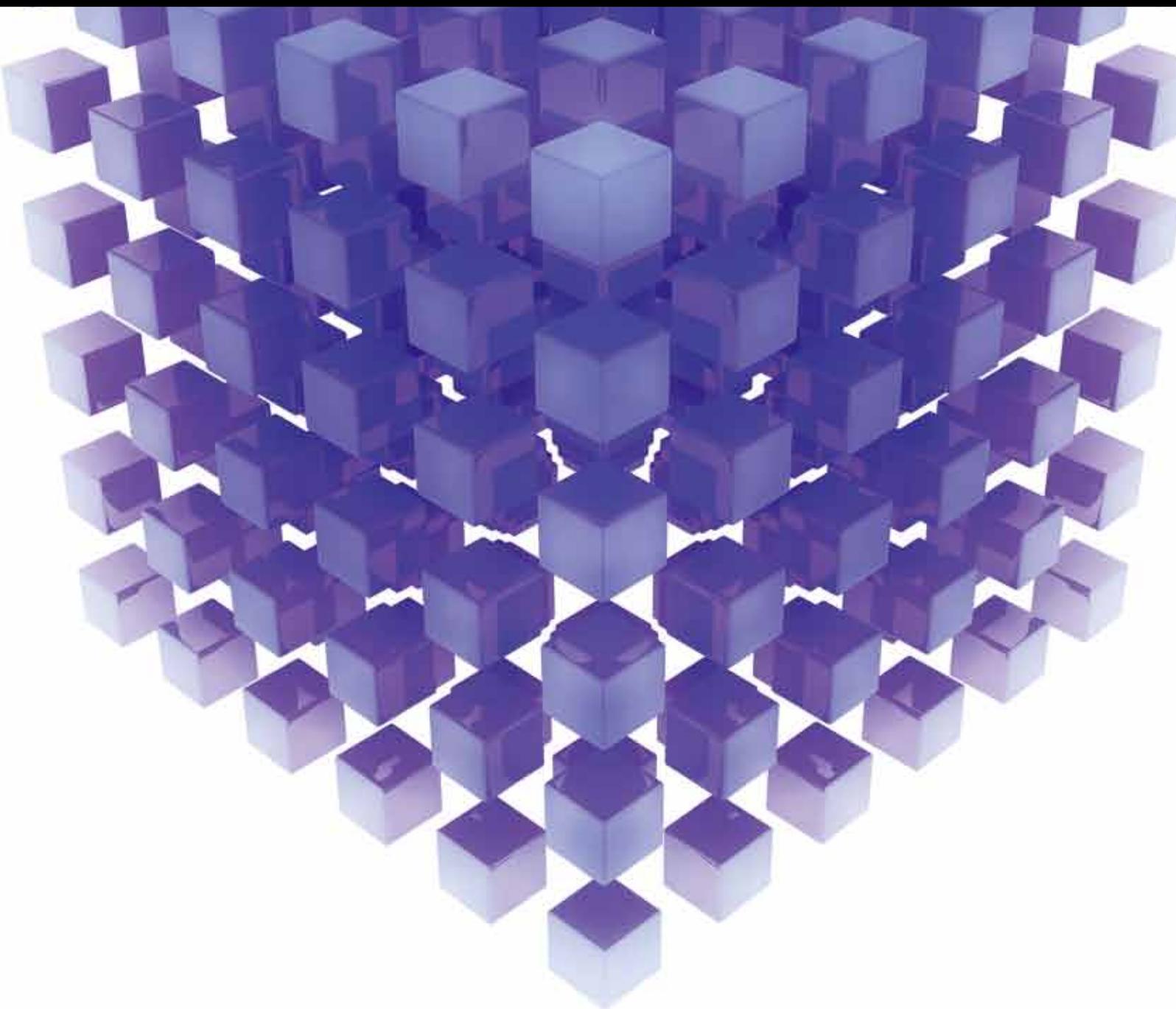


MATHEMATICAL PROBLEMS IN ENGINEERING

RECENT THEORY AND APPLICATIONS ON INVERSE PROBLEMS

GUEST EDITORS: FATİH YAMAN, VALERY G. YAKHNO, AND ROLAND POTTHAST





Recent Theory and Applications on Inverse Problems

Mathematical Problems in Engineering

**Recent Theory and Applications on
Inverse Problems**

Guest Editors: Fatih Yaman, Valery G. Yakhno,
and Roland Potthast



Copyright © 2013 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in “Mathematical Problems in Engineering.” All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

- Mohamed Abd El Aziz, Egypt
Eihab M. Abdel-Rahman, Canada
Rashid K. Abu Al-Rub, USA
Sarp Adali, South Africa
Salvatore Alfonzetti, Italy
Igor Andrianov, Germany
Sebastian Anita, Romania
W. Assawinchaichote, Thailand
Erwei Bai, USA
Ezzat G. Bakhoun, USA
José M. Balthazar, Brazil
R. K. Bera, India
C. Bérenguer, France
Jonathan N. Blakely, USA
Stefano Boccaletti, Spain
Stephane P.A. Bordas, USA
Daniela Boso, Italy
M. Boutayeb, France
Michael J. Brennan, UK
Salvatore Caddemi, Italy
Piermarco Cannarsa, Italy
Jose E. Capilla, Spain
Carlo Cattani, Italy
Marcelo M. Cavalcanti, Brazil
Diego J. Celentano, Chile
Mohammed Chadli, France
Arindam Chakraborty, USA
Yong-Kui Chang, China
Michael J. Chappell, UK
Kui Fu Chen, China
Xinkai Chen, Japan
Kue-Hong Chen, Taiwan
Jyh-Horng Chou, Taiwan
Slim Choura, Tunisia
Cesar Cruz-Hernandez, Mexico
Swagatam Das, India
Filippo de Monte, Italy
Antonio Desimone, Italy
Yannis Dimakopoulos, Greece
Baocang Ding, China
Joao B. R. Do Val, Brazil
Daoyi Dong, Australia
B. Dubey, India
Horst Ecker, Austria
M. Onder Efe, Turkey
Elmetwally Elabbasy, Egypt
A. Elías-Zúñiga, Mexico
Anders Eriksson, Sweden
Vedat S. Erturk, Turkey
Moez Feki, Tunisia
Ricardo Femat, Mexico
Robertt A. F. Valente, Portugal
C. R. Fuerte-Esquivel, Mexico
Zoran Gajic, USA
Ugo Galvanetto, Italy
Furong Gao, Hong Kong
Xin-Lin Gao, USA
Behrouz Gatzmiri, Iran
Oleg V. Gendelman, Israel
Didier Georges, France
Paulo B. Gonçalves, Brazil
Oded Gottlieb, Israel
Fabrizio Greco, Italy
Quang Phuc Ha, Australia
M. R. Hajj, USA
Tony S. Wen Hann, Taiwan
Thomas Hanne, Switzerland
K. R. (Stevanovic) Hedrih, Serbia
M.I. Herreros, Spain
Wei-Chiang Hong, Taiwan
Jaromir Horacek, Czech Republic
Huabing Huang, China
Chuangxia Huang, China
Gordon Huang, Canada
Yi Feng Hung, Taiwan
Hai-Feng Huo, China
Asier Ibeas, Spain
Anuar Ishak, Malaysia
Reza Jazar, Australia
Zhijian Ji, China
Jun Jiang, China
J. J. Judice, Portugal
Tadeusz Kaczorek, Poland
Tamas Kalmar-Nagy, USA
Tomasz Kapitaniak, Poland
Hamid Reza Karimi, Norway
Metin O. Kaya, Turkey
Nikolaos Kazantzis, USA
Farzad Khani, Iran
Kristian Krabbenhoft, Australia
Ren-Jieh Kuo, Taiwan
Jurgen Kurths, Germany
Claude Lamarque, France
Usik Lee, Korea
Marek Lefik, Poland
Stefano Lenci, Italy
Roman Lewandowski, Poland
Shanling Li, Canada
Ming Li, China
Jian Li, China
Shihua Li, China
Teh-Lu Liao, Taiwan
Panos Liatsis, UK
Shueei M. Lin, Taiwan
Yi-Kuei Lin, Taiwan
Jui-Sheng Lin, Taiwan
Yuji Liu, China
Wanquan Liu, Australia
Bin Liu, Australia
Paolo Lonetti, Italy
Vassilios C. Loukopoulos, Greece
Junguo Lu, China
Chien-Yu Lu, Taiwan
Alexei Mailybaev, Brazil
Manoranjan K. Maiti, India
Oluwole D. Makinde, South Africa
Rafael Martinez-Guerra, Mexico
Driss Mehdi, France
Roderick Melnik, Canada
Xinzhu Meng, China
Yuri V. Mikhlin, Ukraine
Gradimir Milovanovic, Serbia
Ebrahim Momoniat, South Africa
Trung Nguyen Thoi, Vietnam
Hung Nguyen-Xuan, Vietnam
Ben T. Nohara, Japan
Sotiris K. Ntouyas, Greece
Gerard Olivar, Colombia
Claudio Padra, Argentina
Bijaya Ketan Panigrahi, India
Francesco Pellicano, Italy
Matjaz Perc, Slovenia
Vu Ngoc Phat, Vietnam
M. do Rosário Pinho, Portugal
A. Pogromsky, The Netherlands

Seppo Pohjolainen, Finland
Stanislav Potapenko, Canada
Sergio Preidikman, USA
Carsten Proppe, Germany
Hector Puebla, Mexico
Justo Puerto, Spain
Dane Quinn, USA
Kumbakonam R. Rajagopal, USA
Gianluca Ranzi, Australia
Sivaguru Ravindran, USA
G. Rega, Italy
Pedro Ribeiro, Portugal
J. Rodellar, Spain
R. Rodriguez-Lopez, Spain
A. J. Rodriguez-Luis, Spain
Ignacio Romero, Spain
Hamid Ronagh, Australia
Carla Roque, Portugal
Rubén R. García, Spain
Manouchehr Salehi, Iran
Miguel A. F. Sanjuán, Spain
Ilmar F. Santos, Denmark
Nickolas S. Sapidis, Greece
E. J. Sapountzakis, Greece
Bozidar Sarler, Slovenia
Andrey V. Savkin, Australia
Massimo Scalia, Italy
Mohamed A. Seddeek, Egypt
A. P. Seyranian, Russia
Leonid Shaikhet, Ukraine

Cheng Shao, China
Bo Shen, Germany
Jian-Jun Shu, Singapore
Zhan Shu, UK
Dan Simon, USA
Luciano Simoni, Italy
Grigori M. Sisoiev, UK
Christos H. Skiadas, Greece
Davide Spinello, Canada
Sri Sridharan, USA
Rolf Stenberg, Finland
Changyin Sun, China
Jitao Sun, China
Xi-Ming Sun, China
Andrzej Swierniak, Poland
Yang Tang, Germany
Allen Tannenbaum, USA
Cristian Toma, Romania
Irina N. Trendafilova, UK
Alberto Trevisani, Italy
Jung-Fa Tsai, Taiwan
Kuppalapalle Vajravelu, USA
Victoria Vampa, Argentina
Josep Vehi, Spain
Stefano Vidoli, Italy
Xiaojun Wang, China
Dan Wang, China
Youqing Wang, China
Yongqi Wang, Germany
Cheng C. Wang, Taiwan

Moran Wang, China
Yijing Wang, China
Gerhard-Wilhelm Weber, Turkey
Jeroen A. S. Witteveen, The Netherlands
Kwok-Wo Wong, Hong Kong
Ligang Wu, China
Zhengguang Wu, China
Gongnan Xie, China
Wang Xing-yuan, China
Xi Frank Xu, USA
Xuping Xu, USA
Jun-Juh Yan, Taiwan
Xing-Gang Yan, UK
Suh-Yuh Yang, Taiwan
Mahmoud T. Yassen, Egypt
Mohammad I. Younis, USA
Bo Yu, China
Huang Yuan, Germany
S.P. Yung, Hong Kong
Ion Zaballa, Spain
Ashraf M. Zenkour, Saudi Arabia
Jianming Zhan, China
Xu Zhang, China
Yingwei Zhang, China
Lu Zhen, China
Liancun Zheng, China
Jian Guo Zhou, UK
Zexuan Zhu, China
Mustapha Zidi, France

Contents

Recent Theory and Applications on Inverse Problems, Fatih Yaman, Valery G. Yakhno, and Roland Potthast
Volume 2013, Article ID 303154, 2 pages

A Survey on Inverse Problems for Applied Sciences, Fatih Yaman, Valery G. Yakhno, and Roland Potthast
Volume 2013, Article ID 976837, 19 pages

A New Concept for Atmospheric Reentry Optimal Guidance: An Inverse Problem Inspired Approach, Davood Abbasi and Mahdi Mortazavi
Volume 2013, Article ID 419409, 13 pages

Reconstruction of Initial Wave Field of a Nonsteady-State Wave Propagation from Limited Measurements at a Specific Spatial Point Based on Stochastic Inversion, S. L. Han and Takeshi Kinoshita
Volume 2013, Article ID 286247, 10 pages

A Fast High-Order Total Variation Minimization Method for Multiplicative Noise Removal, Xiao-Guang Lv, Jiang Le, Jin Huang, and Liu Jun
Volume 2013, Article ID 834035, 13 pages

Conditional Nonlinear Optimal Perturbation of a Coupled Lorenz Model, Lili Yang
Volume 2013, Article ID 846973, 7 pages

Algebraic Reconstruction of Current Dipoles and Quadrupoles in Three-Dimensional Space, Takaaki Nara
Volume 2013, Article ID 413980, 15 pages

Estimation of Open Boundary Conditions for an Internal Tidal Model with Adjoint Method: A Comparative Study on Optimization Methods, Haibo Chen, Chunbao Miao, and Xianqing Lv
Volume 2013, Article ID 802136, 12 pages

A Mollification Regularization Method for a Fractional-Diffusion Inverse Heat Conduction Problem, Zhi-Liang Deng, Xiao-Mei Yang, and Xiao-Li Feng
Volume 2013, Article ID 109340, 9 pages

Stochastic Inverse Identification of Nonlinear Roll Damping Moment of a Ship Moving at Nonzero-Forward Speeds, S. L. Han and Takeshi Kinoshita
Volume 2012, Article ID 769385, 22 pages

A Review of Soft Techniques for Electromagnetic Assessment of Concrete Condition, F. A. A. Queiroz, D. A. G. Vieira, X. L. Travassos, and M. F. Pantoja
Volume 2012, Article ID 935468, 20 pages

Decomposition of the Time Reversal Operator for Target Detection, Chun-xiao Li, Huan-cai Lu, Ming-fei Guo, Hang-fang Zhao, and Jiang-ming Jin
Volume 2012, Article ID 597474, 13 pages

A Direct Solution Approach to the Inverse Shallow-Water Problem, Aleign Gessese and Mathieu Sellier
Volume 2012, Article ID 417950, 18 pages

Exponential Convergence of an Observer Based on Partial Field Measurements for the Wave Equation,

Dominique Chapelle, Nicolae Cîndea, Maya De Buhan, and Philippe Moireau

Volume 2012, Article ID 581053, 12 pages

Design of an Annular Disc Subject to Thermomechanical Loading, Sergei Alexandrov, Elena Lyamina,

and Yeau-Ren Jeng

Volume 2012, Article ID 709178, 9 pages

CT Metal Artifact Reduction Method Based on Improved Image Segmentation and Sinogram

In-Painting, Yang Chen, Yinsheng Li, Hong Guo, Yining Hu, Limin Luo, Xindao Yin, Jianping Gu,

and Christine Toumoulin

Volume 2012, Article ID 786281, 18 pages

Inverse Estimation of Open Boundary Conditions in the Bohai Sea, Zheng Guo, Anzhou Cao,

and Xianqing Lv

Volume 2012, Article ID 628061, 9 pages

Editorial

Recent Theory and Applications on Inverse Problems

Fatih Yaman,¹ Valery G. Yakhno,² and Roland Potthast^{3,4}

¹ *Institut für Theorie Elektromagnetischer Felder, Technische Universität, 64289 Darmstadt, Germany*

² *Department of Electrical and Electronics Engineering, Dokuz Eylul University, 35160 Izmir, Turkey*

³ *Department of Mathematics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 6AX, UK*

⁴ *German Meteorological Service, Deutscher Wetterdienst Research and Development, Head Division FE 12 (Data Assimilation), Frankfurter Straße 135, 63067 Offenbach, Germany*

Correspondence should be addressed to Fatih Yaman; yaman@temf.tu-darmstadt.de

Received 25 July 2013; Accepted 25 July 2013

Copyright © 2013 Fatih Yaman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this special issue is to present some recent developments in theoretical and practical areas for the solution of inverse problems which appear in physical applications. The scope of the issue not only covers a broad range of original studies but also includes two review articles on recent investigations. In addition, three papers are the collections of studies having results depending on experimental investigations with real data. It is also worthy to mention that the paper acceptance rate after the reviewing procedure is approximately 55% for this issue and in the following we give very brief descriptions of the published papers.

F. Yaman et al. contribute to the special issue with a wide range review paper on inverse problems which focus on applied sciences. In the paper, many inversion based engineering applications are introduced and mathematical investigations of some of the important ones are given in detail. The authors employ acoustic, electromagnetic, and elastic waves for presenting different types of inverse problems. Furthermore, an introduction with many links to the literature is given for modern algorithms which combine techniques from classical inverse problems with stochastic tools into ensemble methods both for data assimilation and for forecasting.

D. Abbasi and M. Mortazavi study a classical problem of optimal reentry guidance with optimal control techniques, and in addition they offer a solution for the path constraints inspired by the idea of inverse problems methodology. It is stated that the final resulting algorithm seems suitable for onboard predictive guidance.

S. L. Han and T. Kinoshita apply a stochastic inverse method for the identification of the nonlinear damping moment of a ship moving at nonzero-forward speed. The workability and the accuracy of the method are verified with laboratory tests under controlled conditions. The authors consider two different types of ship rolling motions, time-dependent transient motion and frequency-dependent periodic motion, in their experimental trials. It is illustrated that the method enables the inherent nonlinearity in damping moment to be estimated, including its reliability analysis.

The second paper of S. L. Han and T. Kinoshita in our issue is related to the reconstruction of initial wave fields from the limited measurements on the specific boundaries of the fluid body for two- and three-dimensional problems. The authors adopt a stochastic method based on Bayesian formulation for the solution of the inverse problem and present illustrative simulation results.

X.-G. Lv et al. propose a fast high order total variation minimization method to restore multiplicative noisy images. The proposed method is able to preserve edges and at the same time avoid the staircase effect in the smooth regions. The authors discuss the convergence of the alternating minimization algorithm in the paper and present numerical results which show that the proposed method gives restored images in high quality.

L. Yang employs conditional nonlinear optimal perturbation techniques of a coupled Lorenz model and studied their dependence on the reference state. In the paper, a simple model is considered to demonstrate the technique which

provides promising results. It is concluded that in order to meet the requirements of the realistic applications the more complex models have to be studied for improving the accuracy of weather and climate forecasts with the related results.

T. Nara aims to propose a numerical algebraic reconstruction algorithm for finding number and locations of neural currents from the knowledge of measured magnetic field data in 3D, which has an application in magnetoencephalography. It is assumed that the radial component of the magnetic field is measured on a sphere of a ball consisting of three additional spherical layers which correspond to a scalp, skull, and brain. In the inverse problem second degree algebraic equations for unknown parameters such as sources and the radial magnetic field data are obtained and solved by means of Gröbner bases. The applicability and the effectiveness of the method are supported with the numerical simulation results.

H. Chen et al. are interested in solving a large-scale optimization problem minimizing a cost functional. The main objective of the study is to make a computational investigation on the performance of the L-BFGS method and two versions of gradient descent method through a series of ideal experiments in which the open boundary conditions are inverted by assimilating the interior observations with the adjoint method. In the paper, numerical results are reported which provide comparisons of the methods for the inversion error between the inverted and prescribed open boundary conditions and the cost function normalized by its initial value as well as the gradient of the cost function with respect to the open boundary conditions.

Z.-L. Deng et al. investigate an inverse heat conduction problem on a semi-infinite domain. More precisely, the fractional diffusion equation is solved analytically in the direct problem and a mollification method with Dirichlet kernel is adopted for the solution of the inverse problem. A priori and a posteriori parameter choice rules are proposed to get the corresponding error estimate between the exact solution and its regularized approximation. Furthermore a numerical example is provided to verify the theoretical results.

F. A. A. Queiroz et al. outline the inverse problem considering the GPR assessment of concrete in microwave imaging. The aim of the inverse problem considered in the study is the inclusion reconstruction in concrete slabs. In the paper three reconstruction algorithms are implemented and tests are performed using real experimental data to assess the applicability and efficiency of these algorithms for image reconstruction.

C.-X. Li et al. discuss the detection problem using passive decomposition of the time reversal operator method. The authors derive the generalized likelihood ratio test based on Neyman-Person lemma provided that the signal components can be modeled as a linear combination of basis vectors with an unknown signal subspace. Moreover, authors present results which illustrate the performance of the time reversal operator detector and the energy detector using the real acoustic data collected in the laboratory waveguide experiment.

A. Gessese and M. Sellier report a numerical technique to identify the underlying bedrock topography from given free surface elevation data in shallow open channel flows. The technique described in the paper is based on an explicit finite difference scheme which provides the solution to the inverse problem directly. The algorithm is tested on three numerical examples, and noise data tests are performed.

D. Chapelle et al. show that the exponential convergence can be achieved with an observer strategy based on partial measurements of the solution of a wave equation. The authors adopt the observer strategy "Schur Displacement Feedback" for elasticity-like formulations to the scalar wave equation. Furthermore, numerical results are provided to illustrate the effectiveness of the proposed approach.

S. Alexandrov et al. present two solutions to design a thin annular disc of variable thickness subject to thermomechanical loading. The authors adopted two design criteria, one of which leads to a relation between optimal values of the loading parameters for each specific shape of the disc. The application of the other criterion has shown that the state of required stress and strain appears in the disc of constant thickness at certain values of the loading parameters.

Y. Chen et al. propose a strategy for reducing metallic artifacts in CT images. It is concluded in the paper that, "visual and quantitative analyses on phantom and clinical data show that the proposed correction method provides a substantial reduction of the metallic artifacts in the corrected images. The pixel-wise operations in the pre-filtering and sinogram inpainting steps are greatly accelerated by a CUDA parallelization that makes the algorithm also competitive in computation time."

Z. Guo et al. present a method to estimate open boundary conditions by an adjoint data assimilation approach. The authors assume the open boundary conditions being position dependent and can be approximated by linear interpolation among values at certain independent points. Twin numerical experiments are performed by employing Bohai Sea model in order to verify the feasibility of the method.

Acknowledgment

The guest editors would like to deeply thank all the authors, the reviewers, the editorial board of the Mathematical Problems in Engineering journal, and all the staff involved in the preparation of this issue.

*Fatih Yaman
Valery G. Yakhno
Roland Potthast*

Review Article

A Survey on Inverse Problems for Applied Sciences

Fatih Yaman,¹ Valery G. Yakhno,² and Roland Potthast^{3,4}

¹ *Institut für Theorie Elektromagnetischer Felder, Technische Universität Darmstadt, 64289 Darmstadt, Germany*

² *Department of Electrical and Electronics Engineering, Dokuz Eylül University, 35160 İzmir, Turkey*

³ *Department of Mathematics, University of Reading, Whiteknights P.O. Box 220, Reading RG6 6AX, UK*

⁴ *German Meteorological Service, Deutscher Wetterdienst Research and Development, Head Division FE 12 (Data Assimilation) Frankfurter Straße 135, 63067 Offenbach, Germany*

Correspondence should be addressed to Fatih Yaman; yaman@temf.tu-darmstadt.de

Received 19 March 2013; Accepted 16 June 2013

Academic Editor: Baocang Ding

Copyright © 2013 Fatih Yaman et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this paper is to introduce inversion-based engineering applications and to investigate some of the important ones from mathematical point of view. To do this we employ acoustic, electromagnetic, and elastic waves for presenting different types of inverse problems. More specifically, we first study location, shape, and boundary parameter reconstruction algorithms for the inaccessible targets in acoustics. The inverse problems for the time-dependent differential equations of isotropic and anisotropic elasticity are reviewed in the following section of the paper. These problems were the objects of the study by many authors in the last several decades. The physical interpretations for almost all of these problems are given, and the geophysical applications for some of them are described. In our last section, an introduction with many links into the literature is given for modern algorithms which combine techniques from classical inverse problems with stochastic tools into ensemble methods both for data assimilation as well as for forecasting.

1. Introduction

Finding causes from the knowledge of their effects, which actually constitutes the idea of solving *inverse problems*, is necessary for the identifications of practical applications in different critical areas such as mine detection, medical imaging, remote sensing, nondestructive testing, and geophysical explorations. It is also a crucial ingredient of forecasting in basic areas such as weather prediction or projections for climate change.

However, small changes in the effects might result in large differences in the causes, or the same effect might be obtained from more than one cause. Therefore, it is difficult or sometimes impossible to find actual reason uniquely by observing only effects. This *ill-posed* fundamental characteristic of inverse problems influences the complexity and features of their solutions dramatically [1–6].

It is a generally accepted view that the first mathematical investigation of inverse problems is the study of Abel's on a mechanical problem for finding the curve of an unknown

path in 1826. On the other hand, the invention of radar and sonar during the Second World War inspired researchers to focus especially on inverse scattering problems whose aims are not only to determine locations of the targets from the transmitter/receiver antennas but also to construct their detailed images. This motivation induced the progress in developing new reconstruction methods and their extensions to other research branches such as nondestructive testing, biomedical imaging, seismology, and atmospheric profile inversions. This agenda has led to a large number of methods and tools since 1980, with a continuing flow of new ideas into the field.

Several inversion methods in scattering and tomography have been suggested in the 1980s and 1990s, with qualitative methods arising since around 1996. Since 2000, an increased interest in the treatment of sparsity has led to new regularization tools and new approaches. Also, the need to carry out inversion as a part of dynamical systems and forecasting, also called *data assimilation*, is leading to an increasing synergy of filtering methods from stochastics and inversion tools from

regularization theory. Since around 2000, multiphysics-based methods have become very popular, leading to increased insight for example into the human body or the atmosphere. Nowadays the advent of powerful computers and high technologies made it possible to evaluate and process large volume of data for finding sufficiently accurate solutions of practical inverse problems. For a more detailed information into the different areas of inverse problems we refer to [7–20].

The plan of the paper is as follows. A short summary to the inverse problems in the historical perspective has just been reported in the introduction section. The following two sections are devoted for the presentation of the mathematical basics used in the solution of well-known inverse problems for acoustic, electromagnetic, and elastic waves. Clearly, we cannot aim to cover their entire literature with this survey but describe some interesting and important lines of development. Afterwards, some of important inversion-based practical applications are introduced, and as a selected topic, inverse problems for the neural field equation is described in details. In the final section, conclusions and concluding remarks are given.

2. Inverse Problems for Time-Harmonic Acoustic and Electromagnetic Waves

In this section, we focus on the mathematical investigation of inverse problems whose aims are to reconstruct geometrical and/or physical properties of penetrable/impenetrable objects from the knowledge of scattered acoustic or electromagnetic waves at certain measurement points. Physically speaking, identifications of the unknown and inaccessible objects with certain waves are the main problems of the nondestructive testing, radar/sonar applications, tumor detection, and so forth.

Let us assume a time-harmonic plane wave $u^i = e^{i(kx \cdot d - \omega t)}$, which was insonicated by a sufficiently far acoustic source operating at a constant frequency ω . The acoustic wave propagates in the direction d , with a speed of sound c , in a homogeneous medium having a wave number $k = \omega/c$. In this background medium, we consider a bounded penetrable/impenetrable scatterer with support given by a domain $D \in \mathbb{R}^m$ ($m = 2, 3$) and define ν as the unit outward normal to the boundary ∂D .

The scattered field u^s , which appears from the interaction of the incident wave with the obstacle, satisfies the Sommerfeld radiation condition at infinity

$$\lim_{r \rightarrow \infty} r^{(m-1)/2} \left(\frac{\partial u^s}{\partial r} - iku^s \right) = 0, \quad r = |x|, \quad (1)$$

and ensures that the scattered wave has the form of an outgoing wave. Furthermore, the so-called far-field pattern u_∞ can be obtained from the asymptotic behavior of the scattered wave such that

$$u^s(x) = \frac{e^{ikr}}{r^{(m-1)/2}} \left\{ u_\infty(\hat{x}) + O\left(\frac{1}{r}\right) \right\}, \quad (2)$$

$$\hat{x} = \frac{x}{|x|}, \quad r = |x| \rightarrow \infty.$$

The total field u is simply the superposition of incident and scattered fields $u = u^i + u^s$, which satisfies the homogeneous Helmholtz equation in the exterior domain of the obstacle

$$\Delta u + k^2 u = 0 \quad \text{in } \mathbb{R}^m \setminus \bar{D} \quad \text{for } (m = 2, 3) \quad (3)$$

and the boundary condition depending on scatterer type. Commonly used boundary conditions in the literature are given in the following.

The Dirichlet boundary condition:

$$u = 0 \quad \text{on } \partial D, \quad (4)$$

the Neumann boundary condition:

$$\frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial D, \quad (5)$$

the impedance boundary condition:

$$u + \frac{\eta}{ik} \frac{\partial u}{\partial \nu} = 0 \quad \text{on } \partial D, \quad (6)$$

the conductive boundary condition:

$$u_0 = u, \quad \frac{\partial u}{\partial \nu} - \frac{\partial u_0}{\partial \nu} = \lambda u \quad \text{on } \partial D. \quad (7)$$

Parameters η and λ are the impedance and conductivity functions, respectively. In latest equation, u_0 is the total field inside of the scatterer, and this boundary condition can be considered as a more general form of the transmission conditions.

On the other hand, in order to model acoustic scattering by an inhomogeneous medium, it is sufficient to replace (3)–(7) with

$$\Delta u + k^2 n(x) u = 0 \quad \text{in } \mathbb{R}^m \quad \text{for } (m = 2, 3), \quad (8)$$

where $n(x) = c^2/c_i^2(x) + i\sigma_i(x)$ is the refractive index, c is the sound speed in the homogeneous medium, $c_i(x)$ is the speed of sound in the inhomogeneous medium, and $\sigma_i(x) \geq 0$ models the absorption.

Analogically one can also consider a scattering problem for electromagnetic waves assuming that the electric and magnetic field components of the incident plane wave are given by

$$E^i(x, t) = ik(d \times p) \times de^{i(kx \cdot d - \omega t)}, \quad (9)$$

$$H^i(x, t) = ikd \times pe^{i(kx \cdot d - \omega t)},$$

where $k = \omega\sqrt{\epsilon\mu}$ is the wave number and ϵ and μ are the electric permittivity and magnetic permeability of the host medium, respectively. Here, ω represents the frequency of the wave source, d is the direction, and p is the polarization of the electromagnetic wave. Then, we consider that the wave interacts with a 3D obstacle $D \in \mathbb{R}^3$ in the homogeneous medium and that scattered electromagnetic fields (E^s, H^s) occur similar to the previous case. However, in electromagnetics, the scattered field has to satisfy the Silver-Müller radiation condition

$$\lim_{r \rightarrow \infty} (H^s \times x - rE^s) = 0. \quad (10)$$

The far field pattern of the corresponding scattered field can be obtained by

$$E^s(x) = \frac{e^{ikr}}{r} \left\{ E_\infty(\hat{x}) + O\left(\frac{1}{r}\right) \right\}, \quad (11)$$

$$\hat{x} = \frac{x}{|x|}, \quad r = |x| \rightarrow \infty.$$

Furthermore, the total electric field $E = E^i + E^s$ and the total magnetic field $H = H^i + H^s$ in the medium satisfy the Maxwell equations

$$\text{curl } E - ikH = 0, \quad \text{curl } H + ikE = 0 \quad \text{in } \mathbb{R}^3 \setminus \bar{D} \quad (12)$$

and the chosen boundary conditions on the surface of the scatterer. The boundary conditions which were presented for acoustic waves are now given for electromagnetic waves in (13)–(16).

The Dirichlet-type perfect conductor boundary condition:

$$\nu \times E = 0 \quad \text{on } \partial D, \quad (13)$$

the Neumann-type boundary condition:

$$\nu \times H = 0 \quad \text{on } \partial D, \quad (14)$$

the impedance boundary condition:

$$\nu \times E + \frac{\eta}{ik} (\nu \times (\nu \times H)) = 0 \quad \text{on } \partial D, \quad (15)$$

the conductive boundary condition:

$$\begin{aligned} \nu \times E_0 = \nu \times E, \\ \nu \times (\nu \times H) - \xi \nu \times (\nu \times H_0) = \lambda (\nu \times E) \end{aligned} \quad \text{on } \partial D. \quad (16)$$

The parameter ξ in the last equation is the ratio of the wave numbers $\xi = k_0/k$, where k_0 is the wave number of the object's interior domain.

For the case of modeling scattering of electromagnetic waves by an inhomogeneous medium, (12)–(16) are replaced by

$$\text{curl } E - ikH = 0, \quad \text{curl } H + ikn(x)E = 0 \quad \text{in } \mathbb{R}^3, \quad (17)$$

where $n(x) = 1/\epsilon(\epsilon_i + i\sigma_i(x)/\omega)$ is the refractive index, ϵ is the permittivity of the homogeneous medium, ϵ_i is the permittivity of the inhomogeneous medium, $\sigma_i(x)$ is the conductivity, and ω is the frequency of the wave.

We recommend to the readers the books [9, 21–23] for the detailed investigations of the electromagnetic and acoustic waves with the same notation.

2.1. Boundary Reconstruction Problems of Acoustic Waves. In this section, our main concern is to present mathematical treatments of some available methods for reconstructing shapes of obstacles from the knowledge of scattered field.

Solution of this type of problems with acoustic waves are important from practical point of view since the real life experiments with acoustic waves are easier, safer, and lower in cost.

To provide a simpler mathematical presentation, we consider a two dimensional simply connected object $D \in \mathbb{R}^2$ having the Dirichlet condition on its boundary Γ . It is further assumed that the incident acoustic wave u^i , which interacts with the object, and scattered near/far fields are given. This inverse problem is nonlinear due to the mathematical relation between the scattered wave and the shape of the cylinder, and it is ill-posed since the determination of Γ does not depend continuously on the scattered field [18]. For the solution of this type of inverse problems, there are a variety of methods of which we present a short list in the following without claiming to cover the entire literature.

Iterative Methods

- (i) Landweber iterations [24, 25].
- (ii) Regularized Newton [22, 26–50].
- (iii) Newton-Kantorovich [51, 52].

Decomposition Methods

- (i) Colton-Monk [53–55].
- (ii) Kirsch-Kress [56–58].
- (iii) Angell-Kleinmann-Roach [59–61].
- (iv) Hybrid method [62–67].
- (v) Potthast's point source method [23, 68, 69].

Probe and Sampling Methods

- (i) Linear sampling [70–74].
- (ii) Factorization [75, 76].
- (iii) Singular sources [19, 23, 77].
- (iv) Probe method [78, 79].
- (v) Enclosure method [80, 81].
- (vi) No-response test [82–84].

In the first group of methods [24, 52], the inverse obstacle problem is considered as an ill-posed nonlinear operator equation,

$$A(\Gamma_0) = u_\infty, \quad (18)$$

for an initial boundary Γ_0 . Often, it is assumed to be a star-like parametrization $\Gamma_0 = \{z(t) := r(t)(\cos t, \sin t) : t \in [0, 2\pi)\}$. Here, the operator $A : \Gamma \rightarrow u_\infty$ is actually defined to map the boundary Γ of the scatterer onto the far field u_∞ . Newton's method to solve a nonlinear equation employs successive linearization; that is, we replace (18) by an equation for the update h_n

$$A(z_n) + A'(z_n)h_n = u_\infty, \quad (19)$$

where

$$z_{n+1} = z_n + h_n, \quad n = 0, 1, 2, \dots, \quad (20)$$

leading to a sequence of linear equations (19) and updates (20) with starting value $z_0 := \Gamma_0$. Equation (19) is ill-posed and needs to be regularized. As a simple option, Tikhonov's regularization with regularization parameter $\alpha > 0$ can be applied, leading to

$$\alpha h_n + [A(z_n)']^* A'(z_n) h_n = [A(z_n)']^* [u_\infty - A(z_n)] \quad (21)$$

in each iteration step $n = 0, 1, 2, \dots$. Iterative methods for ill-posed equations need to be stopped after a finite number of iterations to keep the ill-posedness under control; compare, for example, [23, 85].

The main idea of the *decomposition methods* is to split full nonlinear shape reconstruction problem given by (18) into a linear ill-posed equation which is solved first and a nonlinear well-posed equation to be solved in a second step.

In order to discuss the Kirsch-Kress method [56–58] via potential approach let us employ an initial boundary Γ_0 as an approximation of the actual boundary Γ . We assume that Γ_0 is in the interior of the true scatterer.

In this case, the approximate total field $\tilde{u}(x)$ can be approximated by the sum of the incident wave u^i and a single-layer potential,

$$\tilde{u}(x) = u^i(x) + \int_{\Gamma_0} \Phi(x, y) \varphi(y) ds(y), \quad x \in \mathbb{R}^2 \setminus \overline{D}, \quad (22)$$

where $\Phi(x, y)$ is the fundamental solution to the Helmholtz equation in two dimensions and φ is a continuous density (source) function defined over Γ_0 . The far field pattern of the single-layer potential should be measured far field pattern $u^{(\infty)}$, which leads to the linear equation

$$u^{(\infty)}(\hat{x}) = c \int_{\Gamma_0} e^{-ik\hat{x} \cdot y} \varphi(y) ds(y), \quad \hat{x} \in \mathbb{S}, \quad (23)$$

with some constant c depending on the dimension of the space under consideration. The first step consists of solving the ill-posed linear equation (23) to calculate φ .

Then, when we consider a Dirichlet boundary condition, the shape is found in a second nonlinear step as the zero curve of the total field $\tilde{u}(x)$. To find this zero curve, we introduce an operator G , which maps Γ to the values of the approximate total field \tilde{u} , on Γ such that

$$G : \Gamma \longmapsto \tilde{u}|_{\Gamma}. \quad (24)$$

Then the problem is reduced to the solution of the following optimization problem:

$$G(\Gamma) = 0, \quad (25)$$

which can be done by minimizing the defect $\|G(\Gamma)\|_{L^2(\Gamma)}$ in a least square sense.

Linear sampling [70–74] is based on solving the far field equation

$$\int_{\Omega} u_\infty(\hat{x}; d) g(d) ds(d) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} e^{-ik\hat{x} \cdot y} \quad \hat{x} \in \mathbb{S}. \quad (26)$$

Then, the density g satisfies

$$\|g(\cdot, y)\|_{L^2(\Omega)} \longrightarrow \infty, \quad y \longrightarrow \partial D, \quad (27)$$

that is, when the source point y approaches the boundary Γ . This behaviour can be used to visualize the shape of the scatterer from the knowledge of the far field pattern $u_\infty(\cdot, d)$ for all directions $d \in \mathbb{S}$.

We also refer to the *orthogonality sampling* method, which has recently been suggested in [86]. It is particularly suited to deal with multifrequency data as is naturally obtained when acoustic pulses are used to probe an object or region in space. The method has been independently suggested by Ito et al. [87, 88] and successfully applied as a first step in a larger inversion procedure. A convergence analysis of the method in the limit of small scatterers has recently been achieved by Griesmaier [89].

2.2. Parameter on the Boundary Reconstruction Problems of Acoustic Waves. The inverse problem considered in this section is finding a continuous function which is defined on the boundary of the obstacle from the knowledge of the scattered acoustic waves for a given shape in two dimensions. We discuss the methods whose aims are to reconstruct η , in the impedance boundary condition (6), and λ , in the conductive boundary condition (7). Note that conductivity function is employed to model the inhomogeneity which might exist on the boundary of an object in a more realistic way.

In [90–92], 3D obstacles with impedance boundary condition are studied for acoustic case, where in [91] electromagnetic case is included. In the same field, the papers [64, 93–96] are focused on 2D geometries for the less complexity of governing numerical experiments.

In [90], the impedance reconstruction problem is solved theoretically with a technique depending on Backus and Gilbert's method which is applicable to linear moment problems. To this aim, approximate Green's functions are used to reduce the nonlinear problem to two linear moment problems. On the other hand, the study [92] is devoted for the reconstruction of impedance functions via the Kirsch-Kress and Colton-Monk decomposition methods. Furthermore, some interesting papers appeared on the impedance reconstructions, recently [97, 98].

The paper [93] introduced a new method for impedance reconstructions in the spirit of the Kirsch-Kress decomposition method. That is, the scattered field is represented via single-layer potential over the known boundary of the impedance cylinder Γ , instead of defining an auxiliary initially guessed curve. Then, the density function φ is solved from the following ill-posed integral equation

$$u_\infty(\hat{x}) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\Gamma} e^{-ik\hat{x} \cdot y} \varphi(y) ds(y), \quad \hat{x} \in \Omega, \quad (28)$$

through Tikhonov's regularization for known far field, where Ω is a unit circle. From the knowledge of the density function now the total field and the normal derivative of the total field can be computed on the boundary of the obstacle via

jump relations [22]. Finally, η is obtained from (6) in the least squares sense. This method is also extended for the reconstructions of the conductivity functions of the obstacles in free space [99], for the obstacles buried in penetrable cylinders [100] and for a combination of a shape and conductivity function reconstruction problem [101], firstly by Yaman [6].

Moreover, [64, 95, 96] are devoted for the shape and impedance reconstructions of 2D obstacles in acoustics. To do this, the hybrid method is employed by Serranho [64, 66]. In [95], a level set algorithm is combined with boundary integral equations in acoustic case to reconstruct the shape and impedance of 2D obstacles from multi-illuminations, and in [96], it has been shown that the knowledge of the scattered fields corresponding to three incident waves can be used for the determination of the shape and the impedance via integral equation methods and conformal mapping techniques.

3. Inverse Problems for Differential Equations of Elastodynamics

3.1. Differential Equations of Elastodynamics. Let $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ be a 3D space variable, and let $t \in \mathbb{R}$ be a 1D time variable; let $\mathbf{u}(x, t) = (u_1(x, t), u_2(x, t), u_3(x, t))$ be the displacement vector function of an inhomogeneous anisotropic elastic material characterizing by density ρ and the elastic moduli C_{ijkl} . The density ρ and elastic moduli C_{ijkl} are varying functions of position $x = (x_1, x_2, x_3)$. Combining the properties of the strain-energy function with Hooke's law we find [102] that C_{ijkl} satisfy the following property and strong convexity

$$C_{ijkl} = C_{klij} = C_{lki j} = C_{ijlk}, \quad \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 C_{ijkl} \varepsilon_{ij} \varepsilon_{kl} > 0 \quad (29)$$

for any nonzero 3×3 real symmetric matrix $(\varepsilon_{ij})_{3 \times 3}$. Equations for motion in inhomogeneous anisotropic elastic materials are, in our notation (see, e.g., [102]),

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \frac{\partial}{\partial x_j} \left(C_{ijkl} \frac{\partial u_k}{\partial x_l} \right) + f_i, \quad i = 1, 2, 3, \quad (30)$$

where f_i are components of the body forces $\mathbf{f}(x, t) = (f_1(x, t), f_2(x, t), f_3(x, t))$ acting per unit volume on the particle originally at position x at some reference time t .

Let us examine

$$\mathbf{u}(x, t) = \mathbf{U}(t - T(x)) A(x) \quad (31)$$

as an approximate solution of (30) for $\mathbf{f} = 0$. Near a wavefront $t = T(x)$, we assume that components of $\mathbf{U} = (U_1, U_2, U_3)$ are fluctuating much more rapidly than $A(x)$ or C_{ijkl} , and the successive derivatives $\partial \mathbf{U} / \partial t$ and $\partial^2 \mathbf{U} / \partial t^2$ are fluctuating still more rapidly. Substituting (31) into (30) we find (see, e.g., [102])

$$\left(\rho \delta_{ik} - \sum_{j=1}^3 \sum_{l=1}^3 \frac{\partial T}{\partial x_j} \frac{\partial T}{\partial x_l} \right) \frac{\partial^2 U_k}{\partial t^2} A = E_i(\mathbf{U}A), \quad (32)$$

where $E_i(\mathbf{U}A)$ includes merely first-order derivatives of \mathbf{U} , \mathbf{U} itself, the elastic moduli, amplitude function $A(x)$, and gradients of these.

Thus, the left-hand side of (32) must be much smaller than $\partial^2 \mathbf{U} / \partial t^2$. We conclude that the matrix of coefficients of $\partial^2 \mathbf{U}(t - T(x)) / \partial t^2 A(x)$ must be singular:

$$\det \left(\rho \delta_{ik} - \sum_{j=1}^3 \sum_{l=1}^3 \frac{\partial T}{\partial x_j} \frac{\partial T}{\partial x_l} \right) = 0. \quad (33)$$

This equation determines the possible wavefronts in an elastic medium, since it gives a constraint on the function $T(x)$. In an inhomogeneous isotropic medium,

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}), \quad (34)$$

where δ_{ij} is the Kronecker symbol; that is, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$; moreover, $\lambda = \lambda(x)$, $\mu = \mu(x)$ are known as the Lamé functions.

In an inhomogeneous isotropic medium, the special form (34) of C_{ijkl} makes it possible to get (33) as follows:

$$\left(\nabla T \cdot \nabla T - \frac{\rho}{\lambda + 2\mu} \right) \left(\nabla T \cdot \nabla T - \frac{\rho}{\mu} \right)^2 = 0. \quad (35)$$

This is, $T(x)$ satisfies the eikonal equation

$$|\nabla T|^2 = \frac{1}{C_P^2(x)} \quad (36)$$

or eikonal equation

$$|\nabla T|^2 = \frac{1}{C_S^2(x)}, \quad (37)$$

where $C_P^2 = \sqrt{(\lambda + 2\mu)/\rho}$ is the local P -wave speed and $C_S^2 = \sqrt{\mu/\rho}$ is the local S -wave speed.

3.2. Inverse Kinematic Problem of Seismic. Suppose that a point source at position $x^0 \in \mathbb{R}^3$ becomes active at a time chosen to be the origin, $t = 0$. In homogeneous isotropic medium, wavefronts emanate from the source as ever-expanding spheres, with radius $C_P t$ (for P -waves) and $C_S t$ for S -waves, arriving at the general position x at time $t = |x - x^0| / C_P$ and $t = |x - x^0| / C_S$. We introduce the function $T(x, x^0)$ as the travel time required for the wavefront to reach x from x^0 . The function $T(x, x^0)$ satisfies (36) for P -waves and (37) for S -waves.

