the inner was designed to remain unbroken [Foraboschi (2007; 2013)]. The fracture behaviors of laminated glass windows subjected to impulsive and blast loadings were experimentally studied with the laboratory airbag pendulum impact tests and the full-scale field blast tests, respectively [Zhang et al (2015)]. The responses of laminated glass in the lab and field tests, such as failure process, applied pressure and deflections, were used to validate the accuracy of the design standards predictions.

Most of numerical studies mentioned above were preceded in the framework of the finite element method (FEM). The contact problem over the crack surfaces and the fragments dispersion might be present, which were of great importance for the fracture of laminated glass. The essence of material fracture was a complicated transi-

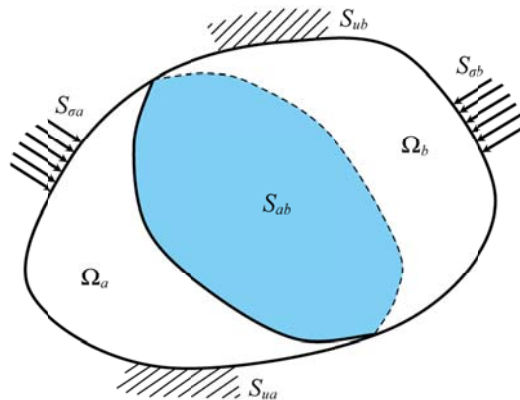
tion from continuum to non-continuum. However, it was very difficult to calculate the contact force over the crack surfaces and capture the dispersing fragments in the FEM framework. Fortunately, it was very simple to deal with the transition process by changing the joint types between the DEs in the framework of the discrete element method (DEM). The DEM proposed by Cundall [Cundall (1971)] had been employed to simulate the fracture behavior of brittle materials [Oda and Yasuda (2007); Griffiths and Mustoe (2001); Zang et al (2007); Shan et al (2009)].

To take full advantages of the both numerical methods mentioned above, a 3D adaptive combined DE/FE algorithm has been developed, which can automatically convert the distorted finite elements into the spherical discrete elements during the impact fracture simulations. The detailed description about the 3D adaptive algorithm is presented in Section 2. The corresponding numerical code has been developed in Fortran 90/95 programming language. In Section 3, the fracture simulation is performed on a three-point bending glass beam and the feasibility of the 3D adaptive algorithm is validated by comparing the fracture procedures simulated by the proposed algorithm and the DEM.

### Adaptive combined DE/FE algorithm

#### *Introduction of the combined DE/FE method*

A continuous elastic solid (solution domain) is considered with the volume  $\Omega$  and the surface  $S$  as shown in Figure 1. It is fragmented into two subdomains,  $\Omega_a$  and  $\Omega_b$  with the surfaces  $S_{\sigma a}$ ,  $S_{ua}$  and  $S_{\sigma b}$ ,  $S_{ub}$ , respectively, which are to be joined together along an interface  $S_{ab}$ . Here  $S_{\sigma a}$  and  $S_{\sigma b}$  are the surfaces prescribed the external forces,  $S_{ua}$  and  $S_{ub}$  are the surfaces prescribed displacements [Lei and Zang (2010)]. In this work the solution domain is divided into the DE and the FE subregions. The constraint condition between both subregions is enforced on the interface  $S_{ab}$  by using the penalty method.

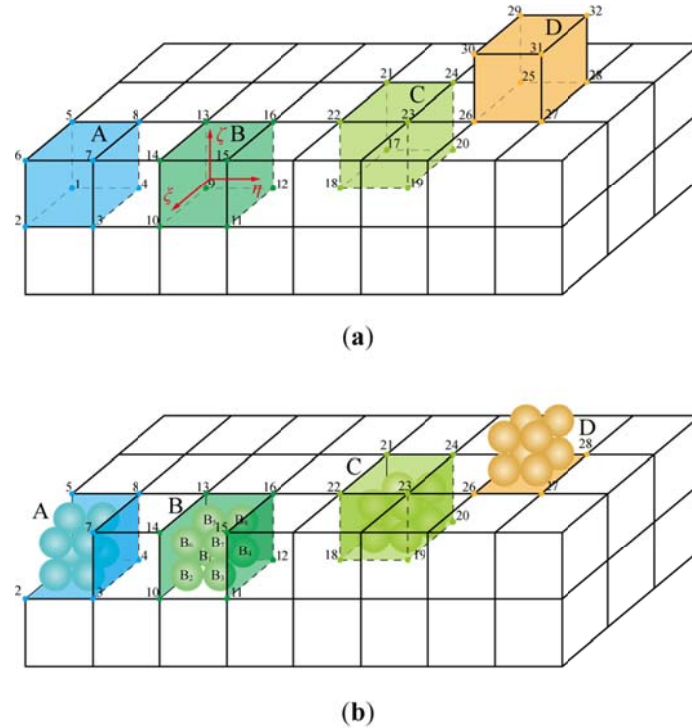


**Figure 1.** Solution domain divided into two subdomains

#### *The adaptive procedure*

Figure 2(a) illustrates a finite element grid with four 8-node cubic FEs (A, B, C and D) on the surface designated as the candidates for conversion into the spherical DEs. In the grid only the nodes 1, 2, 3, ... 32 are numbered for simplicity. The full integration is employed in the FE model. If the maximum tensile stress of a FE, such as the finite element B, exceeds a user-specified criterion, named the conversion stress, the following five steps are taken to convert it into eight spherical DEs. The conversion keeps the mass, the momentum and the energy conservation laws approximately. The deformation of the FE is approximately mirrored by the separation or penetration (the relative displacement in the local coordinate) between the DEs; the stress state of the FE by the internal force between the DEs. The internal force is calculated by the spring constants and the relative displacement between the DEs.





**Figure 2.** Conversion of the distorted finite elements into the spherical discrete elements. **a** The finite elements before conversion, **b** The finite and discrete elements after conversion

(1) The discrete element  $B_i$  ( $i = 1, 2, 3, \dots, 8$ ) is generated at the integration point calculated by the interpolation of the nodal positions of the finite element B,

$$\mathbf{p}^i = \sum_I N_I(\xi_i, \eta_i, \zeta_i) \mathbf{p}^I \quad (1)$$

where  $\mathbf{p}^i$  is the position of the discrete element  $B_i$ ;  $\mathbf{p}^I$  is the nodal position of the finite element B;  $N_I(\xi_i, \eta_i, \zeta_i)$  is the value of the shape function, here  $(\xi_i, \eta_i, \zeta_i)$  is the natural coordinate of the center of the discrete element  $B_i$ . The radius of each DE is  $r = l/4$ , here  $l$  is the original edge length of the finite element B.

(2) The mass of the discrete element  $B_i$  is set to be  $m_f/8$ , in which  $m_f$  is the mass of the finite element B; the translational displacement and velocity of the discrete element  $B_i$  are calculated by the interpolations of the nodal displacements and velocities of the finite element B, respectively:

$$\mathbf{u}^i = \sum_I N_I(\xi_i, \eta_i, \zeta_i) \mathbf{u}^I \quad (2)$$

$$\dot{\mathbf{u}}^i = \sum_I N_I(\xi_i, \eta_i, \zeta_i) \dot{\mathbf{u}}^I \quad (3)$$

in which  $\mathbf{u}^i$  and  $\dot{\mathbf{u}}^i$  are the translational displacement and velocity of the discrete element  $B_i$ , respectively;  $\mathbf{u}^I$  and  $\dot{\mathbf{u}}^I$  are the nodal displacement and velocity of the finite element B, respectively.

(3) The finite element B is removed from the model grid and the nodal mass of the finite element B is also reduced from the FE model. The reduced nodal mass is  $m_f/8$ . If a node does not belong to any finite element, it should be removed from the nodal

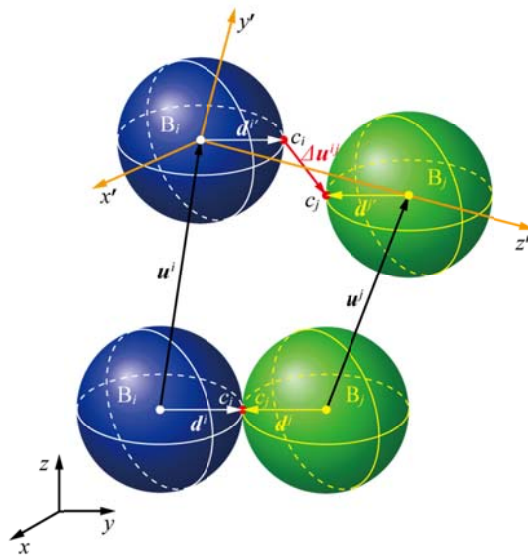
list.

(4) If the surfaces 13-14-15-16 and 10-11-15-14 of the finite element B contacts with the other parts, they are not effective now and should be removed from the contact surface segments. The contact force calculation is replaced by the contact between the discrete elements  $B_5$ - $B_6$ - $B_7$ - $B_8$  and  $B_2$ - $B_3$ - $B_7$ - $B_6$  and the other parts. The discrete elements  $B_1$ ,  $B_2$ ,  $B_6$  and  $B_5$  are combined with the surface 9-10-14-13;  $B_1$ ,  $B_4$ ,  $B_8$  and  $B_5$  with the surface 9-12-16-13;  $B_3$ ,  $B_4$ ,  $B_8$  and  $B_7$  with the surface 11-12-16-17;  $B_1$ ,  $B_2$ ,  $B_3$  and  $B_4$  with the surface 9-10-11-12. The interface force calculation will be introduced in the following subsection.

(5) The connective joint type between the DEs is guaranteed at the moment when the DEs are generated by setting the conversion stress to be about  $(0.90-0.95)\sigma_t$ , slightly smaller than the tensile strength of the brittle material. As shown in Figure 3 the translational displacements of the discrete elements  $B_i$  and  $B_j$  are  $\mathbf{u}^i$  and  $\mathbf{u}^j$ , respectively. The vectors  $\mathbf{d}^i$  and  $\mathbf{d}^j$  are employed to record the rotations of the discrete elements  $B_i$  and  $B_j$ , respectively. The normal and the tangential springs between the DEs are connected by the two points  $c_i$  and  $c_j$  on the discrete elements  $B_i$  and  $B_j$ , respectively. Since the rotation of the DE is neglected for the small deformation problems, the two vector  $\mathbf{d}^i = \mathbf{d}^{i'}$  and  $\mathbf{d}^j = \mathbf{d}^{j'}$ . Finally the relative displacement  $\Delta\mathbf{u}^{ij}$  between the discrete elements  $B_i$  and  $B_j$  in the local coordinate  $x'y'z'$  is calculated by the following formula:

$$\Delta\mathbf{u}^{ij} = \mathbf{N}(\mathbf{u}^j - \mathbf{u}^i) \quad (4)$$

in which  $\mathbf{N}$  is the transform matrix from the global coordinate  $xyz$  to the local coordinate  $x'y'z'$ .



**Figure 3.** Relative displacement between the discrete elements  $B_i$  and  $B_j$

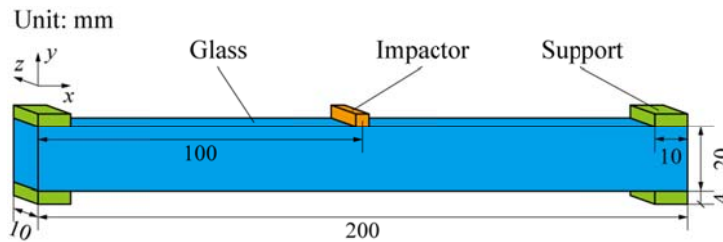
For the conversion of the finite element A, C, and D, most of the steps are similar to those adopted for the finite element B. The final elements configuration after conversion is illustrated in Figure 2(b).

### Numerical simulation

In this part, the impact fracture simulations are performed on a three-point bending glass beam by using the 3D adaptive algorithm and the DEM, respectively. The resulting fracture behaviors are compared with each other to validate the feasibility of

the 3D adaptive algorithm whose computational time is compared with that of the DEM to verify the former's higher computational efficiency.

The geometry of the three-point bending glass beam is illustrated in Figure 4. The dimension of the glass is taken to be 200 mm in span, 20 mm in depth and 10 mm in width. The size of each support is 10 mm  $\times$  4 mm  $\times$  10 mm. The impactor is just on the top of the glass beam, whose initial velocity is  $v_y = -3.13$  m/s, size 4 mm  $\times$  4 mm  $\times$  10 mm and mass 1.0 kg. For the adaptive model the glass is discretized into 4950 8-node cubic FEs with a size of 2 mm and 100 hexahedral FEs with a size of 1 mm  $\times$  2 mm  $\times$  2 mm; each support 50 cubic FEs with a size of 2 mm; the impactor 20 cubic FEs with a size of 2 mm which are the rigid body elements. For the DE model, the glass is discretized into 40000 spherical DEs with a radius of 0.5 mm.



**Figure 4.** Impact fracture model of the three-point bending beam

The material properties of the three-point bending glass beam model are listed in Table 1. In the numerical simulations the conversion stress is assumed to be 56.0 MPa to guarantee the connective joint type in the DE subregion at the moment when the DEs are generated.

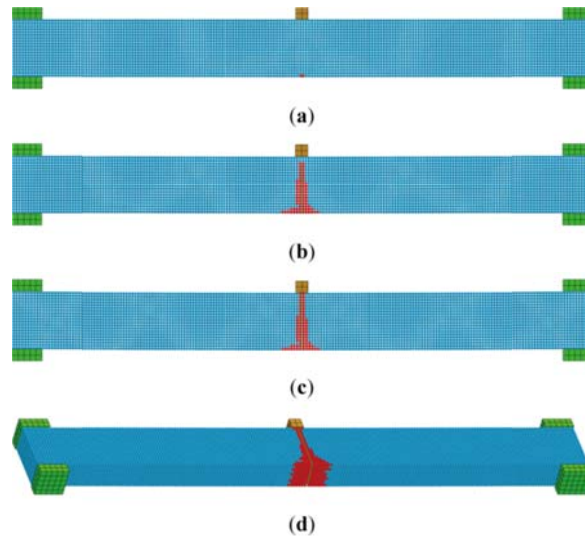
**Table 1.** Material properties used in the simulation

Material	Glass	Support	Impactor
Density $\rho$ (kg/m <sup>3</sup> )	2500.0	2400.0	-
Young's modulus $E$ (GPa)	74.09	5.0	210.0
Poisson's ratio $\nu$	0.2	0.4	0.269
Tensile strength $\sigma_t$ (MPa)	60.0	-	-
Energy release rate $G_f$ (N/m)	10.0	-	-
Shear stress factor $\beta$	0.10	-	-

Based on the theory mentioned above, a numerical code is developed in Fortran 90/95 programming language. The impact fracture procedure of the glass beam is simulated by using the numerical code on a Dell Precision T5810 Workstation. The time step  $\Delta_t$  is  $6 \times 10^{-5}$  ms small enough to ensure calculation stability for both of the DE and the adaptive models. The physical time is about 100  $\mu$ s. The computational time of the 3D adaptive algorithm is about 67s while that of the DEM about 660s. The computational efficiency of the 3D adaptive algorithm is much higher than that of the DEM. Furthermore, almost the same crack patterns are captured as illustrated in Figures 5 and 6.

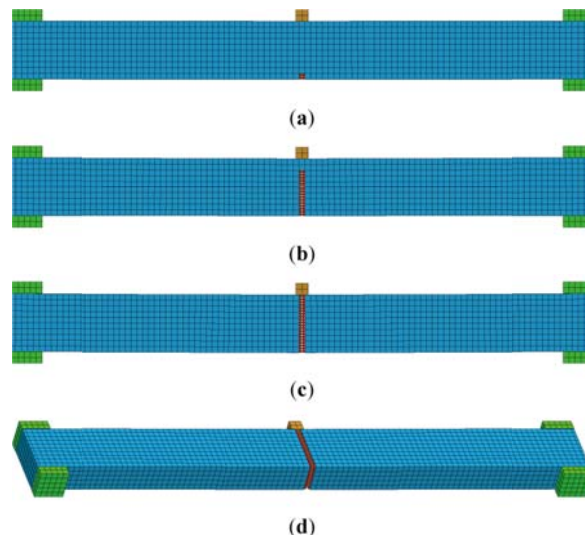
Figure 5 illustrates the fracture procedure of the glass beam with the DE model. In order to illustrate the damage of the glass the DEs' color is set to be red if the joint model between the DEs is changed from the connective model to the cohesive or the contact models and the main crack is marked by green lines. As shown in Figure 5(a), the glass beam begins to damage at the bottom at about 25  $\mu$ s. Subsequently, one main crack progresses upwards in a very short period of time (about 45  $\mu$ s) to the upper of the glass beam as shown in Figures 5(b) and 5(c) and the crack reaches the opposite

side at about  $70\mu\text{s}$ . Figure 5(d) is the oblique view of the final crack pattern at about  $100\mu\text{s}$ , in which the main crack is surrounded by microcracks and the cohesive zone located in the red region.



**Figure 5.** Fracture procedure with the DE model. **a-c** The crack patterns of the middle section by the  $xy$  plane: **a**  $t = 25\mu\text{s}$ , **b**  $t = 50\mu\text{s}$ , **c**  $t = 70\mu\text{s}$ ; **d** The oblique view of the final crack pattern at about  $100\mu\text{s}$

Figure 6 illustrates the fracture procedure of the glass beam with the adaptive model. Five finite elements are converted into forty spherical DEs. Following the conversion procedure, the glass beam begins to fracture at the bottom at about  $26\mu\text{s}$  as shown in Figure 6(a). Subsequently, the conversion procedure progresses upwards in a very short period of time (about  $44\mu\text{s}$ ) to the upper of the glass beam and one main crack propagates only in the DE subdomain as shown in Figures 6(b) and 6(c) and the crack reaches the opposite side at about  $70\mu\text{s}$ . Figure 6(d) is the oblique view of the final crack pattern at about  $100\mu\text{s}$ , in which only the main crack is captured in the adaptive model.



**Figure 6.** Fracture procedure with the adaptive model. **a-c** The crack patterns of the middle section by the  $xy$  plane: **a**  $t = 26\mu\text{s}$ , **b**  $t = 50\mu\text{s}$ , **c**  $t = 70\mu\text{s}$ ; **d** The oblique view of the final crack pattern at about  $100\mu\text{s}$

From the above investigations, we notice that the crack paths and the moment of crack occurrence of the adaptive model in Figure 6 are almost unanimous as those of

the DE model in Figure 5. The differences are small; however, one more important characteristic of the 3D adaptive algorithm is the higher computational efficiency.

## Conclusion

- (1) A 3D adaptive combined DE/FE algorithm is proposed to analyze the impact fracture problem with a higher computational efficiency.
- (2) The impact fracture of a glass beam is simulated with a higher efficiency by the 3D adaptive algorithm than the DEM. Furthermore, almost the same crack patterns are captured with both numerical methods.

In this work, the same time step is adopted in the FE and the DE subregions and only the cubic FEs can be converted into the spherical DEs. In the future we will overcome these obstacles for more efficiently analyzing the impact fracture of laminated glass with an irregular geometry.

## Reference

- Cundall P.A. (1971) A computer model for simulating progressive large scale movements in blocky rock systems. In *Proceedings of the Symposium of the International Society of Rock Mechanics* **1**:129-136.
- Foraboschi P. (2007) Behavior and failure strength of laminated glass beams. *Journal of Engineering Mechanics* **133** (12):1290-1301.
- Foraboschi P. (2013) Hybrid laminated-glass plate: Design and assessment. *Composite Structures* **106**:250-263.
- Hidallana-Gamage H., Thambiratnam D., Perera N. (2014) Influence of structural sealant joints on the blast performance of laminated glass panels. *Journal of Performance of Constructed Facilities*, Doi:10.1061/(ASCE)CF.1943-5509.0000646, 04014151.
- Griffiths D.V., Mustoe G.G.W. (2001) Modelling of elastic continua using a grillage of structural elements based on discrete element concepts. *International Journal for Numerical Methods in Engineering* **50**(7):1759-1775.
- Lei Z., Zang M.Y. (2010) An approach to combining 3D discrete and finite element methods based on penalty function method. *Computational Mechanics* **46**(4):609-619.
- Oda J., Yasuda H. (2007) On impact penetration analysis and evaluation of brittle material plate using three-dimensional DEM (in Japanese). *Journal of Solid Mechanics and Materials Engineering* **1**(3):376-386.
- Shan L., Cheng M., Liu K.X. et al. (2009) New discrete element models for three-dimensional impact problems. *Chinese Physics Letters* **26**(12):120202.
- Zang M.Y., Lei Z., Wang S.F. (2007) Investigation of impact fracture behavior of automobile laminated glass by 3D discrete element method. *Computational Mechanics* **41**(1):73-83.
- Zhang X.H., Hao H., Wang Z.Q. (2015) Experimental study of laminated glass window responses under impulsive and blast loading. *International Journal of Impact Engineering* **78**:1-19.

# Scaled Boundary Finite Element Analysis of three-dimensional crack configurations in laminate structures

†Sascha Hell<sup>1</sup>, \*Wilfried Becker<sup>1</sup>

<sup>1</sup>Department of Structural Mechanics, TU Darmstadt, Germany

\*Presenting author: becker@fsm-tu-darmstadt.de

†Corresponding author: hell@fsm-tu-darmstadt.de

## Abstract

Laminate structures composed of fibre-reinforced plies typically are prone to the formation of inter-fiber cracks because of the given strongly anisotropic stiffness and strength properties. These inter-fiber cracks commonly run through complete plies but are stopped at the ply interfaces. Equally, such laminate structures are prone to the formation of delaminations, e.g. due to the free-edge effect. An inter-fiber crack meeting a delamination forms a non-standard three-dimensional crack configuration with a locally singular stress field that should be investigated in regard of its criticality.

For that purpose, the Scaled Boundary Finite Element Method turns out to be an appropriate and effective analysis method that permits solving linear elastic mechanical problems including stress singularities with comparably little effort. Only the boundary is discretized by two-dimensional finite elements while the problem is considered analytically in the direction of the dimensionless radial coordinate  $\xi$ . A corresponding separation of variables representation for the displacement field employed in the virtual work equation leads to a system of differential equations of Cauchy-Euler type. This differential equation system can be converted into an eigenvalue problem and solved by standard eigenvalue solvers for non-symmetric matrices.

By this kind of analysis, it is revealed that the considered three-dimensional crack configurations may go along with various unexpected non-standard stress singularities, namely singularities that are weaker than the well-known square root stress singularity in linear elastic fracture mechanics, but also singularities that are stronger and which may be called hypersingularities.

**Keywords:** Scaled Boundary Finite Element Method, laminate, inter-fiber crack, delamination, stress singularity

## Introduction

Unidirectionally fiber-reinforced materials as they are used in structural components, exhibit a strongly anisotropic material behavior. They offer very high stiffness and strength properties in the fiber direction but low values in the transverse directions. This makes them prone to the formation of matrix cracks between the fibers, so-called inter-fiber cracks. In a laminated structure of unidirectionally fiber-reinforced plies, these inter-fiber cracks, at first, only lead to some stiffness degradation but also to locally new structural situations with a highly complex failure evolution (see e.g. [Leguillon and Martin (2012)]). Because of this complexity and moreover the lack of predictability of the failure evolution, it is still common practice in industrial composite design to assume laminate failure when the so-called First-Ply-Failure occurs, i.e. when the first inter-fiber cracks emerge. Especially for quasi-isotropic carbon fiber reinforced plastic (CFRP) laminates with a ply lay-up of the kind  $[\pm 45^\circ / 0^\circ / 90^\circ]_s$ , as they are almost exclusively used in the aircraft industry, this assumption dramatically underestimates the true load bearing capacity of such a composite laminate. To achieve further progress in this field, it is crucial to gain a proper understanding of these structural situations and their possible interactions with other defects. Such other defects are e.g. delaminations which commonly occur due to impact loads or also as a consequence of the laminate free-edge effect. A delamination meeting a transverse inter-fiber crack is, thus, a possible crack configuration that needs closer investigation.

A special challenge of such structural situations within the framework of linear elasticity theory is the occurrence of theoretically infinite stresses. Stress singularities typically occur at discontinuities

of geometry and material. The near-field solution at a singular point for the displacements and stresses respectively is usually represented by a power law function series of the kind

$$\bar{u}(r, \varphi, \vartheta) = \sum_{j=1}^{\infty} c_{uj} r^{\lambda_j} (\ln r)^k \bar{\Phi}_{uj}(\varphi, \vartheta) \quad , \quad \boldsymbol{\sigma}(r, \varphi, \vartheta) = \sum_{j=1}^{\infty} c_{\sigma j} r^{\lambda_j-1} (\ln r)^k \boldsymbol{\Phi}_{\sigma j}(\varphi, \vartheta) \quad (1)$$

given in spherical coordinates  $r, \varphi, \vartheta$ . Here,  $\lambda_j$  are complex numbers and  $\bar{\Phi}_{uj}$  and  $\boldsymbol{\Phi}_{\sigma j}$  are vector and tensor functions of the angle coordinates  $\varphi$  and  $\vartheta$ . For  $\text{Re}(\lambda_j) - 1 < 0$  and  $r \rightarrow 0$  the stress tensor becomes singular and the quantities  $\text{Re}(\lambda_j) - 1$  are called stress singularity exponents. The free constants  $c_{uj}$  and  $c_{\sigma j}$  are not independent of each other and need to be determined from the boundary conditions. However, this means that it depends on the boundary conditions whether a stress singularity actually occurs. The exponent  $k$  depends on the geometrical multiplicity of  $\lambda_j$  and is zero in most cases.

Stress singularities can be classified into weak and strong singularities between which the classical crack singularity of  $\text{Re}(\lambda_j) - 1 = -0.5$  constitutes the threshold value. [Leguillon and Sanchez-Palencia (1999)] showed that for 2D as well as 3D situations, weak singularities yield a differential energy release rate of  $\mathcal{G} = 0$ . This means that crack configurations evolving towards a structural situation with a weak stress singularity tend to a crack arrest. This, for example, is the case for a crack under mode I loading growing perpendicularly towards an interface with a stiffer material (e.g. [Leguillon and Martin (2012; 2013)]). On the contrary, strong singularities, which also may be called hypersingularities, are characterized by a differential energy release rate which tends to infinity  $\mathcal{G} \rightarrow \infty$ . This means that structural situations under a loading producing a hypersingularity, but also neighboring crack configurations evolving towards such a structural situation, tend to a further, instable crack growth. This, for example, is the case for a crack under mode I loading growing perpendicularly towards an interface with a less stiff material. In conclusion, weak singularities can be classified as less critical and hypersingularities as more critical than the classical crack singularity.

An appropriate method for linear elastostatic problems, which is capable of both resolving the singularities in 3D structural situations and taking into account different anisotropic materials and interfaces, is the Scaled Boundary Finite Element Method (SBFEM). Its suitability for 2D problems has been demonstrated e.g. by [Song (2006)] and for 3D problems e.g. by [Mittelstedt and Becker (2005)] and [Goswami and Becker (2012)]. In a former work, the SBFEM has furthermore been successfully used by [Hell and Becker (2014)] for the analysis of two meeting transverse inter-fiber cracks in a composite laminate. A very similar method based on an FEM eigenanalysis has also proven to be adequate: [Bažant and Estensoro (1979)], [Somaratna and Ting (1986)], [Gharemani (1991)], [Dimitrov et al. (2001)], [Koguchi and da Costa (2010)], [Korepanova et al. (2013)] are only a few authors who employed that method.

In the following, a brief description of the SBFEM is given before it will be employed for the analysis of a delamination meeting a transverse inter-fiber crack.

### The Scaled Boundary Finite Element Method

The Scaled Boundary Finite Element Method (SBFEM) ([Deeks and Wolf (2002)], [Song and Wolf (1997)], [Wolf (2003)]) is a semi-analytical method which combines the advantages of the Boundary Element Method (BEM) and the Finite Element Method (FEM). Comparable to the BEM, only the boundary, or in some cases even only parts of the boundary, need to be discretized. At the same time, the problem of strongly singular integrands, present in the BEM-approach, does not exist in the SBFEM. This is because the SBFEM is based on a variational principle and does not need any fundamental solutions. As a further consequence of this, arbitrary linear elastic material behavior in three dimensions can be taken into account in a Scaled Boundary Finite Element. This makes the SBFEM a powerful tool for a variety of linear elasticity problems, which has lately also been extended to nonlinear analyses (e.g. [Behnke et al. (2014)], [Ooi et al. (2014)]).

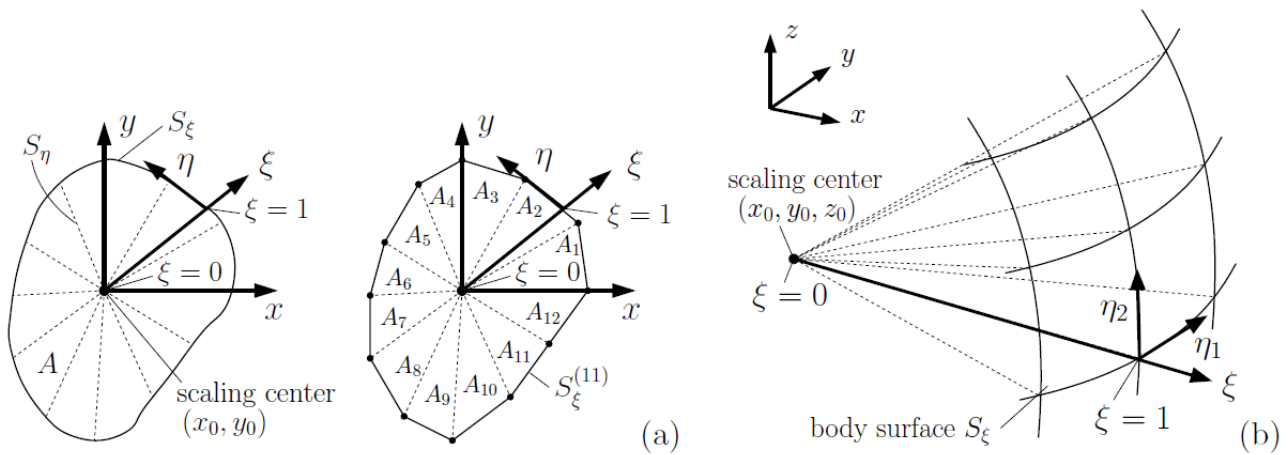


Figure 1: Scaled Boundary Coordinate System and discretization scheme in a (a) 2D and a (b) 3D example.

The *geometrical scalability* is a fundamental requirement of the SBFEM. It requires that the connection of any point on the boundary with the scaling center by a straight line must be possible without any further intersections. Accordingly, a scaled boundary coordinate system is defined having its origin at the scaling center with the coordinates  $(x_0, y_0, z_0)$  in a Cartesian coordinate system. The scaling coordinate  $\xi$  runs from the scaling center  $\xi = 0$  to the boundary  $\xi = 1$ . In case of a 2D problem, a boundary coordinate  $\eta$  runs along the boundary. In case of a 3D problem, two boundary coordinates  $\eta_1$  and  $\eta_2$  describe the surface of the body. Figure 1 illustrates this procedure for the example of an arbitrarily shaped 2D domain and how it can be extended to 3D domains. The Cartesian coordinates are expressed in terms of the scaled boundary coordinates:

$$x = x_0 + \xi \cdot x_\eta(\eta_1, \eta_2) \quad , \quad y = y_0 + \xi \cdot y_\eta(\eta_1, \eta_2) \quad , \quad z = z_0 + \xi \cdot z_\eta(\eta_1, \eta_2). \quad (2)$$

The partial differential operators are calculated by the use of the Jacobian matrix  $\mathbf{J}$ :

$$\begin{bmatrix} \frac{\partial}{\partial \xi} \\ \frac{\partial}{\partial \eta_1} \\ \frac{\partial}{\partial \eta_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial y}{\partial \xi} & \frac{\partial z}{\partial \xi} \\ \frac{\partial x}{\partial \eta_1} & \frac{\partial y}{\partial \eta_1} & \frac{\partial z}{\partial \eta_1} \\ \frac{\partial x}{\partial \eta_2} & \frac{\partial y}{\partial \eta_2} & \frac{\partial z}{\partial \eta_2} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} = \begin{bmatrix} x_\eta & y_\eta & z_\eta \\ x_{\eta_1, \eta_2} & y_{\eta_1, \eta_2} & z_{\eta_1, \eta_2} \\ x_{\eta_2, \eta_1} & y_{\eta_2, \eta_1} & z_{\eta_2, \eta_1} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} = \mathbf{J} \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix}. \quad (3)$$

The notation  $(\cdot)_{,\eta_i}$  is to be read as the partial derivative  $\partial(\cdot)/\partial \eta_i$ . Please note that in this notation the Jacobian  $\mathbf{J}(\eta_1, \eta_2)$  is only a function of the boundary coordinates so that the volumetric differential can be written as

$$dV = \|\mathbf{J}(\eta_1, \eta_2)\| \xi^2 d\xi d\eta_1 d\eta_2. \quad (4)$$

A separation of variables ansatz is made for the displacements and separates dependences of the boundary coordinates  $\eta_1, \eta_2$  from dependences of the scaling coordinate  $\xi$ . However, the resulting equations still cannot be solved analytically so that an approximative approach is needed. Only the boundaries  $S_\xi$  where  $\xi = \text{const}$ , are discretized using isoparametric finite elements and shape functions  $N_j(\eta_1, \eta_2)$ . This sub-divides the body into a number of wedge-shaped domains which are denoted Scaled Boundary Finite Elements (Figure 1). The problem is still considered analytically in the scaling coordinate  $\xi$ . Thus, vector functions  $\vec{u}_j^{(e)}(\xi)$  are introduced for the displacements on rays pointing from the scaling center to the finite element nodes on the boundary. The superscript  $(\cdot)^{(e)}$



denotes the formulation within a Scaled Boundary Finite Element  $e$  where the shape functions are combined to the matrix  $\mathbf{N}(\eta_1, \eta_2)$  and the vector functions  $\bar{u}_j^{(e)}(\xi)$  to  $\bar{U}^{(e)}(\xi)$ . A similar approximative approach is chosen for the virtual displacements:

$$\bar{u}^{(e)}(\xi, \eta_1, \eta_2) = \mathbf{N}(\eta_1, \eta_2) \bar{U}^{(e)}(\xi) \quad , \quad \delta \bar{u}^{(e)}(\xi, \eta_1, \eta_2) = \mathbf{N}(\eta_1, \eta_2) \delta \bar{U}^{(e)}(\xi). \quad (5)$$

Like the FEM, the SBFEM is based on the principle of virtual work

$$\delta W_i = \int_V \boldsymbol{\sigma} : \delta \boldsymbol{\varepsilon} dV = \int_V \bar{\mathbf{f}} \cdot \delta \bar{\mathbf{u}} dV + \int_{S_t} \bar{\mathbf{t}}^* \cdot \delta \bar{\mathbf{u}} dA = \delta W_a \quad (6)$$

where  $\boldsymbol{\sigma}, \delta \boldsymbol{\varepsilon}, \bar{\mathbf{f}}, \delta \bar{\mathbf{u}}, \bar{\mathbf{t}}^*$  are the stress tensor, the virtual strain tensor, body forces, virtual displacements and prescribed boundary tractions respectively.  $\delta W_i$  is the internal and  $\delta W_a$  the external virtual work.  $V$  is the volume of the considered domain and  $S_t$  the part of the domain surface with prescribed traction boundary conditions.

We use the vector notation for stresses and strains and introduce a differential operator  $\mathbf{L}$

$$\mathbf{L}^T = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial x} \\ 0 & 0 & \frac{\partial}{\partial z} & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} & 0 \end{bmatrix} \quad (7)$$

which is used in the equilibrium equations and strain-displacement relations. This differential operator  $\mathbf{L}$  is then transformed into scaled boundary coordinates. To this end,  $\mathbf{L}$  is decomposed into three matrices  $\mathbf{L}_x, \mathbf{L}_y, \mathbf{L}_z$  each associated to one partial differential operator in the Cartesian coordinate system. Then, each partial differential operator can be replaced by its respective counterpart in the scaled boundary coordinate system which yields

$$\mathbf{L} = \mathbf{L}_x \frac{\partial}{\partial x} + \mathbf{L}_y \frac{\partial}{\partial y} + \mathbf{L}_z \frac{\partial}{\partial z} = \mathbf{b}_\xi \frac{\partial}{\partial \xi} + \mathbf{b}_\eta \frac{1}{\xi} \frac{\partial}{\partial \eta_1} + \mathbf{b}_{\eta_2} \frac{1}{\xi} \frac{\partial}{\partial \eta_2}. \quad (8)$$

Employment of the differential operator  $\mathbf{L}$  (7) and the separation of variables ansatz (5) in the strain-displacement relations in vector notation yields:

$$[\boldsymbol{\varepsilon}] = \mathbf{L} \mathbf{N}(\eta_1, \eta_2) \bar{U}(\xi) = \underbrace{\mathbf{b}_\xi \mathbf{N}}_{\mathbf{B}_\xi(\eta_1, \eta_2)} \frac{\partial \bar{U}(\xi)}{\partial \xi} + \underbrace{[\mathbf{b}_{\eta_1} \mathbf{N}_{,\eta_1} + \mathbf{b}_{\eta_2} \mathbf{N}_{,\eta_2}]}_{\mathbf{B}_\eta(\eta_1, \eta_2)} \frac{\bar{U}(\xi)}{\xi}. \quad (9)$$

Additionally, Hooke's law  $[\boldsymbol{\sigma}] = \mathbf{C}[\boldsymbol{\varepsilon}]$  (linear elastic material behavior) with the elasticity matrix  $\mathbf{C}$  is used in the virtual work balance (6). Assuming that the relations for the real quantities are also valid for the virtual ones leads to the virtual work balance in terms of the displacements and virtual displacements in scaled boundary coordinates.

$$\int_V \left[ \delta \bar{U}_{,\xi}^T \mathbf{B}_\xi^T + \delta \bar{U}^T \mathbf{B}_\eta^T \frac{1}{\xi} \right] \mathbf{C} \left[ \mathbf{B}_\xi \bar{U}_{,\xi} + \mathbf{B}_\eta \bar{U} \frac{1}{\xi} \right] dV = \int_V \delta \bar{U}^T \mathbf{N}^T \bar{\mathbf{f}} dV + \int_{S_t} \delta \bar{U}^T \mathbf{N}^T \bar{\mathbf{t}}^* dA \quad (10)$$

Expansion of the product and integration by parts over  $\xi$  in the internal virtual work term  $\delta W_i$  yields an arithmetic expression containing factors which are either dependent of the scaling variable  $\xi$  or of the boundary coordinates  $\eta_1, \eta_2$ . Hence, the integration can be performed separately. The dependence of  $\eta_1$  and  $\eta_2$  actually is one of the introduced shape functions  $\mathbf{N}(\eta_1, \eta_2)$  so that numerical integration over the boundary coordinates can be used. This yields the following matrices within one Scaled Boundary Finite Element  $e$ :

$$\begin{aligned} \mathbf{E}_0^{(e)} &= \int_{S_\xi^{(e)}} \mathbf{B}_\xi^T(\eta_1, \eta_2) \mathbf{C} \mathbf{B}_\xi(\eta_1, \eta_2) \|\mathbf{J}^{(e)}(\eta_1, \eta_2)\| d\eta_1 d\eta_2, \\ \mathbf{E}_1^{(e)} &= \int_{S_\xi^{(e)}} \mathbf{B}_\eta^T(\eta_1, \eta_2) \mathbf{C} \mathbf{B}_\xi(\eta_1, \eta_2) \|\mathbf{J}^{(e)}(\eta_1, \eta_2)\| d\eta_1 d\eta_2, \\ \mathbf{E}_2^{(e)} &= \int_{S_\xi^{(e)}} \mathbf{B}_\eta^T(\eta_1, \eta_2) \mathbf{C} \mathbf{B}_\eta(\eta_1, \eta_2) \|\mathbf{J}^{(e)}(\eta_1, \eta_2)\| d\eta_1 d\eta_2. \end{aligned} \quad (11)$$

In numerical integration, attention must be paid to the Jacobian determinant when the discretized body surface is curved. Then, the numerical integration error needs to be monitored.

The matrices  $\mathbf{E}_0^{(e)}, \mathbf{E}_1^{(e)}, \mathbf{E}_2^{(e)}$  are similarly assembled like in the standard Finite Element Method. This yields the following form of the virtual work balance for the case of a 3D bounded domain  $\xi = 0 \dots 1$ :

$$\begin{aligned} \delta W_i = - \int_{\xi=0}^1 \delta \bar{\mathbf{U}}^T(\xi) \left[ \mathbf{E}_0 \xi^2 \bar{\mathbf{U}}(\xi)_{,\xi\xi} + [2\mathbf{E}_0 + \mathbf{E}_1^T - \mathbf{E}_1] \xi \bar{\mathbf{U}}(\xi)_{,\xi} + [\mathbf{E}_1^T - \mathbf{E}_2] \bar{\mathbf{U}}(\xi) \right] d\xi + \\ \delta \bar{\mathbf{U}}^T(\xi) \left[ \mathbf{E}_0 \xi^2 \bar{\mathbf{U}}(\xi)_{,\xi} + \xi \mathbf{E}_1^T \bar{\mathbf{U}}(\xi) \right]_{\xi=0}^1 = \delta W_a. \end{aligned} \quad (12)$$

This equation contains one integral term and two boundary terms ( $\xi = 0, 1$ ). The boundary terms represent the forces which the continuum exerts on the boundary  $S_\xi$ . Assuming the absence of side face loads, i.e. tractions at the side faces  $S_\eta$  where either  $\eta_1$  or  $\eta_2$  are constant, and that body forces are negligible, the virtual external work can be reduced to  $\delta W_a = \delta \bar{\mathbf{U}}^T(1) \bar{\mathbf{p}}_{\xi=1}$ . The assembled nodal loads vector  $\bar{\mathbf{p}}_{\xi=1}$  gives a contribution to the boundary term in eq. (12). The resulting virtual work balance is valid for arbitrary virtual displacements if and only if the integrand in eq. (12) and the boundary terms are zero each. This yields a homogeneous system of differential equations of Cauchy-Euler type and a linear equation system. A solution fulfilling the homogeneous differential equation system is only approximated over the body surface but analytic in  $\xi$ . The system of linear equations serves for enforcing the boundary conditions on the discretized body surface  $S_\xi$ .

#### *Solution of the homogeneous differential equation system*

By application of a variable transformation  $t = \ln \xi$ , a differential equation system of Cauchy-Euler type can be transformed into an ordinary one with constant coefficients. The introduction of the vector function  $\bar{\mathbf{V}}(t) = \bar{\mathbf{U}}(t)_{,t}$  then allows the further transformation of the differential equation system of second order into one of first order but at the cost of doubling the number of degrees of freedom.

$$\underbrace{\begin{bmatrix} \bar{\mathbf{V}}(t)_{,t} \\ \bar{\mathbf{U}}(t)_{,t} \end{bmatrix}}_{\bar{\mathbf{W}}(t)_{,t}} = \underbrace{\begin{bmatrix} -\mathbf{E}_0^{-1} [\mathbf{E}_1^T - \mathbf{E}_1] - \mathbf{I} & -\mathbf{E}_0^{-1} [\mathbf{E}_1^T - \mathbf{E}_2] \\ \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} \bar{\mathbf{V}}(t) \\ \bar{\mathbf{U}}(t) \end{bmatrix}}_{\bar{\mathbf{W}}(t)} \quad (13)$$

The fundamental solution solving this type of differential equations converts the differential equation system into an eigenvalue problem.

$$\bar{W}_j(t) = \bar{\Phi}_j e^{\lambda_j t} \Rightarrow \lambda_j \bar{\Phi}_j = \mathbf{\kappa} \bar{\Phi}_j \quad (14)$$

By application of established numerical algorithms, even non-symmetric and rather large eigenvalue problems can be solved. Unfortunately, these algorithms suffer from numerical errors, which are not negligible any more if the magnitudes of neighboring eigenvalues move closer together. Nevertheless, small eigenvalues close to zero and associated eigenvectors generally are of good quality and converge appropriately with a discretization refinement.

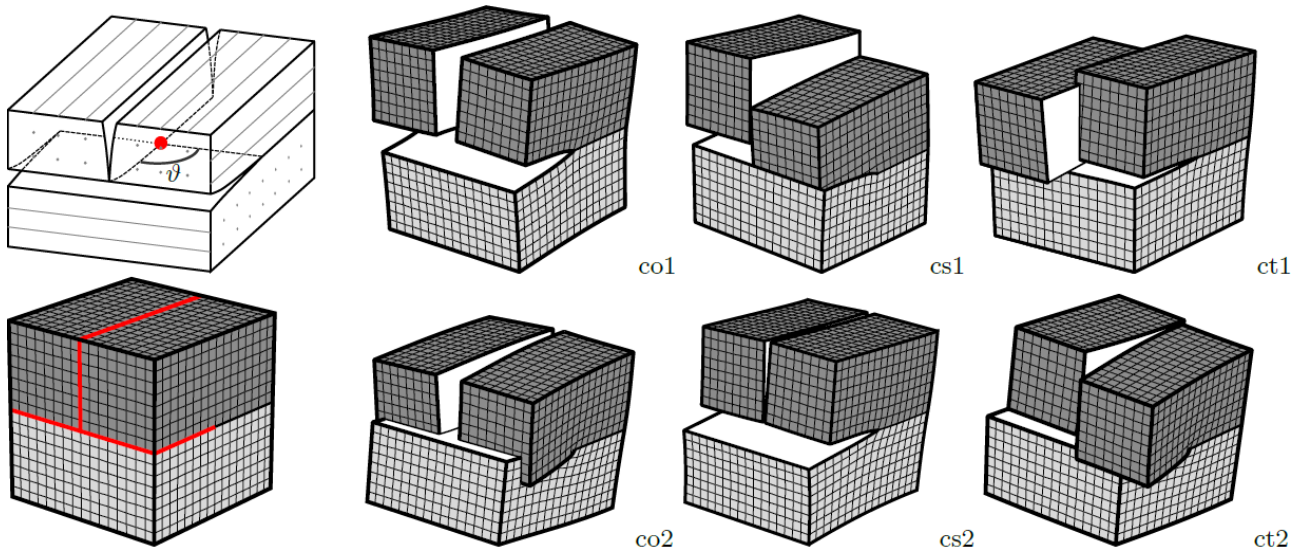
The backtransformation of the fundamental solution ( $t \rightarrow \xi: \bar{W}_j(\xi) = \bar{\Phi}_j \xi^{\lambda_j}$ ) reveals that the eigenvectors can be interpreted as deformation modes and that the eigenvalues are their associated decay ( $\text{Re}(\lambda_j) < 0$ ) or growth rates ( $\text{Re}(\lambda_j) > 0$ ). In fact, only the lower half of the eigenvector  $\bar{\Phi}_{uj}$  represents a deformation mode while the upper half is simply given by  $\bar{\Phi}_{vj} = \bar{\Phi}_{uj} \lambda_j$ . The eigenvalue spectrum in the 3D case is symmetric to  $-0.5$ , which also marks the value of a bounded domain for which the strain energy density tends to infinity. Complex eigenvalues always appear as conjugate pairs  $\lambda = \alpha \pm i\beta$ . In the case of geometric multiplicity, i.e. for a given number of equal eigenvalues, not the same number of linearly independent eigenvectors exists, additional generalized eigenvectors have to be generated resulting in logarithmic deformation modes to complete the solution. Hence, the general solution of the differential equation system with  $N$  different eigenpairs of geometric multiplicity  $n_j$  is

$$\bar{W}(\xi) = \sum_{j=1}^N \sum_{k=0}^{n_j-1} \left[ \text{Re}(\bar{\Phi}_{jk} e^{i\beta_j \ln \xi}) c_{jk1} + \text{Im}(\bar{\Phi}_{jk} e^{i\beta_j \ln \xi}) c_{jk2} \right] \xi^{\alpha_j} (\ln \xi)^k. \quad (15)$$

Here,  $c_{jk1}$  and  $c_{jk2}$  are free constants. In sum, they are of the number of twice the number of degrees of freedom of the approximated boundary value problem. These free constants are determined from regularity and boundary conditions.

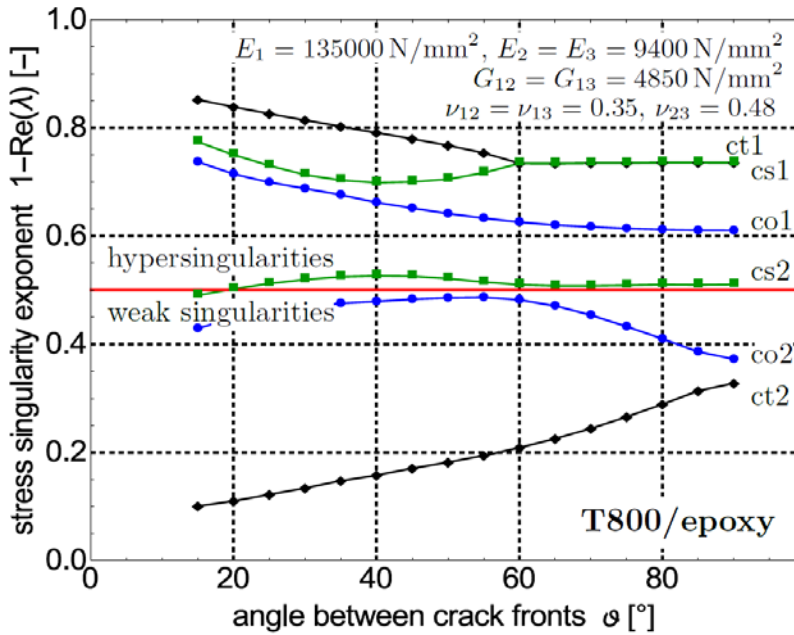
### Delamination Meeting a Transverse Inter-Fiber Crack in a Composite Laminate

The 3D structural situation of a delamination meeting a transverse inter-fiber crack which runs through the complete ply can more generally be described as two plane cracks with straight crack fronts meeting each other at an interface. But in contrast to the structural situation studied in [Hell and Becker (2014)], where the mode I crack growth directions of the two cracks point towards each other representing the situation of two meeting transverse inter-fiber cracks, the mode I crack growth directions here are assumed to be perpendicular to each other. This also implies that the transverse inter-fiber crack can, in a way, be seen as dividing the delamination crack in two parts. Configurations with angles between the crack fronts  $15^\circ < \vartheta < 90^\circ$  and concurrent ply lay-ups  $[(90^\circ - \vartheta)/90^\circ]$  are considered (cf. Figure 2). The stress singularity exponents  $\text{Re}(\lambda_j) - 1$  and their associated deformation modes  $\bar{\Phi}_{uj}(\eta_1, \eta_2)$  are calculated by means of the SBFEM using a spherical boundary mesh for a minimum numerical effort. Between 931 and 1406 bilinear isoparametric SBFEMs for the angles  $\vartheta = 60^\circ$  and  $\vartheta = 35^\circ$ , respectively, are used for the boundary mesh, which is in each case appropriately refined at the crack fronts. The results for the absolute values of the stress singularity exponents  $1 - \text{Re}(\lambda_j)$  are presented in Figure 3 and Figure 4.



**Figure 2: Delamination crack meeting transverse inter-fiber crack in a fiber-reinforced composite laminate with angle  $\vartheta$  between the crack fronts and used boundary mesh. 6 singular deformation modes for  $\vartheta=90^\circ$ : crack opening (*co1/2*), shearing (*cs1/2*) and twisting (*ct1/2*).**

Figure 3 gives the results for a T800/epoxy (a carbon fiber reinforced plastic) laminate revealing 6 stress singularities. The 6 associated deformation modes are shown in Figure 2 for the example of  $\vartheta = 90^\circ$ . For an easier identification, they are presented for a cubic boundary mesh where the cracks implemented as double nodes are marked in red. The upper half (dark shading) represents the upper ply with a  $0^\circ$ -orientation and contains the correspondingly aligned transverse inter-fiber crack. The lower half (light shading) represents the lower ply with  $90^\circ$ -orientation. The delamination crack is located at the interface, which obviously coincides with the plane defined by the present crack fronts. The deformation mode *co1* corresponds to a simultaneous crack opening of the delamination and the inter-fiber crack. On the other hand, deformation mode *co2* corresponds to a crack opening of only one of the cracks and a crack closing of the other. The deformation mode *cs1* corresponds to a crack shearing of the inter-fiber crack, which implies a simultaneous crack opening respectively closing of the delamination crack faces. A crack shearing of the delamination crack can be identified for deformation mode *cs2*. The deformation mode *ct1* corresponds to a crack twisting of the inter-fiber crack, which implies a simultaneous counter-directional crack shearing of the delamination crack. A kind of crack twisting of the delamination crack can be identified for the deformation mode *ct2*. Although these deformation modes actually only correspond to the particular case of  $\vartheta = 90^\circ$ , the wording is kept for all configurations studied. Other crack configurations with different angles  $\vartheta$  between the crack fronts produce different deformation modes for which the individual crack deformations cannot be assigned equally clearly to the classical single-crack deformation modes any more. Please note that deformation mode *cs2* is an exception and constitutes the only deformation mode remaining widely unchanged for all configurations presented in this work. At the same time, *cs2* is the deformation mode most closely related to a pure single-crack deformation mode, namely mode II of the delamination crack. This makes, indeed, perfectly sense as a corresponding mode II delamination crack loading does not require an exchange of forces through the inter-fiber crack faces. In fact, its corresponding stress singularity exponent remains close to the classical crack singularity exponent of  $1 - \text{Re}(\lambda) = 0.5$  for all angles  $\vartheta$  and material combinations considered.

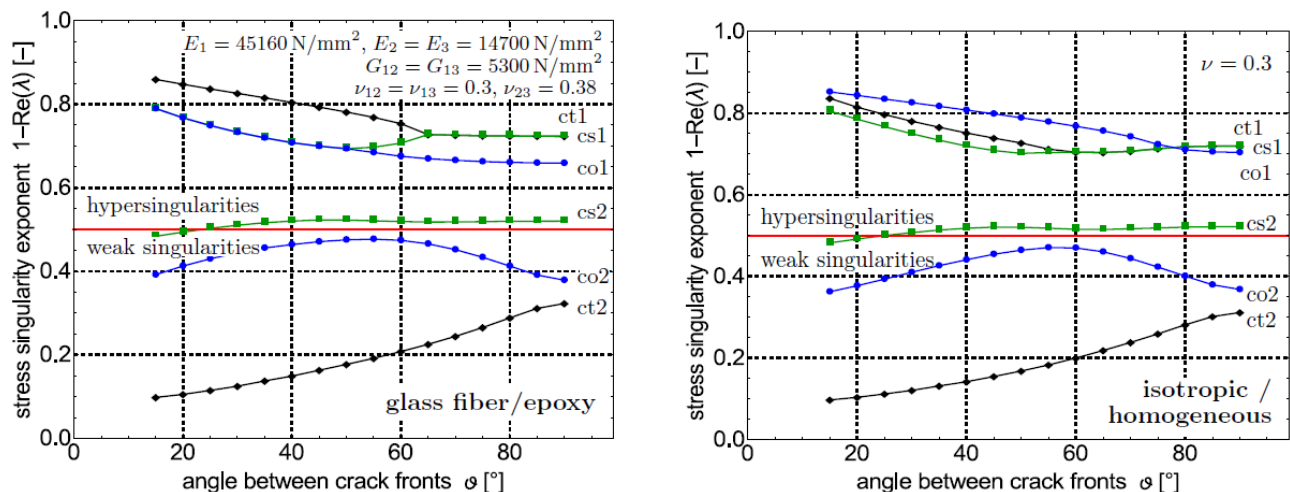


**Figure 3: Stress singularity exponents for a delamination crack meeting a transverse inter-fiber crack in a T800/epoxy CFRP laminate with ply-layups [(90°-9)/90°].**

The classical crack singularity value, marked by a red line in Figure 3 and Figure 4, separates weak singularities from strong singularities, which we also call hypersingularities. It is again emphasized that weak singularities can be classified as less critical than the classical crack singularities, as they yield a differential energy release rate of  $\mathcal{G} = 0$  and, therefore, favor a crack arrest. On the other hand, hypersingularities yield a differential energy release rate of  $\mathcal{G} \rightarrow \infty$ , which obviously favors crack growth. For all crack configurations studied, only two stress singularities are weak, which are the ones associated to deformation modes *co2* and *ct2*. In

contrast, always three hypersingularities are present, namely those associated to deformation modes *co1*, *cs1* and *ct1*. Moreover, it can be stated that strong singularities become even stronger with decreasing angle  $\vartheta$  while the weak singularities decline or at least remain weak. Please also note that the real part of the stress singularity exponents associated to the deformation modes *cs1* and *ct1* are the same for angles  $\vartheta \gtrsim 60^\circ$ . This is because they constitute a pair of complex conjugated stress singularity exponents which makes them occur strongly interconnected (cf. eq. (15)). However, all other stress singularities found for the considered T800/epoxy laminate configurations are not complex.

For comparison, configurations with less anisotropic ply materials were studied: a delamination meeting a transverse inter-fiber crack in a typical [(90° - 9)/90°] glass fiber reinforced plastics (GFRP) laminate and the same geometrical setup but in a homogeneous isotropic body (Figure 4). The material data are given in the figures while the stress singularities present in a homogeneous isotropic body only depend on Poisson's ratio but not on Young's modulus. Here again, all



**Figure 4: Stress singularity exponents for two meeting cracks with perpendicular mode I crack growth directions in a (a) typical GFRP laminate with ply-layups [(90°-9)/90°] and a (b) homogeneous isotropic body.**

configurations considered go along with 3 hypersingularities, 2 weak singularities and one singularity approximately matching the classical crack singularity. From Figure 3 and Figure 4 it can be seen that the material properties mainly affect deformation mode *col*. Its dependence on the elastic contrast between the upper and the lower ply is expected as it also plays a major role for the stress singularity exponent of the crack opening mode of a single crack impinging an interface (e.g. [Bogy (1971); Ting and Hoang (1984)]). The other deformation modes seem to be hardly or only moderately affected by the material properties. Finally, complex singularities can also be found for GFRP laminates and homogeneous isotropic materials. The considered GFRP laminate exhibits complex singularities for angles between the crack fronts of  $50^\circ \gtrsim \vartheta \gtrsim 65^\circ$  and the homogeneous isotropic configuration with a Poisson's ratio of  $\nu = 0.3$  for angles  $\vartheta \gtrsim 60^\circ$ . Although the magnitude of the imaginary part of the stress singularity exponent is always rather small with  $|\text{Im}(\lambda) - 1| < 0.045$ , it is remarkable that, here, a complex singularity can also occur in the homogeneous isotropic case.

## Conclusion

The SBFEM has been used to solve boundary value problems of linear elasticity which contain singular points. Even 3D anisotropic structural situations involving interfaces are treated accurately and efficiently. Such a structural situation is the one of a delamination crack meeting a transverse inter-fiber crack, which has been treated in this contribution. It has been shown that this is a highly critical structural situation involving up to 3 hypersingularities. This strongly motivates a further investigation. It is worth pointing out that the hypersingularities found can be complex singularities - even in the simplified homogeneous isotropic case of this geometrical setup.

## Acknowledgment

The authors highly appreciate the financial support of the German Research Foundation (DFG), project BE1090/35-1.

## References

- Bažant, Z.P. and Estenssoro, L.F. (1979) Surface singularity and crack propagation, *International Journal of Solids and Structures* **15**, 405-426.
- Behnke, R., Mundil, M., Birk, C. and Kaliske, M. (2014) A physically and geometrically nonlinear scaled-boundary-based finite element formulation for fracture in elastomers. *International Journal for Numerical Methods in Engineering* **99**, 966-999.
- Bogy, D.B. (1971) On the plane elastostatic problem of a loaded crack terminating at a material interface. *Journal of Applied Mechanics* **38**, 911-918.
- Deeks, A.J. and Wolf, J. P. (2002) A virtual work derivation of the scaled boundary finite-element method for elastostatics. *Computational Mechanics* **28**, 489-504.
- Dimitrov, A., Andrä, H. and Schnack, E. (2001) Efficient computation of order and mode of corner singularities in 3d-elasticity. *International Journal for Numerical Methods in Engineering* **52**, 805-827.
- Ghahremani, F. (1991) A numerical variational method for extracting 3D singularities. *International Journal of Solids and Structures* **27**, 1371-1386.
- Goswami, S. and Becker, W. (2012) Computation of 3-d stress singularities for multiple cracks and crack intersections by the scaled boundary finite element method. *International Journal of Fracture* **175**, 13-25.
- Hell, S. and Becker, W. (2014) Hypersingularities in three-dimensional crack configurations in composite laminates. *Proceedings in Applied Mathematics and Mechanics* **14**, 157-158.
- Koguchi, H. and da Costa, J.A. (2010) Analysis of the stress singularity field at a vertex in 3D-bonded structures having a slanted side surface. *International Journal of Solids and Structures* **47**, 3131-3140.
- Korepanova, T.O., Matveenko, V.P. and Sevodina, N.V. (2013) Numerical analysis of stress singularity at singular points of three-dimensional elastic bodies. *Acta Mechanica* **224**, 2045-2063.
- Leguillon, D. and Sanchez-Palencia, E. (1999) On 3d cracks intersecting a free surface in laminated composites. *International Journal of Fracture* **99**, 25-40.

- Leguillon, D. and Martin, E. (2012) Crack nucleation at stress concentration points in composite materials - Application to the crack deflection by an interface. *Mathematical Methods and Models in Composites, Computational and Experimental Methods in Structures* **5**, 401-424.
- Leguillon, D. and Martin, E. (2013) The strengthening effect caused by an elastic contrast part i: the bimaterial case. *International Journal of Fracture* **179**, 157-167.
- Mittelstedt, C. and Becker, W. (2005) Semi-analytical computation of 3d stress singularities in linear elasticity. *Communications in Numerical Methods in Engineering* **21**, 247-257.
- Mittelstedt, C. and Becker, W. (2006) Efficient computation of order and mode of three-dimensional stress singularities in linear elasticity by the boundary finite element method. *International Journal of Solids and Structures* **43**, 2868-2903.
- Ooi, E.T., Song, C. and Tin-Loi, F. (2014) A scaled boundary polygon formulation for elasto-plastic analyses. *Computer Methods in Applied Mechanics and Engineering* **268**, 905-937.
- Somaratna, N. and Ting, T.C.T. (1986) Three-dimensional stress singularities in anisotropic materials and composites. *International Journal of Engineering Science* **24**, 1115 -1134.
- Song, C. (2006) Analysis of singular stress fields at multi-material corners under thermal loading. *International Journal for Numerical Methods in Engineering* **65**, 620-652.
- Song, C. and Wolf, J.P. (1997) The scaled boundary finite-element method alias consistent infinitesimal finite-element cell method for elastodynamics. *Computer Methods in Applied Mechanics and Engineering* **147**, 329-355.
- Ting, T.C.T. and Hoang, P.H. (1984) Singularities at the tip of a crack normal to the interface of an anisotropic layered composite. *International Journal of Solids and Structures* **20**, 439-454.
- Wolf, J. P. (2003) *The scaled boundary finite element method*. John Wiley & Sons, Chichester, UK.

# **An investigation of Nano-particle Deposition in Cylindrical Tubes under Laminar condition using Lagrangian transport model**

†\*M. Babaei<sup>1</sup>, P. Talebizadeh<sup>2,3</sup>, K. Inthavong<sup>4</sup>, G. Ahmadi<sup>5</sup>, Z. Ristovski<sup>3</sup>, H. Rahimzadeh<sup>2</sup>, R. Brown<sup>3</sup>

<sup>1</sup>Petroleum and Gas Engineering Division, School of Computing, Science and Engineering, University of Salford, Manchester, United Kingdom

<sup>2</sup>Department of Mechanical Engineering, Amirkabir University of Technology, Iran

<sup>3</sup>Biofuel Engine Research Facility, Queensland University of Technology, Australia

<sup>4</sup>School of Aerospace, Mechanical and Manufacturing Engineering, RMIT University, Australia

<sup>5</sup>Department of Mechanical and Aeronautical Engineering, Clarkson University, United states

\*Presenting author: m.babaie@salford.ac.uk

†Corresponding author: m.babaie@salford.ac.uk

## **Abstract**

Aerosol deposition in cylindrical tubes is a subject of interest to researchers and engineers in many applications of aerosol physics and metrology. Investigation of nano-particles in different aspects such as lungs, upper airways, batteries and vehicle exhaust gases is vital due to the smaller size, adverse health effect and higher trouble for trapping than the micro-particles. The Lagrangian particle tracking provides an effective method for simulating the deposition of nano-particles as well as micro-particles as it accounts for the particle inertia effect as well as the Brownian excitation. However, using the Lagrangian approach for simulating ultrafine particles has been limited due to computational cost and numerical difficulties.

In this paper, the deposition of nano-particles in cylindrical tubes under laminar condition is studied using the Lagrangian particle tracking method. The commercial Fluent software is used to simulate the fluid flow in the pipes and to study the deposition and dispersion of nano-particles. Different particle diameters as well as different flow rates are examined. The point analysis in a uniform flow is performed for validating the Brownian motion. The results show good agreement between the calculated deposition efficiency and the analytic correlations in the literature. Furthermore, for the nano-particles with the diameter more than 40 nm, the calculated deposition efficiency by the Lagrangian method is less than the analytic correlations based on Eulerian method due to statistical error or the inertia effect.

**Keywords:** Nano-particle, Two phase flow, Deposition, Lagrangian particle tracking method, Cylindrical tubes

## **Introduction**

Aerosol deposition in cylindrical tubes is a subject of interest by researchers and engineers in many applications of aerosol physics and metrology. Studies on deposition efficiency in lungs, upper airways, batteries and vehicle exhaust gases are some examples of particle deposition in cylindrical tubes. In studying particle deposition in cylindrical tubes, deposition of nano-particles or ultrafine particles (diameter < 100 nm) are more important among all range of particle diameters. For example, in oral airways, most of particles with the diameter above 1 μm are deposited in the nose and respiratory organ walls; however, nano-particles can pass to the lung airways and compromise the human health (Malet, Alloul et al. 2000).

In the literature, numerous studies have developed theoretical expressions for the particle deposition through a smooth tube in laminar flow. Ingham in 1975 and 1991



developed a model for calculating the deposition efficiency in a fully developed flow in cylindrical tube and in the entrance region of a cylindrical tube, respectively (Ingham 1975, Ingham 1991). Cohen and Asgharian in 1990 developed an empirical expression for the deposition efficiency of particles larger than 10nm (Cohen and Asgharian 1990). Most of these studies are used the mass diffusion equation governing the concentration of particles to find an analytic correlation for the deposition efficiency. Therefore, these models often ignore particle inertia effect for aerosols smaller than 200 nm.

In the absence of inertial effects, a highly efficient Eulerian transport model can be applied that treats the particle phase as a dilute chemical species (Longest and Xi 2007). However, the effects of inertia have not been fully quantified for aerosols in the fine and ultrafine ranges (Longest and Xi 2007). Direct Lagrangian particle tracking may provide an effective method for simulating the deposition of nano-particles which can account particle inertia effect. Furthermore, it has the ability to resolve additional body forces that are applicable to each individual particle (Longest and Xi 2007, Tu, Inthavong et al. 2012).

In this study, direct Lagrangian particle tracking method is used to calculate the deposition of nano-particles in cylindrical tubes under the fully developed laminar condition flow. The deposition efficiency is calculated for different flow rates, different tube lengths and different particle diameters.

### Mathematical modeling

In this paper, the commercial Ansys-Fluent software is used for solving the governing particle equation of motion. For the fluid flow, the exact solution for laminar pipe flow is used for the fluid velocity as a profile at the inlet of the tube and then the fully developed laminar flow is simulated for the entire cylinder. The exact solution for the laminar flow in the cylinder is a parabolic profile for the velocity which is defined as (Longest and Xi 2007):

$$u(r) = 2u_{in} \left(1 - \frac{r^2}{R^2}\right) \quad (1)$$

where  $R$  is the pipe radius and  $u_{in}$  is the inlet velocity.

Then, one-way coupled trajectories of mono-disperse submicron particles ranging in diameter from 5 nm to 100 nm have been calculated based on Lagrangian method by integrating an appropriate form of the particle trajectory equation. In this range of particle diameter, transport of nano-particles is mainly attributed to the Brownian force, therefore, the appropriate equations for spherical particle motion can be expressed as (Wen, Inthavong et al. 2008, Inthavong, Tu et al. 2009):

$$\frac{du_i^p}{dt} = \frac{18\mu}{d_p^2 \rho_p C_c} (u_i^s - u_i^p) + F_{Brownian} \quad (2)$$

where  $u_i^p$  and  $u_i^s$  are the components of the particle and local fluid velocity, respectively.  $\mu$  is the fluid viscosity and  $\rho_p$  is the particle density.  $C_c$  is the Cunningham correction factor to Stokes' drag law which can be calculated as (Zamankhan, Ahmadi et al. 2006, Inthavong, Zhang et al. 2011):

$$C_c = 1 + \frac{2\lambda}{d_p} (1.257 + 0.4e^{-1.1d_p/2\lambda}) \quad (3)$$

where  $\lambda$  is the mean free path of air which is equal to 65 nm.

The amplitude of the Brownian force is defined as (Wang, Inthavong et al. 2009):

$$F_{Brownian} = \zeta \sqrt{\frac{\pi S_0}{\Delta t}} \quad (4)$$

where  $\zeta$  is a zero-mean, unit-variance independent Gaussian random number,  $\Delta t$  is the time-step for particle integration and  $S_0$  is a spectral intensity function defined as (Tian and Ahmadi 2007):

$$S_0 = \frac{216\nu k_B T}{\pi^2 \rho_g d_p^5 \left(\frac{\rho_p}{\rho_g}\right)^2 C_c} \quad (5)$$

$T$  is the absolute temperature of the fluid,  $\nu$  is the kinematic viscosity,  $k_B$  is the Boltzmann constant and  $\rho_g$  is the gas density.

Therefore, the Brwonian force can be obtained as (Inthavong, Tu et al. 2009):

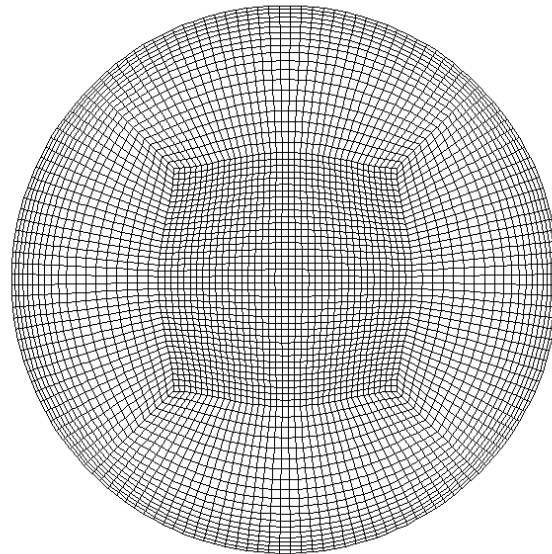
$$F_{Brownian} = \frac{\zeta}{m_d} \sqrt{\frac{1}{\tilde{D}} \frac{2k_B T^2}{\Delta t}} \quad (6)$$

where  $m_d$  is the mass of the particle and  $\tilde{D}$  is the diffusion coefficient which is determined as (Tu, Inthavong et al. 2012):

$$\tilde{D} = \frac{k_B T C_c}{3\pi\mu d_p} \quad (7)$$

### Geometry and mesh structure

A straight pipe is created in Gambit software in this paper as the studied geometry. The diameter of the pipe is 0.45 cm (Longest and Xi 2007). Two different lengths are considered: 3, 5 cm. The structure of the mesh is an important issue for simulating particle deposition. Fig. 1 displays the created mesh at the inlet of the tube.



**Figure 1. Mesh structure on the pipe inlet**

As shown in the figure, dense mesh near the wall is necessary to determine the deposition efficiency correctly (Longest and Vinchurkar 2007). Note that the total number of nodes is almost 900,000.

### Boundary conditions

As mentioned before, the deposition efficiency is calculated in a fully developed flow in this paper. Boundary conditions for the particles were set up as a circular particle release entrained in the flow field. Particles were released from 0.01m from the inlet to prevent any spurious data exiting the inlet upon immediate release. In addition, the radial distance at which a particle was located was not less than 0.1 mm away from

the wall to eliminate artificial immediate deposition on the walls (Wen, Inthavong et al. 2008). Note that 70000 particles are created randomly in order to have the deposition efficiency independent from the particle number. Furthermore, 10 integration steps for Brownian motion is considered as the time step size (Wen, Inthavong et al. 2008). Note that the considered flow rates are 1 and 2 lit/min.

## Results and discussion

Deposition results for the Brownian motion models are first verified by comparing the results with the Ingham equation which proposed an analytic deposition efficiency correlation based on the diffusion parameter. This correlation is defined as (Ingham 1975):

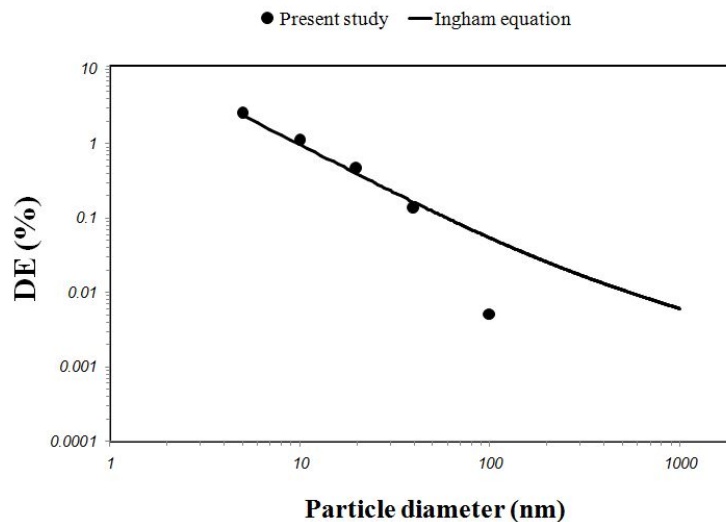
$$DE = 1 - \left( 0.819e^{-14.63\Delta} + 0.0976e^{-89.22\Delta} + 0.0325e^{-228\Delta} + 0.0509e^{-125.9\Delta^{2/3}} \right) \quad (8)$$

where  $\Delta$  is the dimensionless diffusion parameter defined as (Ingham 1975):

$$\Delta = \frac{\tilde{D}L_{pipe}}{4U_{in}R^2} \quad (9)$$

where  $L_{pipe}$  is the pipe length.

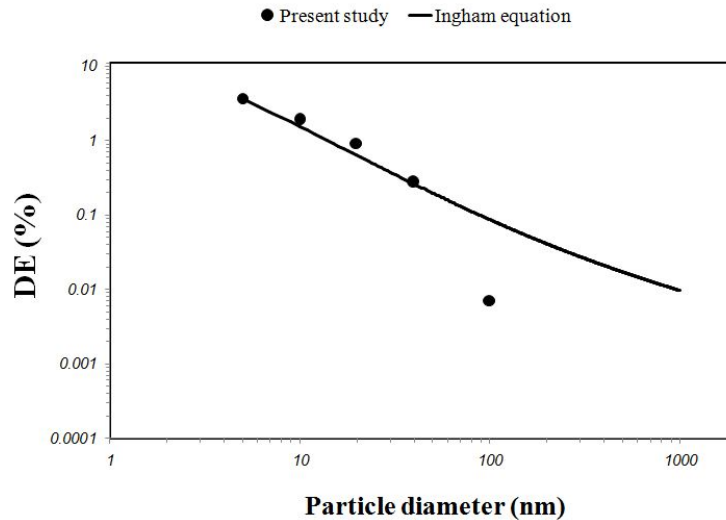
Fig. 2 displays the deposition efficiency calculated in this paper and by Ingham equation for a pipe with the length of 2 cm and the constant inlet velocity of 1 m/s.