One of the first inverse problems, stated and studied in geophysics, was the inverse kinematic problem. The physical interpretation of this problem is the following. Let us assume that Earth is an isotropic inhomogeneous elastic medium and the measurements of the seismic waves, arising from a point source x^0 and propagating in Earth, are given for points on its surface. These measurements contain data of the travel time $T(x, x^0)$ of seismic waves between the point of the source x^0 and any point of the Earth's surface. The inverse kinematic problem is to find the speed of the seismic waves inside of

Earth using the measurement data. Mathematically we can state the inverse kinematic problem as follows. Let D be a domain bounded by the surface S , and let $T(x, x^0)$ be the function of the travel time required by a signal with unknown speed $C(x) > 0$ to reach x from x^0 . Find $C(x)$ for all x from D if the function $T(x, x^0)$ is given for all points $x^0 \in S_1$ and $x \in S_2$, where $S_1 \subseteq S$ and $S_2 \subseteq S$ are subsets of S .

Herglotz [103] and Wiechert and Zoeppritz [104] were the first who studied the inverse kinematic problem in assumptions

$$C(x) \frac{1}{n(r)}, \quad r = |x|, \quad \frac{d}{dr}(rn(r)) > 0, \quad (38)$$

x^0 is a fixed point from S ,

if $T(x, x^0)$ is known for any x from S . Gerver and Markushevich [105] have showed that the condition $(d/dr)(rn(r)) > 0$ can be eliminated, but in this case the inverse kinematic problems have many solutions and the set of these solutions has been described. The first theoretical study of the inverse kinematic problem for a horizontal inhomogeneous medium has been made by Lavrentiev and Romanov [106]. The first result of the study of the multidimensional inverse kinematic problem in a linear approximation, when the function $C(x)$ depends on 3D space variable $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, was obtained by Romanov [107]. The existence, uniqueness, and stability estimate theorems for the multidimensional inverse kinematic problems have been established in some classes of analytic and smooth functions by Anikonov [108], Bukhgeim [109], and Mukhometov [110].

In a recent time, the Earth is modeled as an anisotropic elastic medium which is located in the given 3D domain. The wave speed is given by a symmetric positive definite matrix $G = (g_{ij})_{3 \times 3}$, that is, a Riemannian metric in mathematical terms (see, e.g., [111]). The problem is to determine the metric in a given domain from the lengths of geodesics joining points on the boundary of the domain. The linearization of this problem leads to a problem of the integral geometry [111–113].

The regular study of the problems of finding the isotropic and anisotropic Riemannian metrics and integral geometry problems has been made in the works [111, 114–117]. The modern numerical algorithms for the computation of the inverse kinematic problems of seismic have been developed in the works [118, 119].

3.2.1. Inverse Problems in Ray Statements. Let us note that isotropic inhomogeneous elastic medium is completely characterized by three functions: the density $\rho(x)$ and speeds $C_P(x)$, $C_S(x)$ of P - and S -waves. Using the point source at position x^0 of the boundary S of the given domain D , which becomes active at the time $t = 0$, we measure the function $T(x, x^0)$ of the travel time required for the fronts of P - and S -waves to reach x from x^0 . We use these information for solving two inverse kinematic problems for P - and S -waves. The solutions of these problems are speeds $C_P(x)$, $C_S(x)$. To complete the identification of unknown isotropic inhomogeneous medium we need to determine the last unknown function $\rho(x)$ after finding $C_P(x)$, $C_S(x)$. An

inverse problem to recover $\rho(x)$ in a given bounded domain D , containing an isotropic inhomogeneous elastic medium, has been solved by Yakhno [120–123]. In these works, the displacement fields $\mathbf{u}(x, x^0, t)$ have been measured for all points x and x^0 running the boundary S of D for all times from a time interval containing the time $t = T(x, x^0)$ of arriving of the P -waves.

3.3. One-Dimensional Inverse Dynamic Problems. The vertical inhomogeneous model of Earth is one of the popular models of geophysics [102], and the inverse problems of recovering the density $\rho(x_3)$ and Lamé functions $\mu(x_3)$ and $\lambda(x_3)$, depending on one variable x_3 and appearing in equations of elastodynamics (30) for the case of inhomogeneous isotropic elastic medium, have been studied by many authors [1, 2, 7–15, 21, 22, 24, 26–28, 36–46, 53–55, 59–61, 70–73, 78–81, 85, 87–91, 97, 98, 103, 105, 108, 109, 113, 118, 119, 124–214]. Because the unknown functions depend on one variable, the inverse problems of their recovering are called one-dimensional inverse problems although all differential equations of elasticity contain 3D space variable $x = (x_1, x_2, x_3)$ and 1D time variable t . Alekseev and Dobrinsky [124, 125] were the first who described the importance of one-dimensional inverse problems of elastodynamics in geophysics and studied them as problems of the recovery of smooth functions $\rho(x_3)$, $\mu(x_3)$, and $\lambda(x_3)$ of one variable x_3 . The uniqueness of the solutions of these inverse problems has been studied firstly by Blagoveschenskii [147] and Romanov [215] for the isotropic elastic media and then by Volkova and Romanov [216] for anisotropic elastic media. The regular study of the theory, methods, and applications of one-dimensional inverse problems for dynamical differential equations of isotropic and anisotropic elastic media has been made in works [122, 159, 170, 190, 217–221] and others. The recent development of theory, methods, and applications of one-dimensional inverse problems of dynamic elasticity can be found in the works [150, 211, 222, 223].

We note that a model of Earth as a composite medium consisting of a finite number of different elastic layers is very popular in geophysics. In this case, the one-dimensional inverse problem consists of finding $\rho(x_3)$ and the Lamé functions $\mu(x_3)$ and $\lambda(x_3)$ as functions of one variable x_3 with piecewise constant values. The computation of solutions of this type of one-dimensional inverse problems has been studied in [186, 187, 220]. The modern theory and methods of the construction of solutions of one-dimensional inverse problems for the equations of elastodynamics in elastic composite media can be found in the works [142, 214, 224–228].

3.4. Multidimensional Inverse Dynamic Problems in Linear (Born) Approximation. Linearized multidimensional inverse dynamic problems (or inverse problems in the Born approximation) take an important place through all statements of multidimensional inverse problems for equations of elastodynamics. The statements of these problems have natural physical and mathematical sense. From the physical point of view, an isotropic inhomogeneous elastic body, which is

characterized by the Lamé functions $\mu_1(x)$ and $\lambda_1(x)$ and density $\rho_1(x)$, is included in a vertical inhomogeneous (or homogeneous) elastic medium. Let, for example, the half space $x_3 > 0$ contain this medium, and let the characteristics $\mu_1(x)$, $\lambda_1(x)$, and $\rho_1(x)$ of the elastic body be unknown functions. The linearized inverse problem is to find these unknown functions if we measure the first act of scattering the displacement field on the surface $x_3 = 0$ arising from the forces located on the same surface $x_3 = 0$. From the mathematical point of view, we consider the differential equations of elastodynamics in a half space $x_3 > 0$ with boundary conditions on $x_3 = 0$. We assume that the Lamé functions $\mu(x)$ and $\lambda(x)$ and density $\rho(x)$ appearing in differential equations and boundary conditions can be presented in the form

$$\begin{aligned}\mu(x) &= \mu_0(x_3) + \mu_1(x), \\ \lambda(x) &= \lambda_0(x_3) + \lambda_1(x), \\ \rho(x) &= \rho_0(x_3) + \rho_1(x),\end{aligned}\quad (39)$$

where $\mu_0(x_3)$, $\lambda_0(x_3)$, and $\rho_0(x_3)$ are functions depending on x_3 and characterizing vertical inhomogeneous medium and $\mu_1(x)$, $\lambda_1(x)$, and $\rho_1(x)$ are functions of 3D space variable $x = (x_1, x_2, x_3)$ characterizing the elastic body which is included in the vertical inhomogeneous medium.

We assume that the displacement field $\mathbf{u}(x, t)$ is presented in the form $\mathbf{u}(x, t) = \mathbf{u}_0(x, t) + \mathbf{u}_1(x, t)$, where $\mathbf{u}_0(x, t)$ is the displacement field of the vertical inhomogeneous medium arising from the given forces, and $\mathbf{u}_1(x, t)$ is the first act of scattering $\mathbf{u}_0(x, t)$ on the inhomogeneous inclusion with characteristics $\mu_1(x)$, $\lambda_1(x)$, and $\rho_1(x)$. The equations of elastodynamics with boundary conditions are linearized around $\mu_0(x_3)$, $\lambda_0(x_3)$, $\rho_0(x_3)$, and $\mathbf{u}_0(x, t)$. The unknown functions $\mu_1(x)$, $\lambda_1(x)$, and $\rho_1(x)$ appear in inhomogeneous terms of linearized equations for $\mathbf{u}_1(x, t)$. We need to recover $\mu_1(x)$, $\lambda_1(x)$, and $\rho_1(x)$ if we know $\mathbf{u}_1(x, t)$ for $x_3 = 0$ (see, e.g., [122]). The uniqueness of the solution of a linearized multidimensional inverse problem has been studied by Romanov [215]. The existence theorem and computation of a solutions of a linearized multidimensional inverse problems of elastodynamics have been obtained in the works [122, 229]. The recovery of the function characteristics of an elastic body in linear approximation was a subject of the works [146, 154, 202, 230]. The linearized inverse problems of determining the function characteristics of transversally isotropic elastic media from the measurements of reflected waves have been developed by Sharafutdinov [231, 232]. The linearized inverse problems for nonhomogeneous bodies have been stated and developed by Steinberg [233].

3.5. Multidimensional Inverse Dynamic Problems in the Statements of the Dirichlet-To-Neumann Map. The inverse problems of determining the elastic moduli and density as functions of the space variables in a bounded domain from observed data of the solution on the boundary (or a part of the boundary) of this domain are geophysical motivated. One important class of these problems is inverse problems

for equations of elastodynamics in terms of the Dirichlet-to-Neumann map. The Dirichlet-to-Neumann map models surface measurements by giving the correspondence between a displacement at the boundary S of the given bounded domain D and surface traction

$$\sum_{j=1}^3 \sum_{k=1}^3 \sum_{l=1}^3 \nu_j C_{ijkl} \frac{\partial u_k}{\partial x_l} \Big|_{S \times (0, T)}, \quad i = 1, 2, 3, \quad (40)$$

where $\nu = (\nu_1, \nu_2, \nu_3)$ is the unit outer normal to S , $(0, T)$ is the observed time interval, and u_k , $k = 1, 2, 3$ are components of the displacement vector function $\mathbf{u}(x, t)$ satisfying (30). The details of the use of the Dirichlet-to-Neumann map in modeling surface measurements in inverse problems can be found in [234]. The inverse problems in the Dirichlet-to-Neumann map statements are successfully applied to study the unique determination of the solutions of the inverse problems of elasticity as in the static isotropic and anisotropic cases [235–239] as well as in dynamic case [240].

3.6. Carleman Estimates and Uniqueness of Solutions for Multidimensional Inverse Dynamic Problems. For the study of the inverse problems for the scalar partial differential equations with a finite number of observation, Bukhgeim and Klivanov [160] proposed a remarkable method based on a Carleman estimate. Later, the Carleman estimate method has been generalized to study the uniqueness and stability estimate of the solutions of the inverse problems for equations of elastodynamics by Isakov [208], Ikehata et al. [203], Eller et al. [182], Imanuvilov et al. [204], Imanuvilov and Yamamoto [205, 206], and Lin and Nakamura [241].

In the papers [203–206, 208] the authors assume some geometric constraints on the surface under observation for proving the uniqueness and stability of the solutions of the multidimensional inverse problems using the Carleman estimates. Later, the stability estimate theorem for solutions of a multidimensional inverse problem for equations of elastodynamics has been proven for an arbitrary subboundary by Bellassoued et al. [143].

4. Inversion Based Applications

In the following, we present a *general overview* on some inversion-based *application areas*. We then give more details about particular selected applications in the subsequent sections. Further, different applications whose details are skipped in this section such as remote sensing, nuclear science, and geophysics can be found in [164, 184, 242–247].

Underwater Acoustics and Traveltime Tomography. Inverse problems related to underwater acoustics are a critical research area due to their wide range of important practical applications. Solutions of such problems can provide a heat distribution of seawater, an ocean depth at a point, sediment properties of a seabed, locations and/or shapes of submerged objects (mines, submarines, sunken wrecked ships buried to the seafloor, etc.), positions of underwater cables (communication, gas, petroleum, waste, etc.), profiles of seamounts, and a layered seabed for seismic applications

or paths for current flows and migrations of sea animals [21]. In principle, all these applications can roughly be classified into two groups of problems such as remote sensing of passive marine environments and localization/characterization of acoustic sources which are in the sea or buried in the seafloor [171]. For the solution of these problems, one can use the speed of acoustic wave as a tool which is obtained from traveling time of an acoustic wave between two certain points or the amplitude and/or phase information of acoustic waves reflected/scattered from a target(s). In water-type medium, propagating acoustic waves are collected at a selected number of separate hydrophones to obtain measured field data.

Generally, reconstructions of desired parameters from the knowledge of scattered acoustic waves lead to nonlinear and ill-posed inverse problems. Therefore, it is always a complicated issue to find unique and stable solutions, and one has to apply some additional techniques, that is, regularizations, for the proper treatment of the ill-posedness. More specifically, researchers have been applying the methods of Ray's theory, modal travel time/phase inversions, boundary integral equation approaches, Rayleigh's hypothesis, linear sampling, complete family method, convex scattering support theory, and so forth, in underwater acoustics [21, 71, 153, 165, 171, 189, 248–252].

Nondestructive Testing. The inspection of an object without touching or without changing its characteristic properties is a general definition of nondestructive testing (NDT) in the literature. In practice, NDT is commonly performed by following visual, penetrant, magnetic particle, radiographic, thermal infrared, Eddy's current, and ultrasonic type testing procedures for the identifications of cracks, shapes of surface discontinuities, or corrosion damages in power plants, rails, pipelines, bridges, buildings, airplanes, railroad tank cars, and so forth, [196, 253]. The main idea of these methods, which actually shows the conceptual relation between inverse problems and NDT applications, is the evaluation of the responses obtained from the interaction of a test object with a particular effect, that is, electromagnetic/acoustic field, gamma or X-radiation, fluorescent dye, and so forth. Therefore, approaches which are used for inverse problems, that is, synthetic aperture focussing technique (SAFT, [180, 254]), diffraction tomography (DT, [177–179, 255–257]), multiple signal classification (MUSIC, [129, 155, 167, 176, 198]), linear sampling method (LSM, [70–74]), factorization [75, 76, 258], point source [23, 68, 69], no response test [82–84], and so forth, are also employed for NDT problems [259].

More specifically, SAFT is an algorithm which uses the collection of echo signals over a specific aperture to obtain a reconstruction by performing time shifting and superposition of adjacent signals. DT is based on a linear solution of the wave equation which can be obtained via the Born or Rytov approximations. Here, the linearization approaches define mainly the success and the solution space of the inverse problem. MUSIC was initially employed as a direct imaging algorithm to obtain locations of point scatterers [176] and extended to find also the geometry of targets [198]. The method employs the eigenvalue structure of time-reversal matrix which is obtained from measured data at different

receiver antennas. The main idea of linear sampling method is to find an indicator function such that its value provides whether an arbitrarily tested space coordinate lies inside or outside of the object.

Biomedical Imaging Techniques, Tomography, and EIT. Nowadays medical imaging, which can be considered as one of the most developed area of inverse problems in practice, provides high-resolution reconstructions in the order of millimeters. In the last decades, the main effort is given for the implementation of harmless, fast, cheap, robust, and reliable techniques to use in practice for obtaining high-resolution images in real time. In the similar direction, early studies of bioimaging started with the reconstructions of 2D images of human body parts via the inverse Radon transform [145, 162, 260] of measured X-rays which were attenuated inside the body [168, 261]. Afterwards, 3D images were assembled via computed tomography (CT-scan) from a series of X-ray data measured on different planes (sinograms) in 2D [144, 199, 212, 262, 263]. Even though satisfactory results were obtained with the X-ray radiology especially for the bone structures [131, 264] the method was found not sufficiently efficient due to attenuation characteristics of X-rays and harm risks of using ionized radiation on humans. On the other hand, electrical impedance tomography (EIT), after its main idea and formulation were introduced by Calderon in 1980 [163] and D. Isaacson and E. L. Isaacson [207], has gained high interest both from theoretical and physical point of view. In principle, in EIT low-frequency electrical currents are applied to the body part under investigation. Then electrical properties of body tissues are computed from the measurements of electric currents and voltage at the boundary [126, 136, 151, 152, 156, 166, 197, 265–268]. EIT is successfully applied for diagnosis of breast cancer, monitoring brain and gastrointestinal functions, detection of blood clots in the lungs, and so forth [138]. Furthermore, electroencephalography (EEG) and magnetoencephalography (MEG) are used for passive monitoring of neuron activities in the brain from the weak electric or magnetic fields, respectively, [130, 137, 169, 191, 269, 270].

MRI, PET, and SPECT. A different approach which is based on using properties of subatomic particles with the connection to electromagnetism opened a new area for obtaining high spatial resolutions in bioimaging, for example, magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). In MRI, the patient stays in a tunnel under a strong magnetic field typically 0.2–1.5 T. This large static magnetic field aligns protons of many atoms either parallel or antiparallel existing in the body. In the meantime, weak radio frequency fields are applied systematically to the patient for altering the alignment of the magnetization. As a result of this procedure, rotating magnetic fields induce a voltage at the receiver coils of the magnet which is used to reconstruct the image of the scanning area [131, 192, 271–274]. In PET and in SPECT, chosen molecules are labeled with radioactive atoms having short half-life and injected to the patients' bloodstream at very low concentrations in order not to violate radiation exposure limits. For PET, labeled atoms are chosen

to emit positrons, and for SPECT, they are chosen to emit photons when they are decaying. Gamma rays, which occur when an electron and an emitted positron annihilate in PET, and photons which are released in SPECT can be visualized out of body by using scintigraphic detectors for clinical applications of oncology, cardiology, pharmacology, and so forth [127, 175, 275–282].

MW/Ultrasound Tomography, Optical Imaging, and Cognition. Another group of techniques such as microwave tomography, ultrasound, and optical imaging, which are commonly used for solutions of inverse problems in different areas, are also applied to biomedical applications especially for investigating soft body tissues [135, 148, 149, 185, 210, 283–289]. For instance, optical tomography is used for the detection of cancerous cells in breast and brain. Acoustic waves are employed for the imaging of liver, kidney, fetus in pregnant women, and so forth, and microwaves are used in mammography and diagnosis of leukemia [131, 144].

Data Assimilation. Over the past two decades, it has become feasible to simulate atmospheric and geophysical processes from large-scale atmospheric flow down to convective processes on a kilometer scale. This led to the need to determine initial conditions for simulations and forecasts from a collection of diverse direct and indirect measurements, and the field of *data assimilation* arose.

Inverse Problems in Biological and Environmental Applications. Inverse problems are of growing importance in many parts of medicine or biology as well as in environmental applications. Here, we will provide a brief introduction into two areas, first into recent results of neural field theory and second into the basic setup of data assimilation as it is, for example, used in *numerical weather prediction* or for *climate projections*, which usually incorporates various inverse problems.

4.1. Inverse Neural Field Theory. Neural activity is often modelled by the activity potential $u(x)$ in some domain D . The activity potential u satisfies some integro-differential equations, which in its simplest form has been suggested by Wilson and Cowan [290, 291] and Amari [128]:

$$\tau \frac{\partial u(x, t)}{\partial t} = -u(x, t) + \int_D w(x, y) f(u(y, t)) dy, \quad (41)$$

$$x \in D, \quad t \geq 0,$$

where f is some nonlinear function which usually is used to model the firing rate of neurons in dependence of their current action potential and $\tau > 0$ is a constant. The kernel $w(x, y)$ models the strength of the influence of an excitation at point y to the neural field at point x . The first term of the right-hand side of (41) generates the decay of the activity potential u in the absence of excitation or inhibition. Neural fields have been widely studied in recent years, with applications to a wide range of medical phenomena starting

from electroencephalogram (EEG) and magnetoencephalogram (MEG) rhythms to robotic behaviour—for an extensive literature list we refer to [172].

Neural network, that is, the discrete version of neural fields, has attracted strong interest over many years and is a standard tool today in many applied parts of science. Neural networks have also been used to solve inverse problems, compare, for example, [292, 293] for some recent papers and further citations. However, here we want to look into inverse problems which arise in the modelling of neural activities themselves. Training of neural networks or neural fields is, in general, an ill-posed inverse problem, as we will see in due course.

Inverse neural field theory investigates the construction of connectivity kernels w given some dynamics $u(x, t)$ for $x \in D$ and some time intervals $t \in [0, T]$. This so-called *full-field neural inverse problem* is linear and ill-posed in the sense of Hadamard, as can be readily seen by the following transform. We define

$$\begin{aligned} \psi(x, t) &:= \tau \frac{\partial u(x, t)}{\partial t} + u(x, t), \quad x \in D, \quad t \in [0, T], \\ \varphi(y, t) &:= f(u(y, t)), \quad y \in D, \quad t \in [0, T]. \end{aligned} \quad (42)$$

If we further define the operator

$$(W\varphi)(x) := \int_D w(x, y) \varphi(y) dy, \quad x \in D, \quad (43)$$

equation (41) obtains the form

$$\psi(\cdot, t) = W\varphi(\cdot, t), \quad t \in [0, T]. \quad (44)$$

The task is to find the operator W given the family of states $\varphi(\cdot, t)$ and corresponding images $\psi(\cdot, t)$ for $t \in [0, T]$. Changing our perspective slightly, introducing the operator

$$(Kg)(t) := \int_D \varphi(y, t) g(y) dy, \quad t \in [0, T], \quad (45)$$

we transform (41) or (44) into

$$\psi(x, t) = Kw(x, \cdot)(t), \quad t \in [0, T] \quad (46)$$

for each fixed $x \in D$. For each $x \in D$, (46) is an integral equation of the first kind for the function $w(x, \cdot)$. Usually we need some smoothness of the potential u with respect to its arguments $y \in D$ and $t \in [0, T]$. In this case, the operator K is an integral operator with continuous kernel, which is known to be *compact* in either $L^2(D)$ or $C(D)$ (cf. [22]) and cannot have a bounded inverse. Thus, the inverse neural field problem is ill-posed.

The particular form (46) provides a basis for kernel construction, that is, for the solution of the neural field problem. We may apply spectral regularization schemes as described in [1, 22, 85], for example, Tikhonov regularization in a similar sense of (21)

$$R_\alpha := (\alpha I + K^* K)^{-1} K^*, \quad \alpha > 0, \quad (47)$$

with regularization parameter $\alpha > 0$ as regularized inverse to calculate $w(x, \cdot)$ from the knowledge of ψ and φ . For smooth dynamics, the problem is exponentially ill-posed. We refer to [139–141, 294] for the analysis and many examples for the inverse neural field problem. The problem of the ill-posedness of the neural inverse problem is addressed in [295], where a dimensional reduction approach is suggested to decompose the large and strongly ill-posed full problem into more stable individual tasks.

4.2. Data Assimilation with an Application to Numerical Weather Prediction (NWP). In inverse problems, we are usually given measured data, and the task is to gain insight into some unknown parameter functions insight of an inaccessible body or area of space or to reconstruct the shape of scatterers or inclusions, but often, the quantities under reconstruction are not static but *dynamic* and change over time. Then, the inverse task is not only carried out once but the reconstruction is repeated over time with some cycling frequency given by a time interval δT .

Let $f_k \in Y$ for $k \in \mathbb{N}$ be an element of a Hilbert space Y representing our measurement data. The task is to reconstruct some state $x_k \in X$ in a Hilbert space X , where the measurement is described by an *observation operator* $H_k : X \rightarrow Y$. An underlying *dynamical system* is given by some operator $M_k : X \rightarrow X$, mapping the state x_k at time t_k onto the state x_{k+1} at time t_{k+1} . With the reconstruction $x_{k-1}^{(a)}$ from the previous time step, we can calculate a state $x_k^{(b)}$ at time t_k , which serves as a priori knowledge for the current reconstruction when data f_k are given.

The goal of data assimilation in its simplest form is to employ measurement data f_k , the operator H_k , and the background state x_k to calculate a so-called *analysis* $x_k^{(a)}$, which is the best possible estimate of the true state $x_k^{(\text{true})}$ of the dynamical system at time t_k under the given assumptions.

Data assimilation is needed as soon as dynamic situations are studied. For example, the reconstruction of a current density from magnetic fields leads to static magnetic tomography, as, for example, solved in [193–195, 296], but when the underlying currents show a dynamic behaviour, we need to carry out *dynamic magnetic tomography* [297, 298], which is a data assimilation problem.

The data assimilation scheme can also use measurements from an interval of time steps, for example, a window consisting of L time steps with data $f_{k-L}, f_{k-L+1}, \dots, f_k$. In this case, the task is to fit the trajectory given by

$$\begin{aligned} & Hx_{k-L}, \\ Hx_{k-L+1} &= HM_{k-L}x_{k-L}, \\ Hx_{k-L+2} &= HM_{k-L+1}M_{k-L}x_{k-L}, \dots, \\ Hx_k &= HM_{k-1}M_{k-2} \cdots M_{k-L}x_{k-L} \end{aligned} \quad (48)$$

to the measured data f_{k-L}, \dots, f_k given the a priori information $x_{k-L}^{(b)}$.

Let us denote the analysis or reconstruction, respectively, for this window $[t_{k-L}, t_k]$ at time t_{k-L} by $\tilde{x}_{k-L}^{(a,L)}$. Then, it leads to a state estimate at time t_k by

$$x_k^{(a)} := M_{k-1}M_{k-2} \cdots M_{k-L}\tilde{x}_{k-L}^{(a,L)}, \quad k \in \mathbb{N}. \quad (49)$$

An algorithm, which calculates $\tilde{x}_{k-L}^{(a)}$ at the beginning of the window $[t_{k-L}, t_k]$, is usually denoted as *smoother*. The calculation of $x_k^{(a)}$ at the end of the interval is called a *filter*. There is a wide range of literature about data assimilation, in particular from the perspective of filtering. We refer to [133, 134, 173, 209, 213, 299–304] for introduction and further details, in particular in atmospheric sciences.

We can reformulate the data assimilation task as a minimization, here written in its usual finite-dimensional version. If we carry out minimization independently at every time t_k , the scheme is known as *three-dimensional variational assimilation (3dVar)* and minimizes

$$J(x_k) := \|x_k - x_k^{(b)}\|_{B^{-1}}^2 + \|f_k - Hx_k\|_{R^{-1}}^2, \quad (50)$$

where B and R are weight matrices in X and Y , respectively. In a stochastic framework, B is the covariance matrix of background distribution of states in X and R is the covariance matrix of the data error in Y .

The 3dVar algorithm is basically a version of the well-known Tikhonov regularization in a Hilbert space where weighted norms are used. Also, it can be seen as a version of the Bayes approach for a Gaussian densities; compare [209] for an introduction and [188] for a recent survey.

When a window of measurements is used, the corresponding minimization problem is given by

$$J(x_{k-L}) = \|x_{k-L} - x_{k-L}^{(b)}\|_{B^{-1}}^2 + \sum_{\ell=1}^L \|f_{k-L+\ell} - Hx_{k-L+\ell}\|_{R^{-1}}^2, \quad (51)$$

where $x_{k-L+\ell}$ is given by (48). The minimization of (51) is denoted as *four-dimensional variational assimilation (4dVar)*. The 4dVar algorithm is known to be expensive both in terms of computing time and programming effort, since usually the adjoint tangent linear model of the dynamics needs to be calculated and implemented, but 4dVar has turned out to be the algorithm which provides best scores in the area of numerical weather forecasting (NWP); compare the scores which are available from the *European Centre for Medium Range Weather Forecast (ECMWF)* [174].

For linear systems the minimization of the 4dVar functional (51) can be carried out iteratively with individual steps as in (50). The key difference to 3dVar is a dynamic update of the covariance matrix B in every time step. The method is the well-known *Kalman Filter (KF)* [213]. The Kalman filter is often studied in a stochastic framework, for example, [209, 304], where it calculates the maximum probability estimator (MAP) for the Bayes formula in a Gaussian framework and when the prior density has a covariance matrix B and mean $x_k^{(b)}$. A simple algebraic equivalence proof of the Kalman smoother/filter to 4dVar is presented, for example,

in [188]. We need to remark that for large-scale systems, the update of the B matrix in the Kalman filter is not feasible, which has led to the development of ensemble and particle methods.

Ensemble or particle methods which are well-known in stochastic estimation since around 1980 (Markov Chain Monte Carlo methods, MCMC) have become very popular in the area of data assimilation. The basic idea here is to estimate the covariance matrix B by a stochastic covariance estimator, based on a set of ensembles or particles. Then, an assimilation step as in (50) is carried out. The method is known as *ensemble Kalman filter* (EnKF). Here, a particle or ensemble member, respectively, is basically a state x_{k-1} , and its trajectory $x_k = M_{k-1}x_{k-1}$. When a distribution of states is given, we can calculate mean and covariance by standard tools. The propagation of these states through time using the dynamical system M leads to an estimate of mean and covariance of the propagated distribution at a later point in time.

The estimation of the covariance matrix B by an ensemble with 40, 100, or 200 members in a high-dimensional space leads to spurious covariances, which make the ensemble approach useless for such systems. As a consequence, ensemble data assimilation systems have employed *localization* [132, 181, 200, 305, 306]. Here, calculations are carried out in a local region only, not the global space [307, 308]. The local ensemble Kalman filter (LEKF) is studied in [306], and its computationally more efficient version has been developed by Hunt et al. [201], the *local ensemble transformed Kalman filter* (LETKF).

More recently, there has been increasing interest in the study of stability of data assimilation algorithms over time, in particular when ill-posed operators are involved in the observation process; compare [157, 158, 309, 310]. The key object here is to study the mapping of the data sequence onto the analysis sequence

$$(f_k)_{k \in \mathbb{N}} \mapsto (x_k^{(a)})_{k \in \mathbb{N}}. \quad (52)$$

Stability—estimating $\|x_k^{(a)}\|$ for $k \rightarrow \infty$ —has been derived under particular conditions, leading to tools to control the stability over time by scaling the background covariance matrix B (or alternatively the data covariance matrix R) that have been developed [311].

5. Conclusions

In this paper, inverse problems are reviewed starting with a historical perspective. Mathematical backgrounds for the fundamental inverse problems in acoustics, electromagnetics, and elastics are given. The main ideas of many algorithms whose aims to find locations, shapes, and boundary-parameters of obstacles are introduced by studying two dimensional acoustic cases for the sake of simplicity of the theoretical investigations.

Furthermore, the results of the study of the inverse problems for the time-dependent differential equations of isotropic and anisotropic elasticity are presented chronologically from one-dimensional kinematic and dynamic inverse

problems with essentially over-determined data up to multidimensional inverse problems with a finite number of observations. The physical and geophysical interpretations are supplied for almost all reviewed works.

We note that the study of one-dimensional and linearized multidimensional inverse problems for equations of elastodynamics contains the theory (existence, uniqueness, and stability theorems) and computational methods of finding solutions. The study of the multidimensional dynamic problems for the differential equations of elasticity contains mainly the investigations on the uniqueness of the solutions as well as stability to the variation of data. At the same time, computational methods of solving the multidimensional inverse dynamic problems are not as widely developed. The theory and methods of solving the direct and inverse problems for classical equations of elastodynamics have multiple applications in geophysics and engineering.

In the last part of the paper, practical applications of inverse problems are summarized. Moreover, a brief introduction into recent results of neural field theory is provided since inverse problems arising in cognitive sciences have become popular in an interdisciplinary community of cognitive neuroscience. As a final application basic setup of data assimilation as it is, for example, used in numerical weather prediction or for climate projections is investigated in details.

References

- [1] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Vieweg Mathematics for Scientists and Engineers, Friedrich Vieweg & Sohn, Braunschweig, Germany, 1993.
- [2] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, Conn, USA, 1923.
- [3] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, vol. 120 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1996.
- [4] R. G. Newton, "Inverse problems in physics," *SIAM Review*, vol. 12, pp. 346–356, 1970.
- [5] R. L. Parker, "Understanding inverse theory," *Annual Review of Earth and Planetary Sciences*, vol. 5, pp. 35–64, 1977.
- [6] F. Yaman, *Inverse scattering problems for the objects buried in penetrable cylinders [Ph.D. thesis]*, Istanbul Technical University, 2009.
- [7] W. M. Boerner, A. K. Jordan, and I. W. Kay, "Introduction to the special issue on inverse methods in electromagnetics," *IEEE Transactions on Antennas and Propagation*, vol. 29, no. 2, pp. 185–189, 1981.
- [8] F. Cakoni, "Recent developments in the qualitative approach to inverse electromagnetic scattering theory," *Journal of Computational and Applied Mathematics*, vol. 204, no. 2, pp. 242–255, 2007.
- [9] F. Cakoni and D. Colton, *Qualitative Methods in Inverse Scattering Theory*, Interaction of Mechanics and Mathematics, Springer, Berlin, Germany, 2006.
- [10] D. Colton, J. Coyle, and P. Monk, "Recent developments in inverse acoustic scattering theory," *SIAM Review*, vol. 42, no. 3, pp. 369–414, 2000.
- [11] D. Colton, "Inverse acoustic and electromagnetic scattering theory," in *Inside Out: Inverse Problems and Applications*, vol.

- 47 of *Mathematical Sciences Research Institute Publications*, pp. 67–110, Cambridge University Press, Cambridge, UK, 2003.
- [12] D. Colton and R. Kress, “Using fundamental solutions in inverse scattering,” *Inverse Problems*, vol. 22, no. 3, pp. R49–R66, 2006.
- [13] C. W. Groetsch, “Integral equations of the first kind, inverse problems and regularization: a crash course,” *Journal of Physics*, vol. 73, no. 1, Article ID 012001, pp. 1–32, 2007.
- [14] A. Hasanov, “Some new classes of inverse coefficient problems in non-linear mechanics and computational material science,” *International Journal of Non-Linear Mechanics*, vol. 46, pp. 667–684, 2011.
- [15] S. I. Kabanikhin, “Definitions and examples of inverse and ill-posed problems,” *Journal of Inverse and Ill-Posed Problems*, vol. 16, no. 4, pp. 317–357, 2008.
- [16] R. Kress, “Numerical methods in inverse obstacle scattering,” in *Inverse Problems of Wave Propagation and Diffraction*, vol. 486 of *Lecture Notes in Physics*, pp. 93–106, Springer, Berlin, Germany, 1997.
- [17] R. Kress, “Integral equation methods in inverse obstacle scattering,” *The ANZIAM Journal*, vol. 42, no. 1, pp. 65–78, 2000.
- [18] R. Kress, “Uniqueness and numerical methods in inverse obstacle scattering,” *Journal of Physics*, vol. 73, no. 1, Article ID 012003, pp. 1–16, 2007.
- [19] R. Potthast, “A survey on sampling and probe methods for inverse problems,” *Inverse Problems*, vol. 22, no. 2, pp. R1–R47, 2006.
- [20] O. Scherzer, Ed., *Handbook of Mathematical Methods in Imaging*, Springer, New York, NY, USA, 2010.
- [21] J. L. Buchanan, R. P. Gilbert, A. Wirgin, and Y. S. Xu, *Marine Acoustics: Direct and Inverse Problems*, SIAM Society for Industrial and Applied Mathematics, Philadelphia, Pa, USA, 2004.
- [22] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, vol. 93 of *Applied Mathematical Sciences*, Springer, Berlin, Germany, 2nd edition, 1998.
- [23] R. Potthast, *Point Sources and Multipoles in Inverse Scattering Theory*, vol. 427 of *Chapman & Hall/CRC Research Notes in Mathematics*, Chapman & Hall/CRC, Boca Raton, Fla, USA, 2001.
- [24] M. Hanke, A. Neubauer, and O. Scherzer, “A convergence analysis of the Landweber iteration for nonlinear ill-posed problems,” *Numerische Mathematik*, vol. 72, no. 1, pp. 21–37, 1995.
- [25] R. Kress, *Linear Integral Equations*, vol. 82 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2nd edition, 1999.
- [26] F. Hettlich, “An iterative method for the inverse scattering problem from sound-hard obstacles,” in *Proceedings of the International Council for Industrial and Applied Mathematics (ICIAM '95)*, O. Mahrenholtz and R. Mennicken, Eds., vol. 2, Akademie, Berlin, Germany, 1995.
- [27] T. Hohage, “Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem,” *Inverse Problems*, vol. 13, no. 5, pp. 1279–1299, 1997.
- [28] T. Hohage, “Convergence rates of a regularized Newton method in sound-hard inverse scattering,” *SIAM Journal on Numerical Analysis*, vol. 36, no. 1, pp. 125–142, 1998.
- [29] A. Kirsch, “The domain derivative and two applications in inverse scattering theory,” *Inverse Problems*, vol. 9, no. 1, pp. 81–96, 1993.
- [30] R. Kress, “A Newton method in inverse obstacle scattering,” in *Inverse Problems in Engineering Mechanics*, H. D. Bui and M. Tanaka, Eds., pp. 425–432, Balkema, Rotterdam, The Netherlands, 1994.
- [31] R. Kress, “Integral equation methods in inverse obstacle scattering,” *Engineering Analysis with Boundary Elements*, vol. 15, no. 2, pp. 171–179, 1995.
- [32] R. Kress, “Integral equation methods in inverse acoustic and electromagnetic scattering,” in *Boundary Integral Formulations for Inverse Analysis*, D. B. Ingham and L. C. Wrobel, Eds., Advances in Boundary Elements Series, pp. 67–92, Computational Mechanics, Southampton, UK, 1997.
- [33] L. Mönch, “A Newton method for solving the inverse scattering problem for a sound-hard obstacle,” *Inverse Problems*, vol. 12, no. 3, pp. 309–323, 1996.
- [34] W. Tobocman, “Inverse acoustic wave scattering in two dimensions from impenetrable targets,” *Inverse Problems*, vol. 5, no. 6, pp. 1131–1144, 1989.
- [35] S. L. Wang and Y. M. Chen, “An efficient numerical method for exterior and interior inverse problems of Helmholtz equation,” *Wave Motion*, vol. 13, no. 4, pp. 387–399, 1991.
- [36] C. Farhat, R. Tezaur, and R. Djellouli, “On the solution of three-dimensional inverse obstacle acoustic scattering problems by a regularized Newton method,” *Inverse Problems*, vol. 18, no. 5, pp. 1229–1246, 2002.
- [37] H. Harbrecht and T. Hohage, “Fast methods for three-dimensional inverse obstacle scattering problems,” *Journal of Integral Equations and Applications*, vol. 19, no. 3, pp. 237–260, 2007.
- [38] T. Hohage, *Iterative methods in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems [Dissertation]*, Linz, Austria, 1999.
- [39] O. Ivanyshyn, R. Kress, and P. Serranho, “Huygens’ principle and iterative methods in inverse obstacle scattering,” *Advances in Computational Mathematics*, vol. 33, no. 4, pp. 413–429, 2010.
- [40] O. Ivanyshyn, “Shape reconstruction of acoustic obstacles from the modulus of the far field pattern,” *Inverse Problems and Imaging*, vol. 1, no. 4, pp. 609–622, 2007.
- [41] O. Ivanyshyn, *Nonlinear boundary integral equations in inverse scattering [Ph.D. thesis]*, University of Gottingen, 2007.
- [42] O. Ivanyshyn and R. Kress, “Identification of sound-soft 3D obstacles from phaseless data,” *Inverse Problems and Imaging*, vol. 4, no. 1, pp. 131–149, 2010.
- [43] O. Ivanyshyn and T. Johansson, “Nonlinear integral equation methods for the reconstruction of an acoustically sound-soft obstacle,” *Journal of Integral Equations and Applications*, vol. 19, no. 3, pp. 289–308, 2007.
- [44] O. Ivanyshyn and R. Kress, “Nonlinear integral equations in inverse obstacle scattering,” in *Mathematical Methods in Scattering Theory and Biomedical Engineering*, D. I. Fotiatis and C. V. Massalas, Eds., pp. 39–50, World Scientific, Hackensack, NJ, USA, 2006.
- [45] O. Ivanyshyn and R. Kress, “Nonlinear integral equations for solving inverse boundary value problems for inclusions and cracks,” *Journal of Integral Equations and Applications*, vol. 18, no. 1, pp. 13–38, 2006.
- [46] T. Johansson and B. D. Sleeman, “Reconstruction of an acoustically sound-soft obstacle from one incident field and the far-field pattern,” *IMA Journal of Applied Mathematics*, vol. 72, no. 1, pp. 96–112, 2007.
- [47] R. Kress and W. Rundell, “A quasi-Newton method in inverse obstacle scattering,” *Inverse Problems*, vol. 10, no. 5, pp. 1145–1157, 1994.

- [48] R. Kress and W. Rundell, "Inverse obstacle scattering using reduced data," *SIAM Journal on Applied Mathematics*, vol. 59, no. 2, pp. 442–454, 1999.
- [49] R. Kress and W. Rundell, "Nonlinear integral equations and the iterative solution for an inverse boundary value problem," *Inverse Problems*, vol. 21, no. 4, pp. 1207–1223, 2005.
- [50] R. Potthast, "On the convergence of a new Newton-type method in inverse scattering," *Inverse Problems*, vol. 17, no. 5, pp. 1419–1434, 2001.
- [51] R. D. Murch, D. G. H. Tan, and D. J. N. Wall, "Newton-Kantorovich method applied to two-dimensional inverse scattering for an exterior Helmholtz problem," *Inverse Problems*, vol. 4, no. 4, pp. 1117–1128, 1988.
- [52] A. Roger, "Newton-Kantorovitch algorithm applied to an electromagnetic inverse problem," *IEEE Transactions on Antennas and Propagation*, vol. 29, no. 2, pp. 232–238, 1981.
- [53] D. Colton and P. Monk, "A novel method for solving the inverse scattering problem for time-harmonic acoustic waves in the resonance region. II," *SIAM Journal on Applied Mathematics*, vol. 46, no. 3, pp. 506–523, 1986.
- [54] D. Colton and P. Monk, "A novel method for solving the inverse scattering problem for time-harmonic acoustic waves in the resonance region," *SIAM Journal on Applied Mathematics*, vol. 45, no. 6, pp. 1039–1053, 1985.
- [55] D. Colton and P. Monk, "The numerical solution of the three-dimensional inverse scattering problem for time harmonic acoustic waves," *Society for Industrial and Applied Mathematics*, vol. 8, no. 3, pp. 278–291, 1987.
- [56] A. Kirsch and R. Kress, "On an integral equation of the first kind in inverse acoustic scattering," in *Inverse Problems*, J. R. Cannon and U. Hornung, Eds., vol. 77 of *Internationale Schriftenreihe zur Numerischen Mathematik*, pp. 93–102, Birkhäuser, Basel, Switzerland, 1986.
- [57] A. Kirsch and R. Kress, "A numerical method for an inverse scattering problem," in *Inverse and Ill-Posed Problems*, H. W. Engl and C. W. Groetsch, Eds., vol. 4 of *Notes and Reports in Mathematics in Science and Engineering*, pp. 279–289, Academic Press, Boston, Mass, USA, 1987.
- [58] A. Kirsch and R. Kress, "An optimization method in inverse acoustic scattering," in *Boundary Elements IX, Fluid Flow and Potential Applications*, C. A. Brebbia, W. L. Wendland, and G. Kuhn, Eds., vol. 3, pp. 3–18, Computational Mechanics, Southampton, UK, 1987.
- [59] T. S. Angell, R. E. Kleinman, and G. F. Roach, "An inverse transmission problem for the Helmholtz equation," *Inverse Problems*, vol. 3, no. 2, pp. 149–180, 1987.
- [60] T. S. Angell, R. E. Kleinman, B. Kok, and G. F. Roach, "A constructive method for identification of an impenetrable scatterer," *Wave Motion*, vol. 11, no. 2, pp. 185–200, 1989.
- [61] D. S. Jones and X. Q. Mao, "The inverse problem in hard acoustic scattering," *Inverse Problems*, vol. 5, no. 5, pp. 731–748, 1989.
- [62] R. Kress, "Newton's method for inverse obstacle scattering meets the method of least squares," *Inverse Problems*, vol. 19, no. 6, pp. S91–S104, 2003, Special section on imaging.
- [63] R. Kress and P. Serranho, "A hybrid method for sound-hard obstacle reconstruction," *Journal of Computational and Applied Mathematics*, vol. 204, no. 2, pp. 418–427, 2007.
- [64] P. Serranho, "A hybrid method for inverse scattering for shape and impedance," *Inverse Problems*, vol. 22, no. 2, pp. 663–680, 2006.
- [65] P. Serranho, "A hybrid method for inverse scattering for sound-soft obstacles in R^3 ," *Inverse Problems and Imaging*, vol. 1, no. 4, pp. 691–712, 2007.
- [66] P. Serranho, *A hybrid method for inverse obstacle scattering problems [Ph.D. thesis]*, University of Gottingen, 2007.
- [67] F. Yaman, "Location and shape reconstructions of sound-soft obstacles buried in penetrable cylinders," *Inverse Problems*, vol. 25, no. 6, Article ID 065005, 17 pages, 2009.
- [68] R. Potthast, "A fast new method to solve inverse scattering problems," *Inverse Problems*, vol. 12, no. 5, pp. 731–742, 1996.
- [69] R. Potthast, "A point source method for inverse acoustic and electromagnetic obstacle scattering problems," *IMA Journal of Applied Mathematics*, vol. 61, no. 2, pp. 119–140, 1998.
- [70] T. Arens, "Why linear sampling works," *Inverse Problems*, vol. 20, no. 1, pp. 163–173, 2004.
- [71] D. Colton and A. Kirsch, "A simple method for solving inverse scattering problems in the resonance region," *Inverse Problems*, vol. 12, no. 4, pp. 383–393, 1996.
- [72] D. Colton, M. Piana, and R. Potthast, "A simple method using Morozov's discrepancy principle for solving inverse scattering problems," *Inverse Problems*, vol. 13, no. 6, pp. 1477–1493, 1997.
- [73] D. Colton and P. Monk, "A linear sampling method for the detection of leukemia using microwaves," *SIAM Journal on Applied Mathematics*, vol. 58, no. 3, pp. 926–941, 1998.
- [74] R. Kress and L. Kühn, "Linear sampling methods for inverse boundary value problems in potential theory," *Applied Numerical Mathematics*, vol. 43, no. 1-2, pp. 161–173, 2002.
- [75] A. Kirsch, "Characterization of the shape of a scattering obstacle using the spectral data of the far field operator," *Inverse Problems*, vol. 14, no. 6, pp. 1489–1512, 1998.
- [76] A. Kirsch and N. Grinberg, *The Factorization Method for Inverse Problems*, vol. 36 of *Oxford Lecture Series in Mathematics and its Applications*, Oxford University Press, Oxford, UK, 2008.
- [77] R. Potthast, "Stability estimates and reconstructions in inverse acoustic scattering using singular sources," *Journal of Computational and Applied Mathematics*, vol. 114, no. 2, pp. 247–274, 2000.
- [78] M. Ikehata, "Reconstruction of obstacle from boundary measurements," *Wave Motion*, vol. 30, no. 3, pp. 205–223, 1999.
- [79] M. Ikehata, "Reconstruction of an obstacle from the scattering amplitude at a fixed frequency," *Inverse Problems*, vol. 14, no. 4, pp. 949–954, 1998.
- [80] M. Ikehata, "Enclosing a polygonal cavity in a two-dimensional bounded domain from Cauchy data," *Inverse Problems*, vol. 15, no. 5, pp. 1231–1241, 1999.
- [81] M. Ikehata, "Reconstruction of a source domain from the Cauchy data," *Inverse Problems*, vol. 15, no. 2, pp. 637–645, 1999.
- [82] D. R. Luke and R. Potthast, "The no response test—a sampling method for inverse scattering problems," *SIAM Journal on Applied Mathematics*, vol. 63, no. 4, pp. 1292–1312, 2003.
- [83] R. Potthast, "A set-handling approach for the no-response test and related methods," *Mathematics and Computers in Simulation*, vol. 66, no. 4-5, pp. 281–295, 2004.
- [84] R. Potthast, "On the convergence of the no response test," *SIAM Journal on Mathematical Analysis*, vol. 38, no. 6, pp. 1808–1824, 2007.
- [85] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375 of *Mathematics and its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1996.