**Figure 2. The deposition efficiency for the cylinder with the length of 2cm and the constant inlet velocity of 1m/s**

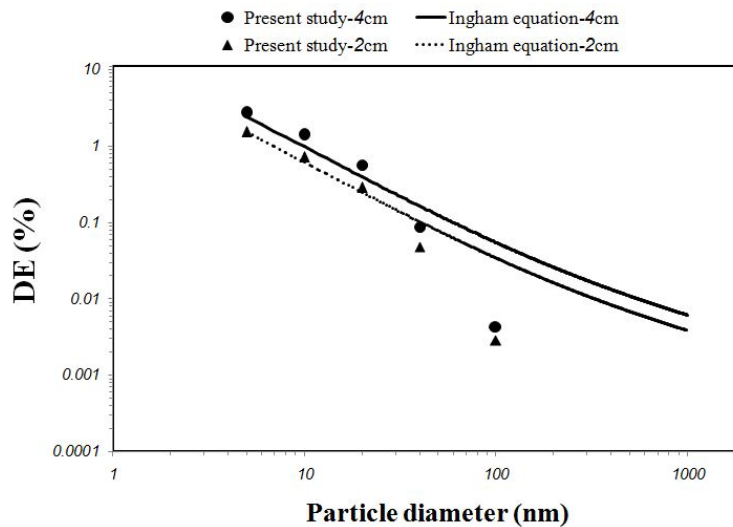
As shown, the results have a good agreement with the Ingham equation. It should be noted that for large particles (40 and 100 nm), due to the inertia effect, the deposition efficiency decreases especially for 100 nm particles (Longest and Xi 2007). As mentioned before, the inertia effect cannot be considered in the Eulerian method or mass diffusion equation and this is another advantage of direct Lagrangian method (Longest and Xi 2007).

Fig. 3 displays the deposition efficiency for both present study and Ingham equation for a 4 cm cylinder with the constant inlet velocity of 1 m/s. As shown, again for 100nm particles, due to the inertia effect, the calculated deposition efficiency is less than the value calculated from Ingham equation.



**Figure 3. The deposition efficiency for the cylinder with the length of 4cm and the constant inlet velocity of 1m/s**

Fig. 4 shows the calculated deposition efficiency in this paper in compare with the Ingham equation for different particle diameter for both tube lengths of 2 and 4 cm for the constant inlet velocity of 2 m/s. As shown, by increasing the inlet velocity, the inertia effect is more effective and for 40 nm particles, the difference between the calculated deposition efficiency and Ingham equation can be seen (Longest and Xi 2007).



**Figure 4. The deposition efficiency for cylinders with the lengths of 2 cm and 4 cm and the constant inlet velocity of 2 m/s**

### Conclusion

In this paper, the direct Lagrangian particle tracking method was employed to determine the deposition efficiency of nano-particles in cylindrical tubes. Different particle diameters, different flow rates and various pipe lengths were examined. The results showed a good agreement with the existed analytic correlations in the literature. Furthermore, by increasing the particles diameter and inlet velocity, due to the inertia effect, a difference in the calculated deposition efficiency by the Lagrangian method and by the analytic correlation based on diffusion can be seen.

## References

- Cohen, B. S. and B. Asgharian (1990) Deposition of ultrafine particles in the upper airways: An empirical analysis, *Journal of Aerosol Science* **21**(6): 789-797.
- Ingham, D. B. (1975) Diffusion of aerosols from a stream flowing through a cylindrical tube, *Journal of Aerosol Science* **6**(2): 125-132.
- Ingham, D. B. (1991) Diffusion of aerosols in the entrance region of a smooth cylindrical pipe, *Journal of Aerosol Science* **22**(3): 253-257.
- Inthavong, K., J. Tu and G. Ahmadi (2009) Computational Modelling of Gas-Particle Flows with Different Particle Morphology in the Human Nasal Cavity, *The Journal of Computational Multiphase Flows* **1**(1): 57-82.
- Inthavong, K., K. Zhang and J. Tu (2011) Numerical modelling of nanoparticle deposition in the nasal cavity and the tracheobronchial airway, *Computer Methods in Biomechanics and Biomedical Engineering* **14**(7): 633-643.
- Longest, P. W. and S. Vinchurkar (2007) Effects of mesh style and grid convergence on particle deposition in bifurcating airway models with comparisons to experimental data, *Medical Engineering & Physics* **29**(3): 350-366.
- Longest, P. W. and J. Xi (2007), Computational investigation of particle inertia effects on submicron aerosol deposition in the respiratory tract, *Journal of Aerosol Science* **38**(1): 111-130.
- Longest, P. W. and J. Xi (2007) Effectiveness of Direct Lagrangian Tracking Models for Simulating Nanoparticle Deposition in the Upper Airways, *Aerosol Science and Technology* **41**(4): 380-397.
- Malet, J., L. Alloul, N. Michielsen, D. Boulaud and A. Renoux (2000) Deposition of nanosized particles in cylindrical tubes under laminar and turbulent flow conditions, *Journal of Aerosol Science* **31**(3): 335-348.
- Tian, L. and G. Ahmadi (2007) Particle deposition in turbulent duct flows—comparisons of different model predictions, *Journal of Aerosol Science* **38**(4): 377-397.
- Tu, J., K. Inthavong and G. Ahmadi (2012) *Computational Fluid and Particle Dynamics in the Human Respiratory System*, Springer.
- Wang, S. M., K. Inthavong, J. Wen, J. Y. Tu and C. L. Xue (2009) Comparison of micron- and nanoparticle deposition patterns in a realistic human nasal cavity, *Respiratory Physiology & Neurobiology* **166**(3): 142-151.
- Wen, J., K. Inthavong, J. Tu and S. Wang (2008) Numerical simulations for detailed airflow dynamics in a human nasal cavity, *Respiratory Physiology & Neurobiology* **161**(2): 125-135.
- Zamankhan, P., G. Ahmadi, Z. Wang, P. K. Hopke, Y.-S. Cheng, W. C. Su and D. Leonard (2006), Airflow and Deposition of Nano-Particles in a Human Nasal Cavity, *Aerosol Science and Technology* **40**(6): 463-476.

## Parametric reliability sensitivity analysis using failure probability ratio function

†Wei Pengfei

School of Mechanics, Civil Engineering and Architecture, Northwestern Polytechnical University, Xi'an, China

†Corresponding author: [wpf0414@163.com](mailto:wpf0414@163.com); [pengfeiwei@nwpu.edu.cn](mailto:pengfeiwei@nwpu.edu.cn)

### Abstract

Reducing the failure probability is an important task in the design of engineering structures. In this paper, a reliability sensitivity analysis technique, called failure probability ratio function, is firstly developed for providing the analysts quantitative information on failure probability reduction while one or a set of distribution parameters of model inputs are changed. The proposed failure probability ratio function can be especially useful for failure probability reduction, reliability-based optimization and reduction of the epistemic uncertainty of parameters. The Monte Carlo simulation (MCS), Importance sampling (IS) and Truncated Importance Sampling (TIS) procedures, which need only a set of samples for implementing them, are introduced for efficiently computing the proposed sensitivity indices. A numerical example is introduced for illustrating the engineering significance of the proposed sensitivity indices and verifying the efficiency and accuracy of the MCS, IS and TIS procedures.

**Keywords:** Sensitivity analysis; Failure probability ratio function; Importance sampling

### 1. Introduction

In design of engineering structure, the analysts often build a lot of computational models (e.g. finite element model, FEM) for simulating the behavior of real structures. However, due the extensively existing uncertainty (aleatory or epistemic), the performance of the structure turn out to be unsteady, which often prevents the analysts from understanding the behavior of structure. In the probabilistic framework, the model inputs are often treated as random variables and represented by probability density function (PDF). The distribution parameters are either fixed at constant value (only aleatory uncertainty is presented) or characterized by confidence interval (due to epistemic uncertainty). Under these assumptions, the analysts' two main concerns are reliability analysis and safety improvement.

The reliability analysis aims at assessing the failure probability of existing structure. During the past several decades, many methods have been developed by researchers for this purpose such as the Monte Carlo Simulation (MCS), First-order Second-moment (FOSM) method [Hasofer and Lind (1974)], Importance Sampling (IS) [Au and Beck (2002); Harbitz(1986); Melchers(1989); Melchers(1990);], Subset Simulation (SS) [Au and Beck (2001)], Line Sampling (LS) [Schuëller et al. (2004)], directional sampling [Ditlevsen et al. (1990)] response surface method [Faravelli (1989)]. However, analysts still find it difficult to employ these methods especially when the epistemic uncertainty is presented in model inputs.

Safety improvement focuses on reducing the failure probability of existing structure via selecting optimal values for distribution parameters (if possible). This is often

dealt with in two ways: reliability-based optimization (RBO) and sensitivity analysis. Up to now, three groups of methods are available for RBO: double-loop method [Enevoldsen and Sorensen (1994); Tu et al. (1999)], single-loop method [Chen et al. (1997); Kuschel and Rackwitz (1997)], decoupling method [Au (2005); Royset et al. (2001); Zou and Mahadevan (2006)]. Compared with the methods for reliability analysis, these RBO methods are generally more time-consuming. One can refer to Schueller and Jensen (2008) and Valdebenito and Schueller (2010) for overviews of the RBO methods. Sensitivity analysis techniques for safety improvement can be divided into three groups: local, global and regional sensitivity analysis.

Local reliability sensitivity analysis indices are generally defined as the partial derivatives of failure probability or reliability index with respect to distribution parameters [Bjerager and Krenk (1989); Lu et al. (2008); Melchers and Ahammed (2004); Wu and Mohanty (2006)]. These indices measure the change on failure probability while the parameters are perturbed at one given point. If the sensitivity index of one parameter is positive, then the failure probability tends to decrease if one reduces this parameter at the given point, otherwise, if the index is negative, then the failure probability tends to increase. The higher the absolute value of the local sensitivity index is, the more dramatically the failure probability will change. Theoretically, a parameter has significant effect on failure probability at one point doesn't mean that it is influential at each point, similarly, although the index of one parameter is close to zero at one point, one cannot think that this parameter is non-influential anywhere. Therefore, the local reliability sensitivity indices cannot tell the analysts the global sensitivity information of the distribution parameters of model inputs to the failure probability.

The global reliability sensitivity indices, which aims at measuring the contribution of individual or a set of inputs to the failure probability by investigating their full distribution ranges, are developed by Wei et al. (2012) based on Sobol's indices [Homma and Saltelli (1996) and Sobol' (1993)]. The higher the main effect index of one inputs is, the more reduction of failure probability can be obtained while one reduce the uncertainty of this input, otherwise, as the total effect index is close to zero, the failure probability will not change obviously while one reducing the uncertainty of this input. The global reliability sensitivity indices can be estimated by the methods developed for Sobol's indices such as MCS [Saltelli (2002); Saltelli et al. (2010); Sobol' (2001)], Fourier Amplitude sensitivity Test (FAST) [Xu and Gertner (2008)] and meta-modelling method [Ratto (2007)], thus can be easily implemented. The global reliability sensitivity indices can only tell the analysts which inputs to focus on so that the failure probability can be reduced efficiently and cheaply, but cannot tell the analysts the amount of failure probability reduction due to specific reduced uncertainty of model inputs.

In this paper, the failure probability ratio function is introduced for measuring the change on failure probability while the distribution parameters of model inputs vary in intervals. The proposed sensitivity index is similar to the function of failure probability developed by Au (2005) to some extent. The failure probability ratio function has important significance in many engineering application. Firstly, it can tell the analysts the amount of failure probability reduction while one change the distribution parameters of inputs to any specific ones, thus can help the analysts reducing the failure probability efficiently and quantitatively. Second, it can provide

plenty of information for RBO. After the failure probability ratio function been obtained, the RBO problem can be transformed to a deterministic one. Third, for model with epistemic uncertainty (due to lack of information (data), the distribution parameters of model inputs are represented by an confidence interval), it is helpful for selecting the inputs distribution parameters which are most valuable for collecting more information.

For numerically estimating the failure probability ratio function, the MCS procedure, which needs only a set of samples for implementing it, is firstly introduced. For problem with relatively large failure probability, the MCS procedure is accurate and efficient. For problem with small failure probability, we suggest using the Importance Sampling (IS) and Truncated Importance Sampling (TIS) procedures for reducing the computational burden.

## 2. Failure probability ratio function

Let  $Y = g(\mathbf{X})$  denote the limit state function of the structure under investigation, where  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  is the  $n$ -dimensional inputs vector. The joint PDF of the input vector  $\mathbf{X}$  is given as  $f_{\mathbf{X}}(\mathbf{x})$ , and the marginal PDF of the input  $X_i$  is denoted by  $f_i(x_i)$ . In this paper, we assume that the failure of structure happens when the model output  $Y$  is less than zero, thus the failure domain  $F$  is defined as:

$$F = \{\mathbf{x} : g(\mathbf{x}) < 0\} \quad (1)$$

and the failure probability  $P_f$  of the structure can be derived as:

$$P_f = P(F) = \int_F f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = \int I_F(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} = E_f(I_F(\mathbf{x})) \quad (2)$$

where  $P(\cdot)$  denotes the probability measure,  $E_f(\cdot)$  indicates the expectation with respect to the joint PDF  $f_{\mathbf{X}}(\mathbf{x})$  and  $I_F(\mathbf{x})$  stands for the indicator function of the failure domain, which is given as:

$$I_F(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in F \\ 0 & \mathbf{x} \notin F \end{cases} \quad (3)$$

Generally, the failure probability is related with the distribution parameters of model inputs such as the mean, variance and correlation of multivariate normal distribution. If one of these distribution parameters is changed, then the failure probability will also be changed. Let  $p$  denote one of these parameters, e.g., variance. Suppose now the parameter  $p$  varies in the interval  $[p^{(l)}, p^{(u)}]$ , then let  $p = (p^{(u)} - p^{(l)})q + p^{(l)}$ , where  $q$  is a variable. While the parameter  $p$  varies in the interval  $[p^{(l)}, p^{(u)}]$ ,  $q$  varies in the unit interval  $[0, 1]$ . If we keeps all the other parameters of model inputs constant, then denote the updated joint PDF of model inputs due to changed  $p$  as  $f_{\mathbf{X}}^*(\mathbf{x}; q)$ . For example, suppose the  $n$ -dimensional inputs vector follows independent

normal distribution, i.e.,  $f_{\mathbf{X}}(\mathbf{x}) = \prod_{j=1}^n f_j(x_j)$ , where  $f_j(x_j) = e^{-\frac{(x_j - \mu_j)^2}{2\sigma_j^2}} / \sqrt{2\pi\sigma_j^2}$ . Let  $p$  denote the variance  $\sigma_i^2$  of  $X_i$  and  $p$  varies in the interval  $[0, \sigma_i^2]$ , then the joint PDF  $f_{\mathbf{X}}(\mathbf{x})$  can be updated as:



$$f_X^*(\mathbf{x}; q) = \frac{1}{\sqrt{2\pi q \sigma_i^2}} e^{-\frac{(x-\mu_i)^2}{2\sigma_i^2}} \prod_{j=1, j \neq i}^n \frac{1}{\sqrt{2\pi \sigma_j^2}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}}, \quad q \in [0,1] \quad (4)$$

The failure probability with respect to the updated joint PDF  $f_X^*(\mathbf{x}; q)$  is computed as:

$$P_f^*(q) = \int I_F(\mathbf{x}) f_X^*(\mathbf{x}; q) d\mathbf{x}, \quad q \in [0,1] \quad (5)$$

The univariate failure probability ratio function  $RPF_p(q)$  is defined as:

$$RPF_p(q) = \frac{P_f^*(q)}{P_f} \quad (6)$$

By definition,  $RPF_p(q)$  measures the ratio between residual and total failure probabilities while the parameter  $p$  is changed to  $(p^{(u)} - p^{(l)})q + p^{(l)}$ , where  $q \in [0,1]$ .

Similarly, we can develop the multivariate failure probability ratio function for measuring the reduction of failure probability while multiple distribution parameters of model inputs are changed. Suppose now we have  $m$  distribution parameters  $p_k$  ( $k=1,2,\dots,m$ ) of model inputs, each of which varies in a interval  $[p_k^{(l)}, p_k^{(u)}]$ . Let  $p_k = (p_k^{(u)} - p_k^{(l)})q_k + p_k^{(l)}$ , where  $q_k \in [0,1]$ . Then, similarly, we can obtain the updated joint PDF  $f_X^*(\mathbf{x}; q_1, q_2, \dots, q_m)$ , and define the failure probability  $P_f^*(q_1, q_2, \dots, q_m)$  with respect to  $f_X^*(\mathbf{x}; q_1, q_2, \dots, q_m)$  as:

$$P_f^*(q_1, q_2, \dots, q_m) = \int I_F(\mathbf{x}) f_X^*(\mathbf{x}; q_1, q_2, \dots, q_m) d\mathbf{x}, \quad q_1, q_2, \dots, q_m \in [0,1] \quad (7)$$

The  $m$ -dimensional failure probability ratio function  $RPF_{p_1, p_2, \dots, p_m}(q_1, q_2, \dots, q_m)$  is defined as

$$RPF_{p_1, p_2, \dots, p_m}(q_1, q_2, \dots, q_m) = \frac{P_f^*(q_1, q_2, \dots, q_m)}{P_f} \quad (8)$$

The local reliability sensitivity index  $S_p$  of the parameter  $p$  is defined as the derivative of the failure probability with respect to  $p$  [Wu and Mohanty (2006)]:

$$S_p = \frac{\partial P_f / P_f}{\partial p / p} \Big|_{p=p^*} = \frac{p}{P_f} \cdot \frac{\partial P_f}{\partial p} \Big|_{p=p^*} \quad (9)$$

where  $p^*$  is a constant value (often chosen as the true value of the parameter  $p$ ). The partial derivative  $\partial P_f / \partial p \Big|_{p=p^*}$  indicates the change of failure probability while one perturbing the parameter  $p$  at point  $p^*$ .  $p/P_f$  is a normalization factor that makes the sensitivity index dimensionless.

Then it can be proved that (see Appendix for proof):

$$\frac{\partial RPF_p(q)}{\partial q} \Big|_{q=q^*} = \frac{p^{(u)} - p^{(l)}}{p^*} S_p \quad (10)$$

where  $p^* = (p^{(u)} - p^{(l)})q^* + p^{(l)}$  is a constant. Eq. (10) indicates that the derivative of

the probability ratio function with respect to the parameter  $p$  is proportional to the local reliability sensitivity index of  $p$ . The local sensitivity index  $S_p$  only reflect the sensitivity information of the parameter  $p$  at one given point  $p^*$ , whereas, the failure probability ratio function  $RPF_p(q)$  measures the reduction of failure probability when the parameter  $p$  is fixed at any point, thus it provides much more information on failure probability reduction than the local sensitivity index.

### 3. Estimators for Failure probability ratio function

#### 3.1 Monte Carlo simulation

By definition

$$P_f^*(q_1, q_2, \dots, q_m) = \int I_F(\mathbf{x}) \frac{f_{\mathbf{x}}^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{f_{\mathbf{x}}(\mathbf{x})} f_{\mathbf{x}}(\mathbf{x}) d\mathbf{x} = E_f \left( I_F(\mathbf{x}) \frac{f_{\mathbf{x}}^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{f_{\mathbf{x}}(\mathbf{x})} \right) \quad (11)$$

Eq. (13) indicates that  $P_f^*(q_1, q_2, \dots, q_m)$  can be expressed as the expectation with respect the original joint PDF  $f_{\mathbf{x}}(\mathbf{x})$ , thus can be estimated by MCS procedure using one set of samples. The Monte Carlo estimator is given as:

$$\hat{P}_f^*(q_1, q_2, \dots, q_m) = \frac{1}{N} \sum_{j=1}^N I_F(\mathbf{x}^{(j)}) \frac{f_{\mathbf{x}}^*(\mathbf{x}^{(j)}; q_1, q_2, \dots, q_m)}{f_{\mathbf{x}}(\mathbf{x}^{(j)})} \quad (12)$$

where  $\mathbf{x}^{(j)}$  ( $j=1, 2, \dots, N$ ) stands for the  $j$ th sample of model inputs generated by using the original joint PDF  $f_{\mathbf{x}}(\mathbf{x})$ .

Apparently,  $\hat{P}_f^*(q_1, q_2, \dots, q_m)$  is an unbiased estimator of  $P_f^*(q_1, q_2, \dots, q_m)$ . In this paper, we use the mean square error (MSE) for quantifying the error of the estimate. Take the univariate failure probability ratio function  $RPF_p(q)$  as an example, the MSE of the estimate  $\widehat{RPF}_p(q)$  is given as:

$$MSE_p = \sqrt{\int_0^1 (\widehat{RPF}_p(q) - RPF_p(q))^2 dq} \quad (13)$$

where  $RPF_p(q)$  is the reference result. In this paper, the MSE is estimated via bootstrap, thus the reference result  $RPF_p(q)$  is computed by averaging the repeated estimates.

The samples can be generated using many methods such as simple random sampling, Latin-hypercube sampling [Helton and Davis (2003); Loh (1996)] and Sobol's sequence [Sobol' (1976)]. In this paper, the Sobol's sequence is recommended since it leads to better convergence rate and lower discrepancy of estimates especially for input dimension less than a few hundred [Sobol' (1976); Varet et al. (2012)].

The above MCS procedure needs only a set of sample for estimating the failure probability ratio function. For structure with relative large failure probability, it is efficient and accurate. However, for small failure probability ( $<10^{-3}$ ), the MCS procedure need more samples for promising some of them dropping in the failure domain so that the failure probability can be correctly estimated, thus the computational cost increases heavily.

#### 3.2 Importance Sampling

To reduce the computational burden of simulation method for computing the small failure probability, many researchers have suggested using the IS procedure [Harbitz (1986); Melchers (1989); Melchers (1990)]. The basic idea of the IS procedure is

choosing an importance sampling density (ISD) for generating samples so that more samples drop into the failure domain.

Denote the ISD as  $h_x(\mathbf{x})$ , then Eq. (13) can be written as:

$$P_f^*(q_1, q_2, \dots, q_m) = \int I_F(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x})} h_x(\mathbf{x}) d\mathbf{x} = E_h \left( I_F(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x})} \right) \quad (14)$$

where  $E_h(\cdot)$  is the expectation operator taken with respect to the ISD  $h_x(\mathbf{x})$ . Eq. (16) indicates that  $P_f^*(q_1, q_2, \dots, q_m)$  can be expressed as an expectation with respect to  $h_x(\mathbf{x})$ , thus can be estimated by using the sample mean, i.e.,

$$\hat{P}_f^*(q_1, q_2, \dots, q_m) = \frac{1}{N} \sum_{k=1}^N \left( I_F(\mathbf{x}^{(k)}) \frac{f_x^*(\mathbf{x}^{(k)}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x}^{(k)})} \right) \quad (15)$$

Similarly, one can verify that Eq. (17) provides an unbiased estimator for the failure probability ratio function  $P_f^*(q_1, q_2, \dots, q_m)$ . The efficiency of the IS procedure is greatly affected by the choice of the ISD  $h_x(\mathbf{x})$ . Theoretically, the optimal ISD can be identified by minimizing the variance of the estimator in Eq. (17), i.e.,

$$\min_h V_h \left( I_F(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x})} \right) \quad (16)$$

where the subscript in  $V_h(\cdot)$  indicates that the variance is computed with respect to  $h_x(\mathbf{x})$ . It can be proved that the solution of the optimization problem is Eq. (18) is [Au and Beck (2002)]:

$$h_{x,opt}(\mathbf{x}) = \frac{I_F(\mathbf{x}) f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{\int I_F(\mathbf{x}) f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m) d\mathbf{x}} = \frac{I_F(\mathbf{x}) f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{P_f^*(q_1, q_2, \dots, q_m)} \quad (17)$$

By using the optimal ISD in Eq. (19), the variance of the estimates in Eq. (17) can be derived as:

$$V_{h_{opt}}(\hat{P}_f^*(q_1, q_2, \dots, q_m)) = \frac{1}{N} V_{h_{opt}} \left( I_F(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_{x,opt}(\mathbf{x})} \right) = \frac{1}{N} V_{h_{opt}}(P_f^*(q_1, q_2, \dots, q_m)) = 0 \quad (18)$$

Although the optimal ISD  $h_{x,opt}(\mathbf{x})$  can be derived by Eqs. (18) and (19), it is not available in practical application. One reason is that the identification of  $h_{x,opt}(\mathbf{x})$  involves the information of the failure probability  $P_f^*(q_1, q_2, \dots, q_m)$ , which is what to be estimated from the simulation. Another reason is due to the fact that the indicator function  $I_F(\mathbf{x})$  is unknown in advance. Even though the optimal ISD  $h_{x,opt}(\mathbf{x})$  can be constructed via some numerical method, it is often a complicated and time-consuming task to generate samples from the constructed ISD especially for high dimensional ISD [Au and Beck (2002)].

In practical application, researchers often attempt to construct the approximate optimal ISD under some assumptions. One of the most common used methods is assuming the ISD belongs to one family of distribution, and then choosing the optimal ISD via optimizing the distribution parameters. Au and Beck (2002) use the cross-entropy for searching the optimal distribution parameters of ISD belongs to assumed distribution family. Others suggested generating the ISD by shifting the center of the original joint PDF  $f_x(\mathbf{x})$  to the design point  $\mathbf{x}^*$  with the highest probability density in the failure region [Harbitz (1986); Melchers (1989)]. In the

examples of this paper, for simplicity, the later method is used. For problem with multiple failure modes, the ISD can be constructed by mixing multiple PDFs, each of which is centered at one design point. The design points can be identified by using many optimization algorithms such as FOSM [Hasofer and Lind (1974)] and genetic algorithm [Obadage and Harnpornchai (2006)].

Compared with the MCS procedure, the above IS procedure is more suitable for problem with small failure probability since the ISD  $h_x(\mathbf{x})$  allows more sample drop into the failure domain. One note that, as the distribution parameters vary, the failure domain will not change since the limit state function remains unchanged. Those samples within the failure domain will always stay in the failure domain no matter how the distribution parameters change. Therefore, one need only one set of samples generated by one pre-identified ISD  $h_x(\mathbf{x})$  for computing the failure probability ratio function  $RPF_{p_1, p_2, \dots, p_m}(q_1, q_2, \dots, q_m)$  at any points.

The main drawback of the above IS procedure is that it may not always be suitable for high-dimensional (up to a few hundred) nonlinear problems since that the identification of an approximate fixed ISD is practically impossible [Katafygiotis and Zuev (2008)].

### 3.3 Truncated Importance Sampling

In subsection 3.2, the introduction of the ISD constructed using the design point has substantially reduced the computational burden for computing the failure probability ratio function. In fact, the computational cost can be further reduced by using the TIS procedure [Grooteman (2008)].

In a standard Gaussian space, the reliability index  $\beta$  is in fact the distance from the design point  $\mathbf{x}^*$  to the origin of coordinate, as shown by Figure 1. Then a hypersphere with radius  $\beta$  can be obtained. We denote this hypersphere as  $\beta$ -sphere. Define the indicator function  $I_\beta(\mathbf{x})$  of the  $\beta$ -sphere as:

$$I_\beta(\mathbf{x}) = \begin{cases} 1 & \|\mathbf{x}\| \geq \beta \\ 0 & \|\mathbf{x}\| < \beta \end{cases} \quad (19)$$

It is shown in Fig. 1 that the failure region outside the  $\beta$ -sphere, and there is no failure point drop in the region of  $\beta$ -sphere, then the failure probability  $P_f^*(q_1, q_2, \dots, q_m)$  can be further derived as:

$$\begin{aligned} P_f^*(q_1, q_2, \dots, q_m) &= \int I_F(\mathbf{x}) I_\beta(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x})} h_x(\mathbf{x}) d\mathbf{x} \\ &= E_h \left( I_F(\mathbf{x}) I_\beta(\mathbf{x}) \frac{f_x^*(\mathbf{x}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x})} \right) \end{aligned} \quad (20)$$

Then  $P_f^*(q_1, q_2, \dots, q_m)$  can be estimated by:

$$\hat{P}_f^*(q_1, q_2, \dots, q_m) = \frac{1}{N} \sum_{j=1}^N \left( I_F(\mathbf{x}^{(j)}) I_\beta(\mathbf{x}^{(j)}) \frac{f_x^*(\mathbf{x}^{(j)}; q_1, q_2, \dots, q_m)}{h_x(\mathbf{x}^{(j)})} \right) \quad (21)$$

In Eq. (23), if the sample point  $\mathbf{x}^{(j)}$  drop into the  $\beta$ -sphere, then  $I_\beta(\mathbf{x}^{(j)}) = 0$ , further  $I_F(\mathbf{x}^{(j)}) I_\beta(\mathbf{x}^{(j)}) f_x^*(\mathbf{x}^{(j)}; q_1, q_2, \dots, q_m) / h_x(\mathbf{x}^{(j)}) = 0$ , thus one needs not to compute the value of limit state function at the point  $\mathbf{x}^{(j)}$ . By this way, the computational

burden is further reduced.

The above TIS procedure can further reduce the computational cost by introducing the  $\beta$ -sphere. However, since the design point  $\mathbf{x}^*$  is computed numerically, as some computational error exists, the  $\beta$ -sphere may include some non-negligible part of the failure region. This will further leads to computational error of the estimate of  $P_f^*(q_1, q_2, \dots, q_m)$  [Wei et al. (2012)].

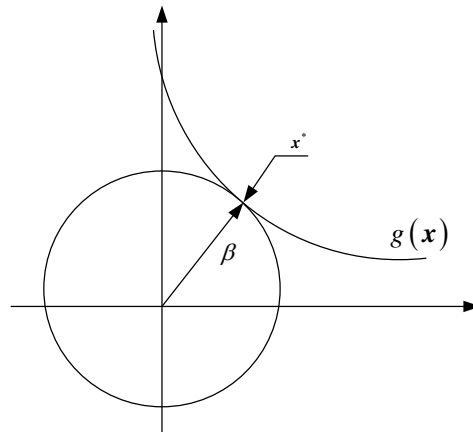


Figure 1. Schematic illustration of the  $\beta$ -sphere.

#### 4. Numerical test case

In this section, we use a numerical example for illustrating the engineering significance of the failure probability ratio function, and verifying the efficiency and accuracy of the proposed numerical methods. The limit state function of the structure is represented by

$$Y = g(\mathbf{X}) = 4 - X_1 - X_2 - X_1 X_3 \quad (22)$$

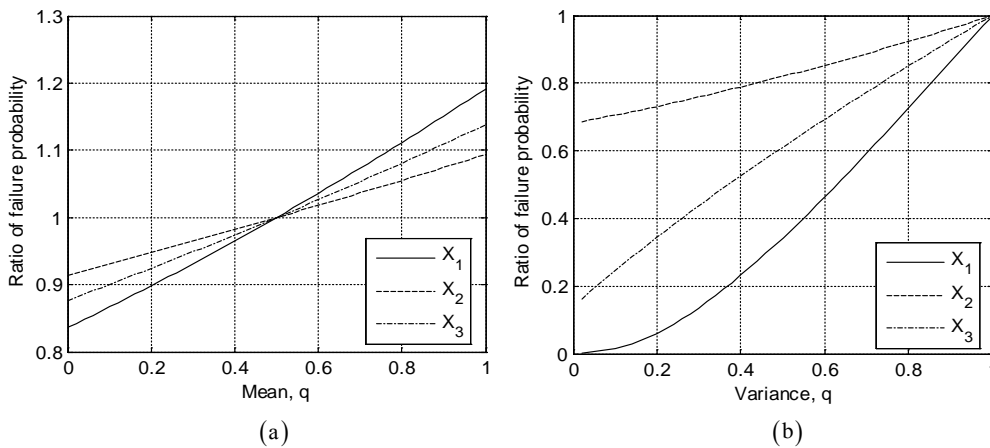
where  $\mathbf{X} = (X_1, X_2, X_3)$  is a vector including three inputs, which are assumed to follow standard normal distribution, i.e.,  $X_i \sim N(0,1)$  for  $i=1,2,3$ .

In this test case, we consider the sensitivity of the mean and variance of the model inputs to the failure probability. We assume that the means and variances of the three inputs vary in the interval  $[-0.1, 0.1]$  and  $[0, 1]$  respectively. Then the mappings from  $q$  to the mean  $\mu$  and variance  $\sigma^2$  are given as  $\mu = 0.2q + 0.1$  and  $\sigma_i^2 = q$ , where  $q \in [0, 1]$ .

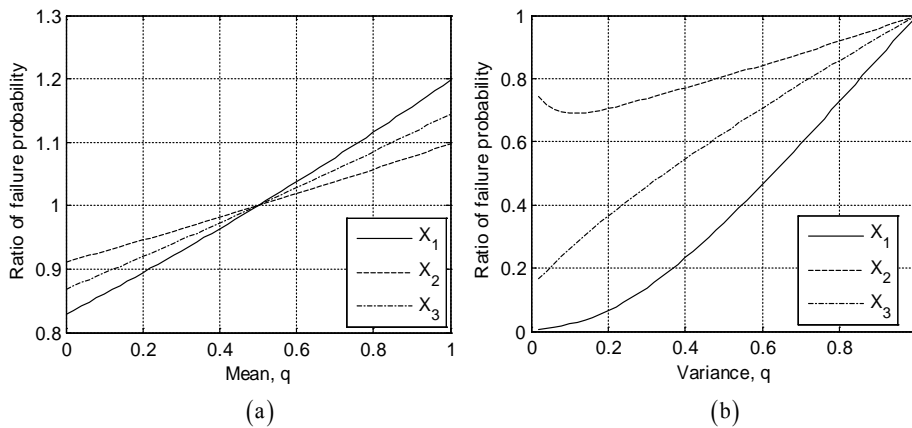
The univariate failure probability ratio functions with respect to the means and variances of the three inputs are computed by MCS, IS and TIS procedures using one set of sample, and the estimates are plotted in Figures 2-4. The total number of function evaluations of the MCS, IS and TIS procedure are  $10^6$ , 1022 and 719, thus the estimates of the MCS procedure can be regarded as the exact solution. It is shown that the estimates of the IS and TIS procedures are in good agreement with their exact solutions. Compared with the MCS procedure, the computation burden of the IS procedure is sufficiently reduced, and due to the introduction of the  $\beta$ -sphere, the computational cost is further reduced without affecting the accuracy of estimates. For

further investigating the convergence of the IS procedure, we plot the MSEs of the estimates with respect to the sample size in Figure.5. It can be seen that the MSE of each estimate tends to zero as the sample size increases.

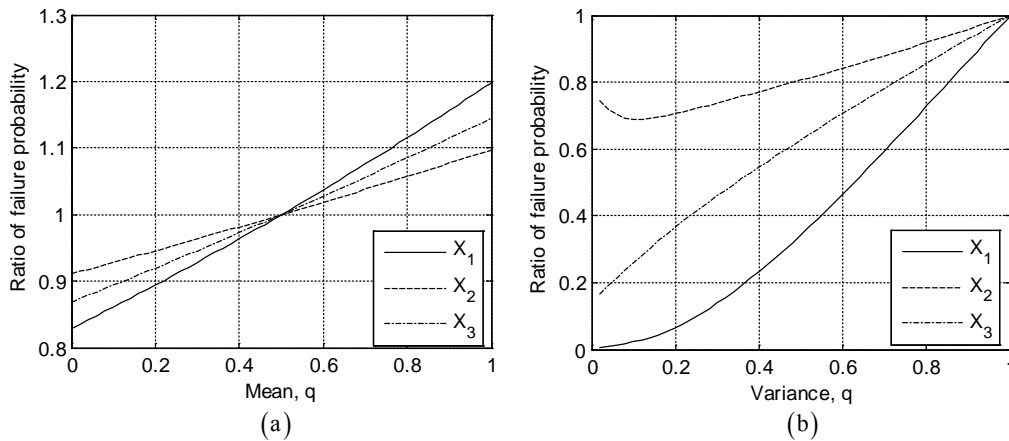
It is shown by Fig.2(a) that, the failure probability increases monotonically and linearly with respect to the means of the three inputs. As one decrease the means of the model inputs, the failure probability tends to reduce linearly, and the slopes of the failure probability ratio functions in Fig.2(a) indicate the rate of reduction. As can be seen, as reducing the same amount of means,  $X_1$  leads to the most reduction of failure probability, followed by  $X_3$ , and then  $X_2$ .



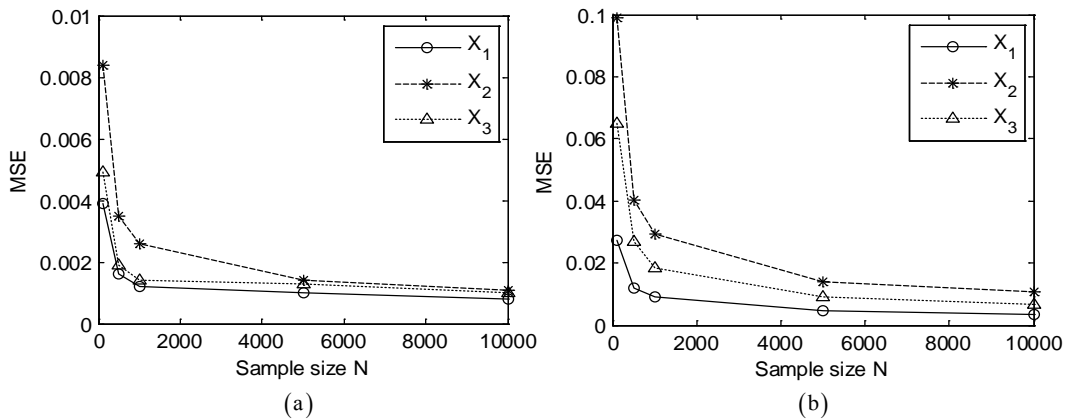
**Figure 2. Univariate failure probability ratio functions computed by MCS procedure: (a) with respect to means of inputs; (b) with respect to variances of inputs.**



**Figure 3. Univariate failure probability ratio functions computed by IS procedure: (a) with respect to means of inputs; (b) with respect to variances of inputs.**



**Figure 4. Univariate failure probability ratio functions computed by TIS procedure: (a) with respect to means of inputs; (b) with respect to variances of inputs.**

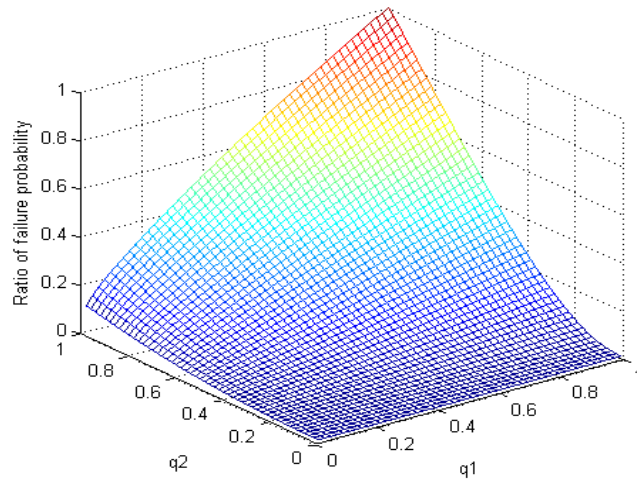


**Figure 5. Convergence plots of the univariate failure probability ratio functions with respect to (a): mean of inputs; (b) variance of inputs.**

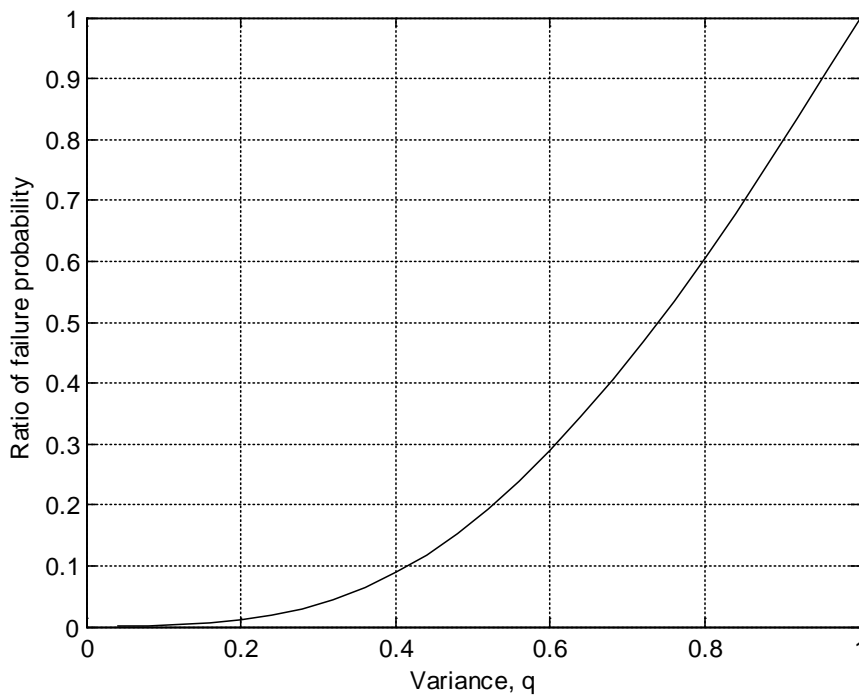
In many engineering application, the failure probability is reduced via decreasing the dispersion (variance) of model inputs. Suppose now our target is to reduce the failure probability by 80%. As one can see in Figure.2(b), the failure probability tends to decrease as the variances of inputs are reduced. It can be read from Figure.2(b) that, to achieve our targeted reduction of failure probability, one need to reduce the variance of  $X_1$  by 63% or that of  $X_3$  by 94% individually. It is possible to achieve our target by reducing the variance of  $X_2$ .

In most cases, it is difficult to reduce the variance of one input by 63%. For reducing the failure probability by 80%, one needs to use the bivariate failure probability ratio function. Since that reducing the variances of  $X_1$  and  $X_3$  leads to more failure probability reduction that reducing the variance of  $X_2$ , we plot the bivariate failure probability ratio function (computed by MCS procedure) with respect to the variances of  $X_1$  and  $X_3$  in Figure.6. The diagonal line of this bivariate failure probability

ratio function measures the ratio between the residual and total failure probability while both the variances of  $X_1$  and  $X_3$  are reduced to  $q$ , thus we plot it in Figure 7. As can be seen that, by reducing the variances of  $X_1$  and  $X_3$  by 48% simultaneously, we can achieve our targeted reduction of failure probability.



**Figure 6. 3D plots of the bivariate failure probability ratio function with respect to the variances of the pair  $(X_1, X_3)$ , where  $q_1$  and  $q_2$  indicate the variance reduction of  $X_1$  and  $X_3$  respectively**



**Figure 7. Diagonal line of the failure probability ratio function with respect to the variance of the input pair  $(X_1, X_3)$**



## Acknowledgement

This work is supported by National Natural Science Foundation of China (NSFC 51475370)

## References

- Au, S.K. and Beck J.L. (2001) Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics* **16**: 263-277
- Au, S.K. and Beck, J.L. (2002) Important sampling in high dimensions. *Structural Safety* **25**: 139-163
- Au, S.K. (2005) Reliability-based design sensitivity by efficient simulation. *Computers and Structures* **83**: 1048-1061.
- Bjerager, P. and Krenk, S. (1989). Parametric sensitivity in first order reliability theory. *Journal of Engineering Mechanics* **115**: 1577-1582
- Chen, X., Hasselman, T. and Neill, D. (1997) Reliability-based structural design optimization for practical applications. In: Proceedings of the 38th AIAA structures, structural dynamics, and materials conference, Florida.
- Ditlevsen, O., Melchers, R.E. and Gluwer H. (1990) General multi-dimensional probability integration by directional simulation. *Computers and Structures* **36**: 355-368
- Enevoldsen, I. and Sorensen, J.D. (1994) Reliability-based optimization in structural engineering. *Structural Safety* **15**:169-196.
- Faravelli, L. (1989) Response-surface approach for reliability analysis. *Journal of Engineering Mechanics* **115**: 2736-2781
- Grooteman, F. (2008). Adaptive radial-based importance sampling method for structural reliability. *Structural Safety* **30**: 533-542
- Harbitz, A. (1986) An efficient sampling method for probability of failure calculation. *Structural Safety* **3**:109-115.
- Hasofer, A.M. and Lind, N.C. (1974) An exact and invariant first order reliability format. *Journal of Engineering Mechanics* **100**: 111-121.
- Helton, J.C. and Davis G.J. (2003). Latin hypercube sampling and the propagation of uncertainty in analysis of complex systems. *Reliability Engineering and System Safety* **81**: 23-69
- Homma, T. and Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and system Safety* **52**: 1-17
- Katafygiotis, L.S. and Zuev K.M. (2008). Geometric insight into the challenges of solving high-dimensional reliability problems. *Probabilistic Engineering Mechanics* **23**:208-218
- Kuschel, N. and Rackwitz, R. (1997) Two basic problems in reliability-based structural optimization. *Mathematical Methods of Operations Research* **46**: 309–333.
- Loh W.L. (1996). On Latin hypercube sampling. *The Annals of Statistics* **24**: 2058-2080
- Melchers, R.E. (1989). Importance sampling in structural system. *Structural Safety* **6**: 3-10.
- Melchers, R.E. (1990) Search-based Importance Sampling. *Structural Safety* **9**: 117-128
- Melchers, R.E. and Ahammed, M. (2004) A fast approximate method for parameter sensitivity estimation in Monte Carlo structural reliability. *Computers and Structures* **82**: 55-61
- Obadage, A.S. and Harnpornchai, N. (2006). Determination of point of maximum likelihood in failure domain using genetic algorithms. *International Journal of Pressure Vessels and Piping* **83**:276-282
- Ratto, M., Pagano, A. and Young, P. C. (2007). State dependent parameter metamodelling and sensitivity analysis. *Computer Physics Communications* **177**: 863–876.
- Royset, J.O., Der Kiureghian, A. and Polak, E. (2001) Reliability-based optimal structural design by the decoupling approach. *Reliability Engineering and System Safety* **73**:213-221
- Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications* **145**: 280-297
- Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and Tarantola, S. (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity indices. *Computer Physics Communications* **181**: 259–270
- Schuëller, G.I., Pradlwarter, H.J. and Koutsourelakis, P.S. (2004) A critical appraisal of reliability estimation procedures for high dimensions. *Probabilistic Engineering Mechanics* **19**: 463-473
- Schuëller, G.I. and Jensen, H.A. (2008). Computational methods in optimization considering uncertainties- an overview. *Computer Methods in Applied Mechanics and Engineering* **198**:2-13
- Sobol', I.M. (1976). Uniformly distributed sequences with additional uniformity properties. USSR

- Computational Mathematics and Mathematical Physics 16: 236-242
- Sobol, I.M. (1993). Sensitivity analysis for non-linear mathematical models. *Mathematical Modelling and Computational Experiment* 1:407-414
- Sobol, I.M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55:271-280
- Tu, J., Choi, K.K. and Park, Y.H. (1999) A new study on reliability-based design optimization. *Journal of Mechanical Design* 121: 557-564.
- Valdebenito, M.A. and Schuëller, G.I. (2010). A survey on approaches for reliability-based optimization. *Structural and Multidisciplinary Optimization* 42: 645-663
- Varet, S., Lefebvre, S., Durand, G., Roblin, A., & Cohen, S. (2012) Effective discrepancy and numerical experiments. *Computer Physics Communications* 183: 2536-2541
- Wei, P.F., Lu, Z.Z., Hao W.R., Feng, J., Wang B.T. (2012) Efficient sampling methods for global reliability sensitivity analysis. *Computer Physics Communications* 183:1728-1743
- Wu, Y.T. and Mohanty, S. (2006). Variable screening and ranking using sampling-based sensitivity measures. *Reliability Engineering and System Safety* 91: 643-647
- Xu, C., and Gertner G.Z. (2008). A general first-order global sensitivity analysis method. *Reliability Engineering and System Safety* 93:1060-1071
- Zou, T. and Mahadevan, S. (2006) A direct decoupling approach for efficient reliability-based design optimization. *Structural and Multidisciplinary Optimization* 31: 190-200

## The Clenshaw-Curtis quadrature in isogeometric analysis

†X.F. Shi<sup>1</sup>, P. Xi<sup>1</sup>, Y.W. Song<sup>2</sup> and X.M. Cai<sup>3</sup>

<sup>1</sup>School of Mechanical Engineering and Automation, BeiHang University, Beijing, China.

<sup>2</sup>Energy & Power Engineering School, North China Electric Power University, Beijing, China

<sup>3</sup>The human resources and social security bureau of Heze, Shandong, China

†Corresponding author: xfshi@me.buaa.edu.cn

### Abstract

Isogeometric analysis (IGA) is a relatively new method and receiving much attention recently, the efficient quadrature in which is an important branch far from mature. We introduce the Clenshaw-Curtis quadrature into IGA and give the corresponding algorithms. The estimated computation cost for both rules and of the whole isogeometric approximation are proposed, and through which we compare it with the optimal standard Gauss rule. It is found that the Clenshaw-Curtis rule have better efficiency than the Gauss for spline degree of 2. Better accuracy of CC than Gauss for low spline degrees are also found through the applications of both rules in the boundary value problems.

**Keywords:** isogeometric analysis, quadrature, Clenshaw-Curtis, Gauss, NURBS

### 1 Introduction

Isogeometric analysis (IGA) is a recently proposed subject which is receiving a great deal of attention amongst the computational mechanics community. isogeometric analysis is a technique of numerical analysis that uses the same basis functions commonly found in description of Computer Aided Design (CAD) geometries to represent both geometry and physical fields in the solution of problems governed by partial differential equations (PDE)[Hughes et al. (2005); Cottrell et al.(2009)]. Based on its initial intends of bridging the gap between the CAD and the Finite Element Analysis (FEA), IGA has the potential to have a profound effect and the promise of overcoming some bottleneck issues that plagued computer aided engineering for decades.

The use of the most popular Non-Uniform Rational B-Spline (NURBS) basis function applied for the geometry description in the solution field therefore leads to elimination of geometric-approximation error in even the coarsest mesh. In this way, the isoparametric concept is maintained but more significantly, the geometry of the problem is preserved exactly. The increased continuity of the NURBS basis has led to significant numerical advantages over traditional Lagrange polynomials and other  $C^0$  inter-element continuity based FEA, e.g. it can possess high regularity across mesh elements, leading to a higher accuracy per degree-of-freedom (DOF) basis [Cottrell et al.(2009)]; it also has better robustness and system condition number than FEA [Bazilevs et al.(2006)]. Many researchers have applied B-splines and NURBS as the basis for IGA applications such as fluid dynamics [Bazilevs et al.(2006); Bazilevs et al.(2012)], structural mechanics [Kiendl et al.(2009); Lipton et al.(2010); Benson et al.(2011)], thermal analysis [Anders et al.(2012)], shape optimization [Qian (2010)], electromagnetics [Buffa and Sangalli (2010)] and so on.

However, several challenges remain for IGA to be fully accepted as industrial-strength analysis technology. One of them is the design of efficient and adaptive quadrature rules. The quadrature scheme of the IGA is accomplished over individual non-zero knot spans of the underlying B-spline based geometry, which is different from performing the numerical quadrature on individual finite elements in the FEA. In fact, the widely used Gauss quadrature on each element in the IGA computation is a choice far from being optimal [Auricchio et al. (2012)]. An optimal quadrature rule which exactly integrates B-spline basis functions with the minimum number of function evaluations for IGA was initially constructed in [Hughes et al. (2010)]. It significantly improves the computational efficiency despite that sometimes it is difficult to solve for high polynomial degrees and numbers of elements due to a global ill-conditioned equation system. [Auricchio et al. (2012)] developed an efficient algorithm through which can obtain nearly optimal rules. That algorithm is proved to be much easier to construct.

In this paper, we discuss the quadrature in IGA from another aspect. As mostly used rules are commonly Gaussian, we introduced an existing non-Gaussian rule named Clenshaw-Curtis into IGA and explored its new features. The Clenshaw-Curtis rule uses Chebyshev points instead of optimal nodes of Gauss quadrature. The computation of a cosine transformation and the arithmetic cost of this were prohibitive and thus limited the use of this rule before the FFT transformation was used. It has particular advantages such as easier implementation [Gentleman (1972a; 1972b)], most similar convergence rate [Calabrò and Esposito (2009)] and in fact, for most integrands, about equally accurate [Trefethen (2008)] compared to the Gauss quadrature. We know its own merits, but how it performs in IGA – this is what we discuss in this paper. We discuss its convergence and efficiency through comparisons with standard Gauss rule and find some interesting points. Note that, there are several variations on this theme (see [Trefethen (2008); Clenshaw and Curtis (1960)]). What we use in this paper is commonly called “practical” Clenshaw-Curtis formula.

The paper is organized as follows. Section 2 gives some of the preliminaries on IGA and Clenshaw-Curtis rules. Section 3 studies the integration of quadrature rules into IGA and makes discussions on computational cost of both rules. Section 4 exploits the Clenshaw-Curtis rules to numerically solve boundary value problems in Poisson’s and elasticity problems and makes verifications of Section 3. In this paper, we took advantage of the open-source codes of GeoPDEs (<http://geopdes.sourceforge.net>) and modified the corresponding parts.

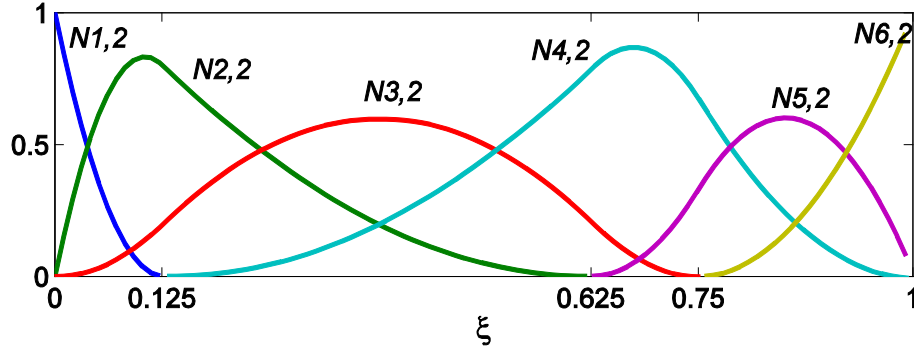
## 2. Preliminaries on IGA

We start with a brief review of some technical aspects of B-spline and NURBS bases for IGA. More detailed introduction can be found in the fundamental works proposed by [Hughes et al. (2005); Cottrell et al.(2009)].

As aforementioned, similar with the isoparametric concept of standard FEM, isogeometric analysis uses higher degree smooth spline functions, in particular B-splines and NURBS. A univariate B-spline function of polynomial degree  $m$  is specified by  $n$  basis functions  $N_{i,m}(\xi)$  ( $N_{i,m}$ , for short), ( $i=1,\dots,n$ ) in the parametric space  $\xi$ . The non-decreasing set of  $(n+m+1)$  coordinates  $\xi_i$  are so-called knots and subdivide the parametric space into  $(n+m)$  knot spans forming a patch [Hughes et al. (2005); Cottrell et al.(2009)].

$$\Xi = \{\xi_1, \xi_2, \dots, \xi_{n+m+1}\} \quad (1)$$

Piecewise polynomial B-spline functions are defined over  $m+1$  knot spans with  $C^{m-1}$  continuity between the spline elements. Repeated knots decrease the continuity between the knot spans and make the B-splines interpolatory at the knots. For a repetition of the first and last knot the knot span is said to be open. Knot spans with non-zero extension will in the following be referred to as knot-span elements. The 1D patch of Figure 1 consists of four knot-span elements. The B-spline basis functions are constructed recursively by the *Cox-de Boor* formula [Piegl and Tille (1997)]



**Figure1. 1D non-uniform NURBS patch.**

NURBS basis functions and geometric entities are then immediately obtained from the previous B-Spline spaces. In brief, a positive weight  $\omega_i$  can be associated to each B-Spline basis function  $N_{i,m}$ , and the corresponding NURBS basis function is defined as

$$R_{i,m}(\xi) = \frac{\omega_i N_{i,m}(\xi)}{\sum_{j=1}^n \omega_j N_{j,m}(\xi)} \quad (2)$$

Such a definition is easily generalized to the two- and three-dimensional cases by means of tensor products. For instance, in the trivariate case, given the degrees  $p_d$ , the integers  $n_d$  and the knot vectors  $\Xi$ ,  $H$  and  $\Gamma$ , the corresponding B-spline and NURBS basis functions are

$$B_{i,j,k}^{p,q,r}(\xi, \eta, \gamma) = N_{i,p}(\xi) M_{j,q}(\eta) L_{k,r}(\gamma) \quad (3)$$

and

$$R_{i,j,k}^{p,q,r}(\xi, \eta, \gamma) = \frac{\omega_{i,j,k} B_{i,j,k}^{p,q,r}(\xi, \eta, \gamma)}{\sum_{\hat{i}\hat{j}\hat{k}} \omega_{\hat{i}\hat{j}\hat{k}} B_{\hat{i}\hat{j}\hat{k}}^{p,q,r}(\xi, \eta, \gamma)} \quad (4)$$

B-spline or NURBS curves, surfaces and volumes are then defined as

$$\mathbf{C}(\xi, \eta, \gamma) = \sum \mathbf{P}_{ijk} R_{i,j,k}^{p,q,r}(\xi, \eta, \gamma) \quad (5)$$

### 3. Clenshaw-Curtis Quadrature in the element of IGA

Let  $n > 1$  be a given fixed integer, and define  $n+1$  quadrature nodes on the standard interval  $[-1, 1]$  as the extremes of the Chebyshev polynomial  $T_n(x)$ , augmented by the boundary points,

$$x_k := \cos \vartheta_k, \quad \vartheta_k := k \frac{\pi}{n}, \quad k = 0, 1, \dots, n. \quad (6)$$

Given a spline function  $f$ , an  $n$ -point interpolatory quadrature rule is a choice of  $n$  ordered points and weights such that

$$\int_{-1}^1 f(x)dx = \sum_{k=0}^n \omega_k f(x_k) + R_n \tag{7}$$

where  $R_n$  is the approximation error, and  $\omega_k$  are the quadrature weights, which can be obtained by integrating the  $n$ -th-degree polynomial interpolating the  $n+1$  discrete points  $(x_k, f(x_k))$ . Applying this procedure to the nodes eq.(6) directly yields the Clenshaw-Curtis rules. [Davis and Rabinowitz (1984)] summarized the explicit expressions for the Clenshaw-Curtis weights  $\omega_k^{cc}$

$$\omega_k^{cc} = \frac{c_k}{n} \left( 1 - \sum_{j=1}^{\lfloor n/2 \rfloor} \frac{b_j}{4j^2 - 1} \cos(2j\theta_k) \right), \quad k = 0, 1, \dots, n, \tag{8}$$

where the coefficient  $b_j, c_k$ , are defined as

$$b_j = \begin{cases} 1, & j = n/2 \\ 2, & j < n/2, \end{cases} \quad c_k = \begin{cases} 1, & k = 0 \pmod n \\ 2, & \textit{otherwise} \end{cases} \tag{9}$$

Eq. (8) holds for every even or odd integer  $n > 1$ , which together with the definition (9) of  $c_k$  implies

$$\omega_0^{cc} = \omega_n^{cc} = \frac{1}{n^2 - 1 + \text{mod}(n, 2)} \tag{10}$$

We give the detailed constructions proposed in [Waldvogel (2006)], which are given by the *inverse discrete Fourier transform* of the vector  $\mathbf{v} + \mathbf{g}$ , where  $\mathbf{v}$  and  $\mathbf{g}$  is defined in eq.(11) and eq.(12), respectively. The evaluation is particularly fast if  $n$  is a power of 2.

$$v_k = \frac{2}{1 - 4k^2}, \quad k = 0, 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor - 1, \\ v_{\lfloor \frac{n}{2} \rfloor} = \frac{n-3}{2 \left\lfloor \frac{n}{2} \right\rfloor - 1} - 1, \\ v_{n-k} = \bar{v}_k, \quad k = 1, 2, \dots, \left\lfloor \frac{n-1}{2} \right\rfloor; \tag{11}$$

$$g_k = -\omega_0^{cc}, \quad k = 0, 1, \dots, \left\lfloor \frac{n}{2} \right\rfloor - 1, \\ g_{\lfloor \frac{n}{2} \rfloor} = \omega_0^{cc} [(2 - \text{mod}(n, 2))n - 1], \\ g_{n-k} = \bar{g}_k, \quad k = 1, 2, \dots, \left\lfloor \frac{n-1}{2} \right\rfloor. \tag{12}$$

where  $\omega_0^{cc}$  is defined in eq.(10) with  $\omega_n^{cc} := \omega_0^{cc}$ . The superscripts in defining  $v_{n-k}$  and  $g_{n-k}$  refer to complex conjugation.

We take one-dimensional parametric domain as an example, two and three-dimensional can be easily obtained by means of tensor product. We assume the interval  $[0,1]$  with a uniform subdivision into  $l$  unitary elements  $[\frac{i-1}{l}, \frac{i}{l}]$ , where  $i = 1, \dots, l$ .

**Remark 3.1:** We call  $S_q^m$  the space spanned by the B-spline basis functions with global  $C^q$ -continuity, which is associated with a knot vector having internal knots with multiplicity  $r = m - q$ .

Given a spline function  $f \in S_q^m$ , which is a polynomial of degree  $m$  in each element  $[\frac{i-1}{l}, \frac{i}{l}]$  with  $f \in C^q([0,1])$ . In the computation,  $m+1$  coefficients on each of the  $l$  elements with  $q+1$  continuity requirements on the  $n-1$  internal points need to be assigned. The dimension of the space  $S_q^m$  is therefore  $l(m-q)+q+1$ , where  $-1 \leq q \leq m-1$  as  $q=-1$  refers to the discontinuous case. In particular,  $q=0$  refers to the case of functions continuous on the whole parametric domain and piecewise polynomial on each single element, which is similar to the FEA approximation.

In the elements defined above, Gauss quadratures are commonly used in numerical integration of functions in  $S_q^m$ . Compared with the standard Gauss rule, which has its quadrature points totally implemented within each element boundaries, the Clenshaw-Curtis have points at element boundaries. We apply Clenshaw-Curtis quadrature points eq.(6) to each element (i.e. subinterval)  $[\frac{i-1}{l}, \frac{i}{l}]$  and thus the first and the last quadrature points are exactly the  $x_1^i = \frac{i-1}{l}$  and  $x_{m-1}^i = \frac{i}{l}$ . If there are same integration points number  $n_l$  in each element, the translation and scaling method

$$\tilde{x}_k^i = x_k^{[-1,1]} \cdot (x_k^i - x_k^{i-1}) + x_k^{i-1}, \quad k = 1, \dots, n_l \quad (13)$$

are always used to save the computation, see Figure 2a; where  $\tilde{x}_k^i$  is the unknown quadrature point;  $x_k^{[-1,1]}$  is the  $k$ -th point in the biunit interval  $[-1,1]$ ;  $x_k^i$  is the  $k$ -th point in the  $i$ -th element. Sometimes quadrature points need to be enriched in certain elements or parametric direction in the sense of product tensor, as we need high accuracy there than the acceptable general accuracy of the whole domain. For this case, different point numbers are implemented and the corresponding weights are usually evaluated separately. Figure 2b shows that the quadrature points are added to 4 in the subinterval  $[0.125, 0.625]$  where only 2 Gauss points are needed to obtain the exact integration for the basis functions of degree 2. A pseudo-code for implementation of the standard Clenshaw-Curtis quadrature points is proposed below according to the two cases aforementioned, which are the respective using of uniform quadrature points and different quadrature points in subintervals. We fix the spline degree and full regularity for simplicity. The computational domain is considered  $k$ -dimensional.

**Input:**  $1 \times k$  integer array: spline\_degree

**for** parametric direction:  $idir=1, \dots, k$

> find out the unique knots:  $uniq\_knots(1:n)$ ;

> record the first  $n-1$  elements of  $uniq\_knots(1:n)$ :  $uniq\_knots\_el(1:n-1)$ ;

> calculate the differences adjacent elements of  $uniq\_knots(1:n)$ :  $du(1:n-1)$ ;

**if** case 1, **then**

> calculate quadrature points in unit interval  $[-1,1]$ :  $q\_points\_temp(idir, points, weights) = CC(spline\_degree(idir))$ ;

**for** the  $i$ -th node,  $inode = 1:n-1$

```

> calculate quadrature points in current parametric direction:  $q\_points(idir, inode) = (q\_points\_temp(idir, :, weights) + 1) / 2 * du(inode) + uniq\_knots\_el(inode)$ ;
> calculate quadrature weights in current parametric direction:  $q\_weights(idir, inode) = q\_points\_temp(idir, points, :)$ ;
endfor
elseif case 2, then
  for the  $i$ -th node,  $inode = 1:n-1$ 
    > calculate quadrature points in unit interval  $[-1, 1]$ :  $q\_points\_temp(idir, inode, points, weights) = CC(spline\_degree(idir, inode))$ ;
    > calculate quadrature points in current parametric direction:  $q\_points(idir, inode) = (q\_points\_temp(idir, inode, :, weights) + 1) / 2 * du(inode) + uniq\_knots\_el(inode)$ ;
    > calculate quadrature weights in current parametric direction:  $q\_weights(idir, inode) = q\_points\_temp(idir, points, :)$ ;
  endfor
endif
endfor
Return:  $n$ -points quadrature rule  $(x_n, \omega_n)$  in the whole parametric direction.

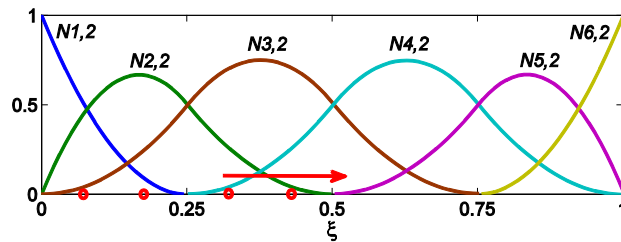
```

in which the subroutine CC is used for evaluation of the Clenshaw-Curtis points and weights in the unit interval. A pseudo-code for subroutine CC is listed below.

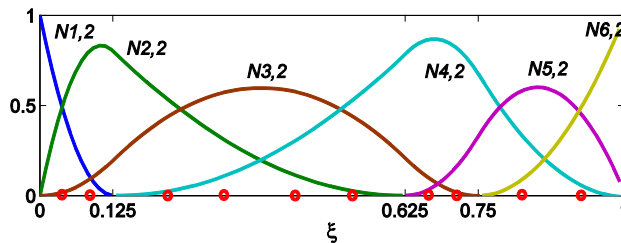
```

Input: number points minus 1:  $n-1$ 
> calculate quadrature points  $x_k$  defined in eq.(6);
> calculate the vector  $\mathbf{v}$  and  $\mathbf{g}$  defined in eq.(11) and eq.(12);
> perform the inverse discrete Fourier transform of the vector  $\mathbf{v} + \mathbf{g}$ ;
Return:  $n$ -points quadrature rule  $(x_n^{[-1,1]}, \omega_n^{[-1,1]})$  in biunit interval  $[-1, 1]$ .

```



**Figure 2a.** Uniform quadrature points (red circles) used in subintervals, which are implemented by the translation and scaling method.

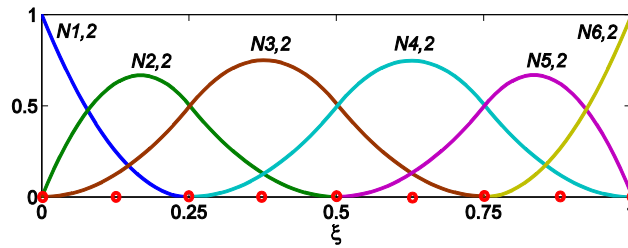


**Figure 2b.** Different quadrature points (red circles) used in subintervals, where the translation and scaling method are invalid.



For  $m$ -degree basis functions,  $\frac{m+1}{2}$  or  $\frac{m+2}{2}$  (for  $m$  is odd or even, respectively) quadrature points for Gauss rules and  $m+1$  for Clenshaw-Curtis rules per element are needed in order to exactly integrate functions in space. Note that, for Gauss rules, all the  $\frac{m+1}{2}$  or  $\frac{m+2}{2}$  points are within each knot span; while for Clenshaw-Curtis, there are  $m-1$  rather than  $m+1$  points within each element as the two boundary points are the knots themselves (Figure 3). Here we denote the minimum number of the quadrature points needed in Clenshaw-Curtis and Gauss by  $n_{I_{\min}}^{CC}$  and  $n_{I_{\min}}^G$  respectively, which expressed as,

$$n_{I_{\min}}^{CC} = m-1, \quad n_{I_{\min}}^G = \begin{cases} (m+1)/2 & \text{for } m = \text{odd} \\ (m+2)/2 & \text{for } m = \text{even} \end{cases} \quad (14)$$



**Fig.3 Setting the quadrature points (red circles) using Clenshaw-Curtis rules in one parametric direction.**

For smaller values of quadrature points number  $n_I$ , the well known Gauss efficiency of the factor of 2 cannot be achieved (i.e., size of  $n_I$  needed to achieve a certain accuracy). [Trefethen (2008)] pointed out that, for functions that are not analytic in stable neighborhood of  $[-1,1]$ , the Clenshaw-Curtis rule too comes close to  $|I - I_{n_I}| \approx E_{2n_I}^*$ , see Remark 3.2. Accordingly, for common low-degree ( $m \leq 4$ ) spline functions in IGA, the minimum number of the quadrature points  $n_{I_{\min}}^{CC}$  needed in Clenshaw-Curtis rule can be decreased. Furthermore, in the lower degree case that the quadrature points and weights need to be calculated in each element, Clenshaw-Curtis rule should be more efficient than Gauss rule in a literal sense; as the former can be done in  $O(n_I \log n_I)$  operations comparing with operations in  $O(n_I^2)$  for the latter. For simplicity, we assume an open knot vector with the form

$$U = \{\underbrace{a, \dots, a}_{m+1}, \xi_{m+1}, \dots, \xi_{l-m-1}, \underbrace{b, \dots, b}_{m+1}\} \quad (15)$$

where the first and the last knots have a multiplicity of  $m+1$  and 1 for other knots,  $m$  is the spline degree,  $l+1$  is the knot number, thus the element number  $k_e$  is

$$k_e = l - 2m \quad (16)$$

and the basis function number is

$$n_B = l - m \quad (17a)$$

or

$$n_B = k_e + m \quad (17b)$$

The estimated operations needed in IGA for these two rules are

$$\text{for CC: } O\{n_{I_{\min}}^{CC} \log n_{I_{\min}}^{CC} + n_{I_{\min}}^{CC} n_B n_{op}\} \quad (18a)$$

$$\text{for Gauss: } O\{n_{I_{\min}}^G{}^2 + n_{I_{\min}}^G n_B n_{op}\} \quad (18b)$$

where  $n_{op}$  is the times of evaluations of basis functions at quadrature points, which is determined by the feature of the problem and the experience of the programmer. The first terms of these two equations are related to the evaluations of quadrature points and weights in non-uniform case; the second terms are related to the evaluations of basis functions of the discrete space at quadrature points computed above. Substitute eq.(17a) and eq.(17b) into eq.(18a) and eq.(18b) yield the estimated operations needed in IGA for this two rules,

$$\text{for CC: } O\{[k_e(m-1)+1]\log[k_e(m-1)+1]+[k_e(m-1)+1](k_e+m)n_{op}\} \quad (19a)$$

$$\text{for Gauss: } \begin{cases} O\left\{\left[k_e\left(\frac{m+1}{2}\right)\right]^2 + \left[k_e\left(\frac{m+1}{2}\right)\right](k_e+m)n_{op}\right\} & \text{for } m = \text{odd} \\ O\left\{\left[k_e\left(\frac{m+2}{2}\right)\right]^2 + \left[k_e\left(\frac{m+2}{2}\right)\right](k_e+m)n_{op}\right\} & \text{for } m = \text{even} \end{cases} \quad (19b)$$

We give the estimated results in Figure 4a - Figure 4c for  $n_{op}$  equaling to  $1 \times 10^{1-3}$  respectively. In each figure the variations of computational cost (operation times) with the spline degree is plotted. We find that the CC rule has higher efficiency for spline degree of 2 than Gauss, which is applicable for all of three cases. However, the Gauss rule is faster for degrees larger than 3 with a larger  $n_{op}$  (for  $n_{op}$  of significantly greater than 10, e.g. 100 and 1000 as shown in figure, the Gauss needs less operations; for  $n_{op}$  of near 10, both are almost the same). Another surprising finding about CC rule in IGA for the spline degree of 2 will be elaborate in section 4. The minimum numbers of the quadrature points for all cases of 1D are shown in table 1. The interesting observation about  $n_t^{CC} \leq n_{t_{min}}^{CC}$  will be shown in section 4.

**Remark 3.2:** We use the definition of  $E_{n_t}^*$  in [Trefethen (2008)]: Given  $f \in C[-1,1]$  and  $n_t \geq 0$ , let  $p_{n_t}^*$  be the unique best approximation to  $g$  on  $[-1,1]$  of degree  $\leq n_t$  with respect to the supremum norm  $\|\cdot\|_\infty$ , and define  $E_{n_t}^* = \|f - p_{n_t}^*\|$ .

Table 1. Minimum number of points for exact integration for standard Gauss rule vs. Clenshaw-Curtis rule.

Degree $m$	2 elements		3 elements		4 elements	
	Gauss	CC	Gauss	CC	Gauss	CC
1	2	3	3	4	4	5
2	4	5	6	7	8	9
3	4	7	6	10	8	13
4	6	9	9	13	12	17

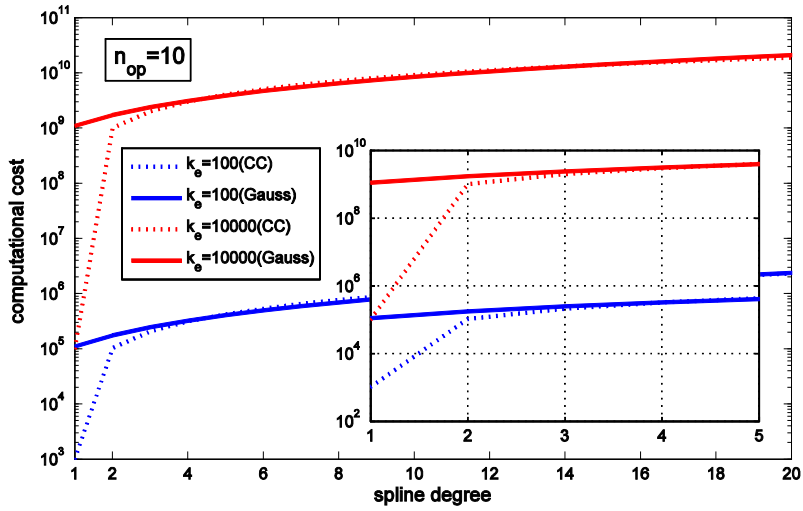


Figure 4a.

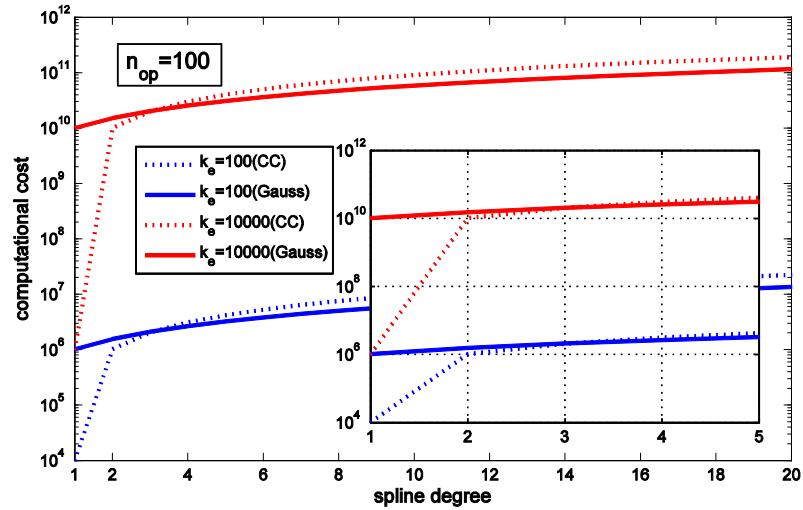


Figure 4b.

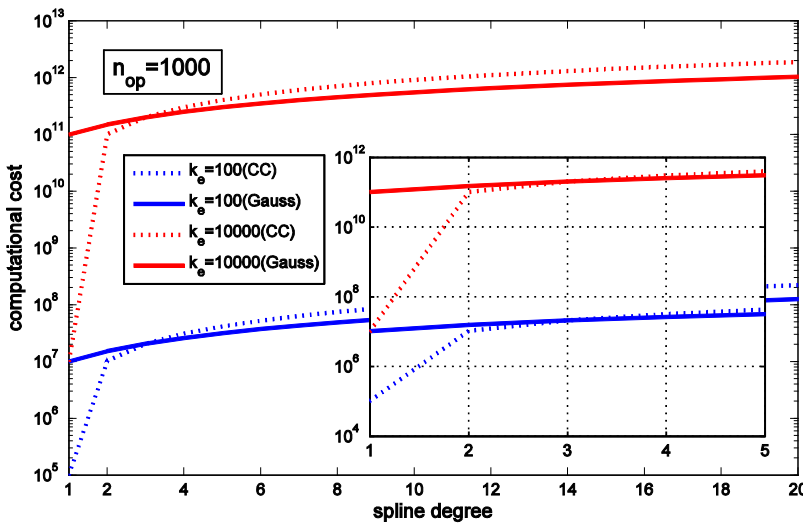


Figure 4c.