- [86] R. Potthast, "A study on orthogonality sampling," *Inverse Problems*, vol. 26, no. 7, Article ID 074015, 2010.
- [87] K. Ito, B. Jin, and J. Zou, "A direct sampling method to an inverse medium scattering problem," *Inverse Problems*, vol. 28, no. 2, Article ID 025003, 11 pages, 2012.
- [88] K. Ito, B. Jin, and J. Zou, "A Two-stage method for inverse medium scattering," *Journal of Computational Physics*, vol. 237, pp. 211–223, 2013.
- [89] R. Griesmaier, "Multi-frequency orthogonality sampling for inverse obstacle scattering problems," *Inverse Problems*, vol. 27, no. 8, Article ID 085005, 23 pages, 2011.
- [90] D. Colton and A. Kirsch, "The determination of the surface impedance of an obstacle from measurements of the far field pattern," *SIAM Journal on Applied Mathematics*, vol. 41, no. 1, pp. 8–15, 1981.
- [91] H. Haddar and R. Kress, "On the Fréchet derivative for obstacle scattering with an impedance boundary condition," *SIAM Journal on Applied Mathematics*, vol. 65, no. 1, pp. 194–208, 2004.
- [92] A. W. Kedzierawski, "The determination of the surface impedance of an obstacle," *Proceedings of the Edinburgh Mathematical Society. Series II*, vol. 36, no. 1, pp. 1–15, 1993.
- [93] I. Akduman and R. Kress, "Direct and inverse scattering problems for inhomogeneous impedance cylinders of arbitrary shape," *Radio Science*, vol. 38, no. 3, pp. 1055–1064, 2003.
- [94] R. Kress, F. Yaman, A. Yapar, and I. Akduman, "Inverse scattering for an impedance cylinder buried in a dielectric cylinder," *Inverse Problems in Science and Engineering*, vol. 17, no. 4, pp. 473–488, 2009.
- [95] L. He, S. Kindermann, and M. Sini, "Reconstruction of shapes and impedance functions using few far-field measurements," *Journal of Computational Physics*, vol. 228, no. 3, pp. 717–730, 2009.
- [96] R. T. Smith, "An inverse acoustic scattering problem for an obstacle with an impedance boundary condition," *Journal of Mathematical Analysis and Applications*, vol. 105, no. 2, pp. 333–356, 1985.
- [97] B. Hassen, F. B. Yosra, and H. Haddar, "Application of the linear sampling method to identify cracks with impedance boundary conditions," *Inverse Problems in Science and Engineering*, vol. 21, no. 2, 2013.
- [98] O. Ivanyshyn and R. Kress, "Inverse scattering for surface impedance from phase-less far field data," *Journal of Computational Physics*, vol. 230, no. 9, pp. 3443–3452, 2011.
- [99] F. Yaman, "Numerical solution of an inverse conductive boundary value problem," *Radio Science*, vol. 43, no. 6, Article ID RS6004, 2008.
- [100] F. Yaman, "Inverse scattering of an arbitrarily shaped buried scatterer with conductive boundary condition," in *Proceedings of the 8th World Congress on Computational Mechanics and 5th European Congress on Computational Methods in Applied Sciences and Engineering*, Venice, Italy, 2008.
- [101] F. Yaman, "A combination of shape and conductivity function reconstruction methods for an inverse boundary value problem," *Wave Motion*, vol. 47, no. 4, pp. 253–263, 2010.
- [102] K. Aki and P. G. Richards, *Quantitative Seismology: Theory and Methods*, vol. 1, W.H. Freedman and Company, San Francisco, Calif, USA, 1980.
- [103] G. Herglotz, "Über die elastizität der erde bei beruecksichtigung ihrer variablen dichte," *Zeitschrift für Mathematik und Physik*, vol. 52, pp. 275–299, 1905.
- [104] E. Wiechert and K. Zoeppritz, "Über Erdbebenwellen," *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und von der Georg-Augusts-Universität*, vol. 4, pp. 415–549, 1907.
- [105] M. L. Gerver and V. M. Markushevich, "Investigation of nonuniqueness under determining of a seismic wave propagation velocity from godo-graph," *Doklady Akademii Nauk SSSR*, vol. 163, no. 6, pp. 1377–1380, 1965 (Russian).
- [106] M. M. Lavrent'ev and V. G. Romanov, "Three linearized inverse problems for hyperbolic equations," *Doklady Akademii Nauk SSSR*, vol. 171, pp. 1279–1281, 1966 (Russian).
- [107] V. G. Romanov, "On reconstruction of a function from integrals along a family of curves," *Siberian Mathematical Journal*, vol. 10, no. 6, pp. 1364–1474, 1969.
- [108] Yu. E. Anikonov, *Some Methods of Investigation of Multidimensional Inverse Problems for Differential Equations*, Nauka, Novosibirsk, Russia, 1978.
- [109] A. L. Bukhgeim, *Volterra Equations and Inverse Problems*, Nauka, Novosibirsk, Russia, 1983.
- [110] R. G. Mukhometov, "On one problem of Riemann's metric reconstruction," *Siberian Mathematical Journal*, vol. 22, no. 3, pp. 119–135, 1981.
- [111] G. Uhlmann, "Travel time tomography," *Journal of the Korean Mathematical Society*, vol. 38, no. 4, pp. 711–722, 2001.
- [112] R. G. Mukhometov and V. G. Romanov, "On the problem of finding an isotropic Riemannian metric in an n -dimensional space," *Doklady Akademii Nauk SSSR*, vol. 243, no. 1, pp. 41–44, 1978 (Russian).
- [113] I. N. Berntein and M. L. Gerver, "On problem of integral geometry for family of geodesic lines and an inverse kinematic problems of seismics," *Doklady Akademii Nauk SSSR*, vol. 243, pp. 302–306, 1978 (Russian).
- [114] V. A. Sharafutdinov, *Integral Geometry of Tensor Fields*, Inverse and Ill-posed Problems Series, VSP, Utrecht, The Netherland, 1994.
- [115] V. A. Sharafutdinov, "Finiteness theorem for the ray transform on a Riemannian manifold," *Inverse Problems*, vol. 11, no. 5, pp. 1039–1050, 1995.
- [116] P. Stefanov and G. Uhlmann, "Rigidity for metrics with the same lengths of geodesics," *Mathematical Research Letters*, vol. 5, no. 1–2, pp. 83–96, 1998.
- [117] V. Sharafutdinov and G. Uhlmann, "On deformation boundary rigidity and spectral rigidity of Riemannian surfaces with no focal points," *Journal of Differential Geometry*, vol. 56, no. 1, pp. 93–110, 2000.
- [118] Yu. E. Anikonov, V. V. Bogdanov, E. Yu. Derevtsov, V. L. Miroshnichenko, N. B. Pivovarova, and L. B. Slavina, "Some approaches to a numerical solution for the multidimensional inverse kinematic problem of seismics with inner sources," *Journal of Inverse and Ill-Posed Problems*, vol. 17, no. 3, pp. 209–238, 2009.
- [119] Yu. E. Anikonov, Yu. S. Volkov, S. B. Gorshkalev, Yu. E. Derevtsov, and S. V. Mal'tseva, "On a horizontal homogeneity criterion in the inverse kinematic problem of seismic," *Vestnik Novosibirskogo Gosudarstvennogo Universiteta. Seriya: Matematika, Mekhanika, Informatika*, vol. 11, no. 3, pp. 3–19, 2011 (Russian).
- [120] V. G. Yakhno, "A two-dimensional inverse problem for a system of dynamical Lamé equations," *Soviet Physics Doklady*, vol. 34, no. 7, pp. 616–617, 1989.
- [121] V. G. Yakhno, "A multidimensional inverse dynamic problem of isotropic elasticity," *Soviet Physics Doklady*, vol. 34, no. 1, pp. 35–36, 1989.

- [122] V. G. Yakhno, *Inverse Problems for Differential Equations of Elasticity*, Nauka, Novosibirsk, Russia, 1990.
- [123] V. G. Yakhno, "Multidimensional inverse problems for hyperbolic equations with point sources," in *Ill-Posed and Inverse Problems*, pp. 443–468, VSP, Zeist, The Netherlands, 2002.
- [124] A. S. Alekseev, "Inverse dynamic seismic problems," in *Some Methods and Algorithms for Interpretation of Geophysical Data*, pp. 9–84, Nauka, Moscow, Russia, 1967.
- [125] A. S. Alekseev and V. I. Dobrinsky, "Some questions of applying inverse dynamical seismology problem," *Mathematical Problems in Geophysics*, vol. 6, no. 2, pp. 7–53, 1975 (Russian).
- [126] G. Alessandrini, "Stable determination of conductivity by boundary measurements," *Applicable Analysis*, vol. 27, no. 1-3, pp. 153–172, 1988.
- [127] A. Alessio and P. Kinahan, "PET image reconstruction," in *Nuclear Medicine*, R. E. Henkin, D. Bova, G. L. Dillehay, S. M. Karesh, J. R. Halama, and R. H. Wagner, Eds., Elsevier, Philadelphia, Pa, USA, 2nd edition, 2006.
- [128] S. Amari, "Dynamics of pattern formation in lateral-inhibition type neural fields," *Biological Cybernetics*, vol. 27, no. 2, pp. 77–87, 1977.
- [129] H. Ammari, E. Iakovleva, and D. Lesselier, "A MUSIC algorithm for locating small inclusions buried in a half-space from the scattering amplitude at a fixed frequency," *Multiscale Modeling & Simulation*, vol. 3, no. 3, pp. 597–628, 2005.
- [130] H. Ammari, G. Bao, and J. L. Fleming, "An inverse source problem for Maxwell's equations in magnetoencephalography," *SIAM Journal on Applied Mathematics*, vol. 62, no. 4, pp. 1369–1382, 2002.
- [131] H. Ammari, *An Introduction to Mathematics of Emerging Biomedical Imaging*, vol. 62 of *Mathématiques & Applications*, Springer, Berlin, Germany, 2008.
- [132] J. L. Anderson, "An ensemble adjustment Kalman filter for data assimilation," *Monthly Weather Review*, vol. 129, no. 12, pp. 2884–2903, 2001.
- [133] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, New York, NY, USA, 1979.
- [134] H. J. Andrew, *Stochastic Processes and Filtering Theory*, Academic Press, New York, NY, USA, 1970.
- [135] S. R. Arridge, "Optical tomography in medical imaging," *Inverse Problems*, vol. 15, no. 2, pp. R41–R93, 1999.
- [136] D. C. Barber and B. H. Brown, "Applied potential tomography," *Journal of Physics E*, vol. 17, no. 9, article no. 002, pp. 723–733, 1984.
- [137] J. S. Barlow, "Computerized clinical electroencephalography in perspective," *IEEE Transactions on Biomedical Engineering*, vol. 26, no. 7, pp. 377–391, 1979.
- [138] R. Bayford and A. Tizzard, "Bioimpedance imaging: an overview of potential clinical applications," *Analyst*, vol. 137, pp. 4635–4643, 2012.
- [139] P. b. Graben, D. Pinotsis, D. Saddy, and R. Potthast, "Language processing with dynamic fields," *Cognitive Neurodynamics*, vol. 2, no. 2, pp. 78–88, 2008.
- [140] P. b. Graben and R. Potthast, "Inverse problems in dynamic cognitive modeling," *Chaos*, vol. 19, no. 1, Article ID 015103, 21 pages, 2009.
- [141] P. b. Graben and R. Potthast, "A dynamic field account to language-related brain potentials," in *Principles of Brain Dynamics: Global State Interactions*, M. Rabinovich, K. Friston, and P. Varona, Eds., MIT Press, Cambridge, Mass, USA, 2012.
- [142] M. I. Belishev and S. A. Ivanov, "Reconstruction of the parameters of a system of connected beams from dynamic boundary measurements," *Zapiski Nauchnykh Seminarov (POMI)*, vol. 324, pp. 20–42, 2005.
- [143] M. Bellassoued, O. Imanuvilov, and M. Yamamoto, "Inverse problem of determining the density and two Lamé coefficients by boundary data," *SIAM Journal on Mathematical Analysis*, vol. 40, no. 1, pp. 238–265, 2008.
- [144] M. Bertero and M. Piana, "Inverse problems in biomedical imaging: modeling and methods of solution," in *Complex Systems in Biomedicine*, A. Quarteroni, L. Formaggia, and A. Veneziani, Eds., pp. 1–33, Springer, Berlin, Germany, 2006.
- [145] G. Beylkin, "The inversion problem and applications of the generalized Radon transform," *Communications on Pure and Applied Mathematics*, vol. 37, no. 5, pp. 579–599, 1984.
- [146] G. Beylkin and R. Burridge, "Linearized inverse scattering problems in acoustics and elasticity," *Wave Motion*, vol. 12, no. 1, pp. 15–52, 1990.
- [147] A. S. Blagoveschenskii, "On the inverse problem in the theory of propagation of seismic waves," *Problems of Mathematical Physics*, vol. 1, pp. 68–81, 1966 (Russian).
- [148] J. C. Bolomey and C. Pichot, "Microwave tomography: from theory to practical imaging systems," *International Journal of Imaging Systems and Technology*, vol. 2, pp. 144–156, 1990.
- [149] J. Bolomey Ch., A. Izadnegahdar, and L. Jofre, "Microwave diffraction tomography for biomedical applications," *IEEE Transactions on Microwave Theory and Techniques*, vol. 30, no. 11, pp. 1998–2000, 1982.
- [150] M. Bonnet and A. Constantinescu, "Inverse problems in elasticity," *Inverse Problems*, vol. 21, no. 2, pp. R1–R50, 2005.
- [151] L. Borcea, "Electrical impedance tomography," *Inverse Problems*, vol. 18, no. 6, pp. R99–R136, 2002.
- [152] L. Borcea and G. C. Papanicolaou, "Low frequency electromagnetic fields in high contrast media," in *Surveys on Solution Methods for Inverse Problems*, D. Colton, H. W. Engl, A. Louis, J. R. McLaughlin, and W. Rundell, Eds., pp. 195–233, Springer, Vienna, Austria, 2000.
- [153] L. Bourgeois, C. Chambeyron, and S. Kusiak, "Locating an obstacle in a 3D finite depth ocean using the convex scattering support," *Journal of Computational and Applied Mathematics*, vol. 204, no. 2, pp. 387–399, 2007.
- [154] W. E. Boyse and J. B. Keller, "Inverse elastic scattering in three dimensions," *Acoustical Society of America. Journal*, vol. 79, no. 2, pp. 215–218, 1986.
- [155] M. Brühl, M. Hanke, and M. S. Vogelius, "A direct impedance tomography algorithm for locating small inhomogeneities," *Numerische Mathematik*, vol. 93, no. 4, pp. 635–654, 2003.
- [156] W. R. Breckon, *Image reconstruction in electrical impedance tomography [Ph.D. thesis]*, Oxford Brookes Polytechnic, 1990.
- [157] C. E. A. Brett, K. F. Lam, K. J. H. Law, D. S. McCormick, M. R. Scott, and A. M. Stuart, "Accuracy and stability of filters for dissipative PDEs," *Physica D*, vol. 245, pp. 34–45, 2013.
- [158] C. E. A. Brett, K. F. Lam, K. J. H. Law, D. S. McCormick, M. R. Scott, and A. M. Stuart, "Stability of filters for the Navier-Stokes equation," <http://arxiv.org/abs/1110.2527>.
- [159] K. P. Bube and R. Burridge, "The one-dimensional inverse problem of reflection seismology," *SIAM Review*, vol. 25, no. 4, pp. 497–559, 1983.
- [160] A. L. Bukhgeim and M. V. Klibanov, "Global uniqueness of class of multidimensional inverse problems," *Soviet Mathematics Doklady*, vol. 24, pp. 244–247, 1981.

- [161] G. Burgers, P. J. Van Leeuwen, and G. Evensen, "Analysis scheme in the ensemble Kalman filter," *Monthly Weather Review*, vol. 126, no. 6, pp. 1719–1724, 1998.
- [162] Y. Cai, K. Shen, and J. Wang, "Application of Radon transform in CT image matching," in *Proceedings of the 16th World Conference on Non-Destructive Testing (WCNDT '04)*, p. 347, 2004.
- [163] A. P. Calderon, "On an inverse boundary value problem," in *Proceedings of the Seminar on Numerical Analysis and its applications to Continuum Physics*, pp. 65–73, Sociedade Brasileira de Matemática, Rio de Janeiro, Brazil, 1980.
- [164] K. Chadan and P. C. Sabatier, *Inverse Problems in Quantum Scattering Theory*, Springer, New York, NY, USA, 1977.
- [165] N. R. Chapman, "Inverse problems in underwater acoustics," in *Handbook of Signal Processing in Acoustics*, D. Havelock, S. Kuwano, and M. Vorladder, Eds., pp. 1723–1735, Springer, New York, NY, USA, 2009.
- [166] M. Cheney, D. Isaacson, and J. C. Newell, "Electrical impedance tomography," *SIAM Review*, vol. 41, no. 1, pp. 85–101, 1999.
- [167] M. Cheney, "The linear sampling method and the MUSIC algorithm," *Inverse Problems*, vol. 17, no. 4, pp. 591–595, 2001.
- [168] Z. H. Cho, J. P. Jones, and M. Singh, *Foundations of Medical Imaging*, Wiley, New York, NY, USA, 1993.
- [169] J. R. Cox Jr., F. M. Nolle, and R. M. Arthur, "Digital analysis of the electroencephalogram, the blood pressure wave, and the electrocardiogram," *Proceedings of the IEEE*, vol. 60, no. 10, pp. 1137–1164, 1972.
- [170] S. Coen, "On the elastic profiles of a layered medium from reflection data," *Journal of the Acoustical Society of America*, vol. 70, no. 1, pp. 172–175, 1981.
- [171] M. D. Collins and W. A. Kuperman, "Inverse problems in ocean acoustics," *Inverse Problems*, vol. 10, no. 5, pp. 1023–1040, 1994.
- [172] S. Coombes, P. B. Graben, and R. Potthast, "Tutorial on neural field theory," in *Neural Field Theory*, P. B. Graben, S. Coombes, R. Potthast, and J. J. Wright, Eds., Springer, New York, NY, USA, 2013.
- [173] P. Courtier, E. Andersson, W. Heckley et al., "The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: formulation," *Quarterly Journal of the Royal Meteorological Society*, vol. 124, no. 550, pp. 1783–1807, 1998.
- [174] "Current and historical scores," <http://www.ecmwf.int>.
- [175] M. Defrise, "A short reader's guide to 3D tomographic reconstruction," *Computerized Medical Imaging and Graphics*, vol. 25, no. 2, pp. 113–116, 2001.
- [176] A. J. Devaney, "Super-resolution processing of multi-static data using time reversal and MUSIC," *Journal of the Acoustical Society of America*. In press.
- [177] A. J. Devaney, "A filtered backpropagation algorithm for diffraction tomography," *Ultrasonic Imaging*, vol. 4, no. 4, pp. 336–350, 1982.
- [178] A. J. Devaney, "A computer simulation study of diffraction tomography," *IEEE Transactions on Biomedical Engineering*, vol. 30, no. 7, pp. 377–386, 1983.
- [179] A. J. Devaney, "Geophysical diffraction tomography," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 22, no. 1, pp. 3–13, 1984.
- [180] S. R. Doctor, T. E. Hall, and L. D. Reid, "SAFT—the evolution of a signal processing technology for ultrasonic testing," *NDT International*, vol. 19, no. 3, pp. 163–167, 1986.
- [181] J. W. Dower, J. W. D. Houtekamer, J. W. D. P. L., and H. L. Mitchell, "Data assimilation using an ensemble Kalman filter technique," *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, 1998.
- [182] M. Eller, V. Isakov, G. Nakamura, and D. Tataru, "Uniqueness and stability in the Cauchy problem for Maxwell and elasticity systems," in *Nonlinear Partial Differential Equations and Their Applications*, vol. 31 of *Studies in Mathematics and Its Applications*, pp. 329–349, North-Holland, Amsterdam, The Netherlands, 2002.
- [183] G. Evensen, "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," *Journal of Geophysical Research*, vol. 99, no. 5, pp. 10–162, 1994.
- [184] L. D. Faddeev, "Inverse problem of quantum scattering theory. II," *Journal of Soviet Mathematics*, vol. 5, no. 3, pp. 334–396, 1976.
- [185] E. C. Fear, S. C. Hagness, P. M. Meaney, M. Okoniewski, and M. A. Stuchly, "Enhancing breast tumor detection with near-field imaging," *IEEE Microwave Magazine*, vol. 3, no. 1, pp. 48–56, 2002.
- [186] V. A. Feofanova and V. G. Yakhno, "Computation of the elastic moduli and density of a two-layer elastic medium using geophysical data on a surface," *Soviet Mining Science*, vol. 25, no. 4, pp. 327–333, 1989.
- [187] V. A. Feofanova and V. G. Yakhno, "Elastic parameters of a laminar mass," *Soviet Mining Science*, vol. 26, no. 6, pp. 493–500, 1990.
- [188] M. Freitag and R. Potthast, "Synergy of inverse problems and data assimilation techniques," in *Large Scale Inverse Problems—Computational Methods and Applications In the Earth Sciences*, M. Cullen, M. A. Freitag, S. Kindermann, and R. Scheichl, Eds., vol. 13 of *Radon Series on Computational and Applied Mathematics*, 2013.
- [189] R. P. Gilbert, T. Scotti, A. Wirgin, and Y. S. Xu, "The unidentified object problem in a shallow ocean," *Journal of the Acoustical Society of America*, vol. 103, no. 3, pp. 1320–1327, 1998.
- [190] S. H. Gray and W. W. Symes, "Stability considerations for one-dimensional inverse problems," *Geophysical Journal*, vol. 80, no. 1, pp. 149–163, 1985.
- [191] M. Hämmäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain," *Reviews of Modern Physics*, vol. 65, no. 2, pp. 413–497, 1993.
- [192] E. M. Haacke, R. W. Brown, M. R. Thompson, and R. Venkatesan, *Magnetic Resonance Imaging, Physical Principles and Sequence Design*, John Wiley & Son, New York, NY, USA, 1999.
- [193] K.-H. Hauer, L. Kühn, and R. Potthast, "On uniqueness and non-uniqueness for current reconstruction from magnetic fields," *Inverse Problems*, vol. 21, no. 3, pp. 955–967, 2005.
- [194] K.-H. Hauer, R. Potthast, T. Wüster, and D. Stolten, "Magnetotomography—a new method for analysing fuel cell performance and quality," *Journal of Power Sources*, vol. 143, no. 1-2, pp. 67–74, 2005.
- [195] K.-H. Hauer and R. Potthast, "Magnetic tomography for fuel cells—current status and problems," *Journal of Physics*, vol. 73, no. 1, Article ID 012008, p. 17, 2007.
- [196] C. J. Hellier, *Handbook of Nondestructive Evaluation*, McGraw-Hill, New York, NY, USA, 2003.
- [197] D. Holder, Ed., *Electrical Impedance Tomography: Methods, History and Applications*, Institute of Physics Publishing, 2004.

- [198] S. Hou, K. Solna, and H. Zhao, "A direct imaging algorithm for extended targets," *Inverse Problems*, vol. 22, no. 4, pp. 1151–1178, 2006.
- [199] G. N. Hounsfield, "Computerized transverse axial scanning (tomography): I. Description of system," *British Journal of Radiology*, vol. 46, no. 552, pp. 1016–1022, 1973.
- [200] P. L. Houtekamer and H. L. Mitchell, "A sequential ensemble Kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.
- [201] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, "Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter," *Physica D*, vol. 230, no. 1–2, pp. 112–126, 2007.
- [202] M. Ikehata, "Inversion formulas for the linearized problem for an inverse boundary value problem in elastic prospection," *SIAM Journal on Applied Mathematics*, vol. 50, no. 6, pp. 1635–1644, 1990.
- [203] M. Ikehata, G. Nakamura, and M. Yamamoto, "Uniqueness in inverse problems for the isotropic Lamé system," *The University of Tokyo. Journal of Mathematical Sciences*, vol. 5, no. 4, pp. 627–692, 1998.
- [204] O. Imanuvilov, V. Isakov, and M. Yamamoto, "An inverse problem for the dynamical Lamé system with two sets of boundary data," *Communications on Pure and Applied Mathematics*, vol. 56, no. 9, pp. 1366–1382, 2003.
- [205] O. Yu. Imanuvilov and M. Yamamoto, "Carleman estimates for the non-stationary Lamé system and the application to an inverse problem," *ESAIM. Control, Optimisation and Calculus of Variations*, vol. 11, no. 1, pp. 1–56, 2005.
- [206] O. Imanuvilov and M. Yamamoto, "Carleman estimates for the non-stationary Lamé system and the application to an inverse problem," in *Book Control Theory of Partial Differential Equations*, vol. 242 of *Lecture Notes Pure and Applied Mathematics*, pp. 337–374, Chapman and Hall-CRC, Boca Raton, Fla, USA, 2005.
- [207] D. Isaacson and E. L. Isaacson, "Comment on Calderon's Paper: on an inverse boundary value problem," *Mathematics of Computation*, vol. 52, no. 186, pp. 553–559, 1989.
- [208] V. Isakov, "A nonhyperbolic Cauchy problem for $\square^b \square^c$ and its applications to elasticity theory," *Communications on Pure and Applied Mathematics*, vol. 39, no. 6, pp. 747–767, 1986.
- [209] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2005.
- [210] L. Jofre, M. S. Hawley, A. Broquetas, E. De Los Reyes, M. Ferrando, and A. R. Elias-Fuste, "Medical imaging with a microwave tomographic scanner," *IEEE Transactions on Biomedical Engineering*, vol. 37, no. 3, pp. 303–312, 1990.
- [211] S. I. Kabanikhin, *Inverse and Ill-Posed Problems: Theory and Applications*, vol. 55 of *Inverse and Ill-posed Problems Series*, Walter de Gruyter, Berlin, Germany, 2012.
- [212] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, IEEE Press, New York, NY, USA, 1988.
- [213] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [214] A. L. Karchevsky, "Reconstruction of pressure velocities and boundaries of thin layers in thinly-stratified layers," *Journal of Inverse and Ill-Posed Problems*, vol. 18, no. 4, pp. 371–388, 2010.
- [215] V. G. Romanov, *Inverse Problems of Mathematical Physics*, VNU Science Press, Utrecht, The Netherlands, 1987.
- [216] E. A. Volkova and V. G. Romanov, "Inverse dynamic problems for an anisotropic elastic medium," *Doklady Akademii Nauk SSSR*, vol. 267, no. 4, pp. 780–783, 1982 (Russian).
- [217] A. E. Yagle and B. C. Levy, "A layer stripping solution of the inverse problem for a one dimensional elastic medium," *Geophysics*, vol. 50, no. 3, pp. 425–433, 1985.
- [218] V. G. Yakhno, "One-dimensional inverse problem of anisotropic elasticity for string sources," *Soviet Physics Doklady*, vol. 30, no. 11, pp. 931–932, 1985.
- [219] M. Meadows and S. Coen, "Exact inversion of plane-layered isotropic and anisotropic elastic media by the state-space approach," *Geophysics*, vol. 51, no. 11, pp. 2031–2050, 1986.
- [220] P. E. Sacks and W. W. Symes, "The inverse problem for a fluid over a layered elastic half-space," *Inverse Problems*, vol. 6, no. 6, pp. 1031–1054, 1990.
- [221] P. E. Sacks and V. G. Yakhno, "The inverse problem for a layered anisotropic half space," *Journal of Mathematical Analysis and Applications*, vol. 228, no. 2, pp. 377–398, 1998.
- [222] A. G. Megrabov, *Forward and Inverse Problems for Hyperbolic, Elliptic, and Mixed Type Equations*, Inverse and Ill-Posed Problems Series, VSP, Utrecht, The Netherlands, 2003.
- [223] A. G. Ramm, *Inverse Problems*, Mathematical and Analytical Techniques with Applications to Engineering, Springer, New York, NY, USA, 2005.
- [224] A. Morassi, G. Nakamura, and M. Sini, "An inverse dynamical problem for connected beams," *European Journal of Applied Mathematics*, vol. 16, no. 1, pp. 83–109, 2005.
- [225] S. Nagayasu, "An inverse problem for the elastic equation in plane-stratified media," *SUT Journal of Mathematics*, vol. 41, no. 2, pp. 117–135, 2005.
- [226] A. Morassi, G. Nakamura, K. Shirota, and M. Sini, "A variational approach for an inverse dynamical problem for composite beams," *European Journal of Applied Mathematics*, vol. 18, no. 1, pp. 21–55, 2007.
- [227] V. G. Romanov, "Determination of the parameters of a layered piecewise-constant medium for an unknown form of an impulse source," *Siberian Mathematical Journal*, vol. 48, no. 6, pp. 1074–1084, 2007.
- [228] V. G. Romanov, "On the problem of determining the parameters of a layered elastic medium and an impulse source," *Siberian Mathematical Journal*, vol. 49, no. 5, pp. 919–943, 2008.
- [229] V. G. Yakhno, "The linearized inverse multidimensional problem of Lamb," *Soviet Mathematics. Doklady*, vol. 29, no. 3, pp. 517–520, 1984.
- [230] R. H. Stolt and A. B. Weglein, "Migration and inversion of seismic data," *Geophysics*, vol. 50, no. 12, pp. 2458–2472, 1985.
- [231] V. A. Sharafutdinov, "Quasi-isotropic approximation in dynamical elasticity and some problems of geotomography," *Russian Geology and Geophysics*, vol. 35, pp. 68–82, 1994.
- [232] V. A. Sharafutdinov, "Linearized inverse problems of determining the parameters of transversally isotropic elastic media from measurements of refracted waves," *Journal of Inverse and Ill-Posed Problems*, vol. 4, no. 3, pp. 245–266, 1996.
- [233] L. G. Steinberg, "The linearization of inverse spectral problems for nonhomogeneous elastic bodies," *Inverse Problems*, vol. 12, no. 4, pp. 483–497, 1996.
- [234] G. Uhlmann, "Inverse boundary value problems and applications," *Astérisque*, vol. 207, no. 6, pp. 153–211, 1992.
- [235] J. Sylvester and G. Uhlmann, "A global uniqueness theorem for an inverse boundary value problem," *Annals of Mathematics*, vol. 125, no. 1, pp. 153–169, 1987.

- [236] J. Sylvester and G. Uhlmann, "Inverse problems in anisotropic media," in *Inverse Scattering and Applications*, vol. 122 of *Contemporary Mathematics*, pp. 105–117, American Mathematical Society, Providence, RI, USA, 1991.
- [237] G. Nakamura and G. Uhlmann, "Identification of Lamé parameters by boundary measurements," *American Journal of Mathematics*, vol. 115, no. 5, pp. 1161–1187, 1993.
- [238] G. Nakamura and G. Uhlmann, "Global uniqueness for an inverse boundary problem arising in elasticity," *Inventiones Mathematicae*, vol. 118, no. 3, pp. 457–474, 1994.
- [239] G. Nakamura, K. Tanuma, and G. Uhlmann, "Layer stripping for a transversely isotropic elastic medium," *SIAM Journal on Applied Mathematics*, vol. 59, no. 5, pp. 1879–1891, 1999.
- [240] L. V. Rachele, "An inverse problem in elastodynamics: uniqueness of the wave speeds in the interior," *Journal of Differential Equations*, vol. 162, no. 2, pp. 300–325, 2000.
- [241] C.-L. Lin and G. Nakamura, "Conditional stability for the hexagonal anisotropic elastic dynamical systems," *Communications in Partial Differential Equations*, vol. 34, no. 10–12, pp. 1251–1264, 2009.
- [242] V. I. Kukulkin and R. S. Mackintosh, "The application of inversion to nuclear scattering," *Journal of Physics G*, vol. 30, no. 2, pp. R1–R55, 2004.
- [243] L. R. Lines and S. Treitel, "Tutorial: a review of least-squares inversion and its application to geophysical problems," *Geophysical Prospecting*, vol. 32, no. 2, pp. 159–186, 1984.
- [244] R. S. Mackintosh, "Inverse scattering: applications to nuclear physics," <http://arxiv.org/abs/1205.0468>.
- [245] R. Snieder and J. Trampert, "Inverse problems in geophysics," in *Wave-Field Inversion*, A. Wirgin, Ed., pp. 119–190, Springer, New York, NY, USA, 1999.
- [246] S. Twomey, *Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurement*, Dover, New York, NY, USA, 1996.
- [247] Y. Wang, "Regularization for inverse models in remote sensing," *Progress in Physical Geography*, vol. 36, no. 1, pp. 38–59, 2012.
- [248] C. Rozier, D. Lesselier, T. S. Angell, and R. E. Kleinman, "Shape retrieval of an obstacle immersed in shallow water from single-frequency farfields using a complete family method," *Inverse Problems*, vol. 13, no. 2, pp. 487–508, 1997.
- [249] T. Scotti and A. Wirgin, "Shape reconstruction of an impenetrable scattering body via the Rayleigh hypothesis," *Inverse Problems*, vol. 12, no. 6, pp. 1027–1055, 1996.
- [250] M. Taroudakis and G. Makrakis, Eds., *Inverse Problems in Underwater Acoustics*, Springer, New York, NY, USA, 2001.
- [251] M. I. Taroudakis, "Inverse problems in underwater acoustics," in *Proceedings of the 2nd EAA International Symposium on Hydroacoustics*, A. Stepnowski, R. Salamon, and E. Kozaczka, Eds., pp. 281–290, 1999.
- [252] Y. Xu and Y. Yan, "Boundary integral equation method for source localization with a continuous wave sonar," *Journal of the Acoustical Society of America*, vol. 92, no. 2 I, pp. 995–1002, 1992.
- [253] R. Prakash, *Non-Destructive Testing Techniques*, New Age Science, 2009.
- [254] A. Seydel, "Ultrasonic synthetic aperture focusing techniques in NDT," in *Research Techniques in Nondestructive Testing*, R. S. Sharpe, Ed., vol. 6, The New York Academy of Sciences, New York, NY, USA, 1983.
- [255] C. Pintavirooj and M. Sangworasil, "Ultrasonic diffraction tomography," *The International Journal of Applied Biomedical Engineering*, vol. 1, no. 1, pp. 34–40, 2008.
- [256] M. Slaney, A. C. Kak, and L. E. Larsen, "Limitations of imaging with first-order diffraction tomography," *IEEE Transactions on Microwave Theory and Techniques*, vol. 32, no. 8, pp. 860–874, 1984.
- [257] R. S. Wu and M. N. Toksoz, "Diffraction tomography and multi-source holography applied to seismic imaging," *Geophysics*, vol. 52, no. 1, pp. 11–25, 1987.
- [258] A. Kirsch, "The MUSIC algorithm and the factorization method in inverse scattering theory for inhomogeneous media," *Inverse Problems*, vol. 18, no. 4, pp. 1025–1040, 2002.
- [259] R. Marklein, J. Miao, M. Rahman, and K. J. Langenberg, "Inverse scattering and imaging in NDT: recent applications and advances," in *Proceedings of the European Conference on Non-Destructive Testing (ECNDT '06)*, Berlin, Germany, 2006.
- [260] J. Radon, "Über die bestimmung von funktionen durch ihre integralwerte längs gewisser mannigfaltigkeiten," *Berichte Über die Verhandlungen der Königlich-Sächsischen Akademie der Wissenschaften Zu Leipzig, Mathematisch-Physische Klasse*, vol. 69, pp. 262–277, 1917.
- [261] S. Webb, *From the Watching of Shadows: The Origins of Radiological Tomography*, Adam Hilger, Bristol, New York, USA, 1990.
- [262] F. Natterer, *The Mathematics of Computerized Tomography*, vol. 32 of *Classics in Applied Mathematics*, SIAM, Philadelphia, Pa, USA, 2001.
- [263] M. J. Paulus, S. S. Gleason, S. J. Kennel, P. R. Hunsicker, and D. K. Johnson, "High resolution X-ray computed tomography: an emerging tool for small animal cancer research," *Neoplasia*, vol. 2, no. 1–2, pp. 62–70, 2000.
- [264] S. Webb, *The Physics of Medical Imaging*, Taylor & Francis, New York, NY, USA, 1988.
- [265] W. R. B. Lionheart, "EIT reconstruction algorithms: pitfalls, challenges and recent developments," *Physiological Measurement*, vol. 25, no. 1, pp. 125–142, 2004.
- [266] K. Paulson, W. Breckon, and M. Pidcock, "Electrode modelling in electrical impedance tomography," *SIAM Journal on Applied Mathematics*, vol. 52, no. 4, pp. 1012–1022, 1992.
- [267] G. J. Saulnier, R. S. Blue, J. C. Newell, D. Isaacson, and P. M. Edic, "Electrical impedance tomography," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 31–43, 2001.
- [268] A. D. Seagar, D. C. Barber, and B. H. Brown, "Electrical impedance imaging," *IEE Proceedings A*, vol. 134, no. 2, pp. 201–210, 1987.
- [269] E. Niedermeyer and F. L. da Silva, Eds., *Electroencephalography: Basic Principles, Clinical Applications and Related Fields*, Lippincott Williams & Wilkins, Philadelphia, Pa, USA, 5th edition, 2005.
- [270] M. Teplan, "Fundamentals of EEG measurements," *Measurement Science Review*, vol. 2, pp. 1–11, 2002.
- [271] G. Katti, S. A. Ara, and A. Shireen, "Magnetic resonance imaging (MRI)—a review," *International Journal of Dental Clinics*, vol. 3, no. 1, pp. 65–70, 2011.
- [272] T. Schaeffter and H. Dahnke, "Magnetic resonance imaging and spectroscopy," in *Molecular Imaging I*, vol. 185 of *Handbook of Experimental Pharmacology*, pp. 75–90, Springer, Berlin, Germany, 2008.
- [273] M. T. Vlaardingerbroek and J. A. Boer, *Magnetic Resonance Imaging: Theory and Practice*, Springer, New York, NY, USA, 3rd edition, 2003.
- [274] C. Westbrook, C. K. Roth, and J. Talbot, *MRI in Practice*, Wiley-Blackwell, Oxford, UK, 4th edition, 2011.

- [275] R. M. Lewitt and S. Matej, "Overview of methods for image reconstruction from projections in emission computed tomography," *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1588–1611, 2003.
- [276] M. T. Madsen, "Recent advances in SPECT imaging," *Journal of Nuclear Medicine*, vol. 48, no. 4, pp. 661–673, 2007.
- [277] T. F. Massoud and S. S. Gambhir, "Molecular imaging in living subjects: seeing fundamental biological processes in a new light," *Genes and Development*, vol. 17, no. 5, pp. 545–580, 2003.
- [278] A. Rahmim and H. Zaidi, "PET versus SPECT: strengths, limitations and challenges," *Nuclear Medicine Communications*, vol. 29, no. 3, pp. 193–207, 2008.
- [279] E. Tanaka, "Recent progress on single photon and positron emission tomography—from detectors to algorithms," *IEEE Transactions on Nuclear Science*, vol. NS-34, no. 1, pp. 313–320, 1986.
- [280] D. W. Townsend, R. A. Isoardi, and B. Bendriem, *The Theory and Practice of 3D PET*, Kluwer Academic Publishers, New York, NY, USA, 1998.
- [281] S. Vandenberghe, Y. D'Asseler, R. Van De Walle et al., "Iterative reconstruction algorithms in nuclear medicine," *Computerized Medical Imaging and Graphics*, vol. 25, no. 2, pp. 105–111, 2001.
- [282] G. L. Zeng, "Image reconstruction—a tutorial," *Computerized Medical Imaging and Graphics*, vol. 25, no. 2, pp. 97–103, 2001.
- [283] R. A. Lemons and C. F. Quate, "Acoustic microscopy: biomedical applications," *Science*, vol. 188, no. 4191, pp. 905–911, 1975.
- [284] B. J. Roth, P. J. Basser, and J. P. Wikswo Jr., "A theoretical model for magneto-acoustic imaging of bioelectric currents," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 8, pp. 723–728, 1994.
- [285] S. Semenov, "Microwave tomography: review of the progress towards clinical applications," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 367, no. 1900, pp. 3021–3042, 2009.
- [286] S. Y. Semenov, R. H. Svenson, A. E. Boulyshev et al., "Microwave tomography: two-dimensional system for biological imaging," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 9, pp. 869–877, 1996.
- [287] S. Y. Semenov, R. H. Svenson, A. E. Bulyshev et al., "Three-dimensional microwave tomography: initial experimental imaging of animals," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 1, pp. 55–63, 2002.
- [288] W. Tabbara, B. Duchêne, Ch. Pichot, D. Lesselier, L. Chommeloux, and N. Joachimowicz, "Diffraction tomography: contribution to the analysis of some applications in microwaves and ultrasonics," *Inverse Problems*, vol. 4, no. 2, pp. 305–331, 1988.
- [289] L. Wang and H.-I. Wu, *Biomedical Optics: Principles and Imaging*, John Wiley & Sons, New York, NY, USA, 2007.
- [290] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1–24, 1972.
- [291] H. R. Wilson and J. D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetik*, vol. 13, no. 2, pp. 55–80, 1973.
- [292] F. Yaman and S. Şimşek, "Neural Network approach to determine non smooth one-dimensional profiles in inverse scattering theory," *Microwave and Optical Technology Letters*, vol. 49, no. 12, pp. 3158–3162, 2007.
- [293] F. Yaman and S. Şimşek, "Shape reconstruction of an inhomogeneous impedance cylinder via potential and Neural Network approach," *Microwave and Optical Technology Letters*, vol. 51, no. 1, pp. 119–124, 2009.
- [294] R. Potthast and P. b. Graben, "Inverse problems in neural field theory," *SIAM Journal on Applied Dynamical Systems*, vol. 8, no. 4, pp. 1405–1433, 2009.
- [295] R. Potthast and P. b. Graben, "Dimensional reduction for the inverse problem of neural held theory," *Frontiers in Neuroscience*, vol. 3, article 17, 2009.
- [296] R. Kress, L. Kühn, and R. Potthast, "Reconstruction of a current distribution from its magnetic field," *Inverse Problems*, vol. 18, no. 4, pp. 1127–1146, 2002.
- [297] B. A. Marx, *Dynamic Magnetic Tomography*, Der Andere, 2011.
- [298] B. A. Marx and R. Potthast, "Data assimilation algorithms for dynamic magnetic tomography," In press.
- [299] W. Lahoz, R. Swinbank, and B. Khattatov, *Data Assimilation: Making Sense of Observations*, Springer, New York, NY, USA, 2010.
- [300] J. M. Lewis, S. Lakshmivarahan, and S. Dhall, *Dynamic Data Assimilation: A Least Squares Approach*, vol. 104 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, UK, 2006.
- [301] A. C. Lorenc, S. P. Ballard, R. S. Bell et al., "The Met. Office global three-dimensional variational data assimilation scheme," *Quarterly Journal of the Royal Meteorological Society*, vol. 126, no. 570, pp. 2991–3012, 2000.
- [302] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [303] C. D. Rodgers, *Inverse Methods for Atmospheric Sounding: Theory and Practice*, vol. 2 of *Series on Atmospheric, Oceanic and Planetary Physics*, World Scientific, River Edge, NJ, USA, 2000.
- [304] A. M. Stuart, "Inverse problems: a Bayesian perspective," *Acta Numerica*, vol. 19, pp. 451–559, 2010.
- [305] J. S. Whitaker and T. M. Hamill, "Ensemble data assimilation without perturbed observations," *Monthly Weather Review*, vol. 130, no. 7, pp. 1913–1924, 2002.
- [306] E. Ott, B. R. Hunt, I. Szunyogh et al., "A local ensemble Kalman filter for atmospheric data assimilation," *Tellus A*, vol. 56, no. 5, pp. 415–428, 2004.
- [307] I. Szunyogh, E. J. Kostelich, G. Gyarmati et al., "Assessing a local ensemble Kalman filter: perfect model experiments with the national centers for environmental prediction global model," *Tellus A*, vol. 57, no. 4, pp. 528–545, 2005.
- [308] I. Szunyogh, E. J. Kostelich, G. Gyarmati et al., "A local ensemble transform Kalman filter data assimilation system for the NCEP global model," *Tellus A*, vol. 60, no. 1, pp. 113–130, 2008.
- [309] B. A. Marx and R. Potthast, "On instabilities in data assimilation algorithms," *GEM-International Journal on Geomathematics*, vol. 3, no. 2, pp. 253–278, 2012.
- [310] R. W. E. Potthast, A. J. F. Moodey, A. S. Lawless, and P. J. van Leeuwen, "On error dynamics and instability in data assimilation," MPS-2012-05, In press, <http://www.reading.ac.uk/mathsand-stats/research/math-preprints.aspx>.
- [311] A. J. F. Moodey, A. S. Lawless, R. W. E. Potthast, and P. J. van Leeuwen, "Nonlinear error dynamics for cycled data assimilation methods," *Inverse Problems*, vol. 29, no. 2, Article ID 025002, 18 pages, 2013.

Research Article

A New Concept for Atmospheric Reentry Optimal Guidance: An Inverse Problem Inspired Approach

Davood Abbasi and Mahdi Mortazavi

Aerospace Department, Center of Excellence in Computational Aerospace, Amirkabir University of Technology, P.O. Box 15875/4413, Teheran, Iran

Correspondence should be addressed to Mahdi Mortazavi; mortazavi@aut.ac.ir

Received 27 July 2012; Revised 29 April 2013; Accepted 21 May 2013

Academic Editor: Fatih Yaman

Copyright © 2013 D. Abbasi and M. Mortazavi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a new concept for atmospheric reentry online optimal guidance and control using a method called MARE G&C that exploits the different time scale featured by reentry dynamics. The new technique reaches a quasi-analytical solution and simplified computations, even considering both lift-to-drag ratio and aerodynamic roll as control variables; in addition, the paper offers a solution for the challenging path constraints issue, getting inspiration from the inverse problem methodology. The final resulting algorithm seems suitable for onboard predictive guidance, a new need for future space missions.

1. Introduction

Atmospheric reentry guidance and control (G&C) has been [1, 2] a significant and ongoing [3–6] research interest in the aerospace arena connected to several domains of engineering and science. The reentry dynamic exhibits four main complicating features. (a) It is governed by a group of time-varying nonlinear differential equations; (b) the states (trajectories) must satisfy some physical-operational constraints like heating rate, heat load, dynamic pressure, and maximum deceleration; (c) the controls have rational limits; and (d) parametric uncertainty is present.

The almost unique reentry features have made it a typical case study for engineers, mathematicians, and scientists who want to show the power of their solving theories or techniques. On the other hand, manned atmospheric reentry G&C is a common denominator of human space exploration, space station operations, and space tourism: all “industries of the future,” that are attracting growing investments from countries worldwide [3, 6].

After some earlier works on finding simple reentry solutions, modern requirements made the reentry problem more difficult and only “finding a solution” for its G&C is not sufficient. Nowadays, there is the need to optimize the reentry trajectory in “some sense” (maximizing downrange, reducing

control use, etc.) and minimize, above all, the final landing point error relative to the previsions.

The other complications arise from the new guidance requirements. The typical Apollo/Shuttle-era strategy (also called path following or drag-based strategy) [2] was the nominal trajectory guidance (the controller tries to reach a predetermined and offline-designed trajectory represented in the drag-velocity plane). Although it had shown to be feasible in several missions (and so used in the ARD project by EU) [7], this method involves a great amount of prelaunch measurements (so great cost) and has little robustness against mission variations and failures, disturbances, and off-design flight parameters.

Therefore, predictive guidance (PG, the controller computes a new trajectory in advance based on the actual flight conditions) [8] has been considered; computing power and numerical algorithms now allow admissible guidance trajectories and control to be computed online, and this is the so-called path-generating strategy. PG methods are a more versatile tool than nominal trajectory guidance and are more flexible to disturbances, or entry conditions and vehicle parameter variations. Additionally, these techniques provide a systematic tool for instantaneously satisfying state and control constraints in the online trajectory generation [9].

To tackle the present challenging problem, two main approaches are developed until now. The direct method relies only on numerical methods; this family of solutions is probably represented in the best way by the works of Betts [10] and Milam [4] and some other authors [11, 12]. The optimal control problem is transformed in a parametric optimization one and then solved by numerical methods. These works are mainly treating the general constrained optimal control problem and they appear still too time-consuming and they do not satisfy the new requirements explained; in addition, they do not take account of the advantages that the physical nature of the problem considered can offer in the solution; for example, in atmospheric reentry, the two-time scale can be very useful.

The other approach, which is theoretical and numerical, is the so-called “indirect approach.” It derives the standard optimal control problem with the aid of the Hamiltonian, the Pontryagin’s minimum principle, and the costate or Lagrange multipliers. Then, the problem is solved by a numerical shooting method [13].

As stated very well by Trélat [14] survey paper, both approaches still present some disadvantages, and extensive efforts are underway to better the time efficiency and the accuracy of these methods, usually combining them with other ones [15]. For the second approach, a main problem is, for example, to guarantee that the global minimum is reached by the solution. For the first, the computation conversion time still appears to be the main challenge. In addition, although the path constraint issue is considered in some papers by these approaches, avoiding these limits especially in path-generating scenarios appears difficult to implement, especially in real problems.

On the other hand, inverse problem methodology, a key approach in the modern science that has attracted worldwide attention, has shown great force in attacking the most difficult engineering problems including reentry. The general idea in this approach is to start from the solution or the typical desired trajectory (so the control) and to obtain the state of the system or its mathematical representation [16]. This is similar to the direct approach discussed previously.

Using inverse problem mentality and so considering the reentry data published and available, one can understand that the states of the reentry dynamic equations exhibit a two-scalar behavior. The mathematical examination of the reentry dynamic equations confirms the fast and slow variables presence. Different methods have tried to employ this useful physical peculiarity of reentry dynamics; singular perturbation theory (SPT) is one of them. In addition, inverse problem approach can be efficiently implied for the solution of one of the main problems of reentry; inverse methodology can be used to consider the limits created by the path constraints similar to how it is used previously for solving other problems [5, 17, 18].

Different methods developed inside SPT have been extensively used in the last decades in the field of optimal control [6, 19]. Shen [20] demonstrated the usefulness of the SPT procedures in reentry path optimization; Shi [21] and Vinh et al. [22] used one of the SPT methods known as matched asymptotic expansion (MAE) to optimize the reentry of a hypersonic vehicle.

The SPT approach presented good and quick solutions and was bettered and reapplied to the reentry problem and other aerospace problems more recently [23, 24]. In most of these works, the equations were priori simplified with considering only the lift or the lift-to-drag ratio as the control. Then, according to MAE method, the general equations were split into outer (Keplerian region) and inner (aerodynamic-predominated region) equations. Finally, the two solutions were matched asymptotically by subtracting their common part, the so-called common solution.

Our work develops the standard optimal control problem and the co-state equations, and therefore it can be classified into the indirect methods family; however, opposite to common indirect methods, it does not use shooting numerical methods. It uses the two-time identity of the physics of the problem (inspired by inverse problem idea) to present for the first time a quasi-analytical solution of the optimal reentry control problem with two controls (lift-to-drag ratio and aerodynamic roll). Thanks to inverse mentality and MAE method, the solution preserves the advantages of the earlier solutions: encouraging precision and speed in solution. The solution considers also the path constraints; to avoid violating them, the technique used here is completely derived from “inverse problem” universe and previously used for different purposes in some Russian references [5, 18].

The resulting controller will be used to introduce the concept of an online PG reentry scheme that is tolerant to uncertainties because it generates the optimum trajectories during the path based on the actual flight conditions; we call it matched asymptotic reentry guidance and control (MARE G&C).

Where required, we will use the data published about ARD [7] capsule successfully experimented by ESA in October 1998, with a shape that seems to be the base of the future space programs for manned reentry [25].

2. A Glance at the Problem Physics

Following the inverse problem approach, we first consider and compare the states of the flight of the typical reentry missions such as ARD [25].

The result of such comparisons between the time change of AoA, velocity, longitude, and latitude suggests that there may be a difference in the rate of change of the different states examined in the reentry dynamics. In fact, such a feature can be very useful for the problem simplification. In other words, considering this and other reentry data available, using the inverse problem mentality, we look for a solution that uses the natural features of the reentry dynamics to make it easily solvable.

3. Reentry Model & Problem Statement

The equations of motion of a nonthrusting vehicle entering a planetary atmosphere assumed to be at rest around a spherical, nonrotating planet (this assumption was made for simplicity, but a similar guidance law can be derived based on

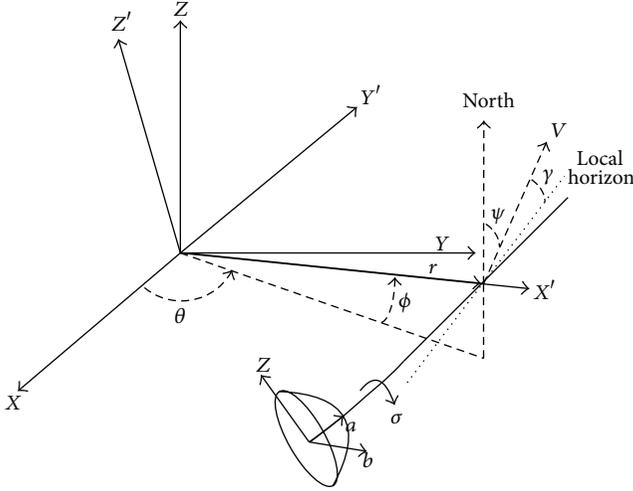


FIGURE 1: Coordinate system.

more precise equations of motion including Coriolis terms due to earth rotation) are first considered (Figure 1) [22]

$$\begin{aligned}
 \frac{dr}{dt} &= V \sin \gamma, \\
 \frac{d\theta}{dt} &= \frac{V \cos \gamma \cos \psi}{r \cos \phi}, \\
 \frac{d\phi}{dt} &= \frac{V \cos \gamma \sin \psi}{r}, \\
 \frac{dV}{dt} &= -\frac{D}{m} - g \sin \gamma, \\
 \frac{d\gamma}{dt} &= \frac{L \cos \sigma}{m} - \left(g - \frac{V^2}{r} \right) \cos \gamma, \\
 \frac{d\psi}{dt} &= \frac{L \sin \sigma}{mV \cos \gamma} - \left(\frac{V}{r} \right) \cos \gamma \cos \psi \tan \phi.
 \end{aligned} \tag{1}$$

An accurate mathematical study of this model confirms our doubts that rose in the previous section. Accurate mathematical theories [6, 23] show that the previous, system exhibits two-time scale properties. Nevertheless, using the experience [23, 24] we look for a more suitable system dynamics.

The first convenient transformation is to rewrite the equations considering height as a new independent variable instead of time:

$$\begin{aligned}
 r &= r_s + y = r_s (1 + h), \\
 \frac{dh}{dt} &= \frac{V \sin \gamma}{r_s},
 \end{aligned} \tag{2}$$

where r_s is the reference height and y is the elevation above it.

Before introducing h as the new independent variable, one must assume that h is absolutely decreasing during the

path. This depends on the initial conditions of reentry, especially initial flight path angle, and on the lift characteristics of the reentry vehicle. It must be clarified that in general the condition that happens for reentry paths that are not steep is that the elevation is absolutely decreasing; moreover, if in a reentry the decreasing elevation becomes increasing, it means that the reentry vehicle had suffered a heavy load; therefore, it is assumed that the mission design makes the elevation decreasing.

To reach a better equations form, we consider lift and drag, gravitational field, and the atmosphere as below. We also introduce some dimensionless quantities:

$$\begin{aligned}
 D &= 0.5\rho SC_D V^2, & L &= 0.5\rho SC_L V^2, & g &= \frac{g_s r_s}{r^2} = \frac{\mu}{r^2}, \\
 \rho &= \rho_s \exp[-\beta(r - r_s)], & \varepsilon &= \frac{1}{\beta r_s}, & v &= \frac{V^2}{r_s g_s}, \\
 \lambda &= \frac{C_L}{C_D}, & B &= \frac{\rho_s SC_D}{2m\beta}.
 \end{aligned} \tag{3}$$

Here, ρ_s is the atmospheric density at the reference height, S is a reference area from the geometry of the reentry vehicle, and β is the reciprocal of the scale height. For the earth atmosphere, the mean βr_s is a large number, almost 900, and hence the parameter ε is almost 0.0011, the famous small or perturbation parameter that let us apply the singular perturbation approach. The difference in the time scale of the observed states in reentry equations is the physical result of the “mathematical presence” of this small value in the equations.

λ is the first control and the aerodynamic roll angle σ the second; this is one of the main differences of the present problem in comparison with others defined previously [21]:

$$\begin{aligned}
 \frac{d\theta}{dh} &= \frac{\cot \gamma \cos \psi}{(1 + h) \cos \phi}, \\
 \frac{d\phi}{dh} &= \frac{\cot \gamma \sin \psi}{1 + h}, \\
 \frac{dv}{dh} &= -\frac{2Bv \exp[-h/\varepsilon]}{\varepsilon \sin \gamma} - \frac{2}{(1 + h)^2}, \\
 \frac{d\gamma}{dh} &= \frac{B\lambda \cos \sigma \exp[-h/\varepsilon]}{\varepsilon \sin \gamma} + \cot \gamma \left[\frac{1}{1 + h} - \frac{1}{v(1 + h)^2} \right], \\
 \frac{d\psi}{dh} &= \frac{B\lambda \sin \sigma \exp[-h/\varepsilon]}{\varepsilon \sin \gamma \cos \gamma} - \frac{\cos \psi \tan \phi \cot \gamma}{1 + h}.
 \end{aligned} \tag{4}$$

In addition, instead of ϕ , ψ , and θ , three new variables are introduced from celestial mechanics. They are I , Ω , and α , respectively, the inclination of the plane of the osculating orbit, the longitude of the ascending node, and the angle

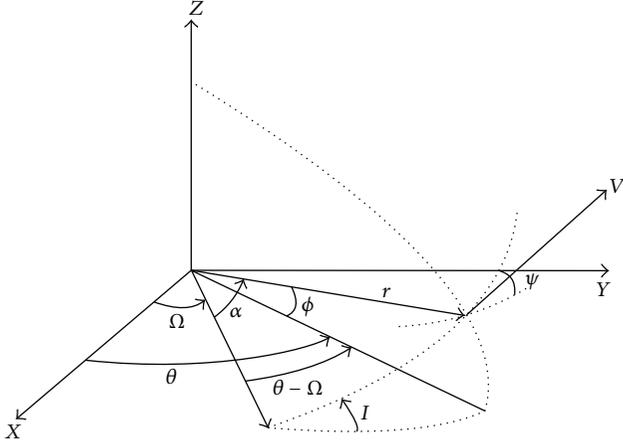


FIGURE 2: Orbital elements in Keplerian motion.

between the line of the ascending node and the position vector (Figure 2):

$$\begin{aligned}\cos \phi \cos \psi &= \cos I, \\ \sin(\theta - \Omega) &= \frac{\tan \phi}{\tan I}, \\ \cos \alpha &= \cos \phi \cos(\theta - \Omega).\end{aligned}\quad (5)$$

These lead to the final equations ready for MARE G&C application.

We have named fast term (FT) a term that is common between all the equations and has an interesting physical meaning that we will explain later in the text:

$$\begin{aligned}\text{FT} &= \frac{B\lambda \sin \sigma \exp[-h/\varepsilon]}{\varepsilon \sin \gamma \cos \gamma}, \\ \frac{d\alpha}{dh} &= -\frac{\sin \alpha}{\tan I} \{\text{FT}\} + \frac{\cot \gamma}{1+h}, \\ \frac{d\Omega}{dh} &= \frac{\sin \alpha}{\sin I} \{\text{FT}\}, \\ \frac{dI}{dh} &= \cos \alpha \{\text{FT}\}, \\ \frac{dv}{dh} &= -\frac{2v \cos \gamma}{\lambda \sin \sigma} \{\text{FT}\} - \frac{2}{(1+h)^2}, \\ \frac{d\gamma}{dh} &= \cos \gamma \cot \sigma \{\text{FT}\} + \cot \gamma \left[\frac{1}{1+h} - \frac{1}{v(1+h)^2} \right].\end{aligned}\quad (6)$$

4. Optimal Problem Statement

We state the optimal control as finding the desired trajectory for the motion between an initial known fix condition and a desired arrival point in the space. Precisely, we assume to know the initial conditions for (6) and the desired final conditions except for velocity that we will try to minimize by the performance measure. We also want to minimize the error in the final position and flight path angle considering

that the MARE G&C is devised for a manned reentry capsule. In manned reentry, precision is of vital importance, especially if the capsule is designed to land on the earth and not in the sea.

The performance measure is

$$\begin{aligned}J &= K_1 v(h_f) + K_2 [\alpha(h_f) - \alpha_f]^2 + K_3 [\Omega(h_f) - \Omega_f]^2 \\ &+ K_4 [I(h_f) - I_f]^2 + K_5 [\gamma(h_f) - \gamma_f]^2\end{aligned}\quad (7)$$

in which $K_{i=1,\dots,5}$ are nonnegative coefficients that lead to regulate the importance of each term in the trajectory generation; h_f stands for the final height (our independent variable is height), $i(h_f)$ are the actual final values of the states, and i_f the desired final values.

Obviously, the method can be applied for performance measures that are more general and the solution strategy; it is clearly independent from the cost function selected. It must be noted that sometimes the criterion considered by aerospace engineers is to minimize the total heat flux or at least to minimize the wear, the deterioration of the engine. However, minimizing the final velocity is important especially in manned reentry missions; the physical motivation of this decision is to “soften” the reentry considering the presence of crew and considering the key role that velocity has in some other constraints such as heating and dynamic pressure.

The Hamiltonian is as follows:

$$\begin{aligned}H &= p^T a(x, u, h), \\ H &= p_1 \left[-\frac{\sin \alpha}{\tan I} \{\text{FT}\} + \frac{\cot \gamma}{1+h} \right] \\ &+ p_2 \left[\frac{\sin \alpha}{\sin I} \{\text{FT}\} \right] + p_3 [\cos \alpha \{\text{FT}\}] \\ &+ p_4 \left[-\frac{2v \cos \gamma}{\lambda \sin \sigma} \{\text{FT}\} - \frac{2}{(1+h)^2} \right] \\ &+ p_5 \left[\cos \gamma \cot \sigma \{\text{FT}\} + \cot \gamma \left[\frac{1}{1+h} - \frac{1}{v(1+h)^2} \right] \right].\end{aligned}\quad (8)$$

The Hamiltonian exhibits the two-time-scale identity, and this is made clearly visible by the insertion of “FT” term. We finally need the costates to end the optimal control problem statement

$$\begin{aligned}\lambda_i^T &= -\frac{\partial H}{\partial x_i}, \\ \dot{p}_1 &= -p_1 \left[\frac{-\cos \alpha}{\tan I} \{\text{FT}\} \right] - p_2 \left[\frac{\cos \alpha}{\sin I} \{\text{FT}\} \right] \\ &- p_3 [-\sin \alpha \{\text{FT}\}], \\ \dot{p}_2 &= 0,\end{aligned}$$

$$\begin{aligned}
\dot{p}_3 &= -p_1 \left[\sin \alpha \frac{1 + \tan^2 I}{\tan^2 I} \{FT\} \right] \\
&\quad - p_2 \left[-\frac{\cos I}{\sin^2 I} \sin \alpha \{FT\} \right], \\
\dot{p}_4 &= -p_4 \left[-\frac{2 \cos \gamma}{\lambda \sin \sigma} \{FT\} \right] - p_5 \left[\frac{\cot \gamma}{v^2(1+h)^2} \right], \\
\dot{p}_5 &= -p_1 \left[-\frac{\sin \alpha}{\tan I} 2 \cos 2\gamma \{FT\} - \frac{(1 + \cot^2 \gamma)}{1+h} \right], \\
&\quad - p_2 \left[\frac{\sin \alpha}{\sin I} 2 \cos 2\gamma \{FT\} \right] - p_3 \left[(\cos \alpha) 2 \cos 2\gamma \{FT\} \right] \\
&\quad - p_4 \left[\frac{2v \cos^2 \gamma}{\lambda \sin \gamma \sin \sigma} \{FT\} \right] \\
&\quad - p_5 \left[-\frac{\cos^2 \gamma \cot \sigma}{\sin \gamma} \{FT\} - (1 + \cot^2 \gamma) \right. \\
&\quad \quad \left. \times \left[\frac{1}{1+h} - \frac{1}{v(1+h)^2} \right] \right].
\end{aligned} \tag{9}$$

Considering the problem statement, we have five initial conditions known and we will have five final values for the co-states that are obtained using the so-called transversality condition.

5. Outline of the Method

As suggested in [6], instead of solving the previous difficult problem, we use the SPT theory to break the solution into two phases, the outer (slow) and the inner (fast) solutions; so the optimal control problem is solved in two phases and finally matched according to matched asymptotic expansion rules. In other words, we must subtract from the addition of the two solutions a common part.

At first, with the degeneration (expanding based on ε and considering the zero-order terms), the outer or Keplerian region equations are obtained; they can be integrated analytically.

Then, the equations related to the inner or aerodynamic-predominated region (obtained after introducing the new stretched independent variable, expanding for ε and considering the zero-order terms) are also integrated analytically except for three co-state equations. The numerical integration of them is easy and straightforward because their equations are simpler than the original ones due to the effect of time-scale separation. Finally, the initial conditions for the state equations are obtained with the aid of the matching condition.

The matching condition [6] says that the limit of the inner solution for $h \rightarrow \infty$ must equal the value of the limit of the outer solution for $h \rightarrow 0$. The key assumption is that in the initial point, say h_i , the influence of the inner solution on the states is negligible. In other word the assumption is that in the Keplerian region, the influence of the inner or

Aerodynamic originated part of the equations is negligible. In this way, by equating the initial known conditions with the equations obtained from the solution of the outer equations, the outer solution can be found. Then, with the aid of the matching condition stated previously, the initial value for the integration of the inner equations can be obtained. In other words, these initial conditions are almost equal to the value of the inner equations for $h \rightarrow \infty$ and so are equal to the limit of the outer (known) equations for $h \rightarrow 0$.

The physical insight gained here is of great importance. We have identified in the reentry optimal equations the terms related to the fast time scale. Considering separately the solution of the fast and the slow equation terms, we achieved a quasi-analytic solution, ideal for a guidance scheme.

However, it is important to understand that this separation can help any reentry solution, based on any technique; in other words, in this way, we are a step nearer to the solution only by the help of the problem physics and related mathematics.

6. Outer (Keplerian) Solution

At first, the states are expanded for ε

$$\begin{aligned}
\alpha &= \alpha_o(h) + \varepsilon \alpha_1(h) + \dots, \\
\Omega &= \Omega_o(h) + \varepsilon \Omega_1(h) + \dots, \\
I &= I_o(h) + \varepsilon I_1(h) + \dots, \\
v &= v_o(h) + \varepsilon v_1(h) + \dots, \\
\gamma &= \gamma_o(h) + \varepsilon \gamma_1(h) + \dots, \\
p_1 &= p_{1o}(h) + \varepsilon p_{11}(h) + \dots, \\
p_2 &= p_{2o}(h) + \varepsilon p_{21}(h) + \dots, \\
p_{3o} &= p_{3o}(h) + \varepsilon p_{31}(h) + \dots, \\
p_{4o} &= p_{4o}(h) + \varepsilon p_{41}(h) + \dots, \\
p_{5o} &= p_{5o}(h) + \varepsilon p_{51}(h) + \dots.
\end{aligned} \tag{10}$$

This part of the solution describes the zone near to the vacuum where the gravitational force is predominant and the velocity is so high that aerodynamic controls are practically ineffective. Substituting the outer expansions (10) in (6) and (9) and equating the ε zero-order terms (we can alternatively let the small parameter ε tends to zero) we obtain (11)

$$\begin{aligned}
\frac{d\alpha_o}{dh} &= \frac{\cot \gamma_o}{1+h}, \\
\frac{d\Omega}{dh} &= 0, \\
\frac{dI}{dh} &= 0, \\
\frac{dv_o}{dh} &= -\frac{2}{(1+h)^2},
\end{aligned}$$

$$\begin{aligned}
\frac{d\gamma_o}{dh} &= \cot\gamma_o \left[\frac{1}{1+h} - \frac{1}{v(1+h)^2} \right], \\
\dot{p}_{1o} &= 0, \\
\dot{p}_{2o} &= 0, \\
\dot{p}_{3o} &= 0, \\
\dot{p}_{4o} &= -p_{5o} \left[\frac{\cot\gamma_o}{v_o^2(1+h)^2} \right], \\
\dot{p}_{5o} &= -p_{1o} \left[-\frac{(1+\cot^2\gamma_o)}{1+h} \right] \\
&\quad - p_{5o} \left[-(1+\cot^2\gamma_o) \left[\frac{1}{1+h} - \frac{1}{v_o(1+h)^2} \right] \right]. \tag{11}
\end{aligned}$$

The subscript “o” is showing that these are the outer equations. The fast terms are vanished (with the atmosphere effect in them) and we are going to solve the slow part of the original equations; in other words, the outer solution corresponds to a vanishing atmosphere.

Considering the Hamiltonian, we see that the control-including terms have vanished, and so there is no need to find the optimal controls. The result is quite logical because in the Keplerian region (for which the outer equations are written) the aerodynamic controls considered in the problem have negligible effect on the motion. Equation (11) can be solved analytically:

$$\begin{aligned}
v_o &= 2C_1 + \frac{2}{1+h}, \\
\cos\gamma_o = q_o &= \frac{C_2}{(1+h)\sqrt{v_o}}, \\
\cos(\alpha_o - C_3) &= \frac{C_2^2/(1+h)^{-1}}{\sqrt{1+2C_1C_2^2}}, \\
I_o &= C_4, \\
\Omega_o &= C_5, \\
p_{1o} &= C_6, \\
p_{2o} &= C_7, \\
p_{3o} &= C_8, \\
p_{4o} &= A(C_1, C_2, C_6, C_9, C_{10}, h), \\
\dot{p}_{5o} &= C_{10} \left\{ [1+h][2+2C_1(1+h)] - C_2^2 \right\}^{1/2} \\
&\quad - \frac{C_6(2+2C_1C_2^2) + 2C_1C_6(1+h)}{1+2C_1C_2^2}. \tag{12}
\end{aligned}$$

In (12), the p_{4o} is slightly long and the expression for is included in the appendix.

7. Inner (Aerodynamic Predominated) Solution

The inner expansions are related to the so-called boundary layer [6] of the solution that in this problem has the physical meaning of the region near to the earth surface.

A new variable, the so-called “stretched” [6] independent variable, is introduced in (6) and (9)

$$\bar{h} = \frac{h}{\varepsilon}. \tag{13}$$

Using another time the expansion method (like (10)), we obtain the zeroth-order inner approximation

$$\begin{aligned}
FT' &= \frac{B\lambda \sin\sigma \exp[-\bar{h}]}{\sin\bar{\gamma}_o \cos\bar{\gamma}_o}, \\
\frac{d\bar{\alpha}_o}{d\bar{h}} &= -\frac{\sin\bar{\alpha}_o}{\tan\bar{I}_o} \{FT'\}, \\
\frac{d\bar{\Omega}_o}{d\bar{h}} &= \frac{\sin\bar{\alpha}_o}{\sin\bar{I}_o} \{FT'\}, \\
\frac{d\bar{I}_o}{d\bar{h}} &= \cos\bar{\alpha}_o \{FT'\}, \\
\frac{d\bar{v}_o}{d\bar{h}} &= -\frac{2\bar{v}_o \cos\bar{\gamma}_o}{\lambda \sin\sigma} \{FT'\}, \\
\frac{d\bar{\gamma}_o}{d\bar{h}} &= \cos\bar{\gamma}_o \cot\sigma \{FT'\}, \\
\dot{\bar{p}}_{1o} &= -\bar{p}_{1o} \left[\frac{-\cos\bar{\alpha}_o}{\tan\bar{I}_o} \{FT'\} \right] \\
&\quad - \bar{p}_{2o} \left[\frac{\cos\bar{\alpha}_o}{\sin\bar{I}_o} \{FT'\} \right] - \bar{p}_{3o} [-\sin\bar{\alpha}_o \{FT'\}], \\
\dot{\bar{p}}_{2o} &= 0, \\
\dot{\bar{p}}_{3o} &= -\bar{p}_{1o} \left[\sin\bar{\alpha}_o \frac{1+\tan^2\bar{I}_o}{\tan^2\bar{I}_o} \{FT'\} \right] \\
&\quad - \bar{p}_{2o} \left[-\frac{\cos\bar{I}_o}{\sin^2\bar{I}_o} \sin\bar{\alpha}_o \{FT'\} \right], \\
\dot{\bar{p}}_{4o} &= -\bar{p}_{4o} \left[-\frac{2\cos\bar{\gamma}_o}{\lambda \sin\sigma} \{FT'\} \right], \\
\dot{\bar{p}}_{5o} &= -\bar{p}_{1o} \left[-\frac{\sin\bar{\alpha}_o}{\tan\bar{I}_o} 2\cos 2\bar{\gamma}_o \{FT'\} \right] \\
&\quad - \bar{p}_{2o} \left[\frac{\sin\bar{\alpha}_o}{\sin\bar{I}_o} 2\cos 2\bar{\gamma}_o \{FT'\} \right] \\
&\quad - \bar{p}_{3o} [(\cos\bar{\alpha}_o) 2\cos 2\bar{\gamma}_o \{FT'\}].
\end{aligned}$$

$$\begin{aligned}
& -\bar{p}_{4o} \left[\frac{2\nu \cos^2 \bar{\gamma}_o}{\lambda \sin \bar{\gamma}_o \sin \sigma} \{FT'\} \right] \\
& -\bar{p}_{5o} \left[-\frac{\cos^2 \bar{\gamma}_o \cot \sigma}{\sin \bar{\gamma}_o} \{FT'\} \right].
\end{aligned} \tag{14}$$

The superscripts “-” are showing that these are the inner equations. Note that the controls, assumed to be originated from the out of the system, are not “stretched.” Equation (14) is solved as follows:

$$\begin{aligned}
\bar{v}_o &= \bar{C}_1 \exp \left[\frac{-2\bar{\gamma}_o}{\lambda \cos \sigma} \right], \\
\bar{q}_o &= \cos \bar{\gamma}_o = B\lambda \cos \sigma \exp \left[-\bar{h} \right] + \bar{C}_2, \\
\sin(\bar{\alpha}_o) \sin(\bar{I}_o) &= \sin \bar{C}_3, \\
\cos \bar{\alpha}_o &= \cos \bar{C}_3 \cos(\bar{C}_4 - \bar{\Omega}_o), \\
\cos \bar{I}_o &= \cos \bar{C}_3 \cos \left\{ \tan \sigma \log \left[\tan \left(\frac{\bar{\gamma}_o}{2} + \frac{\pi}{4} \right) \right] + \bar{C}_5 \right\}, \\
\bar{p}_{2o} &= \bar{C}_7, \\
\bar{p}_{4o} &= \frac{\bar{C}_9}{\bar{v}}.
\end{aligned} \tag{15}$$

For the states, the sequence of the solutions is $\bar{\gamma}_o$ at first and then \bar{v}_o , \bar{I}_o , $\bar{\alpha}_o$, and finally $\bar{\Omega}_o$. For \bar{p}_{1o} , \bar{p}_{3o} , and \bar{p}_{5o} , we have not found yet an analytical solution (new methods are under study), and we choose numerical solution. It is important to note that the initial difficult problem requiring reiterative numerical solution is now reduced to an analytical part and a small numerical straightforward solution, which must be added to it.

8. Matched or Final Solution

Matching principle [6] says that the inner and the outer solutions must have an interface region. As stated in the outline of the method, we can find this common solution in two ways. The common solution is the inner limit of the outer expansion and the outer limit of the inner expansion. As shown in the previous work [21], the matching principle provides a relation between C_i and \bar{C}_i ; moreover, it helps us to find the final solution.

Considering that we have the initial states conditions, we force the outer solution to satisfy them at h_i . In this way, we will find C_i for $i = 1, 5$. Then, using the matching principle (inner limit of the outer solution), we find \bar{C}_i for $i = 1, 5$. For the co-states, the method is the same but from the other side. In fact, we force the inner co-state equations to satisfy the final value that we have for the co-states obtaining \bar{C}_i from $i = 6, 10$ (we integrate inversely the three equations that are not solved analytically), and then using the matching principle, we find C_i for $i = 6, 10$.

9. Optimal Controls

The optimal controls are obtained according to the Pontryagin's minimum principle

$$\bar{H}(\lambda^*, \sigma^*) \leq \bar{H}(\lambda, \sigma). \tag{16}$$

At this step, it is important to remember that both controls are constrained. For aerodynamic roll,

$$\begin{aligned}
A &= \tan^{-1} \frac{-\bar{p}_{5o} \cos \bar{\gamma}_o}{[-\bar{p}_{1o}(\sin \bar{\alpha}_o / \tan \bar{I}_o) + \bar{p}_{2o}(\sin \bar{\alpha}_o / \sin \bar{I}_o) + \bar{p}_{3o} \cos \bar{\alpha}_o]} \\
\sigma^* &= \begin{cases} \sigma_{\text{MIN}} & \text{if } A < \sigma_{\text{MIN}}, \\ A & \text{if } \sigma_{\text{MIN}} < A < \sigma_{\text{MAX}}, \\ \sigma_{\text{MAX}} & \text{if } A > \sigma_{\text{MAX}}, \end{cases} \\
\lambda^* &= \begin{cases} \lambda_{\text{MIN}} & \text{if } \bar{p}_{5o} < 1, \\ \lambda_{\text{MAX}} & \text{if } \bar{p}_{5o} > 1, \\ \text{Sing. Int.} & \text{if } \bar{p}_{5o} = 0. \end{cases}
\end{aligned} \tag{17}$$

The different form of the optimal controls derives from the different dependence of $\bar{H}(\lambda^*, \sigma^*)$ from the controls; aerodynamic roll is present in $\bar{H}(\lambda^*, \sigma^*)$ in trigonometric form, but the dependence of $\bar{H}(\lambda^*, \sigma^*)$ from λ is only linear. The discussion about the singular interval that might occur is left as a possible future work to improve the method.

10. Path Constraints

The path constraints are a serious matter that must be addressed for a reentry; however, the major part of the reentry-related papers in the literature do not discuss them.

Heating rate is the first driving factor to be considered [26]:

$$\sqrt{\frac{\rho}{\rho_{\text{ref}}}} \left[\frac{\sqrt{R_s g_s}}{v_{\text{ref}}} \right]^3 \left[\frac{V}{\sqrt{R_s g_s}} \right]^3 \leq \frac{\dot{q}_{\text{MAX}}}{C_q}, \tag{18}$$

where $\rho_{\text{ref}} = 1 \text{ kg/m}^3$, $v_{\text{ref}} = 1 \text{ m/s}$, \dot{q}_{MAX} is the maximum stagnation point heating rate (W/m^2), and C_q is the heat flux transmission coefficient; the last two parameters can be found for a specific reentry vehicle. This equation can be introduced in the drag equation and it takes the form of (19):

$$\frac{D}{m} = \frac{1}{2m} C_D S \rho_{\text{ref}}^6 v_{\text{ref}}^6 \frac{\dot{q}_{\text{MAX}}^2}{C_q^2 V^4}. \tag{19}$$

Having the optimal control λ and the vehicle data [25], we can have the drag coefficient and so plot the drag force versus the velocity. A typical graph will be like Figure 3.

Maximum deceleration or alternatively speaking the maximum normal load is the maximum load constraint in the body-normal direction [26].

Considering that both lift and drag forces contribute to the normal direction (z direction in Figure 4), the constraint is defined as follows:

$$\bar{N}_Z = \frac{L}{m} \cos \alpha + \frac{D}{m} \sin \alpha = \frac{C_L}{C_D} \frac{D}{m} \cos \alpha + \frac{D}{m} \sin \alpha. \tag{20}$$

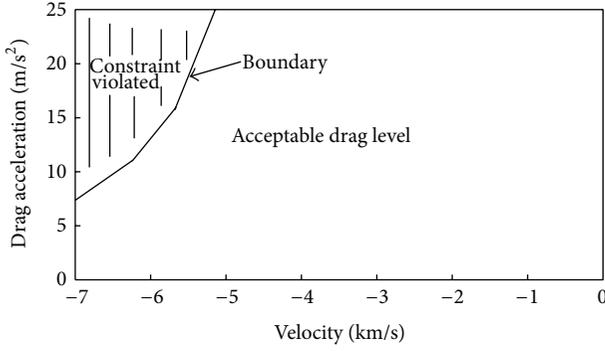


FIGURE 3: Heat rate constraint shown in drag-velocity plane.

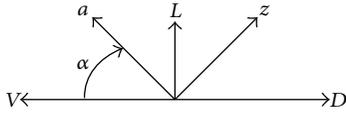


FIGURE 4: Different directions for reentry vehicle.

It is typical to rearrange (20) as follows, considering a maximum sustained value for N_Z that for a manned mission must be of the order of 2.5 g [2]:

$$\frac{D}{m} \leq \frac{\bar{N}_{Z,MAX}}{(C_L/C_D) \cos \alpha + \sin \alpha}. \quad (21)$$

Plotting (21), we obtain something like Figure 5.

The dynamic pressure constraint can be directly seen from drag equation as follows and leads to Figure 6:

$$\frac{D}{m} \leq q_{MAX} \frac{C_D S}{m}. \quad (22)$$

In addition, an equilibrium glide constraint is normally considered that is related to the minimum drag case motion (23). It leads to a graph similar to Figure 7:

$$\frac{D}{m} \geq \frac{C_D}{C_L} \left(-\frac{V^2}{R_s + h} + g \right). \quad (23)$$

Then, heat load constraint must be addressed and is practically the whole heat rate absorbed by the vehicle during reentry. Considering the four constraints discussed previously the final corridor for reentry shown in drag-velocity plane will be like Figure 8. Figure 9 shows the new reentry corridor with a safety margin; this margin is considered to avoid dangerous situations, in other words, to avoid extreme nearing to the boundaries.

We must note that in a reentry not all the constraints are necessarily critical for the missions; in other words, there is no risk of violating all of them, but usually only some of them depend on the shape of the reentry vehicle (essentially that it is slender or blunt).

The algorithm devised for nonviolating the constraints is based on the “inverse problem” methodology, and it will be discussed in the next section.

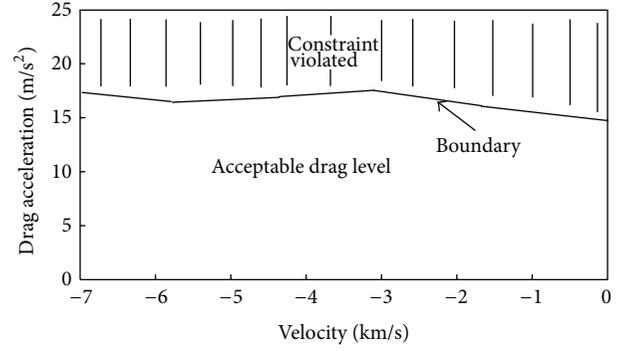


FIGURE 5: Normal load constraint in drag-velocity plane.

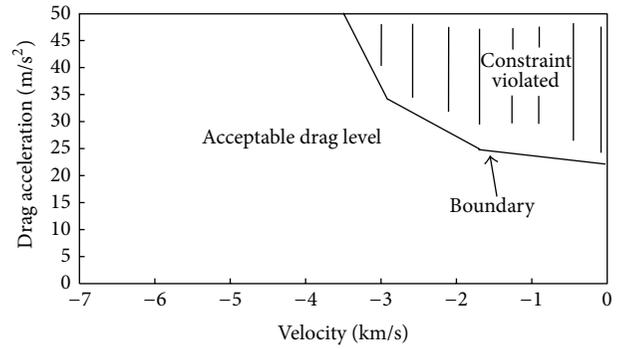


FIGURE 6: Dynamic pressure constraint in drag-velocity plane.

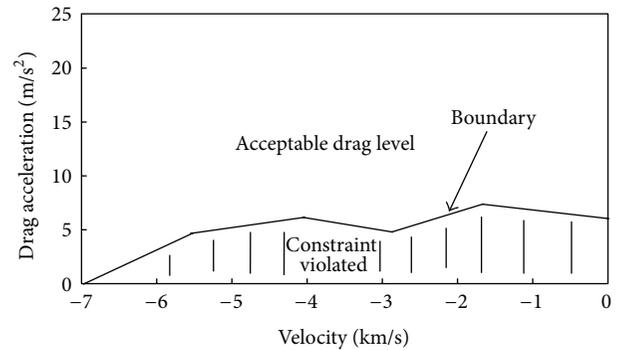
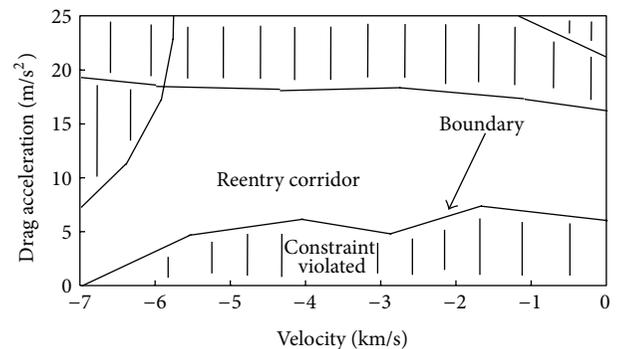


FIGURE 7: Equilibrium glide constraint in drag-velocity plane.

FIGURE 8: D - V plane reentry trajectory corridor.

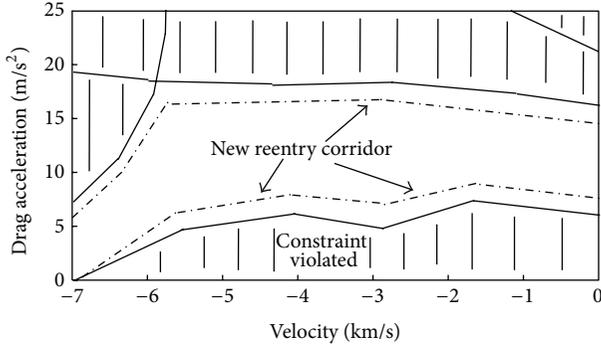


FIGURE 9: New reentry trajectory corridor defined for major safety.

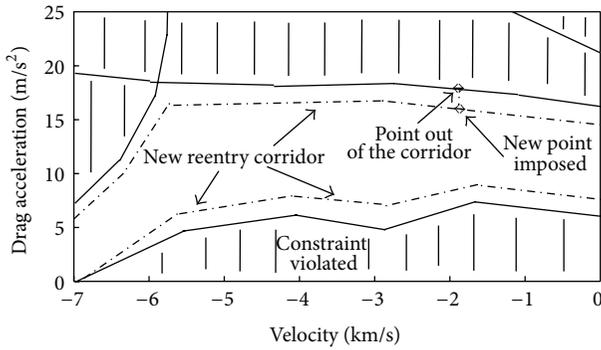


FIGURE 10: Imposing a new point to the problem based on constraint boundary.

11. Path Constraints Problem Solution via “Inverse Problem” Methodology

The inverse problem methodology is a general approach to solving problems in engineering, which can be loosely defined as the art of finding a problem solution via the “observation” of some “desired” or “typical” answers of the problem.

In other words, we get inspiration from the answers of the problem that we have acquired in some way before seeing their equations or something else. In the first part of the present work, although our solution method has the general name of SPT, we must stress that it has used the inverse problem mentality. As stated at the beginning, “a glance to the problem physics” suggests that the problem must have a two-time scale nature.

Moreover, in the optimal control arena, “inverse problem” means having the optimal controls and finding the desired trajectory. This has to be done by some researchers by defining a new cost function that minimizes the error of the found trajectory with respect to the desired one. In the guidance field, it consists in predicting the controls that are compatible with a desired trajectory [17]. A survey about the inverse problem approach in optimal trajectory generation for guidance, both in Russia and in the United States, has been presented by Yakimenko [18].

On the other hand, in the statement of the optimal reentry problem, considering directly the path constraints,

so transforming them in additional equations, increases the complexity of the problem and makes it computationally inefficient. The inverse problem methodology has been used in this part of the problem in the sense that we obtain the optimal trajectories without considering the constraints. Then, if they respect the constraints, they are admissible, and so they are used. Alternatively, we consider the “emergency mode path” derived from the reentry trajectory corridor, and starting from the path given by it, we find the controls, precisely based on the inverse problem methodology [5].

We have considered a so-called “emergency mode” to avoid constraints violation. The first step is to define an entry trajectory corridor inside the previous entry corridor presented for major safety of the mission.

In the guidance scheme of the problem that predicts fast and high speed solution of the reentry problem during the reentry vehicle flight (this is made possible by the analytical and simple form of the solution), we can find in each computation cycle the optimal controls to be applied to the vehicle.

Applying these controls to the equations, we can “predict” the future states of the system and also the velocity and the drag force. If they are in the defined reentry trajectory corridor, as shown previously, the meaning is that the application of the optimal controls obtained will not violate the path constraints, and therefore they are considered as the control input.

Alternatively, if we predict that the found optimal controls will result in drag and velocity out of the corridor (so some constraints will be violated), we consider a point on the entry trajectory boundary, and so we have an “emergency mode” drag and velocity. Figure 10 is showing this situation.

The known values of drag and force are in other words imposing new “emergency mode” controls, which are computed as follows.

We consider

$$D = K_1, \quad V = K_2. \quad (24)$$

Considering the velocity solution obtained according to the inner, outer, and common solutions

$$K_2 = V = 2C_1 + \frac{2}{1+h} + \bar{C}_1 \exp \left[\frac{-2\bar{y}_o}{\lambda \cos \sigma} \right] - V_c, \quad (25)$$

C_1 and \bar{C}_1 are known values. h , the scale height, is also known for a particular point of the reentry path. For \bar{y}_o , we substitute the value of the previous step. Therefore, in (25), we can easily find the value $\lambda \cos \sigma$. So, we have

$$\lambda \cos \sigma = \frac{C_L}{C_D} \cos \sigma = K_3. \quad (26)$$

On the other hand, we know the drag force value, so

$$D = \frac{1}{2} \rho V^2 S C_D = K_1. \quad (27)$$

According to (27), we can also find C_D because the other values are all known. In the different reentry problem statements, we usually have drag and lift coefficients as a function

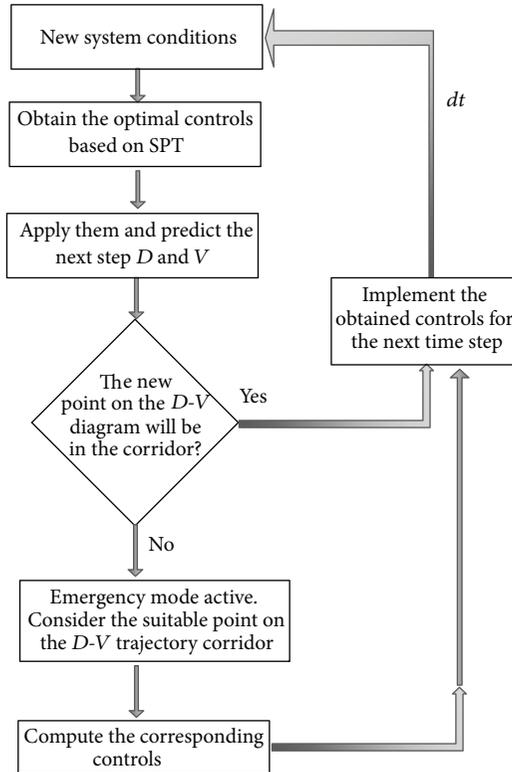


FIGURE 11: Path constraints implementation flowchart based on inverse problem.

of AoA and Mach in the form of tables or we have close form equations for them. In both cases, having C_D , we can find C_L and thus via (25) the value of σ .