Figure 4 The computational cost (operation times) of CC and Gauss versus the spline degree in cases of different  $n_{op}$ : a,  $n_{op}=10$ ; b,  $n_{op}=100$ ; c,  $n_{op}=1000$ .

#### 4. Numerical applications and results

To verify the rules presented in the previous section, the Poisson' problem and plane strain elasticity are numerically solved. NURBS based physical domains used as test samples to investigate the applicability, accuracy and efficiency of Clenshaw-Curtis rules in IGA. The geometric parameterization  $\mathcal{P}$  is defined as

$$\mathcal{P}: \hat{\Omega} \rightarrow \Omega, \quad \hat{\mathbf{x}} \rightarrow \mathbf{x} = F(\hat{\mathbf{x}}) := \sum_{i \in \mathbf{I}} N_{i,p}(\hat{\mathbf{x}}) C_i \quad (20)$$

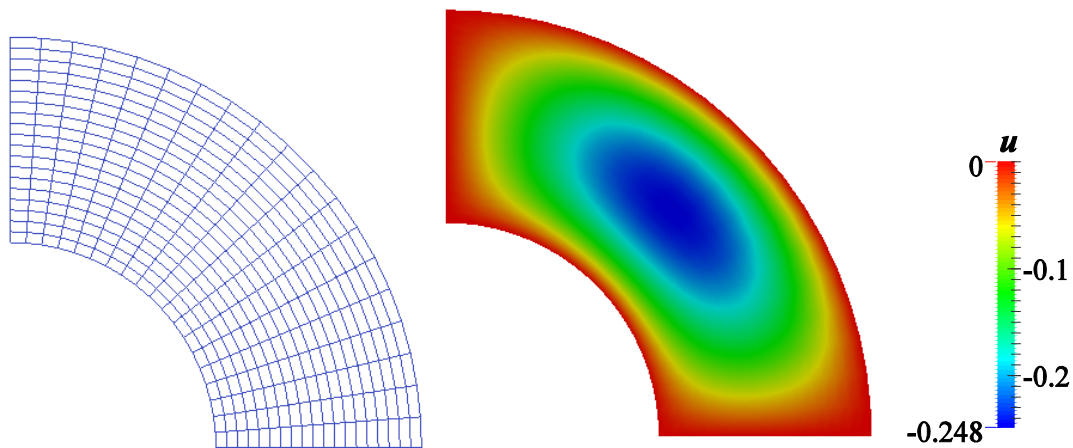
where  $\hat{\Omega}$  is the parametric domain described through the parameterization  $F$ ,  $\Omega$  is the physical domain.  $N_{i,p}(\hat{\mathbf{x}})$  is the NURBS basis function;  $C_i$  is the corresponding control points.

##### 4.1 Poisson's problems on a quarter of annulus with Dirichlet boundary conditions

A Poisson's problem is defined in a single 2D NURBS patch which forms a quarter of annulus (see Figure 5). The domain has an internal radius of 1 and an external radius of 2. For simplicity, homogeneous Dirichlet boundary conditions are imposed on the whole boundary. The problem in their variational formulation with the source term read as

$$\int_{\Omega} \nabla u \nabla v dx = \int_{\Omega} f v dx + \int_{\Gamma_N} g v d\Gamma \quad \forall v \in H_{0,\Gamma_D}^1(\Omega)$$

$$\text{with } \begin{cases} u = 0 & \text{on } \Gamma_D \\ f = \frac{(8 - 9\sqrt{x^2 + y^2}) \sin(2 \arctan \frac{y}{x})}{x^2 + y^2} \end{cases} \quad (21)$$



**Figure 5. Solution of the Poisson' problem with Clenshaw-Curtis rules. On the left: geometry sketch (with elements) of the domain  $\Omega$ . On the right: contours of the solution.**

In this case, the exact solution is given by

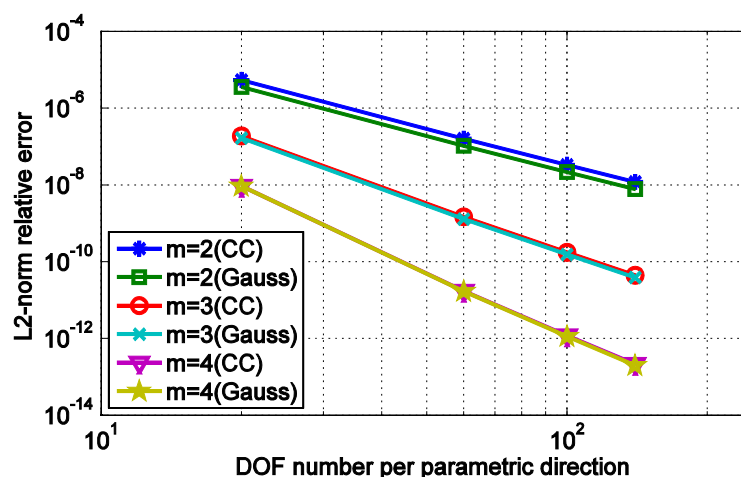
$$u = (x^2 + y^2 - 3\sqrt{x^2 + y^2} + 2) \sin(2 \arctan \frac{y}{x}) \quad (22)$$

Such a problem is approximated by a standard isogeometric Galerkin method for basis function degrees  $m$  ranging from 2 to 4 in both parametric directions. In each direction, the higher regularity  $q$  of  $m-1$  is used.

The Clenshaw-Curtis (CC for short in figures and tables) rule is compared with a standard element-wise Gauss quadrature in Figure 6 by giving the convergence curves for the  $L^2$ -norm of the relative error with respect to (w.r.t.) the exact solution. Such convergence curves are plotted with control points (or degree of freedoms) varying from 20 to 140 in each parametric direction and full quadrature is used. It can be seen that differences between the two kinds of errors are negligible w.r.t. the approximation error. The convergence rate of the Clenshaw-Curtis is almost the same as the Gauss. Besides, the order of accuracy increases evidently with the increment of the degree. The detailed error information is listed in Table 2.

**Table 2. The results of the Poisson's problem:  $L^2$ -norm of the relative error w.r.t. the exact solution in the case of different number of control points for the Clenshaw-Curtis and standard element-wise Gauss quadrature. Full quadratures are evaluated. See section 4.1 for detailed computation setup.**

DOF number per direction	Rule	Spline degree of 2	Spline degree of 3	Spline degree of 4
20	CC	5.431343771566488e-06	1.894233786199076e-07	9.430382212269052e-09
	Gauss	3.572914474223246e-06	1.680628530304760e-07	9.366661545768601e-09
60	CC	1.613192377810281e-07	1.471198622014350e-09	1.742675787352948e-11
	Gauss	1.054458637452052e-07	1.299645295684350e-09	1.678912535123116e-11
100	CC	3.342719810125322e-08	1.752487255178711e-10	1.203087010291692e-12
	Gauss	2.183982466163931e-08	1.547715389168774e-10	1.130145581122103e-12
140	CC	1.196987407020663e-08	4.403201788880789e-11	2.157064347426942e-13
	Gauss	7.819640850756566e-09	3.888441670882580e-11	1.978973380670083e-13



**Figure 6. Convergence curves in solving the Poisson's model for the  $L^2$ -norm of the relative error (double log-scale). Full regularity is assumed and minimum numbers of quadrature points for exact integration are used.**

The comparison between the convergence rate w.r.t. the number of quadrature points of the Clenshaw-Curtis and that of Gauss rules is plotted in Figure 7a- Figure 7c. In these three figures, convergence curves of the relative error w.r.t. the exact solution is considered with the control nets fixed at  $20 \times 20$  and basis function degrees varying from 3 to 5 respectively. It can be found that the Gauss converges faster than CC w.r.t. the same number of quadrature point. For higher degrees it is more evident. In other words, to converge to certain accuracy, the Gauss needs less quadrature points due to the “factor of 2”.

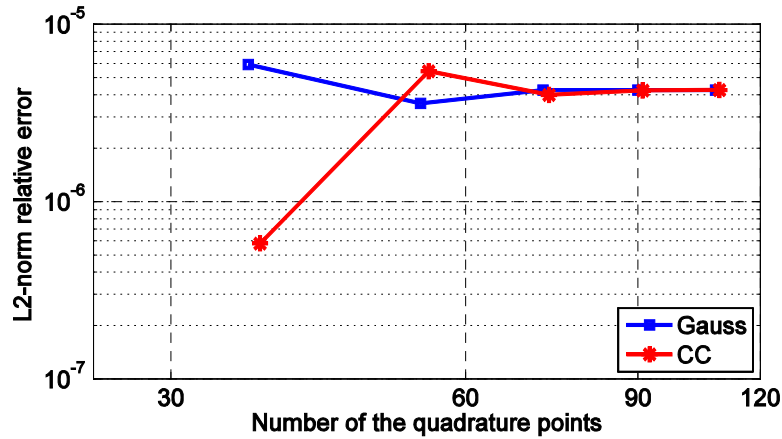


Fig.7a

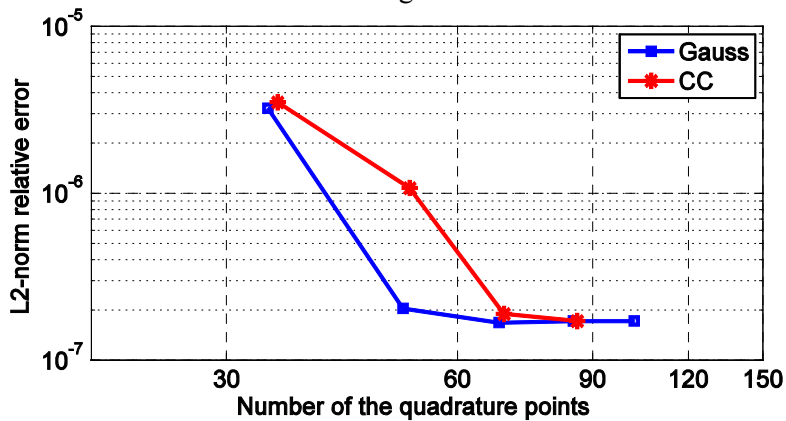


Fig.7b

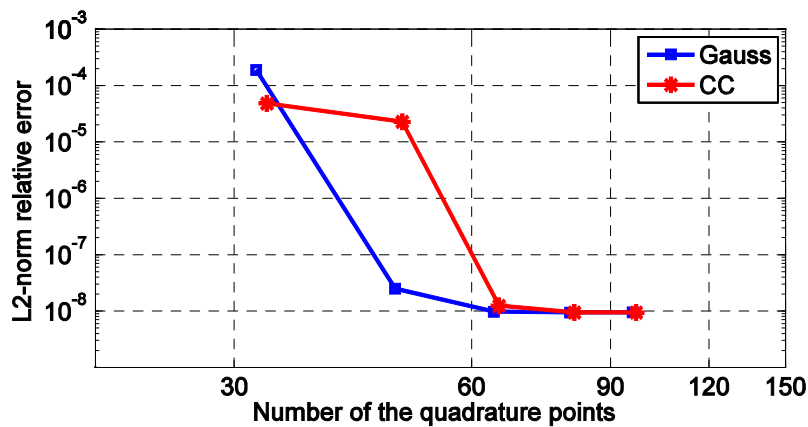


Fig.7c

**Figure 7. The computational cost of the Clenshaw-Curtis rule vs. that of Gauss rules. a. with a spline degree of 3; b. with a spline degree of 4; c. with a spline degree of 5.**

The number of quadrature points for each whole parametric direction and averaged in each element (in brackets) are shown in Table 3, respectively. The bold data is the minimum point number for exact quadrature, corresponding to Table 1. Given spline degree of 3, the minimum point number needed in Gauss and CC for exact integration are 2 and 4 respectively, corresponding total number are 34 and 52. However, the accuracy has already been achieved by 3 CC points (35 in total, which is 1 point more than the optimal Gauss.), rather than the 52 points required. Given spline degree of 4, the CC has a better accuracy than Gauss if the functions are under integrated (the point number used is less than the minimum, and thus the exact quadrature is not achieved), which can be seen from the case of 3 and 2 points for CC and Gauss, respectively. Besides, the Clenshaw-Curtis essentially never requires many more function evaluations than Gauss to converge to a prescribed accuracy [Trefethen (2008)]. In the plane strain problem which will be presented at section 4.2, we will report the similar results.

**Table 3. The results of the Poisson's problem:  $L^2$ -norm of the relative error w.r.t. the exact solution in the case of different quadrature points for the Clenshaw-Curtis and standard element-wise Gauss quadrature. See section 4.1 for detailed computation setup. Integers before and in bracket refer to the number of quadrature points in each parametric direction and in each element, respectively. The bold data relates to the minimum points for exact quadrature, corresponding to Table 1.**

Rule	Number of quadrature points	Spline degree of 2	Number of quadrature points	Spline degree of 3	Number of quadrature points	Spline degree of 4
CC	<b>37(3)</b>	<b>5.810937748390370e-07</b>	35(3)	3.507832956193244e-06	33(3)	4.815269168741496e-05
Gauss	<b>36(2)</b>	<b>5.910381033719918e-06</b>	<b>34(2)</b>	<b>3.237004970763237e-06</b>	32(2)	1.883677307485299e-04
CC	55(4)	5.431343771566488e-06	<b>52(4)</b>	<b>1.076046179389082e-06</b>	49(4)	2.267120441916495e-05
Gauss	54(3)	3.572914474223246e-06	51(3)	2.043586698351132e-07	<b>48(3)</b>	<b>2.487016794552956e-08</b>
CC	73(5)	3.984412401849476e-06	69(5)	1.894233786199076e-07	<b>65(5)</b>	<b>1.242345091306609e-08</b>
Gauss	72(4)	4.254679062985767e-06	68(4)	1.680628530304760e-07	64(4)	9.726530536491443e-09
CC	91(6)	4.221771456414136e-06	86(6)	1.719740835113939e-07	81(6)	9.430382212269052e-09
Gauss	90(5)	4.254426013745271e-06	85(5)	1.720555724776312e-07	80(5)	9.366661545768601e-09
CC	109(7)	4.254443331292910e-06	120(7)	1.719868621478127e-07	97(7)	9.400145380498706e-09
Gauss	108(6)	4.254426057907695e-06	102(6)	1.720541618939914e-07	96(6)	9.386569470718392e-09
CC	127(8)	4.254430377691922e-06	137(8)	1.720541968799450e-07	113(8)	9.390607008627601e-09
Gauss	126(7)	4.254426057908292e-06	119(7)	1.720541622130770e-07	112(7)	9.386561687283082e-09

Table 4. The results of the plane strain problem:  $L^2$ -norm of the relative error w.r.t. the exact solution in the case of different number of control points for the Clenshaw-Curtis and standard element-wise Gauss quadrature. Full quadratures are evaluated. See section 4.1 for detailed computation setup.

DOF number per direction	Rule	Spline degree of 2	Spline degree of 3	Spline degree of 4
20	CC	7.580838422425806e-05	5.709802889966272e-05	1.473215457485271e-06
	Gauss	3.486429685645471e-04	1.721340273127870e-04	2.049928251576902e-06
60	CC	6.898433383401487e-07	7.835584539592522e-07	3.435163758080126e-09
	Gauss	1.021598276698650e-05	2.387179634052311e-06	6.585591990210384e-09
100	CC	8.431769959773612e-08	1.206946575532456e-07	2.774539343390426e-10
	Gauss	2.114898963028304e-06	3.686934132431995e-07	5.763877757080713e-10
140	CC	2.140789685445826e-08	3.590101948398621e-08	5.551247202533930e-11
	Gauss	7.571314444567726e-07	1.098048534810133e-07	1.197411811057983e-10

Another finding is that the minimum CC points needed in exact quadrature lead to the highest accuracy for the degree of 2. It can be seen from Table 3 that the minimum error is obtained when the CC points number is 3, which is about 10 to the -7; however, for the increased number of quadrature points, all the errors are about 10 to the -6 including those reaching steady state. We can also see from this column of data that, the minimum CC points for degree of 2 still yield a better result than each point number case of Gauss rule, which has a power of -6 without exception. This phenomenon reappeared in the benchmark problem of plane strain problem (in section 4.2).

#### 4.2 Plane strain problem with Dirichlet boundary conditions

As another example, we define a plane strain problem which is linear elastic and isotropic in a two dimensional square ( $1 \times 1$ ) region. For comparison, both of the quadrature methods are used in the solution of the problem. The problem in its variational formulation is expressed in eq.(23). Again for simplicity, homogeneous Dirichlet boundary conditions are imposed on the whole boundary and the external force term  $\mathbf{f}$  is defined in eq.(23).

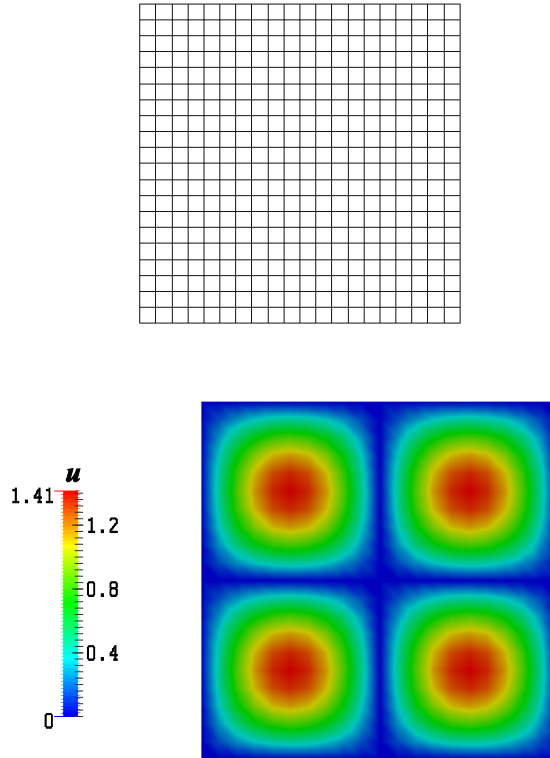
$$\begin{aligned}
 & \text{Find } \mathbf{u} \in V = \left( H_{0, \Gamma_D}^1(\Omega) \right)^2 \text{ such that} \\
 & \int_{\Omega} (\lambda \nabla(\mathbf{u}) \nabla(\mathbf{v}) + 2\mu \varepsilon(\mathbf{u}) : \varepsilon(\mathbf{v})) = \int_{\Omega} \mathbf{f} \cdot \mathbf{v} + \int_{\Gamma_N} \mathbf{g} \cdot \mathbf{v} \quad \forall \mathbf{v} \in V \quad (23) \\
 & \text{with } \begin{cases} \mathbf{u} = 0 & \text{on } \Gamma_D \\ f_x = f_y = -4\pi^2 [-3\mu + \lambda \sin(2\pi x) \sin(2\pi y) \\ \quad + (\mu + \lambda) \cos(2\pi x) \cos(2\pi y)] \end{cases}
 \end{aligned}$$

where  $\mu$  and  $\lambda$  are the Lamé parameters of the material. The geometry of the domain and the displacement magnitude for a material with Young modulus  $E = 1$  and Poisson ratio  $\nu = 0.3$  is plotted in Figure 8 The exact solution of this model is shown in eq.(24)



$$u_x = u_y = \sin(2\pi x)\sin(2\pi y) \tag{24}$$

In order to compare it with the presented Poisson’s equation, this problem is still approximated by a standard isogeometric Galerkin method for basis function degrees  $m$  ranging from 2 to 4 in both parametric directions. In each direction, the higher regularity  $q$  of  $m-1$  is used.



**Figure 8. Solution of the plane strain problem. The geometry sketch of the domain with element (top) and the contour plot of displacements (bottom).**

The convergence curves for the  $L^2$ -norm of the relative error with respect to (w.r.t.) the exact solution for both rules are shown in Figure 9. Similar to the conclusion aforementioned, the Clenshaw-Curtis has almost the same convergence rate as the Gauss rules. The comparison between the convergence rate w.r.t. the number of quadrature points of the Clenshaw-Curtis and that of Gauss rules is showed in Table 5 and not plotted. Similarly, the Gauss’s convergence is faster than CC w.r.t. the same number of quadrature point and more evidently for higher degrees; it not necessarily needs 4 CC points (52 in total) for a degree of 3, actually, 3 points (35 in total) are enough for exact quadrature; the 3 points CC has a better accuracy than 2 points Gauss for degree of 4. From the column of degree 2, the identical conclusion that minimum 3 CC points for exact quadrature lead to the highest accuracy has been found.

**Table 5. The results of the plane strain problem:  $L^2$ -norm of the relative error w.r.t. the exact solution in the case of different quadrature points for the Clenshaw-Curtis and standard element-wise Gauss quadrature. See section 4.2 for detailed computation setup. Integers before and in bracket refer to the number of quadrature points in each parametric direction and in each element,**

respectively. The bold data relates to the minimum points for exact quadrature, corresponding to Table 1.

Rule	Number of quadrature points	Spline degree of 2	Number of quadrature points	Spline degree of 3	Number of quadrature points	Spline degree of 4
CC	<b>37(3)</b>	<b>7.580838422425806e-05</b>	35(3)	1.856112366410516e-04	33(3)	7.889085258275995e-04
Gauss	<b>36(2)</b>	<b>3.486429685645471e-04</b>	<b>34(2)</b>	<b>1.721340273127870e-04</b>	32(2)	0.002869708867170
CC	55(4)	3.213220416249692e-04	<b>52(4)</b>	<b>5.709802889966272e-05</b>	49(4)	3.905908393191598e-04
Gauss	54(3)	2.159149300119856e-04	51(3)	2.122049676131301e-05	<b>48(3)</b>	<b>2.049928251576902e-06</b>
CC	73(5)	2.391125067420231e-04	69(5)	1.973816028233426e-05	<b>65(5)</b>	<b>1.473215457485271e-06</b>
Gauss	72(4)	2.544164579534698e-04	68(4)	1.763946743470982e-05	64(4)	1.508026371023410e-06
CC	91(6)	2.525513107997929e-04	86(6)	1.802195523770241e-05	81(6)	1.474188307465811e-06
Gauss	90(5)	2.544061479534598e-04	85(5)	1.802810705630391e-05	80(5)	1.458299097251400e-06
CC	109(7)	2.544068533940345e-04	103(7)	1.800176949085848e-05	97(7)	1.463032143415990e-06
Gauss	108(6)	2.544061494631728e-04	102(6)	1.802801880799568e-05	96(6)	1.461314539284407e-06
CC	127(8)	2.544063254948461e-04	120(8)	1.802145721232466e-05	113(8)	1.461841366010776e-06
Gauss	126(7)	2.544061494630725e-04	119(7)	1.802801881926631e-05	112(7)	1.461313960405898e-06

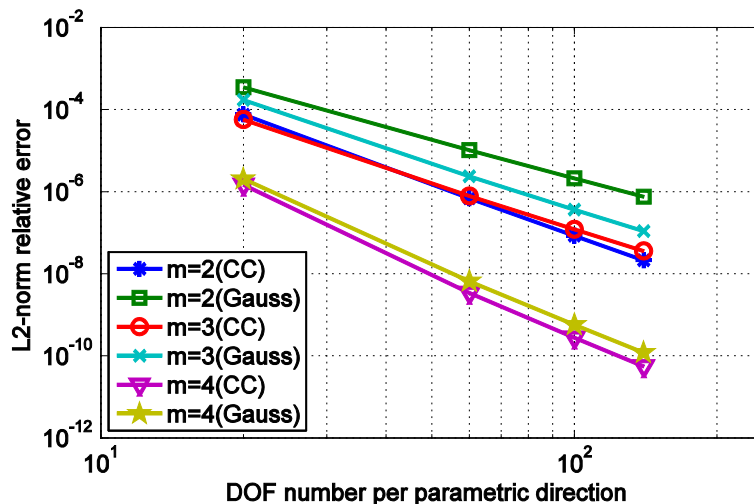


Fig.9 Convergence curves in solving the plane strain model for the  $L^2$ -norm of the relative error (double log-scale). Full regularity is assumed and minimum numbers of quadrature points for exact integration are used.

## 5. Conclusions

We introduced the Clenshaw-Curtis quadrature into the IGA scheme and compared its accuracy and efficiency with that of the optimal standard Gauss rule. We found that for exact quadrature and higher spline degrees ( $m \geq 3$ ), the Gauss has advantages in both accuracy and efficiency due to its “factor of 2”; while for under integration (points number is less than the minimum required), the Clenshaw-Curtis is better. For lower spline degrees ( $m \leq 3$ ), the exact quadrature can be achieved for the Clenshaw-Curtis rule when the functions are under integrated and thus it has an improved efficiency. Moreover, the degree of 2 requires the least points to obtain the highest accuracy for Clenshaw-Curtis rule.

Considering the overall operations needed in isogeometric approximation (Galerkin method is used in this paper), the Gauss also proves its higher efficiency in solving problems with high spline degrees. Considering all the indefinite factors due to the externals (such as PDE types, the programmer’s experience, etc.), we quantify those factors as a coefficient  $n_{op}$  and take values from 10 to  $1 \times 10^3$ . However, all these values yield a same result: for lower spline degrees ( $m \leq 3$ ), the Clenshaw-Curtis has a better efficiency than the Gauss rules.

## References

- [1] Hughes, T. J., Cottrell, J. A., & Bazilevs, Y. (2005). Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer methods in applied mechanics and engineering*, 194(39), 4135-4195.
- [2] Cottrell, J. A., Hughes, T. J., & Bazilevs, Y. (2009). *Isogeometric analysis: toward integration of CAD and FEA*. John Wiley & Sons.
- [3] Bazilevs, Y., Calo, V. M., Zhang, Y., & Hughes, T. J. (2006). Isogeometric fluid–structure interaction analysis with applications to arterial blood flow. *Computational Mechanics*, 38(4-5), 310-322.
- [4] Bazilevs, Y., Hsu, M. C., & Scott, M. A. (2012). Isogeometric fluid–structure interaction analysis with emphasis on non-matching discretizations, and with application to wind turbines. *Computer Methods in Applied Mechanics and Engineering*, 249, 28-41.
- [5] Bazilevs, Y., Hsu, M. C., Kiendl, J., Wüchner, R., & Bletzinger, K. U. (2011). 3D simulation of wind turbine rotors at full scale. Part II: Fluid–structure interaction modeling with composite blades. *International Journal for Numerical Methods in Fluids*, 65(1 - 3), 236-253.
- [6] Kiendl, J., Bletzinger, K. U., Linhard, J., & Wüchner, R. (2009). Isogeometric shell analysis with Kirchhoff–Love elements. *Computer Methods in Applied Mechanics and Engineering*, 198(49), 3902-3914.
- [7] Lipton, S., Evans, J. A., Bazilevs, Y., Elguedj, T., & Hughes, T. J. (2010). Robustness of isogeometric structural discretizations under severe mesh distortion. *Computer Methods in Applied Mechanics and Engineering*, 199(5), 357-373.
- [8] Benson, D. J., Bazilevs, Y., Hsu, M. C., & Hughes, T. J. R. (2011). A large deformation, rotation-free, isogeometric shell. *Computer Methods in Applied Mechanics and Engineering*, 200(13), 1367-1378.
- [9] Anders, D., Weinberg, K., & Reichardt, R. (2012). Isogeometric analysis of thermal diffusion in binary blends. *Computational Materials Science*, 52(1), 182-188.
- [10] Qian, X. (2010). Full analytical sensitivities in NURBS based isogeometric shape optimization. *Computer Methods in Applied Mechanics and Engineering*, 199(29), 2059-2071.
- [11] Buffa, A., Sangalli, G., & Vázquez, R. (2010). Isogeometric analysis in electromagnetics: B-splines approximation. *Computer Methods in Applied Mechanics and Engineering*, 199(17), 1143-1152.
- [12] Auricchio, F., Calabro, F., Hughes, T. J. R., Reali, A., & Sangalli, G. (2012). A simple algorithm for obtaining nearly optimal quadrature rules for NURBS-based isogeometric analysis. *Computer Methods in Applied Mechanics and Engineering*, 249, 15-27.
- [13] Hughes, T. J., Reali, A., & Sangalli, G. (2010). Efficient quadrature for NURBS-based isogeometric analysis. *Computer methods in applied mechanics and engineering*, 199(5), 301-313.
- [14] Gentleman, W. M. (1972a). Implementing Clenshaw-Curtis quadrature, I methodology and

- experience. *Communications of the ACM*, 15(5), 337-342.
- [15] Gentleman, W. M. (1972b). Implementing Clenshaw-Curtis quadrature, II computing the cosine transformation. *Communications of the ACM*, 15(5), 343-346.
- [16] Calabrò, F., & Esposito, A. C. (2009). An evaluation of Clenshaw–Curtis quadrature rule for integration wrt singular measures. *Journal of computational and applied mathematics*, 229(1), 120-128.
- [17] Trefethen, L. N. (2008). Is Gauss quadrature better than Clenshaw-Curtis?. *SIAM review*, 50(1), 67-87.
- [18] Clenshaw, C. W., & Curtis, A. R. (1960). A method for numerical integration on an automatic computer. *Numerische Mathematik*, 2(1), 197-205.
- [19] Piegler, L., & Tiller, W. (1997). *The NURBS book*. 1997. Monographs in Visual Communication.
- [20] Davis, P. J., & Rabinowitz, P. (2007). *Methods of numerical integration*. Courier Corporation.
- [21] Waldvogel, J. (2006). Fast construction of the Fejér and Clenshaw–Curtis quadrature rules. *BIT Numerical Mathematics*, 46(1), 195-202.

# Fluid Models and Parameter Sensitivities: Computations and Applications

Lisa Davis<sup>1</sup>, Monika Neda<sup>2</sup>, Fran Pahlevani<sup>3</sup>, and Jiajia Waters<sup>4</sup>

<sup>1</sup>Department of Mathematics, Montana State University, Bozeman, MT 59717

<sup>2</sup>Department of Mathematics, University of Nevada, Las Vegas, NV 89154

<sup>3</sup>Division of Science and Engineering, Penn State Abington, Abington, PA 19001

<sup>4</sup>T-3 Fluid and Solid Mechanics Division, Los Alamos National Laboratory, NM 87544

## Abstract

Fluid models were developed as an alternative to the Navier-Stokes equations to avoid computational complexity especially in case of turbulent flows. Model errors due to the sensitivity of a model to user-elected parameters become an immediate concern. Quantifying this error and assessing the reliability of the model given a parameter value are essential to understanding and using model predictions within an engineering design process. This paper presents an overview of sensitivity computations of three fluid models namely the Eddy Viscosity Model, the Leray-Alpha Model, and the Time Relaxation Model to the variations of different model parameters. The 2D Cavity problem is used to numerically illustrate the application of sensitivity computations in identifying the range of parameter values for which the fluid model can be considered a reliable approximation. In addition, testing on the 2D flow around a cylinder, our numerical results support the idea that sensitivity information can incorporate the effects of unresolved scales on flow functionals that leads to an improved estimation. Investigations of flow sensitivity for non-isothermal Stokes flow are presented as well.

## 1 Introduction

Numerous types of fluid flows are formulated by Navier-Stokes Equations (NSE) based on the fluid velocity and pressure. When we solve these equations numerically, the flow velocity consists of different scales and eddies. Numerical simulations of NSE are used for two major purposes. One is to understand the physical mechanism of the fluid and the other is to predict the flow characteristics in applications. Both cases require a numerical simulation producing data of very high accuracy. Since the precision of generated data depends on the level of selected resolution, for the best possible numerical result the simulation has to take into account all the space-time scales in the fluid dynamics. As known by Kolmogorov's law, the required number of mesh points in space per time step in a three dimensional flow is related to the Reynolds number,  $Re$ , and it is  $O(Re^{9/4})$ . This leads to the fact that fluid flows with large enough  $Re$  are expensive simulations regarding both the required storage and the running time. Technically, Direct Numerical Simulation (DNS) is computationally infeasible

especially in the case of turbulent flows when the range of velocity scales is very large. As an alternative, regularization models of NSE were developed to allow for computational efficiency in case of high  $Re$  numbers. These models are mostly based on a technique that uses a filtering procedure on NSE, ultimately solving the equations for only large scale velocities. Removing some scales from the fluid system not only affects the accuracy of the numerical data but also their reliability leading to the major issue of model errors and uncertainty in model predictions. This raises concerns particularly in applications where important decisions are made; see [34, 21, 32, 3]. The model reliability becomes an issue especially when the fluid model is sensitive with respect to the variation of a user-elected parameter. Such parameters appear to be inevitable in the process of modeling. Filter length scale is a simple example to be named. To this end, parameter sensitivity analysis is considered a technique to assess the reliability of the computed flow solution using a fluid model. Sensitivity analysis of a flow system is defined as the computation of the derivative of flow variables with respect to model parameters upon which the response of the flow system explicitly and/or implicitly depends [5, 17]. A natural approach to obtain flow sensitivity known as Continuous Sensitivity Equation Method (CSEM) is to form a continuous equation for the designated sensitivity and then to numerically solve it. CSEM has been used in sensitivity calculations of flows with respect to various flow-related parameters; see for example [4, 6, 15].

This paper provides a summary first on the use of CSEM in computing sensitivity of three specific fluid models, namely a subgrid Eddy Viscosity Model (EVM), Leray-Alpha Model (LAM), and Time Relaxation Model (TRM) with respect to a model parameter, second the use of sensitivity information in quantifying the model reliability, and last on the application of the sensitivity computation in improving flow functionals. An extensive study of these topics are presented by authors in [31, 28, 29]. Herein, we also present the investigations of flow sensitivity of the time relaxation parameter for the incompressible non-isothermal Stokes equations.

## 2 Continuous Sensitivity Equation

In this section we introduce the equations for EVM, LAM, and TRM, and derive the sensitivity equations with respect to variations of a model parameter for each. In the first two models, the parameter of consideration is the filter length scale and in the latter, it is the time relaxation coefficient whose value specifies how strongly the growth of fluctuations are truncated.

The discussed subgrid EVM in this paper was first introduced by Layton [23]. The analysis and numerical computations of two first-order semi-implicit schemes for EVM and NSE are presented in [10]. An error analysis of this model using discontinuous polynomial approximations can be found in [22]. An explicit sensitivity study of this model with application to quantifying model reliability is given in [31].

In 1934, for the first time Leray introduced a regularization of NSE on the nonlinear term using a Gaussian filter and proved the existence and uniqueness of strong solutions to his model [26, 27]. In a reexamination of the Leray model, the Gaussian filter was replaced by a differential filter and the theory and computations of this new model, LAM, were studied by different group of scientists [8, 9, 19, 36, 25, 7]. A computational study on the sensitivity of LAM with respect to the filter width is presented in [28].

TRM was originally developed from regularized Chapman-Enskog expansion of conserva-

tion laws by Rosenau [33], Schochet and Tadmor [35]. The given TRM in this paper uses the van Cittert deconvolution (in the regularization term) proposed by Stolz, Adams and Kleiser [1] who also extensively tested the model on compressible flows with shocks and turbulent flows [1, 2, 13]. An analysis of a discrete numerical scheme using a continuous finite element method can be found in [13]. Both EVM and TRM regularize the NSE by adding a stabilization term; however, LAM applies a regularization to the non-linear term in NSE. EVM is obtained by applying a filtering operator to the NSE that is an  $L^2$ -orthogonal projection while the other two fluid models use a differential filter.

In the following equations for EVM, LAM, and TRM,  $\mathbf{u}$  and  $p$  represent velocity and pressure respectively,  $\mathbf{f}$  is the body force, and  $\nu > 0$  is the kinematic viscosity, which is inversely proportional to  $Re$ . In the corresponding sensitivity equations for all models,  $\mathbf{s}$  and  $r$  represent the velocity and pressure sensitivities with respect to the designated parameter, and  $\bar{\mathbf{u}}$  stands for the average velocity with  $\mathbf{w}$  being its sensitivity. In all the equations,  $\Omega$  is considered to be a bounded, simply connected two- or three-dimensional domain with polygonal boundary  $\partial\Omega$ .

**Definition 2.1.** *Let  $\mathbf{u}$ ,  $\bar{\mathbf{u}}$ , and  $p$  be the flow variables as velocity, average velocity, and pressure, respectively. The sensitivity of these variables to variations of a designated model parameter  $\eta$  is defined to be the flow variable derivative with respect to  $\eta$ .*

$$\mathbf{s} = \frac{\partial \mathbf{u}}{\partial \eta}, \mathbf{w} = \frac{\partial \bar{\mathbf{u}}}{\partial \eta}, \text{ and } r = \frac{\partial p}{\partial \eta}$$

Note that all the models are for incompressible flows with zero boundary condition and initial condition given as,

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0, \quad \text{in } \Omega \times [0, T] \\ \mathbf{u} &= 0, \quad \text{on } \partial\Omega \times [0, T] \\ \mathbf{u}(x, 0) &= \mathbf{u}_0(x), \quad \text{in } \Omega. \end{aligned}$$

Assuming that the velocity initial condition is independent of the designated model parameter, the sensitivity of the above equations appears as the following,

$$\begin{aligned} \nabla \cdot \mathbf{s} &= 0, \quad \text{in } \Omega \times [0, T] \\ \mathbf{s} &= 0, \quad \text{on } \partial\Omega \times [0, T] \\ \mathbf{s}(x, 0) &= 0, \quad \text{in } \Omega. \end{aligned}$$

In the following, we introduce the equations for EVM, LAM, and TRM.

## 2.1 Eddy Viscosity Model

The EVM over the time interval  $[0, T]$  is outlined as following

$$\mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p - \alpha \nabla \cdot (\nabla \mathbf{u} - \bar{\mathbf{u}}) = \mathbf{f}, \quad \text{in } \Omega \times (0, T]. \quad (2.1)$$

Here for any  $\mathbf{v} \in (L^2(\Omega))^d$ ,  $d = 2$  or  $3$ ,  $\bar{\mathbf{v}} = P(\nabla \mathbf{v})$ , where  $P : L^2(\Omega) \rightarrow L$  is an  $L^2$ -orthogonal projection, defined on a chosen subspace of  $L^2(\Omega)$  [23], denoted by  $L$ . The parameter  $\alpha$  known as the eddy viscosity coefficient, corresponds to the filter length scale. Therefore its values vary between 0 and 1 with  $\alpha = 0$  corresponding to the Navier-Stokes equations.

For simplicity, it is assumed that the  $L^2$ -orthogonal projection  $P$  is differentiable with respect to parameter  $\alpha$ . Since this operator is a linear operator, using the chain rule, it can be easily shown that the operator  $P$  commutes with differentiation with respect to  $\alpha$ . Implicitly differentiating (2.1) with respect to  $\alpha$  produces the following sensitivity equation,

$$\mathbf{s}_t + \mathbf{u} \cdot \nabla \mathbf{s} + \mathbf{s} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{s} + \nabla r - \alpha \nabla \cdot (\nabla \mathbf{s} - \bar{\mathbf{s}}) = \nabla \cdot (\nabla \mathbf{u} - \bar{\mathbf{u}}), \quad \text{in } \Omega \times (0, T]. \quad (2.2)$$

## 2.2 Leray-Alpha Model

The regularization of the NSE by LAM is formulated as follows

$$\mathbf{u}_t + \bar{\mathbf{u}} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p = \mathbf{f}, \quad \text{in } \Omega \times (0, T] \quad (2.3)$$

where  $\bar{\mathbf{u}}$  is obtained from the differential filter,

$$\begin{aligned} -\alpha^2 \Delta \bar{\mathbf{u}} + \bar{\mathbf{u}} &= \mathbf{u}, \quad \text{in } \Omega \\ \bar{\mathbf{u}} &= 0, \quad \text{on } \partial\Omega. \end{aligned} \quad (2.4)$$

Sensitivity equations of LAM with respect to variations of parameter  $\alpha$  are given as

$$\mathbf{s}_t + \mathbf{w} \cdot \nabla \mathbf{u} + \bar{\mathbf{u}} \cdot \nabla \mathbf{s} + \nabla q - \nu \Delta \mathbf{s} = \mathbf{f}, \quad \text{in } \Omega \times (0, T] \quad (2.5)$$

In (2.5),  $\mathbf{w}$  is obtained from the sensitivity equation of the differential filter (2.4),

$$\begin{aligned} -\alpha^2 \Delta \mathbf{w} + \mathbf{w} &= -\frac{2}{\alpha}(\mathbf{u} - \bar{\mathbf{u}}) + \mathbf{s}, \quad \text{in } \Omega \\ \mathbf{w} &= 0, \quad \text{on } \partial\Omega \end{aligned} \quad (2.6)$$

## 2.3 Time Relaxation Model

Similar to EVM, TRM consists of the Navier-Stokes equations with the addition of a stabilization term to the momentum equation and it is defined by

$$\mathbf{u}_t + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nu \Delta \mathbf{u} + \chi(\mathbf{u} - G_N \bar{\mathbf{u}}) = \mathbf{f}, \quad \text{in } \Omega \times (0, T] \quad (2.7)$$

Here,  $\bar{\mathbf{u}}$  stands for an averaged function of  $\mathbf{u}$  by filter width  $\alpha$  satisfying the differential filter given in (2.4). The operator  $G_N$  is the continuous van Cittert deconvolution operator, where  $N$  denotes the deconvolution order, and for any  $\mathbf{v} \in H_0^1(\Omega)$  is defined as following [24],

$$G_N \mathbf{v} := \sum_{n=0}^N (I - G)^n \mathbf{v}.$$

For the zeroth and first order of deconvolution, we have the van Cittert deconvolution as  $G_0 \mathbf{v} = \mathbf{v}$ , and  $G_1 \mathbf{v} = 2\mathbf{v} - \bar{\mathbf{v}}$ , respectively. As discussed in [12], higher order of deconvolution



produces more accurate approximations but it becomes costly in terms of computational time. All the studies in this paper are carried out for the fundamental case when the order of deconvolution is zero, i.e.  $N = 0$ . The action of the term  $\chi(\mathbf{u} - G_N \bar{\mathbf{u}})$  is to drive fluctuations of size lower than  $O(\alpha)$  to zero as  $t \rightarrow \infty$ .

Differentiating TRM implicitly with respect to the parameter  $\chi$  gives the following equation for sensitivity,

$$\mathbf{s}_t + \mathbf{u} \cdot \nabla \mathbf{s} + \mathbf{s} \cdot \nabla \mathbf{u} + \nabla r - \nu \Delta \mathbf{s} + (\mathbf{u} - \bar{\mathbf{u}}) + \chi(\mathbf{s} - \mathbf{w}) = 0, \quad \text{in } \Omega \times [0, T] \quad (2.8)$$

where  $\mathbf{w}$  is the solution of the following sensitivity equation obtained by differentiating (2.4) with respect to parameter  $\chi$ ,

$$\begin{aligned} -\alpha^2 \Delta \mathbf{w} + \mathbf{w} &= \mathbf{s}, \quad \text{in } \Omega, \\ \mathbf{w} &= 0, \quad \text{on } \partial\Omega. \end{aligned} \quad (2.9)$$

### 3 The Algorithms and Discretizations

This section is devoted to introducing the basis for deriving a finite element approximation of  $\mathbf{u}$  in (2.1), (2.3)-(2.6), and (2.7) as well as  $\mathbf{s}$  in (2.2), (2.5)-(2.6), and (2.8)-(2.9). As it can be seen in (2.2), (2.5)-(2.6), and (2.8)-(2.9), velocity  $\mathbf{u}$ , and its average  $\bar{\mathbf{u}}$  appear in the sensitivity equations. Therefore in computing sensitivities one needs to couple the sensitivity equations with the corresponding model. Given  $\mathbf{u}$ , and  $\bar{\mathbf{u}}$  the sensitivity equations are linear equations. Therefore sensitivity can be calculated in a very inexpensive manner once the numerical method for computing  $\mathbf{u}$  in each model is constructed. The bulk of the work for the computation is in the implementation of equations (2.1), (2.3), and (2.7). Once that is accomplished, the incorporation of the sensitivity equations in (2.2), (2.5), and (2.8) into a numerical algorithm that computes both  $\mathbf{u}$  and  $\mathbf{s}$  is straightforward. As one can easily observe, all the sensitivity data structures are virtually the same or very similar to one computed from the fluid model.

Next we introduce the notation for function spaces used in finite element theory,

$$\begin{aligned} \mathbf{X}^h \subset X &= H_0^1(\Omega) := \{\mathbf{v} \in H^1(\Omega) : \mathbf{v}|_{\partial\Omega} = 0\}, \\ Q^h \subset Q &= L_0^2(\Omega) := \{q \in L^2(\Omega) : \int_{\Omega} q = 0\}. \end{aligned}$$

For the treatment of convective and diffusive terms in the variational formulation of equations, we use the following bilinear and trilinear forms.

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= (\nabla \mathbf{u}, \nabla \mathbf{v}) \\ b(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= (\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{w}) \\ b^*(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{v}, \mathbf{w}) - \frac{1}{2}(\mathbf{u} \cdot \nabla \mathbf{w}, \mathbf{v}). \end{aligned}$$

Next we apply the following classical steps to EVM, LAM, TRM, and their sensitivity equations.

- *Variational Formulation:* The equations (2.1)- (2.2), (2.3)-(2.6), and (2.7)-(2.9) are reformulated in a weak form after multiplication by a suitable set of test functions,  $\mathbf{v} \in X$  and  $\lambda \in Q$ , and performing an integration upon the domain. At this stage, the integration by parts is used to reduce the order of differentiation for solutions,  $\mathbf{u}$  and  $\mathbf{s}$ .
- *Discretization in Space:* Let  $h \in (0, 1]$ , tending to zero, be the spatial mesh size, then  $V^h = \{\mathbf{v} \in X^h : (\lambda, \nabla \cdot \mathbf{v}) = 0, \text{ for all } \lambda \in Q^h\}$  is a finite dimensional subspace of  $X^h$ . Since  $(V^h, Q^h)$  fulfills the inf-sup or Babuska-Brezzi stability condition, by selecting the test functions from these spaces the pressure  $p_h$  can be eliminated from the system in its discrete form; see [16].
- *Discretization in Time:* We start with partitioning the time interval  $[0, T]$  into  $N$  sub-intervals  $[t^n, t^{n+1}]$  of length  $\Delta t = \frac{T}{N}$ . Then at each time level  $t^n$ , an approximation to  $\mathbf{u}$  and  $\mathbf{s}$ , denoted by  $\mathbf{u}_h^n$  and  $\mathbf{s}_h^n$  respectively, are obtained.

In the resulting equations,  $\mathbf{u}_h$  and  $\mathbf{s}_h$ , i.e. the velocity variables, are solved in  $V^h$  and their test functions belong to that space as well, while the pressure variables with their test functions are in  $Q^h$ .

### 3.1 Eddy Viscosity Model

Here we specifically explain how the stabilization term in EVM and its sensitivity are estimated in our calculations. By definition  $\bar{\mathbf{v}}$  is an  $L^2$ -orthogonal projection of  $\nabla \mathbf{v}$  onto  $L$ , therefore it can be obtained by the following equation

$$(\nabla \mathbf{v} - \bar{\mathbf{v}}, l) = 0, \quad \forall \mathbf{v} \in (L^2(\Omega))^d, l \in L. \quad (3.1)$$

In the spatial discretization form of (3.1), a multiscale spatial discretization is applied. Let  $h$  and  $H$  denote two different mesh widths ( $h < H$ ). Then the space  $L^H \subset L^2(\Omega)^{d \times d}$ ,  $d = 2$  or  $3$ , is considered as the space of large scales of the velocity that are numerically solved by EVM since  $H$  represents the coarse mesh size.

In the fully discrete form of EVM, a semi-implicit numerical scheme is applied. The convective term is computed using a backward-forward time-stepping method. Thus the equation reads as: Given  $\mathbf{u}_h^n$ , we seek  $\mathbf{u}_h^{n+1}$  satisfying

$$\begin{aligned} (\nabla \mathbf{u}_h^n - \bar{\mathbf{u}}_H^n, l) &= 0, \\ \frac{1}{\Delta t}(\mathbf{u}_h^{n+1} - \mathbf{u}_h^n, \mathbf{v}) + (\nu + \alpha)a(\mathbf{u}_h^{n+1}, \mathbf{v}) + b(\mathbf{u}_h^n, \mathbf{u}_h^{n+1}, \mathbf{v}) - \alpha(\bar{\mathbf{u}}_H^n, \nabla \mathbf{v}) &= (f^{n+1}, \mathbf{v}). \end{aligned} \quad (3.2)$$

Similarly, given  $\mathbf{s}_h^n$ , we find  $\mathbf{s}_h^{n+1}$  such that

$$\begin{aligned} (\nabla \mathbf{s}_h^n - \bar{\mathbf{s}}_H^n, l) &= 0, \\ \frac{1}{\Delta t}(\mathbf{s}_h^{n+1} - \mathbf{s}_h^n, \mathbf{v}) + (\nu + \alpha)a(\mathbf{s}_h^{n+1}, \mathbf{v}) + b(\mathbf{s}_h^{n+1}, \mathbf{u}_h^{n+1}, \mathbf{v}) + b(\mathbf{u}_h^{n+1}, \mathbf{s}_h^{n+1}, \mathbf{v}) \\ - \alpha(\bar{\mathbf{s}}_H^n, \nabla \mathbf{v}) &= (\bar{\mathbf{u}}_H^n - \nabla \mathbf{u}_h^n, \nabla \mathbf{v}). \end{aligned} \quad (3.3)$$

An extensive numerical analysis of EVM given in (3.2) and its sensitivity equation in (3.3) with further numerical tests can be found in [10, 31].

### 3.2 Leray-Alpha Model

In the discretization of the time derivative of LAM, the Crank-Nicolson method is used. For clarity in notation, we let  $\mathbf{v}(t^{n+1/2}) = \mathbf{v}((t^{n+1} + t^n)/2)$  for the continuous variable and  $\mathbf{v}^{n+1/2} = (\mathbf{v}^{n+1} + \mathbf{v}^n)/2$  for both, continuous and discrete variables.

Discrete approximation solutions of LAM, given by (2.3)-(2.4), on the time interval  $(0, T]$ , is to find  $\mathbf{u}_h^{n+1}$  and  $\bar{\mathbf{u}}_h^n$  such that

$$\begin{aligned} \alpha^2(\bar{\mathbf{u}}_h^n, \nabla \mathbf{v}) + (\bar{\mathbf{u}}_h^n, \mathbf{v}) - (\mathbf{u}_h^n, \mathbf{v}) &= 0, \\ \frac{1}{\Delta t}(\mathbf{u}_h^{n+1} - \mathbf{u}_h^n, \mathbf{v}) + \nu a(\mathbf{u}_h^{n+1/2}, \mathbf{v}) + b^*(\bar{\mathbf{u}}_h^{n+1/2}, \mathbf{u}_h^{n+1/2}, \mathbf{v}) &= (\mathbf{f}_h^{n+1/2}, \mathbf{v}). \end{aligned} \quad (3.4)$$

Discrete approximation to the sensitivity equations (2.5) and (2.6) on the time interval  $(0, T]$ , is to find  $\mathbf{s}_h^{n+1}$  and  $\mathbf{w}_h^n$  such that

$$\begin{aligned} \alpha^2(\nabla \mathbf{w}_h^n, \nabla v) + (\mathbf{w}_h^n, v) - (\mathbf{s}_h^n, \mathbf{v}) + \frac{2}{\alpha}(\mathbf{u}_h^n - \bar{\mathbf{u}}_h^n, \mathbf{v}) &= 0, \\ \frac{1}{\Delta t}(\mathbf{s}_h^{n+1} - \mathbf{s}_h^n, \mathbf{v}) + \nu a(\mathbf{s}_h^{n+1/2}, \mathbf{v}) + b^*(\mathbf{w}_h^{n+1/2}, \mathbf{u}_h^{n+1/2}, \mathbf{v}) + b^*(\bar{\mathbf{u}}_h^{n+1/2}, \mathbf{s}_h^{n+1/2}, \mathbf{v}) &= 0. \end{aligned} \quad (3.5)$$

The numerical analysis of (3.4) can be found in [25].

### 3.3 Time Relaxation Model

Similar to LAM, the Crank-Nicolson numerical scheme is applied to TRM and its sensitivity equations. Therefore we obtain the following discretized finite element variational formulations.

Find  $\mathbf{u}_h^{n+1}$  and  $\bar{\mathbf{u}}_h^n$  satisfying:

$$\begin{aligned} \alpha^2(\nabla \bar{\mathbf{u}}_h^n, \nabla \mathbf{v}) + (\bar{\mathbf{u}}_h^n, \mathbf{v}) &= (\mathbf{u}_h^n, \mathbf{v}), \\ \frac{1}{\Delta t}(\mathbf{u}_h^{n+1} - \mathbf{u}_h^n, \mathbf{v}) + \nu a(\mathbf{u}_h^{n+1/2}, \mathbf{v}) + b^*(\mathbf{u}_h^{n+1/2}, \mathbf{u}_h^{n+1/2}, \mathbf{v}), \\ + \chi(\mathbf{u}_h^{n+1/2} - \bar{\mathbf{u}}_h^{n+1/2}, \mathbf{v}) &= (\mathbf{f}_h^{n+1/2}, \mathbf{v}), \end{aligned} \quad (3.6)$$

and for the sensitivity solution, find  $\mathbf{s}_h^{n+1}$  and  $\mathbf{w}_h^n$  satisfying:

$$\begin{aligned} \alpha^2(\nabla \mathbf{w}_h^n, \nabla \mathbf{v}) + (\mathbf{w}_h^n, \mathbf{v}) &= (\mathbf{s}_h^n, \mathbf{v}), \\ \frac{1}{\Delta t}(\mathbf{s}_h^{n+1} - \mathbf{s}_h^n, \mathbf{v}) + \nu a(\mathbf{s}_h^{n+1/2}, \mathbf{v}) + b^*(\mathbf{s}_h^{n+1/2}, \mathbf{u}_h^{n+1/2}, \mathbf{v}) + b^*(\mathbf{u}_h^{n+1/2}, \mathbf{s}_h^{n+1/2}, \mathbf{v}), \\ + (\mathbf{u}_h^{n+1/2} - \bar{\mathbf{u}}_h^{n+1/2}, \mathbf{v}) + \chi(\mathbf{s}_h^{n+1/2} - \mathbf{w}_h^{n+1/2}, \mathbf{v}) &= 0. \end{aligned} \quad (3.7)$$

The numerical analysis of (3.6) and (3.7) can be found in [11, 29].

## 4 The Interval of Reliability

In this numerical study, we aim to show that the flow sensitivity calculated from sensitivity equations (3.3), (3.5), or (3.7) can be used to quantify the reliability of the flow solution computed using (3.2), (3.4), or (3.6) respectively as the user-elected model parameter takes different values. Let  $\eta$  be the designated model parameter, then one can look at the following difference quotient for the sensitivity,

$$\mathbf{s} = \frac{\partial \mathbf{u}}{\partial \eta} \approx \frac{\mathbf{u}(\eta) - \mathbf{u}(0)}{\eta} \quad (4.1)$$

Considering  $\mathbf{u}$  as an implicit function of parameter  $\eta$ ,  $\mathbf{u}(0)$  indicates the true solution of Navier-Stokes equations while  $\mathbf{u}(\eta)$  for  $\eta > 0$  denotes the corresponding flow model approximation. In all the discussed fluid models in this paper, EVM, LAM, and TRM, the flow solution is an accurate approximation to the Navier-Stokes solution when  $\|\mathbf{u}(\eta) - \mathbf{u}(0)\|$  is small, and according to (4.1) the accuracy of the model approximation can be estimated by measuring  $\eta \|\mathbf{s}\|$ . As noted in Section 3, the sensitivity calculation can be coupled with that of the original fluid model simulation. The computations for sensitivity equations are easily added as all of the data structures and filter calculations are very similar to that of the corresponding fluid model. Thus after computing a model simulation with a given set of parameters, the sensitivity computation can be done with only a nominal extra cost and a quantitative measure of reliability can be then calculated. All the programs have been implemented using the software package FreeFem++; see [18] for details and examples.

Note that in cases where the model parameter  $\eta$  corresponds to the filter width e.g. EVM and LAM, then  $0 < \eta \leq 1$ , by selecting larger values of  $\eta$ , a larger set of velocity scales is removed. Hence, an approximated flow solution corresponding to large values of  $\eta$  may not be considered to be a reliable approximation to a solution of the Navier-Stokes model because too much of the small scale structure could be lost. This situation is especially tenuous for the case of high Reynolds numbers where the velocity contains a large number of small scales. Therefore, it is crucial to find the optimal balance between choosing a value of  $\eta$  that is small enough to provide a reliable approximation to the Navier-Stokes flow while choosing a value of  $\eta$  that is large enough so that the computation of the large scale velocity  $\mathbf{u}$  is feasible. This leads us to identify a range of  $\eta$  values for the interval of reliability for which both  $\eta$  and  $\eta \|\mathbf{s}\|$  are small. With  $\eta$  as the parameter corresponding to filter width to determine the upper end of the interval of reliability, one can use the Taylor expansion taking  $\eta$  values so that  $O(\eta^2)$  is of a certain precision,

$$\mathbf{u}(0) = \mathbf{u}(\eta) - \eta \mathbf{s} + O(\eta^2). \quad (4.2)$$

Next we present two numerical experiments with 2D Cavity problem where we specify reliable parameter values for LAM and TRM using the sensitivity computations as discussed in this section.

### 4.1 2D Cavity Problem

In the following experiments, two fluid models, LAM and TRM, and their sensitivity equations are numerically solved on the domain defined by  $\Omega = [0, 1] \times [0, 1]$ . The upper boundary

condition is chosen to be  $\mathbf{u} = (16x^2(1-x)^2, 0)^T$  and zero everywhere else. The initial data is  $\mathbf{u}(0, x, y) = (3y^2 - 2y, 0)^T$  in  $\Omega$ . Since initial and boundary conditions are independent from the model parameter  $\eta$ , they are set to zero for the sensitivity  $\mathbf{s}$ .

All the computations are carried out with a fixed mesh size  $h = \frac{1}{36}$ , and a uniform time step  $\Delta t = 0.01$  using the Taylor-Hood finite elements. Note that in the following computations, the sensitivity of the approximated velocity  $\mathbf{u}$  with respect to the variations of the model parameter is tested by computing  $\|\mathbf{s}\|_{L^2(0,T;L^2)}$  (where  $\|\mathbf{v}\|_{L^2(0,T;L^2)} = \left[ \Delta t \sum_{i=0}^N \|\mathbf{v}(i\Delta t)\|_{L^p}^q \right]^{1/q}$ ) for final time  $T = 1$ .

#### 4.1.1 Leray-Alpha Model

The sensitivity computations in this section are performed for different viscosities corresponding to Reynolds numbers of 5000, 10000, and 50000. In addition for each tested  $Re$  value, computations are carried out for variations of filter width  $\alpha$ , where the values are chosen based on the spatial mesh size as  $\alpha = kh$ , for  $k = \frac{1}{4}, \frac{1}{2}, 1, 2, 4$ .

The numerical results obtained from these computations are shown in Figure 1. Note that by selecting large values for the parameter  $\alpha$ , e.g. larger than  $4h$ , all the velocity scales that are less or equal to  $\alpha$  are filtered. Using (4.2), the reliable  $\alpha$  values are restricted to values with  $O(\alpha^2)$  less than 0.01. Considering the fact that a higher sensitivity for smaller values of parameter  $\alpha$  indicates the rise of computational complexity, the numerical results in Figure 1 suggest the following interval of  $\alpha$  values as the optimal choice for the tested Reynolds numbers. As indicated in Table 1, the interval of optimal values of  $\alpha$  for the best accuracy and computational complexity is smaller for higher Reynolds number.

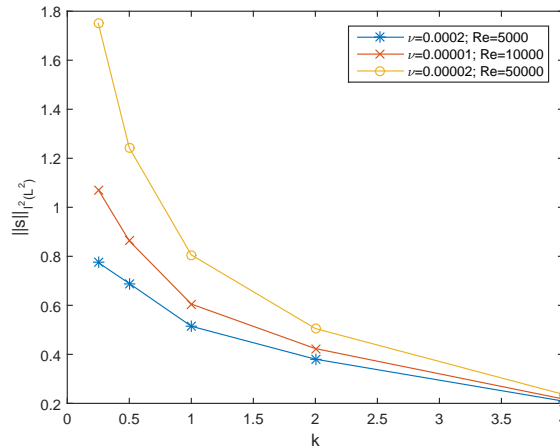


Figure 1: Sensitivity over the time interval of  $[0, 1]$

Furthermore we present the sensitivity computations using the Forward Finite Difference (FFD) quotient  $\frac{\mathbf{u}(\alpha+\Delta\alpha)-\mathbf{u}(\alpha)}{\Delta\alpha}$ , by computing  $\mathbf{u}$  from (3.4) for two inputs  $\alpha + \Delta\alpha$  and  $\alpha$ . The goal is to illustrate a comparison of the sensitivity values obtained from FFD versus that from sensitivity equation in (3.5) denoted by SEM for simplicity. The computations for sensitivity norm via both methods,  $\|\mathbf{s}_{SEM}(t)\|_{L^2(\Omega)}$  and  $\|\mathbf{s}_{FFD}(t)\|_{L^2(\Omega)}$ , are performed for different  $\alpha$

Table 1: The interval of optimal values for parameter  $\alpha$ 

Re	Interval of $\alpha$ values
5000	$\frac{1}{4}h \leq \alpha \leq 4h$
10000	$\frac{1}{2}h \leq \alpha \leq 4h$
50000	$h \leq \alpha \leq 4h$

values with  $\Delta\alpha = 0.001$  at times  $t = 0.1$ , and 1. Figures 2-4 display these computational results.

Note that for all  $\alpha$ , sensitivities computed via FFD is overall larger than the one computed via SEM in all the tested Reynolds numbers. Sensitivity norm in both methods increases as  $\alpha$  takes on values closer to 0 demonstrating a higher sensitivity of the approximated velocity  $\mathbf{u}$  with respect to smaller values of  $\alpha$ . One also observes that as time has progressed from  $t = 0.1$  to  $t = 1$ , sensitivities become larger in scale. In addition, larger Reynolds numbers show larger sensitivities especially at the final time.

The difference between  $\|\mathbf{s}_{SEM}\|_{L^2(\Omega)}$  and  $\|\mathbf{s}_{FFD}\|_{L^2(\Omega)}$  for different values of Reynolds number at  $t = 1$  is presented in Table 2. Sensitivity values for  $\alpha = 2h$ , and  $4h$  are apart up to 0.25 in all cases and there is an increase in the difference as  $\alpha$  decreases. For high Reynolds number, i.e.  $Re = 10000$ , and  $50000$ , the difference in sensitivity norms is noticeable for  $\alpha = \frac{1}{2}h$ , and  $\frac{1}{4}h$ .

 Table 2: Difference between  $\|\mathbf{s}_{SEM}\|$  and  $\|\mathbf{s}_{FFD}\|$  at  $t = 1$  with  $h = \frac{1}{36}$ 

$\alpha$	$Re = 5000$	$Re = 10000$	$Re = 50000$
$4h$	0.1586	0.1603	0.1586
$2h$	0.2458	0.2335	0.1955
$h$	0.4118	0.4315	0.3685
$\frac{1}{2}h$	0.4732	0.6689	0.9310
$\frac{1}{4}h$	0.4394	0.7133	1.4550

#### 4.1.2 Time Relaxation Model

In this experiment, tested Reynolds numbers are 1000, 5000, and 10000, for different values of time relaxation parameter  $\chi = 0.01, 0.1, 1$ , and 10. As seen in Table 3,  $\chi\|\mathbf{s}\|_{l^2(0,1;L^2(\Omega))}$  values via both methods take larger values for larger  $Re$  with any selected value of parameter  $\chi$ . For  $Re = 1000$ , we suggest  $\chi \leq 1$  as the best choice of accuracy while for larger  $Re$  values, we select a smaller interval of  $\chi$  values, that is  $\chi \leq 0.1$ .

**Remark 4.1.** *In this experiment, we chose  $\chi$  values for which  $\chi\|\mathbf{s}\|_{l^2(0,1;L^2(\Omega))} \leq 0.01$  for the best accuracy. The smaller the  $\chi$  value becomes, the more accurate the calculation of*

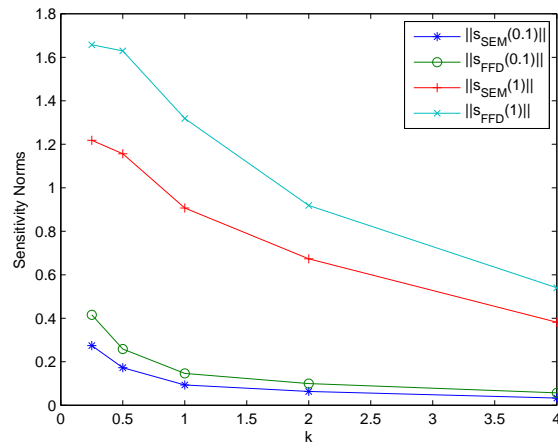


Figure 2: Sensitivity norms via SEM and FFD for  $Re = 5000$

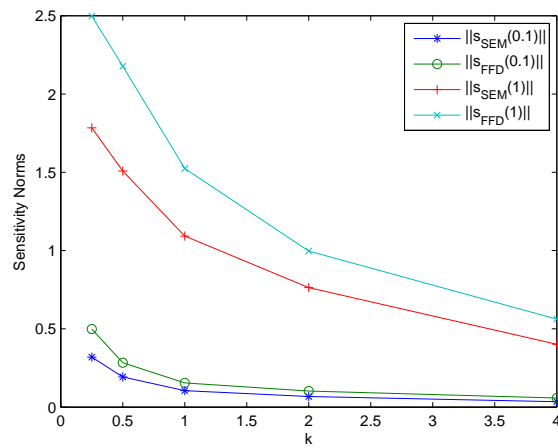


Figure 3: Sensitivity norms via SEM and FFD for  $Re = 10000$

Table 3: Sensitivity values  $\chi \|s\|_{L^2(0,1;L^2)}$

$\chi$	$Re = 1000$	$Re = 5000$	$Re = 10000$
0.01	0.000106756	0.000191289	0.0002248
0.1	0.00103614	0.00184001	0.00215508
1	0.00803338	0.0131381	0.0149263
10	0.0503827	0.060064	0.0621485

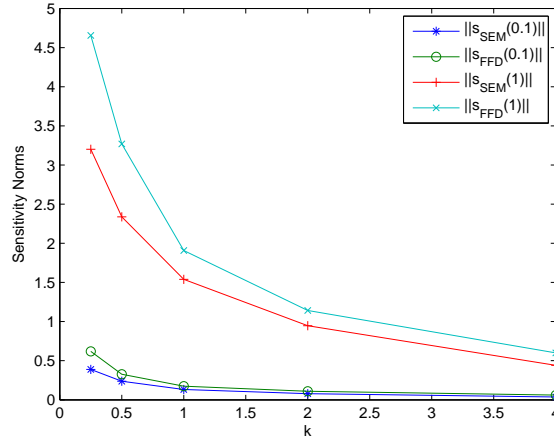


Figure 4: Sensitivity norms via SEM and FFD for  $Re = 50000$

approximated velocity  $\mathbf{u}$  becomes. However very small values of the parameter  $\chi$  result in increasing the complexity of flow structures/scales that cannot be supported by the grid/mesh and thus numerical pollution of the computed velocity begins. Therefore, the user must consider the trade-off between increased accuracy and computational flow complexity when choosing the  $\chi$  value.

The data listed in Tables 4-5 displays the maximum sensitivity values, i.e.  $\|\mathbf{s}\|_{L^\infty(0,1;L^2(\Omega))}$ , over the time interval  $[0, 1]$  for both  $Re = 1000$  and  $Re = 10000$  with different  $\chi$  values as the spatial mesh size is refined. It is worth mentioning that the maximum sensitivity norm happens at the final time for any mesh size as well as any selected  $\chi$  values. One observes a decrease in  $\|\mathbf{s}\|_{L^\infty(0,1;L^2(\Omega))}$  as the spatial mesh size is refined for  $\chi \leq 1$ . In both tables,  $\|\mathbf{s}\|_{L^\infty(0,1;L^2(\Omega))}$  values for  $\chi = 10$  stay close through the mesh refinement.

Table 4: Sensitivity computations for  $Re = 1000$  with mesh refinement

$\chi$	$h = \frac{1}{9}$	$h = \frac{1}{18}$	$h = \frac{1}{36}$
0.01	0.0586237	0.035002	0.0157713
0.1	0.0541932	0.0330563	0.0151647
1	0.0250335	0.0199503	0.0110461
10	0.00893807	0.00966404	0.00747537

Overall, for the 2D cavity problem, both models share stronger sensitivity at higher  $Re$  and at lower values of the elected parameter  $\alpha$  for LAM and  $\chi$  for TRM. Sensitivities can provide useful information for other modeling needs such as in the following section where it is shown that sensitivities can improve some essential flow functionals such as drag and lift.



Table 5: Sensitivity computations for  $Re = 10000$  with mesh refinement

$\chi$	$h = \frac{1}{9}$	$h = \frac{1}{18}$	$h = \frac{1}{36}$
0.01	0.0808053	0.0595574	0.0363773
0.1	0.0747136	0.0557121	0.0344382
1	0.0347436	0.0307948	0.02162
10	0.0094678	0.0106285	0.0090533

## 5 Improving Flow Functionals

The standard procedure of computing a flow functional is to first compute the approximated flow velocity  $\mathbf{u}$  from the fluid model, then use  $\mathbf{u}$  as the fluid velocity input into the given flow functional. Let  $\eta$  be the model parameter upon which  $\mathbf{u}$  depends implicitly with property that  $\mathbf{u}(\eta) \rightarrow \mathbf{u}(0)$  as  $\eta \rightarrow 0$ . Let  $J(\mathbf{u}(0)) = J(\mathbf{u})$  be a flow functional and that  $\mathbf{u}(0)$  is extremely computationally expensive to obtain directly. The natural approach to compute a less expensive approximation of  $J(\mathbf{u})$  is to calculate  $J(\mathbf{u}(\eta))$  for a non-zero  $\eta$ . Note that  $J(\mathbf{u}(\eta))$  is a good approximation provided that  $\mathbf{u}(\eta)$  is an acceptable approximation and that the unresolved scales do not influence the functional. In this section, we discuss how the approximation of  $J(\mathbf{u})$  can be improved by integrating sensitivities into the computations of the flow functional. The idea is simply based on the first order Taylor expansion of the flow functional around a non-zero  $\eta$  value. Expanding  $J(\mathbf{u})$  around a non-zero  $\eta$  implies that

$$J(\mathbf{u}) \approx J(\mathbf{u}(\eta)) - \eta J'(\mathbf{u}(\eta)) \cdot \mathbf{s} \quad (5.1)$$

Replacing  $J'$  by  $J$ , given  $J' = J$  for linear functionals, and incorporating the pressure into the above formula, the approximation (5.1) is rewritten as

$$J(\mathbf{u}, p) \approx J(\mathbf{u}(\eta), p(\eta)) - \eta J(\mathbf{s}, r) = J(\mathbf{u}(\eta) - \eta \mathbf{s}, p(\eta) - \eta r) \quad (5.2)$$

By (5.2), a flow functional can be approximated using sensitivities as the first order correction term for both the velocity and pressure. This idea was proposed by Anitescu and Layton for LES models and was tested on the Smagorinsky model in [3]. Next we provide a numerical support for the idea that uses drag computations in a channel with a cylinder.

### 5.1 2D Flow around Cylinder

In this numerical experiment, we consider estimating the drag functional using EVM on the standard test problem of two-dimensional flow in a channel around a cylinder. The lift and drag functional for Navier-Stokes equations is given by

$$J(\mathbf{u}, p) = \int_D \hat{n} \cdot (pI - 2\nu \nabla^s \mathbf{u}) \cdot \hat{a} ds \quad (5.3)$$

where  $\hat{n}$  denotes the normal vector on the cylinder boundary  $D$  directing into the channel,  $\nabla^s \mathbf{u}$  presents the deformation tensor and is  $\frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$ , the unit vector  $\hat{a}$  in the positive direction of  $x$ -axis or negative direction of  $y$ -axis yield the drag or lift flow functional.

Figure 5 displays the geometry of the channel with the cylinder. The channel is a rectangle with height and width as 0.41m and 2.2m respectively. The cylinder, denoted by  $D$ , is of radius 0.05m, and its center is placed at  $(0.2, 0.2)$ .

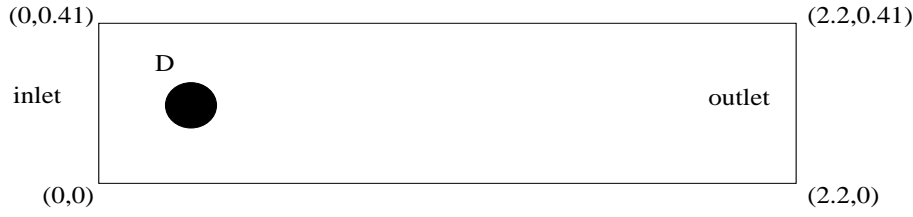


Figure 5: Geometry of 2D-flow around cylinder

The numerical approximation to the solution of EVM in (2.1) are computed for  $0 \leq t \leq 4$  with the inflow conditions given below that are parabolic in space and periodic in time

$$\begin{aligned} \mathbf{u}_1(t, 0, y) &= \frac{6}{(0.41)^2} y(0.41 - y) \sin(\pi t) \\ \mathbf{u}_2(t, 0, y) &= 0. \end{aligned} \quad (5.4)$$

A "do-nothing" condition is used for the outflow boundary condition, and the remaining boundary and initial conditions are given by

$$\begin{aligned} \mathbf{u}_1(t, x, 0) &= \mathbf{u}_2(t, x, 0) = 0 \\ \mathbf{u}_1(t, x, 0.41) &= \mathbf{u}_2(t, x, 0.41) = 0 \\ \mathbf{u}_1(t, x, y) |_{\partial D} &= \mathbf{u}_2(t, x, y) |_{\partial D} = 0 \\ \mathbf{u}_1(0, x, y) &= \mathbf{u}_2(0, x, y) = 0. \end{aligned}$$

### 5.1.1 Eddy Viscosity Model

A non-uniform mesh that is finer around the cylinder  $D$  is used for the triangulation of the domain in Figure 5. A given mesh is constructed using two sizes,  $h_1$  for the sides of the channel, and  $h_2$  for the boundary of  $D$ . Therefore the mesh is identified using the ordered pair  $(h_1, h_2)$ . For the numerical computation of the projection operator in (3.2) and (3.3), the applied coarse mesh has the same structure and is always chosen as  $(H_1 = \sqrt{h_1}, H_2 = \sqrt{h_2})$ . An example of a mesh of size  $(\frac{1}{36}, \frac{1}{49})$  is indicated in Figure 3.

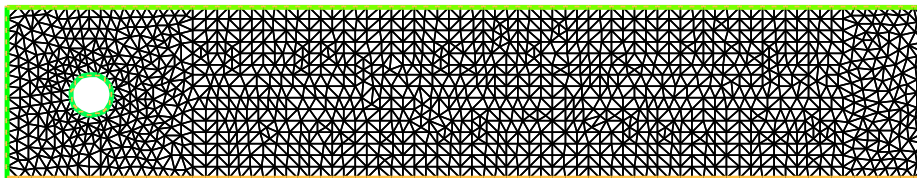


Figure 6: Mesh in a channel of size  $(\frac{1}{36}, \frac{1}{49})$

The reference value of drag for this test problem is calculated using the DNS method on a fine mesh of size  $(\frac{1}{100}, \frac{1}{121})$  for  $0 \leq t \leq 4$ . Figure 7 presents a sample of the scaled velocity vector field for the case when  $Re = 1000$  and  $\alpha = 0$  at  $t = 0.5$ . Note that this graph contains only the portion of the domain surrounding the cylinder where the interesting flow behavior occurs.