The emergency mode controls (λ and σ) are considered as the input control of the reentry vehicle for one time step, and then in the new observed condition (at the beginning of the next time step), we start again the predictive process.

We compute optimal controls, we check for D - V point that they will create, and we restart the explained algorithm.

The method is better shown in Figure 11.

12. Implementation in a Guidance Scheme (MARE G&C)

Singular perturbation methods were previously used to generate online the optimal trajectories for aircrafts [27]. Earlier works suggest a similar approach for the reentry case. Traditionally, a guidance scheme that tries to make the error of the flight path zero with respect to a preknown path has been considered and used for the past studies and missions. This is the famous path-following approach. The method developed in the present work solves almost analytically the optimal reentry problem with two controls minimizing a cost function considered for a manned reentry mission. The capacity of generating fast solutions recommends the method as a suitable predictive guidance scheme; in this scheme, new paths are generated in each time step after each observation, and the reentry vehicle tries to follow the “new” real optimal trajectory generated at the previous time step. The method

called MARE G&C can compute online the “new” optimal trajectory at each new point of the trajectory that can be out of the initial optimal path. In other words, the reentry vehicle, during its motion, receives new control inputs for flying on the new optimal path, obtained with the data transmitted to the onboard computer at the last sampling time. This property improves intrinsically the robustness and accuracy of the method with respect to the path-following methods as it will be explained in the next section. Figure 12 shows the general flight guidance scheme considered in the present work.

13. Accuracy, Time Efficiency, and Robustness

The sources that can affect the precision of the results are the implicit error of the solution method and the other factors like observation errors and parameter variations, and in other words structured and unstructured uncertainties.

We left observation errors as an issue related to the reentry vehicle mission, but we discuss on the intrinsic method error, the effect of the variation of aerodynamic coefficients, and disturbances created by wind.

In the singular perturbation section of the solution (see Figure 12) due to the degeneration made for deriving inner and outer solutions, we have an $O(\varepsilon)$ error [6]. The solution is less precise with respect to those obtained in path-following methods, but it is very fast and its power stays in this feature. The effect of the method error is mitigated and cannot build up during the mission because the optimal path is found at each time step and the reentry vehicle maneuvers according to this new optimal path. One must consider that this feature can be a powerful instrument also against the aerodynamic coefficient variation (the variation of lift coefficient is the more important issue) and wind gusts. In both cases, the computations in the onboard computer are made based on parameters and states that are not the real ones, but the fast repeating of the path generation prevents the vehicle from going far from an admissible reentry trajectory, and this can even mean saving the whole mission.

In addition, to ensure sufficient robustness, the guidance scheme derived can be “analyzed” in the robust control sense [28]. In other words, disturbances and uncertainties can be considered, and then one can determine if the controller designed exhibits the proper characteristics by an LFT method. Considering the solution obtained and rearranging the system in the general format below (Figure 13) are sufficient to know an upper bound for the magnitude of the uncertainties to model them and account for their effect.

This may affect the solution time, and in this case, we are losing the speed, but we can ensure in an alternative way more robustness.

Some previous simulations [29] have shown in some extent that even with a simpler method but without considering the path constraints the general approach based on singular perturbation method can sufficiently mitigate the cited errors and has a precision comparable with the steepest descent or multiple shooting. The typical time efficiency for the code written is less than 10 ms with a normal PC with 2.4 GHz speed.

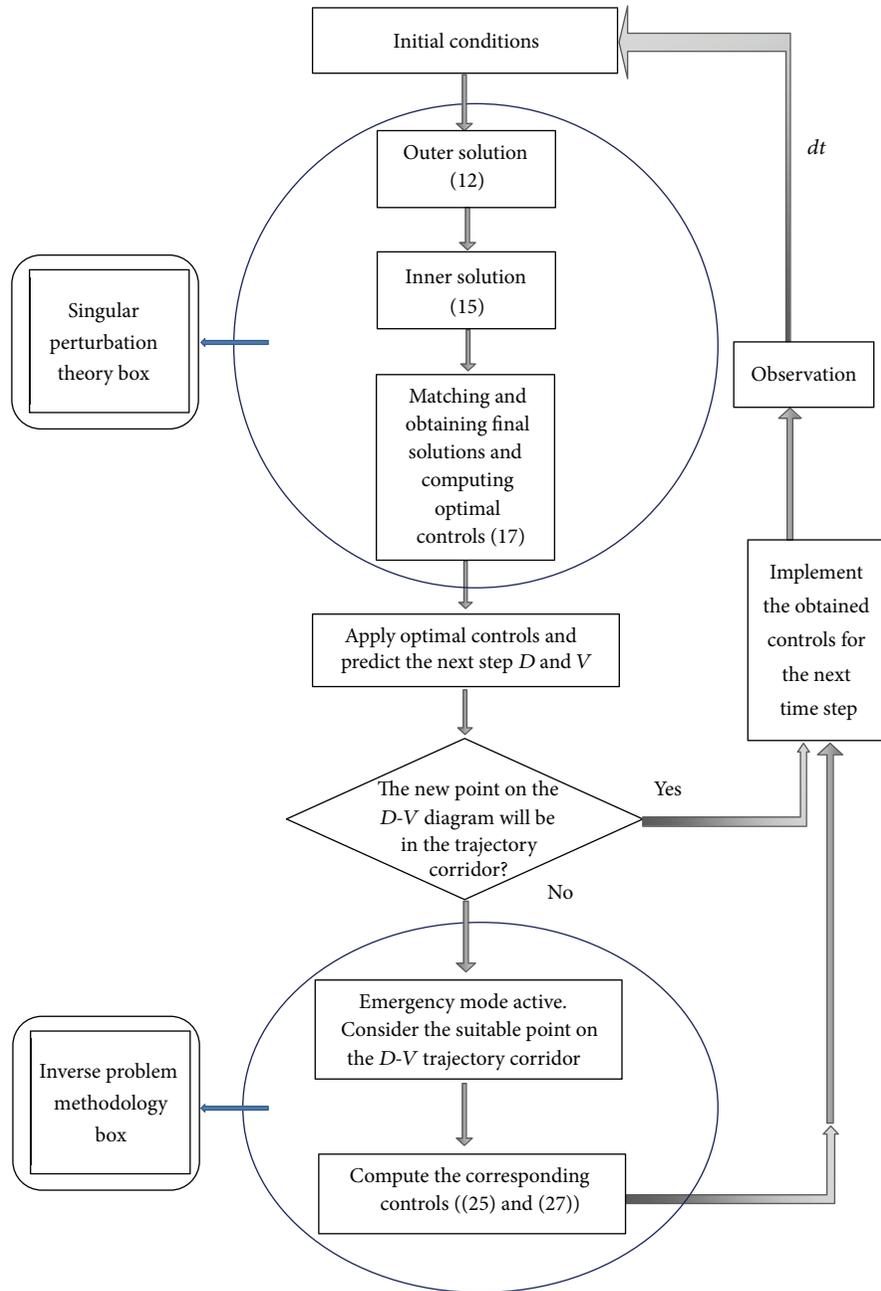


FIGURE 12: MARE G&C guidance scheme.

14. Concluding Remarks

We have addressed the solution of a 3-DOF optimal reentry problem with two controls (aerodynamic roll and lift-to-drag-ratio). Inverse mentality approach leads us to use the different time scales in the state of the dynamics of the problem. Following this suggestion from the physics, we have reached a new quasi-analytical solution based on the matched asymptotic expansion method that includes outer, inner, and common solutions, having reduced considerably the computational cost and efforts. Moreover, we tackled

the path constraints limit directly using an inverse problem methodology. Finally, we suggested a guidance concept based on this method, MARE G&C. Reaching such a form of PG means an improvement compared to the path-following traditional methods. With the online trajectory generation, all missions can be saved even with failures and out of design conditions; the vehicle can follow the real optimal path considering its actual position and is not obliged to follow the initial and ideal path that is not optimal in the remainder of the motion. Moreover, the continuous propagation of trajectories reduces the error and does not permit the errors

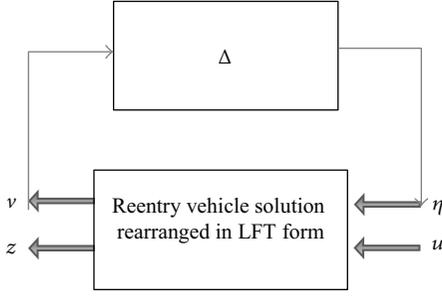


FIGURE 13: General LFT framework to model uncertainties.

to grow. This characteristic of MARE G&C reduces also the undesired effect of parametric uncertainty. The result is more precise in the arrival and more reliable, in general, for the reentry mission.

Appendix

The equation for p_{4o} is as follows:

$$\begin{aligned}
 p_{4o} = & C_9 - \frac{C_6 \ln(C_1 + C_1 h + 1)}{4C_1^2 C_2^2 + 2C_1 C_2} \\
 & + \left\{ \ln \left[4C_1 + 4(2C_1 C_2^2 + 1)^{3/2} + 4C_1 h \right. \right. \\
 & \quad + 19C_1 C_2^2 + 17C_1^2 C_2^2 + 28C_1^2 C_2^4 \\
 & \quad + 20C_1^3 C_2^4 + 12C_1^3 C_2^6 + 4C_1^4 C_2^6 \\
 & \quad + 4C_1 (2C_1 C_2^2 + 1)^{3/2} + 17C_1^2 C_2^2 h \\
 & \quad + 20C_1^3 C_2^4 h + 4C_1^4 C_2^6 h + 4C_1 h (2C_1 C_2^2 + 1)^{3/2} \\
 & \quad + 7C_1 C_2^2 (2C_1 C_2^2 + 1)^{3/2} + 5C_1^2 C_2^2 (2C_1 C_2^2 + 1)^{3/2} \\
 & \quad \left. + C_1^2 C_2^4 (2C_1 C_2^2 + 1)^{3/2} \right. \\
 & \quad \left. + 5C_1^2 C_2^2 h (2C_1 C_2^2 + 1)^{3/2} + 4 \right] \\
 & \quad \times \left[\frac{C_6}{4} + \frac{C_6 (2C_1 C_2^2 + 1)^{3/2}}{4} + \frac{C_1 C_2^2 C_6}{2} \right] \left. \right\} \\
 & \times (4C_1^3 C_2^5 + 4C_1^2 C_2^3 + C_1 C_2)^{-1}
 \end{aligned}$$

$$\begin{aligned}
 & + \left\{ \ln \left[4C_1 - 4(2C_1 C_2^2 + 1)^{3/2} + 4C_1 h \right. \right. \\
 & \quad + 19C_1 C_2^2 + 17C_1^2 C_2^2 + 28C_1^2 C_2^4 \\
 & \quad + 20C_1^3 C_2^4 + 12C_1^3 C_2^6 + 4C_1^4 C_2^6 \\
 & \quad - 4C_1 (2C_1 C_2^2 + 1)^{3/2} + 17C_1^2 C_2^2 h \\
 & \quad + 20C_1^3 C_2^4 h + 4C_1^4 C_2^6 h - 4C_1 h \\
 & \quad \times (2C_1 C_2^2 + 1)^{3/2} - 7C_1 C_2^2 (2C_1 C_2^2 + 1)^{3/2} \\
 & \quad - 5C_1^2 C_2^2 (2C_1 C_2^2 + 1)^{3/2} - C_1^2 C_2^4 \\
 & \quad \times (2C_1 C_2^2 + 1)^{3/2} - 5C_1^2 C_2^2 h \\
 & \quad \left. \left. \times (2C_1 C_2^2 + 1)^{3/2} + 4 \right] \right. \\
 & \quad \times \left[\frac{C_6}{4} - \frac{C_6 (2C_1 C_2^2 + 1)^{3/2}}{4} + \frac{C_1 C_2^2 C_6}{2} \right] \left. \right\} \\
 & \times (4C_1^3 C_2^5 + 4C_1^2 C_2^3 + C_1 C_2)^{-1} \\
 & - \frac{(C_2 C_{10} + 2C_1 C_2 C_6 + 2C_1 C_2^3 C_{10})}{(C_1 (4C_1 + 8C_1 C_2^2 + h (8C_1^2 C_2^2 + 4C_1) + 8C_1^2 C_2^2 + 4))}.
 \end{aligned} \tag{A.1}$$

Acronyms

AoA:	Angle of attack
ARD:	Atmospheric reentry demonstrator
ESA:	European Space Agency
EU:	European Union
G&C:	Guidance and control
MAE:	Matched asymptotic expansion
MARE G&C:	Matched asymptotic reentry guidance and control
PG:	Predictive guidance
SPT:	Singular perturbation theory
TPBVP:	Two-point boundary value problem.

Nomenclature

C_D :	Drag coefficient
C_L :	Lift coefficient
D :	Drag force
I :	Orbital inclination
L :	Lift force
S :	Reference area
V :	Velocity
g :	Earth gravity
h :	Dimensionless height above Earth surface
m :	Mass

r : Radial distance from the Earth center
 t : Time
 v : Dimensionless velocity
 y : Height above Earth surface
 α : Angle between ascending node line and position vector
 β : Scale height reciprocal
 ε : Degeneration parameter
 φ : Latitude angle
 γ : Flight path angle
 λ : Lift-to-drag ratio
 θ : Longitude angle
 σ : Aerodynamic roll angle
 ψ : Heading angle
 Ω : Longitude of the ascending node.

Acknowledgment

The authors wish to express their sincere appreciation to the anonymous reviewers that helped to improve this work with their valuable suggestions and their constructive criticism.

References

- [1] R. C. Wingrove, "Survey of atmosphere re-entry guidance and control methods," *AIAA Journal*, vol. 1, no. 9, pp. 2019–2029, 1963.
- [2] J. C. Harpold and C. A. Graves Jr., "Shuttle entry guidance," *The Journal of the Astronautical Sciences*, vol. 27, no. 3, pp. 239–268, 1979.
- [3] R. Monti and D. M. Paterna, "A low risk reentry: looking backward to step forward," *Aerospace Science and Technology*, vol. 10, no. 2, pp. 156–167, 2006.
- [4] M. Milam, *Real time optimal trajectory generation for constrained dynamical systems [Ph.D. thesis]*, California Institute of Technology, 2002.
- [5] V. T. Taranenko and V. G. Momdzhj, *Direct Method of Modification in Flight Dynamic Problems*, Mashinostroenie, Moscow, Russia, 1986.
- [6] V. P. Kokotovic, H. K. Khalil, and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis & Design*, Academic Press, 1986.
- [7] P. H. Tran, J. C. Pault, and P. Boukhobza, "Re-entry flight experiments lessons learned—the atmospheric reentry demonstrator ARD," NATO Report RTO-EN-AVT-130, 2007.
- [8] Z. Liwei and J. Wuxing, "Application of predictive guidance to reentry vehicle," in *Proceedings of the 1st International Symposium on Systems and Control in Aerospace and Astronautics*, pp. 659–663, January 2006.
- [9] J. M. Carson III, M. S. Epstein, D. G. Macmynowski, and R. M. Murray, "Optimal nonlinear guidance with inner-loop feedback for hypersonic re-entry," in *Proceedings of the American Control Conference*, pp. 5782–5787, June 2006.
- [10] J. T. Betts, *Practical Methods for Optimal Control and Estimation Using Nonlinear Programming*, vol. 19 of *Advances in Design and Control*, SIAM, Philadelphia, Pa, USA, 2nd edition, 2010.
- [11] T. Neckel, C. Talbot, and N. Petit, "Collocation and inversion for a reentry optimal control problem," *Proceedings of the 5th International Conference on Launcher Technology*, 2003.
- [12] M. H. Grablin, J. Telaar, and U. M. Schöttle, "Ascent and reentry guidance concept based on NLP-methods," *Acta Astronautica*, vol. 55, no. 3–9, pp. 461–471, 2004.
- [13] B. Bonnard, L. Faubourg, and E. Trelat, "Optimal control of the atmospheric arc of a space shuttle and numerical simulations with multiple-shooting method," *Mathematical Models and Methods in Applied Sciences*, vol. 15, no. 1, pp. 109–140, 2005.
- [14] E. Trélat, "Optimal control and applications to aerospace: some results and challenges," *Journal of Optimization Theory and Applications*, vol. 154, no. 3, pp. 713–758, 2012.
- [15] A. Verma, X. Peng, K. Vadakkeveedu, R. Mayer, M. W. Oppenheimer, and D. B. Doman, "Neural dynamic trajectory design for reentry vehicles," in *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, pp. 2164–2180, August 2007.
- [16] A. Naghash, R. Esmaelzadeh, M. Mortazavi, and R. Jamilnia, "Near optimal guidance law for descent to a point using inverse problem approach," *Aerospace Science and Technology*, vol. 12, no. 3, pp. 241–247, 2008.
- [17] M. Borri, C. L. Bottasso, and F. Montelaghi, "Numerical approach to inverse flight dynamics," *Journal of Guidance, Control, and Dynamics*, vol. 20, no. 4, pp. 742–747, 1997.
- [18] O. A. Yakimenko, "Direct method for rapid prototyping of near-optimal aircraft trajectories," *Journal of Guidance, Control and Dynamics*, vol. 23, pp. 865–875, 2000.
- [19] D. S. Naidu, *Singular Perturbation Methodology in Control Systems*, Stevenage Herts, England, UK, 1988.
- [20] Y. C. Shen, "Series solution of equations of reentry vehicles with variable lift and drag coefficients," *AIAA Journal*, vol. 1, no. 11, pp. 2487–2490, 1963.
- [21] Y. Y. Shi, "Matched asymptotic solution for optimum lift controlled atmospheric entry," *AIAA Journal*, vol. 9, no. 11, pp. 2229–2238, 1971.
- [22] N. X. Vinh, A. Busemann, and R. D. Culp, "Optimum three-dimensional atmospheric entry," *Acta Astronautica*, vol. 2, no. 7–8, pp. 593–611, 1975.
- [23] D. S. Naidu and A. J. Calise, "Singular perturbations and time scales in guidance and control of aerospace systems: a survey," *Journal of Guidance, Control, and Dynamics*, vol. 24, no. 6, pp. 1057–1078, 2001.
- [24] D. S. Naidu, "Analysis of non-dimensional forms of singular perturbation structures for hypersonic vehicles," *Acta Astronautica*, vol. 66, no. 3–4, pp. 577–586, 2010.
- [25] A. Viviani and G. Pezzella, "Computational flow field analysis over a blunt-body reentry vehicle," *Journal of Spacecraft and Rockets*, vol. 47, no. 2, pp. 258–270, 2010.
- [26] P. Lu, "Entry guidance and trajectory control for reusable launch vehicle," *Journal of Guidance, Control, and Dynamics*, vol. 20, pp. 143–149, 1997.
- [27] J. L. Vian and J. R. Moore, "Trajectory optimization with risk minimization for military aircraft," *Journal of Guidance, Control, and Dynamics*, vol. 12, no. 3, pp. 311–317, 1989.
- [28] K. Zhou and J. C. Doyle, *Essentials of Robust Control*, Prentice Hall, Upper Saddle River, NJ, USA, 1998.
- [29] M. Mortazavi and D. Abbasi, "Online optimal reentry guidance via matched asymptotic expansion," *Journal of Space Science and Technology*, vol. 3, no. 3, pp. 11–18, 2010.

Research Article

Reconstruction of Initial Wave Field of a Nonsteady-State Wave Propagation from Limited Measurements at a Specific Spatial Point Based on Stochastic Inversion

S. L. Han and Takeshi Kinoshita

Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Correspondence should be addressed to Takeshi Kinoshita; kinoshit@iis.u-tokyo.ac.jp

Received 1 August 2012; Revised 11 March 2013; Accepted 14 March 2013

Academic Editor: Fatih Yaman

Copyright © 2013 S. L. Han and T. Kinoshita. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper studies an inverse problem that can be used for reconstructing initial wave field of a nonsteady-state wave propagation. The inverse problem is ill posed in the sense that small changes in the input data can greatly affect the solution of the problem. To address the difficulty, the problem is formulated in the form of an inference problem in an appropriately constructed stochastic model. It is shown that the stochastic inverse model enables the initial surface disturbance to be reconstructed, including its confidence intervals given the noisy measurements. The reconstruction procedure is illustrated through applications to some simulated data for two- and three-dimensional problem.

1. Introduction

Water waves are one of important subjects in ocean-related fields. They are generated on the free surface, between air and water, from various sources such as a storm at sea, moving ship in calm water, and impact from a falling object. Once they are generated, wave propagation occurs on the surface of water away from the source. The propagation of water wave is a little more complex than many other kinds of waves in nature due to its dispersion.

In this paper, we are interested in a nonsteady-state water wave propagation generated by an initial surface disturbance. The classical problem in water waves is the computation of the time and spatial evolution in a whole body of water with this given initial wave field. That is, the aim of such problems is to predict the propagation of water waves as effects of the initial free surface disturbance, which is the cause of this phenomenon. This kind of problem has been widely studied because of its importance [1–6]. We name it forward or direct problem. Although much progress has been achieved in the analysis of resulting wave flow, there are only few studies [7–11] on inverse problems, whose aim is to find the cause of the wave propagation.

In this study, our attention is focused on the inverse problem of reconstructing initial wave field from the limited measurement on the specific boundary of the fluid body for two- and three-dimensional problems. Based on the assumption of linear dispersive wave theory, the reconstructing problem can be formulated as an inverse problem. The problem requires solving the Fredholm integral equation of the first kind, which causes severe ill posedness when numerically approximating. As a result, the inverse problem (reconstruction of initial wave field) is mathematically and numerically much more challenging than the forward problem (simulation of future wave field given initial disturbance). For such problems, the solution may not exist when noise is present in the measurements. Moreover, it may also lose stability although it exists, that is, lacks continuous dependence on the data. Thus, a small perturbation of the data may yield erroneous results of the inverse solution.

To overcome the difficulties, we adopt a stochastic inverse method, which has recently attracted much research attention in diverse fields [12–15]. Stochastic inverse methods solve the inverse problem in a systematic way by modeling all parameters of interests as random variables with joint probability distributions. This randomness can be considered as

parameter variability since it is related to the uncertainty of the true values. The key idea behind stochastic inverse method is to restate the original inverse problem in the form of probabilistic description of the unknown parameters, that is, stochastic inverse problem. Accordingly, the solution of the inverse problem is represented by a probability distribution of random unknowns. This probability distribution is known as the posterior probability distribution function incorporating the degree of confidence of unknown parameter given the measurement.

Two aspects of this study are worth noting. Firstly, the reconstruction procedure was derived from the limited measurement on the specific boundary of the fluid body. Secondly, a stochastic inverse method is introduced to enable the stable reconstruction procedure, since it requires solving the ill-posed problem. It also provides a way of quantifying an uncertainty in the solution, as well as smoothing the solution by expanding original solution spaces into probability spaces.

This study is organized as follows. Section 2 reviews two- and three-dimensional problems regarding a nonsteady-state water wave propagation due to initial surface disturbance. In Section 3, a procedure is described for the reconstruction of the initial wave field based on the limited measurements at a fixed spatial point. The stochastic inversion method is illustrated in Section 4 for the stable solution procedure. In Section 5, the reconstruction procedure is validated through applications to some digitally simulated data. Finally, we make concluding remarks in Section 6.

2. Water Waves due to an Initial Impulse

2.1. Mathematical Description of the Problem. It is assumed here that the motion of water waves is described by the Euler equations [1–6]. That is, the fluid is homogeneous, incompressible, and inviscid. The water depth is constant $z = -h$, and the surface tension is neglected. In this context, the corresponding fluid motion $\varphi(x, y, z, t)$ is governed by the Laplace equation

$$\nabla^2 \varphi = 0 \text{ in the water} \quad (1)$$

with the linearized boundary conditions

$$\begin{aligned} \frac{\partial \eta}{\partial t} &= \frac{\partial \varphi}{\partial z} \text{ on the surface,} \\ \frac{\partial \varphi}{\partial t} + g\eta &= 0 \text{ on the surface,} \\ \frac{\partial \varphi}{\partial z} &= 0 \text{ on the bottom,} \end{aligned} \quad (2)$$

where η, g denote the surface elevation, gravitational acceleration, respectively.

The Euler equations can mathematically be viewed as a Cauchy problem [6], studying a partial differential equation that satisfies certain conditions which are given on a hypersurface in the domain. Given the initial spatial surface disturbance at time zero $\eta(x, y, 0)$ or $\varphi(x, y, z, 0)$, the above equations give a way to determine the whole behavior of fluid with respect to time and space.

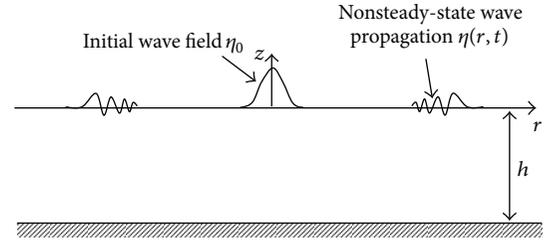


FIGURE 1: Wave propagation generated by the initial disturbance η_0 on the water surface.

2.2. Surface Waves due to Localized Impulse. If the initial wave disturbance η_0 , as shown in Figure 1, is generated by the localized sudden disturbances such as underwater explosions or sudden vertical seafloor displacement, the resulting nonsteady-state waves can be described by the axisymmetric form [1–6, 8, 9]

$$\eta(r, t) = \int_0^\infty a(k) k J_0(kr) \cos \omega t dk, \quad (3)$$

where $r = \sqrt{x^2 + y^2}$, $J_0(\cdot)$ is the usual Bessel function of the first kind of order zero, k is the wave number, and ω is wave frequency. Two parameters k, ω are related to each other through the dispersion relation $\omega^2 = kg \tanh(kh)$. The coefficient function $a(k)$ is determined with the initial wave disturbance η_0 , which is defined by spatial wave field when $t = 0$, by the Hankel transform:

$$a(k) = \int_0^\infty \eta_0(r) r J_0(kr) dr. \quad (4)$$

If the Hankel transform of the initial wave disturbance η_0 is known, the resulting wave motion can be predictable with (3).

For a two-dimensional system (x, z, t) , the problem for the nonsteady-state waves can be simply described [1–6] by

$$\eta(x, t) = \int_{-\infty}^\infty a(k) \exp\{i(kx \pm \omega t)\} dk, \quad (5)$$

where the coefficient function defined by the Fourier transform of $\eta(x, 0)/2$ is

$$a(k) = \frac{1}{2\pi} \int_{-\infty}^\infty \frac{\eta_0}{2} e^{-ikx} dx. \quad (6)$$

2.3. Forward and Inverse Problems. If the initial wave disturbance η_0 is specified, the coefficient function $a(k)$ can be directly obtained by the Hankel transform (4) for the three-dimensional problem and the Fourier transform (6) for the two-dimensional problem. The wave evolution η with respect to space and time variables is then simply computed by integrating (3) or (5) for the given coefficient function. The computation of the wave evolution from knowledge of the initial wave disturbance is defined as the forward problem.

On the other hand, the inverse problem of reconstructing initial wave field on the basis of the measured wave elevation on the water surface is not simple. Such inverse procedures

often involve ill-posed problems whose solution does not exist or lacks continuous dependence on the given data [7–11]. Thus, the problems are mathematically and numerically much more challenging than the forward problem. In the present study, the aim is focused to reconstruct the initial wave field given a wave history data recorded at a specific spatial point.

3. Reconstruction of Initial Wave Field

3.1. Reconstructing Procedure. The inverse procedure for reconstructing initial wave field can be achieved by two steps: (i) estimating the coefficient function $a(k)$ based on the measurement wave elevation at a fixed point, and (ii) transforming the estimated coefficient function into the initial wave field.

Formally, the first step for the reconstruction can be expressed as an operator equation:

$$H\{a(k)\} = \eta_M(t), \quad (7)$$

where $a(k)$ denotes the unknown coefficient function with a finite interval $|k| < N$ and $\eta_M(t)$ denotes the wave elevation measured at fixed-spatial point during the time interval $0 < t < T$ and the operator $H : a \rightarrow \eta_M$ is defined by

$$H\{\cdot\} = \begin{cases} \int_{-N}^N e^{i(kX \pm \omega t)} \{\cdot\} dk, & \text{for 2D problem,} \\ \int_0^N k J_0(kR) \cos(\omega t) \{\cdot\} dk, & \text{for 3D problem.} \end{cases} \quad (8)$$

In (8), X and R represent specific measurement points for 2D problem and 3D problem, respectively.

Let us denote a surface elevation measured at time t_i by η_{Mi} ; that is, $\eta_{Mi} = \eta_M(t_i)$, $i = 1, \dots, m$. The problem (7) is then approximated by the following uniformly discretized model:

$$\eta_{Mi} = \sum_{j=1}^n w_j H_{ij} a_j + \epsilon_i, \quad j = 1, \dots, n, \quad (9)$$

where w_j are weights for a quadrature rule, a_j are a vector representation for the coefficient function $a(k)$, that is, $a_j = a(k_j)$; ϵ_i are an additive random noise, and the elements of H_{ij} are given by

$$H_{ij} = \begin{cases} \exp\{i(k_j X \pm \omega t_i)\}, & \text{for 2D problem,} \\ k_j J_0(k_j R) \cos(\omega t_i), & \text{for 3D problem.} \end{cases} \quad (10)$$

Once the solution a_j is obtained by inverting (9), the initial wave field can be reconstructed by the second step:

$$\hat{\eta}_0 = \begin{cases} 2\mathcal{F}^{-1}(a), & \text{for 2D problem,} \\ \mathcal{H}^{-1}(a), & \text{for 3D problem,} \end{cases} \quad (11)$$

where \mathcal{F}^{-1} , \mathcal{H}^{-1} are the inverse Fourier transform and the inverse Hankel transform, respectively. The accent hat symbol means estimated value since it cannot be identical with the true value.

3.2. Instability. It is worth noting that the numerical system (9) is given by discretizing the integral operator (8), which is classified as Fredholm integral operator of the first kind. It is known that the approximation of such integral operator with a regular kernel yields a highly ill-conditioned system [7–11, 16–18]. As a result, the system (9) is ill conditioned and the inverse is greatly affected by small changes in input data (measurement). More specifically, this kind of system causes the following main difficulties in computational procedures [19]: (i) very large condition number of the matrix system, (ii) no guarantee about a useful solution from the replacement into a well-conditioned matrix.

Therefore, standard methods for obtaining the inverse solution cannot be simply used. Instead, more sophisticated methods should be applied to obtain a physically meaningful solution. In this study, we will address the difficulty by formulating the problem as an inference problem in an appropriately constructed stochastic model.

4. Stochastic Model

4.1. Bayesian Formulation. The aim of stochastic modeling is to derive the probability distribution, which is called the posterior probability density function (PPDF), for the unknown parameter conditional on available data according to Bayes' formula [18]:

$$P(a | \eta_M) \propto P(\eta_M | a) P(a), \quad (12)$$

where $P(\eta_M | a)$ is the likelihood function and $P(a)$ is the prior density function.

Bayesian formulation solves the ill-posed problem in a systematic way by modeling all parameters of interests as random variables with joint probability distributions. The probability distribution (12) can provide the information on the unknown parameter with the degree of confidence given the measurement.

The PPDF (12) can be specified through the appropriate probabilistic modeling for each component. Since the distribution (12) is proportional to the product of its likelihood function $P(\eta_M | a)$ and the prior density $P(a)$, the probabilistic modeling can be achieved by separately modeling these two probability distributions.

If we assume that the random error ϵ in the measurement is an additive Gaussian random noise with zero mean and standard deviation σ , the likelihood function $P(\eta_M | a)$ can be modeled [12–15, 18] as

$$P(\eta_M | a) \propto (\sigma^2)^{-m/2} \exp\left(-\frac{\|Ha - \eta_M\|_2^2}{2\sigma^2}\right), \quad (13)$$

where $\|\cdot\|_2$ refers to Euclidean norm and m is the number of measurements. It is worth emphasizing that other model can also be used to represent the likelihood function. But the Gaussian model (12) is the most commonly used since it is easy to handle and well fit the actual distribution.

Next, we need to consider the prior distribution $p(a)$ reflecting the knowledge of the unknowns before the data is measured. For the present study, there is no information

about prior probability. For this case, a noninformative prior such as uniform distribution can be a possible choice. However, a noninformative prior may not provide sufficient regularity for the solution of an inverse problem which is severely ill posed. In this study, the following pairwise Markov random field (MRF) model, which has been successfully used in various applications [12–15], is used for the smoothness of the unknown coefficient function a :

$$p(a) \propto \lambda^{n/2} \exp\left(-\frac{\lambda}{2} a^T W a\right), \quad (14)$$

where the matrix $W \in \mathbb{R}^{n \times n}$ is determined by

$$W_{ij} = \begin{cases} n_i, & i = j, \\ -1, & i \sim j, \\ 0, & \text{otherwise,} \end{cases} \quad (15)$$

where n_i is the number of neighbors for the point i and $i \sim j$ means that i and j are adjacent. The MRF is a model for a basis of contextual constraints in the prior distribution. It provides a way of synthesizing and capturing the characteristics of a stochastic process.

With the likelihood (13) and the prior models (14), the PPDF in (12) now can be expressed by

$$P(a | \eta_M) \propto (\sigma^2)^{-m/2} \exp\left(-\frac{\|Ha - \eta_M\|_2^2}{2\sigma^2}\right) \lambda^{n/2} \times \exp\left(-\frac{\lambda}{2} a^T W a\right). \quad (16)$$

Here, it is worth noting that the expression (16) is dependent on some parameters σ and λ , which are used for modeling likelihood function and the prior density. If these parameters are given, then the desired unknown parameter can be estimated by exploring the PPDF (16). However, in most cases, the choice of these parameters is also nondeterministic. In the stochastic model, these parameters are also modeled as random variables and have their own priors. The PPDF is then formulated as the following hierarchical model:

$$p(a, \lambda, \sigma | \eta_M) \propto (\sigma^2)^{-(m/2 + \alpha_2 + 1)} \times \exp\left(-\frac{\|Ha - \eta_M\|_2^2 + 2\beta_2}{2\sigma^2}\right) \lambda^{(n/2 + \alpha_1 - 1)} \times \exp\left(-\frac{\lambda(a^T W a + \beta_1)}{2}\right), \quad (17)$$

where (α_1, β_1) is a pair of gamma distribution for λ and (α_2, β_2) is a pair of inverse gamma distribution for the prior of σ .

4.2. Markov Chain Monte Carlo. Now the objective is to extract information on the unknown parameter a from the constructed PPDF (17). The resultant posterior distribution

can be analyzed using a Markov chain Monte Carlo (MCMC) sampling techniques. The PPDF (17) can be explored by the following hybrid algorithm which is designed by mixing Metropolis-Hastings steps in the Gibbs sampler [12–15].

- (i) Initialize $a^{(0)} = \{a_1^{(0)}, a_2^{(0)}, \dots, a_n^{(0)}\}$, $\lambda^{(0)}$, and $\sigma^{(0)}$,
- (ii) For $i = 0 : N_{\text{MCMC}} - 1$
 - Sample $a_1^{(i+1)} \sim p(a_1 | a_2^{(i)}, a_3^{(i)}, \dots, a_n^{(i)}, \lambda^{(i)}, \sigma^{(i)})$
 - Sample $a_2^{(i+1)} \sim p(a_2 | a_1^{(i+1)}, a_3^{(i)}, \dots, a_n^{(i)}, \lambda^{(i)}, \sigma^{(i)})$
 - ⋮
 - Sample $a_n^{(i+1)} \sim p(a_n | a_1^{(i+1)}, a_2^{(i+1)}, \dots, a_{n-1}^{(i+1)}, \lambda^{(i)}, \sigma^{(i)})$
 - Sample $u_1 \sim U(0, 1)$
 - Sample $\lambda^{(*)} \sim q(\lambda^{(*)} | \lambda^{(i)})$
 - if $u_1 < \min\{1, p(\lambda^{(*)} | a^{(i+1)}, \sigma^{(i)})q(\lambda^{(i)} | \lambda^{(*)})/p(\lambda^{(i)} | a^{(i+1)}, \sigma^{(i)})q(\lambda^{(*)} | \lambda^{(i)})\}$, $\lambda^{(i+1)} = \lambda^{(*)}$
 - else $\lambda^{(i+1)} = \lambda^{(i)}$
 - Sample $u_2 \sim U(0, 1)$
 - Sample $\sigma^{(*)} \sim q(\sigma^{(*)} | \sigma^{(i)})$
 - if $u_2 < \min\{1, p(\sigma^{(*)} | a^{(i+1)}, \lambda^{(i+1)})q(\sigma^{(i)} | \sigma^{(*)})/p(\sigma^{(i)} | a^{(i+1)}, \lambda^{(i+1)})q(\sigma^{(*)} | \sigma^{(i)})\}$, $\sigma^{(i+1)} = \sigma^{(*)}$
 - else $\sigma^{(i+1)} = \sigma^{(i)}$

where the full conditionals are derived as

$$p(\lambda | a, \sigma) \sim G\left(\frac{n}{2} + \alpha_1, \frac{a^T W a + 2\beta_1}{2}\right), \quad (18)$$

$$p(\sigma | a, \lambda) \sim IG\left(\frac{m}{2} + \alpha_2, \frac{\|Ha - \eta_M\|_2^2 + 2\beta_2}{2}\right),$$

where G represents gamma distribution and IG represents inverse gamma distribution.

5. Numerical Experiments

5.1. Example I: 2D Problem with Smooth-Type $a(k)$. As the first example, we consider the 2D problem of reconstructing the initial wave field with a smooth-type coefficient function $a(k)$. To illustrate the method, the following particular form of the initial surface disturbance is chosen:

$$\eta_0 = 0.2\sqrt{\pi}e^{-x^2/4}. \quad (19)$$

The transform (4) of (19) yields the smooth-type coefficient function

$$a(k) = 0.1e^{-x^2}. \quad (20)$$

Figure 2 shows the true initial disturbance (19) and the exact coefficient function (20).

The forward problem is first solved with the given initial wave field. The resulting nonsteady-state waves can be simply

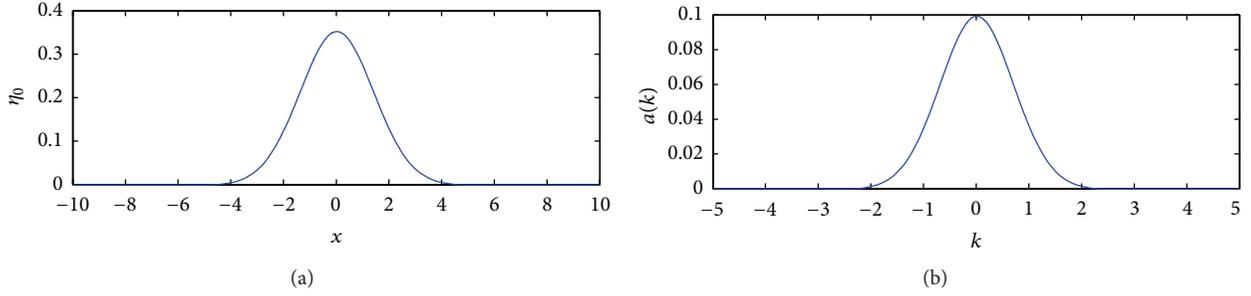


FIGURE 2: (a) True initial surface disturbance and (b) exact coefficient function in 2D example.

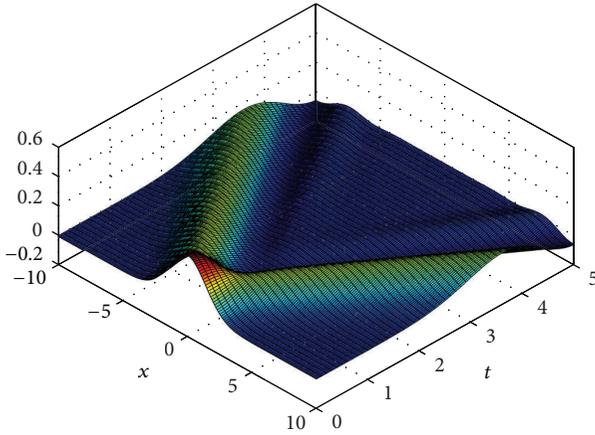
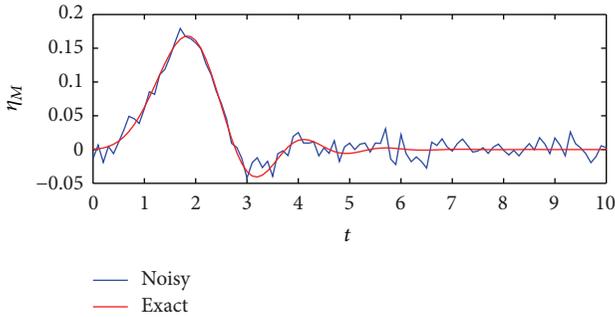


FIGURE 3: Spatial-time evolution of a 2D wave propagation.


 FIGURE 4: Time history of wave elevation collected at $x = 5$.

computed through (5) with the coefficient function (20). Figure 3 shows the computed spatial-time evolution of a nonsteady-state water wave propagation generated by the initial wave disturbance (19). For the forward calculation, the integral operator is directly discretized by Simpson's integration rule. The water depth h is normalized and $\Delta x = 0.1$, $\Delta t = 0.1$ are considered as time and space steps on the physical domain $[-10, 10] \times [0, 10]$. To discretize the integral operator (8), the integration limit N and Δk are taken to be 5 and 0.05, respectively.

From now, we consider the inverse problem of reconstructing the initial surface disturbance η_0 from the wave history data at a specific spatial point $x = X$. For illustrating purpose, a particular time history of wave elevation, η_M , is

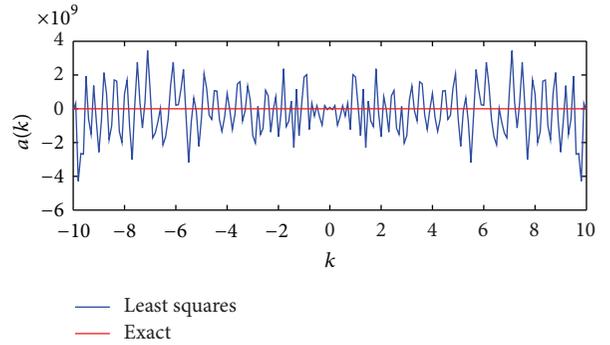


FIGURE 5: The coefficient function obtained through the least squares estimate.

collected at $x = 5$ and is shown in Figure 4. To mimic a practical situation, the Gaussian random noise ϵ is generated with the density $\sigma^2 = 0.01$ and added to the collected data. The resulting data with a random noise is considered as input data for the inverse problem.

As explained earlier, the system is severely ill posed and thus the inverse is mathematically and numerically much more challenging than the forward calculation. To examine the cause and the degree of instability of the system, the discretized system is decomposed as

$$\eta_M = \sum_{i=1}^n \sigma_i (v_i^T a) u_i, \quad (21)$$

where σ_i 's are the singular values ordered, $\sigma_1 \geq \sigma_2 \geq \dots > \sigma_\infty$, and $\{u_i\}_1^n$, $\{v_i\}_1^n$ are orthonormal singular vectors. The preceding decomposition is well known as the singular value decomposition [16–19]. The unknown a can then be obtained by the least squares estimate, that is, $\min_a \|Ka - \eta_M\|_2$:

$$a_{\text{LSQ}} = \sum_{i=1}^n \frac{u_i^T \eta_M}{\sigma_i} v_i. \quad (22)$$

With the measurement data in Figure 4, we first obtained the inverse solution using the least squares estimate (22). It can be seen from Figure 5 that the computed inverse solutions are totally unstable and thus useless. The cause and the degree of instability in the inverse solution can be understood by analyzing the behavior of singular systems, more specifically,

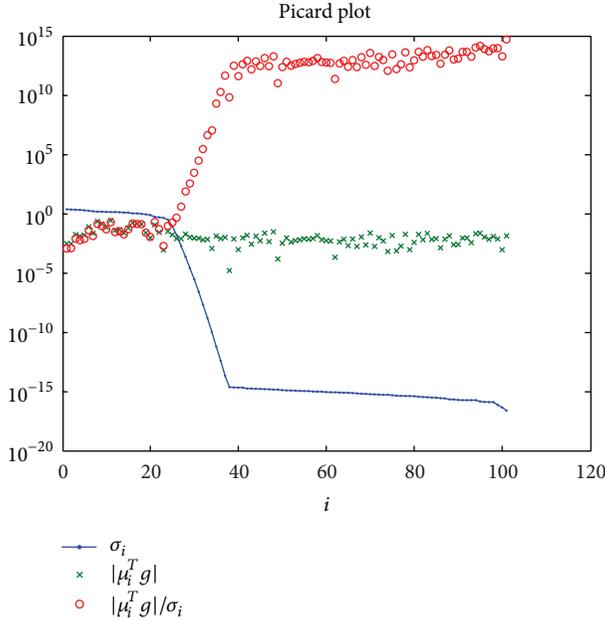


FIGURE 6: Picard plots.

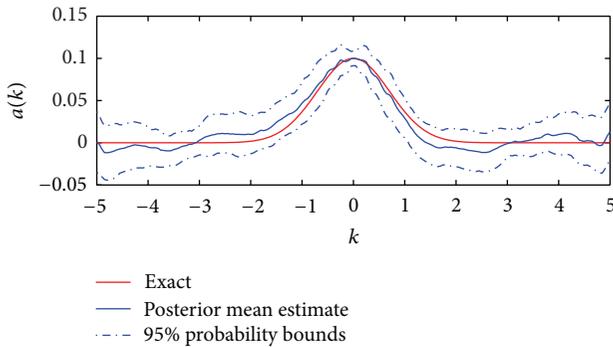
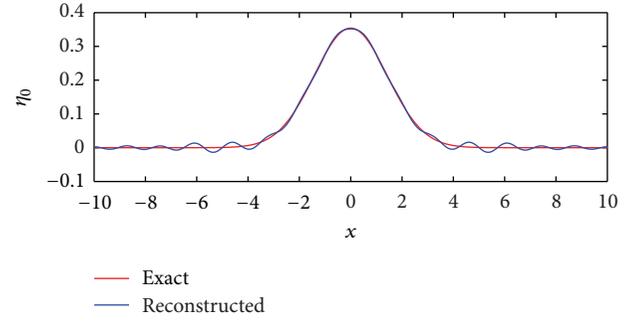


FIGURE 7: Posterior mean estimate for the coefficient function.

each component in the right-hand-side of (22). Figure 6 shows the Picard plot [19] that shows behaviors of $|u_i^T \eta_M|$, σ_i , and the ratio $|u_i^T \eta_M|/\sigma_i$. It can be found that the coefficients $|u_i^T \eta_M|$ become larger than most of singular values σ_i . Thus, the least squares solution for a , which is the sum of the ratio $|u_i^T \eta_M|/\sigma_i$, is dominated by the small singular values. This explains the reason for the unstable solution in Figure 5. The instability, frequent sign changes, may be accelerated and amplified with the increasing amount of noise in data.

With this analysis in mind, we next construct the stochastic inverse model (17) in terms of the measured wave history data. To extract useful information from the probability density, MCMC algorithm described in Section 4 is used. The initial guess for λ , σ , and a are taken to be 1, 0.1, and zero vector, respectively. The pairs of parameters (α_1, β_1) , (α_2, β_2) are taken to be $(10^{-1}, 10)$, $(10^{-1}, 10)$, respectively. The total number of samples N_{MCMC} for the algorithm was taken to be 50,000 and the last 25,000 samples were used to estimate statistics for a since it allows the beginning of

FIGURE 8: Reconstructed initial wave disturbance η_0 for the 2D example.

the Markov chain. After obtaining random samples from the MCMC algorithm, it is necessary to examine the issues such as convergence and mixing of the chain to determine whether simulated set of samples can reasonably be treated as a set of random realizations from the target distribution. In this study, marginal trace plots are used to check if the chain works properly. All the components have the stationary distribution as its marginal distribution.

The true and estimated coefficient functions are shown in Figure 7. The upper and lower dotted lines denote the 95% credible interval. The credible interval quantifies the degree of uncertainties in the solution. The mean estimate of the posterior density is fairly stable and accurate compared with the results in Figure 8. Finally, the initial wave field η_0 is reconstructed using the posterior mean of the coefficient function. The reconstructed initial surface disturbance in Figure 8 represents accurate approximation to the true initial disturbance.

5.2. Example II: 3D Problem with Nonsmooth-Type $a(k)$. As the second example, we consider the 3D problem of reconstructing the initial wave field with a nonsmooth-type coefficient function $a(k)$, which is believed to be a more challenging case than the smooth type. For the purpose, we first specify the initial surface disturbance for 3D problem as

$$\eta_0 = \frac{1}{r} J_1(r), \quad (23)$$

where J_1 is Bessel function of the first kind of order one and $r = \sqrt{x^2 + y^2}$. The transform of (21) yields the nonsmooth type coefficient function:

$$a(k) = \begin{cases} 1, & 0 \leq k \leq 1, \\ 0, & 1 < k < \infty. \end{cases} \quad (24)$$

Figure 9 shows the true initial disturbance (21) and the exact coefficient function (22).

In the similar manner of the previous example, the forward problem is first solved with the given initial surface disturbance. For the calculation, the water depth h is normalized and $\Delta r = 0.1$, $\Delta t = 0.1$ are considered as time and space steps on the physical domain $[0, 10] \times [0, 10]$. When discretizing the integral operator (8), the integration limit N

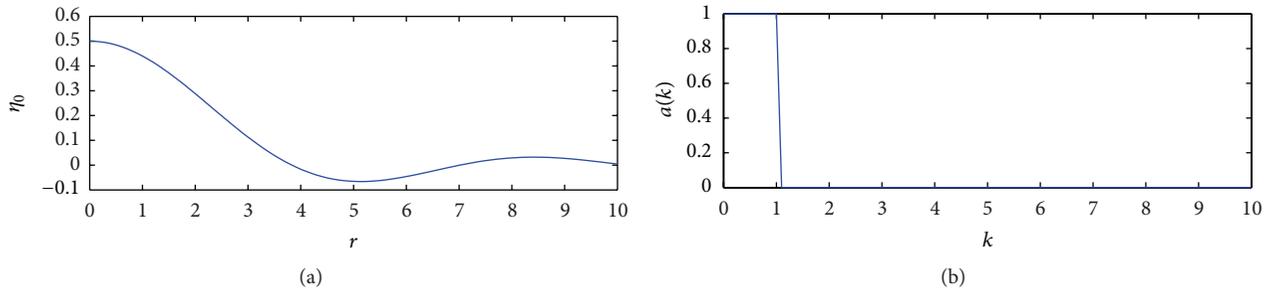


FIGURE 9: (a) True initial wave disturbance and (b) exact coefficient function in 3D example.

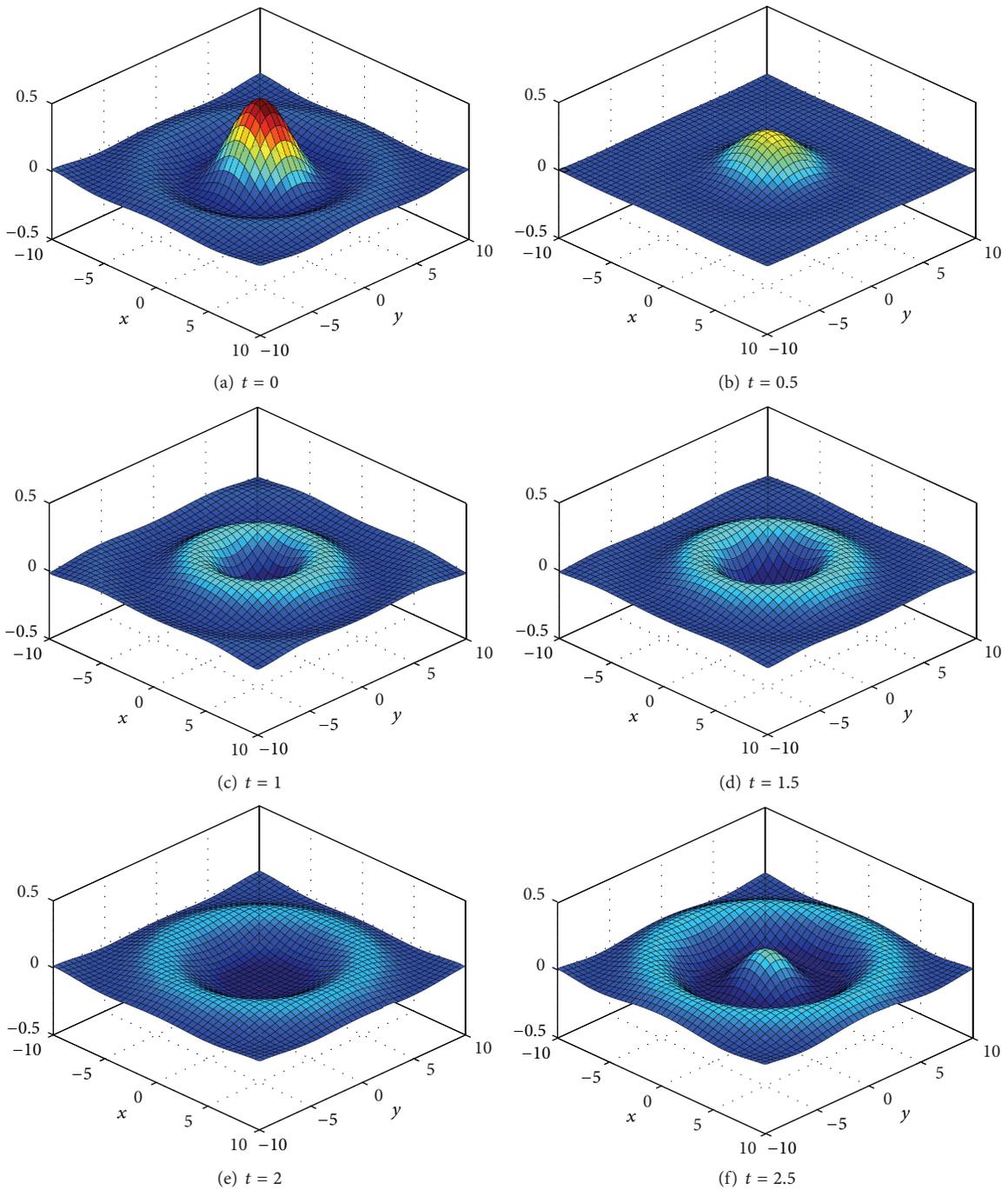


FIGURE 10: Spatial-time evolution of 3D wave propagation.

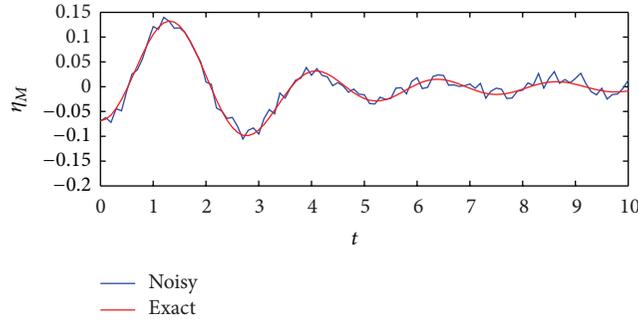
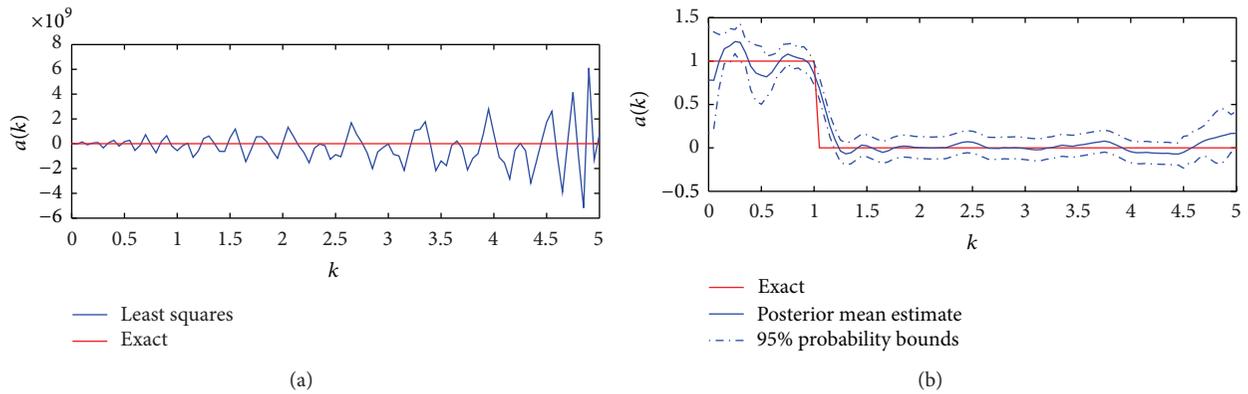
FIGURE 11: Time history of wave elevation collected at $r = 5$.

FIGURE 12: (a) Least squares estimate and (b) Posterior mean estimate of the MCMC samples for the coefficient function.

and Δk are taken to be 5 and 0.05, respectively. Figure 10 shows the computed spatial-time evolution for 3D problem.

The wave history data in Figure 11, which is recorded at $r = 5$, is used to illustrate the reconstructing procedure. To generate noisy data, the random noise ϵ with the variance $\sigma^2 = 0.01$ is added to the collected data. With this measurement, the coefficient function is estimated through the MCMC algorithm. The same conditions as in the first example are used for the MCMC algorithm. The posterior mean estimate of the simulated random samples is shown in Figure 12. For the purpose of comparison, the least squares estimation of the coefficient function is also illustrated. The result indicates that the posterior mean is in good agreement with the exact coefficient function while the result from least squares estimate is totally unstable. Using this posterior mean estimate, the initial wave field is reconstructed and shown in Figure 13. It can be found that the estimated result accurately approximate to the true initial disturbance.

It is also worth to compare the performances of the existing algorithms and the stochastic inverse approach illustrated in this paper to underline the advantages and disadvantages. For this purpose, Tikhonov regularization method [8, 11, 16–19], which is known as a representative deterministic inverse approach, was employed to solve the second example with the same conditions. Tikhonov regularized solution for $a(k)$ is the minimizer of the functional

$$F(a) = \|Ha - \eta_M\|_2^2 + \delta \|a\|_2^2, \quad (25)$$

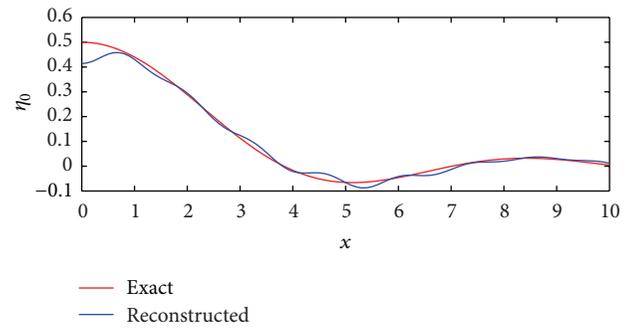


FIGURE 13: Reconstruction of the initial wave disturbance for the 3D example.

where $\delta > 0$ is the regularization parameter.

In deterministic inverse approach, the choice of the regularization parameter is very important because it controls the performance of the regularization. However, selecting an appropriate regularization parameter is not trivial since it depends on both the regularization method itself and the mathematical model of interest in a very complex way. Generally, the regularization parameter is selected by the heuristic method. First, the inverse problem is solved with a set of regularization parameters. The solution remains practically unchanged for a certain range of the parameter. The regularization parameter can then be chosen from this range. Figure 14 shows the Tikhonov regularized solutions to

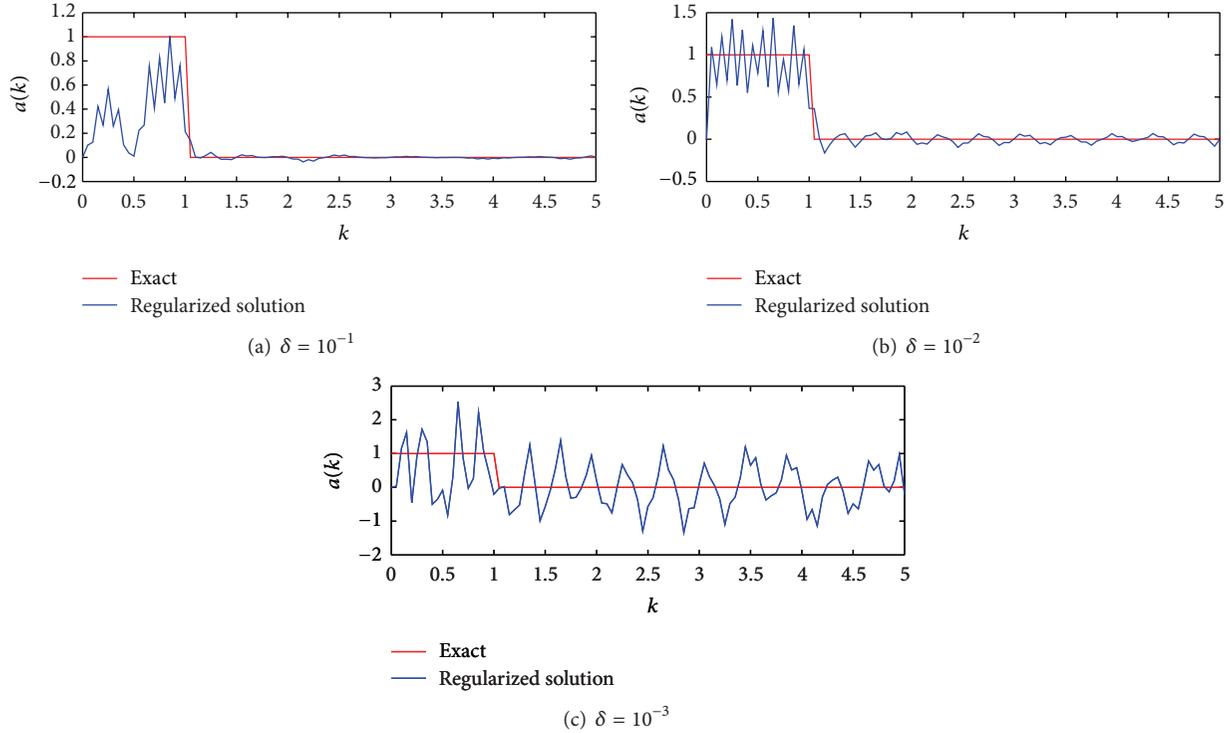


FIGURE 14: Tikhonov regularized solution for three different regularization parameters.

three different regularization parameters. It can be observed that the poor solution is obtained when the regularized solution to the regularization parameter is large. Furthermore, the solution loses its stability when the regularization parameter is small. Therefore, it is necessary to pay attention to the behavior of the regularized solution, when choosing the regularization parameter. It is seen by comparing with Figure 12 that the proposed stochastic inverse approach is smoother than the Tikhonov method.

Compared with the deterministic inverse methods for inverse problems, the proposed stochastic inverse approach has some advantages. Firstly, it provides a natural framework for quantifying the uncertainties by considering a complete probabilistic description of the inverse problem of interest. Generally, there exist many possible candidates for the solution, because of the ill-posed nature of the inverse problem. But, the deterministic inverse methods provide only a single estimate taken from these possible candidates based on the choice of the regularization parameter. Secondly, it also provides a flexible way to choose a nondeterministic parameter by formulating the hierarchical model.

However, there also exist disadvantages. In this work, a computational framework for reconstructing the initial wave field was built based on the Markov chain Monte Carlo technique. In computational mathematics, the efficient computation is important. The key behind the Markov chain is to generate samples from the target distribution based on the iteration. To ensure the chain convergence, quite large numbers of samples are required. Thus, it is expected that the solver will be computationally very expensive because the computational algorithm of the stochastic inverse model

requires the direct simulation. In contrast, the deterministic inverse methods can remarkably reduce the computational time by avoiding the time-consuming iterations.

6. Conclusions

The conclusions can be summarized as follows.

- (1) The fluid motion has been taken to be homogeneous, incompressible, and inviscid such that the velocity potential exists and satisfies Laplace's equation. On this basis, it has been shown that, for a given initial surface disturbance, the surface wave motion can be mathematically expressed as an integral transform, which can be used to compute the time-spatial evolution of the wave motion.
- (2) The inverse problem of reconstructing the initial surface disturbance was formulated with a time history data of surface elevation which is collected at a specific spatial point. The inverse problem requires solving an ill-posed system that causes numerical instability. It was shown that, for such system, a standard method such as least squares estimate fails to obtain the stable solution. To address the difficulty, the system was formulated as an inference problem in an appropriately constructed stochastic model.
- (3) The reconstruction procedure based on the stochastic inverse model has been evaluated through an analysis of some digitally simulated data for numerical examples related to two- and three-dimensional fluid

motion. It has been shown that the proposed procedure can be used to not only reconstruct the initial surface disturbance but also detect uncertainties in solution arising from measurement error.

In this study, the quality of the reconstruction has been illustrated by two examples, where the simulated data fulfilled the assumptions about the noise. Future work on this study will also include the application of real data for ensuring robustness of the method.

References

- [1] C. C. Mei, *The Applied Dynamics of Ocean Surface Waves*, World Scientific, London, UK, 1989.
- [2] J. J. Stoker, *Water Waves*, Wiley Classics Library, John Wiley & Sons, New York, NY, USA, 1992.
- [3] G. B. Whitham, *Linear and Nonlinear Waves*, Pure and Applied Mathematics, John Wiley & Sons, New York, NY, USA, 1999.
- [4] B. L. Mehaute and S. Wang, *Waves Generated by Underwater Explosion*, World Scientific, 1996.
- [5] J. Lighthill, *Waves in Fluids*, Cambridge University Press, Cambridge, UK, 1978.
- [6] A. R. Osborne, *Nonlinear Ocean Waves and the Inverse Scattering Transform*, vol. 97 of *International Geophysics Series*, Elsevier/Academic Press, Boston, Mass, USA, 2010.
- [7] T. S. Jang, S. H. Kwon, and B. J. Kim, "Solution of an unstable axisymmetric Cauchy-Poisson problem of dispersive water waves for a spectrum with compact support," *Ocean Engineering*, vol. 34, no. 5-6, pp. 676–684, 2007.
- [8] T. S. Jang and S. L. Han, "Application of Tikhonov's regularization to unstable water waves of the two-dimensional fluid flow: spectrum with compact support," *Ships and Offshore Structures*, vol. 3, no. 1, pp. 41–47, 2008.
- [9] T. S. Jang, S. L. Han, and T. Kinoshita, "An inverse measurement of the sudden underwater movement of the sea-floor by using the time-history record of the water-wave elevation," *Wave Motion*, vol. 47, no. 3, pp. 146–155, 2010.
- [10] T. S. Jang, H. G. Sung, and J. Park, "A determination of an abrupt motion of the sea bottom by using snapshot data of water waves," *Mathematical Problems in Engineering*, vol. 2012, Article ID 472575, 16 pages, 2012.
- [11] S. H. Kwon, C. H. Kim, and T. S. Jang, "An identification of wave propagation based on a single-point measurement," *Ocean Engineering*, vol. 34, no. 10, pp. 1405–1412, 2007.
- [12] J. Wang and N. Zabarar, "Hierarchical Bayesian models for inverse problems in heat conduction," *Inverse Problems*, vol. 21, no. 1, pp. 183–206, 2005.
- [13] J. Wang and N. Zabarar, "A Markov random field model of contamination source identification in porous media flow," *International Journal of Heat and Mass Transfer*, vol. 49, no. 5-6, pp. 939–950, 2006.
- [14] S. L. Han and T. Kinoshita, "Nonlinear damping identification in nonlinear dynamic system based on stochastic inverse approach," *Mathematical Problems in Engineering*, vol. 2012, Article ID 574291, 20 pages, 2012.
- [15] S. L. Han and T. Kinoshita, "Investigation of a stochastic inverse method to estimate an external force: applications to a wave-structure interaction," *Mathematical Problems in Engineering*, vol. 2012, Article ID 175036, 25 pages, 2012.
- [16] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, vol. 120 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1996.
- [17] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Informatica International, 1993.
- [18] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, vol. 160 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 2005.
- [19] P. C. Hansen, *Rank-Deficient and Discrete ill-Posed Problems*, SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa, USA, 1998.

Research Article

A Fast High-Order Total Variation Minimization Method for Multiplicative Noise Removal

Xiao-Guang Lv,¹ Jiang Le,^{1,2} Jin Huang,² and Liu Jun²

¹ School of Science, Huaihai Institute of Technology, Lianyungang, Jiangsu 222005, China

² School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 610054, China

Correspondence should be addressed to Xiao-Guang Lv; xiaoguangle@126.com

Received 17 October 2012; Revised 10 March 2013; Accepted 13 March 2013

Academic Editor: Fatih Yaman

Copyright © 2013 Xiao-Guang Lv et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Multiplicative noise removal problem has received considerable attention in recent years. The total variation regularization method for the solution of the noise removal problem can preserve edges well but has the sometimes undesirable staircase effect. In this paper, we propose a fast high-order total variation minimization method to restore multiplicative noisy images. The proposed method is able to preserve edges and at the same time avoid the staircase effect in the smooth regions. An alternating minimization algorithm is employed to solve the proposed high-order total variation minimization problem. We discuss the convergence of the alternating minimization algorithm. Some numerical results show that the proposed method gives restored images of higher quality than some existing multiplicative noise removal methods.

1. Introduction

Image denoising problem has been widely studied in the areas of image processing. The problem includes additive noise removal and multiplicative noise removal. Most of the literature deals with the additive noise model [1–4]. Given a noisy image $g = u + v$, where u is the original image and v is the noise, the denoising problem is to recover u from the observed image g . The additive noise model has been extensively studied. We refer the reader to [5–11] and references therein for a review of various methods.

In this paper, we consider the problem of seeking the original image from a multiplicative noisy image. In the multiplicative noise model, the recorded image g is the multiplication of the original image u and the noise v :

$$g = uv. \quad (1)$$

This problem arises in ultrasound imaging, synthetic aperture radar (SAR) and sonar (SAS), laser imaging, and magnetic field inhomogeneity in MRI [12, 13]. In this work, we concentrate on the assumption that v follows a Gamma distribution, which commonly occurs in SAR. With respect to a given resolution cell of the imaging device, the complex SAR image

is computed after the SAR system receives the coherent sum of reflected monochromatic microwaves. The complex SAR image at one point is usually modeled by a complex zero-mean circular Gaussian density (i.e., the real and imaginary parts are independent Gaussian variable with a common variance). The intensity image is defined as the square of the magnitude of the corresponding complex SAR image; its variables follow exponentially distribution independently [14, 15]. Since the complex observation is acquired in highly coherent system, the intensity image occurs to have peculiar granular appearance (known as speckle). The conventional SAR system reduces the speckle effect of the intensity image by the multilooking process (L -look, in the case of L -looks), that is, averaging the observed intensity images from slightly different angles of the same resolution cell [16].

Due to the coherent nature of these image acquisition processes, nearly all the information of the original image may vanish when it is distorted by the multiplicative noise. In general, the standard additive denoising method, so prevalent in image processing, is inadequate. Therefore, it is necessary to devise efficient and reliable algorithms for recovering the true image from the observed multiplicative noisy image. Until the past decade, a few variational approaches devoted to

multiplicative noise removal have been proposed. The early variational approach for multiplicative noise removal is the one by Rudin et al. [17]. According to the statistical properties of the multiplicative noise v , the recovery of the image u was based on solving the following constrained optimization problem [17]:

$$\begin{aligned} \min_u \quad & \int_{\Omega} |Du| \, dx \, dy \\ \text{subject to:} \quad & \int_{\Omega} \frac{g}{u} \, dx \, dy = 1, \\ & \int_{\Omega} \left(\frac{g}{u} - 1 \right)^2 \, dx \, dy = \sigma^2, \end{aligned} \quad (2)$$

where $\int_{\Omega} |Du| \, dx \, dy$ is denoted by the seminorm on boundary variation (BV) space that coincides with $\int_{\Omega} \sqrt{u_x^2 + u_y^2} \, dx \, dy$ when u is smooth. The two constraints state that the mean of the noise is equal to 1, and the variance is equal to σ^2 . In (2), only basic statistical properties, the mean, and the variance of the noise v are considered, which somehow limits the restored results.

By using a maximum a posteriori (MAP) estimator, Aubert and Aujol [18] proposed a function whose minimizer corresponds to the denoised image to be recovered. This function is

$$\min_u \int_{\Omega} \left(\log u + \frac{g}{u} \right) \, dx \, dy + \alpha \int_{\Omega} |Du| \, dx \, dy, \quad (3)$$

where the total variation of u is utilized as the regularization term, and α is the regularization parameter which controls the trade-off between a good fit of g and a smoothness requirement due to the total variation regularization. Although the functional they proposed is not convex, they still proved the existence of a minimizer and showed the capability of their model on some numerical examples.

As a result of the drawback of the objective function (3) that is not convex for all u , the obtained solution is likely not the global optimal solution of (3). To overcome this problem, Huang et al. [19] proposed and studied a strictly convex objective function for multiplicative noise removal problem. In [19], the authors introduced an auxiliary variable $z = \log u$ in (3). With the new variable, the proposed unconstrained total variation denoising problem is described as follows:

$$\min_{z,w} \int_{\Omega} (z + ge^{-z}) \, dx \, dy + \alpha_1 \|z - w\|_{L^2}^2 + \alpha_2 \int_{\Omega} |Dw| \, dx \, dy, \quad (4)$$

where α_1 and α_2 are positive regularization parameters. The parameter α_1 measures the trade-off between an image obtained by a maximum likelihood estimation from the first term and a total variation denoised image w . The parameter α_2 measures the amount of regularization to a denoised image w . The main advantage of the proposed method is that the total variation norm can be used in the noise removal process in an efficient manner. Therefore the method proposed in [19] has the ability to preserve edges very well in the denoised image.

Recently, Bioucas-Dias and Figueiredo [20] proposed an efficient multiplicative noise removal method by using variable splitting and constrained optimization. After converting the multiplicative noise model into an additive one by taking logarithms, they used variable splitting to obtain an equivalent constrained problem and then solved this optimization problem by using the augmented Lagrangian method. A set of experiments has shown that the proposed method, which they named MIDAL (multiplicative image denoising by augmented Lagrangian), yields state-of-the-art results both in terms of speed and denoising performance. In [21], Steidl and Teuber considered a variational restoration model consisting of the I -divergence as data fitting term and the total variation seminorm or nonlocal means as regularizer for removing multiplicative noise. They proposed to compute the minimizers of their restoration functionals by using Douglas-Rachford splitting techniques. For a particular splitting, they presented a semi-implicit scheme to solve the involved nonlinear systems of equations and proved its Q -linear convergence.

It is known that the total variation denoising method is a PDE-based technique that preserves edges well but has the sometimes undesirable staircase effect, namely, the transformation of smooth regions (ramps) into piecewise constant regions (stairs) [22–26]. Therefore, the approaches involving the traditional total variation often cause staircase effect since they favor solutions that are piecewise constant. Staircase solutions fail to satisfy the evaluation of visual quality and they can develop false edges that do not exist in the true image. To attenuate the staircase effect, there is a growing interest in the literature for replacing the total variation norm by the high-order total variation norm. The motivation behind such attempt is to restore potentially a wider of images, which comprise more than merely piecewise constant regions. The majority of the high-order norms involve second-order differential operators because piecewise-vanishing second-order derivatives lead to piecewise-linear solutions that better fit smooth intensity changes; see [27] for more details. There are two classes of second-order regularization methods for image restoration problems. The first class employs a second-order regularizer in a standalone way. For example, in [28], the authors considered a fourth-order partial differential equations (PDE) model for noise removal and employed the dual algorithm of Chambolle for solving the high-order problems. In [29], Hu and Jacob applied higher degree total variation (HDTV) regularization for image recovery. The second class is to combine the TV norm with a second-order regularizer. For example, a technique in [30, 31] combining the TV filter with a fourth-order PDE filter was proposed to preserve edges and at the same time avoid the staircase effect in smooth regions for noise removal. In [32], Papafitsoros and Schönlieb considered a high-order model involving convex functions of the TV and the TV of the first derivatives for image restoration problems and used the split Bregman method to solve numerically the corresponding discretized problem.

In this paper, we present a fast high-order total variation minimization method for multiplicative noise removal problem. This technique substantially reduces the staircase

effect, while preserving sharp jump discontinuities (edges). An alternating minimization algorithm is employed to solve the proposed high-order total variation minimization model. We discuss the convergence of the proposed alternating minimization algorithm. Some numerical results show that the proposed method gives restored images of higher quality than some existing multiplicative noise removal methods.

The organization of this paper is outlined as follows. In the next section, we present the high-order total variation minimization method for multiplicative noise removal problem. In Section 3, an alternating minimization algorithm is employed to find the minimizer of the proposed minimization problem. We analyze the convergence of the proposed alternating minimization algorithm in Section 4. Some numerical experiments are given to illustrate the performance of the proposed algorithm in Section 5. Concluding remarks are given in the last section.

2. A High-Order Total Variation Multiplicative Denoising Model

In this section, our aim is to propose a strictly convex objective function for restoring images distorted by the multiplicative noise. The most important is that we apply a high-order total variation to the objective function to achieve sharp edges and staircase reduction efficiently. We introduce our multiplicative denoising model from the statistical perspective using the Bayesian formula.

Let g , u , and v denote samples of instances of some random variables F , U , and V . Moreover, we assume that the random variable v on each pixel is mutually independent and identically distributed (i.i.d). Moreover, the random variable v in each pixel follows Gamma distribution; that is, its probability density function is

$$P_V(v; L, M) = \frac{M^L v^{L-1}}{\Gamma(L)} e^{-Mv} \quad \text{for } v > 0, \quad (5)$$

where $\Gamma(\cdot)$ is the usual Gamma-function, and L and M denote the shape scale and inverse parameters in the Gamma distribution, respectively. We note that the mean of a gamma-distributed variable is L/M and the variance is L/M^2 . As the multiplicative noise, in this work we assume that the mean of v equals 1, and then we have $L = M$.

According to the posteriori estimation, the restored image u can be determined by

$$u = \arg \max_u P_{U|G}(u | g). \quad (6)$$

From the Bayes rule, we have $P_{U|G}(u | g) = P_{G|U}(g | u)P_U(u)/P_G(g)$. It then follows that

$$u = \arg \max_u \frac{P_{G|U}(g | u) P_U(u)}{P_G(g)}. \quad (7)$$

Using the proposition in [18]: $P_V(g/u)(1/u) = P_{G|U}(g | u)$, we obtain

$$P_{G|U}(g | u) = \frac{L^L g^{L-1}}{u^L \Gamma(L)} e^{-Lg/u}. \quad (8)$$

Taking the logarithm transformation into account, we assume that the image prior $P_U(u)$ is given by the following:

$$P_U(u) = P_{U|W}(u | w) P_W(w) \quad (9)$$

with

$$P_{U|W}(u | w) \propto e^{-\alpha_1 \|\log u - w\|_2^2}, \quad (10)$$

$$P_W(w) = e^{-\alpha_2 \|w\|_{\text{HTV}}}, \quad (11)$$

where α_1 and α_2 are two positive scalars and $\|w\|_{\text{HTV}}$ is a high-order total variation of w . Here we assume that the difference between $\log u$ and w follows a Gaussian distribution, and w follows a high-order total variation prior. Thus, maximizing $P_{U|G}(u | g)$ amounts to minimizing the log-likelihood:

$$\begin{aligned} -\log(P_{U|G}(u | g)) &= -\log(P_{G|U}(g | u)) \\ &\quad -\log(P_{U|W}(u | w)) - \log(P_W(w)) \\ &\quad + \log(P_G(g)). \end{aligned} \quad (12)$$

Since $P_G(g)$ is constant, (6) can be rewritten as the following optimization problem:

$$\begin{aligned} u = \arg \min_u & \left(-\log(P_{G|U}(g | u)) - \log(P_{U|W}(u | w)) \right. \\ & \left. - \log(P_W(w)) \right). \end{aligned} \quad (13)$$

The previous computation leads us to propose the following high-order total variation model for restoring images corrupted by the Gamma noise by introducing a new variable $z = \log u$:

$$\begin{aligned} \min_{z,w} \mathcal{J}(z, w) &\doteq \min_{z,w} \int_{\Omega} (z + ge^{-z}) dx dy + \alpha_1 \|z - w\|_{L_2}^2 \\ &\quad + \alpha_2 \int_{\Omega} |D^2 w| dx dy, \end{aligned} \quad (14)$$

where α_1 and α_2 are positive regularization parameters. In the above model, $D^2 w$ denotes the Hessian of w , and $|D^2 w| = \sqrt{w_{xx}^2 + w_{xy}^2 + w_{yx}^2 + w_{yy}^2}$ is just the Frobenius norm of the Hessian $D^2 w$. The parameter α_1 measures the trade-off between an image obtained by a maximum likelihood estimation from the first term and a high-order total variation denoised image w . The parameter α_2 measures the amount of regularization to a denoised image w . The main advantage of the proposed method is that the high-order total variation norm can be used in the noise removal process in an efficient manner. Therefore, the proposed method has the ability to preserve edges very well and reduce staircase effect in the denoised image.

Moreover, it is obvious that when u includes an edge, z also contains an edge at the same point; that is, the logarithm transformation preserves image edges; see [19] for more

details. Based on this idea, we can view z as an image in the logarithm domain. We can apply a high-order total variation regularization to z to recover its edges and reduce staircase effect. It is interesting to note that z in the proposed model can be positive, zero, or negative, and the corresponding $u = e^z$ is still positive, although u in the objective function (3) should be positive. Especially, one can find that when z in the proposed model is a large negative value, u is just close to zero.

3. The Alternating Minimization Algorithm

In this section, an alternating minimization algorithm is employed to solve the proposed high-order total variation minimization problem efficiently. To solve (14), we need to consider the following two minimization subproblems.

(i) For w fixed, determine the solution of

$$\min_z \int_{\Omega} (z + ge^{-z}) dx dy + \alpha_1 \|z - w\|_{L_2}^2. \quad (15)$$

(ii) For z fixed, determine the solution of

$$\min_w \alpha_1 \|z - w\|_{L_2}^2 + \alpha_2 \int_{\Omega} |D^2 w| dx dy. \quad (16)$$

In the discrete setting, minimizing the first problem amounts to solve the following optimization problem:

$$\min_z \sum_1^{n^2} (z_i + ge^{-z_i}) + \alpha_1 \|z - w\|_2^2. \quad (17)$$

Straightforward computations can show that the minimizer of the above problem is the solution of the following nonlinear systems:

$$1 - ge^{-z_i} + 2\alpha_1 (z_i - w_i) = 0 \quad \text{for } i = 1, 2, \dots, n^2. \quad (18)$$

There are n^2 decoupled nonlinear equations to be solved. The second derivative with respect to z of the first term in (17) is equal to ge^{-z} , which is always greater than zero. Here we assume that $g > 0$ in the multiplicative noise model (1). Therefore, the first term of the objective function is strictly convex for all z . Hence, we know that the objective function (17) is also strictly convex. Therefore, the corresponding nonlinear equation (18) has a unique solution and the solution can be determined by using the Newton method very efficiently.

Subproblem (16) is a high-order total variation regularization process for image denoising. Some effective algorithms for the minimizer of the high-order total variation problem have been proposed in the literature. For example, a lot of PDE-based methods are widely used to preserve edges and reduce the staircase effect for the high-order total variation problem; see [6, 22, 23, 30] for more details. In particular, inspired by the work from Chambolle [5], a dual algorithm and its convergence analysis for minimization of the high-order model are presented in [24, 28]. The algorithm overcomes the numerical difficulties related to the

nondifferentiability of the high-order model. Some numerical results show that the convergence of the dual algorithm is faster than the gradient descent time-marching algorithm.

In the present paper, we employ the dual algorithm proposed by [24, 28] for the high-order total variation denoising problem (16). The discrete version of problem (16) is represented as

$$\min_w \frac{\alpha_1}{\alpha_2} \|z - w\|_2^2 + \|w\|_{\text{HTV}}. \quad (19)$$

In the dual algorithm, we need to solve the following minimization problem:

$$\begin{aligned} \min_p \quad & \|z - \beta \operatorname{div}^2 p\|, \\ \text{subject to} \quad & |p_{j,k}| \leq 1, \quad \forall 1 \leq j, k \leq n, \end{aligned} \quad (20)$$

where $\beta = \alpha_2/2\alpha_1$ and $p = \begin{pmatrix} p^{11} & p^{12} \\ p^{21} & p^{22} \end{pmatrix}$ with $p^{hk} \in R^{n \times n}$ for $h, k = 1, 2$. This problem can be solved by a semi-implicit fixed point iteration:

$$p_{j,k}^{(i+1)} = \frac{p_{j,k}^{(i)} - \tau A_{j,k}^{(i)}}{1 + \tau |A_{j,k}^{(i)}|}, \quad (21)$$

where $p^0 = 0$, $A_{j,k}^{(i)} = (\nabla^2(\operatorname{div}^2 p^{(i)} - z/\beta))_{j,k}$, and $\tau \leq 1/64$ is the step size. The operator ∇^2 is defined as $\nabla^2 w = \begin{pmatrix} w_{xx} & w_{xy} \\ w_{yx} & w_{yy} \end{pmatrix}$. Here, $\operatorname{div}^2 p_{j,k}^{(i)} = D_{xx}((p^{11})_{j,k}^{(i)}) + D_{yx}^-((p^{12})_{j,k}^{(i)}) + D_{xy}^+((p^{21})_{j,k}^{(i)}) + D_{yy}((p^{22})_{j,k}^{(i)})$; see [24, 28] for more details. After the minimizer p^* of the constrained optimization problem in (20) is determined, the denoised image w can be computed by $w = z - \beta \operatorname{div}^2 p^*$. In addition, it is shown that if $\tau \leq 1/64$, then $\operatorname{div}^2 p^{(i)}$ is convergent and $p^{(i)}$ with the dual method is the solution of the minimization problem (20) as $i \rightarrow \infty$.

Starting from an initial guess $w^{(0)}$, the alternating minimization method computes a sequence of iterates

$$z^{(1)}, w^{(1)}, z^{(2)}, w^{(2)}, \dots, z^{(k)}, w^{(k)}, \dots \quad (22)$$

such that

$$\begin{aligned} z^{(k)} \doteq \mathcal{S}_h(w^{(k-1)}) &= \arg \min_z \sum_1^{n^2} (z_i + ge^{-z_i}) \\ &\quad + \alpha_1 \|z - w^{(k-1)}\|_2^2, \end{aligned} \quad (23)$$

$$w^{(k)} \doteq \mathcal{S}_p(z^{(k)}) = \arg \min_w \frac{\alpha_1}{\alpha_2} \|z^{(k)} - w\|_2^2 + \|w\|_{\text{HTV}}$$

for $k = 1, 2, \dots$

We are now in a position to describe the alternating minimization algorithm for the multiplicative noise removal problem.

Algorithm 1. The alternating minimization algorithm for multiplicative noise removal:

Choose an initial guess $w^{(0)}$, and $k = 1$.

- (1) For $w^{(k-1)}$ fixed, employ the Newton iterative method to compute

$$z^{(k)} = \arg \min_z \sum_1^{n^2} (z_i + ge^{-z_i}) + \alpha_1 \|z - w^{(k-1)}\|_2^2. \quad (24)$$

- (2) For $z^{(k)}$ fixed, use a fixed point iteration to compute

$$p^* = \arg \min_p \|z^{(k)} - \beta \operatorname{div}^2 p\| \quad (25)$$

and obtain

$$w^{(k)} = z^{(k)} - \beta \operatorname{div}^2 p^*. \quad (26)$$

- (3) Check the stopping criteria. If a stopping criterion is satisfied, then exit with an approximate restored image $u^{(k)} = e^{z^{(k)}}$; otherwise, let $k = k + 1$ and go to step (1).

4. Convergence Analysis

The main aim of this section is to analyze the convergence of the proposed alternating minimization algorithm. We first remark that $\mathcal{F}(z, w)$ in (14) is strictly convex, as the sum of a convex function and of two strictly convex functions; see [33] for more details. Hence, it suffices to show that there exists a unique minimizer of the objective function $\mathcal{F}(z, w)$. We denote by \mathcal{F}^* the minimum value of $\mathcal{F}(z, w)$. In the following we show that our alternating minimization algorithm gives asymptotically the solution of the discrete problem associated with (14). We have the following theorem.

Theorem 2. For any initial guess $w^{(0)} \in \mathbb{R}^{n^2}$, $(z^{(k)}, w^{(k)})$ generated by Algorithm 1 converges to a minimizer of the objective function $\mathcal{F}(z, w)$ as $k \rightarrow \infty$.

Proof. It follows from the alternating iterative process in Algorithm 1 that

$$\mathcal{F}(z^{(k)}, w^{(k)}) \geq \mathcal{F}(z^{(k+1)}, w^{(k)}) \geq \mathcal{F}(z^{(k+1)}, w^{(k+1)}). \quad (27)$$

It is obvious that the sequence $\mathcal{F}(z^{(k)}, w^{(k)})$ is nonincreasing. We note that it is bounded from below by the minimum value \mathcal{F}^* . Let $\theta_1(z) = \sum_1^{n^2} (z_i + ge^{-z_i})$, $\theta_2(w) = \|w\|_{\text{HTV}}$ and $D_\Phi(z, w) = \alpha_1 \|z - w\|_2^2$. Since θ_1 and θ_2 are convex, Φ is convex and differentiable and D_Φ is the associated Bregman distance, we know from [34, 35] that $\mathcal{F}(z, w) = \theta_1(z) + \theta_2(w) + D_\Phi(z, w)$ has the five-point property. Hence, $\mathcal{F}(z^{(k)}, w^{(k)})$ thus converges in \mathbb{R} to \mathcal{F}^* . So we have

$$\lim_{k \rightarrow \infty} \mathcal{F}(z^{(k)}, w^{(k)}) = \mathcal{F}^*. \quad (28)$$

We note that if the objective function $\mathcal{F}(z, w)$ is coercive, the sequence $(z^{(k)}, w^{(k)})$ is bounded since the sequence $\mathcal{F}(z^{(k)}, w^{(k)})$ converges.

Now we need to show that the objective function $\mathcal{F}(f, u, v)$ is coercive. Let D_{xx} , D_{xy} , D_{yx} , and D_{yy} be the second order difference matrices in the xx , xy , yx , and yy direction, respectively, and

$$S = \begin{pmatrix} D_{xx} \\ D_{xy} \\ D_{yx} \\ D_{yy} \end{pmatrix}. \quad (29)$$

It is not difficult to obtain the lower bound of the discrete high-order total variation as follows:

$$\begin{aligned} \|w\|_{\text{HTV}} &= \sum_{1 \leq j, k \leq n} \sqrt{((w_{xx})_{j,k})^2 + ((w_{xy})_{j,k})^2 + ((w_{yx})_{j,k})^2 + ((w_{yy})_{j,k})^2} \\ &\geq \frac{1}{2} \sum_{1 \leq j, k \leq n} \left(|(w_{xx})_{j,k}| + |(w_{xy})_{j,k}| + |(w_{yx})_{j,k}| + |(w_{yy})_{j,k}| \right) \\ &= \frac{1}{2} \|Sw\|_1. \end{aligned} \quad (30)$$

Let $\mathcal{F}_1(z, w) = \alpha_1 \|z - w\|_2^2$ and $\mathcal{F}_2(z, w) = (\alpha_2/2) \|Sw\|_1$. Denote

$$\Omega = \{(z, w) \mid \mathcal{F}_1(z, w) + \mathcal{F}_2(z, w) = 0\}. \quad (31)$$

If $(z, w) \in \Omega$, we obtain $z = w$. As $\|(z, w)\|_2 \rightarrow \infty$, it is easy to show that

$$\mathcal{F}(z, w) = \sum_i^{n^2} (z_i + ge^{-z_i}) \rightarrow \infty. \quad (32)$$

On the other hand, we obtain

$$\sum_i^{n^2} (z_i + ge^{-z_i}) \geq \sum_i^{n^2} (\log g_i + 1) \quad (33)$$

since it is a strictly convex function. By using the above inequality, we have

$$\mathcal{F}(z, w) \geq \sum_i^{n^2} (\log g_i + 1) + \mathcal{F}_1(z, w) + \mathcal{F}_2(z, w). \quad (34)$$

Hence, if $(z, w) \notin \Omega$ with $\|(z, w)\|_2 \rightarrow \infty$, it is not difficult to obtain that $\mathcal{F}_1(z, w) + \mathcal{F}_2(z, w) \rightarrow \infty$. Thus, we get that $\mathcal{F}(z, w)$ also tends to infinity. It follows from the definition of coercive that $\mathcal{F}(z, w)$ is coercive.