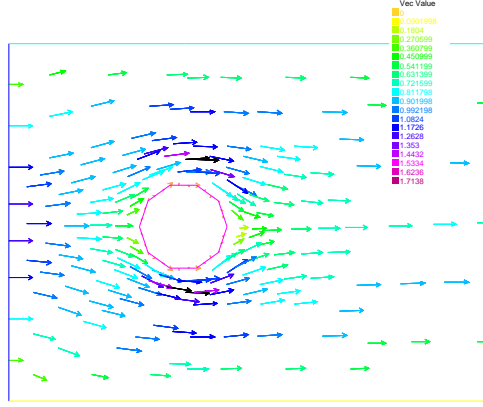


Figure 7: Velocity vector field for  $Re = 1000$  and  $\alpha = 0$

Table 6 lists the reference values of maximum drag, and the error in its estimation using the approximated large eddy velocity and pressure  $(\mathbf{u}(\alpha), p(\alpha))$  and  $(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r)$  in drag calculation by (5.3) for different values of  $Re$ . In this experiment, the approximated flow variables and their sensitivities are obtained from (3.2) and (3.3) with  $\alpha = 0.00125$  and a mesh size of  $(\frac{1}{49}, \frac{1}{64})$ . As indicated in this table, computed drag values by  $(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r)$  are more accurate for all  $Re$ , especially for  $Re \geq 100$ . In addition the computed drag values using  $(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r)$  shows only a small improvement in comparison to the ones computed using  $(\mathbf{u}(\alpha), p(\alpha))$  when  $Re \leq 10$ . However, for  $Re \geq 100$ , the errors incurred by using  $(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r)$  improve significantly and decrease by a full order of magnitude.

Table 6: Maximum drag values and the errors

	Max. Drag	Error using $(\mathbf{u}(\alpha), p(\alpha))$	Error using $(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r)$
$Re$	$J(\mathbf{u}, p)$	$ J(\mathbf{u}, p) - J(\mathbf{u}(\alpha), p(\alpha)) $	$ J(\mathbf{u}, p) - J(\mathbf{u}(\alpha) - \alpha \mathbf{s}, p(\alpha) - \alpha r) $
1	63.7703	0.4037	0.3702
10	41.1958	0.3628	0.3555
100	36.0677	0.389	0.0152
1000	35.29035	0.28095	0.02095
10000	35.1186	0.4354	0.0154

We examined the norm of the sensitivity quantities for the same range of  $Re$  values in Table 6. The sensitivity norms in Table 7 are negligible for  $Re \leq 10$  indicating that the ap-

proximated flow solution is accurate for that range of  $Re$  values. As also reflected in Table 6, there is a nominal error in the drag value approximations using  $(\mathbf{u}(\alpha), p(\alpha))$  for  $Re \leq 10$ . According to Table 7, for large values of  $Re$ , i.e.  $Re \geq 100$ , the flow becomes more sensitive, and using sensitivity information improves the estimated values of the drag functional significantly.

Table 7: Sensitivity for different values of  $\nu$ 

$Re$	$\alpha \ \mathbf{s}\ _{L^\infty(0,T;L^2)}$
1	7.19057e-06
10	2.88244e-04
100	0.00483735
1000	0.0155576
10000	0.0201101

### 5.1.2 Time Relaxation Model

Lift and drag computations were also performed with the same channel geometry using TRM. These calculated computations use the improved velocity and pressure values by their improved sensitivities, i.e.  $J(\mathbf{u}(\chi) - \chi \mathbf{s}, p(\chi) - \chi r)$  in (5.3). The inflow follows a scaled form of (5.4) given as

$$\mathbf{u}_1(t, 0, y) = \frac{6}{(0.41)^2} y(0.41 - y) \sin\left(\frac{\pi t}{8}\right)$$

The numerical approximations are obtained for  $Re = 1000$ , and  $0 \leq t \leq 8$  for different values of parameter  $\chi = 0.1, 0.01$ , and  $0.001$ . Delaunay triangulation is used to create the mesh with mesh size  $h$  such that  $0.00650741 < h < 0.0340958$ . Figure 8 illustrates the correct development of lift and drag values over the time for the tested parameter values. The velocity streamlines at different time steps for  $\chi = 0.01$  are shown in Figure 9 and its sensitivity in Figure 10. As it can be seen in the figure, the velocity is most sensitive around the cylinder, and this conclusion is the same for all the tested  $\chi = 0.1, 0.01$ , and  $0.001$  values.

## 6 Non-isothermal Stokes Sensitivity

Non-isothermal stokes flow is another interesting problem to test for sensitivity, [20, 14, 30]. This differentially heated cavity is defined on the unit square with no slip boundary conditions for the velocity and mixed Dirichlet/Neumann conditions for the temperature and Prandtl

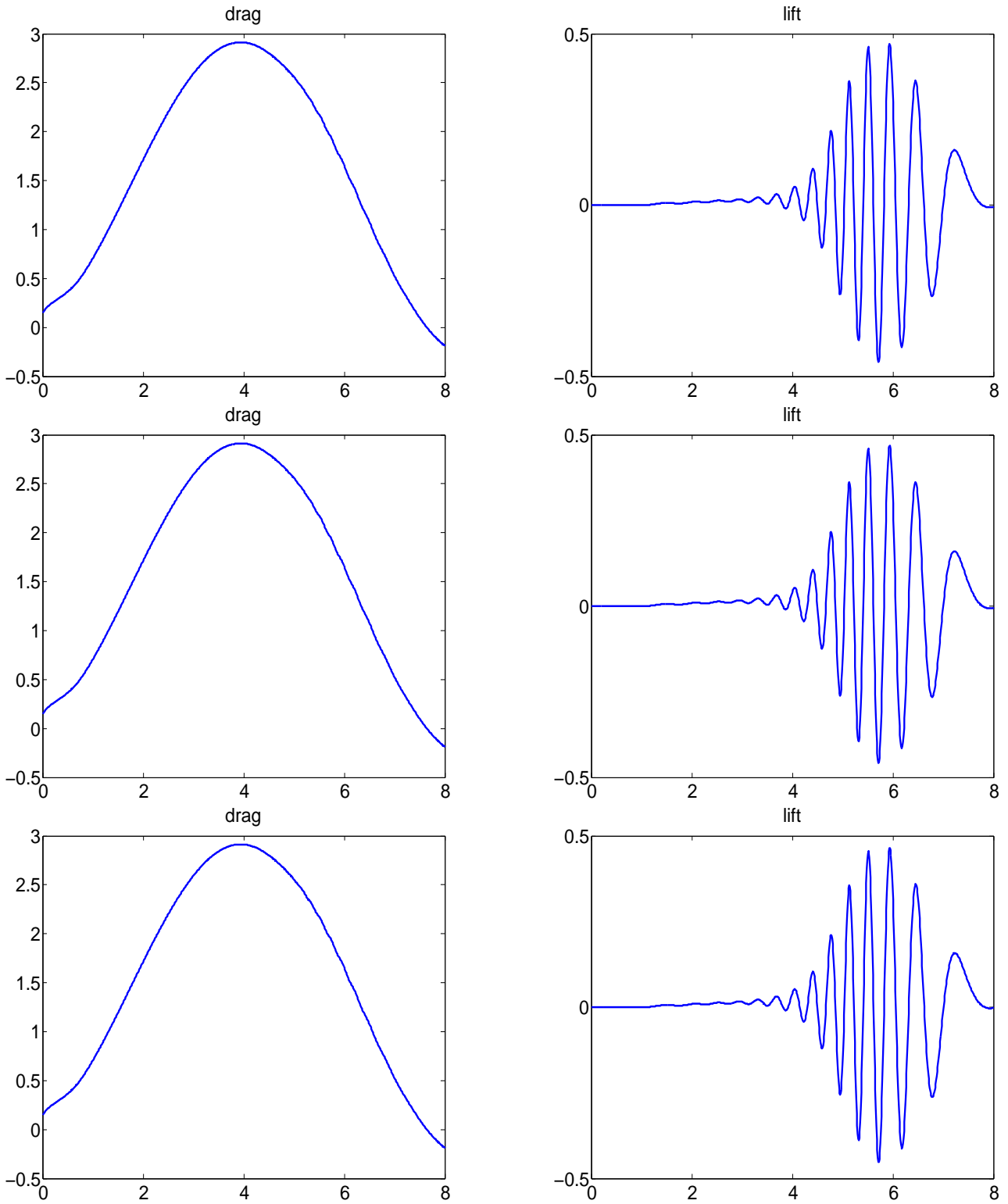


Figure 8: The development of drag and lift (left to right from  $t \in [0, 8]$  with  $\text{DOF}=16\text{K}$ ,  $dt = 0.01$ , and  $\chi = 10^{-3}, 10^{-2}, 10^{-1}$  (top to bottom).

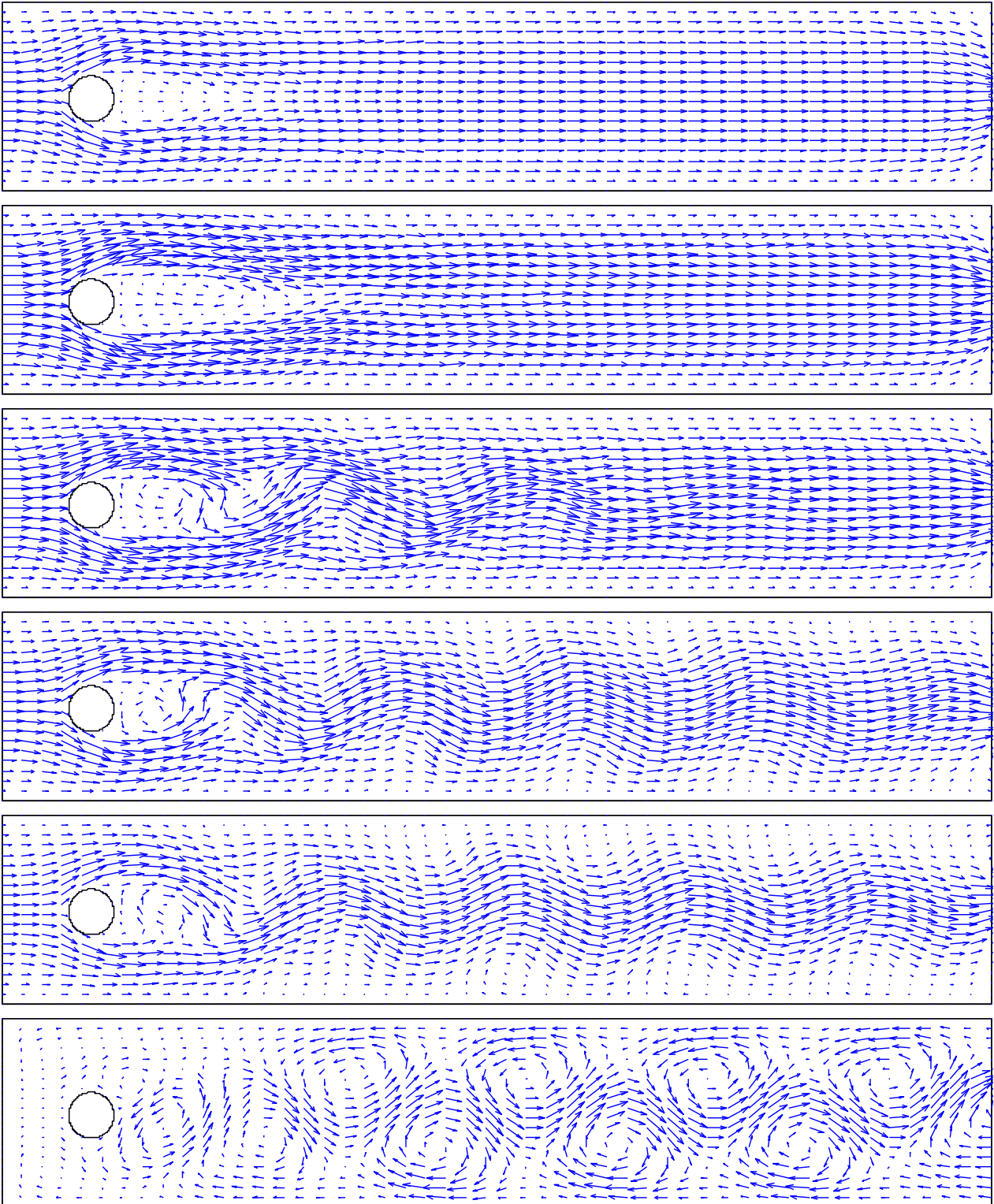


Figure 9: The velocity field at  $t = 2, 4, 5, 6, 7,$  and  $8$  with  $\text{DOF}=16\text{K}$ ,  $dt = 0.01$ , and  $\chi = 10^{-2}$ .

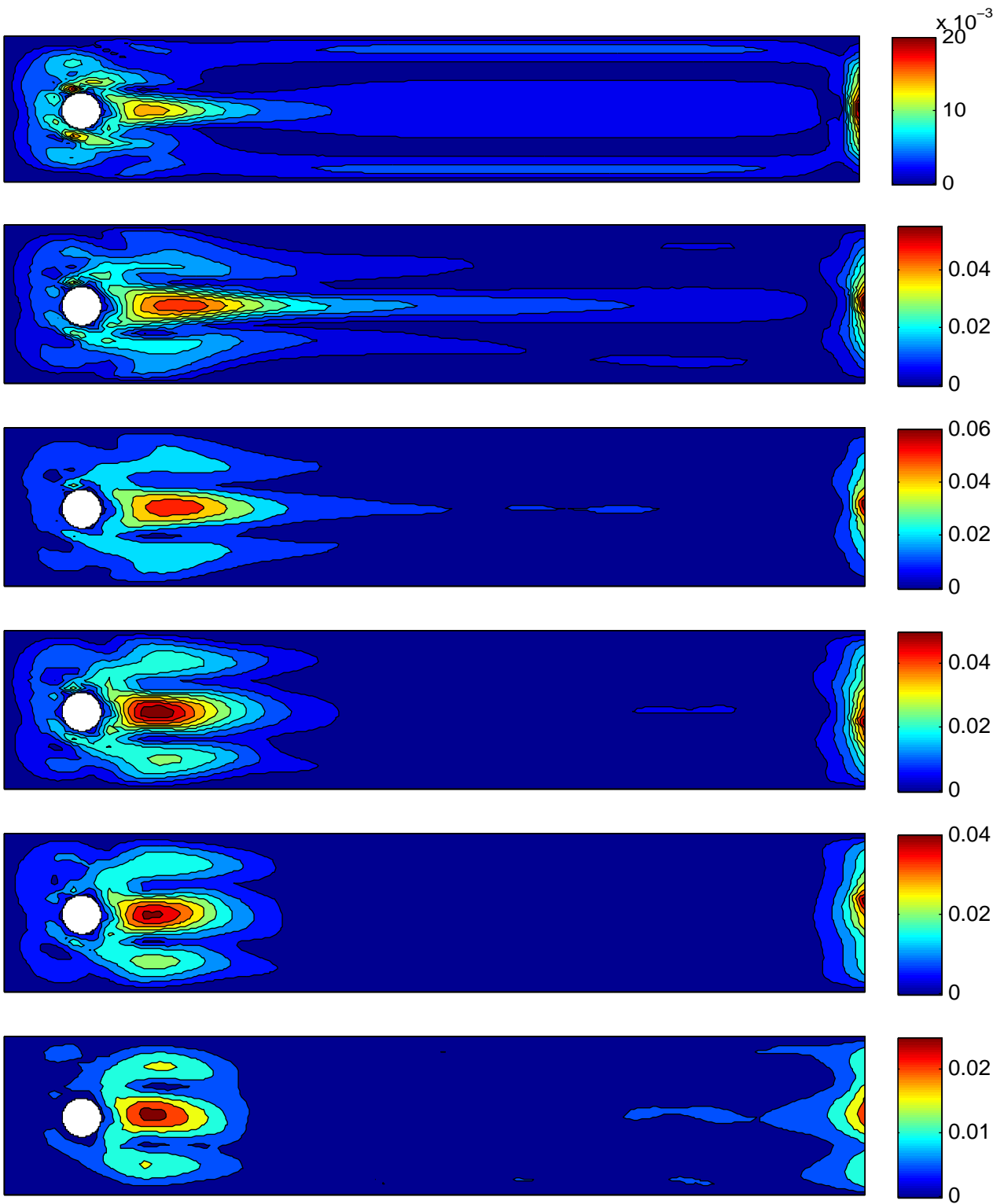


Figure 10: The velocity sensitivity at  $t = 2, 4, 5, 6, 7,$  and  $8$  with  $\text{DOF}=16\text{K}$ ,  $dt = 0.01$ , and  $\chi = 10^{-2}$ .

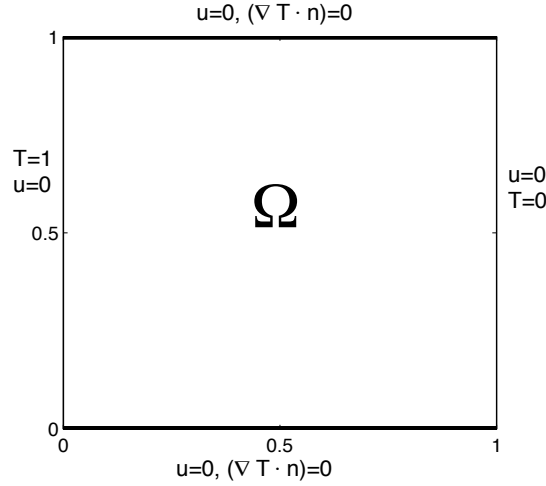


Figure 11: The domain and boundary conditions for the natural convection problem.

number  $Pr = \infty$ . The system of equations governing this flow is given by

$$\begin{aligned}
 -\nu \Delta \mathbf{u} + \nabla p &= (0, RaT)^T, \text{ in } \Omega, \\
 \nabla \cdot \mathbf{u} &= 0, \text{ in } \Omega, \\
 -\Delta T + \mathbf{u} \cdot \nabla T &= 0, \text{ in } \Omega \\
 \mathbf{u} &= \mathbf{0} \text{ on } \partial\Omega, \\
 T &= 0 \text{ on } \Gamma_D, \\
 \nabla T \cdot n &= 0 \text{ on } \Gamma_N,
 \end{aligned}$$

with  $\partial\Omega = \Gamma_D \cup \Gamma_N$ , and  $Ra$  is the Rayleigh number. The domain and boundary conditions are given in Figure 11.

Modeling this system using TRM, we get

$$\begin{aligned}
 -\nu \Delta \mathbf{u} + \chi(\mathbf{u} - \bar{\mathbf{u}}) + \nabla p &= (0, RaT)^T, \text{ in } \Omega, \\
 \nabla \cdot \mathbf{u} &= 0, \text{ in } \Omega, \\
 -\Delta T + \mathbf{u} \cdot \nabla T &= 0, \text{ in } \Omega,
 \end{aligned} \tag{6.1}$$

with the above boundary conditions. The corresponding sensitivity equations are listed as

$$\begin{aligned}
 -\nu \Delta \mathbf{s} + \chi(\mathbf{s} - \bar{\mathbf{s}}) + (\mathbf{u} - \bar{\mathbf{u}}) \cdot \nabla r &= (0, Ra k)^T, \text{ in } \Omega, \\
 \nabla \cdot \mathbf{s} &= 0, \text{ in } \Omega, \\
 -\Delta k + \mathbf{s} \cdot \nabla T + \mathbf{u} \cdot \nabla k &= 0, \text{ in } \Omega,
 \end{aligned} \tag{6.2}$$

where  $k$  is the temperature sensitivity, i.e.  $k = \frac{\partial T}{\partial \chi}$ . In equation (6.1), the computations are carried out on a fixed mesh  $h = \frac{1}{36}$  that uses  $(P_2, P_2, P_1)$  Taylor-Hood elements for velocity, pressure and temperature, of which the nonlinearity is resolved using Newton's method with a continuation method on  $Ra$  to get convergence. The same method is applied in computations of  $s$ ,  $r$ , and  $k$  from (6.2).



Tables (8)- (10) list velocity and temperature sensitivities with respect to  $\chi$  for  $Ra=10^4$ ,  $10^5$  and  $10^6$ . As one can observe from the content of tables, sensitivity values decrease as  $\chi$  values increase.

Velocity streamline and the temperature contour plot for  $Ra = 10^5$  and  $\chi = 10^3$  are shown in Figures 12. Sensitivity plots for velocity, Figures 13-15, illustrate a higher sensitivity areas on the outer left and right area near the boundary than the center, while sensitivity plots for temperature, Figures 16-18, show higher sensitivity areas on the left than right wall of the domain where the temperature is higher. These sensitivity areas elongate and become thinner as Ra number is increased.

Table 8: Sensitivity computations for  $Ra=10^4$  with different  $\chi$

$\chi$	Velocity Sensitivity	Temperature Sensitivity
$10^{-1}$	0.00454608	3.98462e-05
1	0.00454216	3.98278e-05
10	0.00450418	3.96366e-05
$10^2$	0.0041456	3.7878e-05
$10^3$	0.00217831	2.66418e-05
$10^4$	0.000149187	4.44796e-06

Table 9: Sensitivity computations for  $Ra=10^5$  with different  $\chi$

$\chi$	Velocity Sensitivity	Temperature Sensitivity
$10^{-1}$	0.00946049	2.33035e-05
1	0.0094524	2.32945e-05
10	0.00938285	2.319e-05
$10^2$	0.00873504	2.22121e-05
$10^3$	0.00497591	1.5945e-05
$10^4$	0.000624519	4.98162e-06

Table 10: Sensitivity computations  $Ra=10^6$  with different  $\chi$

$\chi$	Velocity Sensitivity	Temperature Sensitivity
$10^{-1}$	0.0170702	1.58207e-05
1	0.0170612	1.58158e-05
10	0.0169329	1.56251e-05
$10^2$	0.0160371	1.50129e-05
$10^3$	0.0101746	1.03233e-05
$10^4$	0.00158146	3.19721e-06

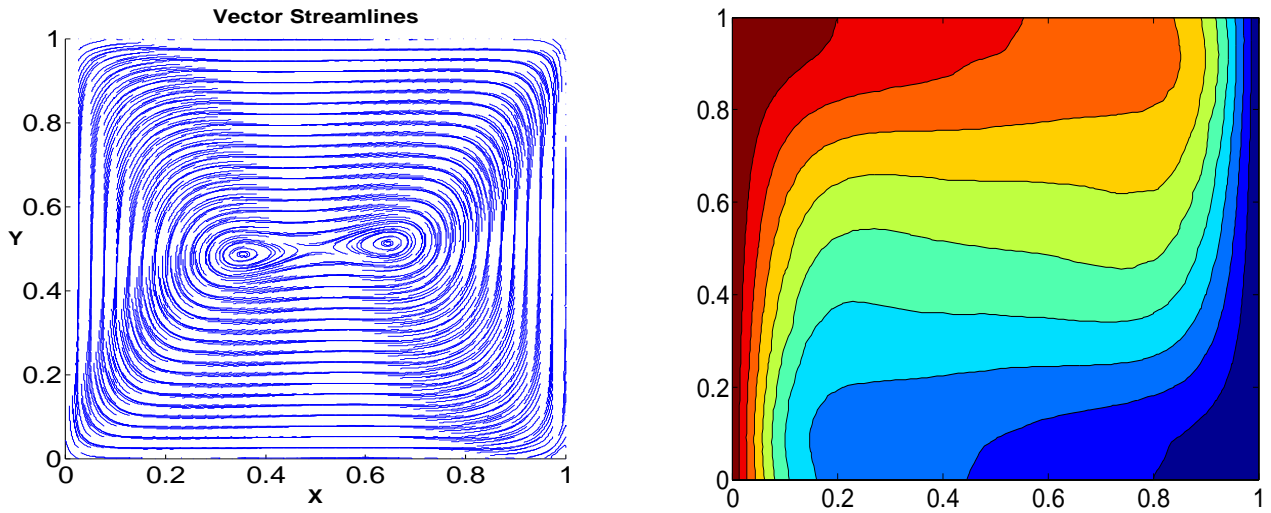


Figure 12: Shown above are velocity streamlines plots (left) and temperature contour (right) for  $Ra=10^5$  and  $\chi = 10^3$ .

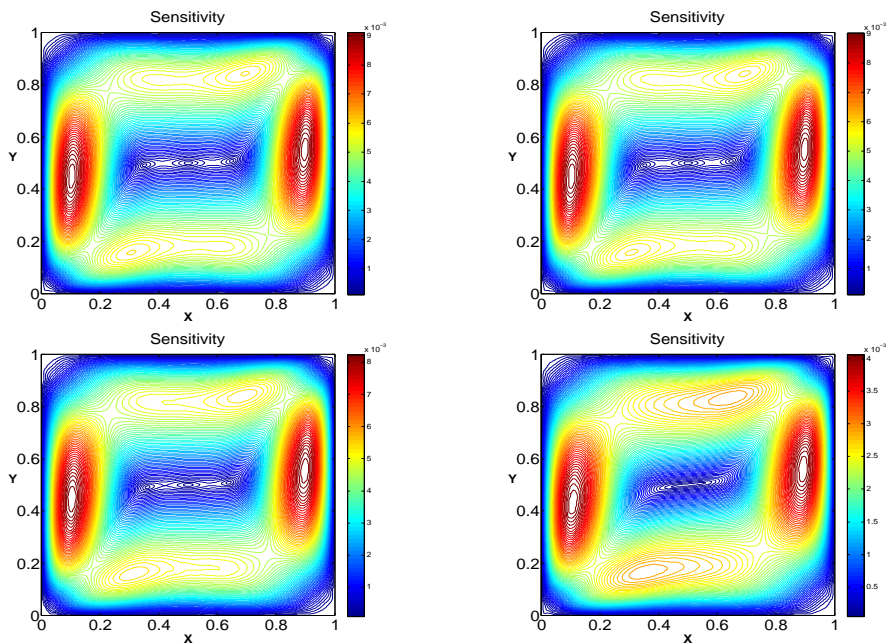


Figure 13: Sensitivity plots for velocity with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Ra=10^4$ .

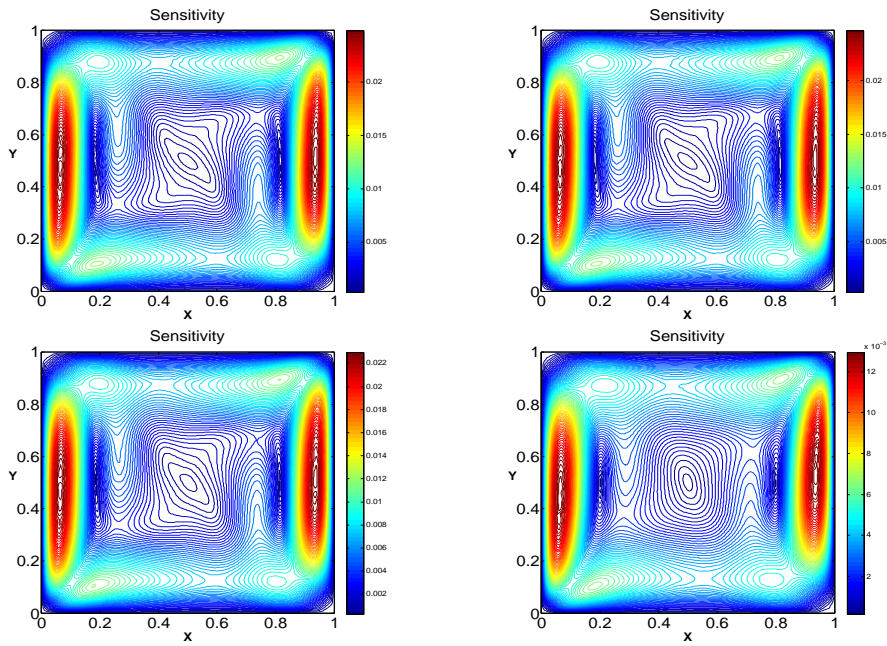


Figure 14: Sensitivity plots for velocity with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Ra=10^5$ .

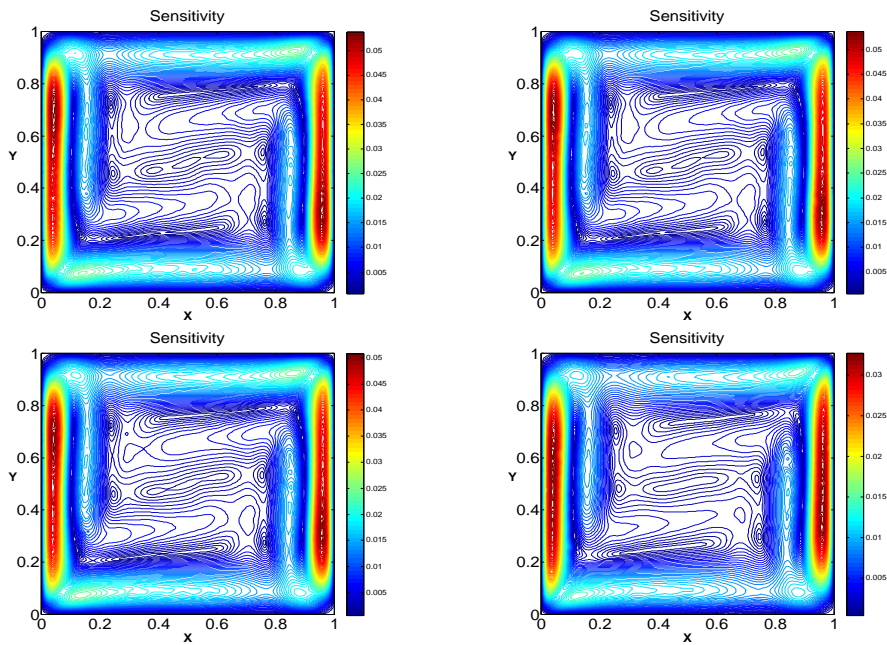


Figure 15: Sensitivity plots for velocity with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Ra=10^6$ .

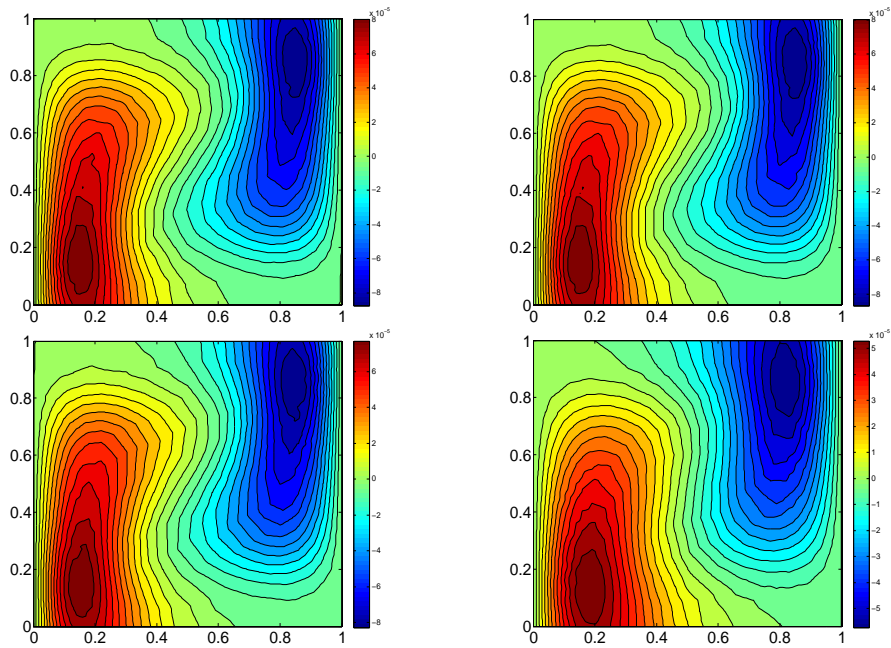


Figure 16: Sensitivity plots for temperature with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Ra=10^4$ .

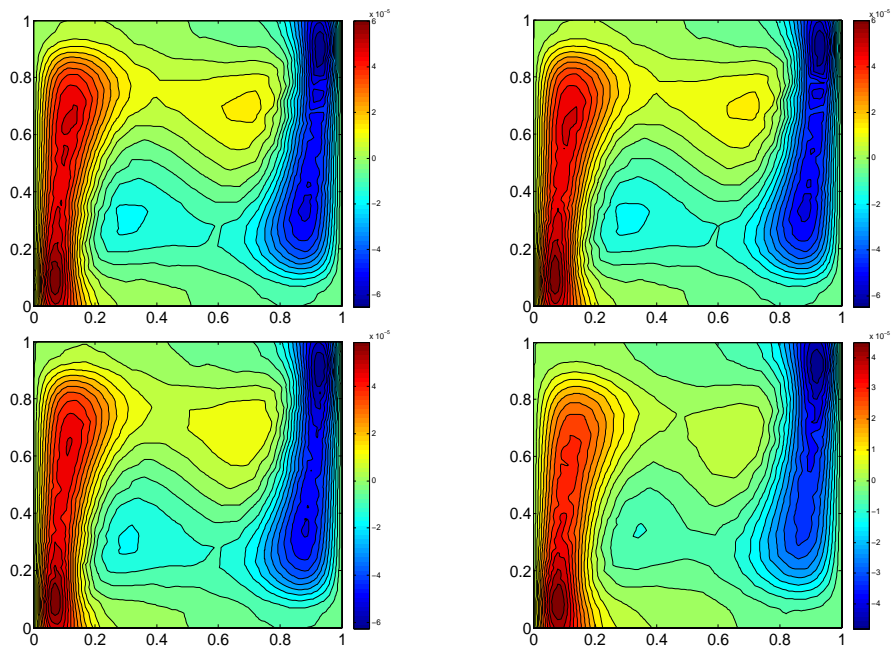


Figure 17: Sensitivity plots for temperature with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Ra=10^5$ .

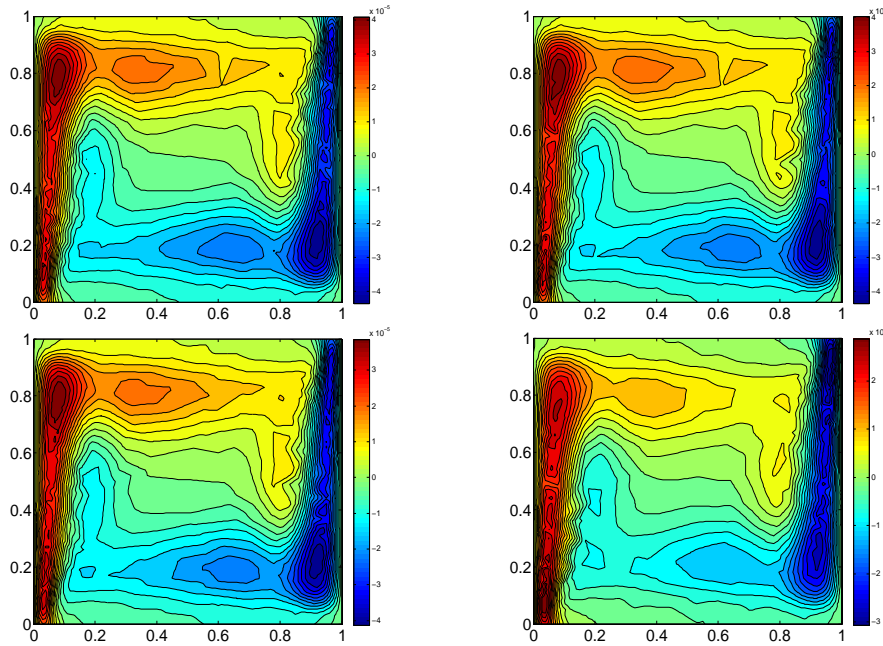


Figure 18: Sensitivity plots for temperature with respect to  $\chi = 1$ ,  $\chi = 10$ ,  $\chi = 10^2$ , and  $\chi = 10^3$  (from left to right and top to bottom) for  $Re=10^6$ .

## 7 Concluding Remarks

In this paper, three fluid models, EVM, LAM and TRM and their corresponding sensitivity equations are presented. This is followed by the numerical schemes for simulating the fluid models and their corresponding sensitivity. Once the numerical algorithm for solving each fluid model is implemented, the sensitivity calculations can be easily added due to the similarity in data structure. Our numerical experiments illustrate the application of sensitivities in quantifying model error arising from the choice of various parameter values and identifying those values that produce a reliable approximated velocity. The numerical results show that a smaller interval of reliable parameter values is obtained for larger values of  $Re$ . In addition, the sensitivity information is shown to be useful in increasing the accuracy of flow functionals for a nominal amount of effort in calculating sensitivities. Future studies can include stochastic finite element discretization that should give more insights into the parameter sensitivity.

## References

- [1] N. A. Adams and S. Stolz, Deconvolution methods for subgrid-scale approximation in large eddy simulation, *Modern Simulation Strategies for Turbulent Flow*, 2001.
- [2] N. A. Adams and S. Stolz, A subgrid-scale deconvolution approach for shock capturing, *Journal of Computational Physics*, **178**, 391-426, 2001.
- [3] M. Anitescu and W. Layton, Sensitivities in Large Eddy Simulation and Improved Estimates of Turbulent Flow Functionals, *SIAM Journal of Scientific Computing*, **29**, 1650-1667, 2007.

- [4] J. Borggaard and J. Burns, A Sensitivity Equation Approach to Shape Optimization in Fluid Flows, *Flow Control*, 49-78, 1995.
- [5] J. Borggaard, D. Pelletier and E. Turgeon, Sensitivity and uncertainty analysis for variable property flows, in *Proceedings of the 39th AIAA Aerospace Sciences Meeting and Exhibit*, 0140, 1993.
- [6] J. Borggaard, D. Pelletier and E. Turgeon, A Continuous Sensitivity Equation Method for Flows with Temperature Dependent Properties, *AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Design*, 4821, 2000.
- [7] A. Bowers and L. Rebholz, Increasing accuracy and efficiency in FE computations of the Leray-deconvolution model, *Numerical Methods for Partial Differential Equations*, to appear.
- [8] V.V.Chepyzhov, E.S. Titi and M.I. Vishik, *On the convergence of the Leray-alpha model to the trajectory attractor of the 3d Navier-Stokes system*, Report, 2005.
- [9] A. Cheskidov, D.D. Holm, E. Olson and E.S. Titi, On a Leray- $\alpha$  model of turbulence, *Royal Society London, Proceedings, Series A, Mathematical, Physical and Engineering Sciences*, **461**, 629-649, 2005.
- [10] L. Davis and F. Pahlevani, *Semi-Implicit Schemes for Transient for Navier-Stokes Equations and Eddy Viscosity Models*, *Journal of Numerical Methods for Partial Differential Equations*, 25 (2009), 212-231.
- [11] S. De, D. Hannasch, M. Neda and E. Nikonova, Numerical analysis and computations of a high accuracy Time Relaxation fluid flow model, *International Journal of Computer Mathematics*, 1-22, 2008.
- [12] A. Dunca and Y. Epshteyn, On the Stolz-Adams deconvolution LES model, *SIAM Journal on Mathematical Analysis*, 2006.
- [13] V. J. Ervin, W. J. Layton and Monika Neda, Numerical Analysis of a Higher Order Time Relaxation Model of Fluids, *International Journal of Numerical Analysis and Modeling*, **4**, 648-670, 2007.
- [14] K. Galvin and A. Linke and L. Rebholz and N. Wilson, Stabilizing poor mass conservation in incompressible flow problems with large irrotational forcing and application to thermal convection, *Computer Methods in Applied Mechanics and Engineering*, **237**, 166-176, 2012.
- [15] A. Godfrey, E. Cliff and W. Eppard, Using Sensitivity Equations For Chemically Reacting Flows, *AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, 4805, 1998.
- [16] V. Girault and P. Raviart, *Finite element approximation of the Navier-Stokes equations*, Springer, (1979).
- [17] M. Gunzburger, Perspectives in Flow Control and Optimization, *Frontiers in Mathematics*, SIAM, Philadelphia, 2003.

- [18] F. Hecht, O. Pironneau and K. Oshtsuka, *Software Freefem++*, <http://www.freefem++.org>, 2003.
- [19] A.A. Ilyin, E.M. Lunasin and E.S. Titi, *A modified Leray-alpha subgrid-scale model of turbulence*, Report, 2005.
- [20] E. Jenkins and V. John and A. Linke and L. Rebholz, On the parameter choice in grad-div stabilization for the Stokes equations, *Advances in Computational Mathematics*, **40**, 491-516, 2014.
- [21] V. John, *Large Eddy Simulation of Turbulent Incompressible Flows*, Springer 2004.
- [22] S. Kaya and B. Riviere, Analysis of a discontinuous Galerkin and eddy viscosity method for Navier-Stokes Equations, *SIAM Journal of Numerical Analysis*, **43**, 1572-1595, 2005.
- [23] W. Layton, *A connection between subgrid scale eddy viscosity and mixed methods*, Applied Mathematics and Computing, **133**, 147-157, 2002.
- [24] W. Layton and L. Rebholz, *Approximate Deconvolution Models of Turbulence: Analysis, Phenomenology and Numerical Analysis*, Springer Lecture Notes in Mathematics, 2012.
- [25] W. Layton, C. Manica, M. Neda and L. Rebholz, Numerical Analysis and Computational Testing of a high-order Leray-deconvolution turbulence model, *Numerical Methods for Partial Differential Equations*, **24**, 555-582, 2008.
- [26] J. Leray, Essay sur les mouvements plans d'une liquide visqueux que limitent des parois, *J. math. pur. appl.*, Paris Ser. IX, **13**, 331-418, 1934.
- [27] J. Leray, Sur les mouvements d'une liquide visqueux emplissant l'espace, *Acta Math.*, **63**, 193-248, 1934.
- [28] M. Neda, F. Pahlevani and J. Waters, Sensitivity computations of the Leray- $\alpha$  model, *Contemporary Mathematics, Recent Advances in Scientific Computing and Applications*, **586**, 2013.
- [29] M. Neda, F. Pahlevani and J. Waters, Sensitivity analysis and computations of Time Relaxation Model, *Advances in Applied Mathematics and Mechanics*, **7**, 89-115, 2015.
- [30] M. Neda, F. Pahlevani, L. Rebholz and J. Waters, Sensitivity analysis of the grad-div stabilization parameter in finite element simulations of incompressible flow, *Journal of Numerical Mathematics*, accepted, 2015.
- [31] F. Pahlevani and L. Davis, Parameter Sensitivity of an Eddy Viscosity Model: Analysis, Computation and its Application to Quantifying Model Reliability, *International Journal of Uncertainty Quantification*, **3**, 397-419, 2013.
- [32] F. Pahlevani, Sensitivity Computations of Eddy Viscosity Models with an Application in Drag Computation, *International Journal for Numerical Methods in Fluids.*, **52**, 381-392, 2006.

- [33] Ph. Rosenau, Extending hydrodynamics via the regularization of the Chapman-Enskog expansion, *Phys. Rev.*, A **40**, 7193, 1989.
- [34] P. Sagaut, *Large Eddy Simulations for Incompressible Flows*, Springer 2006.
- [35] S. Schochet, and E. Tadmor, The regularized Chapman-Enskog expansion for scalar conservation laws, *Arch. Rat. Mech. Anal.*, **119**, 95, 1992.
- [36] M.I. Vishik, E.S. Titi and V.V. Chepyzhov, Trajectory attractor approximations of the 3d Navier-Stokes system by the Leray-alpha model, *Russian Math Dokladi*, **71**, 91-95, 2005.



## A simple lumped mass-damper idealization for dam-reservoir- foundation system for seismic analysis

†\*D.K. Paul<sup>1</sup>, A. Banerjee<sup>2,3</sup>, R.N. Dubey<sup>1</sup>, M.K.Alam-Chowdhury<sup>4</sup>, A. P. Singh<sup>4</sup>

<sup>1</sup>Department of Earthquake Engineering, Indian Institute of Technology Roorkee, India

<sup>2</sup>Department of Mechanical Engineering, University of Auckland, New Zealand

<sup>3</sup>Callaghan Innovation, Auckland, New Zealand

<sup>4</sup> Larsen & Toubro Ltd, Infra IC (Hydel BU), Faridabad, India

\*Presenting author: dpaulfeq@gmail.com

†Corresponding author: dpaulfeq@gmail.com

### Abstract

A simple lumped mass-damper idealization of impounded reservoir water is suggested in this paper. The effect of impounded reservoir water was modeled as virtual lumped mass which simulates the hydrodynamic pressure according to Westergaard (1933), but the lumped mass approach does not consider the surface wave generation and wave transmission through the boundaries. On the other hand, acoustic/ water elements can be employed to model the dam-reservoir interaction which can capture the above mentioned effect, but that needs more computational effort and expertise. In this paper, equivalent dampers together with the virtual mass are lumped at the upstream face to simulate the effect of the impounded reservoir water on the dam. A simplified procedure is proposed for calculating the coefficient of dampers using the geometry of the reservoir and impedance of the foundation and dam materials. A comparison of response of lumped mass-damper model with acoustic water model is also presented.

**Keywords:** Lumped Damper approach; Dam-foundation-reservoir interaction; Concrete gravity dams; Compressible fluid element; Incompressible fluid; Absorbing boundary; Seismic analysis

### Introduction

Concrete gravity dams are one of the most important civil engineering structures. Gravity dam stores a large amount of water in the reservoir, which contributes to the destabilizing forces on the dam. That is why, the modeling of the impounded reservoir water and its interaction with the dam and foundation attracts the attention of many researchers.

Reservoir considerably affects the earthquake response of the gravity dam. There are three ways to consider the reservoirs effect in the seismic response analysis of the dam system, for example, Westergaard, Lagrangian and Eulerian approaches. Westergaard approach considers the virtual lumped mass on the wet surface which stimulates the hydrodynamic effect on the upstream face of the dam [Westergaard, 1933; Zanger, 1953]. In Lagrangian Eulerian approach [Bayraktar *et al.*, 2011; Bleich and Sandler, 1970; Calayir *et al.*, 1996; Dunger, 1978; Kalateh and Attarnejad, 2011; Ross *et al.*, 2009; Zienkiewicz *et al.*, 1983] displacements are the variable in structure and pressures are the variable in the fluid. Special compatibility equation is required to establish the compatibility between the reservoir water and the dam-foundation system. In Lagrangian approach displacement is the variable for both impounded reservoir water and the dam-foundation system. No compatibility equation is required for Lagrangian approach [Hamdan, 1999; Wilson and Khalvati, 1983]. On the other hand, the performance of Westergaard model (incompressible water) and Acoustic water (compressible water) model were compared [Banerjee *et al.*, 2014]. It is observed that when the water is considered incompressible, the response is much higher as compared to when the water is considered compressible. To improve the results of Westergaard model, a simple lumped mass-damper model is proposed. A simplified lumped mass-damper model

From the results of dam response considering incompressible reservoir water and compressible water [Banerjee et al., 2014], it is clear that the lumped mass model (Westergaard model considering water as incompressible) gives almost 18-20% higher response than considering the reservoir water as compressible allowing for wave propagation and transmission of waves through the radiating boundary. Actually, p-waves propagated through the impounded reservoir water from the vibrating dam. As the ratio of energy transmission from one layer to another depends on the impedance ( $Z$ ) of the layers, so the transmitted energy from the reservoir by different boundaries also depends on the shape of the reservoir and the impedance of the boundary. That is why, the wave propagation phenomena can be considered in the model by multiplying the impedance of water with the ratio of energy transmitted from the impounded water with the energy transmitting at the dam-reservoir interface.

*Dam sections*

Two high dams of different heights are designed using Optidam software [Banerjee et al., 2015] are taken as examples. One is on the varying rock strata and another is on the single rock profile. The two dams are shown in Figure.1.

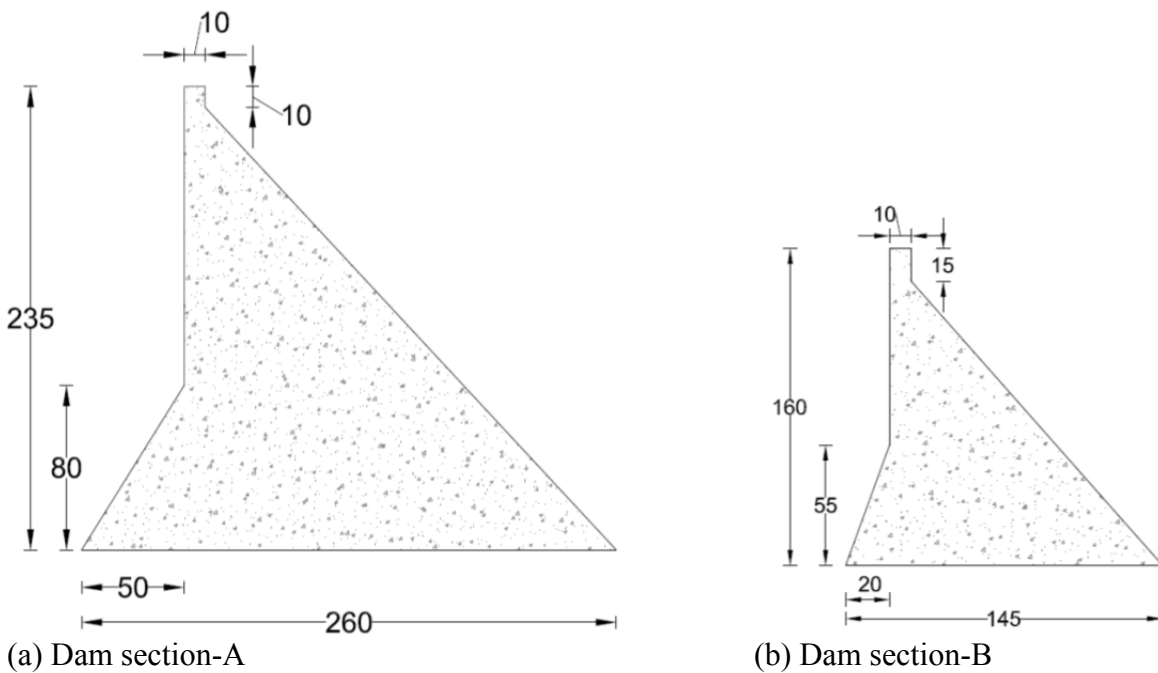


Figure.1 Two dam sections

The geometrical properties of the dams are given in Table-1.

Table-1 Geometric properties

	Height	Base width	Upstream slope	Downstream slope
Dam-A	235	260	1:0.625	1:0.9
Dam-B	160	145	1:0.360	1:0.8

The peak ground acceleration at the site is taken as 0.2g. The dam section –A is assumed to be on the layered rock strata. But the dam section-B is taken on single layer of rock. The property of concrete is given in Table-2.

Table-2 Property of Concrete

	Young Modulus (GPa)	Poisson Ratio	Density (kg/m <sup>3</sup> )
Dam –A	30	0.15	2500
Dam –B	22	0.15	2400

The geotechnical data taken in the analysis is tabulated in Table.3.

Table-3 Geotechnical data

Dam Section	Rock layer	Depth (m)	Young Modulus (GPa)	Poisson Ratio	Density (kg/m <sup>3</sup> )
Dam –A	Alluvium/Boulder	0-5	2	0.30	1800
	Fragmented Rock	5-20	7	0.32	2400
	Rock	20-250	10	0.32	2600
Dam –B	Rock	0-250	17	0.16	2000

There are two Philite strips of 10m width present at 60m and 160m from the upstream face of the dam. The two strips are parallel to each other making 22° with the horizontal. The property of philite strips is same as the fragmented rock layer for Dam-A. Bulk modulus and density of water is taken 2.07 GPa and 1000 kg/m<sup>3</sup>. Damping values of the different material are tabulated in Table-4.

Table.4 Damping Value (%)

Material	Concrete	Alluvium/Boulder	Intact Rock	Fragmented Rock
Dam-A	5	10	7	10
Dam-B	5	-	7	-

#### Computation of dampers coefficient

The energy loss from the non-reflecting boundary of the reservoir is not considered in Westergaard lumped mass model if the mass distribution is as per Eq(1).

$$\left( \begin{array}{l} C_w = \frac{C_m}{2} \left\{ \frac{y}{h} \left( 2 - \frac{y}{h} \right) + \sqrt{\frac{y}{h} \left( 2 - \frac{y}{h} \right)} \right\} \\ M_i = C_w \gamma_w h b (y_i - y_{i-1}) \end{array} \right) \quad (1)$$

To consider the energy loss, lumped dampers are attached at the upstream face of the dam along with the virtual mass due to incompressible water. The coefficient of the damper is calculated based on the energy transmission from different boundaries (except the dam-reservoir interface). Actually, the dampers are used to consider the ratio of energy going out from the radiation boundary and the energy impinging on the dam-reservoir interface. To compute the transmitted energy the following procedure as shown in Fig.2 is adopted.

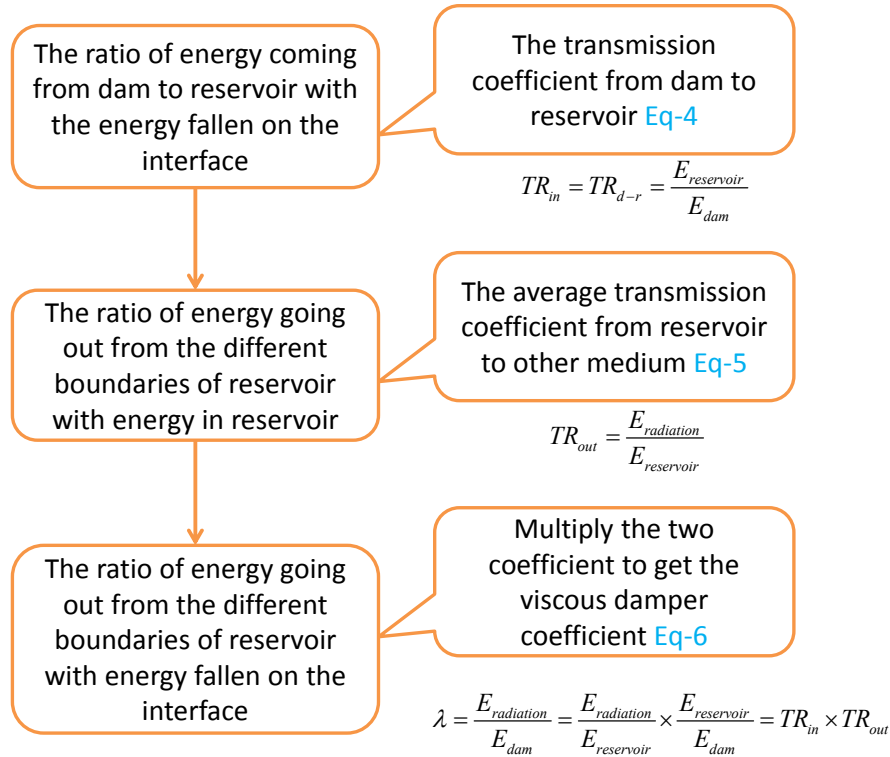


Figure.2 process to calculate the damping coefficient

The ratio of energy transmitted from one medium to the other is computed by

$$TR = 1 - \left( \frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2 \quad (2)$$

where,  $Z$  is the impedance of a medium depends on the p wave velocity ( $V_p$ ) and the density ( $\rho$ ) of the layer. Here s- wave velocity is not considered because s wave cannot propagate through the water.

$$Z = \rho V_p \quad (3)$$

The impedance of various layers are tabulated in Table-5 according to Eq (3) for two dams A & B.

Table-5 Impedance of different layers

Material	Dam-A			Dam-B		
	Density ( $\rho$ ) (kg/m <sup>3</sup> )	P wave velocity ( $V_p$ ) (m/s)	Impedance ( $Z$ ) (kg m <sup>-2</sup> s <sup>-1</sup> ) x10 <sup>6</sup>	Density ( $\rho$ ) (kg/m <sup>3</sup> )	P wave velocity ( $V_p$ ) (m/s)	Impedance ( $Z$ ) (kg m <sup>-2</sup> s <sup>-1</sup> ) x10 <sup>6</sup>
Concrete	2500	3500.0	8.75	2400	3133	7.52
Rock	2600	2342.3	6.09	2000	3010	6.02
Water	1000	1440.0	1.44	1000	1440	1.44

The ratio of total energy transmitted from dam to reservoir  $R_{d-r}$  and the total energy impinges on the interface ( $TR_{in}$ ) should be equal, i.e.

$$TR_{in} = TR_{d-r} \quad (4)$$

The ratio of average energy transmitted out ( $TR_{out}$ ) from the different boundaries of the reservoir is calculated as:

$$TR_{out} = \frac{\sum_{i=1}^n TR_i \cdot l_i}{\sum_{i=1}^n l_i} \quad (5)$$

So, the ratio of energy transmitted from different boundaries of the reservoir and the total energy impinges on the interface of the dam reservoir system ( $\lambda$ ) is calculated by

$$\lambda = TR_{in} \cdot TR_{out} \quad (6)$$

The length and ratio of transmitted energy from different boundaries of reservoir are tabulated in Table.6.

Table-6. Properties of boundary

	Reservoir Boundaries	$Z_1$ (kg m <sup>-2</sup> s <sup>-1</sup> ) x10 <sup>6</sup>	$Z_2$ (kg m <sup>-2</sup> s <sup>-1</sup> ) x10 <sup>6</sup>	$TR$	length (m)	$TR \cdot l$
<b>Dam-A</b>	Free surface	1.44	0.00	0.00	550	0.0
	Infinite boundary	1.44	1.44	1.00	215	215.0
	Reservoir- Dam	1.44	8.75	0.48	230	110.4
	Reservoir- Foundation	1.44	6.09	0.61	512	312.3
<b>Dam-B</b>	Free surface	1.44	0.00	0.00	320	0.0
	Infinite boundary	1.44	1.44	1.00	140	140.0
	Reservoir- Dam	1.44	7.52	0.53	150	79.5
	Reservoir- Foundation	1.44	6.02	0.61	300	183.0

Where  $Z_1$  and  $Z_2$  are the impedances before and after a boundary. The co-efficient of virtual lumped damper at the upstream face is calculated by

$$C = \lambda c \rho_w A \quad (7)$$

where,  $C$  is the coefficient of a damper,  $c$  is the velocity of sound in water,  $\rho_w$  is the density of impounded reservoir water, and  $A$  is the tributary area of single damper. The evaluation of  $C$  is worked out in Table 7.

Table.7 Coefficient of dampers

Dam	$TE_{in}$	$TE_{out}$	$\lambda$	$A$ (m <sup>2</sup> )	$C$ (N s m <sup>-1</sup> ) x10 <sup>6</sup>
Dam-A	0.48	0.42	0.20	75	21.60
Dam-B	0.53	0.44	0.23	75	24.84

#### *Different dam models*

The seismic response obtained by the proposed simplified method is compared with that of the compressible water model as well as incompressible water model. The three dam-reservoir models

are (i) with incompressible reservoir water (Westergaard lumped mass idealization, Fig. 3(a)), (ii) with compressible reservoir water (exact numerical model, Fig. 3(b)) and (iii) lumped mass-damper model (Fig.3(c)), are shown below. In the lumped mass-damper model, the virtual mass of the water is lumped on the wet surface in a similar fashion as if the reservoir water is treated incompressible and the energy used in surface waves and the radiation damping is replaced by equivalent dampers lumped on the wet surface.

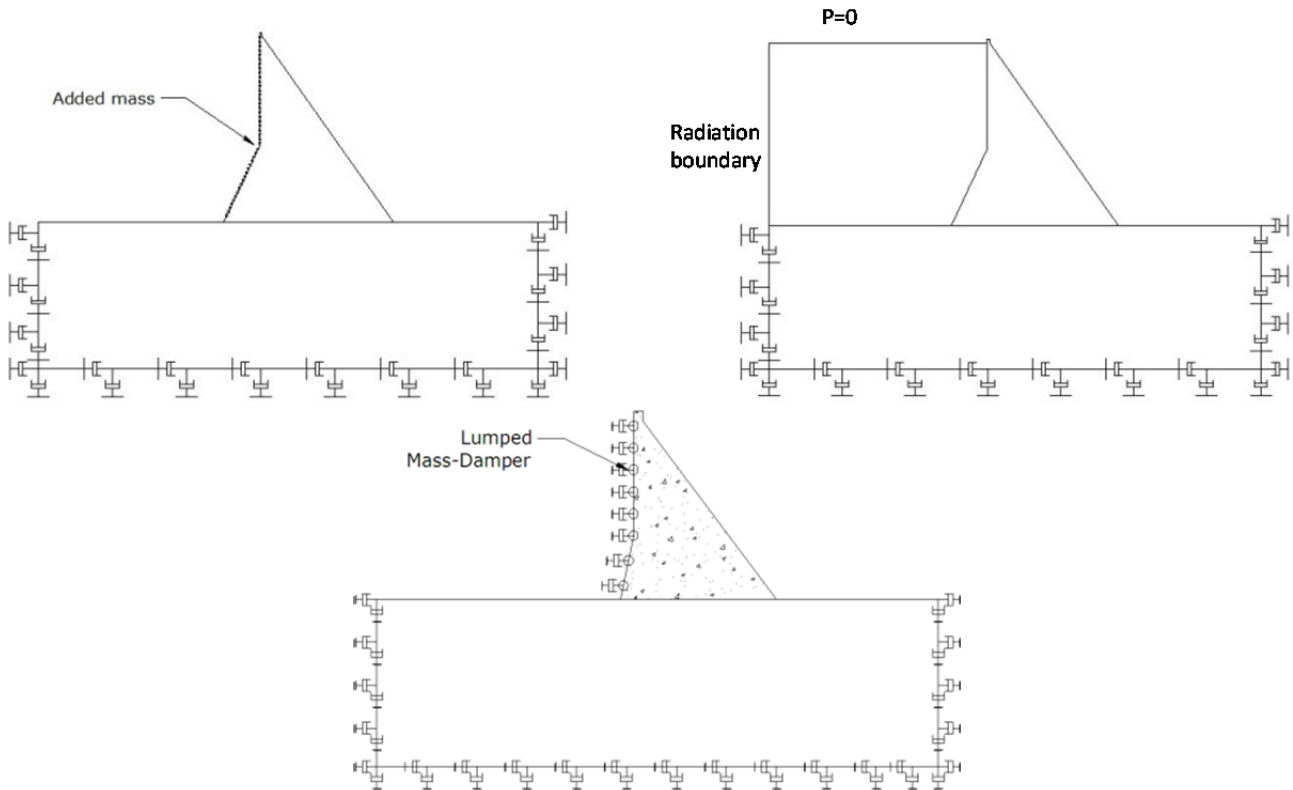
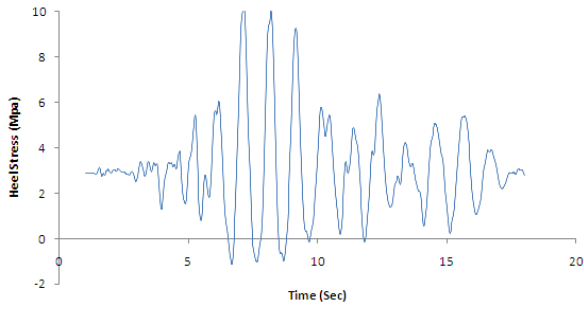


Figure-3 Different dam models

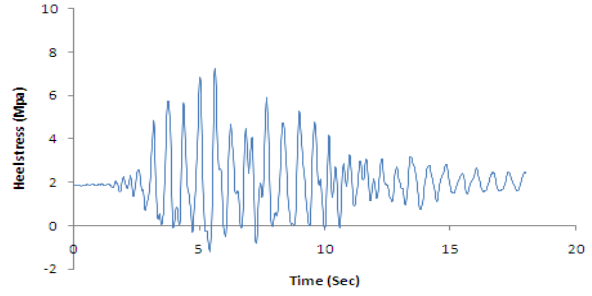
*Response of dams*

Baseline corrected deconvoluted accelerogram is applied at the base of the considered foundation. The stresses at the heel of the two dam models are plotted in Fig.4.

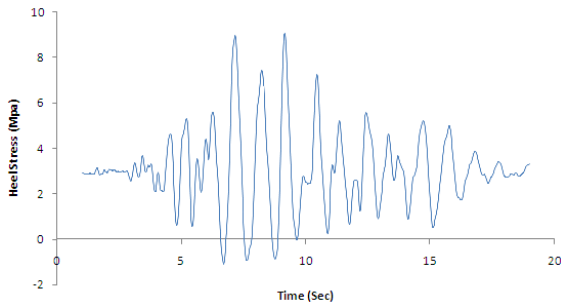
In Model-1, the incompressible water is replaced by virtual added mass on the wet surface where no radiation damping is considered at the truncated reservoir boundary. In Model-2 surface waves can be generated and the energy loss due to the outgoing waves at the truncated reservoir boundary attributes for lesser response when water is considered compressible. In Model-3 impounded reservoir water is replaced by lumped mass to consider the inertia of the water and lumped dampers to stimulate the energy dissipation from the reservoir. The principal stresses at the dam heel and u/s slope change for the Model-3 are comparable with Model-2.



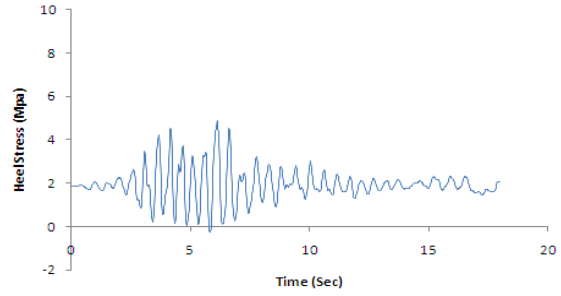
(a) Dam-A, Model-1 (Max 10.10 MPa)



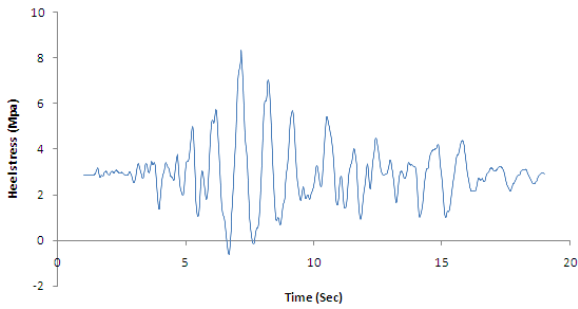
(b) Dam-B, Model-1 (Max 6.85MPa)



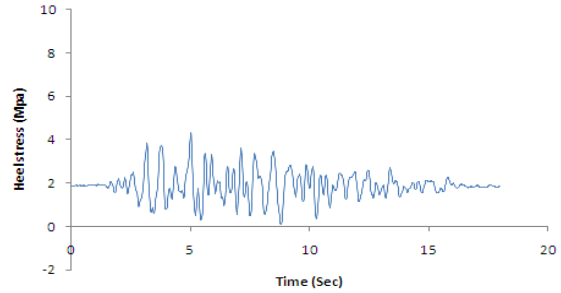
(b) Dam-A, Model-2 (Max 8.90MPa)



(d) Dam-B, Model-2 (Max 4.90MPa)



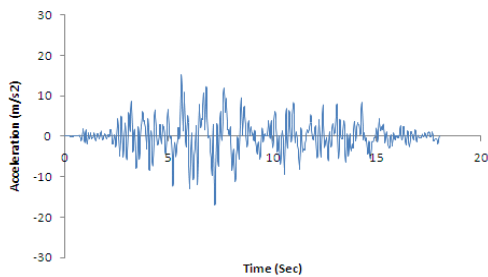
(e) Dam-A, Model-3 (Max 8.40MPa)



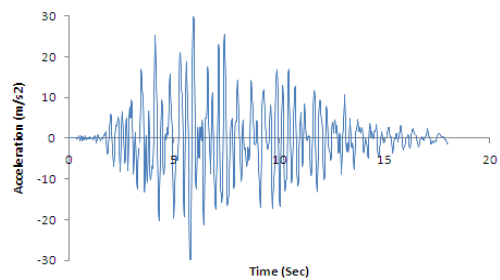
(f) Dam-B, Model-3 (Max 4.30MPa)

Figure-4 Principal stresses at heel

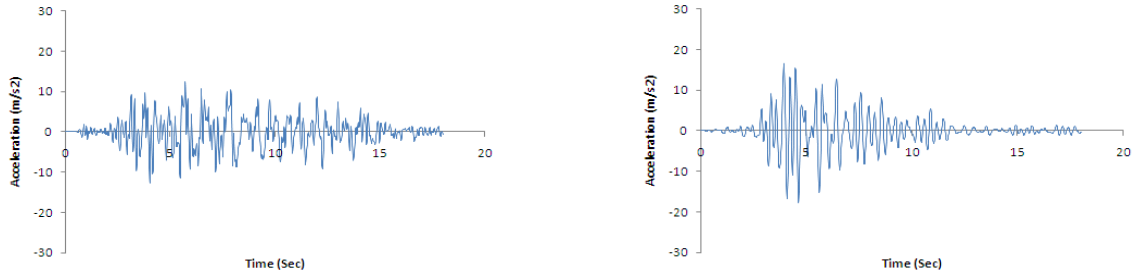
The maximum stresses at the heel and u/s slope change are reduced by 10-15% due to the fluid compressibility and energy dissipation through the boundaries. But the stress at the toe is not so much affected due to the compressibility of impounded water. Plot of dam crest acceleration for different models are plotted in Fig.5.



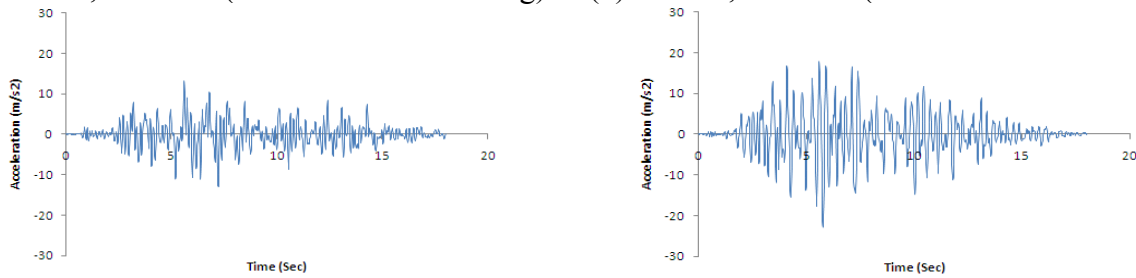
(a) Dam-A, Model-1 (Max acceleration 1.69g)



(b) Dam-B, Model-1 (Max acceleration 2.90g)



(c) Dam-A, Model-2 (Max acceleration 1.27g)      (d) Dam-B, Model-2 (Max acceleration 1.80g)



(e) Dam-A, Model-3 (Max acceleration 1.29g)      (f) Dam-B, Model-3 (Max acceleration 2.10g)

Figure.5 Crest acceleration for different cases

Crest acceleration for Model I is higher but it is almost same for the Model-2 and Model-3. The peak acceleration at different locations of the dam-foundation-reservoir system due to compressible and incompressible water models are tabulated in Table-8.

Table.8 Maximum acceleration and stresses at different locations of dam

Dam model	Model no.	Crest	Upstream slope		Heel	
		Acceleration (g)	Acceleration (g)	Stress (MPa)	Acceleration (g)	Stress (MPa)
Dam-A	Model-1	1.69	0.40	4.3	0.30	10.1
	Model-2	1.27	0.33	2.3	0.26	8.9
	Model-3	1.29	0.32	1.8	0.26	8.4
Dam-B	Model-1	2.90	0.70	4.0	0.43	6.8
	Model-2	1.75	0.28	2.0	0.24	4.9
	Model-3	1.80	0.30	1.9	0.28	4.3

Acceleration at crest, heel and upstream slope change location are reduced almost 18-24% due to the influence of compressibility of water. The virtual lumped mass model (Model I) gives result on a conservative side. The stress computed at the heel of the dam in Model-3 is slightly less than the stress calculated in Model-2 whereas acceleration and stresses at the u/s face are comparable with Model-2.

## Conclusions

In this paper, an approach for approximation dam-reservoir system using a simplified lumped mass-damper model is suggested, where the energy loss due to radiation is modelled as lumped dampers on the upstream wet surface. Radiation damping due to outgoing waves is stimulated by lumping virtual dampers. The computation for the coefficient of lumped dampers depends on the density of water, sound wave velocity through water, tributary area of a single damper and the ratio of incoming energy and outgoing energy from the reservoir. The virtual lumped mass-damper



approach is applicable for both the layered media and homogeneous rock profile. The stresses and accelerations of the compressible water model and virtual lumped mass-damper model are comparable. The wave propagation is considered only in perpendicular to the boundary for the analysis. The simplified virtual lumped mass-damper model is found to be a good approximation for the response of dam reservoir foundation system for preliminary design of dam.

## References

- Banerjee, A., Paul, D.K., Acharyya, A., (2015) Optimization and safety evaluation of concrete gravity dam section, *KSCE Journal of Civil Engineering*, DOI:10.1007/s12205-015-0139-0
- Banerjee, A., Paul, D.K., Dubey, R.N., (2014) Modelling issues in seismic analysis of concrete gravity dam, *Dam Engineering* **XXIV**(2), 87-109
- Bayraktar, A., Sevim, B., Altunisk, A.C., (2011) Finite Element Model Updating Effects on Nonlinear Seismic Response of Arch Dam-reservoir-foundation Systems, *Finite Element in Analysis & Design* **47** (2) 85-97.
- Bleich, H.H., Sandler, I.S., (1970) Interaction Between Structures and Bilinear Fluids', *International Journal of Solids & Structures* **6**(5), 617-639
- Calayir, Y, Dumanoglu, A A & Bayraktar, A, (1996) Earthquake Analysis of Gravity Dam reservoir Systems Using the Eulerian and Lagrangian Approach, *Computers & Structures* **59**(5) 877-890
- Dungar, R, (1978) An Efficient Method of Fluid-structure Coupling in the Dynamic Analysis of Structures, *International Journal for Numerical Methods in Engineering* **13**(1), 93-107
- Hamdan, F H, (1999) Near-field Fluid-structure Interaction Using Lagrangian Fluid Finite Element', *Computers & Structures*, **71**(2), 123-141.
- Kalateh, F & Attarnejad, R, (2011) Finite Element Simulation of Acoustic Cavitation in the Reservoir and Effects on Dynamic Response of Concrete Dams, *Finite Element in Analysis & Design*, **47**(5), 543-558
- Ross, M.R., Sprague, M.A., Felippa, C.A. & Par, K.C., (2009) Treatment of Acoustic Fluid structure Interaction by Localized Lagrange Multipliers and Comparison to Alternative Interface-coupling Methods, *Computer Methods in Applied Mechanics & Engineering*, **198**( 9-12), 986-1005
- Westergaard, H M, (1933) Water Pressures on Dams During Earthquakes, *Transactions,ASCE*, **98**(2), 418-433
- Wilson, E L & Khalvati, M, (1983) Finite Elements for the Dynamic Analysis of Fluid-solid Systems, *International Journal for Numerical Methods in Engineering*, **19**(11), 1657-1668 .
- Zanger, C N, (1953) Hydrodynamic Pressure on Dam Due to Horizontal Earthquake, *Proc. Soc. Exp. Stress Anal.* **10**, 93-102
- Zienkiewicz, O.C., Paul, D.K. & Hinton, E., (1983) Cavitations in Fluid Structure Response (with particular reference to dam on earthquake loading), *Earthquake Engineering & Structural Dynamics*, **11**(4), 463-481

# **Ball's motion, sliding friction and internal load distribution in a high-speed ball bearing subjected to a combined radial, thrust, and moment load**

**Mário César Ricci**

Space Mechanics and Control Division, National Institute for Space Research, Av. dos Astronautas, 1758, 12227-010, São José dos Campos, SP, Brazil

E-mail: mariocesarricci@uol.com.br

## **Abstract**

A set of non-linear algebraic equations, which must to be solved using a numerical procedure, for ball's motion, sliding friction and internal loading distribution computation in a high-speed, single-row, angular-contact ball bearing, subjected to a known combined radial, thrust and moment load, which must be applied to the inner ring's centre of mass, is introduced. For each step of the procedure it is required the iterative solution of  $9Z + 3$  simultaneous non-linear equations – where  $Z$  is the number of the balls – to yield exact solution for contact angles, ball attitude angles, rolling radii, normal contact deformations and axial, radial, and angular deflections of the inner ring with respect the outer ring. The focus of this work is obtaining the steady state forces and moments equilibrium conditions on the balls, under the selected external loading, and to describe the numerical aspects of the procedure. The numerical results derived from the described procedure shall be published later.