Since the sequence $(z^{(k)}, w^{(k)})$ is bounded, we can thus extract a convergent subsequence $\{(z^{(k_i)}, w^{(k_i)})\}$ from $\{(z^{(k)}, w^{(k)})\}$ such that $\lim_{i \rightarrow \infty} \{(z^{(k_i)}, w^{(k_i)})\} = (\tilde{z}, \tilde{w})$. Then

we obtain $m = \mathcal{F}(\hat{z}, \hat{w})$. Moreover, we have, for all $k_i \in N$ and all z ,

$$J(z, w^{(k_i)}) \geq \mathcal{F}(z^{(k_i+1)}, z^{(k_i)}), \quad (35)$$

and for all $k_i \in \mathbb{N}$ and all w ,

$$\mathcal{F}(z^{(k_i+1)}, w) \geq \mathcal{F}(z^{(k_i+1)}, w^{(k_i+1)}). \quad (36)$$

Let us denote by \bar{z} a cluster point of $\{z^{(k_i+1)}\}$. We may immediately obtain from (27) that

$$\mathcal{F}^* = \mathcal{F}(\bar{z}, \hat{w}) = \mathcal{F}(\bar{z}, \hat{w}). \quad (37)$$

Thus, we have that $\sum_i^{n^2} (\bar{z}_i + ge^{-\bar{z}_i}) + \alpha_1 \|\bar{z} - \hat{w}\|_2^2 = \sum_i^{n^2} (\bar{z}_i + ge^{-\bar{z}_i}) + \alpha_1 \|\bar{z} - \hat{w}\|_2^2$. We conclude that $\bar{z} = \hat{z}$ since \bar{z} and \hat{z} are the minimizers of the strictly convex function $\mathcal{S}_h(\hat{w})$ defined in (23). Hence, $\lim_{k_i \rightarrow \infty} z^{(k_i+1)} = \hat{z}$. Following a similar analysis, we can show that $\lim_{k_i \rightarrow \infty} w^{(k_i+1)} = \hat{w}$. By passing to the limit in (35) and (36), we have

$$\begin{aligned} \mathcal{F}(z, \hat{w}) &\geq \mathcal{F}(\bar{z}, \hat{w}), \quad \forall z, \\ \mathcal{F}(\hat{z}, w) &\geq \mathcal{F}(\hat{z}, \hat{w}), \quad \forall w. \end{aligned} \quad (38)$$

We know that (38) can be rewritten as follows:

$$\begin{aligned} \mathcal{F}(\hat{z}, \hat{w}) &= \min_z \mathcal{F}(z, \hat{w}), \\ \mathcal{F}(\hat{z}, \hat{w}) &= \min_w \mathcal{F}(\hat{z}, w). \end{aligned} \quad (39)$$

From the definition of $\mathcal{F}(z, w)$, (39) is equivalent to the following:

$$\begin{aligned} 0 &\in \psi(\hat{z}, \hat{w}), \\ 0 &\in \phi(\hat{z}, \hat{w}), \end{aligned} \quad (40)$$

where $\psi(z, w) = 1 - ge^{-z} + 2\alpha_1(z - w)$ and $\phi(z, w) = 2\alpha_1(z - w) + \alpha_2 \partial(\|w\|_{\text{HTV}})$. The subdifferential of $\mathcal{F}(z, w)$ at (\hat{z}, \hat{w}) is given by

$$\partial \mathcal{F}(\hat{z}, \hat{w}) = \begin{pmatrix} \psi(\hat{z}, \hat{w}) \\ \phi(\hat{z}, \hat{w}) \end{pmatrix}. \quad (41)$$

Therefore, we have

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \partial \mathcal{F}(\hat{z}, \hat{w}), \quad (42)$$

which implies that (\hat{z}, \hat{w}) is the minimizer of $\mathcal{F}(z, w)$. Hence the whole sequence $\mathcal{F}(z^{(k)}, w^{(k)})$ converges towards \mathcal{F}^* , the unique minimum of $\mathcal{F}(z, w)$. We deduce that the sequence $(z^{(k)}, w^{(k)})$ converges to (\hat{z}, \hat{w}) , the minimizer of $\mathcal{F}(z, w)$ as k tends to infinity. \square

5. Numerical Experiments

In this section we provide numerical results from multiplicative noise removal problem to demonstrate the performance

of the proposed regularization method. All computations of the present paper were carried out in Matlab 7.10. The results were obtained by running the Matlab codes on an Intel Core i3 CPU (2.27 GHz, 2.27 GHz) computer with RAM of 2048 M. The initial guess is chosen to be the noisy image in all tests.

The quality of the restoration results with different methods is compared quantitatively by using the Peak-Signal-to-Noise Ratio (PSNR), the relative error (ReErr) and Structural SIMilarity index (SSIM). They are defined as follows:

$$\text{PSNR} = 20 \log_{10} \left(\frac{255n}{\|u^{(k)} - u_{\text{true}}\|_2} \right), \quad (43)$$

$$\text{ReErr} = \frac{\|u^{(k)} - u_{\text{true}}\|_2}{\|u_{\text{true}}\|_2},$$

where u_{true} and $u^{(k)}$ are the ideal image and the restored image of order n respectively, and

$$\text{SSIM} = \frac{(2\mu_u \mu_{\tilde{u}} + C_1)(2\sigma_{u\tilde{u}} + C_2)}{(\mu_u^2 + \mu_{\tilde{u}}^2 + C_1)(\sigma_u^2 + \sigma_{\tilde{u}}^2 + C_2)}, \quad (44)$$

where μ_u and $\mu_{\tilde{u}}$ are averages of u and \tilde{u} , respectively. σ_u and $\sigma_{\tilde{u}}$ are the variance of u and \tilde{u} , respectively. $\sigma_{u\tilde{u}}$ is the covariance of u and \tilde{u} . The positive constants C_1 and C_2 can be thought of as stabilizing constants for near-zero denominator values. In general, a high PSNR-value or a small ReErr-value indicates that the restoration is more accurate. The SSIM is a well-known quality metric used to measure the similarity between two images. This method developed by Wang et al. [36] is based on three specific statistical measures that are much closer to how the human eye perceives differences between images. The whiter SSIM map, the closer two images are. Therefore, we use the SSIM map to reveal areas of high or low similarity between two images in this work.

We note that a straightforward high-order total variation regularization scheme for multiplicative noise removal is to solve the following optimization problem:

$$\min_u \int_{\Omega} \left(\log u + \frac{g}{u} \right) dx dy + \alpha \int_{\Omega} |D^2 u| dx dy, \quad (45)$$

where α is a positive regularization parameter which measures the trade-off between a good fit and a regularized solution. The Euler-Lagrange equation for minimization problem (45) amounts to solve (omitting the boundary conditions):

$$\begin{aligned} \frac{u-g}{\alpha u^2} + \left(\frac{u_{xx}}{|D^2 u|} \right)_{xx} + \left(\frac{u_{xy}}{|D^2 u|} \right)_{yx} + \left(\frac{u_{yx}}{|D^2 u|} \right)_{xy} \\ + \left(\frac{u_{yy}}{|D^2 u|} \right)_{yy} = 0. \end{aligned} \quad (46)$$

Hence, the time-marching (TM) algorithm solving the high-order total variation model is as follows:

$$\begin{aligned} u_t = \frac{g-u}{\alpha u^2} - \left(\frac{u_{xx}}{|D^2 u| + \epsilon} \right)_{xx} - \left(\frac{u_{xy}}{|D^2 u| + \epsilon} \right)_{yx} \\ - \left(\frac{u_{yx}}{|D^2 u| + \epsilon} \right)_{xy} - \left(\frac{u_{yy}}{|D^2 u| + \epsilon} \right)_{yy}, \end{aligned} \quad (47)$$

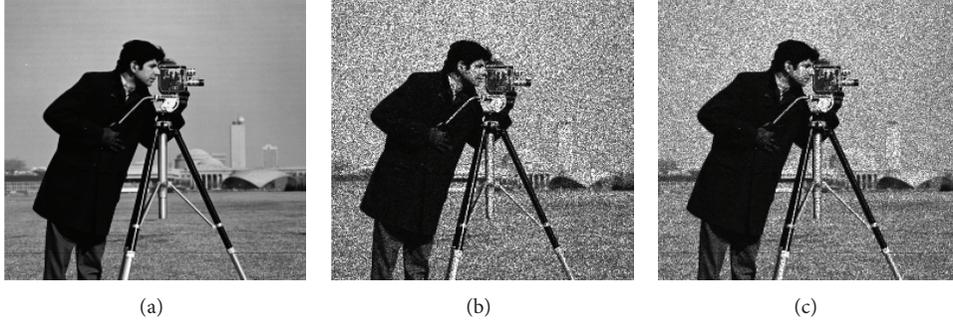


FIGURE 1: The related images in Test 1. (a) Original image. (b) Noisy image with $L = 10$. (c) Noisy image with $L = 20$.

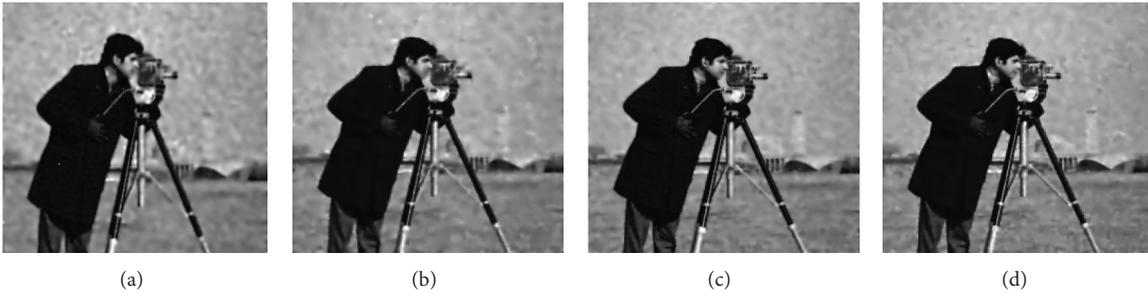


FIGURE 2: Restored results for Test 1. (a) Restored image by the TM algorithm for $L = 10$. (b) Restored image by the high-order method for $L = 10$. (c) Restored image by the TM algorithm for $L = 20$. (d) Restored image by the high-order method for $L = 20$.

where the parameter $\epsilon = 10^{-5}$ is introduced to avoid divisions by zero.

In the first test, we compare our method with the straightforward high-order total variation regularization scheme. In the time-marching algorithm for solving the straightforward high-order total variation regularization problem, we set the step size as 0.1 in order to obtain a stable iterative procedure. The algorithm is stopped when the maximum number of iterations is reached. In addition, we carry out many experiments with different α -values in the model (45), the ones with the best results are presented in this work. For the proposed high-order total variation method (HighTV method), we terminate the iterations for the method and accept $u^{(k)}$ as the computed approximation of the ideal image as soon as the maximum number of allowed outer iterations has been carried out or the relative differences between consecutive iterates $u^{(1)}, u^{(2)}, u^{(3)}, \dots$ satisfy

$$\frac{\|u^{(k+1)} - u^{(k)}\|_2}{\|u^{(k+1)}\|_2} < 10^{-4}. \quad (48)$$

In the proposed HighTV regularization method, there are two regularization parameters. We know that the quality of the restored image is highly depended on the regularization parameters. Similar to [19], we fix $\alpha_1 = 19$ in the proposed method in order to reduce the computational time in search for good regularization parameters. We determine the best value of α_2 such that the relative error of such a restored image with respect to such an ideal image is the smallest.

In the first experiment, we consider the ‘‘Cameraman’’ image of size 256×256 . The original image in Figure 1(a) is distorted by the Gamma noise with $L = 10$ and $L = 20$. The noisy images are shown in Figures 1(b) and 1(c). It is seen from Figures 1(b) and 1(c) that the smaller L is, the more noisy the observed images are. When $L = 10$, the relative error (ReErr) between the noisy image and the original image is 0.3150. When $L = 20$, the ReErr between the noisy image and the original image is 0.2234.

In Figures 2(a)–2(d), we show the restoration results of the two different methods for $L = 10$ and $L = 20$. According to the figures, we see that the proposed algorithm can provide better visual quality of restored images than those by the time-marching algorithm. In Figures 3(a) and 3(b), we compare the speed of convergence of the two algorithms for $L = 10$ and $L = 20$. In the figures, the x -axis is the number of iterations and the y -axis is the the relative error between the restored image and the original image. We see that the speed of convergence of the proposed algorithm is faster and the relative error of the proposed algorithm is smaller. In Table 1, we compare the results of the computational time and the the relative error when the best restored images are obtained. From Table 1, we observe that the computational time required by our method is lesser than that of the time-marching algorithm.

In the following, we compare the proposed HighTV method with the one proposed in [18] (AA method), the one proposed in [19] (HNW method) and the one proposed in [20] (MIDAL method). For the AA method, we use the time-marching algorithm with $\epsilon = 10^{-4}$ to solve the minimization

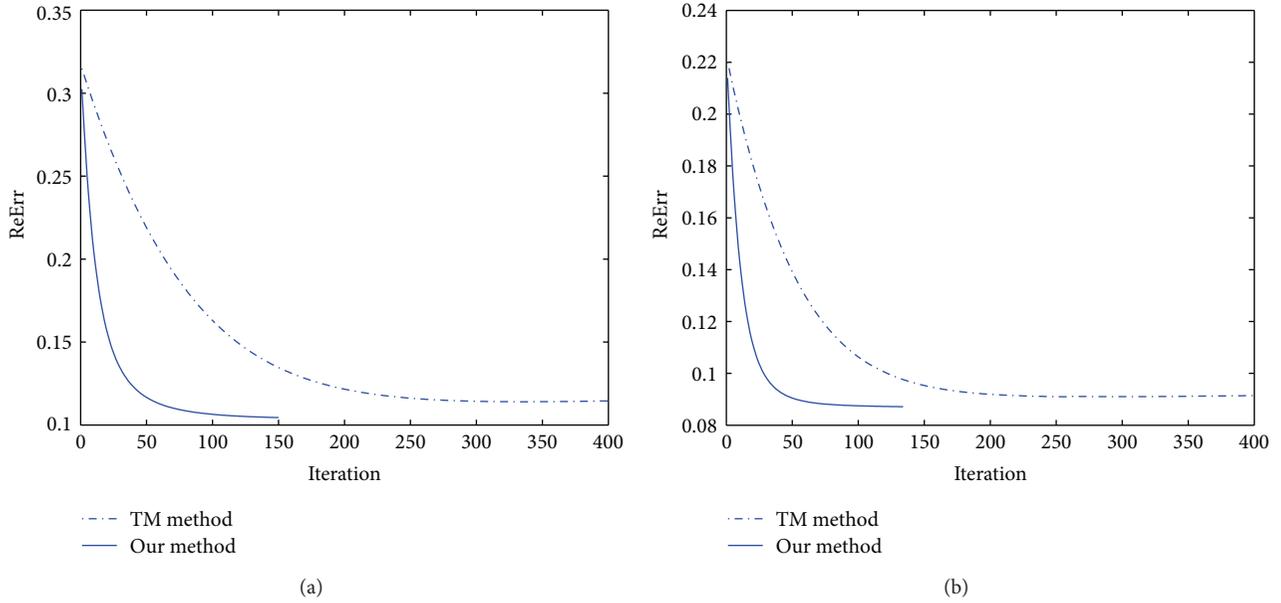


FIGURE 3: Comparison of the TM algorithm and the proposed method for Test 1. (a) $L = 10$. (b) $L = 20$.

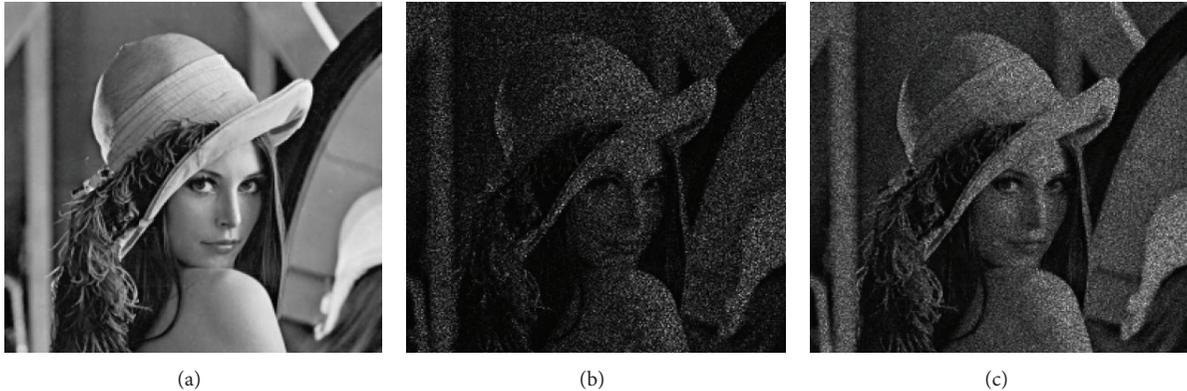


FIGURE 4: The related images in Test 2. (a) Original image. (b) Noisy image with $L = 3$. (c) Noisy image with $L = 13$.

TABLE 1: Restoration results with TM and High TV methods for Test 1.

	Method	Parameter	CPU time	ReErr
$L = 10$	TM	$\lambda = 240$	12.1623	0.1150
	HighTV	$\alpha_2 = 0.005$	5.9285	0.1039
$L = 20$	TM	$\alpha = 800$	12.0281	0.0921
	HighTV	$\alpha_2 = 0.003$	4.6362	0.0871

models as proposed in [18]. Similarly, we set the step size as 0.1 in order to obtain a stable iterative procedure. The algorithm is stopped when the maximum number of iterations is reached. In addition, we carry out many experiments with different α -values in the model (3), the ones with the best results are presented in this work. For the HNW and MIDAL methods, we use the same stopping rule as the HighTV method.

For the MIDAL method, the authors have verified experimentally that setting two parameter be equal yields good results. Therefore, we only select α by searching for the value leading to the smallest relative error of such a restored image with respect to the ideal image. In the HNW method, there are two regularization parameters. In [19], the authors fix $\alpha_1 = 19$ in all tests in order to reduce the computational time in search for good regularization parameters. We use the best value of α_2 such that the relative error of such a restored image with respect to such an ideal image is the smallest. In this way, we have a fair comparison since we compare about the best restorations for four different methods.

The aim of the second test is to give evidence of the effectiveness of employing the propose algorithm over the other three algorithms for multiplicative noise removal problems. Especially, we show that the proposed method can reduce the staircase effect. In this test, we consider the well-known ‘‘Lena’’ image of size 256×256 . The original image in



FIGURE 5: Restored results by different methods for Test 2. (a) Restored image by the AA method for $L = 3$. (b) Restored image by the HNW method for $L = 3$. (c) Restored image by the MIDAL method for $L = 3$. (d) Restored image by our method for $L = 3$. (e) Restored image by the AA method for $L = 13$. (f) Restored image by the HNW method for $L = 13$. (g) Restored image by the MIDAL method for $L = 13$. (h) Restored image by our method for $L = 13$.

TABLE 2: Restoration results with four different methods for Test 2 and Test 3.

Image	Method	Parameter	PSNR	ReErr
Lena ($L = 3$)	AA	$\lambda = 15$	22.96	0.1625
	HNW	$\alpha_2 = 0.02$	23.39	0.1543
	MIDAL	$\alpha = 2$	23.85	0.1451
	HighTV	$\alpha_2 = 0.012$	24.11	0.1425
Lena ($L = 13$)	AA	$\lambda = 300$	26.97	0.1022
	HNW	$\alpha_2 = 0.008$	27.14	0.0992
	MIDAL	$\alpha = 4$	27.31	0.0979
	HighTV	$\alpha_2 = 0.004$	27.62	0.0951
Barche ($L = 6$)	AA	$\lambda = 80$	23.22	0.1271
	HNW	$\alpha_2 = 0.012$	23.69	0.1223
	MIDAL	$\alpha = 3$	23.70	0.1221
	HighTV	$\alpha_2 = 0.008$	23.88	0.1187
Barche ($L = 16$)	AA	$\lambda = 15$	25.14	0.1024
	HNW	$\alpha_2 = 0.007$	25.67	0.0973
	MIDAL	$\alpha = 4.5$	25.66	0.0974
	HighTV	$\alpha_2 = 0.0035$	26.01	0.0939

Figure 4(a) is distorted by the Gamma noise with $L = 3$ and $L = 13$. The noisy images are shown in Figures 4(b) and 4(c). When $L = 3$, the PSNR and the relative error (ReErr) between the noisy image and the original image are 11.95 dB and 0.5768. When $L = 13$, the PSNR and the ReErr between the noisy image and the original image are 18.37 dB and 0.2765.

The restored images by the AA method, the HNW method, the MIDAL method and the proposed method are shown in Figures 5(a)–5(h). From these figures, compared

with the AA method, the HNW method and the MIDAL method, the proposed approach yields better results in image restoration since it avoids the staircase effect of the general total variation methods while at same time preserving edges as well as the total variation methods. In Figure 6, we have enlarged some details of the eight restored images. It is clear that the models involving the general total variation transform smooth regions into piecewise constant regions. As it is seen in the zoomed parts, the proposed method outperforms another three methods since the proposed model process smooth regions better than another three models. For the comparison of the performance quantitatively and the computational efficiency, in Table 2, we give their restoration results in SNRs and ReErrs. We observe from Table 2 that both the SNR and ReErr values of the restored images by the proposed method are better than those by the AA method, HNW method and the MIDAL method.

In the third test, the “barche” image is considered. The original image in Figure 7(a) is distorted by the Gamma noise with $L = 6$ and $L = 16$, see Figures 7(b) and 7(c). When $L = 6$, the PSNR and the ReErr between the noisy image and the original image are 13.20 dB and 0.4090. When $L = 16$, the PSNR and the ReErr between the noisy image and the original image are 17.44 dB and 0.2514.

The information of restored images by the AA method, the HNW method, the MIDAL method and the proposed method are displayed in Figures 8(a)–8(h). From the visual quality of restored images, the proposed regularization method is quite competitive with the other three algorithms. In addition, in Figures 9(a)–9(h) we show the SSIM maps of the restored images. It is not difficult to observe that the SSIM maps by the proposed method are whiter than the

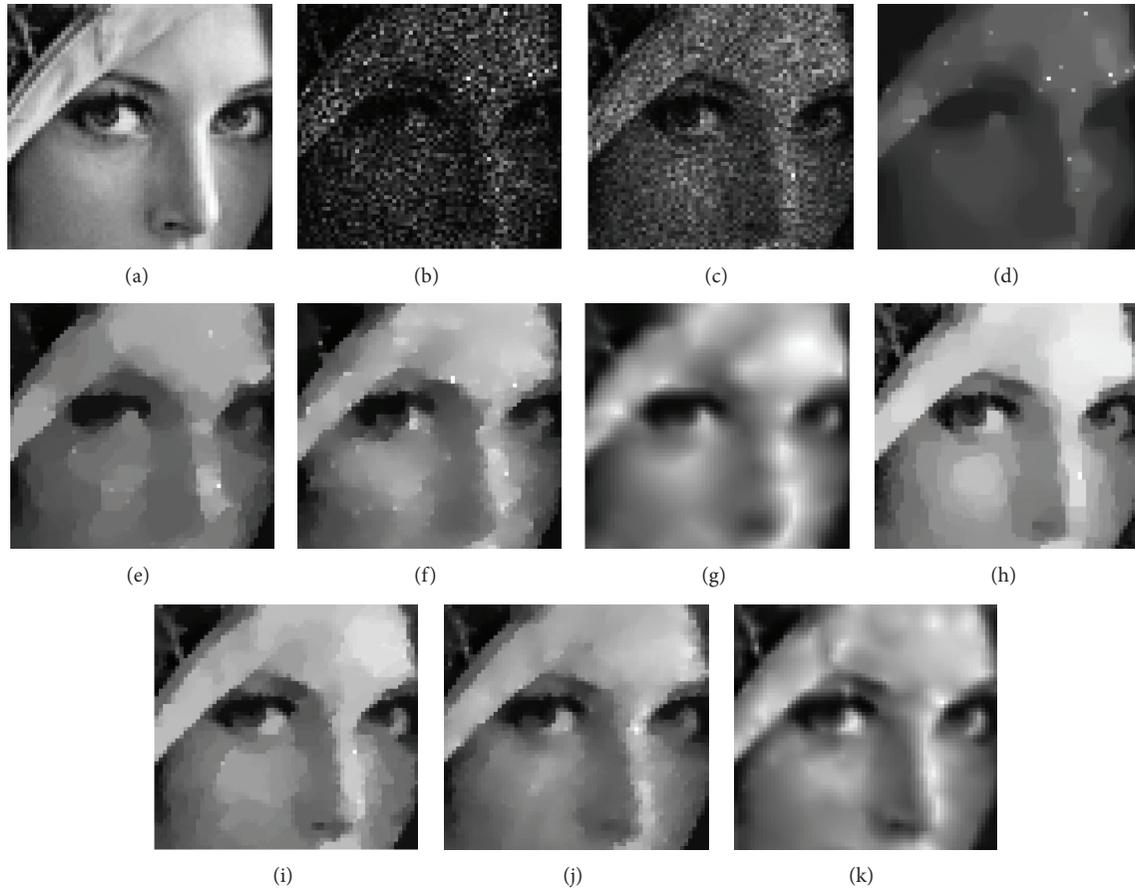


FIGURE 6: A small portion of Lena image is emphasized to show the difference of the four methods for Test 2. (a) Original image. (b) Noisy image with $L = 3$. (c) Noisy image with $L = 13$. (d) Restored image by the AA method for $L = 3$. (e) Restored image by the HNW method for $L = 3$. (f) Restored image by the MIDAL method for $L = 3$. (g) Restored image by our method for $L = 3$. (h) Restored image by the AA method for $L = 13$. (i) Restored image by the HNW method for $L = 13$. (j) Restored image by the MIDAL method for $L = 13$. (k) Restored image by our method for $L = 13$.

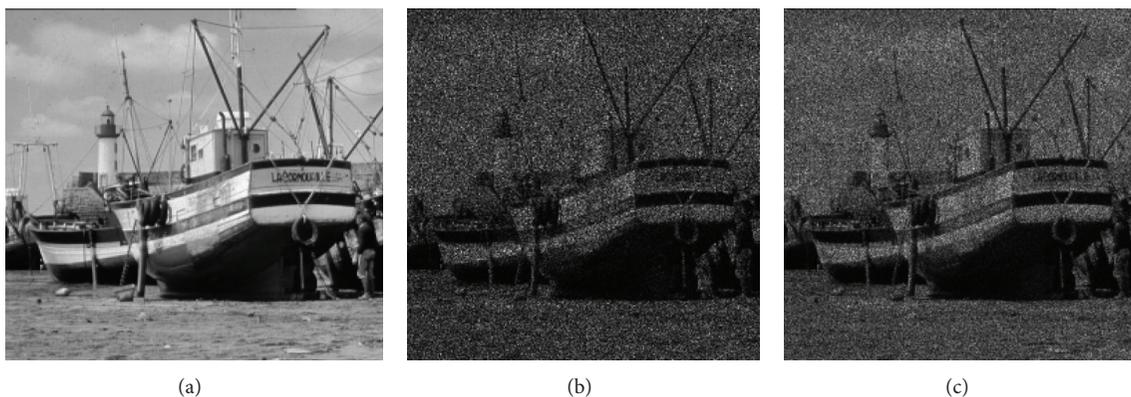


FIGURE 7: The related images in Test 3. (a) Original image. (b) Noisy image with $L = 6$. (c) Noisy image with $L = 16$.

AA method, the HNW method and the MIDAL method, which means that the proposed method behaves slightly better. The restoration results of these four different methods in SNRs, ReErrs are also presented in Table 2. It is shown

from Table 2 that both the SNR and ReErr values of the restored images by the proposed method are better than those by the AA method, the HNW method and the MIDAL method.

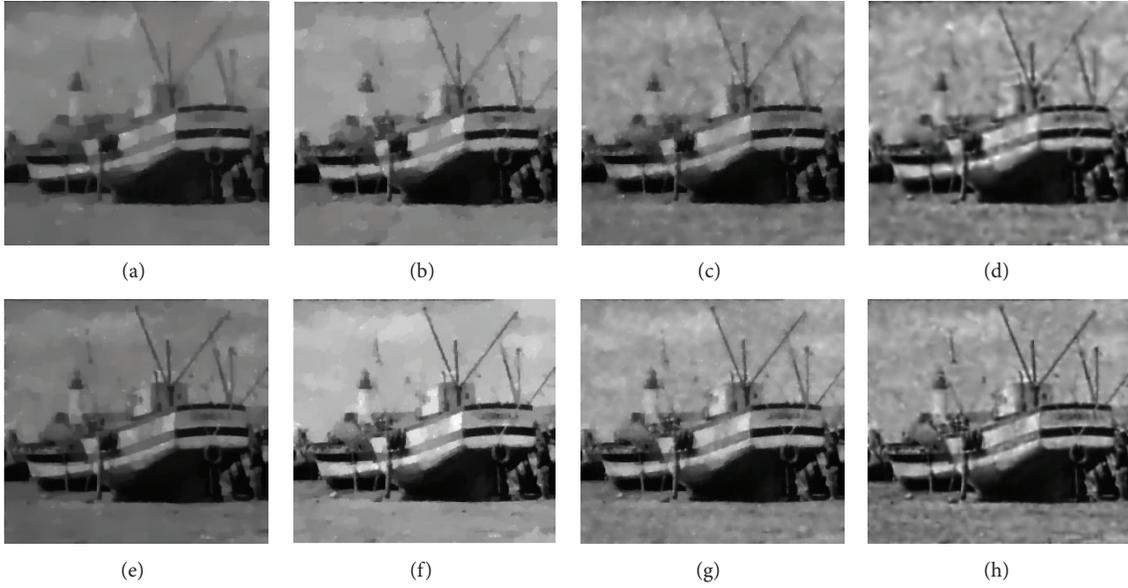


FIGURE 8: Restored results by different methods for Test 3. (a) Restored image by the AA method for $L = 6$. (b) Restored image by the HNW method for $L = 6$. (c) Restored image by the MIDAL method for $L = 6$. (d) Restored image by our method for $L = 6$. (e) Restored image by the AA method for $L = 16$. (f) Restored image by the HNW method for $L = 16$. (g) Restored image by the MIDAL method for $L = 16$. (h) Restored image by our method for $L = 16$.

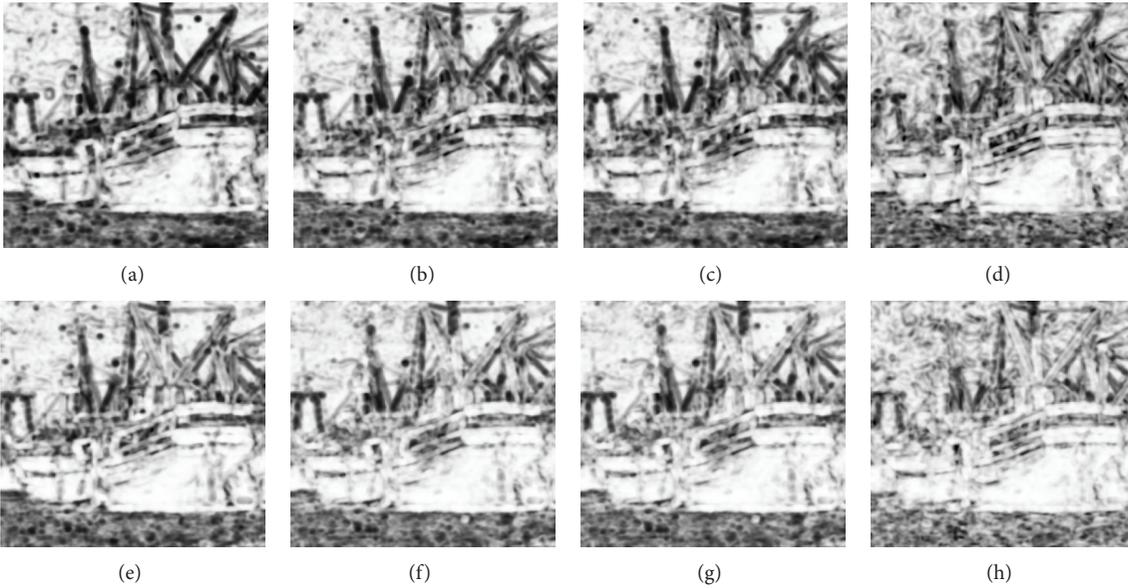


FIGURE 9: Comparison of the SSIM maps for Test 3. (a) The SSIM map by the AA method for $L = 6$. (b) The SSIM map by the HNW method for $L = 6$. (c) The SSIM map by the MIDAL method for $L = 6$. (d) The SSIM map by our method for $L = 6$. (e) The SSIM map by the AA method for $L = 16$. (f) The SSIM map by the HNW method for $L = 16$. (g) The SSIM map by the MIDAL method for $L = 16$. (h) The SSIM map by our method for $L = 16$.

6. Conclusions

In this paper, we investigate a fast and efficient way to restore blurred and noisy images with a high-order total variation minimization technique. An alternating minimization algorithm is employed to solve the high-order total

variation model. The edges in the restored image can be preserved quite well and the staircase effect is reduced effectively in the proposed algorithm. Our experimental results show that the performance of the proposed method is superior to that of some existing total variation restoration methods.

Acknowledgments

This research is supported by NSFC (10871034, 61170311), Sichuan Province Sci. & Tech. Research Project (2011JY0002, 12ZC1802), Jiangsu Province Research Project (12KJD110001) and Research Grant KQ12013 from HuaiHai Institute of Technology.

References

- [1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.
- [2] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.
- [3] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Publishing, London, UK, 1998.
- [4] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, Pa, USA, 2005.
- [5] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.
- [6] T. F. Chan, S. Esedoglu, and F. E. Park, "Image decomposition combining staircase reduction and texture extraction," *Journal of Visual Communication and Image Representation*, vol. 18, no. 6, pp. 464–486, 2007.
- [7] S. Hu, Z. Liao, and W. Chen, "Image denoising based on dilated singularity prior," *Mathematical Problems in Engineering*, vol. 2012, Article ID 767613, 15 pages, 2012.
- [8] S. Chountasis, V. N. Katsikis, and D. Pappas, "Digital image reconstruction in the spectral domain utilizing the Moore-Penrose inverse," *Mathematical Problems in Engineering*, vol. 2010, Article ID 750352, 14 pages, 2010.
- [9] S. Becker, J. Bobin, and E. J. Candès, "NESTA: a fast and accurate first-order method for sparse recovery," *SIAM Journal on Imaging Sciences*, vol. 4, no. 1, pp. 1–39, 2011.
- [10] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.
- [11] Y. L. Wang, J. F. Yang, W. T. Yin, and Y. Zhang, "A new alternating minimization algorithm for total variation image reconstruction," *SIAM Journal on Imaging Sciences*, vol. 1, no. 3, pp. 248–272, 2008.
- [12] L. Xiao, L. L. Huang, and Z. H. Wei, "A weberized total variation regularization-based image multiplicative noise removal algorithm," *Eurasip Journal on Advances in Signal Processing*, vol. 2010, Article ID 490384, 2010.
- [13] S. Durand, J. Fadili, and M. Nikolova, "Multiplicative noise removal using L1 fidelity on frame coefficients," *Journal of Mathematical Imaging and Vision*, vol. 36, no. 3, pp. 201–226, 2010.
- [14] P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.
- [15] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*, SciTech Publishing, 2004.
- [16] J. Goodman, "Some fundamental properties of speckle," *Journal of the Optical Society of America*, vol. 66, pp. 1145–1150, 1976.
- [17] L. Rudin, P.-L. Lions, and S. Osher, "Multiplicative denoising and deblurring: theory and algorithms," in *Geometric Level Sets in Imaging, Vision and Graphics*, S. Osher and N. Paragios, Eds., pp. 103–119, Springer, New York, NY, USA, 2003.
- [18] G. Aubert and J.-F. Aujol, "A variational approach to removing multiplicative noise," *SIAM Journal on Applied Mathematics*, vol. 68, no. 4, pp. 925–946, 2008.
- [19] Y.-M. Huang, M. K. Ng, and Y.-W. Wen, "A new total variation method for multiplicative noise removal," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 20–40, 2009.
- [20] J. M. Bioucas-Dias and M. A. T. Figueiredo, "Multiplicative noise removal using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1720–1730, 2010.
- [21] G. Steidl and T. Teuber, "Removing multiplicative noise by Douglas-Rachford splitting methods," *Journal of Mathematical Imaging and Vision*, vol. 36, no. 2, pp. 168–184, 2010.
- [22] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 503–516, 2000.
- [23] M. Lysaker, A. Lundervold, and X. C. Tai, "Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time," *IEEE Transactions on Image Processing*, vol. 12, no. 12, pp. 1579–1590, 2003.
- [24] G. Steidl, "A note on the dual treatment of higher-order regularization functionals," *Computing*, vol. 76, no. 1-2, pp. 135–148, 2006.
- [25] B. Chen, J.-L. Cai, W.-S. Chen, and Y. Li, "A multiplicative noise removal approach based on partial differential equation model," *Mathematical Problems in Engineering*, vol. 2012, Article ID 242043, 14 pages, 2012.
- [26] J. Zhang, Z. H. Wei, and L. Xiao, "Adaptive fractional-order multi-scale method for image denoising," *Journal of Mathematical Imaging and Vision*, vol. 43, no. 1, pp. 39–49, 2012.
- [27] S. Lefkimmiatis, A. Bourquard, and M. Unser, "Hessian-based norm regularization for image restoration with biomedical applications," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 983–995, 2012.
- [28] H.-Z. Chen, J.-P. Song, and X.-C. Tai, "A dual algorithm for minimization of the LLT model," *Advances in Computational Mathematics*, vol. 31, no. 1–3, pp. 115–130, 2009.
- [29] Y. Hu and M. Jacob, "Higher degree total variation (HDTV) regularization for image recovery," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2559–2571, 2012.
- [30] F. Li, C. M. Shen, J. S. Fan, and C. L. Shen, "Image restoration combining a total variational filter and a fourth-order filter," *Journal of Visual Communication and Image Representation*, vol. 18, no. 4, pp. 322–330, 2007.
- [31] M. Lysaker and X. C. Tai, "Iterative image restoration combining total variation minimization and a second-order functional," *International Journal of Computer Vision*, vol. 66, no. 1, pp. 5–18, 2006.
- [32] K. Papafitsoros and C. B. Schönlieb, "A combined first and second order variational approach for image reconstruction," <http://arxiv.org/abs/1202.6341>.
- [33] D. P. Bertsekas, A. Nedic, and E. Ozdaglar, *Convex Analysis and Optimization*, Athena Scientific, Belmont, Mass, USA, 2003.
- [34] C. L. Byrne, "Alternating minimization as sequential unconstrained minimization: a survey," *Journal of Optimization Theory and Applications*, vol. 156, no. 3, pp. 554–566, 2012.

- [35] H. H. Bauschke, P. L. Combettes, and D. Noll, "Joint minimization with alternating Bregman proximity operators," *Pacific Journal of Optimization*, vol. 2, no. 3, pp. 401–424, 2006.
- [36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

Research Article

Conditional Nonlinear Optimal Perturbation of a Coupled Lorenz Model

Lili Yang

Beijing Jiaotong Vocational Technical College, Deng Zhuang, Changping District, Beijing 102200, China

Correspondence should be addressed to Lili Yang; yanglili-1974@hotmail.com

Received 11 October 2012; Revised 7 February 2013; Accepted 18 February 2013

Academic Editor: Fatih Yaman

Copyright © 2013 Lili Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The conditional nonlinear optimal perturbation (CNOP) technique is a useful tool for studying the limits of predictability in numerical weather forecasting and climate predictions. The CNOP is the optimal combined mode of the initial and model parameter perturbations that induce the largest departure from a given reference state. The CNOP has two special cases: the CNOP-I is linked to initial perturbations and has the largest nonlinear evolution at the time of prediction, while the other case, CNOP-P, is related to the parameter perturbations that cause the largest departure from a given reference state at a given future time. Solving the CNOPs of a numerical model is a mathematical problem. In this paper, we calculate the CNOP, CNOP-I, and CNOP-P of a coupled Lorenz model and study the properties of these CNOPs. We find that the CNOP, CNOP-I, and CNOP-P always locate the boundary of their respective constraints. This property is also demonstrated analytically for the model whose solutions depend continuously on the initial and parameter perturbations, which provides a theoretical basis for testing the accountability of the numerically computed CNOPs. In addition, we analyze the features of the CNOPs for the coupled Lorenz model and explain their structures.

1. Introduction

Error growth dynamics is one of the fundamental problems in the study of numerical weather forecasting and climate prediction. The prediction errors are generally caused by initial and model errors. To study the role of initial and model errors in yielding prediction errors, Lorenz [1] classified two types of predictability problems: one is related to the initial error, assuming a perfect model, and is referred to as the first type of predictability; the other is associated with model errors, assuming a perfect initial field, and consists of the second type of predictability. The former has been largely investigated, and many theories and methods have been proposed or introduced [2–4] in which a singular vector (SV; [2, 5]) is often used. However, SVs are the result of a linearized model and do not consider the effect of nonlinearity. Considering the limitation of the linear theory of SV, Mu et al. [4] proposed the approach of conditional nonlinear optimal perturbation (CNOP) to search for the optimal initial perturbation (denoted by “CNOP-I”) in a given constraint. The CNOP-I considers the effect of nonlinearity and is more effective than the SV at revealing the effect of nonlinearity.

CNOP-I has been applied to predictability studies for weather and climate [6–8] and has played an important role in guiding scientists to develop and improve numerical models and even to propose innovative ideas to increase the forecasting accuracy of weather and climate models [9].

The effect of model error on predictability is related to the second type of predictability [1]. One important aspect in this field is the effect of the uncertainties in model parameters on predictability [10]. In numerical models, there are generally several empirical parameters, and they may simultaneously have uncertainties. In predictability studies, one is required to find the patterns of the parameter perturbations that have the largest effect on predictability. Furthermore, the initial and model parameters’ errors generally coexist in predictions. In this situation, one expects to know which combined mode of initial and parameter errors yields the largest effect on predictions. For this reason, Mu et al. [11] extended the CNOP-I to study parameter uncertainties and proposed the CNOP, which consists of the optimal combined mode of the initial and parametric perturbations. The CNOP induces the largest departure from a given reference state, which can be used to estimate the limit of prediction errors that is caused

by initial and model parameter errors [11]. If we only consider the effect of the initial perturbation, then the CNOP is only related to the initial perturbation, as is the CNOP-I. If we only consider the parametric perturbation, the CNOP is merely associated with the parametric perturbation and acts as the parametric perturbation that causes the largest departure from the reference state. In this case, the CNOP is denoted as CNOP-P.

To use the CNOP approach to study the limits of predictability, the CNOPs of models must be solved, which can result from a mathematical problem. CNOPs are related to several optimization problems. In this paper, we will solve the CNOP, CNOP-I, and CNOP-P of a coupled Lorenz model and study the properties of the CNOPs. In the following section, the CNOP approach is reviewed. The CNOPs of a coupled Lorenz model are explored in Section 3. Finally, a summary is presented in Section 4.

2. Conditional Nonlinear Optimal Perturbation

In this section, we follow Mu et al. [11] and Duan and Zhang [10] to review the CNOP approach. We assume the following evolution equations for the state vector, w :

$$\frac{\partial \mathbf{w}}{\partial t} + F(\mathbf{w}, \mathbf{p}) = 0, \quad \mathbf{w}|_{t=0} = \mathbf{w}_0, \quad \text{in } \Omega \times [0, \tau], \quad (1)$$

where $\mathbf{w}(\mathbf{x}, t) = (w_1(\mathbf{x}, t), w_2(\mathbf{x}, t), \dots, w_n(\mathbf{x}, t))$ and \mathbf{w}_0 is its initial state. $(\mathbf{x}, t) \in \Omega \times [0, \tau]$, Ω is a domain in R^n , $\tau < +\infty$, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and t is the time. Furthermore, $\mathbf{p} = (p_1, p_2, \dots, p_n)$ is a model parameter (vector), and F is a nonlinear operator. Assuming that the dynamic system equations (1) and the initial state are known exactly, the future state can be determined by integrating (1) with the appropriate initial condition. The solution to (1) for the state vector \mathbf{w} at time T is

$$\mathbf{w}(\mathbf{x}, T) = M_T(\mathbf{p})(\mathbf{w}_0). \quad (2)$$

Here, $M_T(\mathbf{p})$ is the propagator of (1) with the parameter vector \mathbf{p} and, as described by (2), ‘‘propagates’’ the initial value to the time T in the future.

Let $U(t)$ and $U(t) + u(t)$ be solutions to (2) with initial values U_0 and $U_0 + u_0$, respectively,

$$\begin{aligned} U(T) &= M_T(\mathbf{p})(U_0), \\ U(T) + u(T) &= M_T(\mathbf{p})(U_0 + u_0), \end{aligned} \quad (3)$$

where u_0 is the initial perturbation of a time-dependent state $U(t)$ (hereafter the ‘‘reference state’’), which describes the nonlinear evolution of this initial perturbation.

Assume that a parameter perturbation \mathbf{p}' is superimposed on a reference parameter \mathbf{P} :

$$\begin{aligned} U(T) &= M_T(\mathbf{P})(U_0), \\ U(T) + u_p(T) &= M_T(\mathbf{P} + \mathbf{p}')(U_0), \end{aligned} \quad (4)$$

where $u_p(T)$ describes the departure from the reference state $U(T)$, which is caused by \mathbf{p}' .

Now, we explore the situation with both initial and parameter perturbations in

$$\begin{aligned} U(T) &= M_T(\mathbf{P})(U_0), \\ U(T) + u_p(T) &= M_T(\mathbf{P} + \mathbf{p}')(U_0 + u_0). \end{aligned} \quad (5)$$

$u_p(T)$ is the departure from the reference state $U(T)$, which is caused by the combined mode $(u_0; \mathbf{p}')$ of the initial perturbation u_0 and the parameter perturbation \mathbf{p}' .

A nonlinear optimization problem is defined in

$$J(u_{0\delta}; \mathbf{p}'_{\sigma}) = \max_{u_0 \in C_{\delta}, \mathbf{p}' \in C_{\sigma}} J(u_0; \mathbf{p}'), \quad (6)$$

where

$$J(u_0; \mathbf{p}') = \frac{1}{2} \left\| M_T(\mathbf{P} + \mathbf{p}')(U_0 + u_0) - M_T(\mathbf{P})(U_0) \right\|^2. \quad (7)$$

$u_0 \in C_{\delta}$, and $\mathbf{p}' \in C_{\sigma}$ are, respectively, the constraining conditions of the initial and parameter perturbations, where the constraining conditions constrain the amplitude of the initial and parameter perturbations. By solving (6), one can obtain the optimal combined mode of the initial and parameter perturbations $(u_{0\delta}; \mathbf{p}'_{\sigma})$ that, for the given constraint, induces the largest departure from the reference state $U(t)$ at time T . This combined mode is the CNOP. When we consider only the initial perturbation or assume that the constraint of the parameter perturbation $\mathbf{p}' = 0$, then (6) becomes (8):

$$J_{u_0}(u_{0\delta}^I) = \max_{u_0 \in C_{\delta}} \frac{1}{2} \left\| M_T(\mathbf{P})(U_0 + u_0) - M_T(\mathbf{P})(U_0) \right\|^2. \quad (8)$$

The initial perturbation $u_{0\delta}^I$ that satisfies (8) is the CNOP-I, which has the largest nonlinear evolution at time T . It is clear that the CNOP-I is a special case of CNOP. Another special case of the CNOP is the CNOP-P. In fact, in (6), we neglect the initial perturbation, that is, the constraint of the initial perturbation is $u_0 = 0$, and we obtain

$$J_p(\mathbf{p}'_{\sigma}^p) = \max_{\mathbf{p}' \in C_{\sigma}} \frac{1}{2} \left\| M_T(\mathbf{P} + \mathbf{p}')(U_0) - M_T(\mathbf{P})(U_0) \right\|^2, \quad (9)$$

in which \mathbf{p}'_{σ}^p is the CNOP-P.