**Keywords:** ball, bearing, high-speed, load, distribution

## **Introduction**

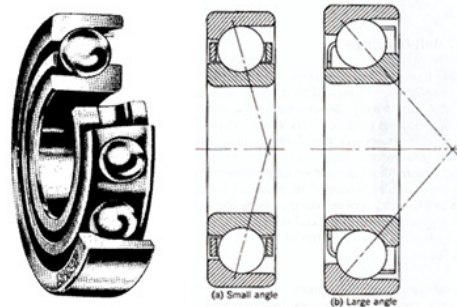
Ball and roller bearings, generically called *rolling bearings*, are commonly used machine elements. They are employed to permit rotary motions of, or about, shafts in simple commercial devices such as bicycles, roller skates, and electric motors. They are also used in complex engineering mechanisms such as aircraft gas turbines, rolling mills, dental drills, gyroscopes, and power transmissions.

The standardized forms of ball or roller bearings permit rotary motion between two machine elements and always include a complement of ball or rollers that maintain the shaft and a usually stationary supporting structure, frequently called *housing*, in a radially or axially spaced-apart relationship. Usually, a bearing may be obtained as a unit, which includes two steel rings each of which has a hardened raceway on which hardened balls or rollers roll. The balls or rollers, also called *rolling elements*, are usually held in an angularly spaced relationship by a *cage*, also called a *separator* or *retainer*.

There are many different kinds of rolling bearings. This work is concerned with *single-row angular-contact ball bearings* – see Fig. 1 – which are designed to support combined radial and thrust loads or heavy thrust loads depending on the *contact angle* magnitude. The bearings having large contact angle can support heavier thrust loads. The figure 1 shows bearings having small and large contact angles. The bearings generally have groove curvature radii in the range of 52-53% of the ball diameter. The contact angle does not usually exceed  $40^\circ$ .

This work is devoted to study of internal load distribution in a *high-speed* angular-contact ball bearing. Several researchers have studied the subject of internal load distribution in a *statically loaded* angular-contact ball bearing (see [Stribeck (1907); Sjöväll (1933); Jones (1946); Rumbarger (1962); Ricci (2009; 2009a; 2009b; 2009c; 2009d; 2010)]). The methods developed by them to calculate distribution of load among the balls and rollers of rolling bearings can be used in most bearing applications because rotational speeds are usually slow to moderate. Under these speed conditions, the effects of rolling element centrifugal forces and gyroscopic moments are negligible. At high

speeds of rotation these body forces become significant, tending to alter contact angles and clearance. Thus, they can affect the static load distribution to a great extension.



**Figure 1. Angular-contact ball bearing**

[Harris (2001)] described methods for internal loading distribution in statically loaded bearings addressing pure radial; pure thrust (centric and eccentric loads); combined radial and thrust load, which uses radial and thrust integrals introduced by Sjöväll; and for ball bearings under combined radial, thrust, and moment load, initially due to Jones.

When an external load is applied to one of the rings of a rolling bearing it is transmitted through rolling elements to the other ring. Because the internal load distribution on the rolling elements is an important operating characteristic of a bearing a great number of authors have addressing the problem. A literature review on the subject can be found in [Tomović (2012; 2012a)], in which a mathematical model for necessary radial displacement between rings, and a mathematical model for external radial load, so that the  $q$ -th rolling element passes to participate in the load transfer were presented.

In [Tomović (2013)], a model was developed, which enables a very simple determination of the number of active rolling elements participating in an external load transfer, depending on the bearing type and internal radial clearance.

In [Tudose et al. (2013)], the theoretical analysis of a single-row radial bearing with radial clearance under constant external radial load was presented. The analysis was focused on finding the rolling element deflection that allows determining the number of active rolling elements that participate in the load transfer. Taking into account the bearing internal geometry, a mathematical model to calculate the rolling elements deflections during the bearing rotation has been derived.

In [Rasolofondraibe et. al. (2012; 2013); Murer et al. (2015)], capacitive probes were inserted into the fixed ring of the bearing such that forms with the raceway a capacitor with variable gap that depends on the transmitted load by the rolling element. A numerical model of this capacitor's capacitance as a function of transmitted load by the rolling element has been established. An experimental prototype has been established in order to precisely measure the probe's capacitance. Finally, this technique has been generalized with a capacitive probe in front of each rolling element. Thus, knowing the load transmitted by each of the rolling elements, the external load on the bearing of the rotating machine can be easily reconstructed.

The evaluation of change in contact angle due to applied load is vital in order to study the load carrying capacity of large diameter bearings. Analytical and numerical procedures have been developed to calculate various design factors such as contact angle, contact stress and deformation. In [Starvin et al. (2011)] the change in contact angle of balls was determined by using FEA. The change in contact angle was compared with analytical, FEA and the Newton-Raphson method. The results show a good agreement with the values calculated using Hertz's relations for deformation. The

FEA method was used to get the nodal solution of contact angle, contact Stress and deflection for various loading conditions.

In [Rajasekhar et al. (2013)] the dynamic modeling of a centrally supported symmetrical disk-shaft bearing system has been analyzed using Timoshenko beam elements. Intermittent ball bearing contact forces and Muszynska's force [Muszynska (1986)] at seal-disk interface were considered in the model to simulate a real-time system. Results show that there was a marked effect of each type of nonlinear excitation on the overall system response.

In [Seong et al. (2014)] a wheel bearing life prediction method, which considers the bearing dynamics characteristics, was proposed. The results were compared with existing formulas and static analyses results from structural dynamics commercial software.

In [Zhenguo et al. (2011)] a unidirectional compression spring was used to model the contact between a rolling element and the raceway of a heavy-duty slewing bearing accounting for the supporting structure flexibility and the plastic deformation of the bearing. The spring constant was determined by the load against elastic-plastic deformation relationship of a single rolling element, which was obtained by finite element contact method. The difference between the traditional Hertz contact results and the FEM results is very obvious for the slewing bearings with plastic deformation, such as contact deflection of the rolling elements and the raceway, load distribution on the rolling elements, stress in the raceway and contact pressure between the rolling elements and the raceway. Therefore, the method based on the Hertz contact mechanics theory is not applicable for the performance analysis of the heavy-duty slewing bearing.

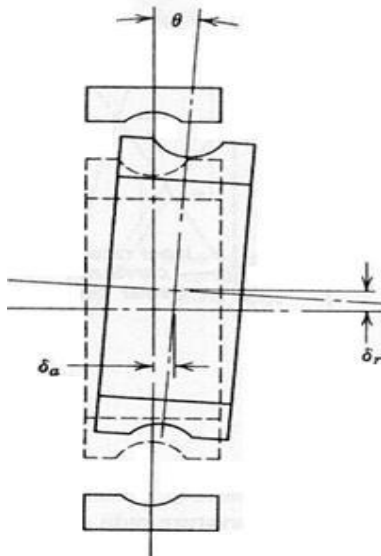
The first great contribution to the study of ball motion, sliding friction and internal load distribution in a *high-speed* angular-contact ball bearing must be credited to A. B. Jones [Jones (1959; 1960)]. Harris describes the orbital, pivotal and spinning ball's motions and load distribution in ball bearings, in general reproducing the Jones's developments. In this work the Jones's works is revisited and differences are introduced under the yoke of critical analysis, which will be detailed. Then, particularly, in this work, a set of non-linear algebraic equations, which must to be solved using a numerical procedure, for ball's motion, sliding friction and internal loading distribution computation in a high-speed, single-row, angular-contact ball bearing, subjected to a known combined radial, thrust and moment load, which must be applied to the inner ring center of mass, is introduced. For each step of the procedure it is required the iterative solution of  $9Z + 3$  simultaneous non-linear equations – where  $Z$  is the number of the balls – to yield exact solution for contact angles, ball attitude angles, rolling radii, normal contact deformations and axial, radial, and angular deflections of the inner ring with respect the outer ring. The focus of this work is obtaining the steady state forces and moments equilibrium conditions on the balls, under the selected external loading, and to describe the numerical aspects of the procedure. The numerical results derived from the described procedure shall be published later.

### Mathematical model

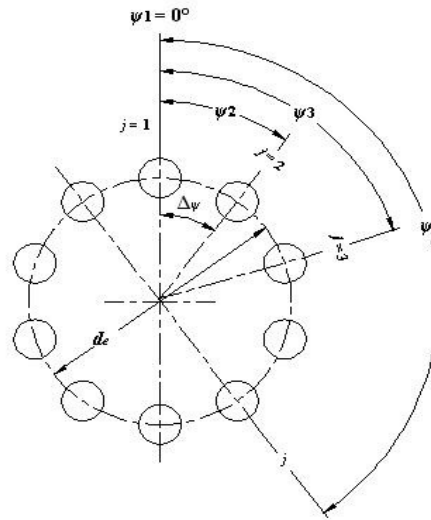
Having defined in other works analytical expressions for bearing geometry and the contact stress and deformations for a given ball or roller-raceway contact (point or line loading) in terms of load (see, e.g., [Harris (2001)]) it is possible to consider how the bearing load is distributed among the rolling elements. In this section a specific internal loading distribution resulting from a combined radial, thrust, and moment external load, which must be applied to the center of mass of the inner ring of a high speed ball bearing, is considered.

The Fig. 2 shows the displacements of an inner ring related to the outer ring due to a generalized loading system including radial, axial, and moment loads. The Fig. 3 shows the relative angular position of each ball in the bearing.

Let a ball bearing with  $Z$  balls, each with diameter  $D$ , symmetrically distributed about a pitch circle according to Fig. 3, to be subjected to a combined radial, thrust, and moment load applied to the inner ring's center of mass. Then, a *relative axial displacement*,  $\delta_a$ , a *relative angular displacement*,  $\theta$ , and a *relative radial displacement*,  $\delta_r$ , between the inner and outer ring raceways may be expected according Fig. 2. Let  $\psi = 0$  to be the angular position of the maximum loaded ball.



**Figure 2. Displacements of an inner ring (outer ring fixed) due to a combined radial, axial, and moment external loading**



**Figure 3. Ball angular positions in the radial plane that is perpendicular to the bearing's axis of rotation,  $\Delta\psi = 2\pi/Z$ ,  $\psi_j = 2\pi(j-1)/Z$ ,  $j = 1 \dots Z$ , in which  $Z$  is the number of balls**

Under zero load the centers of raceway groove curvature radii are separated by a distance  $A$  given by

$$A = (f_o + f_i - 1)D, \tag{1}$$

in which  $f_o, f_i$  are the conformities for outer and inner raceways, respectively.

Under an applied static load, the distance  $s$  between centers will increase from  $A$  to  $A$  plus the amount of the contact deformation  $\delta_i$  plus  $\delta_o$ , as show by Fig. 4. The line of action between centers is collinear with  $A$ . If, however, a centrifugal force acts on the ball, then because the inner and outer raceway contact angles are dissimilar, the line of action between raceway groove curvature radii centers is not collinear with  $A$ , but is discontinuous as indicated by Fig. 5. It is assumed in Fig. 5 that the outer raceway groove curvature center is fixed in space and the inner raceway groove curvature center moves relative to that fixed center. Moreover, the ball center shifts by virtue of the dissimilar contact angles.

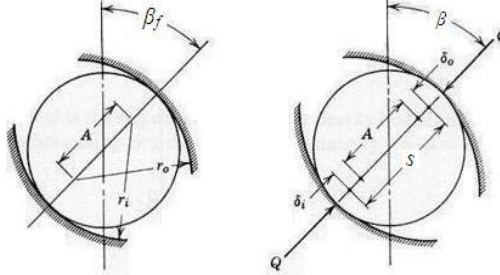
The Fig. 5 when compared with similar figures in [Harris (2001)] and [Jones (1960)] shows minor differences. The inner contact angle must be  $\beta_{ij} + \theta \cos \psi_j$  rather than  $\beta_{ij}$ , to take into account the tilting of the rigid inner ring with respect the rigid outer ring, during the external loading application. Furthermore, since the problem is to be solved numerically, no makes sense to linearize the distances between the final and initial inner raceway groove curvature center positions, as done in previous works.

In accordance with Fig. 5 the distance between the fixed outer raceway groove curvature center and the final position of the ball center at any ball location  $j$  is

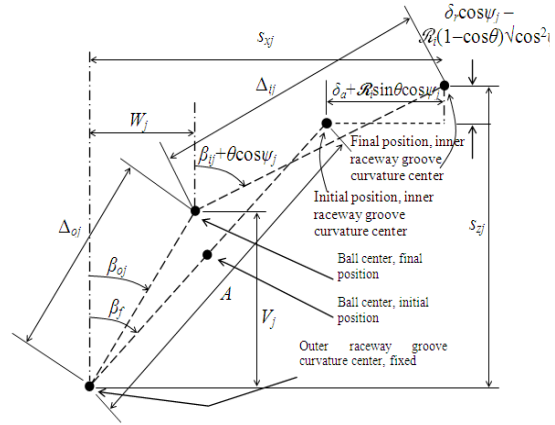
Since  $r_o = f_o D$ ,

$$\Delta_{oj} = r_o - \frac{D}{2} + \delta_{oj}. \quad (2)$$

$$\Delta_{oj} = (f_o - 0.5)D + \delta_{oj}. \quad (3)$$



**Figure 4. (a) Ball-raceway contact before loading; (b) Ball-raceway contact under load.**



**Figure 5. Positions of ball center and raceway groove curvature centers at angular position  $\psi_j$  with and without applied load.**

Similarly, the distance between the moving inner raceway groove curvature center and the final position of the ball center at any ball location  $j$  is

$$\Delta_{ij} = (f_i - 0.5)D + \delta_{ij}, \quad (4)$$

in which  $\delta_{oj}$  and  $\delta_{ij}$  are the normal contact deformations at the outer and inner raceway contacts, respectively.

In accordance with the relative axial displacement between inner and outer rings mass centers,  $\delta_a$ , and the relative angular displacement  $\theta$ , the axial distance between inner and outer raceway groove curvature centers at ball position  $j$  is

$$s_{xj} = A \sin \beta_f + \delta_a + \mathcal{R}_i \sin \theta \cos \psi_j, \quad (5)$$

in which

$$\mathcal{R}_i = \frac{1}{2}d_e + (f_i - \frac{1}{2})D \cos \beta_f \quad (6)$$

is the radius to locus of inner raceway groove curvature centers,  $d_e$  is the unloaded pitch diameter, and  $\beta_f$  is the unloaded contact angle. Further, in accordance with the relative radial displacement between inner and outer rings mass centers,  $\delta_r$ , and the relative angular displacement  $\theta$ , the radial distance between inner and outer groove curvature centers at each ball location  $j$  is

$$s_{zj} = A \cos \beta_f + \delta_r \cos \psi_j - \mathcal{R}_i (1 - \cos \theta) \sqrt{\cos^2 \psi_j}. \quad (7)$$

Since the iterative techniques of the Newton-Raphson method will be used to solve the associated nonlinear equations, the angles  $\beta_{oj}$  and  $\beta_{ij}$  are best stated in terms of the coordinates  $V$  and  $W$ , in Fig. 5. Then

$$\sin \beta_{oj} = \frac{W_j}{(f_o - 0.5)D + \delta_{oj}}, \quad (8)$$

$$\cos\beta_{oj} = \frac{V_j}{(f_o-0.5)D+\delta_{oj}}, \quad (9)$$

$$\sin(\beta_{ij} + \theta\cos\psi_j) = \frac{s_{xj}-W_j}{(f_i-0.5)D+\delta_{ij}}, \quad (10)$$

$$\cos(\beta_{ij} + \theta\cos\psi_j) = \frac{s_{zj}-V_j}{(f_i-0.5)D+\delta_{ij}}. \quad (11)$$

Similarly, the ball angular speed about its own center pitch and yaw angles,  $\alpha_j$  and  $\alpha'_j$ , are best stated in terms of the ball angular velocity components:  $\omega_{x'j}$ ,  $\omega_{y'j}$ , and  $\omega_{z'j}$ ; in which  $x'$ ,  $y'$ , and  $z'$  are the axes of the coordinate frame whose origin is at the ball center;  $x'$  is parallel to the longitudinal axis of the bearing around which the balls circulate in its orbital motion, and  $z'$  is the radial axis. Then

$$\sin\alpha_j = \frac{\omega_{z'j}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}}, \quad (12)$$

$$\cos\alpha_j = \frac{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}}, \quad (13)$$

$$\sin\alpha'_j = \frac{\omega_{y'j}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2}}, \quad (14)$$

$$\cos\alpha'_j = \frac{\omega_{x'j}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2}}. \quad (15)$$

Using the Pythagorean Theorem, it can be seen from figure 5 that

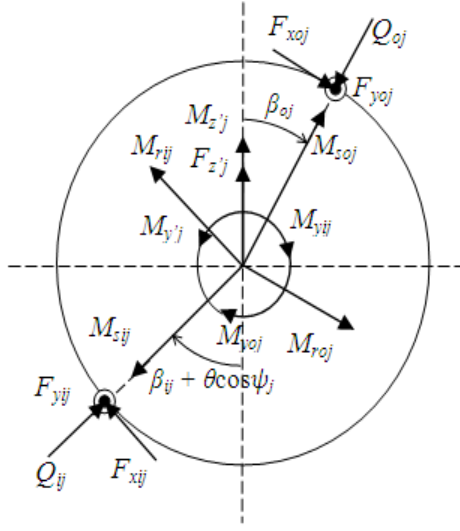
$$(s_{zj} - V_j)^2 + (s_{xj} - W_j)^2 - [(f_i - 0.5)D + \delta_{ij}]^2 = 0 = \epsilon_j, \quad (16)$$

$$V_j^2 + W_j^2 - [(f_o - 0.5)D + \delta_{oj}]^2 = 0 = \epsilon_{j+z}. \quad (17)$$

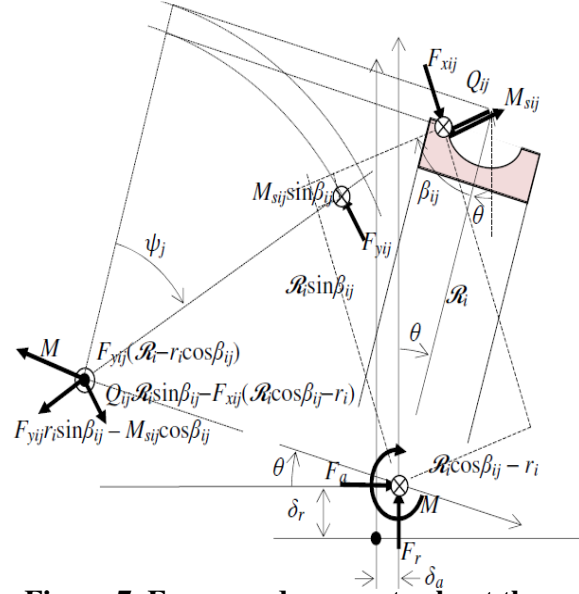
From (12)-(15)

$$\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2 - \omega_{Rj}^2 = 0 = \epsilon_{j+2Z}, \quad \omega_{Rj} = \sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}. \quad (18)$$

For steady state operation of a ball bearing at high speed, the forces and moments acting on each ball are as shown by figure 6.



**Figure 6. Ball loading at angular position  $\psi_j$**



**Figure 7. Forces and moments about the inner ring center of mass**

The normal ball loads are related to normal contact deformations by

$$Q_{oj} = K_{oj}\delta_{oj}^{1.5}, \quad Q_{ij} = K_{ij}\delta_{ij}^{1.5}, \quad (19)$$

in which  $K_{oj}$  and  $K_{ij}$  are functions of contact angles [Harris (2001)].

From Fig. 6 considering the three axes equilibrium forces, yields

$$Q_{ij} \sin(\beta_{ij} + \theta \cos \psi_j) - Q_{oj} \sin \beta_{oj} - F_{xij} \cos(\beta_{ij} + \theta \cos \psi_j) + F_{xoj} \cos \beta_{oj} = 0, \quad (20)$$

$$Q_{ij} \cos(\beta_{ij} + \theta \cos \psi_j) - Q_{oj} \cos \beta_{oj} + F_{xij} \sin(\beta_{ij} + \theta \cos \psi_j) - F_{xoj} \sin \beta_{oj} + F_{z'j} = 0, \quad (21)$$

$$F_{yoj} + F_{yij} = 0 = \epsilon_{j+5Z}, \quad (22)$$

Substituting (8)-(11) and (19) into (20)-(21) yields

$$\frac{F_{xoj}V_j - K_{oj}\delta_{oj}^{1.5}W_j}{(f_o - 0.5)D + \delta_{oj}} + \frac{K_{ij}\delta_{ij}^{1.5}(s_{xj} - W_j) - F_{xij}(s_{zj} - V_j)}{(f_i - 0.5)D + \delta_{ij}} = 0 = \epsilon_{j+3Z}, \quad (23)$$

$$\frac{K_{oj}\delta_{oj}^{1.5}V_j + F_{xoj}W_j}{(f_o - 0.5)D + \delta_{oj}} - \frac{K_{ij}\delta_{ij}^{1.5}(s_{zj} - V_j) + F_{xij}(s_{xj} - W_j)}{(f_i - 0.5)D + \delta_{ij}} - F_{z'j} = 0 = \epsilon_{j+4Z}. \quad (24)$$

From Fig. 6 considering the three axes equilibrium moments, yields

$$-M_{sij} \sin(\beta_{ij} + \theta \cos \psi_j) + M_{soj} \sin \beta_{oj} - M_{rij} \cos(\beta_{ij} + \theta \cos \psi_j) + M_{roj} \cos \beta_{oj} = 0, \quad (25)$$

$$-M_{sij} \cos(\beta_{ij} + \theta \cos \psi_j) + M_{soj} \cos \beta_{oj} + M_{rij} \sin(\beta_{ij} + \theta \cos \psi_j) - M_{roj} \sin \beta_{oj} + M_{z'j} = 0, \quad (26)$$

$$M_{y'j} - M_{yij} - M_{yoj} = 0 = \epsilon_{j+8Z}. \quad (27)$$

Substituting (8)-(11) into (25)-(26) yields

$$\frac{M_{roj}V_j + M_{soj}W_j}{(f_o - 0.5)D + \delta_{oj}} - \frac{M_{sij}(s_{xj} - W_j) + M_{rij}(s_{zj} - V_j)}{(f_i - 0.5)D + \delta_{ij}} = 0 = \epsilon_{j+6Z}, \quad (28)$$



$$\frac{M_{S_{0j}}V_j - M_{R_{0j}}W_j}{(f_o - 0.5)D + \delta_{0j}} - \frac{M_{S_{ij}}(s_{zj} - V_j) - M_{R_{ij}}(s_{xj} - W_j)}{(f_i - 0.5)D + \delta_{ij}} + M_{z'j} = 0 = \epsilon_{j+7Z}. \quad (29)$$

The centrifugal force acting on the ball at angular position  $\psi_j$  is given by [Harris (2001)]

$$F_{z'j} = \frac{1}{2} m d_{mj} \omega_{mj}^2, \quad (30)$$

in which  $m$  is the mass of ball,

$$d_{mj} = d_e + 2[V_j - (f_o - 1/2)D \cos \beta_f] \quad (31)$$

is the operational ball's pitch diameter at position  $j$ , and  $\omega_{mj}$  is the absolute orbital speed of the ball about of the bearing axis.

Substituting the identity  $\omega_{mj}^2 = (\omega_{mj}/\omega)^2 \omega^2$  in (30) gives

$$F_{z'j} = \frac{1}{2} m \omega^2 d_{mj} \left( \frac{\omega_{mj}}{\omega} \right)^2, \quad (32)$$

in which  $\omega$  is the absolute angular velocity of the rotating ring.

For the outer race to be stationary  $\omega_{mj} = -\omega_{oj}$ ,  $\omega = \omega_{ij} + \omega_{mj}$ ,

$$\frac{\omega_{mj}}{\omega} = \frac{1}{\frac{r'_{ij} \left\{ \frac{d_{mj}}{2} [(f_o - 0.5)D + \delta_{0j}] + r'_{oj} V_j \right\} [\omega_{x'j}(s_{zj} - V_j) + \omega_{z'j}(s_{xj} - W_j)]}{1 + \frac{r'_{oj} \left\{ \frac{d_{mj}}{2} [(f_i - 0.5)D + \delta_{ij}] - r'_{ij}(s_{zj} - V_j) \right\} (\omega_{x'j} V_j + \omega_{z'j} W_j)}}}, \quad (33)$$

and

$$\frac{\omega_{Rj}}{\omega} = \frac{-\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}}{\frac{r'_{oj} (\omega_{x'j} V_j + \omega_{z'j} W_j)}{\frac{d_{mj}}{2} [(f_o - 0.5)D + \delta_{0j}] + r'_{oj} V_j} + \frac{r'_{ij} [\omega_{x'j}(s_{zj} - V_j) + \omega_{z'j}(s_{xj} - W_j)]}{\frac{d_{mj}}{2} [(f_i - 0.5)D + \delta_{ij}] - r'_{ij}(s_{zj} - V_j)}}}, \quad (34)$$

in which  $\omega_{ij}$ ,  $\omega_{oj}$  are the angular velocities about the bearing axis of the inner and outer rings with respect to the ball at position  $j$ , and  $r'_{ij}$ ,  $r'_{oj}$  are the inner and outer rolling radii [Harris (2001)].

Likewise, for the inner race to be stationary  $\omega_{mj} = -\omega_{ij}$ ,  $\omega = \omega_{oj} + \omega_{mj}$ ,

$$\frac{\omega_{mj}}{\omega} = \frac{1}{\frac{r'_{oj} \left\{ \frac{d_{mj}}{2} [(f_i - 0.5)D + \delta_{ij}] - r'_{ij}(s_{zj} - V_j) \right\} (\omega_{x'j} V_j + \omega_{z'j} W_j)}{1 + \frac{r'_{ij} \left\{ \frac{d_{mj}}{2} [(f_o - 0.5)D + \delta_{0j}] + r'_{oj} V_j \right\} [\omega_{x'j}(s_{zj} - V_j) + \omega_{z'j}(s_{xj} - W_j)]}}}, \quad (35)$$

and  $\omega_{Rj}/\omega$  is given by (34) with opposite sign.

Similarly, the gyroscopic moments acting on the ball at angular position  $\psi_j$  are given by [Harris (2001)]

$$M_{y'j} = J \omega^2 \left( \frac{\omega_{Rj}}{\omega} \right) \left( \frac{\omega_{mj}}{\omega} \right) \frac{\omega_{z'j}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}}, \quad (36)$$

and

$$M_{z'j} = -J\omega^2 \left(\frac{\omega_{Rj}}{\omega}\right) \left(\frac{\omega_{mj}}{\omega}\right) \frac{\omega_{y'j}}{\sqrt{\omega_{x'j}^2 + \omega_{y'j}^2 + \omega_{z'j}^2}}, \quad (37)$$

in which  $J$  is the ball's mass moment of inertia.

The friction forces due to sliding in the  $x$  and  $y$ -directions of inner and outer ball-raceway elliptical contact areas are given by [Harris (2001)]

$$F_{xij} = \frac{3\mu K_{ij} \delta_{ij}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \sin \gamma_{ij} dy_{ij} dx_{ij}, \quad (38)$$

$$F_{xoj} = \frac{3\mu K_{oj} \delta_{oj}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \sin \gamma_{oj} dy_{oj} dx_{oj}, \quad (39)$$

$$F_{yij} = \frac{3\mu K_{ij} \delta_{ij}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \cos \gamma_{ij} dy_{ij} dx_{ij}, \quad (40)$$

$$F_{yoj} = \frac{3\mu K_{oj} \delta_{oj}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \cos \gamma_{oj} dy_{oj} dx_{oj}, \quad (41)$$

in which  $\mu$  is the friction coefficient;  $a_{ij}$ ,  $b_{ij}$ ,  $a_{oj}$ , and  $b_{oj}$  are semimajor and semiminor-axes of inner and outer pressure ellipses;  $x_{ij}$ ,  $y_{ij}$ ,  $x_{oj}$ ,  $y_{oj}$  are the co-ordinates of an element of area,  $dydx$ , inside the contact ellipse, which has a resultant velocity of slip  $V$  of the race on the ball acting at the angle  $\gamma$  with respect to the  $y$ -direction, which are given by

$$\gamma_{ij} = \tan^{-1} \frac{y_{ij} \omega_{sij} - V_{xij}}{x_{ij} \omega_{sij} + V_{yij}}, \quad \gamma_{oj} = \tan^{-1} \frac{y_{oj} \omega_{soj} - V_{xoj}}{x_{oj} \omega_{soj} + V_{yoj}}. \quad (42)$$

$V_{xij}$ ,  $V_{xoj}$ ,  $V_{yij}$ ,  $V_{yoj}$ ,  $\omega_{sij}$ , and  $\omega_{soj}$  are the relative linear and angular slip velocities of inner and outer races with respect to the ball located at position  $j$ . The terms involving these velocities for use in (42) are given by [Harris (2001)]

$$\frac{V_{xij}}{\omega_{sij}} = \frac{\left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} \right) \left[ (f_i - 0.5)D + \delta_{ij} - \frac{r'_{ij}}{d_{mj}} (s_{zj} - V_j) \right] \omega_{y'j}}{\omega_{x'j} (s_{xj} - W_j) - \omega_{z'j} \left[ s_{zj} - V_j - \frac{(f_i - 0.5)D + \delta_{ij}}{d_{mj}} r'_{ij} \right]}, \quad (43)$$

$$\frac{V_{yij}}{\omega_{sij}} = \frac{\left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} - r'_{ij} \right) \left[ \omega_{x'j} (s_{zj} - V_j) + \omega_{z'j} (s_{xj} - W_j) \right]}{\omega_{x'j} (s_{xj} - W_j) - \omega_{z'j} \left[ s_{zj} - V_j - \frac{(f_i - 0.5)D + \delta_{ij}}{d_{mj}} r'_{ij} \right]}, \quad (44)$$

$$\frac{V_{xoj}}{\omega_{soj}} = \frac{-\left(\sqrt{R_0^2 - x_{oj}^2} - \sqrt{R_0^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2}\right) \left[ (f_0 - 0.5)D + \delta_{oj} + \frac{r'_{oj}}{\frac{d_{mj}}{2}} V_j \right] \omega_{y'j}}{\omega_{x'j} W_j - \omega_{z'j} \left[ V_j + \frac{(f_0 - 0.5)D + \delta_{oj} r'_{oj}}{\frac{d_{mj}}{2}} \right]}, \quad (45)$$

$$\frac{V_{yoj}}{\omega_{soj}} = \frac{\left(\sqrt{R_0^2 - x_{oj}^2} - \sqrt{R_0^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2} - r'_{oj}\right) (\omega_{x'j} V_j + \omega_{z'j} W_j)}{\omega_{x'j} W_j - \omega_{z'j} \left[ V_j + \frac{(f_0 - 0.5)D + \delta_{oj} r'_{oj}}{\frac{d_{mj}}{2}} \right]}, \quad (46)$$

in which  $R_i$  and  $R_o$  are the curvature radii of deformed surfaces, given by

$$R_i = \frac{2f_i D}{2f_i + 1}, \quad R_o = \frac{2f_o D}{2f_o + 1}. \quad (47)$$

The total frictional moments of the friction forces about the normal at the center of the contact ellipse are [Harris (2001)]

$$M_{sij} = \frac{3\mu K_{ij} \delta_{ij}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{x_{ij}^2 + y_{ij}^2} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \cos\left(\gamma_{ij} - \tan^{-1} \frac{y_{ij}}{x_{ij}}\right) dy_{ij} dx_{ij}, \quad (48)$$

$$M_{soj} = \frac{3\mu K_{oj} \delta_{oj}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{x_{oj}^2 + y_{oj}^2} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \cos\left(\gamma_{oj} - \tan^{-1} \frac{y_{oj}}{x_{oj}}\right) dy_{oj} dx_{oj}. \quad (49)$$

The moments of the friction forces about the  $y'$ -axis are [Harris (2001)]

$$M_{yij} = \frac{3\mu K_{ij} \delta_{ij}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} \right) \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \sin\gamma_{ij} dy_{ij} dx_{ij}, \quad (50)$$

$$M_{yoj} = \frac{3\mu K_{oj} \delta_{oj}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \left( \sqrt{R_o^2 - x_{oj}^2} - \sqrt{R_o^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2} \right) \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \sin\gamma_{oj} dy_{oj} dx_{oj}. \quad (51)$$

The frictional moments about an axis through the ball center perpendicular to the line defining the contact angle, which line lies in the  $x'z'$ -plane, are [Harris (2001)]

$$M_{Rij} = \frac{3\mu K_{ij} \delta_{ij}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} \right) \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \cos\gamma_{ij} dy_{ij} dx_{ij}, \quad (52)$$

$$M_{Roj} = \frac{3\mu K_{oj} \delta_{oj}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \left( \sqrt{R_o^2 - x_{oj}^2} - \sqrt{R_o^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2} \right) \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \cos\gamma_{oj} dy_{oj} dx_{oj}. \quad (53)$$

Equations (16)-(18), (22)-(24) and (27)-(29) may be solved simultaneously for  $V_j$ ,  $W_j$ ,  $\delta_{oj}$ ,  $\delta_{ij}$ ,  $r'_{oj}$ ,  $r'_{ij}$ ,  $\omega_{x'j}$ ,  $\omega_{y'j}$ , and  $\omega_{z'j}$  at each ball angular location once values for  $\delta_a$ ,  $\delta_r$ , and  $\theta$  are assumed.

An iterative procedure is to be used to solve the equations simultaneously. Since  $K_{oj}$  and  $K_{ij}$  are functions of contact angle, equations (8)-(11) may be used to establish  $K_{oj}$  and  $K_{ij}$  values iteratively.

To find the values of  $\delta_a$ ,  $\delta_r$ , and  $\theta$ , it remains to establish the equilibrium conditions of forces and moments about the inner ring center of mass, as shown by Fig. 7, which are

$$F_a - \sum_{j=1}^Z \left[ \frac{K_{ij} \delta_{ij}^{1.5} (s_{xj} - W_j) - F_{xij} (s_{zj} - V_j)}{(f_i - 0.5)D + \delta_{ij}} \right] = 0 = \epsilon_{9Z+1}, \quad (54)$$

$$F_r - \sum_{j=1}^Z \left[ \frac{K_{ij} \delta_{ij}^{1.5} (s_{zj} - V_j) + F_{xij} (s_{xj} - W_j)}{(f_i - 0.5)D + \delta_{ij}} \right] \cos \psi_j = 0 = \epsilon_{9Z+2}, \quad (55)$$

$$M - \sum_{j=1}^Z \{ \mathcal{R}_i [K_{ij} \delta_{ij}^{1.5} \sin \beta_{ij} - F_{xij} (\cos \beta_{ij} - r_i / \mathcal{R}_i)] \cos \psi_j - (F_{yij} r_i \sin \beta_{ij} - M_{sij} \cos \beta_{ij}) \sin \psi_j \} = 0 = \epsilon_{9Z+3}, \quad (56)$$

in which  $F_a$ ,  $F_r$ , and  $M$  are external forces and moment applied to the inner ring center of mass.

Having computed values for  $V_j$ ,  $W_j$ ,  $\delta_{oj}$ ,  $\delta_{ij}$ ,  $r'_{oj}$ ,  $r'_{ij}$ ,  $\omega_{x'j}$ ,  $\omega_{y'j}$ , and  $\omega_{z'j}$  at each angular position and knowing  $F_a$ ,  $F_r$ , and  $M$  as input conditions the values of  $\delta_a$ ,  $\delta_r$ , and  $\theta$  may be computed by equations (54)-(56). After obtaining the primary unknown quantities  $\delta_a$ ,  $\delta_r$ , and  $\theta$ , it is necessary to repeat the calculation of  $V_j$ ,  $W_j$ ,  $\delta_{oj}$ ,  $\delta_{ij}$ ,  $r'_{oj}$ ,  $r'_{ij}$ ,  $\omega_{x'j}$ ,  $\omega_{y'j}$ , and  $\omega_{z'j}$ , until compatible values of primary unknown quantities  $\delta_a$ ,  $\delta_r$ , and  $\theta$  are obtained.

### Numerical procedure

Equations (16)-(18), (22)-(24), (27)-(29), and (54)-(56) may be written as

$$\epsilon_g(\delta_h) = 0, \quad g, h = 1, \dots, 9Z + 3, \quad (57)$$

in which  $\delta_1 = V_1, \dots, \delta_Z = V_Z, \delta_{Z+1} = W_1, \dots, \delta_{2Z} = W_Z, \delta_{2Z+1} = \delta_{o1}, \dots, \delta_{3Z} = \delta_{oZ}, \delta_{3Z+1} = \delta_{i1}, \dots, \delta_{4Z} = \delta_{iZ}, \delta_{4Z+1} = r'_{o1}, \dots, \delta_{5Z} = r'_{oZ}, \delta_{5Z+1} = r'_{i1}, \dots, \delta_{6Z} = r'_{iZ}, \delta_{6Z+1} = \omega_{x'1}, \dots, \delta_{7Z} = \omega_{x'Z}, \delta_{7Z+1} = \omega_{y'1}, \dots, \delta_{8Z} = \omega_{y'Z}, \delta_{8Z+1} = \omega_{z'1}, \dots, \delta_{9Z} = \omega_{z'Z}, \delta_{9Z+1} = \delta_a, \delta_{9Z+2} = \delta_r, \delta_{9Z+3} = \theta$ .

The first  $9Z$  equations from (57) must be solved simultaneously for  $\delta_1, \dots, \delta_{9Z}$  once values for  $\delta_{9Z+1}, \dots, \delta_{9Z+3}$  are assumed. If  $\delta_h^0, h = 1, \dots, 9Z$ , is a  $9Z$ -dimensional vector with the initial estimates of the variables  $\delta_1, \dots, \delta_{9Z}$ , improved values are given by

$$\delta'_h = \delta_h^0 - [a_{gh}]^{-1} \{ \epsilon_g \}, \quad (58)$$

in which  $\{ \epsilon_g \}$ ,  $g = 1, \dots, 9Z$ , is the  $9Z$ -dimensional vector with the first  $9Z$  errors functions from (1). The elements of the square  $9Z \times 9Z$ -matrix  $[a_{gh}]$  are

$$a_{jh} = -2(s_{zj} - \delta_j) \frac{\partial \delta_j}{\partial \delta_h} - 2(s_{xj} - \delta_{Z+j}) \frac{\partial \delta_{Z+j}}{\partial \delta_h} - 2[(f_i - 0.5)D + \delta_{3Z+j}] \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (59)$$

$$a_{(j+Z)h} = 2\delta_j \frac{\partial \delta_j}{\partial \delta_h} + 2\delta_{Z+j} \frac{\partial \delta_{Z+j}}{\partial \delta_h} - 2[(f_o - 0.5)D + \delta_{2Z+j}] \frac{\partial \delta_{2Z+j}}{\partial \delta_h}, \quad (60)$$

$$a_{(j+2Z)h} = 2\delta_{6Z+j} \frac{\partial \delta_{6Z+j}}{\partial \delta_h} + 2\delta_{7Z+j} \frac{\partial \delta_{7Z+j}}{\partial \delta_h} + 2\delta_{8Z+j} \frac{\partial \delta_{8Z+j}}{\partial \delta_h}, \quad (61)$$

$$a_{(j+3Z)h} = \frac{\left\{ \left( \frac{\partial \delta_j}{\partial \delta_h} - K_{oj} \delta_{2Z+j}^{1.5} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} - 1.5 K_{oj} \delta_{2Z+j}^{0.5} \delta_{Z+j} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \delta_j \frac{\partial F_{xoj}}{\partial \delta_h} \right) \times \left[ \frac{\partial \delta_{2Z+j}}{\partial \delta_h} - (F_{xo} - 0.5)D + \delta_{2Z+j} \right] - (F_{xoj} \delta_j - K_{oj} \delta_{2Z+j}^{1.5} \delta_{Z+j}) \frac{\partial \delta_{2Z+j}}{\partial \delta_h} \right\}}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} + \frac{\left\{ \left[ F_{xij} \frac{\partial \delta_j}{\partial \delta_h} - K_{ij} \delta_{3Z+j}^{1.5} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} + 1.5 K_{ij} \delta_{3Z+j}^{0.5} (s_{xj} - \delta_{Z+j}) \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - (s_{zj} - \delta_j) \frac{\partial F_{xij}}{\partial \delta_h} \right] \times \left[ \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - [(f_i - 0.5)D + \delta_{3Z+j}] - [K_{ij} \delta_{3Z+j}^{1.5} (s_{xj} - \delta_{Z+j}) - F_{xij} (s_{zj} - \delta_j)] \frac{\partial \delta_{3Z+j}}{\partial \delta_h} \right] \right\}}{[(f_i - 0.5)D + \delta_{3Z+j}]^2}, \quad (62)$$

$$a_{(j+4Z)h} = \frac{\left\{ \left( K_{oj} \delta_{2Z+j}^{1.5} \frac{\partial \delta_j}{\partial \delta_h} + F_{xoj} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + 1.5 K_{oj} \delta_{2Z+j}^{0.5} \delta_j \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \delta_{Z+j} \frac{\partial F_{xoj}}{\partial \delta_h} \right) \times \right.}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} + \frac{\left\{ \left[ K_{ij} \delta_{3Z+j}^{1.5} \frac{\partial \delta_j}{\partial \delta_h} + F_{xij} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - 1.5 K_{ij} \delta_{3Z+j}^{0.5} (s_{zj} - \delta_j) \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - (s_{xj} - \delta_{Z+j}) \frac{\partial F_{xij}}{\partial \delta_h} \right] \times \right.}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} - \frac{\partial F_{z'j}}{\partial \delta_h}, \quad (63)$$

$$a_{(j+5Z)h} = \frac{\partial F_{y'oj}}{\partial \delta_h} + \frac{\partial F_{yij}}{\partial \delta_h}, \quad (64)$$

$$a_{(j+6Z)h} = \frac{\left\{ \left( M_{Roj} \frac{\partial \delta_j}{\partial \delta_h} + M_{soj} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \delta_j \frac{\partial M_{Roj}}{\partial \delta_h} + \delta_{Z+j} \frac{\partial M_{soj}}{\partial \delta_h} \right) \times \right.}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} - \frac{\left\{ \left[ M_{Rij} \frac{\partial \delta_j}{\partial \delta_h} + M_{sij} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - (s_{xj} - \delta_{Z+j}) \frac{\partial M_{sij}}{\partial \delta_h} - (s_{zj} - \delta_j) \frac{\partial M_{rij}}{\partial \delta_h} \right] \times \right.}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} + \frac{\left\{ \left[ (f_o - 0.5)D + \delta_{2Z+j} \right] - (M_{Roj} \delta_j + M_{soj} \delta_{2Z+j}) \frac{\partial \delta_{2Z+j}}{\partial \delta_h} \right.}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} + \frac{\left\{ \left[ (f_i - 0.5)D + \delta_{3Z+j} \right] + [M_{sij}(s_{xj} - \delta_{Z+j}) + M_{rij}(s_{zj} - \delta_j)] \frac{\partial \delta_{3Z+j}}{\partial \delta_h} \right.}{[(f_i - 0.5)D + \delta_{3Z+j}]^2}, \quad (65)$$

$$a_{(j+7Z)h} = \frac{\left\{ \left( M_{soj} \frac{\partial \delta_j}{\partial \delta_h} - M_{Roj} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \delta_j \frac{\partial M_{soj}}{\partial \delta_h} - \delta_{Z+j} \frac{\partial M_{Roj}}{\partial \delta_h} \right) \times \right.}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} - \frac{\left\{ \left[ M_{sij} \frac{\partial \delta_j}{\partial \delta_h} - M_{rij} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - (s_{zj} - \delta_j) \frac{\partial M_{sij}}{\partial \delta_h} + (s_{xj} - \delta_{Z+j}) \frac{\partial M_{rij}}{\partial \delta_h} \right] \times \right.}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} + \frac{\partial M_{z'j}}{\partial \delta_h}, \quad (66)$$

$$a_{(j+8Z)h} = \frac{\partial M_{y'j}}{\partial \delta_h} - \frac{\partial M_{yij}}{\partial \delta_h} - \frac{\partial M_{y'oj}}{\partial \delta_h}. \quad (67)$$

The forces  $F_{z'j}$ ,  $F_{xij}$  and  $F_{xoj}$  to be used in (62) and (63) are given by (32), (38) and (39) and their differentiation with respect to  $\delta_h$ ,  $h = 1, \dots, 9Z$ , yields

$$\frac{\partial F_{z'j}}{\partial \delta_h} = m\omega^2 \left[ \left( \frac{\omega_{mj}}{\omega} \right)^2 \frac{\partial \delta_j}{\partial \delta_h} + d_{mj} \left( \frac{\omega_{mj}}{\omega} \right) \frac{\partial \left( \frac{\omega_{mj}}{\omega} \right)}{\partial \delta_h} \right], \quad (68)$$

$$\frac{\partial F_{xij}}{\partial \delta_h} = \frac{3\mu K_{ij} \delta_{3Z+j}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \cos \gamma_{ij} \frac{\partial \gamma_{ij}}{\partial \delta_h} dy_{ij} dx_{ij} + \frac{3}{2} \frac{F_{xij}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (69)$$

$$\frac{\partial F_{xoj}}{\partial \delta_h} = \frac{3\mu K_{oj} \delta_{2Z+j}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \cos \gamma_{oj} \frac{\partial \gamma_{oj}}{\partial \delta_h} dy_{oj} dx_{oj} + \frac{3}{2} \frac{F_{xoj}}{\delta_{2Z+j}} \frac{\partial \delta_{2Z+j}}{\partial \delta_h}. \quad (70)$$

The forces  $F_{yij}$  and  $F_{y'oj}$  to be used in (64) are given by (40) and (41) and their differentiation with respect to  $\delta_h$ ,  $h = 1, \dots, 9Z$ , yields

$$\frac{\partial F_{yij}}{\partial \delta_h} = -\frac{3\mu K_{ij} \delta_{3Z+j}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \sin \gamma_{ij} \frac{\partial \gamma_{ij}}{\partial \delta_h} dy_{ij} dx_{ij} + \frac{3}{2} \frac{F_{yij}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (71)$$

$$\frac{\partial F_{y'oj}}{\partial \delta_h} = -\frac{3\mu K_{oj} \delta_{2Z+j}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \sin \gamma_{oj} \frac{\partial \gamma_{oj}}{\partial \delta_h} dy_{oj} dx_{oj} + \frac{3}{2} \frac{F_{y'oj}}{\delta_{2Z+j}} \frac{\partial \delta_{2Z+j}}{\partial \delta_h}. \quad (72)$$

The moments  $M_{sij}$ ,  $M_{soj}$ ,  $M_{rij}$ ,  $M_{roj}$  and  $M_{z'j}$  to be used in (65) and (66) are given by (48), (49), (52), (53) and (37) and their differentiation with respect to  $\delta_h$ ,  $h = 1, \dots, 9Z$ , yields

$$\frac{\partial M_{sij}}{\partial \delta_h} = \frac{-3\mu K_{ij} \delta_{3Z+j}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \sqrt{x_{ij}^2 + y_{ij}^2} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \sin \left( \gamma_{ij} - \tan^{-1} \frac{y_{ij}}{x_{ij}} \right) \frac{\partial \gamma_{ij}}{\partial \delta_h} dy_{ij} dx_{ij} + \frac{3}{2} \frac{M_{sij}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (73)$$

$$\frac{\partial M_{soj}}{\partial \delta_h} = \frac{-3\mu K_{oj} \delta_{2Z+j}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \sqrt{x_{oj}^2 + y_{oj}^2} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \sin\left(\gamma_{oj} - \tan^{-1} \frac{y_{oj}}{x_{oj}}\right) \frac{\partial \gamma_{oj}}{\partial \delta_h} dy_{oj} dx_{oj} + \frac{3}{2} \frac{M_{soj}}{\delta_{2Z+j}} \frac{\partial \delta_{2Z+j}}{\partial \delta_h}, \quad (74)$$

$$\frac{\partial M_{Rij}}{\partial \delta_h} = \frac{-3\mu K_{ij} \delta_{3Z+j}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} \right) \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \sin \gamma_{ij} \frac{\partial \gamma_{ij}}{\partial \delta_h} dy_{ij} dx_{ij} + \frac{3}{2} \frac{M_{Rij}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (75)$$

$$\frac{\partial M_{Roj}}{\partial \delta_h} = \frac{-3\mu K_{oj} \delta_{3Z+j}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \left( \sqrt{R_o^2 - x_{oj}^2} - \sqrt{R_o^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2} \right) \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \sin \gamma_{oj} \frac{\partial \gamma_{oj}}{\partial \delta_h} dy_{oj} dx_{oj} + \frac{3}{2} \frac{M_{Roj}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (76)$$

$$\frac{\partial M_{zj}}{\partial \delta_h} = -J\omega^2 \left\{ \frac{\delta_{7Z+j}}{\sqrt{\delta_{6Z+j}^2 + \delta_{7Z+j}^2 + \delta_{8Z+j}^2}} \left[ \left( \frac{\omega_{Rj}}{\omega} \right) \frac{\partial \left( \frac{\omega_{mj}}{\omega} \right)}{\partial \delta_h} + \frac{\partial \left( \frac{\omega_{Rj}}{\omega} \right)}{\partial \delta_h} \left( \frac{\omega_{mj}}{\omega} \right) \right] - \frac{\delta_{6Z+j} \delta_{7Z+j} \frac{\partial \delta_{6Z+j}}{\partial \delta_h} - (\delta_{6Z+j}^2 + \delta_{8Z+j}^2) \frac{\partial \delta_{7Z+j}}{\partial \delta_h} + \delta_{7Z+j} \delta_{8Z+j} \frac{\partial \delta_{8Z+j}}{\partial \delta_h}}{(\delta_{6Z+j}^2 + \delta_{7Z+j}^2 + \delta_{8Z+j}^2)^{\frac{3}{2}}} \left( \frac{\omega_{Rj}}{\omega} \right) \left( \frac{\omega_{mj}}{\omega} \right) \right\} \quad (77)$$

The moments  $M_{yij}$ ,  $M_{yoj}$  and  $M_{y'j}$  to be used in (67) are given by (50), (51) and (36) and their differentiation with respect to  $\delta_h$ ,  $h = 1, \dots, 9Z$ , yields

$$\frac{\partial M_{yij}}{\partial \delta_h} = \frac{3\mu K_{ij} \delta_{3Z+j}^{1.5}}{2\pi a_{ij} b_{ij}} \int_{-a_{ij}}^{a_{ij}} \int_{-b_{ij}}^{b_{ij}} \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2}} \left( \sqrt{R_i^2 - x_{ij}^2} - \sqrt{R_i^2 - a_{ij}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{ij}^2} \right) \sqrt{1 - \frac{x_{ij}^2}{a_{ij}^2} - \frac{y_{ij}^2}{b_{ij}^2}} \cos \gamma_{ij} \frac{\partial \gamma_{ij}}{\partial \delta_h} dy_{ij} dx_{ij} + \frac{3}{2} \frac{M_{yij}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (78)$$

$$\frac{\partial M_{yoj}}{\partial \delta_h} = \frac{3\mu K_{oj} \delta_{3Z+j}^{1.5}}{2\pi a_{oj} b_{oj}} \int_{-a_{oj}}^{a_{oj}} \int_{-b_{oj}}^{b_{oj}} \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2}} \left( \sqrt{R_o^2 - x_{oj}^2} - \sqrt{R_o^2 - a_{oj}^2} + \sqrt{\left(\frac{D}{2}\right)^2 - a_{oj}^2} \right) \sqrt{1 - \frac{x_{oj}^2}{a_{oj}^2} - \frac{y_{oj}^2}{b_{oj}^2}} \cos \gamma_{oj} \frac{\partial \gamma_{oj}}{\partial \delta_h} dy_{oj} dx_{oj} + \frac{3}{2} \frac{M_{yoj}}{\delta_{3Z+j}} \frac{\partial \delta_{3Z+j}}{\partial \delta_h}, \quad (79)$$

$$\frac{\partial M_{y'j}}{\partial \delta_h} = J\omega^2 \left\{ \frac{\delta_{8Z+j}}{\sqrt{\delta_{6Z+j}^2 + \delta_{7Z+j}^2 + \delta_{8Z+j}^2}} \left[ \left( \frac{\omega_{Rj}}{\omega} \right) \frac{\partial \left( \frac{\omega_{mj}}{\omega} \right)}{\partial \delta_h} + \frac{\partial \left( \frac{\omega_{Rj}}{\omega} \right)}{\partial \delta_h} \left( \frac{\omega_{mj}}{\omega} \right) \right] - \frac{\delta_{6Z+j} \delta_{8Z+j} \frac{\partial \delta_{6Z+j}}{\partial \delta_h} + \delta_{7Z+j} \delta_{8Z+j} \frac{\partial \delta_{7Z+j}}{\partial \delta_h} - (\delta_{6Z+j}^2 + \delta_{7Z+j}^2) \frac{\partial \delta_{8Z+j}}{\partial \delta_h}}{(\delta_{6Z+j}^2 + \delta_{7Z+j}^2 + \delta_{8Z+j}^2)^{\frac{3}{2}}} \left( \frac{\omega_{Rj}}{\omega} \right) \left( \frac{\omega_{mj}}{\omega} \right) \right\} \quad (80)$$

As in (38)-(41) and (48)-(53), also in (69)-(76), (78) e (79)  $\gamma_{ij}$ ,  $\gamma_{oj}$  are given by (42). The derivatives of  $\gamma_{ij}$ ,  $\gamma_{oj}$  with respect  $\delta_h$ ,  $h = 1, \dots, 9Z$ , to be used in (69)-(76), (78) e (79) are given by

$$\frac{\partial \gamma_{ij}}{\partial \delta_h} = \frac{-\left(x_{ij} + \frac{V_{yij}}{\omega_{sij}}\right) \frac{\partial \left(\frac{V_{xij}}{\omega_{sij}}\right)}{\partial \delta_h} - \left(y_{ij} - \frac{V_{xij}}{\omega_{sij}}\right) \frac{\partial \left(\frac{V_{yij}}{\omega_{sij}}\right)}{\partial \delta_h}}{\left(y_{ij} - \frac{V_{xij}}{\omega_{sij}}\right)^2 + \left(x_{ij} + \frac{V_{yij}}{\omega_{sij}}\right)^2}, \quad \frac{\partial \gamma_{oj}}{\partial \delta_h} = \frac{-\left(x_{oj} + \frac{V_{yoj}}{\omega_{soj}}\right) \frac{\partial \left(\frac{V_{xoj}}{\omega_{soj}}\right)}{\partial \delta_h} - \left(y_{oj} - \frac{V_{xoj}}{\omega_{soj}}\right) \frac{\partial \left(\frac{V_{yoj}}{\omega_{soj}}\right)}{\partial \delta_h}}{\left(y_{oj} - \frac{V_{xoj}}{\omega_{soj}}\right)^2 + \left(x_{oj} + \frac{V_{yoj}}{\omega_{soj}}\right)^2}, \quad (81)$$

in which  $V_{xij}/\omega_{sij}$ ,  $V_{yij}/\omega_{sij}$ ,  $V_{xoj}/\omega_{soj}$ ,  $V_{yoj}/\omega_{soj}$  are given by (43)-(46). The derivatives of  $V_{xij}/\omega_{sij}$ ,  $V_{yij}/\omega_{sij}$ ,  $V_{xoj}/\omega_{soj}$  and  $V_{yoj}/\omega_{soj}$  with respect  $\delta_h$ ,  $h = 1, \dots, 9Z$ , to be used in (81) are given by







$$\begin{aligned}
 \alpha_{(9Z+3)h} = \frac{\partial M}{\partial \delta_h} - \sum_{j=1}^Z \left\{ \mathcal{R}_i \left( K_{ij} \delta_{3Z+j}^{1.5} c \beta_{ij} + F_{xij} s \beta_{ij} \right) \left[ \frac{(s_{zj} - \delta_j) \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \frac{\partial \delta_{Z+j}}{\partial \delta_h} \right) - (s_{xj} - \delta_{Z+j}) \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \frac{\partial \delta_j}{\partial \delta_h} \right)}{(s_{zj} - \delta_j)^2 + (s_{xj} - \delta_{Z+j})^2} \right. \right. \\
 \left. \left. - c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + 1.5 K_{ij} \delta_{3Z+j}^{0.5} s \beta_{ij} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - (c \beta_{ij} - r_i / \mathcal{R}_i) \frac{\partial F_{xij}}{\partial \delta_h} \right] c \psi_j - \left( F_{yij} r_i c \beta_{ij} + M_{sij} s \beta_{ij} \right) \left[ \frac{(s_{zj} - \delta_j) \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \frac{\partial \delta_{Z+j}}{\partial \delta_h} \right)}{(s_{zj} - \delta_j)^2 + (s_{xj} - \delta_{Z+j})^2} \right. \right. \\
 \left. \left. - \frac{(s_{xj} - \delta_{Z+j}) \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \frac{\partial \delta_j}{\partial \delta_h} \right)}{(s_{zj} - \delta_j)^2 + (s_{xj} - \delta_{Z+j})^2} - c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right] + r_i s \beta_{ij} \frac{\partial F_{yij}}{\partial \delta_h} - c \beta_{ij} \frac{\partial M_{sij}}{\partial \delta_h} \right\} s \psi_j \} \quad (91)
 \end{aligned}$$

Differentiating (16)-(18), (22)-(24) and (27)-(29) with respect  $\delta_h$ ,  $h = 9Z+1, \dots, 9Z+3$ ,  $27Z$  simultaneous linear equations in  $\partial \delta_{kZ+j} / \partial \delta_h$ ,  $k = 0, \dots, 8$ ;  $j = 1, \dots, Z$ , results, which are

$$(s_{zj} - \delta_j) \frac{\partial \delta_j}{\partial \delta_h} + (s_{xj} - \delta_{Z+j}) \frac{\partial \delta_{Z+j}}{\partial \delta_h} + [(f_i - 0.5)D + \delta_{3Z+j}] \frac{\partial \delta_{3Z+j}}{\partial \delta_h} = (s_{xj} - \delta_{Z+j}) \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right) + (s_{zj} - \delta_j) \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right), \quad (92)$$

$$\delta_j \frac{\partial \delta_j}{\partial \delta_h} + \delta_{Z+j} \frac{\partial \delta_{Z+j}}{\partial \delta_h} - [(f_o - 0.5)D + \delta_{2Z+j}] \frac{\partial \delta_{2Z+j}}{\partial \delta_h} = 0, \quad (93)$$

$$\delta_{6Z+j} \frac{\partial \delta_{6Z+j}}{\partial \delta_h} + \delta_{7Z+j} \frac{\partial \delta_{7Z+j}}{\partial \delta_h} + \delta_{8Z+j} \frac{\partial \delta_{8Z+j}}{\partial \delta_h} = 0, \quad (94)$$

$$\begin{aligned}
 \left[ \frac{F_{xoj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{F_{xij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_j}{\partial \delta_h} + \frac{\delta_j}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial F_{xoj}}{\partial \delta_h} - \frac{(s_{xj} - \delta_{Z+j})}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial F_{xij}}{\partial \delta_h} - \left[ \frac{K_{oj} \delta_{2Z+j}^{1.5}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{K_{ij} \delta_{3Z+j}^{1.5}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_{Z+j}}{\partial \delta_h} - \\
 \left\{ \frac{1.5 K_{oj} \delta_{2Z+j}^{0.5} \delta_{Z+j} [(f_o - 0.5)D + \delta_{2Z+j}] + F_{xoj} \delta_j - K_{oj} \delta_{2Z+j}^{1.5} \delta_{Z+j}}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} \right\} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \left\{ \frac{1.5 K_{ij} \delta_{3Z+j}^{0.5} (s_{xj} - \delta_{Z+j}) [(f_i - 0.5)D + \delta_{3Z+j}] + F_{xij} (s_{zj} - \delta_j) - K_{ij} \delta_{3Z+j}^{1.5} (s_{xj} - \delta_{Z+j})}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} \right\} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} = \\
 \frac{-K_{ij} \delta_{3Z+j}^{1.5} \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right) + F_{xij} \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right)}{(f_i - 0.5)D + \delta_{3Z+j}}, \quad (95)
 \end{aligned}$$

$$\begin{aligned}
 \left[ \frac{K_{oj} \delta_{2Z+j}^{1.5}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{K_{ij} \delta_{3Z+j}^{1.5}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_j}{\partial \delta_h} + \frac{\delta_{Z+j}}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial F_{xoj}}{\partial \delta_h} - \frac{(s_{xj} - \delta_{Z+j})}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial F_{xij}}{\partial \delta_h} + \left[ \frac{F_{xoj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{F_{xij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_{Z+j}}{\partial \delta_h} + \\
 \left\{ \frac{1.5 K_{oj} \delta_{2Z+j}^{0.5} \delta_j [(f_o - 0.5)D + \delta_{2Z+j}] - K_{oj} \delta_{2Z+j}^{1.5} \delta_j - F_{xoj} \delta_{Z+j}}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} \right\} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} - \\
 \left\{ \frac{1.5 K_{ij} \delta_{3Z+j}^{0.5} (s_{zj} - \delta_j) [(f_i - 0.5)D + \delta_{3Z+j}] - K_{ij} \delta_{3Z+j}^{1.5} (s_{zj} - \delta_j) - F_{xij} (s_{xj} - \delta_{Z+j})}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} \right\} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} - \frac{\partial F_{xij}}{\partial \delta_h} = \\
 \frac{K_{ij} \delta_{3Z+j}^{1.5} \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right) + F_{xij} \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right)}{(f_i - 0.5)D + \delta_{3Z+j}}, \quad (96)
 \end{aligned}$$

$$\frac{\partial F_{yoj}}{\partial \delta_h} + \frac{\partial F_{yij}}{\partial \delta_h} = 0, \quad (97)$$

$$\begin{aligned}
 \left[ \frac{M_{roj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{M_{rij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_j}{\partial \delta_h} + \frac{\delta_j}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial M_{roj}}{\partial \delta_h} + \frac{\delta_{Z+j}}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial M_{soj}}{\partial \delta_h} + \\
 \left[ \frac{M_{soj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{M_{sij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_{Z+j}}{\partial \delta_h} - \frac{(s_{xj} - \delta_{Z+j})}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial M_{sij}}{\partial \delta_h} - \frac{(s_{zj} - \delta_j)}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial M_{rij}}{\partial \delta_h} - \frac{M_{roj} \delta_j + M_{soj} \delta_{Z+j}}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \\
 \frac{M_{sij} (s_{xj} - \delta_{Z+j}) + M_{rij} (s_{zj} - \delta_j)}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} = \frac{M_{sij} \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right) + M_{rij} \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right)}{(f_i - 0.5)D + \delta_{3Z+j}}, \quad (98)
 \end{aligned}$$

$$\begin{aligned}
 \left[ \frac{M_{soj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{M_{sij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_j}{\partial \delta_h} + \frac{\delta_j}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial M_{soj}}{\partial \delta_h} - \frac{\delta_{Z+j}}{(f_o - 0.5)D + \delta_{2Z+j}} \frac{\partial M_{roj}}{\partial \delta_h} - \\
 \left[ \frac{M_{roj}}{(f_o - 0.5)D + \delta_{2Z+j}} + \frac{M_{rij}}{(f_i - 0.5)D + \delta_{3Z+j}} \right] \frac{\partial \delta_{Z+j}}{\partial \delta_h} + \frac{(s_{xj} - \delta_{Z+j})}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial M_{rij}}{\partial \delta_h} - \frac{(s_{zj} - \delta_j)}{(f_i - 0.5)D + \delta_{3Z+j}} \frac{\partial M_{sij}}{\partial \delta_h} - \frac{M_{soj} \delta_j - M_{roj} \delta_{Z+j}}{[(f_o - 0.5)D + \delta_{2Z+j}]^2} \frac{\partial \delta_{2Z+j}}{\partial \delta_h} + \\
 \frac{M_{sij} (s_{zj} - \delta_j) - M_{rij} (s_{xj} - \delta_{Z+j})}{[(f_i - 0.5)D + \delta_{3Z+j}]^2} \frac{\partial \delta_{3Z+j}}{\partial \delta_h} + \frac{\partial M_{zrj}}{\partial \delta_h} = \frac{M_{sij} \left( c \psi_j \frac{\partial \delta_{9Z+3}}{\partial \delta_h} - \mathcal{R}_i |c \psi_j| s \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right) - M_{rij} \left( \frac{\partial \delta_{9Z+3}}{\partial \delta_h} + \mathcal{R}_i c \psi_j c \delta_{9Z+3} \frac{\partial \delta_{9Z+3}}{\partial \delta_h} \right)}{(f_i - 0.5)D + \delta_{3Z+j}}, \quad (99)
 \end{aligned}$$

$$\frac{\partial M_{y'j}}{\partial \delta_h} - \frac{\partial M_{yij}}{\partial \delta_h} - \frac{\partial M_{yoj}}{\partial \delta_h} = 0. \quad (100)$$

The derivatives of  $F_{z'j}$ ,  $F_{xij}$ ,  $F_{xoj}$ ,  $F_{yij}$ ,  $F_{yoj}$ ,  $M_{sij}$ ,  $M_{soj}$ ,  $M_{rij}$ ,  $M_{roj}$ ,  $M_{z'j}$ ,  $M_{yij}$ ,  $M_{yoj}$  and  $M_{y'j}$  with respect  $\delta_h$ ,  $h = 9Z+1, \dots, 9Z+3$ , to be used in (89)-(100) are given by (68)-(80). In

(69)-(76), (78) and (79)  $\gamma_{ij}$ ,  $\gamma_{oj}$  are given by (42). The derivatives of  $\gamma_{ij}$ ,  $\gamma_{oj}$  with respect  $\delta_h$ ,  $h = 9Z+1, \dots, 9Z+3$ , to be used in (69)-(76), (78) and (79) are given by (81), with  $V_{xij}/\omega_{sij}$ ,  $V_{yij}/\omega_{sij}$ ,  $V_{xoj}/\omega_{soj}$  and  $V_{yoj}/\omega_{soj}$  given by (43)-(46). The derivatives of  $V_{xij}/\omega_{sij}$ ,  $V_{yij}/\omega_{sij}$ ,  $V_{xoj}/\omega_{soj}$  and  $V_{yoj}/\omega_{soj}$  with respect  $\delta_h$ ,  $h = 9Z+1, \dots, 9Z+3$ , to be used in (81) are given by (82)-(85).

For outer race to be stationary  $\omega_{mj}/\omega$  and  $\omega_{Rj}/\omega$  are given by (33) and (34), and for inner race to be stationary are given by (35) and (34), the last with opposite sign. The derivatives of (33) and (34) with respect  $\delta_h$ ,  $h = 9Z+1, \dots, 9Z+3$ , to be used in (68), (77) and (80) are given by (30) and (31); and for (35)-(34), the last with opposite sign, are given by (32) and (31), the last with opposite sign.

The linear system's solutions of the equations (92)-(100)  $-\partial\delta_{kZ+j}/\partial\delta_h$ ,  $k = 0, \dots, 8; j = 1, \dots, Z -$  are to be used in (89)-(91) for the new estimates of  $\delta_{9Z+1}$ ,  $\delta_{9Z+2}$  and  $\delta_{9Z+3}$ .

### Acknowledgements

To FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) for financial support.

### References

- Harris, T. (2001) *Rolling Bearing Analysis*, 4<sup>th</sup> ed, John Wiley & Sons, New York.
- Jones, A. (1946) *Analysis of Stresses and Deflections*, New Departure Engineering Data, Bristol, CT.
- Jones, A. B. (1959) Ball motion and sliding friction in ball bearings, *ASME J. Basic Eng.* **3**, 1-12.
- Jones, A. B. (1960) A General Theory for Elastically Constrained Ball and Radial Roller Bearings Under Arbitrary Load and Speed Conditions, *J. Fluids Eng.* **82(2)**, 309-320.
- Murer, S., Bogard, F., Rasolofondraibe, L., Pottier, B. and Marconnet, P. (2015) Determination of loads transmitted by rolling elements in a roller bearing using capacitive probes: Finite element validation, *Mech. Syst. Signal Process.* **54-5** 306-313.
- Rajasekhar, M., Srinivas, J. and Divekar, A. (2013) Dynamic Analysis of Aero-engine Rotors Supported on Ball bearing system, *Proc 1st Int. and 16th Nat. Conf. on Machines and Mechanisms (iNaCoMM2013)* 941-946.
- Rasolofondraibe, L., Pottier, B., Marconnet, P. and Chimentin, X. (2012) Capacitive sensor device for measuring loads on bearings, *IEEE Sens. J.* **12** 2186-2191.
- Rasolofondraibe, L., Pottier, B., Marconnet, P. and Perrin, E. (2013) Numerical Model of the Capacitive Probe's Capacitance for Measuring the External Loads Transmitted by the Bearing's Rolling Elements of Rotating Machines, *IEEE Sens. J.* **13** 3067-3072.
- Ricci, M. C. (2009) Ball bearings subjected to a variable eccentric thrust load, *Proc. 8<sup>th</sup> Brazilian Conf. on Dynamics, Control and Appl.*, Bauru, Brazil, 18-22 May.
- Ricci, M. C. (2009a) Internal loading distribution in statically loaded ball bearings, *1<sup>st</sup> Int. Conf. on Comp. Contact Mech., Program and Abstracts*, Lecce, Italy, 16-18 Sept., p. 21-22.
- Ricci, M. C. (2009b) Internal loading distribution in statically loaded ball bearings subjected to a combined radial and thrust load, *Proc. 6<sup>th</sup> Int. Congress of Croatian Soc. Mech.*, Dubrovnik, Croatia, Sept. 30 to Oct. 2.
- Ricci, M. C. (2009c) *Proc. 60<sup>th</sup> Int. Astron. Congress*, Daejeon, South Korea, October, 12-16.
- Ricci, M. C. (2009d) Internal loading distribution in statically loaded ball bearings subjected to an eccentric thrust load, *Math. Probl. Eng.*
- Ricci, M. C. (2010) Internal loading distribution in statically loaded ball bearings subjected to a combined radial, thrust and moment load, including the effects of temperature and fit, *Proc. 11<sup>th</sup> Pan-American Congress of Appl. Mech.*, January, 04-10, Foz do Iguaçu, Brazil.
- Rumbarger, J. (1962) Thrust Bearings with Eccentric Loads, *Mach. Des.*
- Starvin, M. S., Christopher, A. S. and Manisekar, K. (2011) Finite Element Analysis of Large Diameter Bearings Subjected to Thrust Load, *CiiT Int. J. Automation and Autonomous Syst.* **3** 389-394.
- Muszynska, A. (1986) Whirl and whip—rotor/bearing stability problems, *J. sound vib.* **110** 443-462.
- Seong, S., Kim, W., Bae, D. and Lee, S. (2014) *Korean Soc. Mech. Eng. Fall Annual Meet.* 776-779.
- Sjoväll, H. (1933) The Load Distribution within Ball and Roller Bearings under Given External Radial and Axial Load, *Teknisk Tidskrift, Mek* **h.9**.
- Stribeck, R. (1907) Ball Bearings for Various Loads, *Trans. ASME* **29**, 420-463.

- Tomović, R. (2012) Calculation of the boundary values of rolling bearing deflection in relation to the number of active rolling elements, *Mechanism and Mach. Theory* **47** 74-88.
- Tomović, R. (2012a) Calculation of the necessary level of external radial load for inner ring support on q rolling elements in a radial bearing with internal radial clearance, *Int. J. Mech. Sci.* **60** 23-33.
- Tomović, R. (2013) Investigation of the Effect of Rolling Bearing Construction on Internal Load Distribution and the Number of Active Rolling Elements, *Adv. Mat. Res.* **633** 103-116.
- Tudose, C., Rusu, F. and Tudose, L. (2013) Influence of rotation angle on bearing rolling bodies load distribution. Part 1: Mathematical Model, *Appl. Math. Mech.* **56** 469-474.
- Zhenguo, S., Huimin, D., Fanhai, M. and Hua, W. (2011) Finite element analysis method of slewing bearing with plastic deformation, *Trans. CSAE* **27** 52-56.

# Damage prediction of carbon-epoxy composites under shear loads using the finite element method

†K.M. Karumbaiah<sup>1\*</sup>, R. Das<sup>1</sup>, and S. Campbell<sup>2</sup>

<sup>1</sup> Centre for Advanced Composite Materials, Department of Mechanical Engineering, University of Auckland, Auckland, New Zealand.

<sup>2</sup> Defence Technology Agency, Auckland, New Zealand

\*Presenting author: Kmal847@aucklanduni.ac.nz

†Corresponding author: Kmal847@aucklanduni.ac.nz

## Abstract

Carbon fiber reinforced polymer composite (CFRP) laminates are used in many applications of the aerospace industry, particularly, in aircraft structural components due to their good stiffness to weight ratios compared to traditionally used aluminium. The ability of accurate prediction of the structural response of composites under complex loadings is crucially important for high-end designs and optimisation of composite structures. The accuracy and predictive capabilities of finite element models in the failure analysis of the carbon-epoxy (IM7/977-3) composite laminates have been evaluated. The cross ply composite laminate with the layup configuration  $[0,90]_{4S}$  was numerically modelled under shear loads, and the commercial finite element program ABAQUS was utilised. The failure due to shear loads were analysed using the Hashin's failure criterion. The numerical results were validated by comparing them against the carefully conducted experimental test data. The difference between the experimental and numerically predicted values of the stress and strain were compared to evaluate the accuracy of the finite element models.

**Keywords:** Composite materials, Carbon-epoxy, Failure criteria, Progressive damage.

## Introduction

Composite laminates developed of fiber-reinforced plies are being progressively used in locomotive, aerospace, marine and defence industries as they acquire higher strength than those of metallic structures, and they can be engineered to attain optimum material properties in anticipated directions. A thought-provoking issue in designing composites is delineating numerous failure modes, such as matrix cracking, fiber breakage, fiber kinking, delamination, and fiber/matrix debonding, due to its anisotropic nature [Reifsnider 1980, Rotem 1998, Chen, Morozov et al. 2014]. The difficulty of the problem is supported by the fact that there are only few successfully theories of failure, such as Hashin's failure criterion [Hashin 1983], Tsai-Hill, Tsai-Wu, and Puck's failure criterion that are incorporated by the researchers to predict the damage and failure of the composite materials [Hinton, Kaddour et al. 2004, Kaddour, Hinton et al. 2004, Kaddour, Hinton et al. 2004]. Overall, the load carrying capability of a structure does not become extinct as soon as either damage or failure arises at a material point. The structure can support supplementary loads before it ultimately fails. Thus, it is significant to enumerate damage caused by the commencement of a failure mode and study its progression and the ultimate failure of a structure with an increase in the applied load.

Damage and failure in composite structures can be analysed by either using a micro-mechanics method that deliberates damage and failure at the constituent level or a continuum damage mechanics (CDM) method in which material properties of the composite have been standardised and damage and failure is studied at the ply/lamina level [Sun, Tan et al. 2011]. Damage studied at the constituent level is not only computationally expensive for a real time problem but also involves extensive investigational characterization to determine the values of material properties in the

damage modelling and analysis. Considering a specific case, a statistical approach can be adapted to study the micro-mechanical behaviour of the composite based on Weibull's distribution and the hypothesis that the inclusion or the reinforcements carries no load after it has debonded from the matrix; they thus included the effect of fiber/matrix debonding in Mori–Tanaka's micromechanics method of deriving effective properties [Benveniste 1986]. The effect of the micromechanics-approach of progressive debonding is considered by gradually reducing the elastic constants of the inclusions [Sun, Tan et al. 2011]. The debonding of the fiber and the matrix can be modelled by reducing strengths of the interface among the reinforcement and the matrix [Nguyen and Khaleel 2004]. Additionally, the micro-mechanics assists in analysing the combined effects of micro-cracks and debonding on the effective properties of a composite [Meraghni, Desrumaux et al. 2002]. Moreover, the micro-mechanics approach has a limitations, such as expensive for real time application and the modelling and simulation requires extensive experimental characterisation to determine the material properties.

One of the way to overcome the aforementioned limitation is to use damage mechanics approach and to study damage and failure at the lamina level. Damage mechanics is related with the illustration, or modelling, of damage of materials that is appropriate for obtaining engineering predictions about the fracture of materials, initiation, and propagation of damage and without resorting to a microscopic report that would be difficult for real-world engineering design and analysis [Krajcinovic and Mastilovic 1995]. Damage mechanics demonstrates the typical engineering methodology to model complex phenomena. One of the final task of engineering research is to deliver the insight of the examined phenomenon, and to provide a rational predictive tool applicable in design [Krajcinovic and Mastilovic 1995]. There is a need of failure criterion to predict the initiation of the damage and to analyse its propagation leading to failure.

Ladeveze and Dantec adapted damage mechanics based approach to degrade elastic properties of the composite due to matrix cracking, and fiber breakage and a plasticity theory considering permanent deformations induced under shear loading [Ladeveze and LeDantec 1992]. Hassan and Batra have used three internal variables, such as material properties, loads and ply stacking sequence to illustrate the behaviour of composite after initiation of damage due to matrix cracking, fiber breakage, and fiber/matrix debonding [Hassan and Batra 2008]. Puck and Schürmann have comprehended Hashin's [Hashin 1983] stress-based failure criterion, and proposed methods to degrade elastic parameters of the lamina consequent to the initiation of a damage [Puck and Schürmann 2004]. To predict the damage in three dimensional (3-D) composite structure Donadon et al. have used a smeared crack approach, developed on the basis of damage mechanics [Donadon, Iannucci et al. 2008]. Clegg et al. have considered plastic deformations of a composite material and have well-defined a damage surface in terms of stresses induced [Clegg, Horsfall et al. 1988]. The development of damage variables is articulated in terms of a fracture energy, critical strain, and fracture stress and a local characteristic dimension, which may help to decrease the reliance of computed results upon the finite element (FE) mesh used to analyse the problem numerically. To model and analyse the failure behaviour of composite laminates under shear loads, the Hashin's failure criterion has been used to analyse the failure of the composite laminate. The damage ensues at any point can be characterized by different modes of failure, such as fiber breakage in tension and compression, and matrix cracking in tension and compression. The internal variables used to characterise different modes of failure depends upon values of stresses in Hashin's failure criteria, which are expressed in terms of the strength parameters for the composite, longitudinal and transverse tensile strength, longitudinal and transverse compressive strength, and shear strengths.

In this paper, a load cases has been considered; the cross ply composite laminate with the layup configuration  $[0,90]_{4S}$  was numerically modelled under shear loads. The failure has been

investigated using the Hashin's failure criterion. The numerical results were validated by comparing them against the wisely conducted experimental test data. The accuracy and the dependency of the finite element models have been discussed by analysing the difference between the experimental and numerically predicted stress and strain behaviour.

### Finite Element Modelling

A finite element simulation was carried out using a commercial software ABAQUS/CAE, which is an engineering tool that is used to solve various engineering problems ranging from linear to non-linear problems that are complex. Finite element modelling of composites is liable on the requirement of the investigation. In ABAQUS, there are numerous methods for composite modelling, such as, macroscopic modelling, microscopic modelling, mixed modelling, discrete reinforcement modelling and sub-modelling. However, the commonly used in finite element simulations of composite material are layered-shells using conventional shell elements. A 16 ply composite laminate was analysed under shear loads.

### Material and properties

The IM7/977-3 carbon epoxy composite materials are widely used in aerospace applications. Wide applications of this composite has resulted to analyse and predict the specific properties of the composite to obtain the long service life with minimum maintenance or repair of composites manufactured of IM7/977-3. Table 1 shows the properties of the composite used to numerically analyse the shear and three point bend loads.

**Table 1. Properties of unidirectional IM7/977-3 carbon/epoxy composites [Clay 2014].**

Property	Magnitude	Description
$E_{1T}$ (GPa)	164.3	Modulus in fibre direction in tension
$S_{11}^+$ (MPa)	2905	Maximum stress in fibre direction in tension
$\epsilon_{1T}$	0.01610	Maximum strain in fibre direction in tension
$\nu_{12}$	0.3197	Poisson's ratio in fibre direction in tension
$E_{2T}$ (GPa)	8.977	Modulus in 90-degree direction in tension
$S_{22}^+$ (MPa)	44.4	Maximum stress in 90-degree direction in tension
$\epsilon_{2T}$	0.00499	Maximum strain in 90-degree direction in tension
$\nu_{21}$	0.0175	Poisson's ratio in 90-degree direction in tension
$E_{1C}$ (GPa)	137.4	Modulus in fibre direction in compression
$S_{11}^-$ (MPa)	1274	Maximum stress in 0-degree direction in compression
$E_{2C}$ (GPa)	8.694	Modulus in 90-degree fibre direction in compression
$S_{22}^-$ (MPa)	247.6	Maximum stress in 90-degree direction in compression
$G_{12}$ (GPa)	4.880	Shear modulus calculated from $\pm 45$ -degree tension test
$\alpha_1$ ( $^{\circ}$ C)	0.01e-06	Coefficient of thermal expansion in fibre direction [Tay, Liu et al. 2008]
$\alpha_2 = \alpha_3$ ( $^{\circ}$ C)	32.73-06	Coefficient of thermal expansion in transverse direction [Tay, Liu et al. 2008]

### Failure criteria

Hashin [Hashin 1980] proposed three dimensional failure criterion of unidirectional fibre-reinforced composites. The criteria are established in terms of quadratic stress polynomials which are

articulated in terms of the transversely isotropic invariants of the applied average stress state. The four failure modes are: tensile fibre failure, compressive fibre failure, tensile matrix failure, and compressive matrix failure.

$S_{11}^+$  = Value of  $\sigma_{11}$  at longitudinal tensile failure

$S_{11}^-$  = Value of  $\sigma_{11}$  at longitudinal compressive failure

$S_{22}^+$  = Value of  $\sigma_{22}$  at transverse tensile failure

$S_{22}^-$  = Value of  $\sigma_{22}$  at transverse compressive failure

$S_{12}$  = Complete value of  $\sigma_{12}$  at longitudinal shear failure

$S_{23}$  = Absolute value of  $\sigma_{23}$  at transverse shear failure

If  $\sigma_{11} \geq 0$ , the Tensile Fibre Failure Criterion is:

$$F_f^+ = \left( \frac{\sigma_{11}}{S_{11}^+} \right)^2 + \alpha \left( \frac{\sigma_{12}}{S_{12}} \right)^2 \geq 1.0 \quad (1)$$

If  $\sigma_{11} < 0$ , the Compressive Fibre Failure Criterion is:

$$F_f^- = \left( \frac{\sigma_{11}}{S_{11}^-} \right)^2 \geq 1.0 \quad (2)$$

If  $\sigma_{22} \geq 0$ , the Tensile Matrix Failure Criterion is:

$$F_m^+ = \left( \frac{\sigma_{22}}{S_{22}^+} \right)^2 + \left( \frac{\sigma_{12}}{S_{12}} \right)^2 \geq 1.0 \quad (3)$$

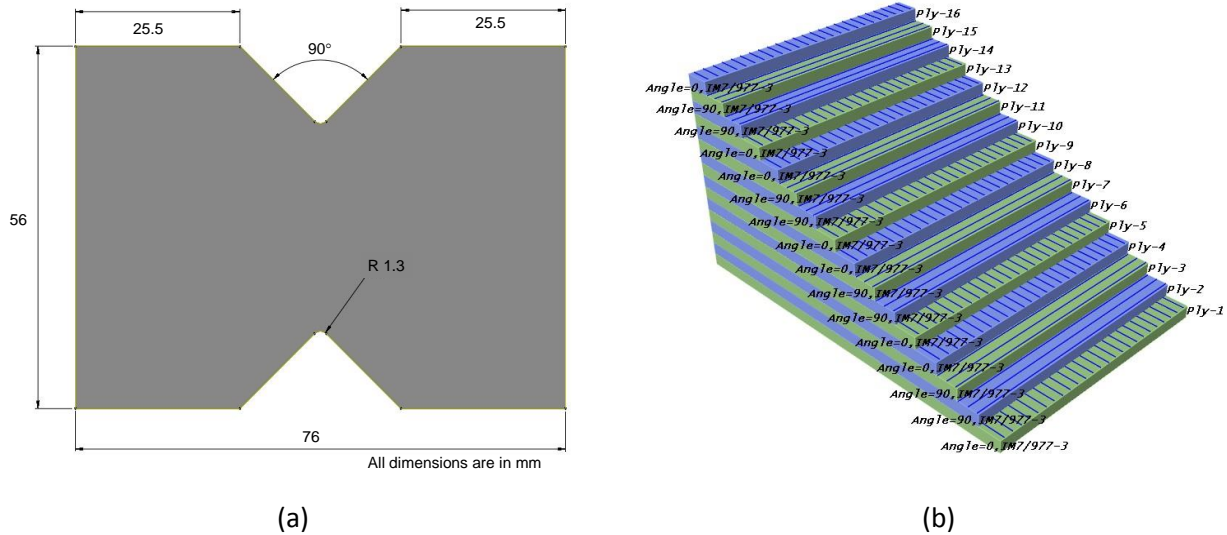
If  $\sigma_{22} < 0$ , the Compressive Matrix Failure Criterion is:

$$F_m^- = \left( \frac{\sigma_{22}}{2S_{23}} \right)^2 + \left[ \left( \frac{S_{22}^-}{2S_{23}} \right)^2 - 1 \right] \frac{\sigma_{22}}{S_{22}^-} + \left( \frac{\sigma_{12}}{S_{12}} \right)^2 \geq 1.0 \quad (4)$$

The Hashin's equations include two user-specified parameters:  $\alpha$  and  $S_{23}$ .  $\alpha$  is the user-specified coefficient that determines the contribution of the longitudinal shear stress to fibre tensile failure. Permissible range is  $0.0 \leq \alpha \leq 1.0$ , and the default value is  $\alpha=0$ .  $S_{23}$  is the transverse shear strength of the composite material. During modelling  $\alpha$  is maintained 1 and  $S_{23}$  has been extracted from the literature [Tay, Liu et al. 2008].  $\alpha$  is a coefficient that determines the contribution of the shear stress to the fibre tensile initiation criterion. Based on Hashin's model proposed in 1980 [Hashin 1983]  $\alpha$  is maintained as 1. The accuracy and predicative capability of the Hashin's failure criterion has been analysed, under shear loads, and the behaviour of the composite laminates are discussed below.

### Model Geometry

A double-notched cross ply composite laminate is of 76 mm length and 56 mm width are analysed under shear loads, with a notch tip radius of 1.3 mm as shown in (Figure 1). The composite laminate is made of 16 ply of 2 mm thick, with equal number of 0 and 90 degree fiber oriented plies, hence the composite layup configuration is balanced and symmetric (Figure 1).



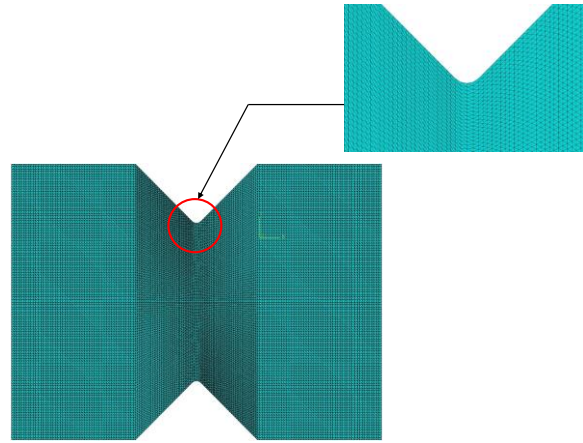
**Figure 1: (a) Model geometry and (b) Layup configuration of the composite laminate**

### Meshing and Boundary Conditions

ABAQUS stress/displacement shell elements use a Lagrangian formulation, where the element relocates; bestowing to the behaviour of the constituent material. Gaussian Quadrature is used to solve for static equilibrium at each integration point within the element. Shell elements are defined in ABAQUS using the \*SHELL SECTION. This specifies the material, shell element thickness and number of through thickness integration points. The optional composite parameter permits the user to define separate layers and orientations. Using this type of section characterize a multiple layered, complex, laminate can be precisely described and allocated to shell elements.

A convergence study is conducted to determine the minimum number of triangular elements required to produce converged mechanical properties and failure patterns (Figure 2). This is achieved by starting with a small number of elements and continuously refining the mesh until a maximum of 5 % difference in the maximum principal stress is observed. The difference in the maximum principal stress between the 0.5 mm and 0.25 mm mesh size is 4.22%, which is less than 5%. It is preferred to select the mesh size 0.5 mm over 0.25 mm, because the convergence difference below 5% is acceptable and the choosing mesh size 0.5 mm saves the computational time and memory.





**Figure 2: Finite element mesh of the composite laminate**

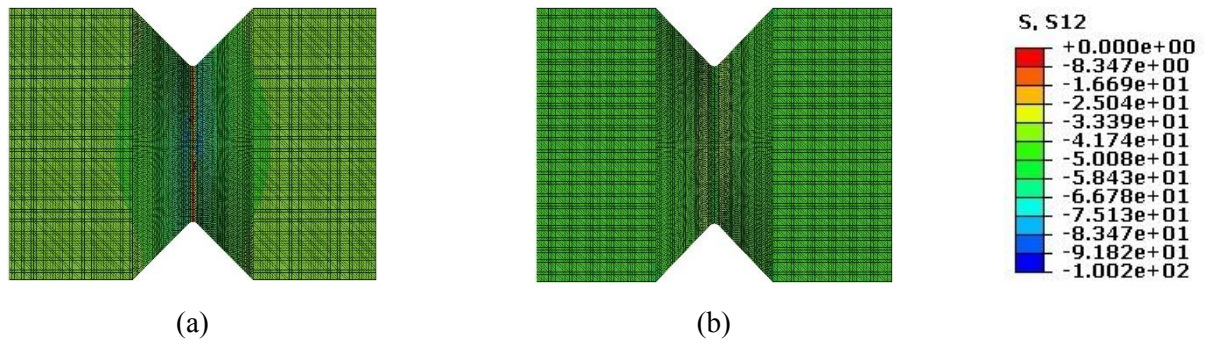
The fiber reinforcement in the double-notch shear test modelled composite is apparently planar, and its principal axis is lateral to the model axis, planes of low shear strength will be existing, comprising the plane that is parallel to the bottoms of the two opposing notches. Under the shear loading, a shearing action is developed along the specimen centre-line between the notch roots, apparently leading to a failure on this shear plane. The loading and the boundary conditions are shown in (Figure 3). The fixture used in experimental test, has been designed to minimize the compressive, tensile and bending influence on the specimen under test. In the similar way the model developed has been constrained to bend and behaviour non-linearly.



**Figure 3: Boundary and loading conditions**

### *Results and Discussions*

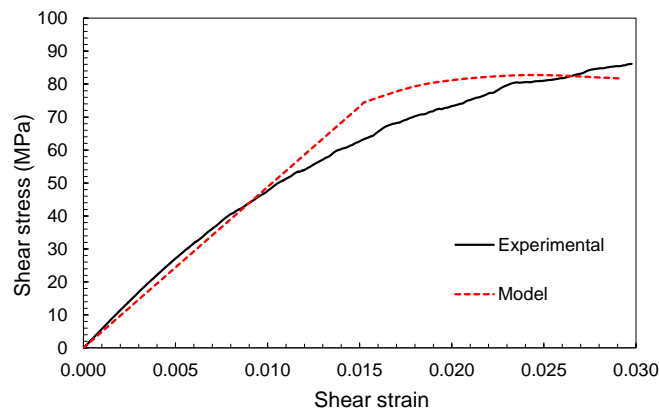
The fundamental problem with the Double-Notch Shear test models is that significant stress concentrations occur at the roots of the notches, resulting in premature failure. The below (Figure 4) shows the increased shear stress concentration in the roots of the notches of both fibers oriented in longitudinal and transverse direction.



**Figure 4: Stress distribution of [0,90]<sub>4S</sub> composite laminate under shear loads (a) 0-degree ply and (b) 90-degree ply**

The shear response is dependent on the fiber orientation in the composite laminate and it decreases from longitudinal direction (0-degree) to transverse direction (90-degree). The finite element model was developed to notice a simple shear state generated on the composite ply with double-notch. Since, the model's parallel faces are in opposite direction, under shear loads the model will develop horizontal cracks and delamination. But the model develop is based on the composite shell elements, hence the model cannot be used to predict delamination

The shear stress strain behaviour for [0/90]<sub>4S</sub> model revealed, after a few increment of loading both the curves showed proportional behaviour at the commencement. Nevertheless, just beyond the comparative limit, they turn into nonlinear due to the build-up of matrix cracks. As stress was increased, the eccentricity from the linearity also increased due to the initiation and propagation of the cracks along the fiber direction and failure of some fibers by tension in the transverse loading direction. The materials exhibited first ply failure at the shear stress of 74 MPa, after which the model under goes series of damage in the elements of the adjacent plies, those are the elements which fail to carry loads. The maximum shear stress of 86 MPa and 81 MPa are observed under experimental and modelled results. The difference between the experimental and modelled results are 6% (Figure 5).



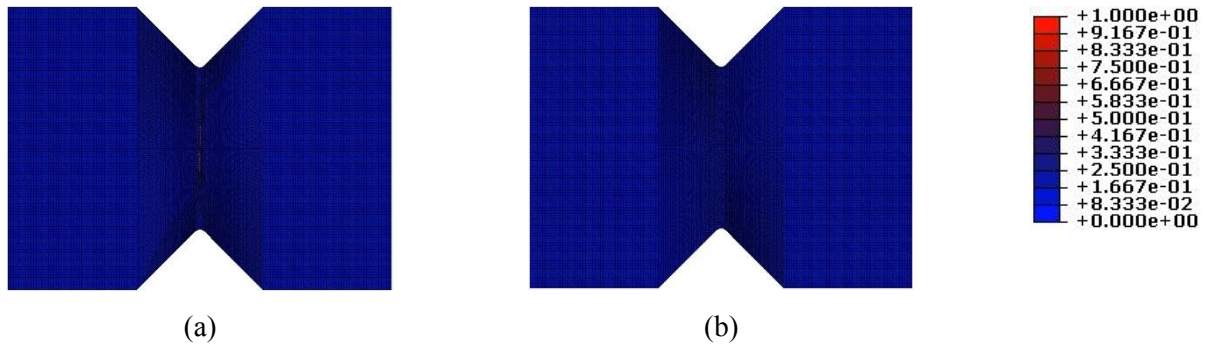
**Figure 5: Stress strain behaviour of [0,90]<sub>4S</sub> composite laminate under shear loads**

**Failure**

The damage and failure prediction helps in knowing the failure pattern and critical elements that undergo more deformation, leading to catastrophic failure. Implementing Hashin's failure criterion the matrix and fiber failure modes are analysed and discussed below. The red coloured area represents the damaged portion.

**Fiber compression**

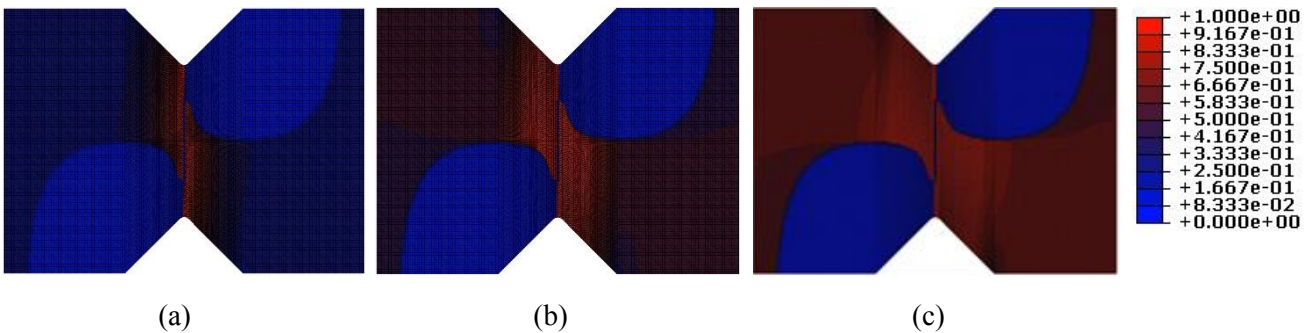
Under shear loads the developed model does not show any significant amount of damage in the fibers under compression. Because the fibers oriented in longitudinal and transverse direction has the potential to withstand compressive loads leading to failure (Figure 6).



**Figure 6: Comparison of Fiber compression failure pattern (a) 0-degree ply and (b) 90-degree ply**

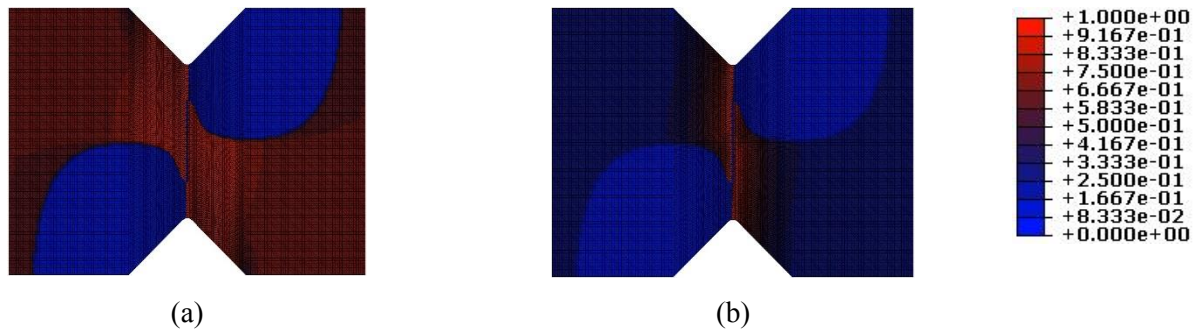
**Fiber tension**

A significant amount of damage is observed in both the plies (0-degree and 90-degree). The fibers oriented in longitudinal direction shows an initiation of the damage; in line with the notch tip. The damage propagates with the increase in load, and the direction of damage propagation is inclined at an angle of 45° to the orientation of fibers. Additionally the damaged areas shows the tensile failure zones in the model (Figure 7).



**Figure 7: Fiber tension failure pattern of [0,90]<sub>4s</sub> composite laminate under shear loads**

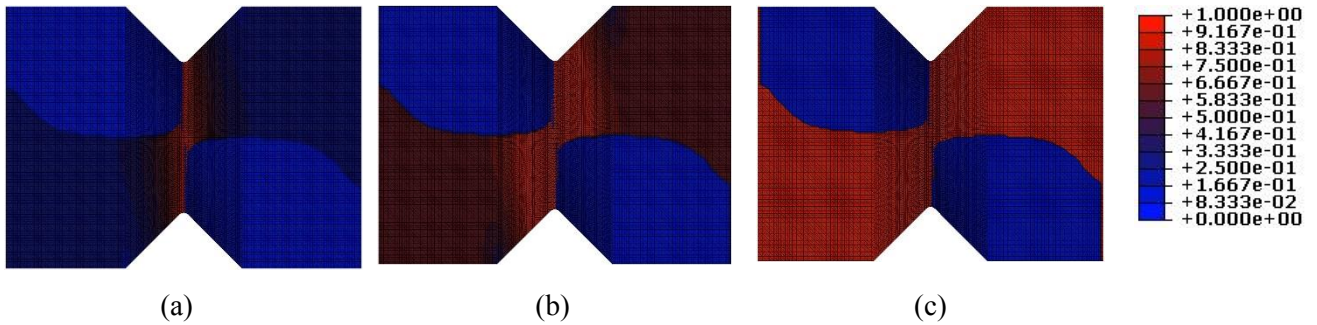
Comparatively, the damage initiation and propagation in 90-degree plies are less, with respect to 0-degree plies (Figure 8). It is mainly due to the shear response is dependent on the fiber orientation, the shear response is maximum in fibers oriented in longitudinal direction, and it decreases as the fibers are oriented in transverse direction. Hence, the use of computational model, and implementing the Hashin’s failure criteria shows the dependence of shear failure upon the orientation of fibers.



**Figure 8: Comparison of Fiber tension failure pattern (a) 0-degree ply and (b) 90-degree ply**

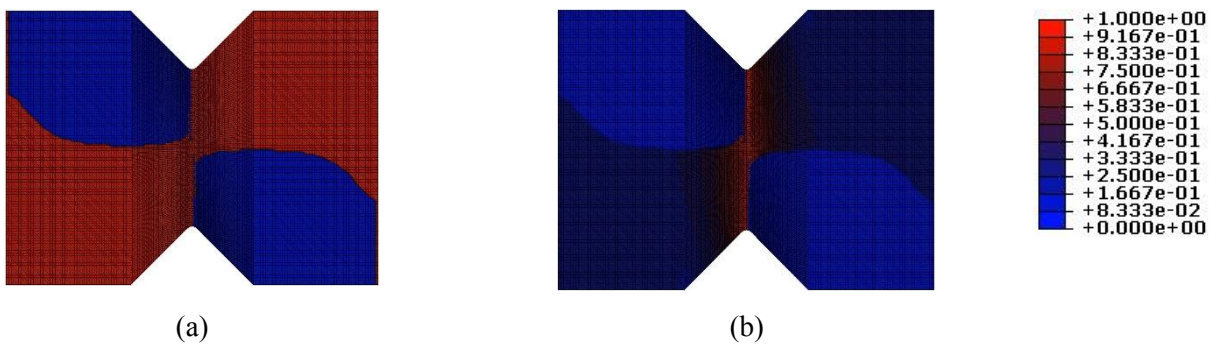
**Matrix compression**

The fiber reinforced polymer matrix composite laminate with v-notches suffers from loss in its stiffness, strength, and service life due to notch-enhanced stresses. The presence of notch results prompting new modes of damage in the matrix of the composite laminate or by accelerating the growth of the existing damage. Figure 9 shows the initiation of the matrix damage at the centre of the composite in line with the notch root. A rapid increase in damage areas has been observed in the subsequent increment between the notch root tips. It is mainly due to global stiffness reduction and decrease in strength in the respective failure areas of the composite as shown in (Figure 9). However, the matrix damage developed by a double-notch is moderately small compared to the effects of a notch on the ultimate strength of the model.



**Figure 9: Matrix compression failure pattern of [0,90]<sub>4s</sub> composite laminate under shear loads**

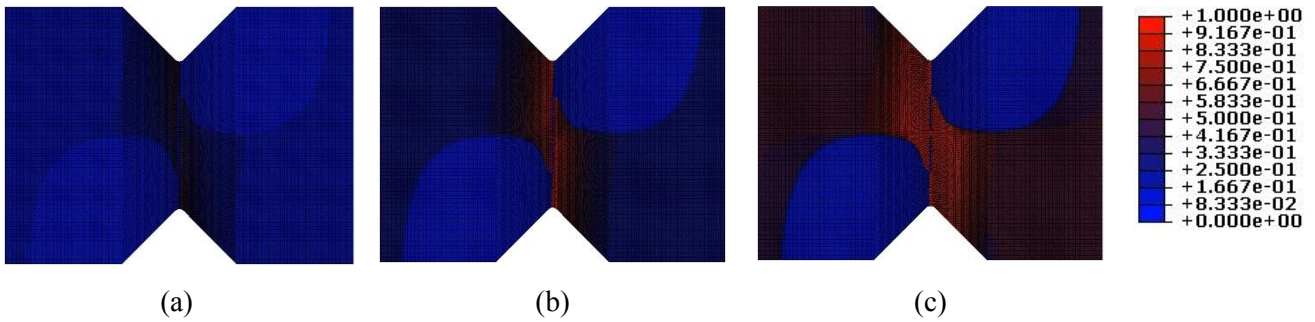
Relatively the matrix damage is more dominantly observed in the fibers oriented in longitudinal direction, but the areas of damage is comparatively same (Figure 10). This conveys the fiber orientation doesn't play a vital role in the regions of the matrix damage.



**Figure 10: Comparison of matrix compression failure pattern (a) 0-degree ply and (b) 90-degree ply**

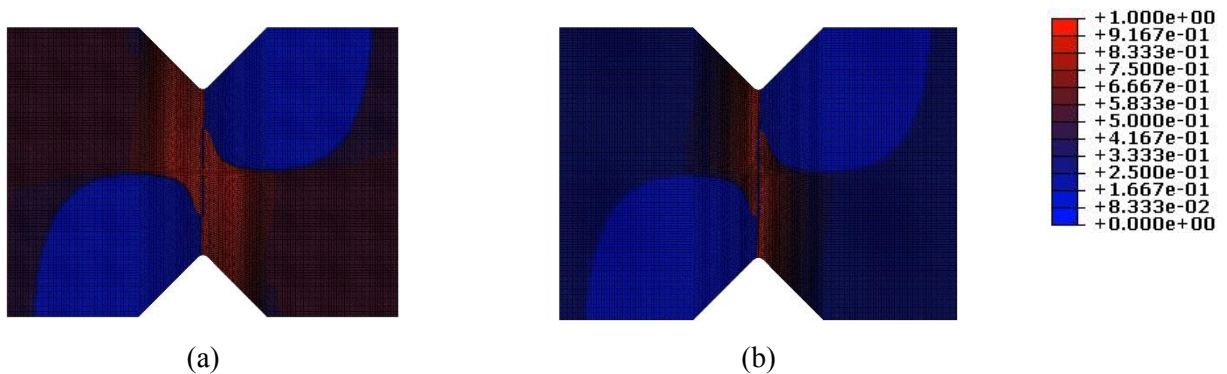
**Matrix tension**

It has been observed from the experiments that the failure mechanisms leading to failure were the matrix cracks in transverse direction. Furthermore increase in loads will result in longitudinal cracks (splits) appearing and propagating in the longitudinal direction. A similar way of damage initiation and propagation has been observed in the finite element model, specifically in the matrix damage. The matrix which fails under tension also exhibits similar pattern of failure as matrix fail under compressive nature of the force, but the location of damage is opposite in nature (Figure 11). The region where matrix are damage is entirely dependent on the loading and the boundary conditions.



**Figure 11: Matrix tension failure pattern of [0,90]<sub>4s</sub> composite laminate under shear loads**

When the matrix in the 0-degree laminate are fully damaged, the matrix in the 90-degree laminate are only 20% damaged. Additionally, it is evident that for both matrix tensile and matrix compressive damage are dominantly observed in 0-degree plies (Figure 12).



**Figure 12: Comparison of matrix tension failure pattern (a) 0-degree ply and (b) 90-degree ply**

The damage patterns observed in simulated results are comparable with the experimental results. Figure 13 shown below shows the failed test specimen. The visible damage pattern is of 0-degree ply, which is comparable with the simulated results (Figure 7, Figure 9 and Figure 11).



**Figure 13: failure of carbon-epoxy specimen under shear loads**

## Conclusions

A progressive failure approach was applied to study the failure in terms of transverse and longitudinal fracture of cross-ply laminates with double-notch. The local matrix and fiber dominated failure was analysed by adapting the Hashin's failure criteria. The shear stress and strain behaviour of the model was analysed by a uniform stiffness reduction approximation. In the initial stages of loading the material behaved linearly, after reaching a yield shear stress of 75 MPa a non-linear behaviour has been observed with the development of damage in transverse direction. Furthermore increment of the load resulted in rapid development and propagation of damage in longitudinal direction. A difference of 6% between the experimental and modelled shear stress has been observed. Additionally, it has been evidently observed that a matrix dominated failure can be predicted in the double-v-notch composite specimen. It was found that the failure in shear model of  $[0, 90]_{4S}$  was due to pure shear along the line between V-notches, where shear cracks were almost parallel to the loading direction, but in the off axis angle  $45^\circ$ , cracks were initiated at roots of notch and propagated in the direction of fibers.

## Reference

- Benveniste, Y. (1986). "On the Mori-Tanaka's method in cracked bodies." *Mechanics Research Communications* 13(4): 193-201.
- Chen, J.-F., E. V. Morozov and K. Shankar (2014). "Simulating progressive failure of composite laminates including in-ply and delamination damage effects." *Composites Part A: Applied Science and Manufacturing* 61(0): 185-200.
- Clay, S. (2014). "Assess and Quantify the Benefits of Applying Damage Tolerant Design Principles to Advanced Composite Aircraft Structures.", FA8650-08-D-3858. U. S. A. F. R. Laboratory.
- Clegg, W. J., I. Horsfall, J. F. Mason and L. Edwards (1988). "The tensile deformation and fracture of Al-"Saffil" metal-matrix composites." *Acta Metallurgica* 36(8): 2151-2159.
- Donadon, M. V., L. Iannucci, B. G. Falzon, J. M. Hodgkinson and S. F. M. de Almeida (2008). "A progressive failure model for composite laminates subjected to low velocity impact damage." *Computers & Structures* 86(11-12): 1232-1252.
- Hashin, Z. (1980). "Failure criteria for unidirectional fiber composites." *Journal of applied mechanics* 47(2): 329-334.
- Hashin, Z. (1983). "Analysis of Composite Materials—A Survey." *Journal of Applied Mechanics* 50(3): 481-505.
- Hassan, N. M. and R. C. Batra (2008). "Modeling damage in polymeric composites." *Composites Part B: Engineering* 39(1): 66-82.
- Hinton, M. J., A. S. Kaddour and P. D. Soden (2004). "A further assessment of the predictive capabilities of current failure theories for composite laminates: comparison with experimental evidence." *Composites Science and Technology* 64(3-4): 549-588.

- Kaddour, A. S., M. J. Hinton and P. D. Soden (2004). Chapter 6.1 - Predictive capabilities of nineteen failure theories and design methodologies for polymer composite laminates. Part B: Comparison with experiments. Failure Criteria in Fibre-Reinforced-Polymer Composites. M. J. H. S. K. D. Soden. Oxford, Elsevier: 1073-1221.
- Kaddour, A. S., M. J. Hinton and P. D. Soden (2004). "A comparison of the predictive capabilities of current failure theories for composite laminates: additional contributions." *Composites Science and Technology* 64(3-4): 449-476.
- Krajcinovic, D. and S. Mastilovic (1995). "Some fundamental issues of damage mechanics." *Mechanics of Materials* 21(3): 217-230.
- Ladeveze, P. and E. LeDantec (1992). "Damage modelling of the elementary ply for laminated composites." *Composites Science and Technology* 43(3): 257-267.
- Meraghni, F., F. Desrumaux and M. L. Benzeggagh (2002). "Implementation of a constitutive micromechanical model for damage analysis in glass mat reinforced composite structures." *Composites Science and Technology* 62(16): 2087-2097.
- Nguyen, B. N. and M. A. Khaleel (2004). "A mechanistic approach to damage in short-fiber composites based on micromechanical and continuum damage mechanics descriptions." *Composites Science and Technology* 64(5): 607-617.
- Puck, A. and H. Schürmann (2004). Chapter 5.6 - Failure analysis of FRP laminates by means of physically based phenomenological models. Failure Criteria in Fibre-Reinforced-Polymer Composites. M. J. H. S. K. D. Soden. Oxford, Elsevier: 832-876.
- Reifsnider, K. (1980). "Fatigue behavior of composite materials." *International Journal of Fracture* 16(6): 563-583.
- Rotem, A. (1998). "Prediction of Laminate Failure with the Rotem Failure Criterion." *Composites Science and Technology* 58(7): 1083-1094.
- Sun, X. S., V. B. C. Tan and T. E. Tay (2011). "Micromechanics-based progressive failure analysis of fibre-reinforced composites with non-iterative element-failure method." *Computers & Structures* 89(11-12): 1103-1116.
- Tay, T., G. Liu, V. Tan, X. Sun and D. Pham (2008). "Progressive failure analysis of composites." *Journal of Composite Materials*.

# **A computational approach for the modelling of rolling shear cracks in cross-laminated timber structures**

**†E.I. Saavedra Flores<sup>1</sup>, K. Saavedra<sup>2</sup>, and R. Das<sup>3</sup>**

<sup>1</sup>Departamento de Ingeniería en Obras Civiles, Universidad de Santiago de Chile. Av. Ecuador 3659, Estación Central, Santiago, Chile.

<sup>2</sup>Departamento de Tecnologías Industriales, Universidad de Talca, Los Niches km 1, Curicó, Chile.

<sup>3</sup>Department of Mechanical Engineering, University of Auckland, Auckland 1010, New Zealand

†Presenting and Corresponding author: erick.saavedra@usach.cl

## **Abstract**

This paper addresses the computational modelling of rolling shear cracks in cross-laminated timber structures. In order to predict the structural response, four spatial scales are interlinked within a multi-scale modelling framework. Material information is taken from the wood cell-wall at the order of few nanometers, wood fibres with dimensions of some micrometers and growth rings described by a few millimeters. A computational homogenisation scheme is adopted to determine the effective mechanical properties at each scale. The homogenised mechanical properties are then used to analyse the fourth (structural) scale represented by a cross-laminated timber plate with dimensions of the order of the meter. In order to simulate the cracking in the material, a cohesive zone model is adopted at the homogenised macroscopic scale. This approach allows us to model interlaminar and inter-fibre cracks. Our numerical simulations reveal the potential predictive capabilities of the present approach to investigate further wood and other natural materials.

**Keywords:** Cross-laminated timber, Rolling shear cracks, Multi-scale modelling

## **Introduction**

Cross-laminated timber (CLT) consists of structural panels made up of several layers of boards stacked crosswise and glued together on their faces. Among its main advantages, we can highlight its favorable seismic performance, its ability to self-protect against fire and its excellent strength, which allows wood to be used in tall buildings with heights up to 30 stories [Fairhurst et al. (2010)].

One important issue in the design of CLT structures which still requires further investigation is the rolling shear failure [Zhou et al. (2014)]. It consists of inter-fibre cracking due to shear strains in the plane perpendicular to the longitudinal axis of the wood fibres. Figure 1 shows a typical rolling shear failure found in the central layer of a CLT plate subject to out-of-plane loads. In particular, the design of CLT floor systems with low span-to-depth ratios is often governed by the rolling shear capacity of CLT plates and therefore, its full understanding is of paramount importance to prevent damage in CLT structures.





**Figure 1: Typical rolling shear failure in a CLT plate subject to bending**

In order to capture rolling shear cracking in a CLT plate, we propose in this paper a modelling strategy which combines the use of a homogenisation-based multi-scale modelling framework to determine the undamaged mechanical properties of wood, and the adoption of cohesive interfaces at the homogenised macroscopic structural scale to model the crack behaviour.

In the context of multi-scale modelling of CLT structures, few attempts have been made in order to predict the CLT structural response [Saavedra Flores et al. (2014); Saavedra Flores et al. (2015a; 2015b)]. We note, however, that despite the increasing interest in this subject, the complete understanding of the mechanical properties of CLT is still an issue which remains open at present. In this new paper, we continue with the line of development started in the above references [Saavedra Flores et al. (2014); Saavedra Flores et al. (2015a; 2015b)] by presenting new numerical results.

### **Computational approach**

In the present paper, we adopt a homogenisation-based multi-scale constitutive framework in which each material scale is associated with a microstructure whose most statistically relevant features are incorporated within a representative volume element (RVE). This RVE is assumed to have a (microscopic) characteristic length much smaller than the macro-continuum, and at the same time, a size large enough to capture the microscopic heterogeneities in an averaged sense.

In this theory it is assumed that the macroscopic or homogenised strain tensor at any arbitrary point of the macroscopic continuum is the volume average of the microscopic strain tensor field over the domain of a representative volume element of material (RVE). Similarly, the macroscopic or homogenised stress tensor field is assumed to be the volume average of the microscopic stress tensor.

This multi-scale framework is adopted to find the homogenised constitutive response at each material scale. In wood, these scales are represented by the wood cell-wall at the order of few nanometers, the wood fibres with cross section dimensions of tens of micrometers, and the growth rings described by some few millimeters. For further information about the morphology and composition of wood at the nano- and microscopic scale level, we refer, for instance, to [Dinwoodie (1981)]. The homogenisation of these three (material) scales (represented by three different RVEs) allows us to predict the response of the fourth (structural) scale, that is, the CLT plate.

The main philosophy behind the present multi-scale strategy is to start from the response of very basic (but fundamental) ingredients at small scales and then, build up an increasingly complex and intricate response as the length scale increases. We note here that the determination of an accurate prediction of this response would (probably) be unfeasible by means of conventional phenomenological models.

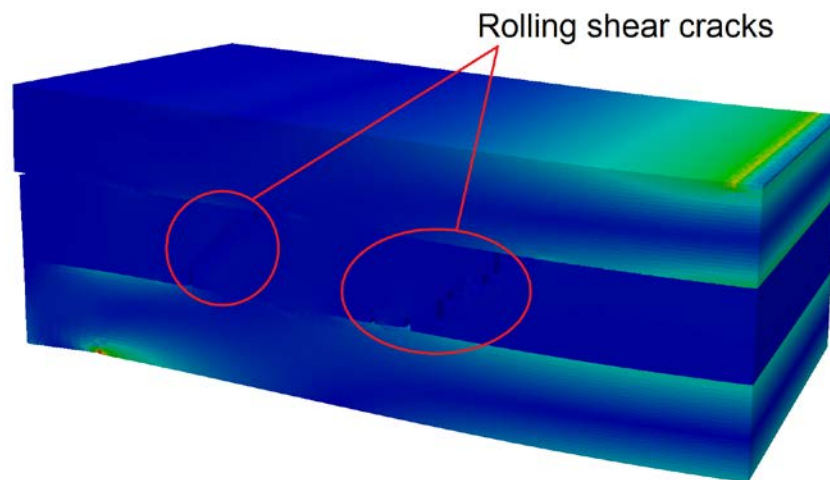
As we are interested in modelling cracking in the material, we adopt a Cohesive Zone Model (CZM) at the macroscopic structural scale. CZM is the simplest model that allows to describe in full a fracture process (i.e. initiation and propagation of the crack) and has been thoroughly used to treat several materials such as concrete, rocks, fibre-reinforced plastics and wood [Allix et al. (1998); Elices et al. (2002), Saavedra et al. (2012)].

The basic idea of the CZM can be described as a zero thickness interface transferring tractions which are related to the displacement jump of the interface  $[u]$  by the meaning of a softening function. This evolution law can be written in terms of a damageable stiffness operator  $k([u])$ . At the beginning, the interface stiffness has no damage ( $k_0$ ). Then, the stiffness decreases with respect to the displacement jump and becomes zero at some critical displacement jump. It is possible to use a damage variable  $d$  to represent the stiffness of the interface, i.e.,  $k = (1-d)k_0$ , with  $d$  ranging from 0 (healthy interface point) to 1 (completely damaged interface point). The area under the entire stress-displacement jump curve is the energy per unit area  $G_f$  [ $\text{J/m}^2$ ] necessary to separate completely the interface at a given point.

## Numerical simulations

This section describes the numerical results obtained by the present multi-scale approach enriched with macroscopic cohesive laws. The CLT specimen consisted of three 4-cm-thick layers with a length of 75 cm and a width of 39 cm. The span length between supports was 60 cm. The outer layers were made up of timber pieces oriented in the strong direction of the panel. The central layer was made up of members oriented in the weak direction. Edge-gluing is considered between the opposite sides of adjacent layers. The macroscopic finite element mesh consists of 861696 linear wedge elements and 2.9 millions of DOFs (because of the symmetry of the problem, only one quarter of the geometry is considered). Further details on the finite element models of the RVEs associated with the corresponding sub-scales can be found in [Saavedra Flores et al. (2015b)]. Cohesive interfaces are used to simulate the rolling shear failure, but a distinction between the interlaminar interfaces (for delamination) and the inner interfaces of the central layer (for inter-fibre cracking) is made. When a cohesive interface is completely delaminated, contact conditions are considered to avoid interpenetration.

Figure 2 shows the computational simulation of the rolling shear failure in the CLT plate subject to three-point-bending. The corresponding contour plot shows the principal stresses. In our simulation, two parallel cracks are predicted consistently with experiments (not detailed here), along with some delaminated regions, mainly between the bottom and central layer. In our numerical predictions, we note that after a cohesive interface has become fully damaged, it is converted into a contact interface as well.



**Figure 2: Computational simulation of rolling shear failure in a CLT plate subject to bending**

Furthermore, our model predicts a critical load of 160000 N when the first crack starts propagating in the material, which coincides with the critical value obtained during the experiment.

## Conclusions

This paper has addressed the modelling of rolling shear cracks in CLT plates subject to three-point-bending. Micromechanical information coming from three scales have been taken into account in order to compute the effective material properties for the analysis of the structural scale. Cracking has been captured by introducing macroscopic cohesive interfaces in the model. The proposed approach has been able to capture key features in the rolling shear failure of CLT plates, revealing the potential applications of our approach on the study of CLT structures under different configurations of layers and loading conditions.

Finally, we remark that studies are currently under way to explore the buckling behavior and buckling-delamination interaction in CLT walls under compressive loads. This will be the subject of a future publication.

## Acknowledgments

E.I. Saavedra Flores acknowledges the financial support from the Chilean National Commission for Scientific and Technological Research (CONICYT), FONDECYT REGULAR research project No 1140245.

K. Saavedra acknowledges the financial support from CONICYT, FONDECYT Initiation into Research project No 11130623.

## References

- Allix, O., Lévêque, D., Perret, L., Identification and forecast of delamination in composite laminates by an interlaminar interface model, *Composites Science and Technology* 58 (5) (1998) 671–678.
- Dinwoodie, J., *Timber – Its nature and behavior*, Von Nostrand Reinhold, New York, 1981.
- Elices, M., Guinea, G., Gómez, J. and Planas, J., The cohesive zone model: advantages, limitations and challenges, *Engineering Fracture Mechanics* 69 (2) (2002) 137 – 163.
- Fairhurst, M., Zhang, X. and Tannert, T., Nonlinear dynamic analyses of a novel timber-steel hybrid system, in: *Proceedings of the World Conference on Timber Engineering*, Canada, 2010.
- Saavedra, K., Allix, O. and Gosselet, P., On a multiscale strategy and its optimization for the simulation of combined delamination and buckling, *International Journal for Numerical Methods in Engineering* 91 (7) (2012) 772–798.
- Saavedra Flores, E.I., DiazDelaO, F.A., González Soto, P., Pérez Pulgar, E., An investigation of cross-laminated timber by means of a multi-scale framework, in: B. H. V. Topping, P. Iványi (Eds.), *Proceedings of the Twelfth International Conference on Computational Structures Technology*, Civil-Comp Press, Stirlingshire, UK, Paper 90, 2014. doi:10.4203/ccp.106.90.
- Saavedra Flores, E.I., Dayyani, I., Ajaj, R., Castro-Triguero, R., DiazDelaO, F., Das, R. and González Soto, P., Analysis of cross-laminated timber by computational homogenisation and experimental validation, *Composite Structures* 121 (0) (2015a) 386 – 394.
- Saavedra Flores, E.I., Saavedra, K., Hinojosa, J., Chandra, Y. and Das, R., Multi-scale modelling of rolling shear failure in cross-laminated timber structures by homogenisation and cohesive zone models. Submitted to *International Journal of Solids and Structures* (2015b).
- Zhou, Q., Gong, M., Chui, Y.H. and Mohammad, M., Measurement of rolling shear modulus and strength of cross laminated timber fabricated with black spruce, *Construction and Building Materials* 64 (2014) 379386.

# An Implicit Algorithm for Finite Volume - Finite Element Coupling

Davor Čokljat<sup>1</sup>, Aleksandar Jemcov<sup>\*2</sup>, and Joseph P. Maruszewski<sup>3</sup>

<sup>1</sup>ANSYS Inc., 6 Europa View, Sheffield, South Yorkshire S9 1XH, UK

<sup>2</sup>Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

<sup>3</sup>ANSYS Inc., 1007 Church Street Evanston, IL 60201, USA

## Abstract

We present an implicit coupling algorithm that is suitable for strongly coupled physical problems that were discretized by heterogeneous numerical schemes, namely finite volume and finite element methods. The primary characteristic of the proposed scheme is an implicit treatment of the heterogeneous schemes through a single matrix approach. The finite element and finite volume parts of the discretized domain exchange information through a coupling boundary and the resulting discretization coefficients are stored in a block matrix. The structure of the matrix is such that the coupling coefficients are stored in the off-diagonal blocks of the matrix, while finite element and finite volume subdomains are stored in the diagonal blocks of the matrix. A suite of efficient linear solvers based on the Krylov subspace methods were developed for the solution of the resulting coupling problem. Several demonstration cases that illustrate the coupling algorithm are presented.

**Keywords:** Finite Element Method (FEM); Finite Volume Method(FVM); Coupling algorithm, Block matrix

## Introduction

Multiphysics problems are prevalent in today's engineering practice. It is hard to imagine a device that does not need structural, thermal and fluid flow analysis in order to design it for the safe operation and the peak efficiency. Today's engineering practice rarely undertakes a numerically integrated approach to analyzing and simulating the proposed designs. The current design practice takes an iterative approach to simulations and analysis through a series of the stages that involve the use of computational fluid dynamics, thermal and stress analysis. Numerical difficulties arise in a staged approach due to increased stiffness of the problem and

the loss of coupling among equations. A typical approach to simulating coupled phenomena involving solids and fluids is done through an exchange of boundary conditions through the coupling boundaries. This process is inherently iterative. The coupling between fields in fluids and solids can be recovered by an iterative procedure in which the fields on the coupling boundaries acting as boundary conditions for different analyses, are updated in this iterative process.

This simulation of coupled problems is further complicated by the use of different simulation practices that involve different discretization methods used for the particular stage of the analysis. In fluid flow, finite volume method is commonly used for the discretization of governing equations. In analyses that involve solid materials finite element is the method of choice for producing the discrete systems of equations. The choice of the discretization method that is being used for a given problem is often dictated by the efficiency and accuracy requirements. The Finite Volume Method (FVM) is often used in computational fluid dynamics (CFD) as a method of choice due to its simplicity and ability to reproduce the conservation laws. The simplicity of the finite volume discretization stems from the fact that the low degree polynomials ( $C^0$ ) are used for the interpolation within finite volumes. The numerical efficiency of the finite volume schemes stems from the fact that only one integration point per face of the finite volume cell is required to evaluate the numeric flux. This approach yields a low storage numerical scheme that produces very sparse matrices. Furthermore,  $C^0$  interpolation functions used for the representation of the variables within a finite volume cell allows the usage of arbitrary shapes of finite volumes cells, thus simplifying the mesh generation for the domains with the complex geometric shapes. In this work we are concerned with the cell-centered finite volume method that stores all variables in cell centers. Matrix coefficients arising in implicit discretization of the finite volume problems are obtained by evaluating fluxes on the cell face centers. This approach is chosen intentionally given the fact that cell-centered finite volume discretization is the dominant approach in the CFD community. Data structures required for the efficient representation of the finite element connectivity on unstructured meshes of arbitrary shapes is of the so-called face-to-cell type. In practical terms, face-to-cell connectivity allows very fast access to the cell data required for the flux evaluation in the face center.

On the other hand, finite element method is a preferred approach for the problem solution in structural analysis and in computational mechanics in general. Finite element approach to discretization to problems in mechanics offers a strong mathematical foundations that allow for error and convergence estimates even for challenging computational problems. Usage of higher order interpolation polynomials enables a higher order of accuracy when compared to finite volumes for the nominally same size of the computational mesh. The finite element method is particularly well suited for the elliptic problems arising in linear elasticity and structural mechanics. Since the finite element data is stored in the nodes of the finite elements, the data structure required for the representation of the finite element discretization on unstructured meshes consists of node-to-cell connectivity information. Matrix coefficients in the finite element discretization are obtained as the collection of the contributions from all finite elements sharing the particular node.

In recent years researchers have started addressing the problem of coupled sim-

ulations. Geiger et al. [4] proposed an algorithm for coupling of nodal based finite volume and finite elements using overlapping groups of finite elements and finite volumes. Galerkin finite elements were used to provide the second order of accuracy interpolation for flux evaluations on finite volume grid. This approach may be classified as a control volume finite element method using dual grids [1]. Lazarov et al. [5] proposed the method for coupling finite volume-finite elements by using the node based finite volume scheme on dual grids. Sardella [3] proposed a mixed finite element/volume method that used finite volume approach to discretizing convective terms while finite element approach was used for the discretization of the diffusion terms in convection diffusion problems. The mixed algorithm was applied to the singularly perturbed problems in fluid mechanics providing the numerical stability to the computations. Gadeschi et al. [2] proposed the coupling method based on hierarchical Cartesian grids for heat transfer between solids and fluids. Vierendeels [6] and Sicklinger et al. [7] recently proposed frameworks for a general coupling between codes. In both works the idea is based on using the Jacobians of the governing equations to create the coupling conditions for coupled simulations. Vierendeels and Sicklinger algorithms are examples of the explicit coupling of black-box solvers within the framework of iterative coupling approach.

It is observed that in the previous attempts to couple finite volume and finite element methods, the approach was to modify one or the other method across the shared interface in order to make the approach more suitable for the discretization. Contrary to that, in this work we maintain the characteristics of each method used for the discretization of the respective part of the domain. Therefore, we propose an approach to coupling that maintains the discretization practices of both finite volume and finite element methods.

A novel implicit coupling algorithm for the mixed discretizations involving finite elements and finite volumes that exchange information along one or more boundaries called coupling interfaces is proposed. The resulting discretization coefficients are stored in a block matrix in which coupling interface coefficients are stored in off-diagonal blocks while the finite element and finite volume discretizations are stored in diagonal blocks. We propose an algorithm for coupled interfaces that uses native information from each discretization scheme to produce the necessary data for coupling of finite element and finite volume discretizations. The method is conservative and there is no loss of mass, energy or momentum across the interface even though substantially different discretization schemes are used on each side of the interface. Linear solver suite capable of handling block matrices arising in coupled finite element-finite volume discretizations is also developed. Two computational examples are presented as an illustration of the coupled algorithm. The first example demonstrates the coupling of finite elements and finite volumes in the case of the energy equation for the solid with an interface separating finite volume from finite element discretizations. The second example is coupled fluid-solid heat transfer problem in which energy equation is coupled through the solid-to-fluid interface.

### **Governing equations and boundary conditions**

Consider the energy equation in the solid shown in Figure (1)

$$k\partial_i^2 T = f \text{ in } \Omega = \Omega_{FE} \cup \Omega_{FV}. \quad (1)$$

Constant heat conduction coefficient independent of spatial location and temperature was assumed. The energy equation is closed by adding the corresponding boundary conditions

$$B(T) = b \text{ on } \Gamma = \Gamma_{FE} \cup \Gamma_{FV}. \quad (2)$$

In addition to the solid energy equation, the coupled system might have a fluid domain. In that case, the energy equation in the fluid is given by the convection-diffusion equation

$$\partial_i(\rho u_i T) = k\partial_i^2 T \text{ in } \Omega = \Omega_{FE} \cup \Omega_{FV}, \quad (3)$$

supplemented by the boundary conditions

$$C(T) = c \text{ on } \Gamma = \Gamma_{FE} \cup \Gamma_{FV}. \quad (4)$$

Equations (1) and (3) represent the steady-state equations of energy transport in solids and fluids. The boundary conditions considered here are of Neumann and Dirichlet type. The computational domain is general is divided in two parts by an internal surface called coupled boundary separating regions where finite volume and finite element discretizations are applied. However, the coupling boundary might be a physical boundary separating fluid from a solid thus representing a physical surface coinciding with the coupling interface.

An internal consistency condition can be added to the governing equations (1) through (4) that enforces the conservation of energy across the coupling interface as illustrated in Figure (1) In simple terms, the energy flux across the coupling interface is preserved and we can write the balance equation

$$\int_{\Gamma^-} f_i^- n_i d\Gamma - \int_{\Gamma^+} f_i^+ n_i d\Gamma = 0. \quad (5)$$

In the case of heat conduction, the consistency condition becomes

$$\int_{\Gamma^-} (k\partial_i T)_i^- n_i d\Gamma - \int_{\Gamma^+} (k\partial_i T)_i^+ n_i d\Gamma = 0. \quad (6)$$

Internal consistency condition is used to produce the consistent interpolations between finite elements and finite volumes for both cases when the interface is an arbitrary surface separating discretization zones and as well as when the interface coincides with both the physical and discretization boundaries. Consistent interpolation is defined here in terms of the energy conservation as well as in terms of bridging the different requirements for the smoothness of the interpolation functions. In other words, the interface consistency condition must allow conservation of energy even when on one side the data variation within a finite volume cell is represented by functions belonging to  $C^0$  space and on the other side the data variation within a finite volume is represented by  $C^1$  space. We limit our discussion here to interfaces that have  $C^0$  to  $C^1$  (and vice versa) transitions. Higher order transitions are the subject of ongoing research.



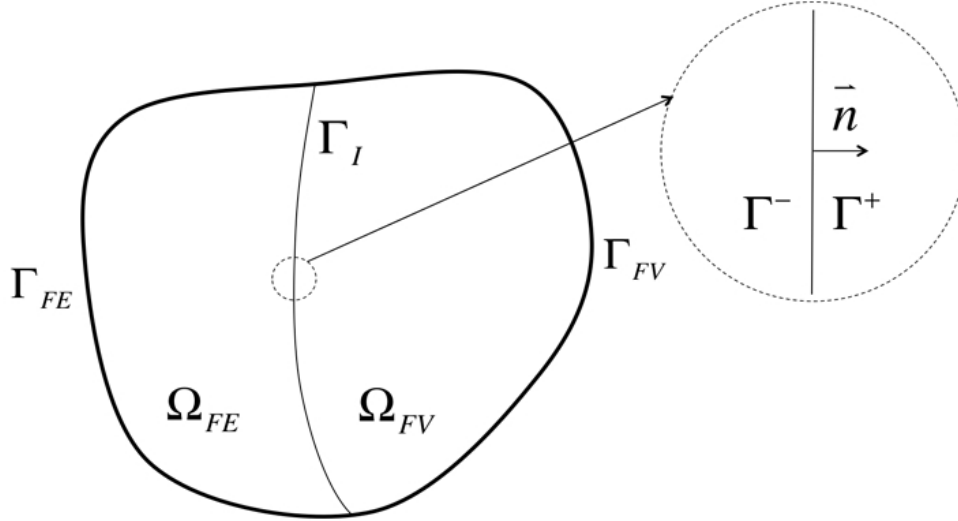


Figure 1: Coupling interface separating either physical domains or arbitrary internal domains discretized with different numerical schemes

### Method of solution

We use two different approaches to discretizing equations (1) and (2), finite volume and finite element methods namely. We consider the computational domain consisting of two parts as depicted in Figure (1). The part denoted by  $\Omega_{FE}$  together with the boundary  $\Gamma_{FE}$  constitute the domain of the discretization by finite elements. Similarly, the part denoted by  $\Omega_{FV}$  together with the boundary  $\Gamma_{FV}$  is discretized by the cell-centered finite volume method. Interface  $\Gamma_I$  between two discretization domains is an internal surface that is used to transfer the information between two discretization methods.

Linear 4-node quadratic finite elements with  $C^1$  continuity are used for the finite element basis while the cell centered finite volume discretization with the  $C^0$  interpolation basis is used for the finite volume part of the domain. Finite element and finite volume discretizations are both nominally of the second order of accuracy. In the case of finite elements, the second order of accuracy is achieved by using the shape functions with  $C^1$  continuity. In the case of the finite volume discretization, the second order of accuracy is achieved by evaluating the weak form of the energy equation at the geometric center of the finite volume cell.

Weak form of the energy equation is used for both finite volume and finite element discretizations. The weak form of the energy equation for the solid is obtained by multiplying both sides of Eq. (1) by a test function  $\xi$  and integrating over the whole domain  $\Omega$

$$\int_{\Omega} k \partial_i^2 T \xi d\Omega = \int_{\Omega} f \xi d\Omega. \quad (7)$$

The cell centered finite volume method is obtained when the test function  $\xi$  is chosen to be constant over the finite volume cell and equal to unity i.e.,  $\xi \in C^0$ . In that case Eq (7) in the domain  $\Omega_{FV}$  is replaced by the conservation law of

energy in the solid

$$\int_{\Gamma_{FV}} k \partial_i T n_i d\Gamma_{FV} = \int_{\Omega_{FV}} f d\Omega_{FV}. \quad (8)$$

Vector of local surface normal is denoted by  $n_i$  and we used Gauss-Ostrogradsky's theorem to obtain the weak form of the conservation law of energy in the solid.

The finite element method for the case of the energy transport in the solid is obtained through integration by parts Eq. (7)

$$\int_{\Omega_{FE}} k \partial_i T \partial_i \xi d\Omega_{FE} = \int_{\Omega_{FE}} f \xi d\Omega_{FE} + \int_{\Gamma_{FE}} \xi k \partial_i T d\Gamma_{FE}. \quad (9)$$

In this work we use test functions and shape functions that are linear i.e.,  $\xi \in C^1$ .

Standard finite volume and finite element discretizations are obtained when integrals in equations (8) and (9) are replaced by the numerical integration performed in face and cell centers for finite volume and nodes for finite element methods. In addition, partial differentials in Eq. (8) are replaced by finite differences for a given neighboring finite volume cells, leading to the following expression for the surface integral

$$\int_{\Gamma_{FV}} k \partial_i T n_i d\Gamma_{FV} \approx k \frac{T_{cR} - T_{cL}}{d} A_{\Gamma_{FV}} + G. \quad (10)$$

Symbol  $d$  is used to represent the distance between two cell centers while  $G$  is the non-orthogonal contribution  $A_{\Gamma_{FV}}$  is the surface area of the interface between finite element and finite volume cell. In a general case the direction between two cell centers does not coincide with any of the cartesian directions and the non-orthogonal contribution of the partial derivative, denoted by  $g_i$ , has to be added to the expression in Eq. (10). However, here we assume that this contribution was lumped into the right-hand-side of Eq. (8) without any loss of generality. Right hand side of Eq. (8) is evaluated in the cell center of the finite volume cell.

The finite element approximation of Eq. (9) is obtained by assuming that the test and shape functions belong to the same space thus yielding the following expression for the left-hand-side of Eq. (9)

$$k \partial_i T \partial_i \xi = \sum_j k T_{nj} (\partial_i S_j)^2. \quad (11)$$

Symbol  $S$  represents a shape function that is in this case chosen to be linear Lagrange polynomial and we have used the linear representation of the temperature field over the finite element  $T = \sum_j S_j T_{nj}$ . Right-hand-side of Eq. (9) is evaluated by computing the contributions of each volumetric integral to the nodes of the given finite element.

Standard discretization practices are easily applied throughout respective discretization domains. However, in order to complete the discretization of the whole domain  $\Omega = \Omega_{FE} \cup \Omega_{FV}$  the question of the discretization along the interface  $\Gamma_I$  must be addressed.

Figure (2) depicts a detail the interface  $\Gamma_I$  where we can see finite element being a neighbor to a finite volume cell connected but the common face  $\Gamma_I^e$ . We use the consistency condition in Eq. (5) to complete the discretization for finite element-finite volume couple. The consistency condition states that the in order to conserve the energy in the domain then the energy flux leaving one domain

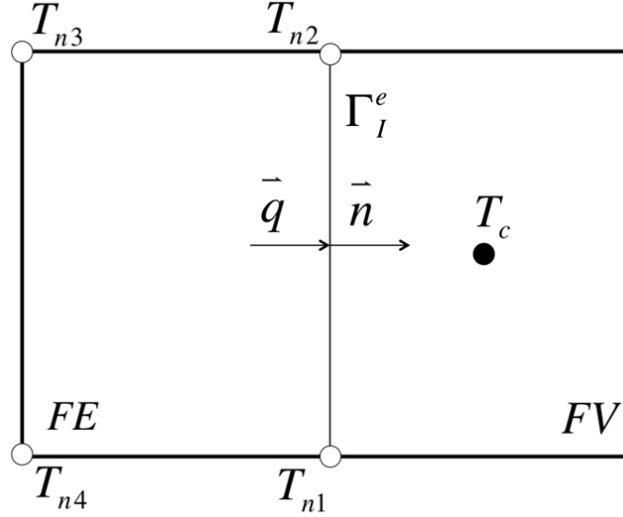


Figure 2: Finite element-finite volume pair sharing a common face that is a part of the  $\Gamma_I$  interface. Black circle represents the finite volume cell center while open circles represent the nodes of finite element.

along the interface  $\Gamma_I$  must be equal to the energy entering the other domain across the same interface  $\Gamma_I$ . Therefore, in order to complete the discretization along the  $\Gamma_I$  interface we must find the expression for the flux across every face connecting finite elements and finite volumes  $\Gamma_I^e$ .

This expression is obtained by approximating the flux across the interface by the finite difference between finite volume and finite element centroid values

$$\int_{\Gamma_I^e} k \partial_i T n_i d\Gamma_I^e \approx k_f \frac{T_c - \sum S_i (T_{ni})_c}{d} A_{\Gamma_I^e} + G. \quad (12)$$

Expression  $\sum (S_i T_{ni})_c$  is the finite element interpolation of the temperature field in the centroid of the finite element,  $T_c$  is the value of the temperature field at the finite volume centroid,  $A_{\Gamma_I^e}$  is the surface area between finite volume and finite element and  $G$  is the non-orthogonal contribution due to misalignment of the face normal and the direction defined by the cell distance. Since Eq. (5) requires that the fluxes on both sides of the interface  $\Gamma_I$  are equal, Eq. (12) is used to complete the discretization on along the interface. It should be noted that the same expression Eq. (12) is used to compute the coupling matrix entries for both finite element and finite volume discretizations. The resulting matrix structure is shown in Figure (3).

The off-diagonal entries in the block matrix  $A_{\Gamma_{C^1}^{C^0}}$  and  $A_{\Gamma_{C^0}^{C^1}}$  are obtained by computing the contributions to the finite element and finite volume system of equations using Eq. (12). The consistency condition given by Eq. (5) produces the contributions to finite element and finite volume side of the interface so that the finite volume side ( $\Gamma^+$ ) has the following entries:

$$A_{FV} = \left[ k_f \frac{T_c}{d} A_{\Gamma_I^e} \right]^{-}, \quad (13)$$

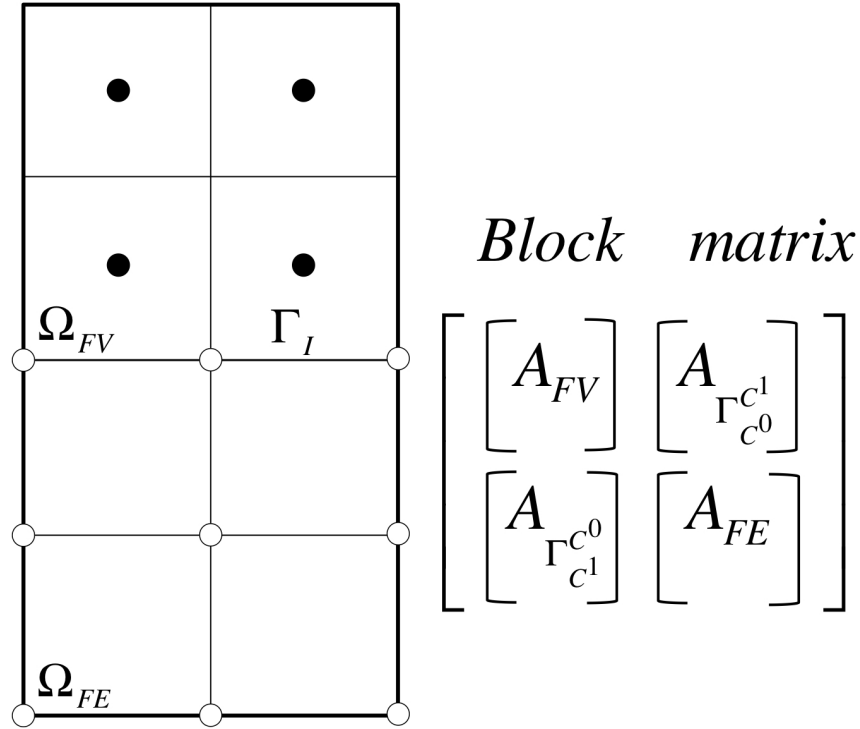


Figure 3: Matrix structure for the coupled finite element-finite volume discretization.

$$A_{\Gamma_{c^0}^{c^1}} = \left[ -k_f \frac{(\sum S_i T_{ni})_c}{d} A_{\Gamma_I^e} \right]^- . \quad (14)$$

The symbol  $k_f$  is the surface heat conduction coefficient evaluated by using the harmonic averaging procedure. It should be noted that the expression in Eq. (13) is the contribution to the the block matrix  $A_{FV}$  at the interface while Eq. (14) is the contribution to the coupling block matrix  $A_{\Gamma_{c^1}^{c^0}}$ . The non-orthogonal contribution  $G$  is added to the right hand side of the finite volume block. Similarly, finite element discretization produces the following contributions:

$$A_{FE} = \left[ -\frac{k_f}{2} \frac{(\sum S_i T_{ni})_c}{d} A_{\Gamma_I^e} \right]^+ , \quad (15)$$

$$A_{\Gamma_{c^1}^{c^0}} = \left[ \frac{k_f}{2} \frac{T_c}{d} A_{\Gamma_I^e} \right]^+ . \quad (16)$$

The finite element discretization of the element produces the contribution to the diagonal block matrix  $A_{FE}$  as well as to the off-diagonal coupling block  $A_{\Gamma_{c^0}^{c^1}}$ . Since the consistency condition specifies the flux for the finite element face of the interface  $\Gamma_I^e$ , the nodes that reside on the interface receive the half of the flux due to linear interpolation as indicated in Eq. (15) and (16). The non-orthogonal correction is similarly interpolated and added as the contribution to the coefficients  $A_{\Gamma_{c^0}^{c^1}}$ .

Eq. (12) implicitly uses the fact that the flux across the face  $\Gamma_f^e$  is evaluated using the pointwise values of the temperature field instead of cell averages. This transition between pointwise and cell averages is possible due to the fact that the pointwise values coincide with the cell averages if the function is evaluate at the cell centroid for the finite volumes of second order of accuracy. Therefore, even though the test functions for finite element and finite volume methods belong to different spaces of continuous functions, the transition from  $C^1$  to  $C^0$  functions is enabled throughout the use of pointwise values in finite volume method thus matching the desired interpolation continuity requirements. Clearly, the higher order transition, for example  $C^2$  to  $C^0$  will require the reconstruction of the pointwise values in the finite volume domain that will recover the desired continuity requirements. This is the subject of the ongoing research work.

Once the off-diagonal coupling coefficient have been computed, the diagonal entries in the block matrix  $A_{FV}$  and  $A_{FE}$  are obtained by applying the standard discretization practices applicable to finite volume and finite element discretization schemes. The resulting system represents a fully coupled system that is solved by the linear algebra suite. Each block in the coupled matrix  $A$  is represented through a sparse matrix structure utilizing the compressed-row format to save the memory. However, it should be noted that the implementation of the linear algebra library allows for storage of dense blocks as well. The linear algebra library implements Krylov subspace algorithms including conjugate gradient (CG), bi-conjugate gradient stabilized (BCGSTB), transpose free qausi-minimum residual (TFQMR). In addition, algebraic multigrid solver (AMG) based on aggregation of neighbors has been implemented to operate on the coupled block matrix. Consistent restriction and prolongation operators have been implemented so that the block matrix can be consistently defined on the progression of coarse levels.

## Results and discussions

In this section we present the results for the two cases of the coupled systems. The first case is represented by the energy transport in the solid with the boundary conditions as depicted in Fig. (4). The upper half of the domain was discretizaed by the cell-centered finite volume method while the lower half of the domain was discretizaed by the linear quadratic finite elements. Since there is no jump in material properties and given the adiabatic conditions on the sides, the solution to this problem is a linear variation of the temperature between 300  $K$  and 400  $K$  in the vertical direction. The mesh is fully orthogonal and there was no non-orthogonal contribution in the discretizaed system of equations. The computed temperature profiles at two locations are shown in Fig. (5) and (6). The agreement between analytical and the b=numerical solution if excellent as the analytical behavior of the temperature is recovered. It should be observed that the temperature was plotted at the nodes of the finite element plot in Fig. (5). For the finite volume portion of the domain the temperature was plotted in the cell centers as shown in (6). This explains the offset of the plot in the  $x$ -direction.

The second case considered in this work is the coupled energy transport between fluid and solid domains as depicted in Fig. (7). Boundary conditions are given in Fig. (7). The fluid portion of the domain was discretized with the cell-centered finite volume schemes while the solid part was discretized using the linear

quadratic finite elements. Since the problem involves the fluid flow, the boundary layer formed at the surface of the solid wall is largely responsible for the heat transfer to and from the fluid. For the given inlet velocity and length of the domain, the Reynolds number based on the length is very low  $Re = 400$ , well below the transitional Reynolds number for the flat plate (500,000). Therefore, the flow over the surface of the solid is laminar. The computed temperature profiles and the comparison to the theoretical one is given in Fig. (8) and (9).

## Conclusion

We introduced a novel algorithm for finite element-finite volume coupling that is based on the native cell-centered finite volume and linear quadratic finite element discretization methods in their respective parts of the computational domain. The coupling coefficients required for the implicit representation of the coupled matrix were defined. The consistency condition that is based on the flux conservation between finite element and finite volume discretizations was defined and used to define the coupling coefficients. Block matrix linear solver based on BCGSTAB was used to solve the coupled solver in one matrix thus producing the implicitly coupled solution. Two examples of the application of the newly defined coupled method were provided. The first example was the energy transport in the solid body discretized in part by finite volume and the other part by finite element technique. The solution was compared to the analytical solution and the excellent agreement was achieved. The second case that was considered consisted of the energy transport between fluid and solid domains. The fluid domain was discretized but the cell-centered finite volume while the solid part was discretized by the linear quadratic finite element scheme. The agreement between analytical and numerical solution, in all cases, is very good.

## References

- [1] B.R. Baliga and S.V. Patankar. A new finite-element formulation for convection diffusion problems. *Numerical Heat Transfer*, 3:393–409, 1980.
- [2] B. Gadeschi L. Schneiders M. Meinke and W. Schroeder. A numerical method for multi physics simulations based on hierarchical cartesian grids. *Journal of Fluid Science and Technology*, 10, 2015.
- [3] M. Sardella. On a coupled finite element-finite volume method for convection diffusion problems. *IMA Journal of Numerical Analysis*, 20:281–301, 2000.
- [4] S. Geiger S. Roberts S.K. Matthai C. Soppou and A. Burri. Combining finite element and finite volume methods for efficient multiphase flow simulations in highly heterogeneous and structurally complex geologic media. *Geofluids*, 44:284–299, 2004.
- [5] R.D. Lazarov J.E. Pasciak P.S. Vassilevsk. Coupling mixed and finite volume discretizations of convection-diffusion-reaction equations on non-matching grids. *Symposium on Finite Volume Methods for Complex Applications*, 1999.

- [6] J. Vierendeels. Strong coupling of partitioned fluid-structure interaction problems with reduced order models. *European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2006*, 2006.
- [7] S. Sicklinger V. Belsky B. Engelmann H. Elmqvist H. Olsson R. Wuchner and K.-U. Bletzinger. Interface jacobian-based co-simulation. *International Journal for Numerical Methods in Engineering*, 98:418–444, 2014.

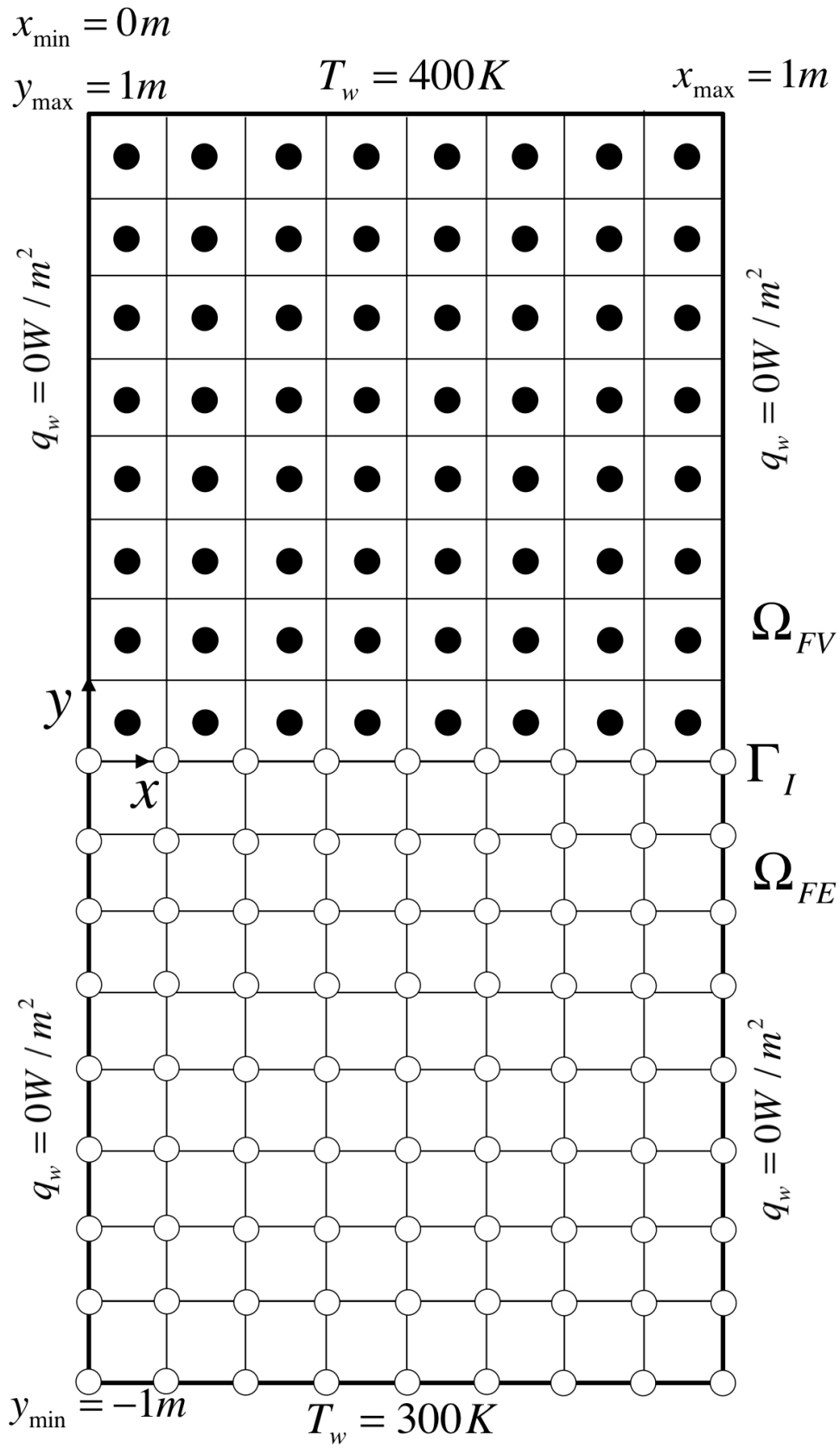


Figure 4: Computational domain for energy transport in a solid.



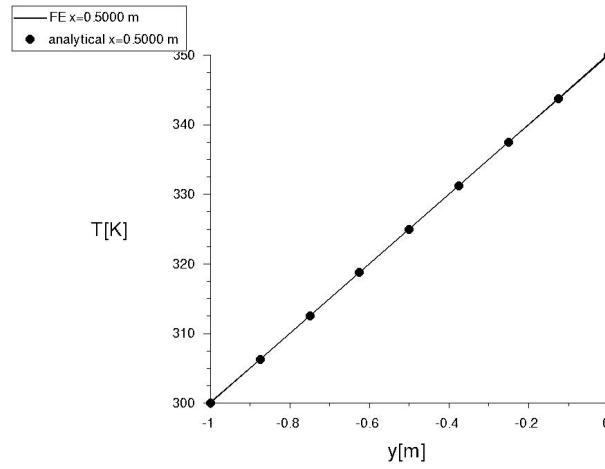


Figure 5: Temperature profile in the vertical direction for the case of the energy transport in the solid for the finite element portion of the domain at the location  $x = 0.5 m$ . Solid line represents the numerical whereas dots represent the analytical solution.

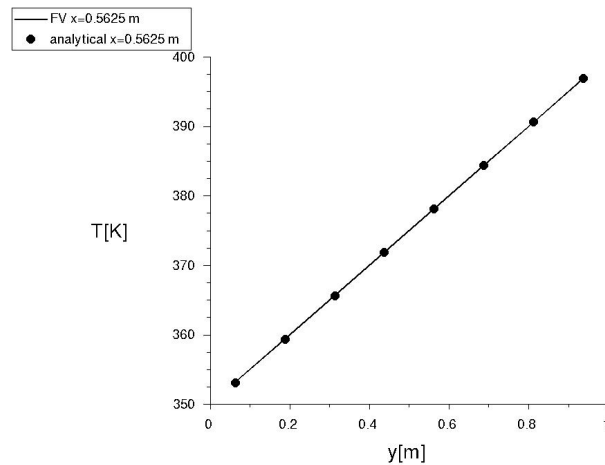


Figure 6: Temperature profile in the vertical direction for the case of the energy transport in the solid for the finite volume portion of the domain at the location  $x = 0.625 m$ . Solid line represents the numerical whereas dots represent the analytical solution.

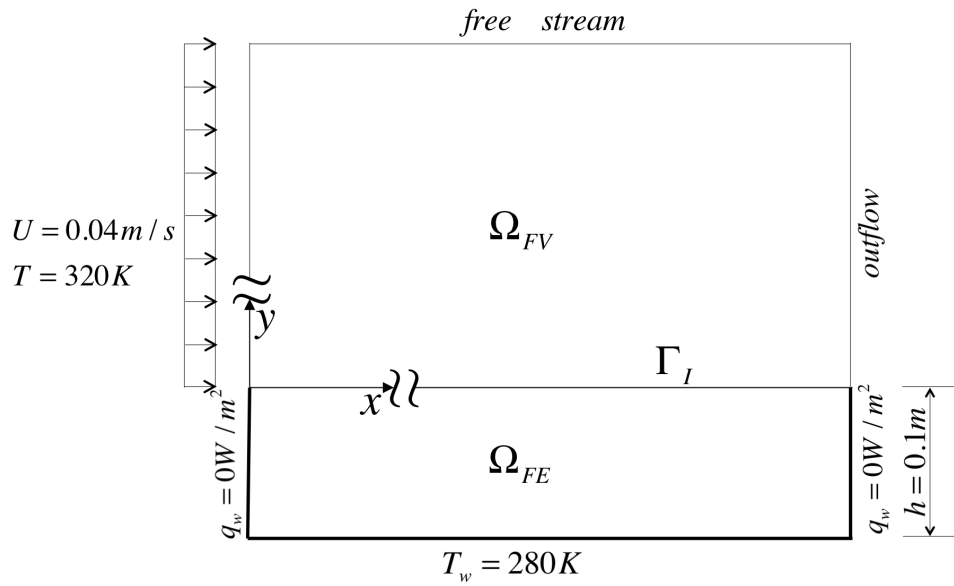


Figure 7: Computational domain for the fluid-solid energy transport.

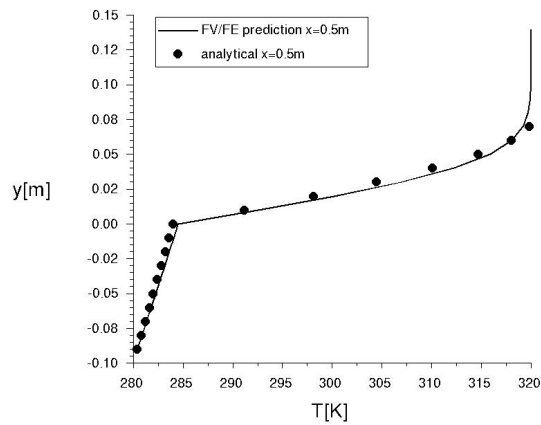


Figure 8: Temperature profile in the vertical direction for the case of the energy transport in the fluid-solid system for the finite element portion of the domain at the location  $x = 0.5m$ . Solid line represents the numerical whereas dots represent the analytical solution.

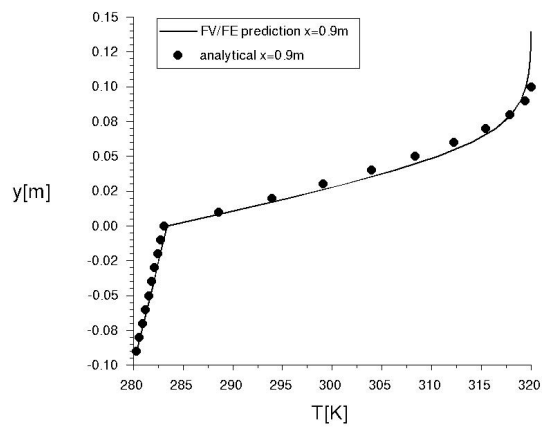


Figure 9: Temperature profile in the vertical direction for the case of the energy transport in the solid for the finite volume portion of the domain at the location  $x = 0.625 m$ . Solid line represents the numerical whereas dots represent the analytical solution.

# Finite element modelling of deformation behavior in incremental sheet forming of aluminum alloy

†,\*C.P. Jiang<sup>1</sup> and T.H. Huang<sup>2</sup>

<sup>1</sup>Department of Power Mechanical Engineering, National Formosa University, Taiwan.

<sup>2</sup> Institute of Mechanical and Electro Mechanical Engineering, National Formosa University, Taiwan.

\*Presenting author: cpjiang@nfu.edu.tw

†Corresponding author: cpjiang@nfu.edu.tw

## Abstract

In this paper, the finite element method (FEM) is used to study the incremental sheet forming process of pyramidal shape. The material used is aluminium alloy 5052. The tool, a hemispherical ball-head with a diameter ( $d = 4$  mm) made of HSS tool steel, is used to press down on the sheet metal causing locally plastic deformation. The comparison between spiral tool path, spiral-step tool path and z-level tool path is carried out. Moreover, the final thickness distribution is investigated. The results indicate that the minimal thickness can be found on the corner of wall angle in SPIF process. Under the same step over, spiral-step tool path can obtain the deepest depth for pyramidal shape. The maximum formability for successful forming of the pyramidal shape with depth 60mm is wall angles  $65^\circ$ .

**Keywords:** Finite element, Incremental sheet forming, Aluminum

## Introduction

Single Point Incremental Forming (SPIF) involves a local and progressive pressing out of the desired shape on a clamped sheet by a round-headed forming tool which follows a continuous path. SPIF offers full flexibility because the use of dedicated tooling isn't required. However, one disadvantage is that the operation time is high. Furthermore, SPIF may still be of use for low volume series and in combination with other forming processes to produce part details.

Thickness reduction is a phenomenon common to any sheet metal forming process that leads to the occurrence of crack and fracture of the resulting shape. In the numerical study of SPIF, the deformation behaviour such as thickness distribution and the mechanical properties of sheet metal on a pyramidal shape was investigated [Dejardin et al. (2010); Emmems et al. (2010)]. These reports present the thickness distribution in the final product is not uniform. Since tool path defines the component geometry, the different tool path with different incremental steps may affect the formability. Spiral tool path and z-level tool path are proposed to test formability for truncated cones [Liu et al. (2013)]. Experimental results show that inclination angle and incremental step of the tool path affect the deformability.

In this paper, the finite element method is used to study the single point incremental forming process of pyramidal shape. The effect of spiral tool path, spiral-step tool path and z-level tool path on pyramidal depth and minimal thickness is investigated.

## Materials and Method

### *Forming Strategy*

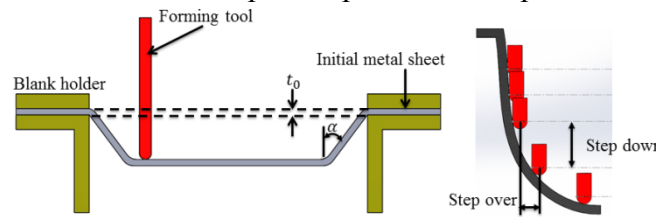
The proposed part is a quadrangular pyramid. The initial dimension of Al5052 is  $86 \times 86 \times 0.6$  mm. The base dimension of pyramidal shape is  $60 \times 60$  mm and the height is 60 mm. The maximal inclination angle is a variation as indicator of forming limit. The composition of this material is shown in table 1.

**Table 1. Composition of Al5052.**

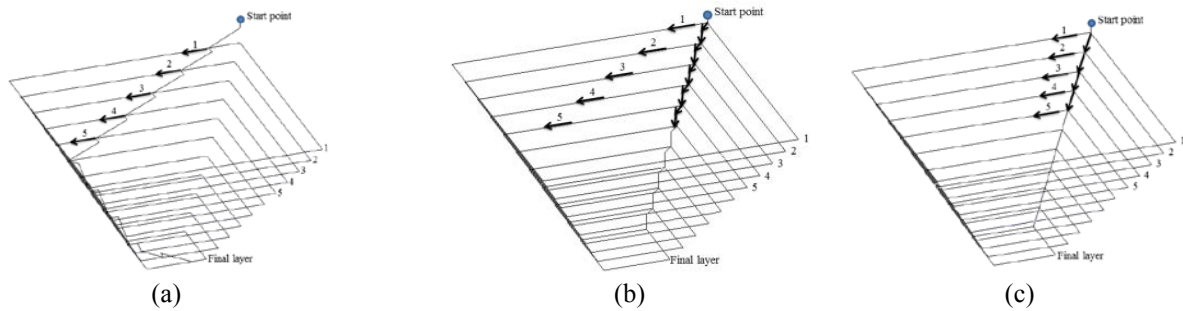
Al	Mg	Cr	Cu	Fe	Mn	Si	Zn
Balance	2.5	0.3	0.05	0.25	0.05	0.05	0.05

*Tool path*

Figure 1 shows the details of single point incremental forming and the illustration of incremental steps. Step over is an adjustable parameter quantity to generate the tool path. For the investigated pyramidal with the variation inclination angle, step over is used to generate three kinds of tool paths (spiral, spiral-step and z-level). The spiral-step tool path in Figure 2 (a) is continuous with an incremental descent of the tool distributed over the entire surface of a part. Wall angle can be calculated by the trigonometric relationship of step over and step down. Figure 2(b)



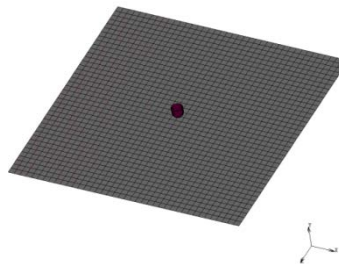
**Figure 1. Details of incremental forming**



**Figure 2. Tool paths for pyramidal shape: (a) spiral tool path, (b) spiral-step tool path and (c) z-level tool path.**

*Finite element model*

A set of forming tool and sheet used to simulate the deformation behavior with different tool path in SPIF process as shown in Figure 3. The sheet material is considered to be isotropic, homogeneous and incompressible. The sheet is meshed by shell element with four nodes and assumed deformable. Nodes on all edges of the sheet are constrained as fixed displacement and rotation during the forming process. The rigid tool shape is a cylinder with hemispherical head whose diameter is 4 mm. The friction between sheet and tool is modelled using the Coulomb friction with a friction coefficient value of 0.13.



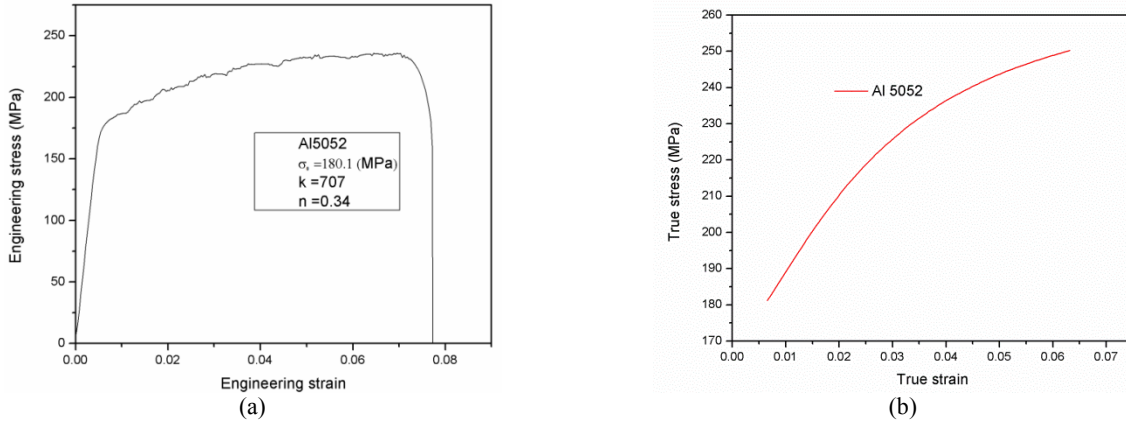
**Figure 3. Numerical model used for incremental sheet metal forming process**

*Material properties*

The material is underwent the tensile test and obtained the engineering stress-strain curve as shown in Figure 4 (a). The power law equation (1) describes the material behaviour in incremental sheet forming:

$$\sigma = k\varepsilon^n \tag{1}$$

where  $\sigma$  is the equivalent stress,  $\varepsilon$  denotes the equivalent strain,  $k$  is the strength coefficient and  $n$  is the strain-hardening exponent. Therefore, the curve of true stress-strain can be shown in Figure 4(b). The required mechanical properties of Al-5052 used in simulation are given in Table 2.



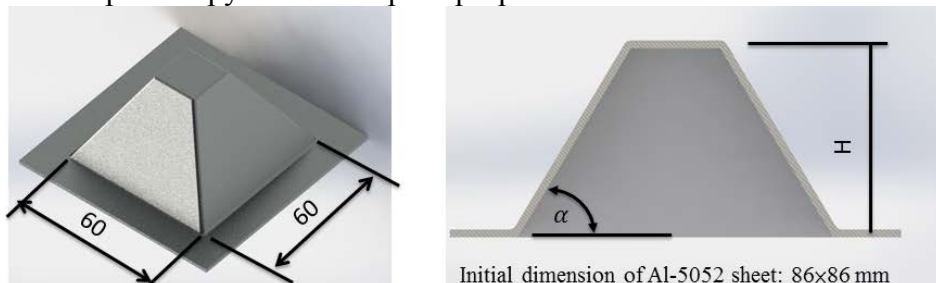
**Figure 4. Engineering stress-strain curve for 5052 aluminium alloy sheet with 0.6 mm**

**Table 2. Mechanical properties for 5052 aluminium alloy sheet.**

k	n	E (GPa)	$\sigma_0$	$\nu$
707	0.34	35.2	180.1	0.3

**Results and discussion**

In SPIF, the formability is higher than the deep drawing and stamping. Research literatures report that wall angle can be a formability indicator. In this study, the maximum wall angle is used as indicator to evaluate the formability of Al-5052 sheet with 0.6 mm thickness considering the effect of three types of tool path. A pyramidal shape is proposed as the benchmark as show in Figure 5.

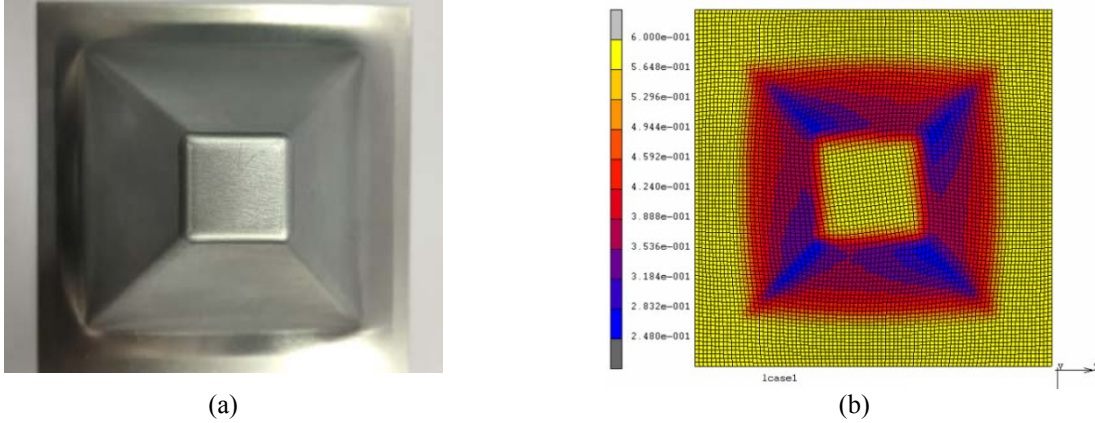


**Figure 5. Benchmark- a pyramidal shape**

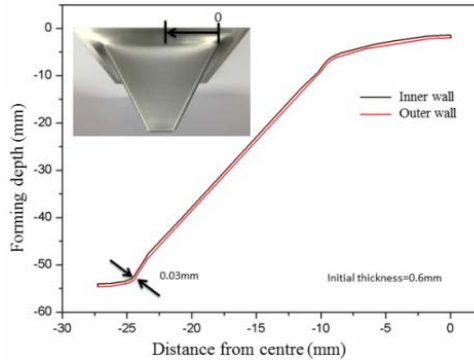
*Effect of tool path on formability*

Figure 6 (a) shows the complete pyramidal shape with base dimension of 60x60 mm and the height of 60 mm. The tool movement is spiral-step as show in Figure 2(b) and step-down is 0.5 mm. The

final wall angle is  $69^\circ$ . The simulation result presents the same shape as shown Figure 6(b). The wall angle is the same and no fracture occurs after the tool completes the movement. Hence, the thickness of wall is compared as shown in Figure 7. The minimal thickness which is 0.03 mm can be found on the corner of wall angle.



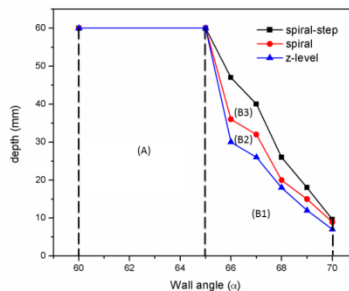
**Figure 6. The complete benchmark shape: experimental (a) and simulation (b) results**



**Figure 7. The comparison of wall thickness between experiment and simulation**

*Forming limit diagram*

In the first test, the step over was set to be a constant value 0.5 mm for all tool paths. A series of pyramidal shapes were simulated with wall angle  $60^\circ$ ,  $65^\circ$ ,  $66^\circ$ ,  $67^\circ$ ,  $68^\circ$ ,  $69^\circ$  and  $70^\circ$  to evaluate the formability. Figure 8 presents the simulation results. Under the same step over, no noticeable different on formability between spiral, spiral-step and z-level tool paths. As can be seen, the maximum formability for successful forming of the pyramidal shape with depth 60mm is wall angles  $65^\circ$ . A safety formability region is identified in area A for taking the required pyramid height into consideration. The area B3 has larger safety formability zone for spiral-step tool path than those for others. Therefore, spiral-step tool path can obtain the deepest depth for pyramidal shape with 0.5 mm step over size.



**Figure 8. Influence of spiral, spiral-step and z-level tool paths on formability vs wall angle ( $\alpha$ ).**

## Conclusions

Deformation behaviour of Al-5052 sheet in an incremental process for various tool paths is investigated. Formability is evaluated to understand the forming mechanism. Some conclusion can be drawn as follows:

- (1) The minimal thickness can be found on the corner of wall angle in SPIF process.
- (2) Under the same step over, spiral-step tool path can obtain the deepest depth for pyramidal shape.
- (3) The maximum formability for successful forming of the pyramidal shape with depth 60mm is wall angles 65°.

## Acknowledgements

Funding from the National Science Council of Taiwan for this project, NSC101-2628-E-150-002-MY3, is gratefully acknowledged.

## References

- Dejardin, S., Thibaud, S., Gelin, J. C., Michel, G. (2010) Experimental investigations and numerical analysis for improving knowledge of incremental sheet forming process for sheet metal parts, *Journal of Materials Processing Technology* **210**, 537–554.
- Emmens, W.C. and Sebastiani, G. and Boogaard, A.H. van den (2010) The technology of Incremental Sheet Forming - a brief review of the history, *Journal of Materials Processing Technology* **210**, 981–997.
- Liu, Z., Li, Y., Meehan, P. A. (2013) Experimental investigation of mechanical properties, formability and force measurement for AA7075-O aluminum alloy sheets formed by incremental forming, *International Journal of Precision Engineering and Manufacturing* **14**, 1891–1899.



## A ghost particle-based coupling approach for combined finite-discrete element method

\*Hu Chen<sup>1</sup>, †Y.X. Zhang<sup>1</sup> and Mengyan Zang<sup>2</sup>

<sup>1</sup>School of Engineering and Information Technology, The University of New South Wales, ADFA, Canberra, ACT 2600, Australia

<sup>2</sup>School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, China

\*Presenting author: hu.chen@student.adfa.edu.au

†Corresponding author: y.zhang@adfa.edu.au

### Abstract

Combined finite-discrete element method takes the advantages of both the finite element (FE) method and the discrete element (DE) method, but a coupling approach is required for effective combination of the two methods. In this paper, a coupling approach is proposed by using ghost particles. The entire domain is decomposed into a FE region and a DE region which are connected by an interface, and ghost particles are constructed inside the boundary of FEs by interpolation and used to connect the boundary of DEs. This interface connection method effectively couples FEs with DEs. A numerical example is computed to demonstrate the effectiveness of the proposed approach.

**Keywords:** Ghost particle, Coupling approach, Combined finite-discrete element method, Multiscale modelling.

### Introduction

The conventional finite element (FE) method has some limitations in numerical modelling of brittle fracture problems since it is based on the continuum mechanics, and special treatments have to be employed, e.g., remeshing strategies [Mediavilla et al. (2006); Belytschko and Black (1999)] or extended finite element method [Belytschko et al. (2001); Moës and Belytschko (2002)]. In contrast, the discrete element (DE) method is able to capture the discontinuous nature of materials, but it is not as computational efficient as the FE method. Combined finite-discrete element method, which combines the two methods, has been developed so as to use the advantages while avoiding the drawbacks of the two numerical methods [Munjiza (2004)].

Essentially, the combination of the two methods can be achieved by the interaction of either contact or coupling. The contact interaction is more suitable for contact occurring between different pieces of an object when fragmentations form [Munjiza et al. (1995)] and between two or more different media [Onate and Rojek (2004); Chen et al. (2015)]. The coupling interaction is more applicable for the connection between different subdomains of an entire domain, which is decomposed for efficient computation via the so-called multiscale modeling. The coupling approaches used along with the combined finite-discrete element method are generally classified into overlapping coupling [Wellmann and Wriggers (2012); Li et al. (2015); Jebahi et al. (2013); Tu et al. (2014)] and interface coupling [Christian et al. (2014); Lei and Zang

(2010)]. For the overlapping coupling compatibility condition is required to be satisfied in a bridging domain [Xiao and Belytschko (2004)], where the total energy is a linear mixing of energies from both FEs and DEs, while only a simple compatibility condition at the interface is required to be satisfied in the interface coupling.

In this paper, a ghost particle-based approach is proposed to couple the continua with discrete elements for the combined finite-discrete element method. The computational domain is decomposed into a FE region and a DE region, which are distinguished by an interface. The two regions are connected by the interface interaction between the boundary DEs and ghost particles which are constructed at the boundary FEs. This approach is conceived from the extension of atomistic-continuum coupling [Kohlhoff et al. (1991); Shilkrot et al. (2002)], in which ghost atoms are used to coincide with the FE nodes in the transition region (harsh mesh requirement), and the number of coincidence layers is dependent on the cut-off radius of atomistic potential. For the present approach, however, only one layer of ghost particles is required as the interaction of DEs only occurs between adjacent particles. Furthermore, to overcome the difficulties of remeshing, the strictly coincident compatibility condition is relaxed by introducing the ghost particles, whose position is determined from the interpolation of nodal coordinates of FEs.

To apply this coupling approach to model fracture problems, an important and essential step is to validate its effectiveness prior to the occurrence of fractures. This paper is therefore confined to the development and validation of this new coupling approach, without exploration and investigation into brittle fractures. The governing equations and DE model will be briefly introduced in next section. Afterwards, the development of the proposed coupling approach is presented in detail, followed by the numerical validation. Finally, conclusions are drawn.

### General formulations

The governing equations for FEs and DEs are briefly described first, followed by the introduction of a DE connective model, which is used to discretize the area of interest.

#### *Governing equations*

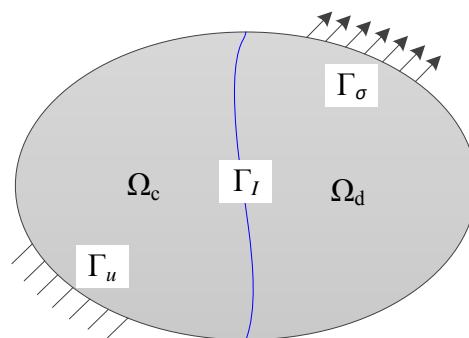


Figure 1. Domain decomposition into a continuum region and a discrete region

A computational domain, subject to prescribed displacement at the Dirichlet boundary  $\Gamma_u$  and prescribed tractions at the Neumann boundary  $\Gamma_\sigma$ , is divided into a continuum subdomain  $\Omega_c$  and a discrete subdomain  $\Omega_d$  as shown in Figure 1. The interface between these two subdomains is defined as  $\Gamma_I = \partial\Omega_c \cap \partial\Omega_d$ .

For the continuum subdomain, the FEs are used for spatial discretization and the governing equation of the FE method is given as

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{K}\mathbf{U} = \mathbf{F} + \mathbf{F}^{\text{cou}} \quad (1)$$

where  $\mathbf{M}$  and  $\mathbf{K}$  are mass and stiffness matrices, respectively;  $\mathbf{U}$  is the displacement vector;  $\mathbf{F}$  is the resultant force vector consisting of external tractions and/or damping force;  $\mathbf{F}^{\text{cou}}$  is the coupling force resulting from the interaction with the discrete subdomain.

The discrete subdomain is discretized by DEs and the equations of motion for each DE particle are expressed as follows

$$\begin{cases} m\dot{\mathbf{v}} = \mathbf{f} + \mathbf{f}^{\text{cou}} \\ I\dot{\boldsymbol{\omega}} = \mathbf{t} + \mathbf{r} \times \mathbf{f}^{\text{cou}} \end{cases} \quad (2)$$

where  $m$  is the mass of the particle;  $\mathbf{v}$  is the centroid velocity of the particle in global coordinate frame;  $\mathbf{f}$  is the resultant force applied to the particle due to (a) external traction, (b) body force, (c) internal force when the DE connective model is used and (d) damping force (linearly proportional to the velocity) to dissipate kinetic energy for computational purpose;  $\mathbf{f}^{\text{cou}}$  is the coupling force exerted on the particle resulting from the interaction with the continuum subdomain;  $I$  is the moment of inertia of the particle;  $\boldsymbol{\omega}$  is the angular velocity of the particle in the local coordinate frame which is embedded into the centroid of the particle;  $\mathbf{t}$  is the resultant moment applied to the particle due to the aforementioned forces  $\mathbf{f}$ ;  $\mathbf{r}$  is the moment arm of the coupling force  $\mathbf{f}^{\text{cou}}$ .

#### Discrete element model

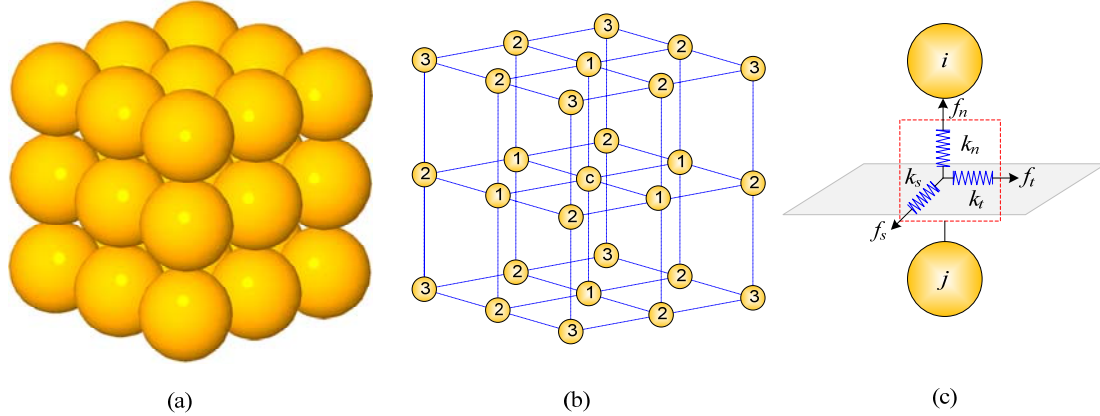


Figure 2. The discrete element connective model. (a) Cubic arrangement of the 27 spherical particles. (b) Linkage relation. (c) Virtual springs connect adjacent particles.

The DE connective model as shown in Figure 2 is employed. This model is in a cubic arrangement (Figure 2a), which can produce a neat surface desirable for domain decomposition. Each particle in this model is connected to its neighboring 26 particles, which are categorized into three groups according to their distances to the central one as depicted by different numbers in Figure 2b. The interaction force  $\mathbf{f}$  between any two adjacent particles is calculated based on their relative displacement  $\mathbf{d}$  and spring stiffness  $k$  (see Figure 2c) as follows

$$\mathbf{f} = k\mathbf{d} \quad (3)$$

where force  $\mathbf{f} = (f_n, f_s, f_t)$ , and  $\mathbf{d} = (d_n, d_s, d_t)$ .

Inside each pair as shown in Figure 2c, there are an orthogonal ( $k_n$ ) and two tangential ( $k_s$  and  $k_t$ ) linear springs invisibly connecting them. Their stiffness are determined based on the energy equivalence between the one stored in the springs and that in solid elasticity, as given by [Yu (2011)]

$$\left\{ \begin{array}{l} k_n^1 = k_n^2 = \frac{2Er}{5(1-2\nu)} \\ k_s^1 = k_t^1 = k_s^2 = k_t^2 = \frac{2Er(1-4\nu)}{5(1-2\nu)(1+\nu)} \\ k_n^3 = k_s^3 = k_t^3 = 0 \end{array} \right. \quad (4)$$

where  $E$ ,  $\nu$  and  $r$  are the Young's modulus, Poisson's ratio and radius of the DE particles.

### Coupling approach

The coupling force is explicitly expressed through the interaction between the boundary DE particles and the ghost particles, which are virtually constructed at the inside of boundary hexahedral elements. Next, the strategy of constructing ghost particles is illustrated first, followed by the determination of natural coordinates and the kinematic relations between DE boundary particles and ghost particles.

#### *Construction of ghost particles*

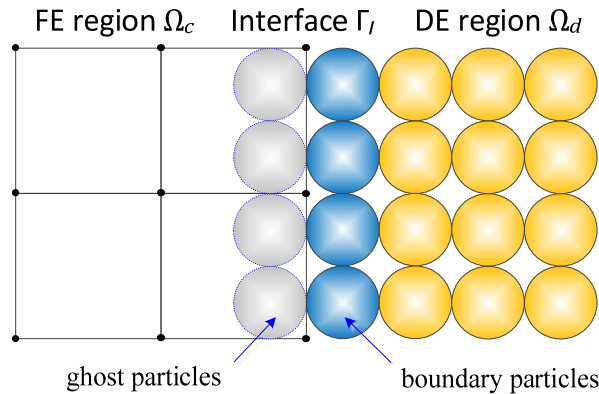


Figure 3. Schematic diagram of the coupling approach

A general rule to decompose domain is that the DE particles are used to spatially discretize the region of particular interest, such as the impact zone or area where fractures possibly occur, whereas the other region is discretized by FEs. Usually, the radius of DE particles is much smaller than the size of FEs to better characterize fracture patterns. The DE region should be large enough to cover the interest zone so as to avoid possible fractures occurring in the continuum region. Therefore, it is not necessary to generate a flexible but uncontrollable interface and a neat interface as shown in Figure 3 is then utilized for easy model preprocessing.

Ghost particles are constructed inside the FEs as shown in Figure 3, and they can be treated as a part of the DE model. Each DE particle only interacts with its adjacent particles; hence one layer of ghost particles is adequate to describe the coupling

interaction between boundary DE particles and ghost particles. The interaction mechanism between them is the same as that for a sole DE model as shown in Eqs. (3) and (4), and particularly, this interaction interconnects the FE model and the DE model. Note that these ghost particles and DE particles should have the same radius for the application of this interaction mechanism to link them together. Because only one layer of ghost particles is required and also the size of them is small, the ghost particles are merely constructed in the boundary FEs.

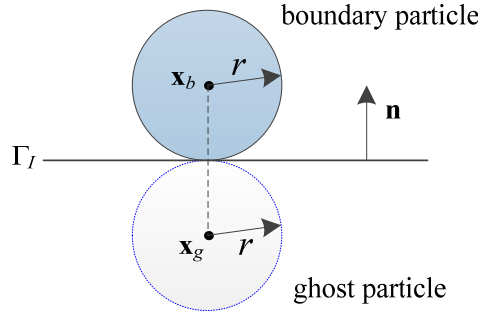


Figure 4. Position determination of the ghost particle

As a virtual part of the corresponding FE, each ghost particle has the same kinematic response as that of the FE when subject to loading or deformation. With this compatibility condition taken into account, a reasonable assumption is made that each ghost particle sticks to the same natural coordinate of the corresponding isoparametric FE. The natural coordinates will be determined next, but before doing so, the physical coordinates of ghost particles need to be acquired. As being symmetric to the boundary DE particles (as shown in Figure 4), the physical coordinates of ghost particles at initial configuration can be immediately determined by

$$\mathbf{x}_g = \mathbf{x}_b - 2r\mathbf{n} \tag{5}$$

where  $\mathbf{x}_g$  and  $\mathbf{x}_b$  are physical coordinates of boundary particles and ghost particles;  $\mathbf{n}$  is the normal vector of interface.

*Determination of natural coordinates*

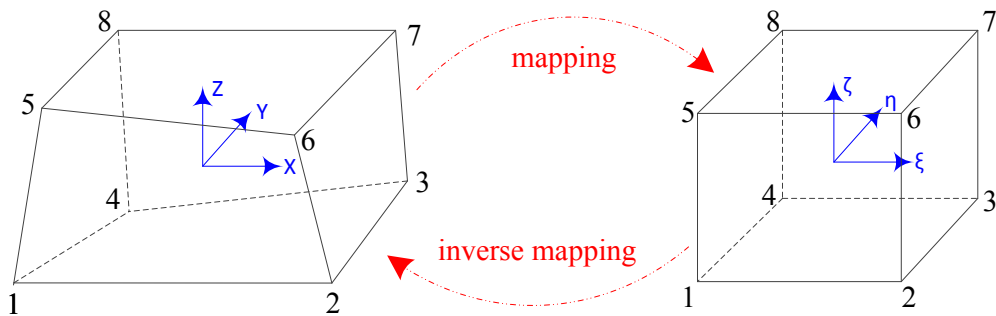


Figure 5. Mapping and inverse mapping between physical coordinates and natural coordinates

For any ghost particles, their physical coordinates can be obtained from Eq. (5), thus the purpose here is to determine corresponding unknown natural coordinates.

For a given point  $\mathbf{x}$  located inside a hexahedron element, the physical coordinate  $(x, y, z)$  of that point can be expressed by the interpolation from nodal coordinates and their associated shape functions  $N(\xi, \eta, \zeta)$  as follows

$$\mathbf{x} = \mathbf{x}(\xi, \eta, \zeta) = \sum_{I=1}^8 N_I(\xi, \eta, \zeta) \mathbf{x}_I \quad (6)$$

The shape functions are specifically given

$$N_I = \frac{1}{8} (1 + \xi_I \xi) (1 + \eta_I \eta) (1 + \zeta_I \zeta) \quad (7)$$

where the natural coordinate for each node  $I$  is shown in Table 1.

Table 1. Natural coordinate for each node,  $\xi_I = (\xi_I, \eta_I, \zeta_I)$

$\xi_1 = (-1, -1, -1)$	$\xi_2 = (+1, -1, -1)$
$\xi_3 = (+1, +1, -1)$	$\xi_4 = (-1, +1, -1)$
$\xi_5 = (-1, -1, +1)$	$\xi_6 = (+1, -1, +1)$
$\xi_7 = (+1, +1, +1)$	$\xi_8 = (-1, +1, +1)$

It is worth noting that Eq. (6) indicates the mapping from physical coordinates to natural coordinates (shown in Figure 5); however, its inverse mapping from natural coordinates to physical coordinates, i.e.,  $\xi = \xi(x, y, z)$ , is generally impossible to be explicitly expressed [Belytschko et al. (2013)]. Note that here we use position vectors to denote specific coordinates, i.e.,  $\mathbf{x} = (x, y, z)$  and  $\xi = (\xi, \eta, \zeta)$ . Therefore, to specifically determine natural coordinates, an alternative numerical approach [Qian et al. (1998)] based on Taylor series is employed.

The position vector of physical coordinates as seen in Eq. (6) is expanded around  $\mathbf{x}_o = (x_o, y_o, z_o)$  with a linear approximation

$$\mathbf{x} \approx \mathbf{x}_o + \frac{\partial \mathbf{x}(\xi_o, \eta_o, \zeta_o)}{\partial \xi} (\xi - \xi_o) \quad (8)$$

where Jacobian matrix  $\mathbf{J} = \partial \mathbf{x} / \partial \xi$  is specifically given by

$$\mathbf{J} = \begin{bmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} & \frac{\partial x}{\partial \zeta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} & \frac{\partial y}{\partial \zeta} \\ \frac{\partial z}{\partial \xi} & \frac{\partial z}{\partial \eta} & \frac{\partial z}{\partial \zeta} \end{bmatrix} \quad (9)$$

and  $\xi_o$  is the corresponding natural coordinates of  $\mathbf{x}_o$ , i.e.,  $\mathbf{x}_o = \mathbf{x}(\xi_o, \eta_o, \zeta_o)$ .

Rearranging Eqs. (8) and (9), an iterative form for natural coordinates is expressed as follows

$$\xi_{k+1} \approx \xi_k + \mathbf{J}_k^{-1} (\mathbf{x} - \mathbf{x}_k) \quad (k = 0, 1, 2, \dots) \quad (10)$$

where  $\mathbf{x}_k = \mathbf{x}(\xi_k, \eta_k, \zeta_k)$ . The iteration is terminated when the  $\mathcal{L}_2$  norm is less than a user-specified tolerance  $\varepsilon$ , i.e.,

$$\|\mathbf{x} - \mathbf{x}_k\|_2 < \varepsilon \quad (11)$$

As the natural coordinates only need to be determined once at the initial configuration when hexahedron elements are not deformed yet, this numerical iteration is very efficient to achieve its convergence.

#### *Interaction between particles*

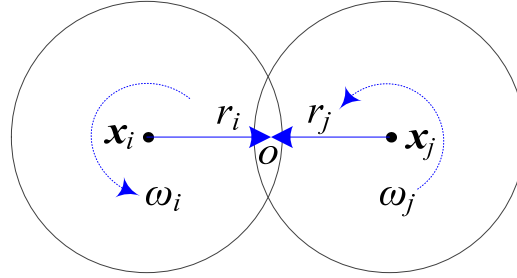


Figure 6. Kinematic relation of two adjacent particles

For any two adjacent particles ( $i$  and  $j$ ) as shown in Figure 6, their interaction force can be calculated by Eq. (3). As the stiffness of springs keeps constant, the practical interest here is to determine their relative displacement. With the central difference method employed for temporal discretization, an incremental form is used to calculate the relative displacement at  $t_{n+1}$

$$\mathbf{d}_{n+1} = \mathbf{d}_n + \Delta \mathbf{d}_{n+1} \quad (12)$$

where the incremental relative displacement  $\Delta \mathbf{d}_{n+1}$  is determined by relative velocity as follows

$$\Delta \mathbf{d}_{n+1} = \dot{\mathbf{x}}_{n+1/2}^{ij} \Delta t_{n+1} \quad (13)$$

At any time, the relative velocity is evaluated at the middle point  $o$  of the line connecting two particles' centroids, as expressed by

$$\dot{\mathbf{x}}^{ij} = \dot{\mathbf{x}}_o^j - \dot{\mathbf{x}}_o^i \quad (14)$$

in which

$$\begin{cases} \dot{\mathbf{x}}_o^i = \dot{\mathbf{x}}^i + \boldsymbol{\omega}^i \times \mathbf{r}^i \\ \dot{\mathbf{x}}_o^j = \dot{\mathbf{x}}^j + \boldsymbol{\omega}^j \times \mathbf{r}^j \end{cases} \quad (15)$$

where  $\dot{\mathbf{x}}^i$  and  $\dot{\mathbf{x}}^j$  are velocities of the particle  $i$  and  $j$ , respectively.

If one of the particles in the pair is a ghost one, its rotational velocity is not taken into account and the translational velocity is interpolated from the nodal velocities of the corresponding FE as follows

$$\dot{\mathbf{x}} = \sum_{l=1}^8 N_l(\xi, \eta, \zeta) \dot{\mathbf{x}}_l \quad (16)$$

where the natural coordinate  $\boldsymbol{\xi} = (\xi, \eta, \zeta)$  is determined using the method mentioned before.

It is worth noting that the interaction force exerted on each ghost particle is distributed into all nodes of the corresponding FE with the use of shape functions as given by Eq (17).

$$\mathbf{f}_I = -N_I(\xi, \eta, \zeta)\mathbf{f} \quad (17)$$

### Numerical validation

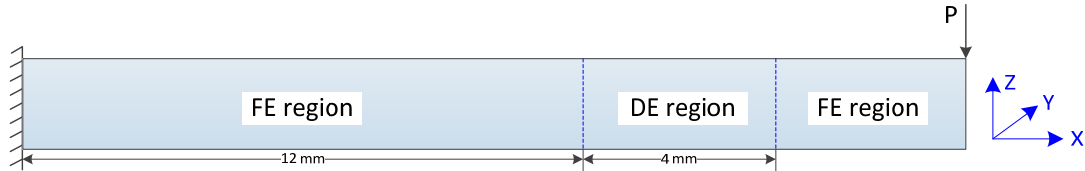


Figure 7. Diagram of the cantilever beam and the particles distribution in the coupled finite-discrete element model.

To validate the proposed numerical method, the coupling approach is implemented in an in-house code named as CDFP [Lei and Zang (2010)]. A cantilever beam with one end fixed and the other subject to constant loading ( $P = 0.4$  N) as shown in Figure 7 is modeled, and the computed deflection at the free end of the beam is compared to the analytical result.

The dimensions of this beam are: length  $L = 20$  mm, width  $W = 4$  mm and height  $H = 2$  mm. The elastic material properties are: mass density  $\rho = 0.001$  g/mm<sup>3</sup>, Young's modulus  $E = 1000$  MPa and Poisson's ratio  $\nu = 0.23$ . The coupled finite element model with the FE and DE regions as allocated in Figure 7 is generated by using different spatial discretization techniques. The mesh size for both the FE model and the coupled model is 0.5 mm and the particle radius for both the DE model and the coupled model is 0.125 mm. To acquire a static solution from the explicit central difference time integration, an appropriate damping is imposed to the nodal and particle velocities.

The deflection at the free end computed from the coupled model is 0.388 mm, which is very close to the analytical solution 0.4 mm. To further demonstrate the effectiveness of the proposed coupling model, the displacement distribution of the beam along the loading direction obtained from the FE model, the DE model and the coupled model are compared in Figure 8, where a good agreement can be observed.

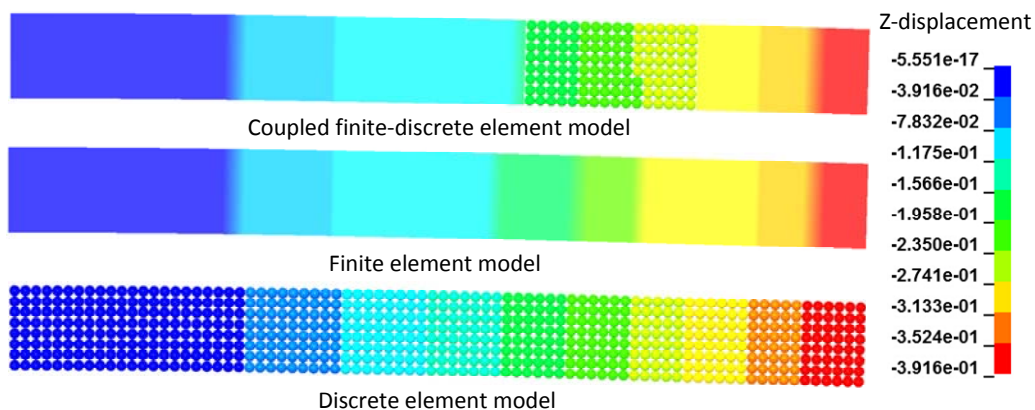


Figure 8. The comparison of displacement distributions along loading direction.



## Conclusions

In this paper a new approach to couple continua with DEs is proposed by using ghost particles. The ghost particles are constructed at the boundary FEs to connect DEs and this connection is achieved through the same interaction mechanism as the DE method. To keep ghost particles constantly sticking to the same natural coordinates of the corresponding FEs, an iterative strategy for inverse mapping from physical coordinates to natural coordinates is employed. The proposed numerical approach is used to model a cantilever beam and the agreement of the computed deflection at the free end with the analytical result validates the numerical method.

## Acknowledgement

The PhD scholarship awarded to Hu Chen from UNSW Canberra and the China Scholarship Council is gratefully acknowledged.

## References

- Mediavilla, J., Peerlings, R. H. J., and Geers, M. G. D. (2006) A robust and consistent remeshing-transfer operator for ductile fracture simulations, *Computers & Structures* **84**, 604-623.
- Belytschko, T. and Black, T. (1999) Elastic crack growth in finite elements with minimal remeshing, *International Journal for Numerical Methods in Engineering* **45**, 601-620.
- Belytschko, T., Moës, N., Usui, S., and Parimi, C. (2001) Arbitrary discontinuities in finite elements, *International Journal for Numerical Methods in Engineering* **50**, 993-1013.
- Moës, N. and Belytschko, T. (2002) Extended finite element method for cohesive crack growth, *Engineering Fracture Mechanics* **69**, 813-833.
- Munjiza, A. (2004) *The combined finite-discrete element method*, John Wiley & Sons, West Sussex.
- Munjiza, A., Owen, D., and Bicanic, N. (1995) A combined finite-discrete element method in transient dynamics of fracturing solids, *Engineering Computations* **12**, 145-174.
- Onate, E. and Rojek, J. (2004) Combination of discrete element and finite element methods for dynamic analysis of geomechanics problems, *Computer Methods in Applied Mechanics and Engineering* **193**, 3087-3128.
- Chen, H., Zhang, Y. X., Zang, M., and Hazell, P. J. (2015) An accurate and robust contact detection algorithm for particle-solid interaction in combined finite-discrete element analysis, *International Journal for Numerical Methods in Engineering*, doi: 10.1002/nme.4913.
- Wellmann, C. and Wriggers, P. (2012) A two-scale model of granular materials, *Computer Methods in Applied Mechanics and Engineering* **205-208**, 46-58.
- Li, M., Yu, H., Wang, J., Xia, X., and Chen, J. (2015) A multiscale coupling approach between discrete element method and finite difference method for dynamic analysis, *International Journal for Numerical Methods in Engineering* **102**, 1-21.
- Jebahi, M., Charles, J. L., Dau, F., Illoul, L., and Iordanoff, I. (2013) 3D coupling approach between discrete and continuum models for dynamic simulations (DEM-CNEM), *Computer Methods in Applied Mechanics and Engineering* **255**, 196-209.
- Tu, F., Ling, D., Bu, L., and Yang, Q. (2014) Generalized bridging domain method for coupling finite elements with discrete elements, *Computer Methods in Applied Mechanics and Engineering* **276**, 509-533.
- Christian, M., Françoise, L. P., and Ludovic, A. (2015) A least-squares coupling method between a finite element code and a discrete element code, *International Journal for Numerical Methods in Engineering* **101**, 731-743.
- Lei, Z. and Zang, M. (2010) An approach to combining 3D discrete and finite element methods based on penalty function method, *Computational Mechanics* **46**, 609-619.
- Xiao, S. P. and Belytschko, T. (2004) A bridging domain method for coupling continua with molecular dynamics, *Computer Methods in Applied Mechanics and Engineering* **193**, 1645-1669.
- Kohlhoff, S., Gumbsch, P., and Fischmeister, H. F. (1991) Crack propagation in b.c.c. crystals studied with a combined finite-element and atomistic model, *Philosophical Magazine A* **64**, 851-878.
- Shilkrot, L. E., Miller, R. E., and Curtin, W. A. (2002) Coupled Atomistic and Discrete Dislocation Plasticity, *Physical Review Letters* **89**, 025501.

- Yu, J., Analysis and application of the algorithm by combining discrete and finite element method in plane. 2011, South China University of Technology: Guangzhou, China.
- Belytschko, T., Liu, W. K., and Moran, B. (2013) *Nonlinear finite elements for continua and structures*.
- Qian, X., Ren, Q., and Zhao, Y. (1998) An algorithm for inverse isoparametric mapping in FEM, *Chinese Journal of Computational Mechanics* **15**, 437-441.

# Improved edge-based smoothed finite element method (IES-FEM) for Mid-frequency acoustic analysis

\*Z.C. He<sup>1,2</sup>, G.Y.Li<sup>1</sup>, Eric. Li<sup>1</sup>, G.Y. Zhang<sup>2</sup>, Xin.Nie<sup>1</sup>

<sup>1</sup> State Key Laboratory of Advanced Design and Manufacturing for Vehicle Body, Hunan University, Changsha, 410082 P. R. China

<sup>2</sup> State Key Laboratory of Structural Analysis for Industrial Equipment, School of Naval Architecture, Dalian University of Technology, Dalian, 116024, P. R. China

\*Corresponding author: hezhicheng815@163.com

## Abstract

The 3D ES-FEM using tetrahedral mesh proposed recently has showed many great features in mechanics, and mid-frequency acoustics analysis, such as  $kh \approx 1$ . When it comes to higher frequencies, the 3D ES-FEM also encounters the dispersion error, which is related to “slightly overly-soft stiffness” induced by the excessive edge-based smoothing operations compared to the continua system. In this paper, an improved 3D edge-based smoothed finite element method (IES-FEM) is proposed by introducing a parameter controlling the ratio of “slightly over-softness” of the ES-FEM and “over-stiffness” of the FEM, and the balance of the discretized system can then be tuned to reduce the dispersion error in higher frequency range, i.e.  $1 \leq kh \leq 2$ . Numerical results demonstrate the advantages of IES-FEM for acoustic problems, in comparison with the ES-FEM using the same mesh.

**Keywords:** Numerical method; Edge-based smoothed finite element method (ES-FEM); Acoustic; Dispersion error;

## 1. Introduction

Acoustic analysis has been a hot topic due to the increasing customer demands on the comfort of automobiles or aircraft, etc. As a traditional numerical modal is used, especially for a linear finite element system, the mid-frequency acoustic analysis often encounters computational difficulties. Researches in [1] reveal that the standard FEM can provide proper results with the restriction of  $kh < 1$ , which is related to the “rule of thumb”. In the numerical analysis, the linear FEM cannot give accurate prediction even if the  $kh < 1$  is satisfied.

In order to explain the root cause of the acoustic error and predict the acoustic field in the mid-frequency range, Ihlenburg [1] found that the acoustic error bounds contain a pollution term that are related to the loss of stability with large wave numbers, and he firstly showed that FEM with higher-order polynomial approximations (the  $hp$  version of the FEM) work well in reduce the acoustic error. The higher order methods such as  $p$ -FEM [2], and the discontinuous enrichment method (DEM) [3] have also been studied for acoustic computations in the mid-frequency regime. Babuška [4] attempted to correct the loss of stability in the Helmholtz operator and designed a Generalized Finite Element Method (GFEM) for the Helmholtz equation such that the pollution effect is minimal. The Galerkin/least-squares finite element method (GLS)[5], the quasi-stabilized finite element method (QSFEM) [4] were also proposed to improve the accuracy of acoustic analysis.

In a standard Galerkin FEM formulation, the discretized system behaves “overly-stiff”, which leads to inaccurate results in acoustic simulation, especially for the linear FEM in the mid-frequency regime. Due to the adaptively of linear elements (such as triangular and tetrahedral elements) in generating mesh in complex domain, the research on the use of low order mesh with high accuracy is of practical significance. Recently, Liu’s group conducted intensive studies on the softness of low-order elements and proposed smoothed finite element methods (S-FEM) by using the gradient smoothing technique together with the finite element methods. Using the node-based

strain smoothing operation, a node-based smoothed point interpolation method (NS-PIM or LC-PIM [6, 7]) and a node-based smoothed finite element method (NS-FEM) [8] have been formulated in the frame of meshfree and FEM, respectively. The node-based smoothed technique can provide a much better gradient solution than the standard FEM, while the smoothing domain contains too many elements leading to “overly-soft” of NS-FEM and instability in solving dynamic and acoustic problems [9,10]. The edge-based smoothed finite element methods (ES-FEM) [10-13] are thus proposed for 2D and 3D mechanics and acoustic problems, respectively. The ES-FEM showed super convergence properties, ultra accuracy and high computational efficiency compared to the traditional FEM using the same set of triangular and tetrahedral meshes, and is very suitable for dynamic and acoustic problems.

In studying the 3D ES-FEM for acoustic problems[13], the ES-FEM provides better results than improved FEM using hexahedral mesh at  $kh \approx 1$ . When it comes to higher frequencies, the ES-FEM also encounters the dispersion error [14], which is rooted at the “slightly over-softness” of stiffness matrix in ES-FEM. In this paper, an improved 3D edge-based smoothed finite element method (IES-FEM) is proposed by introducing a parameter controlling the ratio of “slightly over-softness” of the ES-FEM and “over-stiffness” of the FEM, and the balance of the discretized system can then be tuned to reduce the dispersion error in higher frequency range, *i.e.*  $1 \leq kh \leq 2$ . Numerical results demonstrate the advantages of IES-FEM for acoustic problems, in comparison with the standard ES-FEM using the same elements.

In this work, we mainly focus on the acoustic analysis in mid-frequency regime, and “the  $1 < kh < 2$ ” is adopted as a reference of mid-frequency. The paper is organized as follows: Section 2 briefly describes the mathematical model of the acoustic problems. The idea of IES-FEM method is formulated detailed in Section 3. The numerical example is used to evaluate the performance of the proposed method in Section 4. Finally the conclusions from the numerical results are made in Section 5.

## 2. Acoustic problems and its standard Galerkin discretization

The acoustic pressure  $p$  in a bounded domain  $\Omega$  is governed by the following well-known Helmholtz equation:

$$\Delta p + k^2 p = 0 \quad (1)$$

where  $\Delta$  is the Laplace operator,  $k$  is the wavenumber defined by

$$k = \frac{\omega}{c} \quad (2)$$

where  $c$  and  $\omega$  denote the speed of sound traveling in the homogeneous media and angular frequency, respectively. For general interior acoustic problems with boundary  $\Gamma$ , there are three types of boundary conditions prescribed on the boundary of  $\Gamma_D$ ,  $\Gamma_N$  and  $\Gamma_A$ , where  $\Gamma = \Gamma_D \cup \Gamma_N \cup \Gamma_A$ , and the three sets of boundary conditions are expressed

$$p = p_D \quad \text{on } \Gamma_D \quad (3)$$

$$\nabla p \cdot n = -j\rho\omega v_n \quad \text{on } \Gamma_N \quad (4)$$

$$\nabla p \cdot n = -j\rho\omega A_n p \quad \text{on } \Gamma_A \quad (5)$$

where  $p_D$  is the prescribed acoustic pressure on the boundary  $\Gamma_D$ ,  $v_n$  is the normal velocity on boundary  $\Gamma_N$ ,  $\rho$  is the medium density and  $A_n$  represents the admittance coefficient on boundary  $\Gamma_A$ . In the standard FEM, the discrete of acoustic pressure  $p$  can be expressed in the following form:

$$p = \sum_{i=1}^m N_i p_i = \mathbf{Np} \quad (6)$$

where  $N_i$  are the nodal shape functions obtained using standard finite element procedure and  $p_i$  are the unknown nodal pressures. In standard Galerkin weak form, the shape function  $\mathbf{N}$  is also chosen as the weight function  $w$  and the discretized form for acoustic problems can be obtained as:

$$-\int_{\Omega} \nabla \mathbf{N} \cdot \nabla \mathbf{N} p d\Omega + k^2 \int_{\Omega} \mathbf{N} \cdot \mathbf{N} p d\Omega - j\rho\omega \int_{\Gamma_N} \mathbf{N} \cdot \mathbf{v}_n d\Gamma - j\rho\omega A_n \int_{\Gamma_A} \mathbf{N} \cdot \mathbf{N} p d\Gamma = 0 \quad (7)$$

The discretized system equations can be finally obtained and written in the following matrix form:

$$[\mathbf{K}^{\text{FEM}} - k^2 \mathbf{M}^{\text{FEM}} + j\rho\omega \mathbf{C}] \{\mathbf{P}\} = -j\rho\omega \{\mathbf{F}\} \quad (8)$$

where  $\mathbf{K}^{\text{FEM}}$  is the acoustical stiffness matrix,  $\mathbf{M}^{\text{FEM}}$  is the acoustical mass matrix,  $\mathbf{C}$  is the acoustical damping matrix,  $\{\mathbf{P}\}^T$  is nodal acoustic pressure vector,  $\mathbf{F}$  is the nodal acoustic forces vector, and all of them are described as follows

$$\mathbf{K}^{\text{FEM}} = \int_{\Omega} (\nabla \mathbf{N})^T (\nabla \mathbf{N}) d\Omega \quad (9)$$

$$\mathbf{M}^{\text{FEM}} = \int_{\Omega} \mathbf{N}^T \mathbf{N} d\Omega \quad (10)$$

$$\mathbf{C} = \int_{\Gamma_A} \mathbf{N}^T \mathbf{N} A_n d\Gamma \quad (11)$$

$$\mathbf{F} = \int_{\Gamma_N} \mathbf{N}^T \mathbf{v}_n d\Gamma \quad (12)$$

### 3. The idea of Improved edge-based smoothed finite element method (IES-FEM)

#### 3.1 Brief the Smoothing domain in ES-FEM for acoustic problems

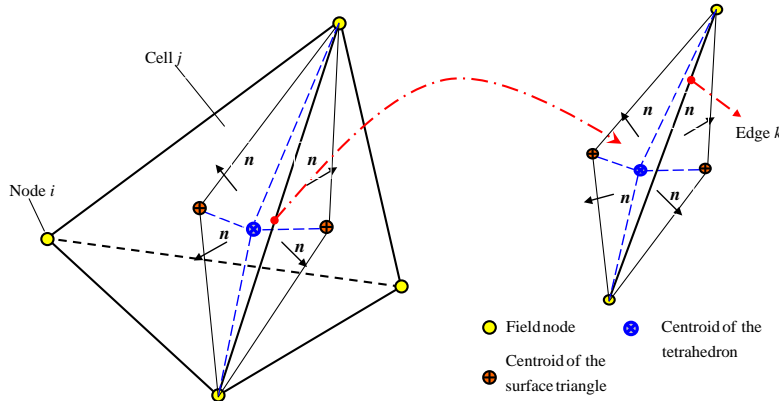


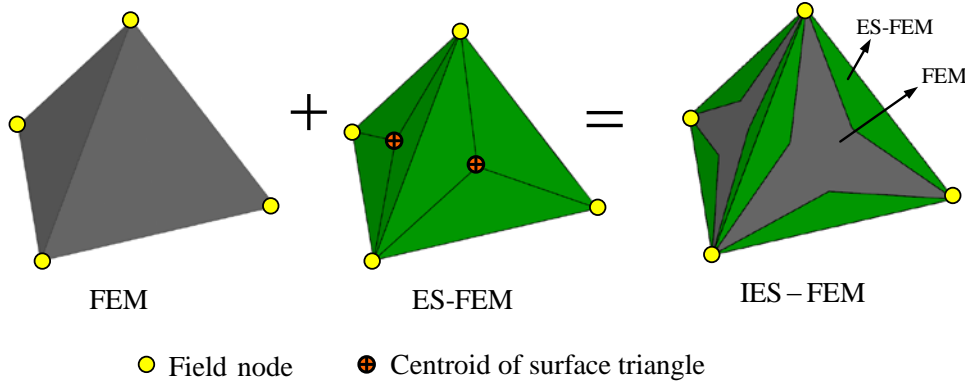
Figure 1 The sub-smoothing-domain of edge  $k$  in cell  $j$

In the scheme of the ES-FEM for 3D problems, the numerical integration of Eq. (9) is not performed based on the tetrahedral elements but the edge-based smoothing domains. In the construction of local smoothing domains, the sub-domain of the smoothing domain  $\Omega_k^s$  for edge  $k$  located in the particular cell  $j$  can be obtained by connecting two end nodes of the edge to the centroids of the surface triangles and the centroid of cell  $j$ , as shown in Fig. 1. The sub-smoothing-domain for edge  $k$  is one sixth region of this tetrahedral element. Finding out other sub-domains located in other elements containing edge  $k$  and the smoothing domain for edge  $k$  can be constructed by uniting all the sub-domains. The number of local smoothing domain is equals to the number of elemental edges. Using edge-based smoothing operation, the gradient component  $\nabla \mathbf{N}$  is replaced by the smoothed item  $\overline{\nabla \mathbf{N}}$ , and the global smoothed acoustic stiffness matrix can be written as:

$$\overline{\mathbf{K}}^{\text{ES-FEM}} = \int_{\Omega} (\overline{\nabla \mathbf{N}})^T (\overline{\nabla \mathbf{N}}) d\Omega \quad \text{The smoothed acoustical stiffness matrix} \quad (13)$$

### 3.2 Improved ES-FEM (IES-FEM) for acoustic problems

Recently, it is found that the ES-FEM also encounters the dispersion error[14] in large wave number, i.e.  $1 \leq kh \leq 2$ , which is related to “slightly overly-soft stiffness” induced by the excessive edge-based smoothing operations compared to the continua system. The numerical method can be tuned to the exact or nearly exact stiffness using appropriate domains as well. Similar as  $\alpha$ -FEM by combining the FEM and NS-FEM, an IES-FEM is also proposed by introducing a parameter alpha making the best use of “over-stiffness” of the FEM model and “slightly over-softness” of the ES-FEM model to achieve the ultimate performance.



**Figure 2** The IES-FEM is formulated by combining of the FEM and ES-FEM

In the 3D problem of IES-FEM, each tetrahedral element is divided into seven portions, as shown in Fig.2: six volumes containing tetrahedral edges have an equal volume of  $\frac{1}{6}\alpha V_e$ , and the remaining part in the middle of the element has a volume of  $(1-\alpha)V_e$ , where the  $V_e$  is the volume of the tetrahedral element. Six volumes containing tetrahedral edges compose a part of six edge-based smoothing domain of ES-FEM, and the middle part is used to calculate the contribution to the local stiffness matrix of IES-FEM using FEM, thus the stiffness of IES-FEM can be formulated as follows:

$$\mathbf{K}^{\text{IES-FEM}} = (1-\alpha) \sum_i^{N^e} \mathbf{K}^{\text{FEM}} + \alpha \sum_i^{N^n} \mathbf{K}^{\text{ES-FEM}} \quad (14)$$

where  $N^e$  is the number of total elements in the entire problem domain and  $N^n$  is the number of total edges in the entire problem domain. The stiffness matrix  $\mathbf{K}^{\text{FEM}}$  and  $\mathbf{K}^{\text{ES-FEM}}$  can be calculated by Eq. (9) and Eq. (13). Note that the parameter alpha which controls the contribution of ES-FEM and FEM can be selected as constant to improve the results of acoustic simulation [14]. In the 3D IES-FEM for acoustic problems, numerical analysis indicates that when alpha equals to 0.82, it can always provide very good results.

## 4. Numerical study

### 4.1 Numerical error for acoustic problems

The pollution is mainly a consequence of the dispersion effect, meaning that the wave number of the numerical solution and the wave number of the exact solution disagree. The dispersion effect was first examined for the wave equation [1]

In the mid-frequency analysis, the standard FEM cannot provide proper results under the condition of  $kh > 1$  due to dispersion error. In this work, we mainly focus on the acoustic analysis in mid-frequency regime, and “the  $1 < kh < 2$ ” is adopted as a reference of mid-frequency. In the

following computations, the numerical global error indicator is given in terms of velocity and can be described as follows:

$$\left| p^e - p^h \right|_1^2 = \int_{\Omega} (\tilde{v}^e - \tilde{v}^h)^T (v^e - v^h) d\Omega \quad (15)$$

where  $\tilde{v}$  is complex conjugate of the velocity  $v$ , the superscript  $e$  denotes the exact solutions and  $h$  denotes the numerical solutions obtained from numerical methods including the present IES-FEM, ES-FEM using low-order elements.

#### 4.2 3D tube with Neumann boundary condition

The numerical example adopted in 3D interior acoustic problem is a cylinder tube filled with air as shown in Fig. 3. The air density  $\rho$  is  $1.225\text{kg/m}^3$  and the speed of sound in the air is  $340\text{m/s}$ . The dimension of this cylinder tube with length  $l=1\text{m}$ , diameter  $d=0.3\text{m}$  is considered. The boundary conditions of this problem are that: the left end of tube is specified normal velocity boundary condition with  $v_n=10\text{m/s}$ , and the right end of tube is a rigid wall with zero velocity. The analytical solutions for this problem can be easily derived and the acoustic pressure and velocity are given by

$$p = -j\rho cv_n \frac{\cos(k(1-\xi))}{\sin(k)} \quad (16)$$

$$v = \frac{v_n \sin(k(1-\xi))}{\sin(k)} \quad (17)$$

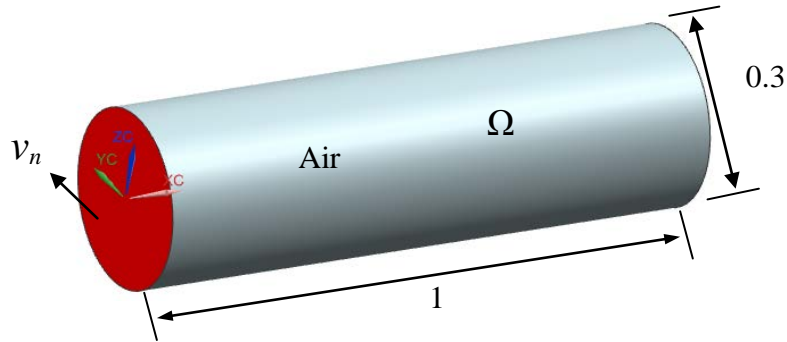


Figure 3 3D acoustic tube with the Neumann boundary condition.

##### 4.2.1 The verification of parameter alpha

In 3D acoustic analysis, the parameter alpha is firstly investigated. The frequency value of  $1352\text{Hz}$  is used. Fig. 4 plots the numerical error of IES-FEM against the parameter alpha from 0 to 1 with a step of 0.01 by the use of four models with nodal spacing of  $0.06\text{m}$ ,  $0.05\text{m}$ ,  $0.04\text{m}$  and  $0.03\text{m}$ . It can be seen from this figure that: (1) at frequency of  $1352\text{Hz}$ , the monotonic convergence can be obtained for the IES-FEM with alpha varies from 0 to 1 with the refinement of mesh; and the IES-FEM can always provide very good results when  $\alpha$  equals to about 0.82.

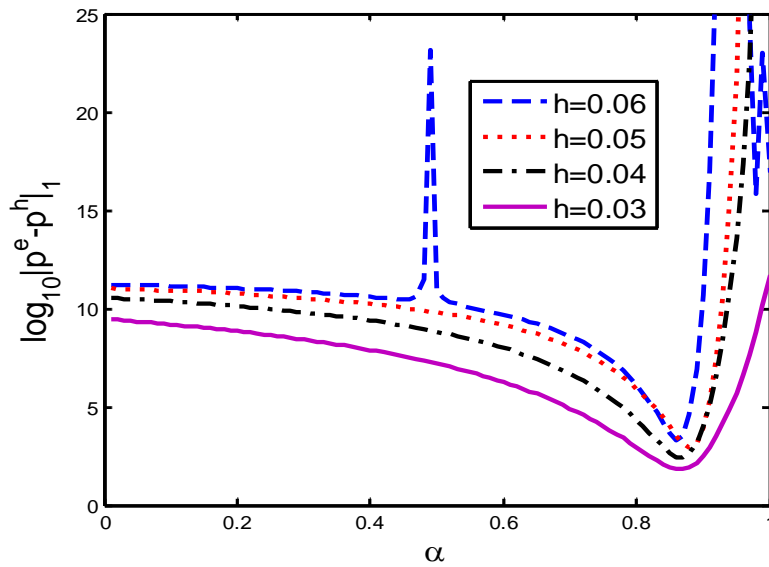


Figure 4 Numerical error for different mesh size by varying the parameter alpha at different frequencies

#### 4.2.2 Acoustic accuracy and convergence

The tube is discretized by unstructured mesh with an average mesh size of 0.06m. The frequency value of 1352Hz ( $kh=1.5$ ) has been studied using present IES-FEM. For the purpose of comparison, ES-FEM[13] solutions are also computed using this set of tetrahedral mesh. The results obtained from these two numerical methods have been plotted in Fig. 5, together with the exact solutions. It can be seen from the plots that: the IES-FEM can give much better results than the ES-FEM, that gives solution departing a lot from the exact one. The convergence property is also investigated by using four models with 588, 990, 1753 and 3894 uniformly distributed nodes with nodal spacing of 0.06m, 0.05m, 0.04m and 0.03m. Fig. 6 presents the convergence curves in terms of global error against the average nodal spacing  $h$  at the frequency of 1352Hz for ES-FEM and IES-FEM. From the figure, it can be observed that the present IES-FEM can give much more accurate gradient results than that of ES-FEM.

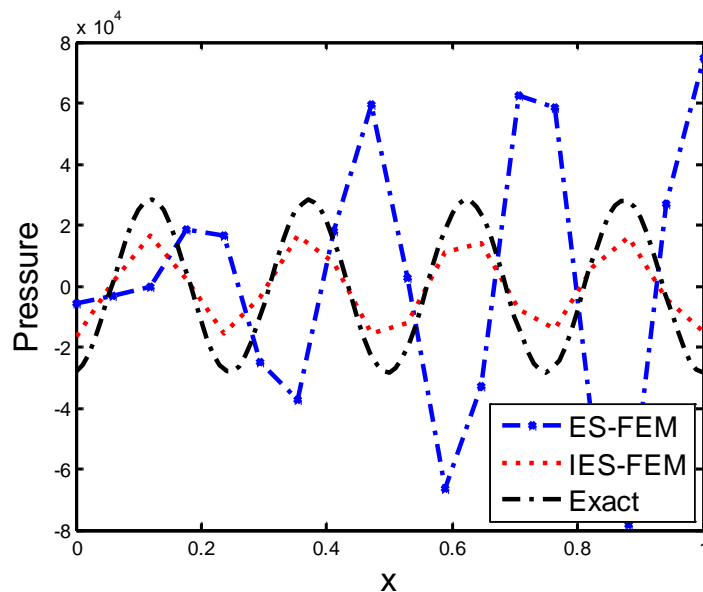


Figure 5 Exact and numerical solutions of acoustic pressure for the 3D acoustic tube at  $kh=1.5$



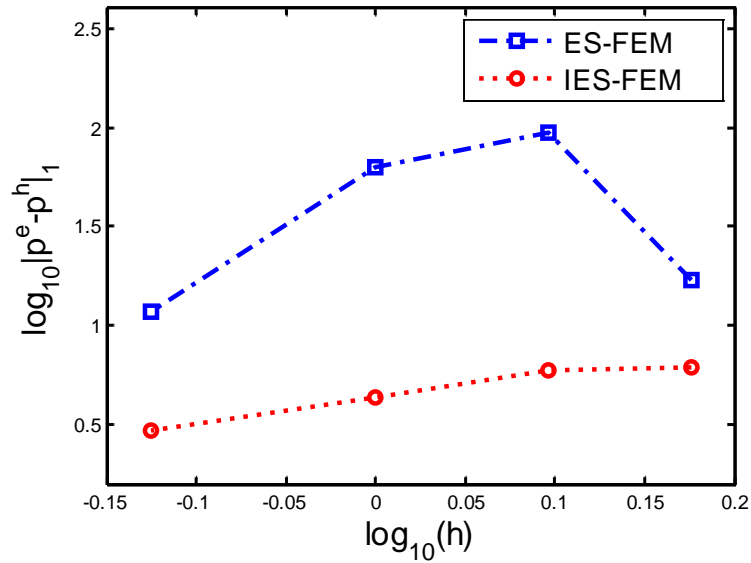


Figure 6 Comparison of accuracy and convergence for 3D acoustic tube

### 4.2.3 Computational time and efficiency study

Fig. 7 compares the computational time between IES-FEM and ES-FEM. Note the bandwidth of IES-FEM is the same as ES-FEM, and the computational time using IES-FEM is almost the same as ES-FEM. However, when it comes to the computational efficiency in terms of CPU time for the same accuracy, the IES-FEM is found to much more efficient than the ES-FEM as shown in Fig. 8.

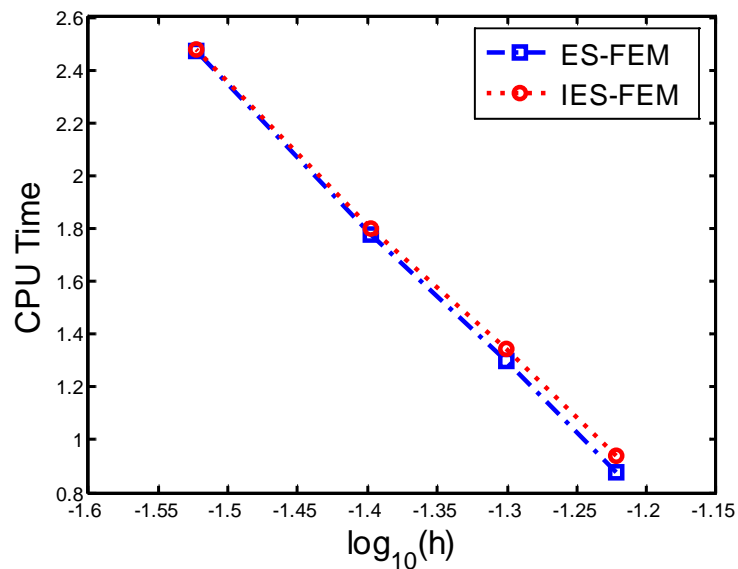
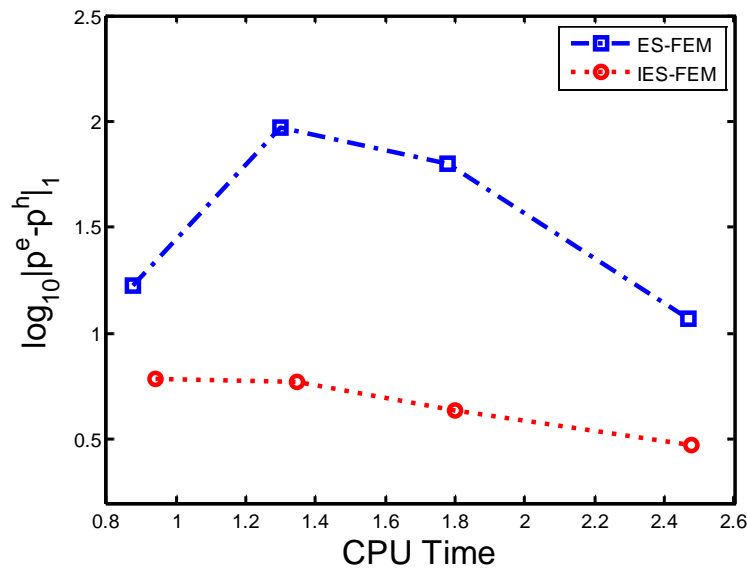


Figure 7 Comparison of the computational time for the IES-FEM and ES-FEM



**Figure 8** Comparison of the efficiency of numerical results in terms of global error

## 5. Conclusions

In this paper, an improved 3D edge-based smoothed finite element method (IES-FEM) is proposed by introducing a parameter controlling the ratio of “slightly over-softness” of the ES-FEM and “over-stiffness” of the FEM, and the balance of the discretized system can then be tuned to reduce the dispersion error in higher frequency range, i.e.  $1 \leq kh \leq 2$ . The determined parameter alpha controls a proper gradient smoothing operation in the IES-FEM, and provides a perfect balance between stiffness and mass in the discrete system matrix, which dramatically reduces the dispersion error. Numerical results demonstrate the advantages of IES-FEM for acoustic problems, in comparison with the ES-FEM using the same mesh.

- The IES-FEM uses the simplest linear tetrahedral mesh, which can be easily generated for any complicated domains of acoustic media, and hence is ideal for automated modeling and simulation.
- In the IES-FEM, no additional parameters or degrees of freedom are introduced, and the present method can be implemented in a straightforward way with little change to the standard FEM code.
- In IES-FEM for acoustic problems, when the parameter alpha equals to 0.82, the numerical studies show that it can always provide much better results and is more efficient than ES-FEM.

## Acknowledgements

The authors wish to thank the support of National Natural Science Foundation of China (Grant No. 11202074, Grant No.61232014), and Research Project of State Key Laboratory of Structural Analysis for Industrial Equipment (Grant No. GZ1403).

## References

- [1] Ihlenburg F. (1998) *Finite Element Analysis of Acoustic Scattering*. New York: Springer-Verlag.
- [2] Petersen S., Dreyer D., Estorff Ov. (2006) Assessment of finite and spectral element shape functions or efficient iterative simulations of interior acoustics, *Computer Methods in Applied Mechanics and Engineering* **195**, 6463-6478.

- [3] Farhat C., Harari I., Franca L.P. (2001) The discontinuous enrichment method, *Computer Methods in Applied Mechanics and Engineering* **190**, 6455 - 6479.
- [2] Babuska I., Ihlenburg F., Paik E., and Sauter S. (1995) A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution, *Computer Methods in Applied Mechanics and Engineering* **128**, 325 - 359.
- [4] Harari I., and Hughes, T.J.R. (1992) Galerkin/least-squares finite element methods for the reduced wave equation with nonreflecting boundary conditions in unbounded domains, *Computer Methods in Applied Mechanics and Engineering* **98**, 411-454.
- [5] Thompson L.L., Pinsky P.M. (1995) A Galerkin least-squares finite element method for the two-dimensional Helmholtz equation. *International Journal for Numerical Methods in Engineering* **38**, 371-397.
- [6] Liu G.R., Zhang G.Y., Dai K.Y., Wang Y.Y., Zhong Z.H., Li G.Y., Han X. (2005) A linearly conforming point interpolation method (LC-PIM) for 2D solid mechanics problems, *International Journal of Computational Methods* **2**: 645-665.
- [7] Zhang G.Y., Liu G.R., Wang Y.Y., Huang H.T., Zhong Z.H., Li G.Y., Han X. (2007) A linearly conforming point interpolation method (LC-PIM) for three-dimensional elasticity problems, *International Journal for Numerical Methods in Engineering* **72**, 1524-1543.
- [8] Liu G.R., Nguyen T.T., Nguyen X.H., Lam K.Y. (2009) A node-based smoothed finite element method for upper bound solution to solid problems. *Computers and Structures* **87**, 14-26.
- [9] Liu G.R., Nguyen T.T., Lam K.Y. (2009) An edge-based smoothed finite element method (ES-FEM) for static, free and forced vibration analysis, *Journal of Sound and Vibration* **320**, 1100-1130.
- [10] He Z.C., Cheng A.G., Liu G.R., Zhong Z.H., Zhang G.Y., Cheng A. G. (2010) Dispersion free analysis of acoustic problems using the alpha finite element method. *Computational Mechanics* **46**, 867-881.
- [11] He Z.C., Liu G.R., Zhong Z.H., Wu S.C., Zhang G.Y., Cheng A.G. (2009) An edge-based smoothed finite element method (ES-FEM) for analyzing three-dimensional acoustic problems, *Computer Methods in Applied Mechanics and Engineering* **199**, 20-33.
- [12] He Z.C., Li G.Y., Zhong Z.H., Cheng A.G., Zhang G.Y. et (2013) An edge-based smoothed tetrahedron finite element method (ES-T-FEM) for 3D static and dynamic problems, *Computational Mechanics* **52**, 221-236.
- [13] He Z.C., Li G.Y., Zhong Z.H., Cheng A.G., Zhang G.Y., Li Eric, Liu G.R. (2012) An ES-FEM for accurate analysis of 3D mid-frequency acoustics using tetrahedron mesh. *Computers and Structures*, **106 - 107**: 125 - 134.
- [14] He Z.C., Li Eric, Li G.Y., Zhong Z.H., Liu G. R., Nie X. (2015) Acoustic simulation using  $\alpha$ -FEM with a general approach for reducing dispersion error, *Engineering Analysis with Boundary Elements* (revised).

## High Resolution Numerical Simulation for corner-turning in LX-17

†C. Wang, \*X.Q. Liu

State Key Lab of Explosion Science and Technology, Beijing Institute of Technology, Beijing, 100081, China

\*Presenting author: 3120100070@bit.edu.cn

†Corresponding author: wangcheng@bit.edu.cn

### Abstract

In this paper, Combined with the third-order TVD Runge-Kutta method, we develop a parallel solver using the fifth-order high-resolution weighted essentially non-oscillatory (WENO) finite difference scheme to simulate detonation diffraction for three-dimensional condensed explosives. Using the explosive LX-17, abrupt corner turning of detonation was investigated as the detonation moved from a near-ideal cylinder of small radius suddenly into a cylinder of large radius. The numerical simulation results revealed the restrictive relationships among the pressure, the density and the reaction progress in the failure regions around the corner. As a result, the detonation cannot turn the corner and subsequently fails, but the shock wave continues to propagate in the unreacted explosive, leaving behind a dead zone. Previously, we have used the PBX-9404 explosive to simulation the detonation diffraction. Comparing with that, the inert explosive LX-17 has the longer distance to detonation along the inner wall. It demonstrated that a larger field of the dead zone can be generated, and it may not close.

**Keywords:** Condensed explosives; Numerical simulation; WENO; High resolution; Corner-Turning; Dead zone

### Introduction

The ability of a detonation wave to continue to propagate after it encounters an abrupt or a more gradual geometry change is very important to understand in practical explosive charge design. The corner leads the detonation wave to curve and spread to the side along the inner wall, and the reaction zone may decouple and lag behind the shock. If extinction occurs, the lead shock degrades into a weak shock followed by a fast flame. The dead zone represents a loss of available energy in the system to do work and must be computed accurately.

Numerous experimental and numerical studies have been performed to study corner turning phenomena. Souers *et al.* [2004] have performed a series of highly-instrumented experiments examining corner-turning of detonation, and they used pulsed X-rays to observe the dead zones in ambient detonating LX-17 with the breakout on the edges measured by streak camera. Cox and Campbell studied the ability of PBX 9502 to turn corners using a long, cylindrical straight section that suddenly changes into a much wider cylinder, which has a slice taken out of one side [Cox and Campbell (1981)]. They concluded that the explosive had a 17mm corner turning radius, and that the explosive inside this radius did not detonate. The most detailed experiments of TATB corner turning and dead zone formation are those fired using proton radiography at Los Alamos National Laboratory [Ferm *et al.* (2002) ; Mader *et al.* (2002)]. The greater penetrating power of high energy protons compared to X-rays allows finer structures of the dead zones to be observed. Kapila [Kapila *et al.* (2007)] simulated the process of detonation diffraction of the explosive LX-17 with different corners by second-order accurate capture scheme to research the changes of the explosive state at the corner wall under the influence of detonation diffraction around the corner and observed the initiation process along the inner wall to be similar to the shock-initiation under low-pressure. Tarver [2010] modeled the hydrodynamics of double corner turning and shock desensitization in two dimensions using the Ignition and Growth LX-17 detonation reactive flow model. They compared the calculated arrival times and axial free surface velocity histories of the top aluminum plates with the experimental measurements, well in agreement. In recent years, Yang G. [Yang *et al.* (2013)] used the smoothed particle hydrodynamics (SPH) method combined with ignition and growth model to give good prediction for the von Neumann spike state of detonation in condensed explosives. The smoothed particle hydrodynamics (SPH) method is more and more widely applied to investigate the detonation phenomena, e.g. [Yang *et al.* (2011); Hu *et al.* (2014)].

In this paper, the process of corner-turning of the double cylinder geometries in LX-17 are numerically simulated by high resolution numerical scheme. Fifth-order WENO scheme [Jiang and

Shu (1996)] and third-order TVD Runge-Kutta method are employed to discretize Euler equations with chemical reaction source, and parallel high resolution code is developed. The code was used to stimulate the process of detonation diffraction at the corner. Respectively, the change rules of pressure, density and reaction rate in the low-pressure region and the inner wall region will be investigated, and the corner effect on the detonation wave propagation will be discussed.

## 1. Governing equations

The detonation process is very short, so the transport effect of viscosity, heat conduction and diffusion can be neglected in the detonation process. The non-stationary compressible Euler equations can be used as the fluid dynamics governing equations. In the form as follows

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y + \mathbf{H}(\mathbf{U})_z = \mathbf{S}(\mathbf{U}) \quad (1)$$

Where the conserved variable vector  $\mathbf{U}$ , the flux vectors  $\mathbf{F}$ ,  $\mathbf{G}$  and  $\mathbf{H}$  as well as the source term  $\mathbf{S}$  are given, respectively, by

$$\mathbf{U} = (\rho, \rho u, \rho v, \rho w, \rho E, \rho \lambda)^T \quad (2)$$

$$\mathbf{F}(\mathbf{U}) = (\rho u, \rho u^2 + p, \rho uv, \rho uw, \rho u(E + p/\rho), \rho u \lambda)^T \quad (3)$$

$$\mathbf{G}(\mathbf{U}) = (\rho v, \rho vu, \rho v^2 + p, \rho vw, \rho v(E + p/\rho), \rho v \lambda)^T \quad (4)$$

$$\mathbf{H}(\mathbf{U}) = (\rho w, \rho wu, \rho wv, \rho w^2 + p, \rho w(E + p/\rho), \rho w \lambda)^T \quad (5)$$

$$\mathbf{S}(\mathbf{U}) = (0, 0, 0, 0, 0, \rho \lambda)^T \quad (6)$$

$$E = e + (1 - \lambda)q + \frac{1}{2}(u^2 + v^2) \quad (7)$$

Here  $u$ ,  $v$ ,  $w$  are the Cartesian component of the particle velocity in the  $x$ ,  $y$ ,  $z$  directions. Respectively,  $\rho$  is density,  $p$  is pressure,  $E$  is total energy per unit volume,  $e$  is internal energy,  $\lambda$  is reaction rate variable.

## 2. Equations of state and reaction rate

The unreacted explosive and the reaction products are both described by the JWL equation of state, but their parameter values are different. For LX-17 explosive, the specific parameter values of equation of state are shown in Table 1 [Zhang and Chen (1991)].

$$\begin{aligned} p &= A \exp(-R_1 V) + B \exp(-R_2 V) + \frac{\omega}{V} C_v T \\ e &= \frac{A}{\rho_0 R_1} \exp(-R_1 V) + \frac{B}{\rho_0 R_2} \exp(-R_2 V) + \frac{C_v T}{\rho_0} \\ V &= v / v_0 \end{aligned} \quad (8)$$

**Table 1. EOS data for the explosive LX-17**

JWL parameters	Unreacted	Products
$A(10^2 \text{GPa})$	77.81	6.547
$B(10^2 \text{GPa})$	-1.5031	0.7124
$C_v(10^2 \text{GPa/K})$	$2.487 \times 10^{-5}$	$1.0 \times 10^{-5}$
$R_1$	11.3	4.45
$R_2$	1.13	1.2
$\omega$	0.8939	0.38
$\rho_0(\text{g/cm}^3)$	1.895	1.895

Respectively,  $v$  is the specific volume, and  $v_0$  is the specific volume at the initial moment.  $A$ ,  $B$ ,  $R_1$ ,  $R_2$ ,  $C_0$  and  $\omega$  are constants, and  $T$  is temperature.

The most commonly used chemical reaction rate model of condensed explosives is the ignition and growth reaction rate model proposed by Lee and Tarver. For LX-17 explosive, the specific parameter values of the chemical reaction rate equation are shown in Table 2 [Zhang and Chen (1991)].

$$\frac{d\lambda}{dt} = I(1-\lambda)^b (\rho / \rho_0 - 1 - a)^x + G_1(1-\lambda)^c \lambda^d p^y + G_2(1-\lambda)^e \lambda^g p^z \quad (9)$$

Here  $I$ ,  $G_1$ ,  $G_2$ ,  $b$ ,  $a$ ,  $x$ ,  $c$ ,  $d$ ,  $y$ ,  $e$ ,  $g$  and  $z$  are constants.

**Table 2. Rate data for the explosive LX-17**

Parameters	Value
$I(\mu\text{sec}^{-1})$	$4.0 \times 10^6$
$G_1(\mu\text{sec}^{-1}(10^2\text{GPa})^{-z})$	0.6
$G_2(\mu\text{sec}^{-1}(10^2\text{GPa})^{-z})$	400
$b$	0.667
$a$	0.22
$x$	7.0
$c$	0.667
$d$	0.111
$y$	1.0
$e$	0.333
$g$	1.0
$z$	3.0

### 3. The numerical method

In this paper, WENO finite difference scheme is employed to discretize Euler equations with chemical reaction source in space. The semi-discrete scheme is as follows:

$$\left( \frac{\partial \mathbf{U}}{\partial t} \right)_{i,j,k} = - \frac{(\tilde{\mathbf{F}}_{i+1/2,j,k} - \tilde{\mathbf{F}}_{i-1/2,j,k})}{\Delta x} - \frac{(\tilde{\mathbf{G}}_{i,j+1/2,k} - \tilde{\mathbf{G}}_{i,j-1/2,k})}{\Delta y} - \frac{(\tilde{\mathbf{H}}_{i,j,k+1/2} - \tilde{\mathbf{H}}_{i,j,k-1/2})}{\Delta z} + \mathbf{S}_{i,j,k} \quad (10)$$

Third-order TVD Runge-Kutta scheme is employed for temporal discretization.

$$\begin{cases} \mathbf{U}_{i,j,k}^{(1)} = \mathbf{U}_{i,j,k}^n + \Delta t L(\mathbf{U}_{i,j,k}^n) \\ \mathbf{U}_{i,j,k}^{(2)} = \frac{3}{4} \mathbf{U}_{i,j,k}^n + \frac{1}{4} \mathbf{U}_{i,j,k}^{(1)} + \frac{1}{4} \Delta t L(\mathbf{U}_{i,j,k}^{(1)}) \\ \mathbf{U}_{i,j,k}^{n+1} = \frac{1}{3} \mathbf{U}_{i,j,k}^n + \frac{2}{3} \mathbf{U}_{i,j,k}^{(2)} + \frac{2}{3} \Delta t L(\mathbf{U}_{i,j,k}^{(2)}) \end{cases} \quad (11)$$

### 4. Results and discussion

The explosive LX-17 is selected as an example, and the model geometry size is shown in figure 1. There is the inflow boundary conditions on the left of the small radius cylinder and all other boundaries are set wall boundary conditions. We take the mesh size  $\Delta=0.1(\text{mm})$ , and the explosive is initiated by C-J (Chapman-Jouguet) condition within 0.5mm distance on the left of the small radius cylinder.

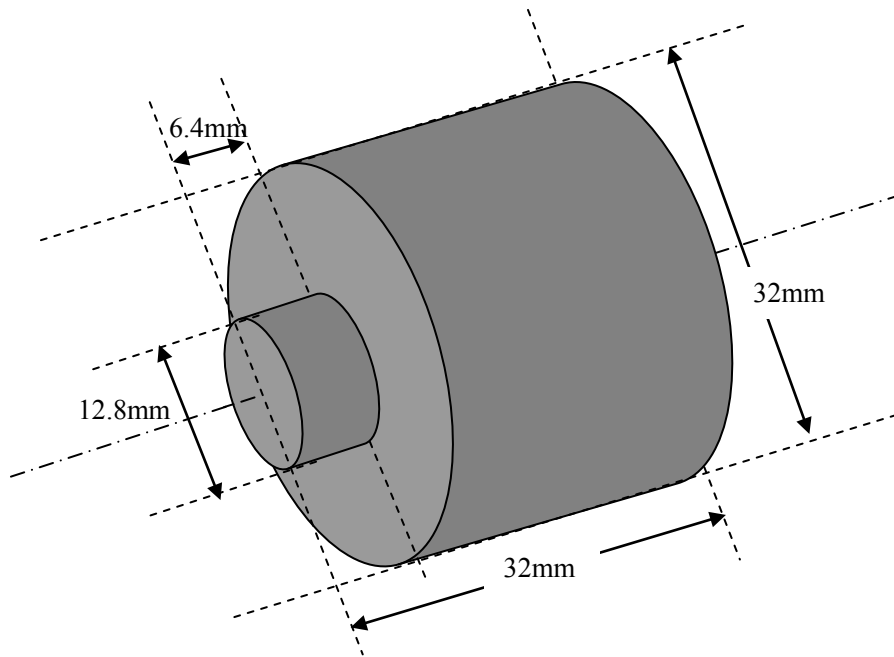
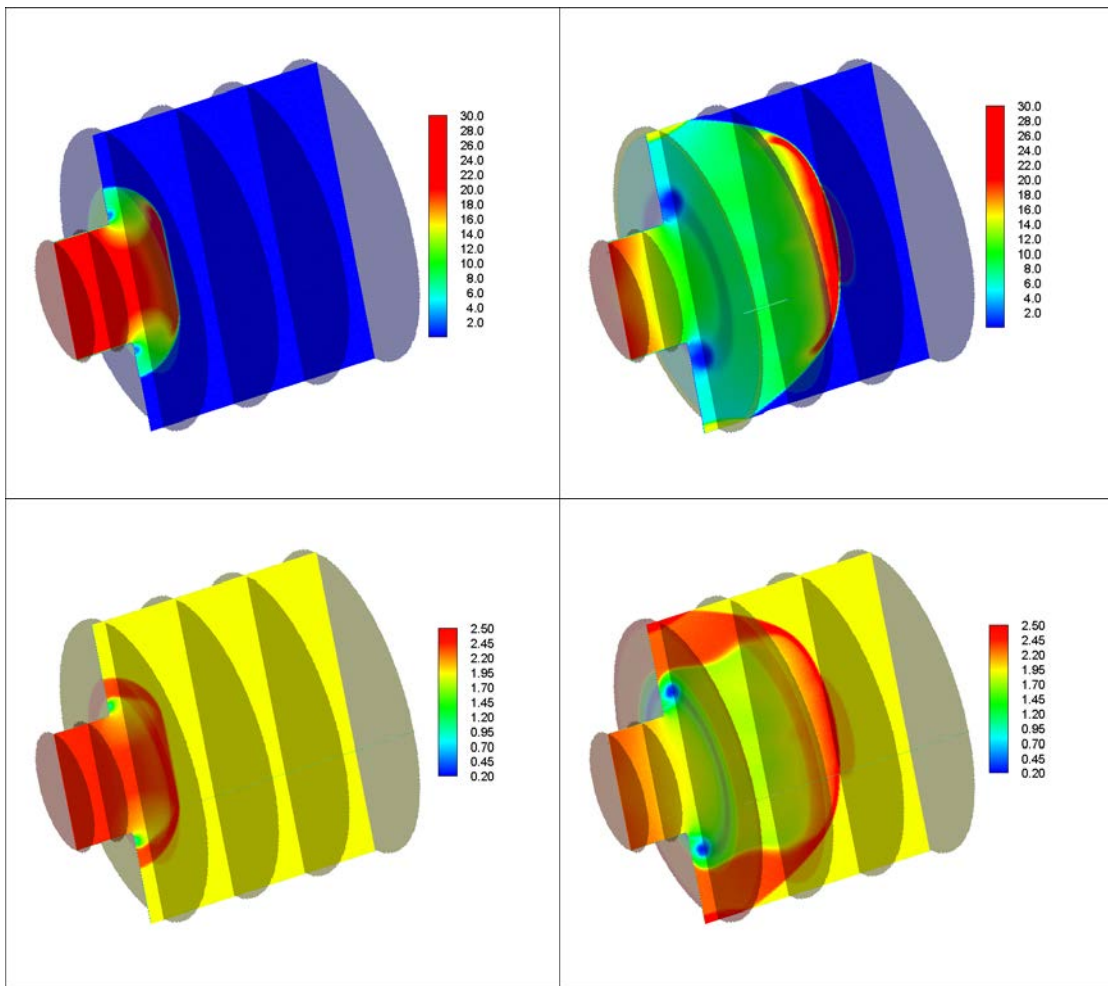
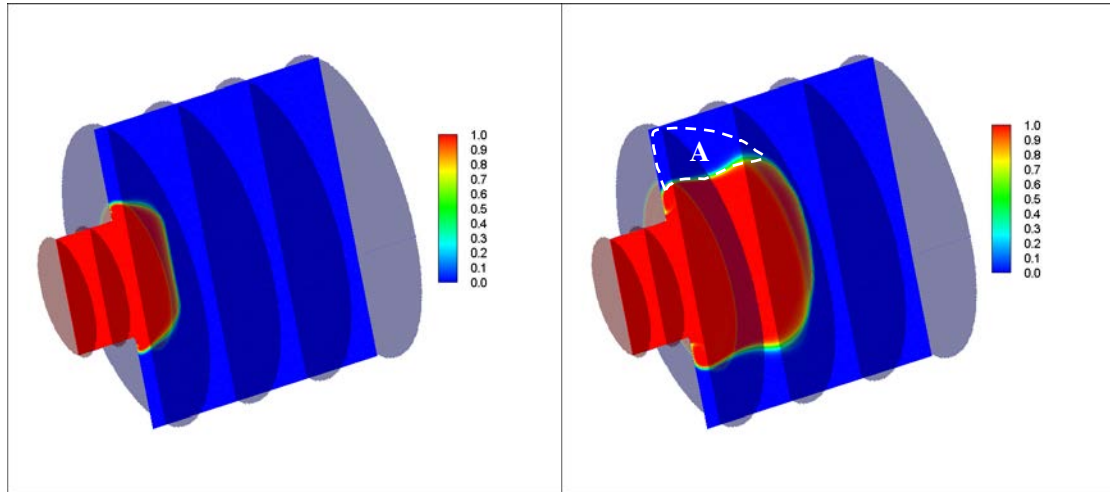


Fig. 1. Schematic for the double-cylinder corner-turning model

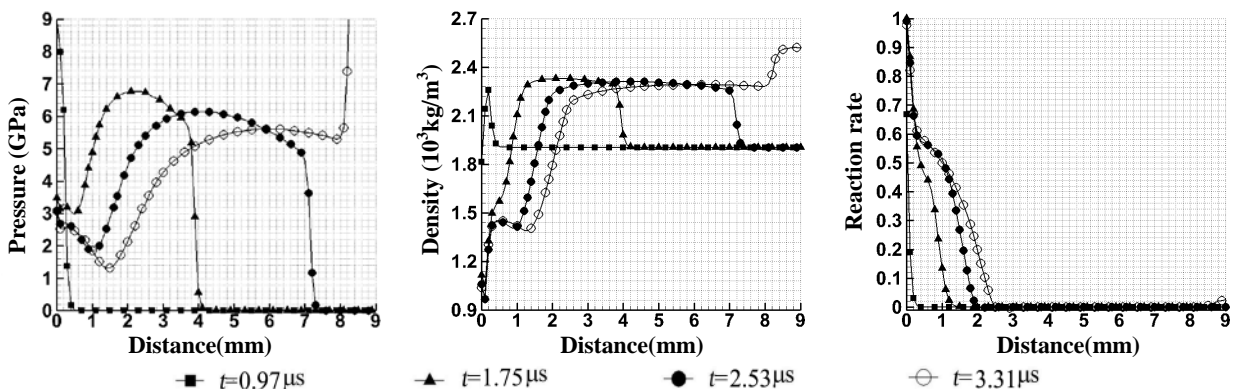




**Fig. 2. Contours of pressure (top), density (middle) and the reaction rate (bottom) of the double-cylinder corner-turning at 1.65 $\mu$ s (left), 3.3 $\mu$ s (right)**

Previously, we have obtained The propagating state of detonation wave around the corner is generally determined by two factors in the other paper. One is the transverse shock wave along the inner wall downwards, and the other is the extending curved detonation wave. In this model, when the detonation wave reaches the corner position, the time span is 0.97 $\mu$ s. From figure 2, because of the effect of corner turning influence, the expansion generated there is felt by the lead shock.

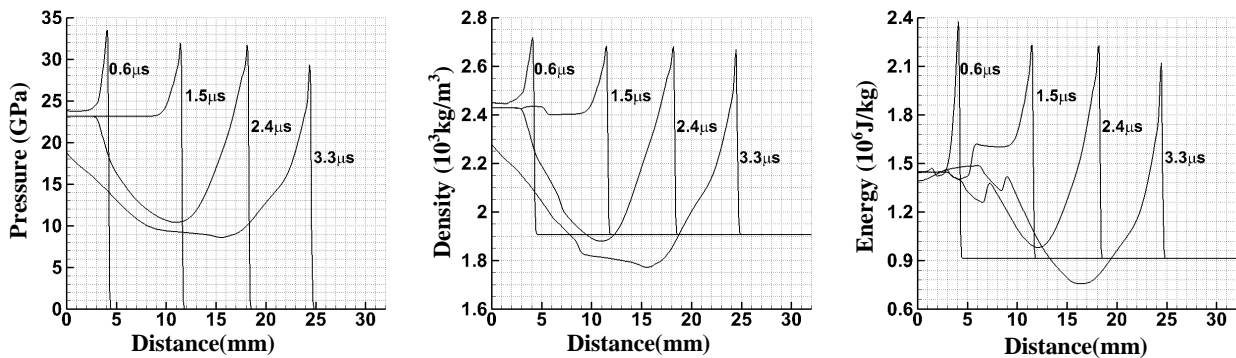
At 1.65 $\mu$ s, the influencing regions domain centred around the corner, and the low-pressure region, low-density region have appeared. Comparing the contour of the reaction rate with the contours of the pressure and the density, we can find that the shock wave continues to propagate in the unreacted explosive, leaving behind unreacted explosive. The dead zone, which is a region of unreacted explosive, generally forms around the right-angle turn. At this moment, the dead zone appears primarily behind the lead shock wave. Then the factor of the transverse shock wave along the inner wall downwards plays a leading role, and the extending curved detonation wave have just extended alightly. In the center position of the model section, there is still the steady detonation wave propagating forward along the central axis of the model. At 3.3 $\mu$ s, as the detonation wave propagates forward and extends to the sides of the head shock wave, the low-pressure region and the low-density region have become increasingly large. Along the inner wall, the lead shock wave has reached the edges of the cylinder, and the reflection can be seen. However, lots of unreacted explosive is left behind the lead shock wave. It demonstrated the failure exits, and the failure/dead zone has been shown in the figure 1 as the region A. It looks like a turnip shape which has a good consistent with the experimental results[Souers *et al.* (2006)].



**Fig. 3. Plots of pressure, density and the reaction rate along the inner wall at different moments**



As showed the plot of pressure in figure 3, at  $0.97\mu\text{s}$ , the detonation wave just arrived at the corner position, and the transverse shock wave causes pressure changes only within  $0.5\text{mm}$  along the inner wall. The maximum is  $9\text{GPa}$  and rapidly decreased to 0. The lead shock wave propagates  $4.1\text{mm}$  and  $7.2\text{mm}$  along the inner wall at  $1.75\mu\text{s}$  and  $2.53\mu\text{s}$  respectively. At  $3.31\mu\text{s}$ , the lead shock wave has reached the edges of the cylinder. The low-pressure region can be clearly shown by the plot of pressure at different moment. And again, the low-density region can be seen from the plot of density, and we can find, in the low-density region, the density has fallen below the initial density. From the plot of the reaction rate, at  $3.31\mu\text{s}$ , we can see that the reaction progress will decline to 0 directly, which suggests only slight retonation generating. With time lapsing, the re-initiation is failure. In the next moment, the extending curved detonation wave reaches the inner wall to re-initiate, otherwise it is detonation failure along the inner wall.



**Fig. 4. Plots of pressure, density and total energy along the the central axis of the model at different moments**

The figure 4 gives the corner-turning influence on the pressure, density and energy in the position of the central axis. After the detonation wave has reached the corner, the detonation wave in the axis position has little effect, and it still keep steady detonation to propagate forward. At  $2.4\mu\text{s}$  and  $3.3\mu\text{s}$ , because of expansion, the troughs appear behind the detonation wave. The trailing expansion region shows some effects of the signals of expansion emanating from the corner. By contrast, the evolution along the wall is more complex, as discussed in detail above.

## Conclusions

In this paper, Fifth-order WENO scheme and third-order TVD Runge-Kutta method are employed to investigate the process of the corner turning of LX-17. The parallel high resolution code is developed. By calculation and detailed analysis for the corner-turning of the double cylinder geometries, we can obtain the following conclusions:

(1) As showed in the example, using the LX-17 explosive, the detonation failure happens along the inner wall, and there is no re-initiation subsequently within a radius of  $9.6\text{mm}$ . The transverse shock wave plays little direct role for initiation explosive which swept by the lead shock wave. In addition, the extending curved detonation wave needs a long time to reach the inner wall. Because of the two factors, there is a great proportion of unreacted explosive remained, forming a large area of dead zone.

(2) The low-pressure, low-density regions are formed in the vicinity of the corner, in which the explosive will react slowly in the subsequent steps, but the released energy cannot support the transverse shock wave propagation forward to re-initiate explosive.

## References

- Souers, P. C., Andreski, H. G., Cook III, C. F., Garza, R., Pastrone, R., Phillips, D., Roeske, F., Vitello, P., Molitoris, J. D. (2004) LX-17 Corner Turning, *Propellants, Explos., Pyrotech* **29**, 359-367.
- Cox, M. and Campbell, A. W. (1981) *Corner-turning in TATB* (No. LA-UR-81-739; CONF-810602-13). Los Alamos Scientific Lab., NM (USA).
- Ferm, E. N., Morris, C. L., Quintana, J. P., Pazuchanic, P., Stacy, H., Zumbro, J. D., Hogan, G. and King, N. (2002, July) Proton radiography examination of unburned regions in PBX 9502 corner turning experiments. *In AIP Conference Proceedings* (No. 2, pp. 966-972). IOP INSTITUTE OF PHYSICS PUBLISHING LTD.
- Mader, C. L., Zumbro, J. D., Ferm, E. N., Gittings, M. and Weaver, R. (2002) Proton radiographic and numerical modeling of colliding, diverging PBX 9502 Detonations. Paper presented at the 12th International Detonation Symposium, August 11–16, San Diego, CA.
- Kapila, A. K., Schwendeman, D. W., Bdzil, J. B. and Henshaw, W. D. (2007) A study of detonation diffraction in the ignition-and-growth model. *Combust. Theor. Model.* **11**(5): 781-822.
- Tarver, C. M. (2010) Corner turning and shock desensitization experiments plus numerical modeling of detonation waves in the triaminotrinitrobenzene based explosive LX-17. *The Journal of Physical Chemistry A* **114**(8): 2727-2736.
- Yang, G., Fu, Y., Hu, D. and Han, X. (2013). Feasibility analysis of SPH method in the simulation of condensed explosives detonation with ignition and growth model. *Computers & Fluids* **88**: 51-59.
- Yang, G., Han, X. and Hu, D. (2011). Computer simulation of two-dimensional linear shaped charge jet using smoothed particle hydrodynamics. *Engineering Computations* **28**(1): 58-75.
- Hu, D., Liu, C. H., Xiao, Y. H. and Han, X. (2014). Analysis of explosion in concrete by axisymmetric FE-SPH adaptive coupling method. *Engineering Computations* **31**(4): 758-774.
- Jiang, G. S. and Shu, C. W. (1996). Efficient implementation of weighted ENO schemes. *J. Comput. Phys* **126**(1): 202-228.
- Zhang, G. R. and Chen, D. N. (1991). *Dynamics of Condensed Explosives Initiation Detonation*. National Defense Industry Press, Beijing.
- Souers, P. C., Andreski, H. G., Batteux, J., Bratton, B., Cabacungan, C., Cook, C. F., ... and Vitello, P. (2006) Dead Zones in LX - 17 and PBX 9502. *Propellants, Explosives, Pyrotechnics* **31**(2), 89-97.

# Three Dimensional High Order Parallel Investigations on Underwater Explosion

†C. Wang, \*J. X. Ding, W. J. Wang and T. Li

State Key Lab of Explosion Science and Technology, Beijing Institute of Technology,  
Beijing, 100081, China

\*Presenting author: dingjianxu@163.com

†Corresponding author: wangcheng@bit.edu.cn

## Abstract

Based on the double shockwave approximation procedure and combining RGFM procedure with level-set method, a local Riemann problem for strong nonlinear equations of state such as JWL equation of state was constructed and then solved to suppress successfully the numerical oscillation caused by high-density ratio and high-pressure ratio across the explosion products and water interface. A fifth order finite difference WENO scheme and the third order TVD Runge-Kutta method were utilized for spatial discretization and time advance, respectively. A novel enclosed type MPI-based parallel methodology for RGFM procedure on uniform structured meshes was presented to realize the parallelization of the three dimensional RGFM-based code for underwater explosion, which had dramatically improved the practical scale of computing model. The overall process of three dimensional bubble pulsations generated by underwater explosion of both TNT and aluminized explosives was successfully simulated with high order numerical scheme. The peak overpressure at different locations of three dimensional underwater explosion for both explosives mentioned above was monitored and analyzed for revealing the influence of aluminum powder combustion on peak overpressure of explosion wave. The numerical results obtained indicated that the attenuation of explosion wave formed by aluminized explosives was slower than that caused by TNT. The influences of aluminum powder combustion on bubble pulsations were also investigated by comparing TNT with aluminized explosives.

**Key words:** Underwater explosion; WENO scheme; RGFM; Local Riemann solver; Parallel computation

## Introduction

Underwater explosion is one of the typical multi-medium problems, in which explosion flow usually consists of detonation products and water. It is obvious that the sharp medium interface with high-density ratio and high-pressure ratio separates the detonation products from water. In the process of numerical simulation on underwater explosion problem, because of the abrupt change with density and pressure close to the interface, the unphysical numerical oscillation is very easy to occur in the neighborhood of the interface. Meanwhile, as the continuous upgrade of explosives, density ratio and pressure ratio increase constantly and greatly. Therefore, tracking and treating the strong discontinuous nonlinear interface of underwater explosion becomes a gradually popular topic and receives considerable attentions in this field.

As for interface tracking technique, many scholars had presented some efficient ways over the past few decades. With particle-in-cell [Amsden (1966)] approach based on the rectangular mesh, the numerical results for the flow field and sharp material interface were achieved by tracking the particle. MAC method [Harlow and Welch (1965)] was often used to track the interface movement and the flow field evolution.

Level-set method [Sethian (1996); Adalsteinsson and Sethian (1999)] used the Hamilton-Jacobi equations to describe the moving interface, which was tracked for later time as the zero level set of the smooth signed distance function instead of the explicit function. Thus, some cases with complex interface, such as crisscross, torsion and separation, can be easily treated.

Fedkiw [Fedkiw et al. (1999)] presented the Ghost Fluid Method (GFM), which can obtain excellent results in treating the interaction between the weak shock wave and interface. Under the condition of strong shock wave, it, however, may get fake physical solution. Liu [Liu et al. (2003)] had put forward the MGFM procedure and a local Riemann problem solver, where the states of ghost fluid across the interface for each phase were defined. Wang [Wang et al. (2006)] presented the RGFM procedure. According to the states of medium across the interface, a local Riemann problem was constructed at first. The solution obtained was then used to redefine the flow states for not only real fluid grids next to the interface but also ghost fluid grids. Thus fewer errors were introduced by RGFM procedure. Considering the impact of the explosion on the interface with high-density ratio and high-pressure ratio and different equations of state, Zhao [Zhao et al. (2013)] proposed a novel interface treatment by combining the original GFM with RGFM procedure.

It can be observed that, since the GFM, MGFM and RGFM were presented, the numerical simulations for multi-medium flow had been applied widely. Simulating the underwater explosion based on the GFM in the attainable literatures generally used the stiff-gas equation of state for explosion products and water, while the complex EOS like JWL was seldom used. Meanwhile, RGFM-based simulation by splitting a multi-medium problem into single medium flows, usually requires that the computer hardware configuration is extremely rigor to satisfy a large-scale calculation. In general, serial computation can't meet the needs of three dimensional large-scale engineering applications.

In this paper, based on double shockwave approximation procedure, we first presented in detail a technology to solve the local Riemann problem with the complex equations of state such as JWL equation of state used in underwater explosion simulations. It effectively avoided the unphysical oscillation occurring at the multi-medium interface in explosion flow field. Combining RGFM method formally transformed multi-medium flow field into pure flow fields with enclosed type parallelization module, the mechanism of three dimensional underwater explosion and the evolutionary laws of bubble pulsation were obtained by using the fifth order finite difference WENO scheme on uniform structured meshes.

## 1. Governing equations

The equations governing three dimensional underwater explosion can be written as

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} + \frac{\partial H(U)}{\partial z} = S, \quad (1)$$

$$U = (\rho, \rho u, \rho v, \rho w, \rho E)^T,$$

$$F = (\rho u, \rho u^2 + p, \rho uv, \rho uw, (\rho E + p)u)^T,$$

$$G = (\rho v, \rho uv, \rho v^2 + p, \rho vw, (\rho E + p)v)^T,$$

$$H = (\rho w, \rho uw, \rho vw, \rho w^2 + p, (\rho E + p)w)^T,$$

$$S = (0, 0, 0, 0, 0)^T,$$

where  $\rho$ ,  $p$  denote the density and pressure, respectively.  $u$ ,  $v$  and  $w$  are the velocity components, and  $E$  is the total energy per unit mass.

The total energy generally consists of internal energy and kinetic energy as follows

$$E = \frac{1}{2}(u^2 + v^2 + w^2) + e, \tag{2}$$

here  $e$  means the internal energy.

To close the above governing equations, the corresponding equations of state for explosion products and water must be introduced.

The explosion products are usually described by the JWL equation of state, which can be expressed in the following form as

$$p = A\left(1 - \frac{\omega\rho}{R_1\rho_0}\right)e^{-\frac{R_1\rho_0}{\rho}} + B\left(1 - \frac{\omega\rho}{R_2\rho_0}\right)e^{-\frac{R_2\rho_0}{\rho}} + \omega\rho e, \tag{3}$$

where the parameters  $A$ ,  $B$ ,  $R_1$ ,  $R_2$ ,  $\omega$  and  $\rho_0$  are material constants of detonation products, which are specifically shown in the following table.

**Table 1. JWL EOS parameters for explosion products**

$\rho_0(\text{kg/m}^3)$	A(MPa)	B(MPa)	C(MPa)	$R_1$	$R_2$	$\omega$
1630.0	373800.0	3747.0	734.0	4.15	0.9	0.35

The compressible water enclosing the explosion product is also our focus. So the stiff-gas equations of state must be required and could be written as

$$p = (\gamma - 1)\rho e - \gamma B, \tag{4}$$

where  $\gamma$  and  $B$  shown in Table 2 are constants.

**Table 2. Stiff-gas EOS parameters for water**

$\gamma$	B(MPa)
7.15	331.0

For aluminized explosives, Miller mathematical model describing the combustion of aluminum powder can be written as

$$\frac{d\lambda}{dt} = a(1 - \lambda)^{\frac{1}{2}} p^{\frac{1}{2}}, \tag{5}$$

where  $\lambda$  is the reaction process variable characterizing the aluminum powder reaction degree in the process of aluminum powder combustion and  $a$  is a material constant.

## 2. The local Riemann problem for underwater explosion

After confirming the specific locations of the explosion products and water interface by advancing implicit Level-set function, RGM-based multi-medium interface treatment should be utilized to change a multi-medium problem into two single medium problems. A local Riemann problem is constructed first at the interface and then solved. In the RGM method, the predicted interface states solved by Riemann problem are assigned to the real fluid nodes just close to the interface in the real fluid. Then normal constant extrapolation by solving extension equation is used to acquire the values of three ghost fluid nodes required for high order WENO scheme. Next, the particular solution to local Riemann problem in underwater explosion will be described briefly.

Taking one dimensional case as an example, with the help of mass conservation equation and momentum conservation equation, the relationship between the velocity and pressure of the interface can be achieved as follows

$$u_* = F_1(p_*) = u_L - \sqrt{(P_* - P_L) \left( \frac{1}{\rho_L} - \frac{1}{\rho_{*L}} \right)}, \quad (6)$$

$$u_* = F_2(p_*) = u_R + \sqrt{(P_* - P_R) \left( \frac{1}{\rho_R} - \frac{1}{\rho_{*R}} \right)}, \quad (7)$$

here  $\rho_{*L}$  and  $\rho_{*R}$  are the densities of both sides of interface, which are unknown quantities for the time being and also need to be solved.

The nonlinear function relationship on interfacial pressure is obtained by the above two equations, which can be written as

$$F(p_*) = F_1(p_*) - F_2(p_*) = 0. \quad (8)$$

Combining energy conservation equation with respective equations of state, the implicit function with respect to  $\rho_{*L}$ ,  $\rho_{*R}$  and  $p_*$  can be deduced and then given as

$$e_L(p_L, \rho_L) - e_{*L}(p_*, \rho_{*L}) - \frac{1}{2}(p_L + p_*) \frac{\rho_L - \rho_{*L}}{\rho_L \rho_{*L}} = 0, \quad (9)$$

$$e_{*R}(p_*, \rho_{*R}) - e_R(p_R, \rho_R) - \frac{1}{2}(p_* + p_R) \frac{\rho_{*R} - \rho_R}{\rho_{*R} \rho_R} = 0. \quad (10)$$

For the local Riemann problem composing of JWL equations of state and Stiff-gas equations of state, the brief process of solving it by Newton iterative method is described below. It is supposed that, for one dimensional Riemann problem, gaseous

detonation products locate at the left hand side of the interface and the water is on the opposite side. Obviously, the Eq. (8-10) is now a closed nonlinear equation consisting of interfacial pressure and both sides of densities close to interface as unknown variables. The classical Newton iterative method is adopted to solve the problem of interest, and it can be written as

$$P_{*(n+1)} = P_{*(n)} - \frac{F_1(P_{*(n)}) - F_2(P_{*(n)})}{F_1'(P_{*(n)}) - F_2'(P_{*(n)})}. \quad (11)$$

The appropriate guess value of the interfacial pressure for (11) is necessary, and should be selected at first. If the value of  $|P_{*(n+1)} - P_{*(n)}|$  is equal or less than a given threshold, Newton iteration will be stop automatically, and the final result  $P_{*(n+1)}$  is the interfacial pressure  $p^*$  to be sought. Thus, the velocity and both sides of densities can be obtained by simultaneously solving the Eq. (6), (9) and (10). At this point, the solution of the local Riemann problem in underwater explosion has been deduced completely.

### 3. Parallelization for RGFM procedure

For a general difference scheme, it is only required to communicate data in current process with adjacent processes in the orthogonal directions, that is to say, the processes in the diagonal directions are not necessary. For the WENO scheme combining with RGFM for underwater explosion, it is slightly different. If the interface is very close to any corner of the current computational process, information such as pressure and density at some nodes of diagonal processes will be inevitably used in the construction of the local Riemann problem. Considering such special requirements for RGFM method to implement data communication, a novel fully enclosed type data communication mode is presented, in which the enclosed communication boundaries are built on the periphery of computational domain of any process.

**Table 3. Parallelization for RGFM**

Process number of receive buffer	Receive buffer number	Send buffer number	Process number of send buffer
n	1	4	n-x-1
n	2	3	n-x+1
n	3	2	n+x-1
n	4	1	n+x+1
n	5	6	n-1
n	6	5	n+1
n	7	8	n-x
n	8	7	n+x

The above table takes the two dimensional case as an example to illustrate the presented enclosed type parallel method. It is obvious that the method can be directly extended to three dimension underwater explosion, and has been implemented in this paper.

**4. Numerical investigations on three dimensional underwater explosion**

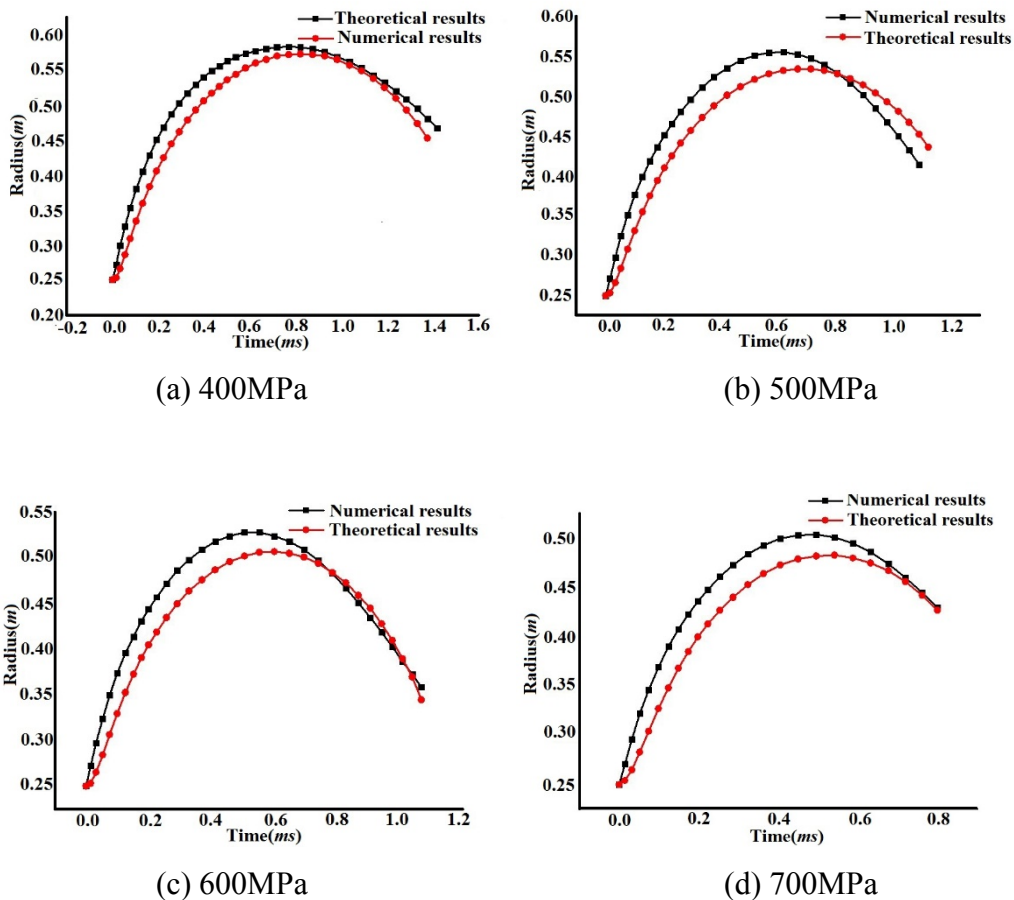
**4.1 Validation**

According to the one dimensional fluid unsteady motion theory, the law of bubble pulsation can be given as

$$R_b'' = \frac{27}{64} P_H \frac{r^9}{\rho R_b^{10}} - \frac{P_\infty}{\rho R_b} - \frac{3}{2} \frac{R_b'^2}{R_b} \tag{12}$$

The corresponding initial condition has the following form

$$\begin{cases} R_b|_{t=0} = r \\ R_b'|_{t=0} = 0 \end{cases} \tag{13}$$



**Figure 1. The comparisons of numerical results with theoretical results on bubble radius under different pressures**



The numerical results of bubble pulsation using RGFM procedure are compared with the theoretical results, which is shown in Fig.1. The maximum bubble radius and its corresponding time are shown in Table 4 and Table 5, respectively. It is evident from Fig.1 and two tables that the numerical results obtained are consistent with the theoretical results, and that the errors of maximum bubble radius and its time don't exceed 10%. So, the RGFM-based high order procedure presented in this paper can exactly describe the whole process of bubble pulsation.

**Table 4. The comparisons of numerical results of maximum bubble radius with theoretical results**

Pressures	Theoretical results(m)	Numerical results(m)	Relative errors
400MPa	0.572	0.583	1.9%
500MPa	0.534	0.555	3.9%
600MPa	0.506	0.527	4.2%
700MPa	0.483	0.504	4.3%

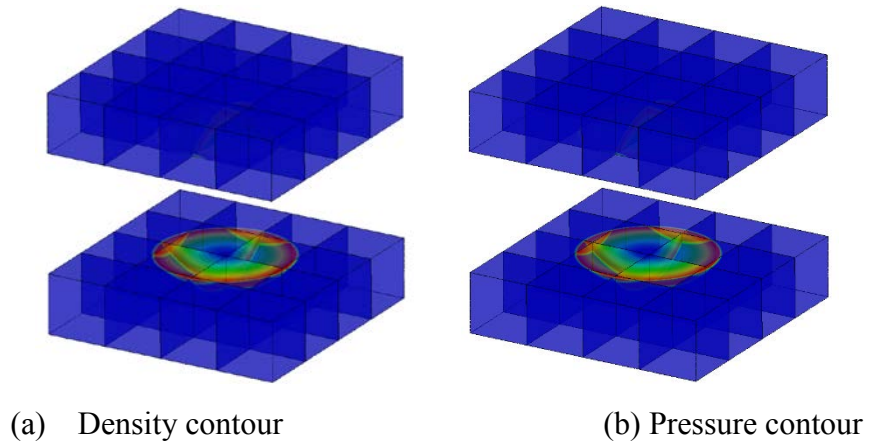
**Table 5. The comparisons of numerical results of the maximum radius time with theoretical results**

Pressures	Theoretical results(m)	Numerical results(m)	Relative errors
400MPa	0.819	0.768	-6.2%
500MPa	0.668	0.618	-7.5%
600MPa	0.602	0.554	-8.0%
700MPa	0.541	0.495	-8.5%

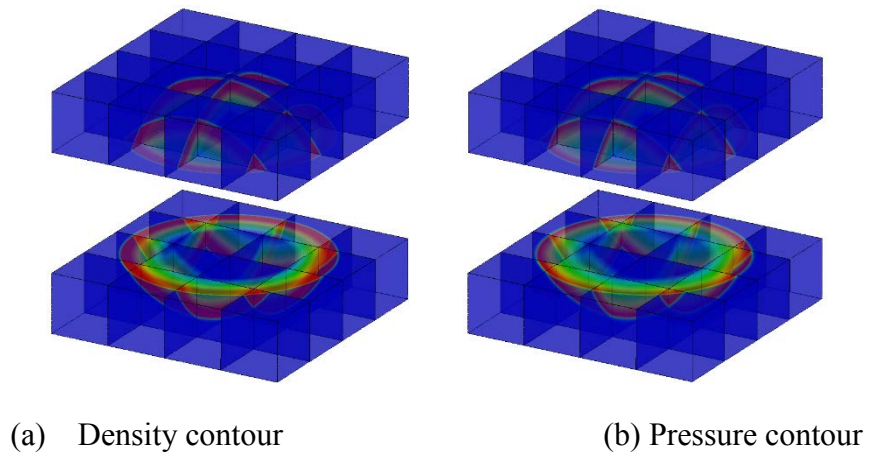
#### ***4.2 Numerical investigations on bubble pulsation of TNT explosive***

The TNT explosive is used with the radius of 0.25m for underwater explosion. Water pressure is 400MPa, 500MPa and 600 MPa, respectively. Apparently, the selected pressure exceeds the usual pressure of underwater explosion. The main purpose is to save the computing time without damaging the essential laws of bubble pulsation. The 64 processes are employed to compute this problem with up to 13.824 million grids. The size of grids is 0.02m×0.02m×0.02m, while the whole physical domain is 4.8m×4.8m×4.8m.

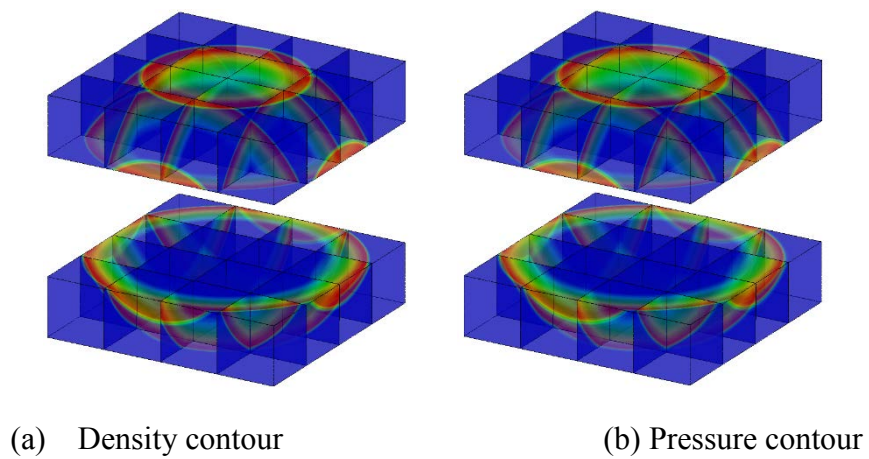
The computational results at some typical moments are shown in Fig. 2-4. The contour of density and pressure are very symmetrical and fake nonphysical oscillation doesn't happen. The distribution of density and pressure in the underwater explosion field can be reflected precisely.



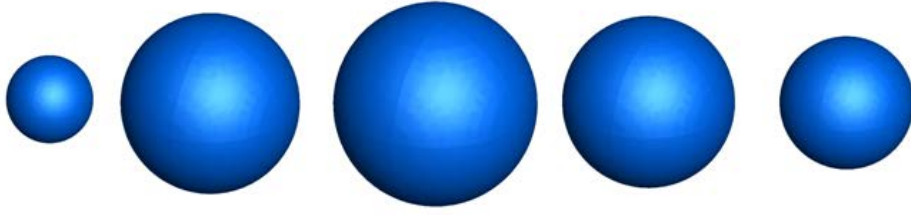
**Figure 2. Numerical results at 0.531ms**



**Figure 3. Numerical results at 0.744ms**

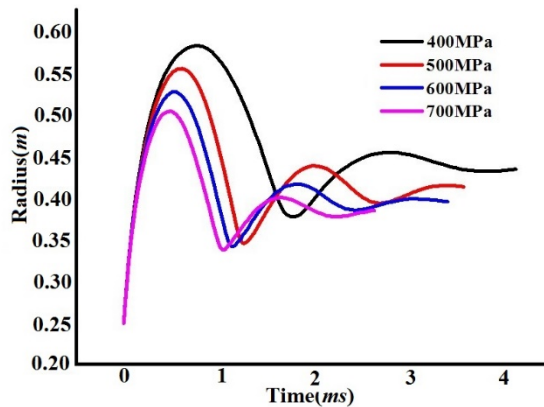


**Figure 4. Numerical results at 0.962ms**



**Figure 5. Numerical results on the evolution of bubble pulsation**

The numerical results shown in Fig. 5 indicate that combining the local Riemann solver with RGFM procedure can exactly simulate the full process of underwater explosion and bubble pulsation, and the interface of the computed bubble consistently maintains smooth without any numerical oscillation.



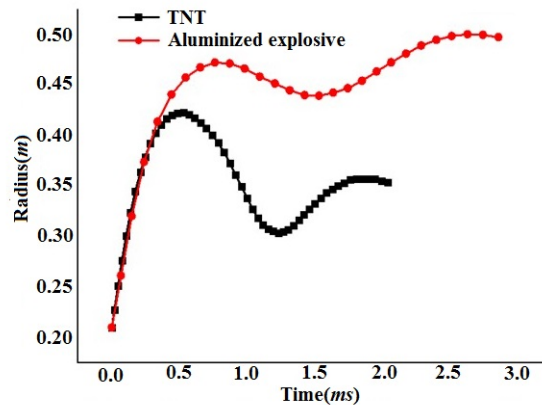
**Figure 6. The evolution of bubble radius under the different pressure conditions**

Fig. 6 shows that maximum bubble radius decreases with the increase of pressure. The maximum radius is 0.583m when water pressure is 400MPa. When its pressure is 700MPa, the maximum radius reduces by 13.6% and decreases to 0.504m. Bubble pulsation period also decreases with the increase of pressure. The pulsation period is 1.749ms when pressure is 400MPa. When pressure is 700MPa, the period reduces by 40.4% and reduce to 1.043ms. So, the influence of water pressure on bubble pulsation period is more obvious.

The numerical results reveal that the physical essence of expansion and contraction of the bubble interface is the movement of contact discontinuity for local Riemann problem at the interface, and the pressure, normal speed and density close to the interface determine the evolution of bubble interface. Thus, solving local Riemann problem is an effective way to quantitatively reveal the formation mechanism of bubble pulsation. Solving local Riemann problem can obtain accurate flow characteristics near the interface and the evolution of bubble interface in the flow field. The complete process of bubble pulsation is captured nicely. Therefore, the unique advantages based on RGFM procedure and local Riemann solver are obvious to simulate bubble pulsation in underwater explosion.

### 4.3 Numerical investigations on bubble pulsation of aluminized explosives

The aluminized explosives is simplified to gas detonation products containing unreacted aluminum powder. The stiff-gas equation of state is used for water. The initial pressure is 500 MPa, and the radius of explosive is 0.21m. The 64 processes are employed for this problem with 13.824 million grids, and the size of grids is 0.02m×0.02m×0.02m, while the whole physical domain is also 4.8m×4.8m×4.8m.



**Figure 7. The comparison of aluminized explosive with TNT on bubble pulsation radius under 500 MPa water pressure**

It is shown in Fig. 7 that the expansion processes of two kinds of explosive are almost the same at the initial stage and the radius-time curve is substantially coincident. But, at the late stage, the appearances of two kinds of explosives are very different. First, the maximum bubble radius and its evolutionary period of aluminized explosives are significantly greater than those of TNT. Second, because explosion wave propagation causes energy loss, the maximum radius for TNT at the later stage becomes small. For aluminized explosives, the maximum radius during the second period is even greater than that of the previous period due to the secondary energy release formed by the combustion of aluminum powder supplementing the energy losses to some extent.

### Conclusions

Combining RGFM multi-medium interface treatment method with fifth order finite difference WENO scheme, the large scale underwater explosion parallel simulations are performed by the novel Riemann problem solver presented for complex nonlinear equations of state and MPI-based enclose type parallel module. The three dimensional whole physical processes of bubble pulsation in underwater explosion are numerically investigated. The important mechanism of underwater explosion is also revealed as follows:

1) Solving local Riemann problem defined by detonation products and water can essentially depicts the motion process of bubble surface caused by both sides of the flow field, by which the complete process of bubble pulsation is captured exactly. The

maximum radius and its corresponding time of bubble obtained by numerical simulations are agreed well with the theoretical results;

2) For bubble pulsation in underwater explosion, the physical essence of expansion and contraction of bubble surface is the propagation of contact discontinuity of local Riemann problem at the explosion products and water interface. The interfacial states including pressure and normal speed at the interface and density on both sides of the interface play a key role in determining the motion behavior of bubble surface;

3) For the same charge configuration, with the increase of water pressure, both maximum bubble radius and its pulsation period decrease. Nevertheless, the influence of water pressure on bubble pulsation period is more obvious than maximum bubble radius;

4) The secondary energy release for aluminized explosives can continuously add the energy loss caused by explosion wave propagation in water. So in the subsequent process of bubble pulsation, the radius and pulsation period for aluminized explosives are obviously larger than those of TNT charge. Maximum radius of aluminized explosive in the process of the second pulsation even exceeds the first pulsation process.

### Acknowledgement

This work is supported by the National Natural Science Foundation of China under grant number 11325209 and 11272056, the Specialized Research Foundation for the Doctoral Program of Higher Education of China under grant 20121101110004, Beijing Natural Science Foundation (8132034)

### References

- Amsden, A. A. (1966) The particle-in cell method for the calculations of the dynamics of compressible fluids, *Los Alamos Scientific Laboratory Report*, LA-3466.
- Harlow, F. H. and Welch, J. E. (1965) Numerical calculation of time dependent viscous incompressible flow of fluid with free surface, *Physics of Fluids* **8**, 2182-2189.
- Sethian, J. A. (1996) Theory, algorithms, and applications of level set methods for propagating interfaces, *Acta Numerica* **5**, 309-395.
- Adalsteinsson, D. and Sethian, J. A. (1999) The fast construction of extension velocities in level set methods, *Journal of Computational Physics* **148**, 2-22.
- Fedkiw, R. P., Aslam, T., Merriman, B. and Osher, S. (1999) A non-oscillatory Eulerian approach to interfaces in multimaterial flows (the ghost fluid method), *Journal of Computational Physics* **152**, 457-492.
- Liu, T. G., Khoo, B. C. and Yeo, K. S. (2003) Ghost fluid method for strong shock impacting on material interface, *Journal of Computational Physics* **190**, 651-681.
- Wang, C. W., Liu, T. G. and Khoo, B. C. (2006) A real ghost fluid method for the simulation of multimediuim compressible flow, *Siam Journal on Scientific Computing* **28**, 278-302.
- Zhao, H. T. and Wang, C. (2013) High resolution numerical simulation of air explosion, *Acta Armamentarii* **34**, 1536-1546.

## Authors Index

Name	Page No
<i>Ahmadi, Goodarz</i>	847
<i>Alam-Chowdhury, M.K.</i>	912
<i>Ali, Amjad</i>	24
<i>Amaya, K.</i>	543
<i>Amirpour, M.</i>	712
<i>Ananthasuresh, G.K.</i>	384
<i>Andò, L.</i>	749
<i>Ansari, Mohammad Sharifuddin</i>	74
<i>Archer, Rosalind</i>	816
<i>Asao, Shinichi</i>	503, 588
<i>Babaei, Meisam</i>	847
<i>Bagde, Ashutosh D.</i>	356
<i>Bai, Ruixiang</i>	696
<i>Banerjee, A.</i>	156, 912
<i>Bao, S.H.</i>	696
<i>Battley, Mark</i>	624
<i>Becker, Wilfried</i>	837
<i>Behl, Ramandeep</i>	362
<i>Bellini, V.</i>	787
<i>Bhumkar, Y.G.</i>	175
<i>Brown, Richard</i>	847
<i>Bui, Ha H.</i>	478
<i>Burbery, N.J.</i>	62
<i>Cai, Xueming</i>	866
<i>Calius, E.P.</i>	156
<i>Campbell, Stephen</i>	939
<i>Capogni, M.</i>	787
<i>Celona, Luigi</i>	749
<i>Chacón, Jorge L.</i>	736
<i>Chai, Yingbin</i>	526
<i>Chassaing, J.C.</i>	725
<i>Chen, Dyi-Cheng</i>	675
<i>Chen, Hu</i>	976
<i>Chen, Wei</i>	702
<i>Chen, Xiang-Zi</i>	11

Name	Page No
<i>Chen, Youming</i>	624
<i>Chen, Zhiyi</i>	702
<i>Chiba, Kazuhisa</i>	117
<i>Chines, F.</i>	749
<i>Chiu, Wing Kong</i>	617, 668
<i>Chiu, Zelia Karmen</i>	668
<i>Cho, Maenghyo</i>	273
<i>Ciavola, G.</i>	749
<i>Cisbani, E.</i>	787
<i>Cokljat, Davor</i>	956
<i>Costa, G.</i>	749
<i>Dahake, Sandeep W.</i>	356
<i>Das, Raj</i>	62, 156, 624, 634, 712, 809, 939, 951
<i>Davis, Lisa</i>	884
<i>De Costa, C.</i>	492
<i>Deng, Qiong</i>	580
<i>Dhandole, S.D.</i>	175
<i>Ding, Jianxu</i>	1002
<i>Dong, Haiping</i>	515
<i>Dubey, R.N.</i>	912
<i>Fan, Hui</i>	702
<i>Fan, Jizhuang</i>	168
<i>Fang, Jianguang</i>	570
<i>Ferguson, W.G.</i>	62
<i>Fernandez, Justin W.</i>	634, 809
<i>Gammino, S.</i>	749
<i>Gao, Renjing</i>	535
<i>Gao, Yongsheng</i>	168
<i>Garcia, Luis F.</i>	736
<i>Gáspár, C.</i>	547
<i>Gelet, Rachel</i>	216
<i>Gerasimov, A.V.</i>	186
<i>Ghoniem, N.</i>	62

<i>Gonçalves, Kleber A.</i>	94	<i>Lee, Jaehun</i>	273
<i>Gong, Zhixiong</i>	526	<i>Leonardi, Ornella</i>	749
<i>Gu, Yuantong</i>	725, 766	<i>Li, Eric</i>	263, 986
<i>Guo, Yongchang</i>	137	<i>Li, G.Y.</i>	986
<i>Gupta, Raj K.</i>	324	<i>Li, Mingli</i>	535
<i>He, Jian-jun</i>	11	<i>Li, Peng,</i>	647
<i>He, Zhicheng</i>	263, 986	<i>Li, Qing</i>	570
<i>Hell, Sascha</i>	837	<i>Li, Shouju</i>	16
<i>Hou, Ming</i>	647	<i>Li, T.</i>	1002
<i>Hou, Peng</i>	515	<i>Li, Wei</i>	526
<i>Huang, Jia</i>	559	<i>Li, Xiao-ting</i>	597
<i>Huang, T.H.</i>	971	<i>Li, Yulong</i>	559
<i>Huang, Ting</i>	606	<i>Liang, X.H.</i>	696
<i>Huang, Wenhua</i>	461	<i>Lin, Frank</i>	233
<i>Huynh, B.P.</i>	255	<i>Lin, G.</i>	413
<i>Hwang, J.-S.</i>	199	<i>Lin, X.</i>	86
<i>Imamura, Taro</i>	117	<i>Lin, Xue-yang</i>	647
<i>Inomoto, Takeshi</i>	503, 588	<i>Liu, Gangfeng</i>	168
<i>Inthavong, Kiao</i>	847	<i>Liu, J.Y.</i>	413
<i>Ishihara, Sadanori</i>	503, 588	<i>Liu, Jian</i>	137
<i>Islam, Mohammad S.</i>	766	<i>Liu, Jun</i>	211, 559
<i>Jaramillo, Julian E.</i>	736	<i>Liu, L.M.</i>	679
<i>Jemcov, Aleksandar</i>	331, 956	<i>Liu, Shuangyan</i>	580
<i>Jia, Zupeng</i>	211	<i>Liu, X.Y.</i>	696
<i>Jiang, Cho-Pei</i>	971	<i>Liu, Xinqiao</i>	995
<i>Jog, C.S.</i>	384	<i>Liu, Yinli</i>	611
<i>Johnson, Michael</i>	233	<i>Maeda, Kengo</i>	687
<i>Kabir, Sumrin Mehak</i>	471	<i>Malhotra, Akshat</i>	809
<i>Kamran</i>	24	<i>Mammoliti, F.</i>	787
<i>Kanazaki, Masahiro</i>	687	<i>Man, Xianfeng</i>	137
<i>Kanber, B.</i>	418	<i>Marketta, S.</i>	749
<i>Karumbaiah, K.M.</i>	939	<i>Maruszewski, Joseph P.</i>	956
<i>Kawazoe, Mitsuhiro</i>	687	<i>Matsuno, Kenichi</i>	503, 588
<i>Khristenko, Yu F.</i>	186	<i>Matsuno, Takashi</i>	687
<i>Kim, H.</i>	199	<i>Mawale, Mahesh B.</i>	356
<i>Kutle, Abhaykumar M.</i>	356	<i>Mendis, P.</i>	398
<i>Kwon, Eppyje E.</i>	634, 809	<i>Meng, Qingjie</i>	461
<i>Lai, Yuehua</i>	515	<i>Mikalajunas, Mike</i>	279

<i>Miller, Ian</i>	794	<i>Qin, Y.P.</i>	775
<i>Miramini, S.</i>	398	<i>Rahimzadeh, Hassan</i>	847
<i>Mirza, Alina</i>	471	<i>Rajic, N.</i>	617
<i>Mondal, Sabyasachi</i>	42	<i>Ramachandran, Prabhu</i>	762
<i>Motsa, S.S.</i>	74, 362	<i>Re, L.</i>	787
<i>Musico, P.</i>	787	<i>Ricci, Mário César</i>	921
<i>Nagpal, Vinod K.</i>	794	<i>Richardson, M.</i>	398
<i>Nakagawa, S.T.</i>	107	<i>Ristovski, Zoran D.</i>	766, 847
<i>Nandkeolyar, Raj</i>	74	<i>Rognon, Pierre</i>	216
<i>Nandy, Arup</i>	384	<i>Rosalie, C.</i>	617
<i>Narangale, Sachin M.</i>	351	<i>Roychowdhury, A.</i>	384
<i>Neda, Monika</i>	884	<i>Russ, Matthias</i>	668
<i>Ng, C.T.</i>	438	<i>Saavedra Flores, Erick I.</i>	712, 951
<i>Nguyen, Chi T.</i>	478	<i>Saavedra, K.</i>	951
<i>Nguyen, Giang D.</i>	216, 478	<i>Saha, Suvash C.</i>	725, 766
<i>Nguyen, Sy-Ngoc</i>	273	<i>Sakamoto, Jiro</i>	830
<i>Nguyen, Trung-Thanh</i>	35	<i>Sanada, Yasuto</i>	117
<i>Nguyen, Vinh Phu</i>	478	<i>Sauret, Emilie</i>	725, 766
<i>Nicolosi, Dario</i>	749	<i>Sheikh, Shahzad Amin</i>	471
<i>Nie, Xin</i>	986	<i>Shi, Xiaofei</i>	866
<i>Niktash, A.</i>	255	<i>Shin, S. -H.</i>	199
<i>Noto, F.</i>	787	<i>Shinde, G.N.</i>	351
<i>Noto, Francesco</i>	749	<i>Sibanda, Precious</i>	42
<i>Ohashi, K.</i>	107	<i>Sideroff, Chris</i>	331
<i>Omori, Taiga</i>	117	<i>Singh, A.P.</i>	912
<i>Ong, Wern Hann</i>	617, 668	<i>Singh, Michael R.</i>	634
<i>Onishi, Y.</i>	543	<i>Soares Jr., Delfim</i>	94
<i>Pahlevani, Fran</i>	884	<i>Song, Gang</i>	679
<i>Pai, Shantaram S.</i>	794	<i>Song, Yuwang</i>	866
<i>Park, Hong- Seok</i>	35	<i>Spinali, S.</i>	787
<i>Patil, Kunal D.</i>	384	<i>Stephens, Darrin W.</i>	331
<i>Paul, D.K.</i>	912	<i>Su, Mengjia</i>	580
<i>Perrino, R.</i>	787	<i>Su, Peng-hui</i>	193
<i>Po, G.</i>	62	<i>Sugita, K.</i>	107
<i>Pradhan, Jitenjaya</i>	175	<i>Sun, Guangyong</i>	570
<i>Pratap, Rudra</i>	384	<i>Sun, Xiaoying</i>	168
<i>Przekwas, Andrzej J.</i>	324	<i>Sun, Z.C.</i>	413
<i>Puri, Kunal</i>	762	<i>Sutera, M.C.</i>	787



<i>Talebizadeh, Pouyan</i>	847	<i>Xu, Fei</i>	597
<i>Tang, Hao</i>	611	<i>Xu, Wei</i>	830
<i>Tang, Zhengyuan</i>	374, 449	<i>Xu, Xu</i>	263
<i>Tang, Zhongbin</i>	559	<i>Xu, Yihang</i>	580
<i>Taylor, Michael C.</i>	634, 809	<i>Yamakawa, Masashi</i>	503, 588
<i>Telles, José Claudio de Faria</i>	94	<i>Yang, Chunhui</i>	606
<i>Tian, Zerun</i>	16	<i>Yang, Yang</i>	597
<i>Torrise, Giuseppe</i>	749	<i>Yavuz, M.M.</i>	418
<i>Tsai, Meng-Hao</i>	403	<i>Yi, Xiaojian</i>	515
<i>Tu, Chia-An</i>	675	<i>You, Ci-Syong</i>	675
<i>Uddin, Marjan</i>	24	<i>Zang, Mengyan</i>	976
<i>Ushijima, Satoru</i>	492	<i>Zarrouk, Sadiq J.</i>	816
<i>Vallabh, Ravin D.</i>	634	<i>Zhang, Chao</i>	816
<i>Wan, Decheng</i>	374, 449, 461	<i>Zhang, Feng</i>	656
<i>Wang, Cheng</i>	995, 1002	<i>Zhang, Guiyong</i>	986
<i>Wang, De-gui</i>	647	<i>Zhang, Liang</i>	193
<i>Wang, J.</i>	679	<i>Zhang, Lihai</i>	398
<i>Wang, Lu</i>	597	<i>Zhang, P.</i>	413
<i>Wang, W.J.</i>	1002	<i>Zhang, Shudao</i>	211, 445
<i>Wang, Wenxuan</i>	656	<i>Zhang, Y.X.</i>	86, 606, 976
<i>Wang, Y.Y.</i>	413	<i>Zhang, Youlin</i>	374, 449
<i>Wang, Yanjin</i>	445	<i>Zhang, Zhixue</i>	559
<i>Wang, Yongfen</i>	611	<i>Zhao, Jian</i>	535
<i>Wang, Zhen</i>	775	<i>Zhao, Long</i>	580
<i>Wang, Zhenhong</i>	233	<i>Zhou, Changcong</i>	656
<i>Waters, Jiajia</i>	884	<i>Zhou, Haibing</i>	211
<i>Wei, Pengfei</i>	853	<i>Zhu, Yanhe</i>	168
<i>White, Nigel</i>	233	<i>Zhu, Zhi-qi</i>	647
<i>Xi, Ping</i>	866	<i>Zou, A.</i>	725
<i>Xie, Longxiang</i>	137	<i>Zou, Li</i>	775
<i>Xiong, Jun</i>	211		