CNOP-I, CNOP-P, and CNOP are generally vectors. Their magnitudes are constrained by the constraint conditions $u_0 \in C_{\delta}$, $\mathbf{p}' \in C_{\sigma}$, and the two of $u_0 \in C_{\delta}$, $\mathbf{p}' \in C_{\sigma}$, respectively (see (6), (8), and (9)). The constraint conditions and their boundaries are generally determined by the physical problems of the investigation. There are a large number of combinations of different components of initial perturbations or/and parametric perturbation vectors that satisfy the constraint conditions. The goal of the CNOP is to determine the combination that causes the largest departure from the reference state. Therefore, in predictability studies, despite the determination of the boundaries of the constraint

conditions, the CNOP must still be calculated to explore the initial errors (or parametric errors, or the combination of the two) that cause the largest prediction errors at the time of prediction. Initial and parametric perturbations represent the initial and parametric errors in weather or climate models, respectively. Initial errors are often referred to as the initial observational errors. Observational errors are usually caused by observational instruments and observational methods and, consequently, have a finite amplitude. The amplitude of observational errors can be regarded as the amplitude of the initial errors. Therefore, initial errors are finite, and the boundaries of the constraint conditions are finite. Model parameters are used to describe the effect of certain physical processes, which, limited by our understanding and realization of physical processes, may be empirical and have uncertainties. Of course, the uncertainties in the model parameters should have finite amplitudes. Otherwise, the corresponding numerical models fail to describe the physical processes to an acceptable extent, which, as a result, cannot be regarded as forecasting models. Therefore, the boundary of the constraint conditions in (6) (i.e., $u_0 \in C_\delta$ or $p' \in C_\sigma$) should have a certain amplitude based on physics, which, mathematically, should be constrained in the neighborhood of reference states.

As mentioned in the introduction, CNOP-I is a natural generalization of LSV in a nonlinear field. Mu et al. [4] compared CNOP-I and LSV and demonstrated that, when CNOP-I is computed with a linearized model, CNOP-I is the LSV. In addition, Mu et al. [4] also demonstrated that when the CNOP-I is computed with a nonlinear model, a CNOP can be approximated by an LSV only when the initial perturbations are sufficiently small; otherwise, the CNOP cannot be approximated by the LSV (also see [12]). In this paper, we calculate not only the CNOP-I but also the CNOP-P and CNOP. The latter two are also related to parameter perturbations. When a nonlinear model is linearized, some parameters may be removed due to linearization of the nonlinear model. In particular, if we linearize the coupled Lorenz model that is adopted in this paper, the parameters α and β will be removed. Thus, we cannot consider the effect of the uncertainties of all of the parameters on the prediction results. Therefore, it is not reasonable to consider the CNOP-P and CNOP for a linearized model. To coordinate CNOP-I with CNOP-P and CNOP, we do not calculate the CNOP-I of the linearized coupled Lorenz model in this paper.

3. CNOP, CNOP-I, and CNOP-P of a Coupled Lorenz Model

In this section, we will illustrate the CNOP, CNOP-I and CNOP-P for a coupled Lorenz model (see [13]).

The coupled Lorenz model is given in

$$\begin{aligned}\dot{X} &= -\sigma X + \sigma Y + \alpha \int_0^t X(t') e^{-\beta(t-t')} dt', \\ \dot{Y} &= -XZ + rX - Y + \alpha \int_0^t Y(t') e^{-\beta(t-t')} dt', \\ \dot{Z} &= XY - bZ.\end{aligned}\quad (10)$$

In this model, there are five parameters; in the numerical experiments, these parameters are set as $\sigma = 10$, $\alpha = 0.12$, $r = 28$, $b = 8/3$, and $\beta = 0.3$ (see [13]). In this paper, the model (10) is discretized by the fourth-order Runge-Kutta scheme, with a time step $dt = 0.01$.

It is known that the Lorenz model has three steady states [14]:

$$\begin{aligned}O &: (X, Y, Z) = (0, 0, 0), \\ C_1 &: (X, Y, Z) = (-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, r-1), \\ C_2 &: (X, Y, Z) = (\sqrt{b(r-1)}, \sqrt{b(r-1)}, r-1).\end{aligned}\quad (11)$$

However, we note that these steady states in the Lorenz model may be unsteady in the coupled Lorenz model. That is, if these states are U_0 and the propagation of (10) is M_t , then there exists the possibility that $U_0 \neq M_t(U_0)$ in the coupled Lorenz model. In fact, the steady states C_1 and C_2 in the Lorenz model will be unsteady in the coupled Lorenz model, whereas the steady state O in the Lorenz model is also a steady state in the coupled Lorenz model. We can regard these states as reference states to compute the CNOP, CNOP-I, and CNOP-P of the coupled Lorenz model. O is a steady reference state, and C_1 and C_2 are unsteady in the coupled model, yielding two time-dependent reference states. For convenience, we call them Ref- O , Ref- C_1 , and Ref- C_2 , respectively.

In the objective function J in (7), we take U_0 as the three states of (X, Y, Z) and P as the five parameters $(\sigma, \alpha, r, b, \beta)$ in the coupled Lorenz model to obtain their CNOP, CNOP-I, and CNOP-P. We signify the initial perturbations as $u_0 = (\delta x_0, \delta y_0, \delta z_0)$ and the parameter perturbations as $p' = (\delta\sigma, \delta\alpha, \delta r, \delta b, \delta\beta)$. Then, the objective function $J(u_0; p') = \sqrt{\delta x(t)^2 + \delta y(t)^2 + \delta z(t)^2}$ and describes the departure from the reference state $(X(t), Y(t), Z(t))$ that is caused by the initial perturbation u_0 and the parameter perturbation p' . The conditions

$$\begin{aligned}C_\delta &= \{(\delta x_0, \delta y_0, \delta z_0) \mid \max\{|\delta x_0|, |\delta y_0|, |\delta z_0|\} \leq \varepsilon\} \\ C_\sigma &= \left\{(\delta\sigma, \delta\alpha, \delta r, \delta b, \delta\beta) \mid \max\left\{\frac{|\delta\sigma|}{\sigma}, \frac{|\delta\alpha|}{\alpha}, \frac{|\delta r|}{r}, \frac{|\delta b|}{b}, \frac{|\delta\beta|}{\beta}\right\} \leq \mu\right\}\end{aligned}\quad (12)$$

are used to constrain the magnitudes of the initial and parameter perturbations, respectively, where ε and μ are two given positive numbers. For the different values of ε and μ , we obtain different magnitudes of constraints.

To compute the CNOP, CNOP-I, and CNOP-P, we need the information from the gradient of the objective function J . Mu et al. [11] and Duan and Zhang [10] derived the formula for the objective function with respect to the initial and parametric perturbations. The gradient of the objective function

TABLE 1: CNOPs for the steady reference state O and the model parameters $\sigma = 10$, $\alpha = 0.12$, $r = 28$, $b = 8/3$, $\beta = 0.3$.

	δ_{x_0}	δ_{y_0}	δ_{z_0}	δ_σ	δ_α	δ_r	δ_b	δ_β
$\varepsilon = 0.1, \mu = 0.05$	0.100	0.100	-0.100	0.500	0.006	1.400	-0.133	-0.015
$\varepsilon = 0.2, \mu = 0.1$	0.045	0.200	-0.200	1.000	0.012	2.800	-0.267	-0.030
$\varepsilon = 0.3, \mu = 0.2$	-0.115	0.300	-0.300	2.000	0.024	5.600	-0.533	-0.060
$\varepsilon = 0.4, \mu = 0.3$	-0.242	0.400	-0.400	3.000	0.036	8.400	-0.800	-0.090

J with respect to the initial and parametric perturbations can be obtained with the following formula:

$$\frac{\partial J}{\partial u_0} = \lambda_1(0),$$

$$\frac{\partial J}{\partial p'} = \lambda_2(0) = \int_0^T \left[\frac{\partial F(U(t)) + u(t); P + p'}{\partial p'} \right]^* \lambda_1(t) dt, \quad (13)$$

where $\lambda_1(t)$ and $\lambda_2(t)$ satisfy

$$\begin{aligned} \frac{\partial \lambda_1}{\partial t} + \left[\frac{\partial F(U(t)) + u(t); P + p'}{\partial u} \right]^* \lambda_1 &= 0, \\ \frac{\partial \lambda_2}{\partial t} + \left[\frac{\partial F(U(t)) + u(t); P + p'}{\partial p'} \right]^* \lambda_1 &= 0, \end{aligned} \quad (14)$$

$$\lambda_1|_{t=T} = u(T),$$

$$\lambda_2|_{t=T} = 0.$$

Equation (14) is the adjoint of the tangent linear model from (1). By integrating (14), we can obtain the gradient for the objective function $J(u_0; p')$ with respect to u_0 and p' . With the gradient $J/\partial u_0 = \lambda_1(0)$ and $J/\partial p' = \lambda_2(0)$, the CNOP-I (CNOP-P) can be computed using optimization solvers, such as the spectral projected gradient 2 (SPG2, [15]), sequential quadratic programming (SQP, [16]), and limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS, [17]). For both gradients $J/\partial u_0 = \lambda_1(0)$ and $J/\partial p' = \lambda_2(0)$, one can calculate the CNOP of the initial perturbation and the parameter perturbation.

Equation (14) suggests that the gradient $(\partial J/\partial u_0, \partial J/\partial p')$ of the objective function J in (7) can be obtained by integrating the adjoint model (14). With this gradient information, we can obtain not only the CNOP but also the CNOP-I and the CNOP-P using an optimization solver.

Next, we investigate the CNOPs that consist of the combined mode of the initial and parameter perturbations for the given reference states Ref-O, Ref-C₁, and Ref-C₂, where the optimization time period T associated with CNOPs is equal to 50, with a time step $dt = 0.01$ (see (10)). According to (14), we construct the tangent linear model for (10) and then the adjoint model. By integrating this adjoint model, we obtain the gradient for the objective function J with respect to the initial and parameter perturbations. Next, we use this gradient to compute the CNOP, CNOP-I, and CNOP-P of the coupled Lorenz model using the SQP algorithm [16]. First, we evaluate the CNOPs for the reference state

Ref-O with the given objective function and the different amplitudes of constraints. It is demonstrated that there exist two CNOPs for the steady reference state Ref-O for each constraint. For example, when the constraints are $\varepsilon = 0.1$ and $\mu = 0.05$, both the perturbations $(u_{0\delta}; p\sigma') = (-0.100, -0.100, 0.100, 0.500, 0.006, 1.400, -0.133, -0.015)$ and $(0.100, 0.100, 0.100, 0.500, 0.006, 1.400, -0.133, -0.015)$ are the CNOPs of the reference state O . It is easily seen that the difference between the perturbations only consists of the components δx_0 and δy_0 . These two CNOPs are symmetric about the z -axis. This relationship is easily determined from (10). In fact, if $(X(t), Y(t), Z(t))$ is a solution to model (10), then $(-X(t), -Y(t), Z(t))$ is also a solution. For simplicity, we only list one of the two CNOPs in Table 1. From Table 1, it is evident that, for the different amplitudes of constraints, the CNOPs always locate the boundary of the constraint.

The reference state Ref-O is steady and, together with the z -axis, acts as a symmetric axis of the system described by the coupled Lorenz model, which induces its two symmetric CNOPs about the z -axis. Note that the reference states Ref-C₁ and Ref-C₂ are unsteady and may not be a symmetric axis, in which case symmetric CNOPs do not exist. Nevertheless, we notice that the Ref-C₁ and Ref-C₂ are symmetric about the z -axis. Are the CNOPs of these two reference states symmetric about z -axis? To address this question, we compute the CNOPs for Ref-C₁ and Ref-C₂.

With the corresponding objective function and constraint condition, we obtain the CNOPs for the reference states Ref-C₁ and -C₂. The results demonstrate that these two reference states have one CNOP; furthermore, the components δx_0 and δy_0 between these two CNOPs are inverses of each other, while the other components remain the same (i.e., these two CNOPs are also symmetric about z -axis). For simplicity, we only display the results for Ref-C₁ in Table 2.

It is clear from Tables 1 and 2 that the CNOPs of these two reference states are different, which indicates that the CNOPs are dependent on the reference states. Furthermore, it is easily shown that the CNOPs of the reference states Ref-C₁ and -C₂ always locate the boundary of the constraint. In numerical experiments, we also adopt other reference states to compute the CNOPs. Similar results are obtained. It is obvious that whether the reference states are steady or unsteady, the CNOPs of the coupled Lorenz model will always locate the boundary of the constraint.

CNOP consists of both initial and parameter perturbations. If one only perturbs the initial condition of a reference state and the remaining parameters are unperturbed, one can obtain a CNOP-I with (8). CNOP-I is one of the special cases of CNOP. For its computation, we still adopt the reference states Ref-O, Ref-C₁, and Ref-C₂ and consider which types of

TABLE 2: CNOPs for the reference state with C_1 as the initial condition and the model parameters $\sigma = 10$, $\alpha = 0.12$, $r = 28$, $b = 8/3$, $\beta = 0.3$.

	δ_{x_0}	δ_{y_0}	δ_{z_0}	δ_σ	δ_α	δ_r	δ_b	δ_β
$\varepsilon = 0.1, \mu = 0.05$	-0.100	-0.100	0.100	-0.500	-0.006	-1.400	-0.133	0.015
$\varepsilon = 0.2, \mu = 0.1$	-0.200	-0.200	0.200	-1.000	-0.012	-2.800	-0.267	0.030
$\varepsilon = 0.3, \mu = 0.2$	-0.300	-0.300	0.300	-2.000	-0.024	-5.600	-0.533	0.060
$\varepsilon = 0.4, \mu = 0.3$	-0.400	0.400	0.400	-2.606	-0.036	-8.400	-0.628	0.090

 TABLE 3: CNOP-Is for initial condition O with different initial constraints.

	δ_{x_0}	δ_{y_0}	δ_{z_0}
$\varepsilon = 0.1$	0.100	0.100	-0.100
$\varepsilon = 0.2$	0.177	0.200	-0.200
$\varepsilon = 0.3$	0.098	0.300	-0.300
$\varepsilon = 0.4$	0.018	0.400	-0.400

 TABLE 4: CNOP-Is for the reference state with initial condition C_1 .

	δ_{x_0}	δ_{y_0}	δ_{z_0}
$\varepsilon = 0.1$	-0.100	-0.100	-0.100
$\varepsilon = 0.2$	-0.200	-0.200	-0.200
$\varepsilon = 0.3$	-0.300	-0.300	-0.300
$\varepsilon = 0.4$	-0.400	-0.400	-0.400

initial perturbations in the constraint $C_\delta = \{(\delta x_0, \delta y_0, \delta z_0) \mid \max\{|\delta x_0|, |\delta y_0|, |\delta z_0|\} \leq \varepsilon\}$ cause the largest nonlinear evolution at the time of prediction. With different values of ε , we calculate the CNOP-I for the given reference states. For the steady state O , we obtain two symmetric patterns of CNOP-I due to the symmetry of the model (10). In Table 3, we list only one CNOP-I for the steady reference state O .

From Table 3, it is apparent that the CNOP-I locates the boundary of the initial constraint. This property of CNOP-I can also be demonstrated by the CNOP-I of the unsteady reference states Ref- C_1 and Ref- C_2 . To illustrate this point, we list in Table 4 the CNOP-I of the Ref- C_1 ; the Ref- C_1 and Ref- C_2 have one CNOP-I. Furthermore, due to the symmetry of (10), the CNOP-I of both Ref- C_2 and Ref- C_1 is symmetric about the z -axis in phase space. For example, when the constraint $\varepsilon = 0.1$, the CNOP-I related to C_1 is $(\delta x_0, \delta y_0, \delta z_0) = (-0.1, -0.1, -0.1)$, while that of the Ref- C_2 is $(0.1, 0.1, -0.1)$. For simplicity, we only show the details for Ref- C_1 .

In addition, we perform a comparison of the results in Tables 1–4. Although the CNOP-I is a special case of the CNOP, it may not have the same pattern as that of the initial perturbation in the CNOP pattern. That is, the CNOP cannot be a simple combination of CNOP-I and CNOP-P. This conclusion indicates that, in predictability studies, it is necessary to investigate the specific combination of initial and parameter errors that can cause the largest prediction error. To further address this result, we now study the CNOP-P for the given reference states.

In (6) and (7), we neglected the initial perturbation and obtained (9), which is related to CNOP-P. With the given reference states and the constraint C_σ of the model parameter perturbation, we can obtain the CNOP-P by solving (9).

 TABLE 5: CNOP-P for the reference state with initial condition C_1 and model parameters $\sigma = 10$, $\alpha = 0.12$, $r = 28$, $b = 8/3$, $\beta = 0.3$.

	δ_σ	δ_α	δ_r	δ_b	δ_β
$\mu = 0.05$	-0.500	-0.006	-1.400	-0.133	0.015
$\mu = 0.1$	-1.000	-0.012	-2.800	-0.267	0.030
$\mu = 0.2$	-2.000	-0.024	-5.600	-0.533	0.060
$\mu = 0.3$	-2.736	-0.036	-8.400	-0.721	0.090

The Ref- O is a steady state for (10), and the objective function J_p is always equal to zero, regardless of the model parameters. That is, parameter perturbation cannot cause a departure from the steady reference state Ref- O . Consequently, there is no CNOP-P for Ref- O . The reference states Ref- C_1 and Ref- C_2 are time dependent and have one CNOP-P; the CNOP-P of Ref- C_1 is the same as that of Ref- C_2 . For simplicity, we list only the CNOP-P related to C_1 (Table 5).

It is clear from Table 5 that the CNOP-P also locates the boundary of the constraint. Furthermore, the perturbation components are not the same as the parameter perturbation components in the CNOP patterns that were obtained above. This finding further demonstrates that the CNOP for a given reference state is not a simple combination of the corresponding CNOP-I and CNOP-P. As a comparison, we list in Table 6 the CNOP of the Ref- C_1 for the constraint $\varepsilon = 0.4$ and $\mu = 0.3$ and the simple combination of the corresponding CNOP-I and CNOP-P. It is shown that the CNOP-I in the simple combination is different from the initial perturbation component in the CNOP pattern. Additionally, the corresponding CNOP-P is not the parameter perturbation component in the ‘‘CNOP.’’ Furthermore, the CNOP causes a larger prediction error than does the simple combination of CNOP-I and CNOP-P (i.e., the value of the objective function $J(u_0; p')$), which emphasizes the importance of CNOP calculation in predictability studies.

4. A Property of CNOP, CNOP-I, and CNOP-P

It has been demonstrated numerically that CNOP, CNOP-I, and CNOP-P for the coupled Lorenz model always locate the boundary of their respective constraints. Furthermore, Liu [18] proved theoretically that CNOP-I locates the boundary of the constraint for the model that depends continuously on initial perturbations. Next, we prove theoretically that the CNOP and CNOP-P also locate the boundaries of their respective constraints.

We rewrite (1) as another initial-value problem:

$$\frac{\partial \mathbf{w}}{\partial t} + F(\mathbf{w}, \tilde{\mathbf{p}}) = 0,$$

TABLE 6: The CNOP of the Ref-C1 for $\varepsilon = 0.4$ and $\mu = 0.3$ and the simple combination of the corresponding CNOP-I and CNOP-P.

	δ_{x_0}	δ_{y_0}	δ_{z_0}	δ_σ	δ_α	δ_r	δ_b	δ_β	$J(u_0; p')$
CNOP	-0.400	0.400	0.400	-2.606	-0.036	-8.400	-0.628	0.090	19.2096
combination	-0.400	-0.400	-0.400	-2.736	-0.036	-8.400	-0.721	0.090	16.5665

$$\begin{aligned}
\frac{\partial \bar{\mathbf{p}}}{\partial t} &= 0, \\
\mathbf{w}|_{t=0} &= \mathbf{w}_0, \\
\bar{\mathbf{p}}|_{t=0} &= \mathbf{p}.
\end{aligned} \tag{15}$$

Then, u_0 and p' in (10) can be regarded as the initial perturbations of (15) with the reference state U_0 and the reference parameter P . Furthermore, their evolutions are governed by the model $M_t(P + p')(U_0 + u_0) - M_t(P)(U_0) = M'_t(u_0; p')$, where M_t is the propagator of (15) and (1) and M'_t is defined as the propagator of the perturbation equation of (15). The perturbation $(u_0; p')$ can be regarded as the initial perturbation of (15). We assume that the model $u(t) = M'_t(u_0; p')$ depends continuously on initial values $(u_0; p')$. Then, we conclude with the following theorem.

Theorem 1. *The CNOP and its two special cases, CNOP-I and CNOP-P, locate the boundary of their respective constraints for the models whose perturbation solutions depend continuously on initial perturbations and parameter perturbations.*

Here, we prove this property of CNOP. Without loss of generality, the above initial perturbations of (15), (u_0, p') , are assumed to be the initial perturbations superimposed on the origin, where the origin is regarded as the equilibrium basic state, that is, when $u_0 = 0$, $p' = 0$, $M'_t(u_0; p') = u(t; u_0, p') = 0$. With this assumption, the optimization problem related to CNOP becomes

$$J(u_{0\delta}, p'_\sigma) = \max_{u_0 \in C_\delta; p' \in C_\sigma} \|u(T; u_0, p')\|. \tag{16}$$

Using (16), the CNOP-I $(u_{0\delta}, p'_\sigma)$ of (15) can be obtained; that is, (16) defines the CNOP-I of (15). In fact, this CNOP-I is the CNOP of (1). Liu [18] proved that CNOP-I locates the boundary of the constraint if the relevant model depends continuously on the initial perturbation. Furthermore, it has been assumed that the perturbation equation $u(t) = M'_t(u_0; p')$ depends continuously on initial values $(u_0; p')$. Therefore, according to Liu [18], the CNOP of (1), that is, the CNOP-I of (15), locates the boundary of the constraint.

In (16), we use $u_0 = 0$ and $p' \in C_\sigma$ to constrain the initial perturbations (u_0, p') . Then, the optimization problem (16) becomes (17):

$$J(p'_\sigma) = \max_{u_0=0; p' \in C_\sigma} \|u(T; 0, p')\| = \max_{p' \in C_\sigma} \|u(T; p')\|. \tag{17}$$

p'_σ in (17) is a CNOP-P of (15), and the perturbation $(0; p'_\sigma)$ is a CNOP of (15), which can be obtained using (16) with the constraints $u_0 = 0$ and $p' \in C_\sigma$. The parameter perturbation

p'_σ is equivalent to $(0; p'_\sigma)$. Thus, the CNOP-P p'_σ locates the boundary of its constraint C_σ .

To facilitate this discussion, if we signify the constraint $Q = \{(u_0, p) \mid u_0 = 0, p \in C_\sigma\}$, then Q is a closed set including the origin, indicating that the initial perturbations of (15) are constrained by the closed set Q . Furthermore, it has been assumed that (15) depends continuously on the initial perturbation. According to Liu [18], the CNOP-I of (15), $(0, p'_\sigma)$, is on the boundary of the constraint Q . Equivalently, the CNOP-P p'_σ locates the boundary of C_σ . Therefore, if we further assume that (1) depends continuously on parameter perturbations, CNOP-P usually locates the boundary of its constraint.

5. Summary

The CNOP is a useful tool for studying the limits of predictability in numerical weather forecasting and climate prediction. The CNOP has two special cases: one is the CNOP-I, which only relies on initial perturbations and has the largest nonlinear evolution at the time of prediction, while the other, CNOP-P, is merely related to the parameter perturbations, which causes the largest departure from a given reference state at a given future time. To find the CNOPs of a model, we reduce the model into an optimization problem. In this paper, we explore the CNOP, CNOP-I, and CNOP-P of a coupled Lorenz model and study their dependence on the reference state. We find that the CNOP, CNOP-I, and CNOP-P always locate the boundary of their respective constraints; we have proved this theorem logistically. Furthermore, we demonstrate that the CNOPs are dependent on the properties of the given reference states. For example, the reference state Ref-O is steady and, together with the z -axis, acts as a symmetric axis for the coupled Lorenz model system, which means that the Ref-O has two symmetric CNOPs about the z -axis. The reference states Ref-C₁ and Ref-C₂ are unsteady but symmetric about the z -axis. This property of these two reference states causes the two reference states to have one CNOP; furthermore, these two CNOPs are symmetric with respect to the z -axis, similar to their reference states. In addition, we demonstrate that the CNOP is not a simple combination of CNOP-I and CNOP-P, which suggests further investigation of the optimal combined mode of the initial perturbations and model perturbations, that is, the CNOP for predictability studies of numerical weather forecasting and climate predictions.

In this paper, we use a simple model to demonstrate CNOPs. In numerical weather forecasting and climate prediction, the models are generally complex. Therefore, to meet the requirements of realistic applications, we should solve the CNOPs of more complex models. The resulting CNOPs should also prove useful in improving the accuracy of weather

and climate forecasts. If mathematicians and meteorologists collaborate, their skills in forecasting weather and climate may be greatly improved.

Acknowledgments

The authors thank the anonymous reviewer very much for his useful comments. This work was sponsored by the Funding Project of Competence Development Program for Beijing VET Teachers and the National Natural Science Foundation of China (no. 41176013).

References

- [1] E. N. Lorenz, "Climate predictability," in *The Physical Basis of Climate Modelling*, Global Atmospheric Research Programme Publication Series, No. 16, pp. 132–136, World Meteorology Organization, Geneva, Switzerland, 1975.
- [2] E. Lorenz, "A study of the predictability of a 28-variable atmospheric model," *Tellus*, vol. 17, pp. 321–333, 1965.
- [3] Z. Toth and E. Kalnay, "Ensemble forecasting at NCEP and the breeding method," *Monthly Weather Review*, vol. 125, no. 12, pp. 3297–3318, 1997.
- [4] M. Mu, W. S. Duan, and B. Wang, "Conditional nonlinear optimal perturbation and its applications," *Nonlinear Processes in Geophysics*, vol. 10, no. 6, pp. 493–501, 2003.
- [5] B. F. Farrell, "Optimal excitation of baroclinic waves," *Journal of the Atmospheric Sciences*, vol. 46, no. 9, pp. 1193–1206, 1989.
- [6] W. Duan, M. Mu, and B. Wang, "Conditional nonlinear optimal perturbations as the optimal precursors for El Niño-Southern Oscillation events," *Journal of Geophysical Research D*, vol. 109, no. 23, Article ID D23105, 12 pages, 2004.
- [7] W. S. Duan, H. Xu, and M. Mu, "Decisive role of nonlinear temperature advection in El Niño and La Nina amplitude asymmetry," *Journal of Geophysical Research*, vol. 113, no. C1, Article ID C01014, 2008.
- [8] W. S. Duan, X. C. Liu, K. Y. Zhu, and M. Mu, "Exploring the initial errors that cause a significant "spring predictability barrier" for El Niño events," *Journal of Geophysical Research*, vol. 114, no. C4, Article ID C04022, 2009.
- [9] W. S. Duan and M. Mu, "Conditional nonlinear optimal perturbation: applications to stability, sensitivity, and predictability," *Science in China D*, vol. 52, no. 7, pp. 883–906, 2009.
- [10] W. S. Duan and R. Zhang, "Is model parameter error related to a significant spring predictability barrier for El Niño events? Results from a theoretical model," *Advances in Atmospheric Sciences*, vol. 27, no. 5, pp. 1003–1013, 2010.
- [11] M. Mu, W. Duan, Q. Wang, and R. Zhang, "An extension of conditional nonlinear optimal perturbation approach and its applications," *Nonlinear Processes in Geophysics*, vol. 17, no. 2, pp. 211–220, 2010.
- [12] W. S. Duan, Y. Yu, H. Xu, and P. Zhao, "Behaviors of nonlinearities modulating the El Niño events induced by optimal precursory disturbances," *Climate Dynamics*, 2012.
- [13] T. N. Palmer, "Predictability of the atmosphere and oceans: from days to decades," in *Proceeding of the ECMWF Seminar on Rredictability*, vol. 1, pp. 143–156, Reading, UK, September 1995.
- [14] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences*, vol. 20, no. 2, pp. 130–141, 1963.
- [15] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1196–1211, 2000.
- [16] M. J. D. Powell, "VMCWD: a Fortran subroutine for constrained optimization," DAMTP Report 1982/NA4, University of Cambridge, Cambridge, UK, 1982.
- [17] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming B*, vol. 45, no. 3, pp. 503–528, 1989.
- [18] Y. M. Liu, "Maximum principle for conditional optimal nonlinear perturbation," *Journal of East China Normal University*, no. 2, pp. 131–134, 2008 (Chinese).

Research Article

Algebraic Reconstruction of Current Dipoles and Quadrupoles in Three-Dimensional Space

Takaaki Nara

Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan

Correspondence should be addressed to Takaaki Nara; nara@mce.uec.ac.jp

Received 27 July 2012; Accepted 28 December 2012

Academic Editor: Valery Yakhno

Copyright © 2013 Takaaki Nara. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents an algebraic method for an inverse source problem for the Poisson equation where the source consists of dipoles and quadrupoles. This source model is significant in the magnetoencephalography inverse problem. The proposed method identifies the source parameters directly and algebraically using data without requiring an initial parameter estimate or iterative computation of the forward solution. The obtained parameters could be used for the initial solution in an optimization-based algorithm for further refinement.

1. Introduction

Inverse source problems for the Poisson equation have a wide variety of applications such as bioelectromagnetic inverse problems. Magnetoencephalography (MEG) and electroencephalography (EEG) are typical examples in which the current source inside the human brain is inversely reconstructed from measurements of the magnetic field outside the head and the electric potential on the head surface [1–3]. Due to the quasistatic property of the magnetic and electric fields, MEG and EEG exhibit high temporal resolution compared to other tools for the noninvasive inspection of brains. Hence, the development of a stable inverse algorithm which reconstructs the source with high spatial resolution is of importance for MEG and EEG.

Conventional algorithms for the MEG inverse problem, whose solution is generally nonunique [4], can be divided into two categories [2]: the parametric and imaging approaches. The parametric approach assumes that the current source inside the brain is expressed by a relatively small number of equivalent current dipoles (ECDs), which guarantees the uniqueness of the solution [4] and finds their number, positions, and dipole moments. A typical method uses optimization-based algorithms to minimize the squared sum of the difference between the forward solution and data measured at a finite number of sensor positions (see, e.g., [5]). The

forward solution is given by either the Geselowitz formula, which assumes that the head consists of arbitrarily-shaped layered domains, or the Sarvas formula, which assumes that the head consists of concentric spheres [1]. A problem with optimization-based algorithms is that they require an initial solution close to the true solution, or the algorithm will often converge to a local minimum. To resolve this problem, algebraic methods have been proposed and are gaining increased importance [6–10]. In these methods, the ECD parameters are reconstructed directly and algebraically from the boundary measurements of the magnetic field.

In contrast, the second category of inverse algorithms, the imaging approach, assumes that the many dipoles are set at the vertices of a mesh overlaying the cortical surface and solves for their dipole moments. An advantage of this approach is that the spatially distributed source can be reconstructed by solving linear equations for the unknown dipole moments. However, it has the disadvantages that the solution is not unique and an oversmoothed solution is often obtained by regularization, such as the minimum L_2 norm solution [11]. A method using the minimum L_1 norm solution has also been proposed to reconstruct a sparse source [12].

Recently, a new method, the multipolar representation of the source, has been developed that incorporates some of the advantages of the above two methods, and has attracted considerable attention [13–18]. In this model, instead of

expressing the current source by an equivalent current dipole, an equivalent dipole and quadrupole [15, 16] or an equivalent dipole and octopole [13, 14] are used where the quadrupole or octopole is determined depending on the spatial extent of the support of the current source. Jerbi et al. showed that the centroid of the spatially distributed source, which they called a patch source, can be estimated more accurately using the dipole and quadrupole model than using the dipole model by means of the nonlinear least squares method [16]. We considered a two-dimensional (2D) problem using a complex analysis framework and proposed an algebraic method to reconstruct the dipole and quadrupole parameters directly from the boundary measurements of the meromorphic function [17]. The aim of the current paper is to extend our algebraic method from the 2D case to the 3D case so that the dipole and quadrupoles, which equivalently represent the neural current, can be reconstructed from the magnetic field data without an initial parameter estimate or iterative computing of the forward solution. In [19, 20], we proposed an algebraic method when the dipoles were distributed in a plane parallel to the xy -plane, which is a very special case and is severely restricted in its practical usage. This paper derives a method for the general case. From a practical viewpoint, our method can provide an estimate of the number of patches as well as a good initial solution close to the true solution for optimization-based algorithms.

This paper is organized as follows. In Section 2, the forward problem with the dipole-quadrupole source model is summarized and the inverse problem is formulated. In Section 3, we propose a method to reconstruct the xy -coordinates of the dipole-quadrupole source by solving simultaneous algebraic equations of second degree. A method to estimate the z -coordinates is also proposed. Section 4 is devoted to numerical simulations for estimating spatially distributed dipoles.

2. Forward and Inverse Problems with the Dipole and Quadrupole Source Model

Let $\Omega_1, \Omega_2, \Omega_3$, and Ω_4 be concentric balls centered at the origin in 3D space, where $\Omega_i \subset \Omega_{i+1}$ for $i = 1, 2, 3$, as shown in Figure 1. Here, $\Omega_1, \Omega_2/\Omega_1$, and Ω_3/Ω_2 represent the brain, skull, and scalp, respectively. Ω_4 represents the head. We assume that the radial component of the magnetic field is measured on the sphere $\Gamma = \partial\Omega_4$ with the radius of R . Although we use this simple head model as well as the spherical sensor surface in this paper, our method can be extended to a more realistic case when the head is modeled by a piecewise homogeneous layered domain and the sensors are set on an arbitrarily shaped surface as assumed in the algebraic method for the dipole source model [10]. Let $\mathbf{J}_p(\mathbf{r})$ represent the neural currents in the innermost domain Ω_1 . The magnetic permeability is assumed to be constant μ_0 in the whole space.

First, we derive the representation of the radial magnetic field in terms of the equivalent dipoles and quadrupoles, which has already been given in [15], to determine how the

quadrupole terms equivalently express the spatial extent of the source.

It is known that with this head model, the radial component of the magnetic field at \mathbf{r} is equal to that when $\mathbf{J}_p(\mathbf{r})$ exists in infinite homogeneous space and is given by the Biot-Savart law:

$$\mathbf{r} \cdot \mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \int_{\Omega_1} \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r} \right) \cdot \mathbf{J}_p(\mathbf{r}') dv'. \quad (1)$$

Let us assume that the neural activity is localized at several domains $D_k (\subset \Omega_1)$ for $k = 1, 2, \dots, N$ so that $\mathbf{J}_p(\mathbf{r})$ is given by

$$\mathbf{J}_p(\mathbf{r}) = \sum_{k=1}^N \mathbf{j}_p(\mathbf{r}) \chi_{D_k}, \quad (2)$$

where χ_{D_k} is the characteristic function. Then, the radial component of the magnetic field is given by

$$\mathbf{r} \cdot \mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \sum_{k=1}^N \int_{D_k} \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv'. \quad (3)$$

Assume that D_k is contained in a ball $B_k \subset \Omega_1$, and there exists a ‘‘centroid’’ \mathbf{r}_k of D_k such that

$$|\mathbf{r} - \mathbf{r}_k| > |\mathbf{r}' - \mathbf{r}_k|, \quad \forall \mathbf{r} \in \Gamma, \mathbf{r}' \in D_k. \quad (4)$$

From the Taylor series expansion, we have

$$\begin{aligned} \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} &= \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}' = \mathbf{r}_k} \\ &+ \nabla' \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}' = \mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) + \dots \end{aligned} \quad (5)$$

Substituting this expansion into the right-hand side of (3) gives

$$\begin{aligned} \mathbf{r} \cdot \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \sum_{k=1}^N \frac{\mathbf{r} - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} \cdot \left(\mathbf{r}_k \times \int_{D_k} \mathbf{j}_p(\mathbf{r}') dv' \right) \\ &+ \frac{\mu_0}{4\pi} \sum_{k=1}^N \frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k) + |\mathbf{r} - \mathbf{r}_k|^2 X_r}{|\mathbf{r} - \mathbf{r}_k|^5} \\ &: \int_{D_k} \mathbf{j}_p(\mathbf{r}') (\mathbf{r}' - \mathbf{r}_k) dv' + \dots, \end{aligned} \quad (6)$$

where X_r is the cross-product tensor defined by

$$X_r \mathbf{a} = \mathbf{r} \times \mathbf{a}, \quad (7)$$

and hence is written as

$$X_r = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix} \quad \text{for } \mathbf{r} = (x, y, z)^T. \quad (8)$$

In (6), ‘‘:’’ represents the tensor contraction defined by $A : B = \sum_{i=1}^3 \sum_{j=1}^3 A_{ij} B_{ij}$ for second-order tensors A and B whose

(i, j) components are given by A_{ij} and B_{ij} , respectively. See Appendix A for the derivation of (6).

Now, let

$$\mathbf{p}_k \equiv \int_{D_k} \mathbf{j}_p(\mathbf{r}') d\mathbf{v}',$$

$$\begin{aligned} Q_k &\equiv \frac{1}{2} \left(\int_{D_k} \mathbf{j}_p(\mathbf{r}') (\mathbf{r}' - \mathbf{r}_k) + (\mathbf{r}' - \mathbf{r}_k) \mathbf{j}_p(\mathbf{r}') \right) d\mathbf{v}' \\ &= \begin{pmatrix} q_{xx,k} & q_{xy,k} & q_{xz,k} \\ q_{yx,k} & q_{yy,k} & q_{yz,k} \\ q_{zx,k} & q_{zy,k} & q_{zz,k} \end{pmatrix}, \end{aligned} \quad (9)$$

then we have

$$\begin{aligned} \mathbf{r} \cdot \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \sum_{k=1}^N (\mathbf{r}_k \times \mathbf{p}_k) \cdot \frac{\mathbf{r} - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} + \frac{\mu_0}{4\pi} \sum_{k=1}^N Q_k \\ &: \frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k) + |\mathbf{r} - \mathbf{r}_k|^2 X_r}{|\mathbf{r} - \mathbf{r}_k|^5} + \dots \end{aligned} \quad (10)$$

Note here that we define Q_k to be symmetric, $q_{ij,k} = q_{ji,k}$, since in (6) only the symmetric part of $\int_{D_k} \mathbf{j}_p(\mathbf{r}')(\mathbf{r}' - \mathbf{r}_k) d\mathbf{v}'$ contributes to $\mathbf{r} \cdot \mathbf{B}$ as shown in [15]. Equation (10) was given in (46) in [15] (when $N = 1$). The first term in (10) is the magnetic field created by an “equivalent current dipole” \mathbf{p}_k , and the second term is the magnetic field created by an “equivalent current quadrupole” Q_k . Note here that \mathbf{p}_k does not depend on the size of D_k , while Q_k depends on the spatial extent of \mathbf{j}_p in D_k . Q_k is a 3×3 tensor of order 2 and is called the quadrupole moment tensor.

Here, we can show that the truncation of (10) up to the second term,

$$\begin{aligned} \mathbf{r} \cdot \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \sum_{k=1}^N \left(\frac{(\mathbf{r}_k \times \mathbf{p}_k) \cdot (\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^3} \right. \\ &\quad \left. + Q_k : \frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k) + |\mathbf{r} - \mathbf{r}_k|^2 X_r}{|\mathbf{r} - \mathbf{r}_k|^5} \right), \end{aligned} \quad (11)$$

is the forward solution for the source model given by

$$\mathbf{J}_p(\mathbf{r}) = \sum_{k=1}^N \mathbf{p}_k \delta(\mathbf{r} - \mathbf{r}_k) + \sum_{k=1}^N Q_k \nabla \delta(\mathbf{r} - \mathbf{r}_k), \quad (12)$$

where δ is the Dirac delta function (see Appendix B). We call (12) the “equivalent dipole and quadrupole source model” or simply the “dipole-quadrupole source model.” We are now ready to state our inverse problem.

Problem 1. Assume that a head Ω_3 is modeled by a spherically layered domain. Assume that the neural current is expressed by (12) supported in the inner-most ball Ω_1 representing the brain. Then, reconstruct the number N and positions \mathbf{r}_k of the dipole-quadrupole source in (12) from the radial component of the magnetic field measured on the spherical boundary Γ which encloses Ω_3 .

Remark 1. Once the number and positions of the dipole-quadrupole source are determined, the dipole moments \mathbf{p}_k and quadrupole moment tensors Q_k can be obtained by solving the relevant linear equation (11). In this paper, we concentrate on reconstructing N and \mathbf{r}_k .

3. Algebraic Reconstruction Method

According to (88) in [15], we can express (1) by the multipole expansion

$$\mathbf{r} \cdot \mathbf{B}(\mathbf{r}) = \mu_0 \sum_{l=0}^{\infty} \sum_{m=-l}^l \frac{l+1}{2l+1} M_{lm} \frac{\widehat{Y}_{lm}^*(\theta, \phi)}{r^{l+1}}, \quad (13)$$

where

$$\begin{aligned} \widehat{Y}_{lm}(\theta, \phi) &= \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} \\ &\times Y_{lm}(\theta, \phi): \text{normalized spherical harmonics,} \end{aligned}$$

$$Y_{lm}(\theta, \phi) = P_l^m(\cos \theta) e^{im\phi}: \text{spherical harmonics,} \quad (14)$$

and $P_l^m(\cos \theta)$ are the associated Legendre polynomials. As shown in (84) in [15], the multipole moment has the following relationship with \mathbf{J}_p :

$$M_{lm} = \frac{1}{l+1} \int_{\Omega_1} [\nabla' r'^l \widehat{Y}_{lm}(\theta', \phi')] \cdot (\mathbf{r}' \times \mathbf{J}_p(\mathbf{r}')) d^3 r'. \quad (15)$$

Substituting (12) into (15) and using (B.1) gives

$$\begin{aligned} M_{lm} &= \frac{1}{l+1} \sum_{k=1}^N \left(\nabla' r'^l \widehat{Y}_{lm}(\theta', \phi') \Big|_{\mathbf{r}'=\mathbf{r}_k} + Q_k \right. \\ &\quad \left. : (\nabla' (\nabla' r'^l \widehat{Y}_{lm}(\theta', \phi') \times \mathbf{r}'))^T \Big|_{\mathbf{r}'=\mathbf{r}_k} \right). \end{aligned} \quad (16)$$

On the other hand, using the orthogonality of the spherical harmonics, the multipole moment is shown to have another relationship with the radial magnetic field on a spherical surface Γ [21]:

$$M_{lm} = \frac{2l+1}{\mu_0 (l+1)} \int_{\Gamma} \mathbf{n}' \cdot \mathbf{B}(\mathbf{r}') R^l \widehat{Y}_{lm}(\theta', \phi') dS', \quad (17)$$

where \mathbf{n}' is the outward unit normal vector at \mathbf{r}' on Γ . Equating (16) and (17) for $l \geq m \geq 0$ gives algebraic equations relating the unknown parameters of the sources to the radial magnetic field on Γ . Although (17) requires the radial components of the magnetic field on the whole boundary Γ , Taulu et al. proposed a new approach for obtaining M_{lm} from the data on the part of Γ by using the truncated spherical harmonic expansion [22]. Using the method, we have algebraic equations about the unknown source parameters with the practically measurable MEG data.

3.1. *Algebraic Reconstruction of the x- and y-Coordinates.* As we derived the algebraic method for the dipole source model in [10], we now use the multipole moments of the sectorial harmonics represented by $l = m = n$, where $n \in \mathbb{N}$. From the relationship $r^n Y_{nm}(\theta, \phi) = (2n - 1)!!(x + iy)^n$ [23], we have

$$\begin{aligned} & \sum_{k=1}^N \left(\nabla(x + iy)^n \Big|_{\mathbf{r}=\mathbf{r}_k} \cdot (\mathbf{r}_k \times \mathbf{p}_k) \right. \\ & \quad \left. + Q_k : \left(\nabla \left(\nabla(x + iy)^n \times \mathbf{r} \right) \right)^T \Big|_{\mathbf{r}=\mathbf{r}_k} \right) \\ & = \frac{2n + 1}{\mu_0} \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n dS. \end{aligned} \quad (18)$$

The primes in both sides are removed for simplicity. Since it holds that

$$\begin{aligned} \nabla(x + iy)^n &= n \binom{1}{i} (x + iy)^{n-1}, \\ (\nabla \left(\nabla(x + iy)^n \times \mathbf{r} \right))^T &= n \begin{pmatrix} i(n-1)z(x+iy)^{n-2} & -(n-1)z(x+iy)^{n-2} & i(x+iy)^{n-1} \\ -(n-1)z(x+iy)^{n-2} & -i(n-1)z(x+iy)^{n-2} & -(x+iy)^{n-1} \\ -in(x+iy)^{n-1} & n(x+iy)^{n-1} & 0 \end{pmatrix}, \end{aligned} \quad (19)$$

we obtain

$$\nabla(x + iy) \Big|_{\mathbf{r}=\mathbf{r}_k} \cdot (\mathbf{r}_k \times \mathbf{p}_k) = n[\mathbf{r}_k \times \mathbf{p}_k]_{x+iy} (x_k + iy_k)^{n-1}, \quad (20)$$

$$\begin{aligned} Q_k : \left(\nabla \left(\nabla(x + iy)^n \times \mathbf{r} \right) \right)^T \Big|_{\mathbf{r}=\mathbf{r}_k} \\ = n(n-1) \left(\left(i(q_{xx,k} - q_{yy,k}) - 2q_{xy,k} \right) z_k \right. \\ \quad \left. + (q_{yz,k} - iq_{xz,k})(x_k + iy_k) \right) (x_k + iy_k)^{n-2}, \end{aligned} \quad (21)$$

where $[\mathbf{r}_k \times \mathbf{p}_k]_{x+iy}$ represents $(x\text{-component of } \mathbf{r}_k \times \mathbf{p}_k) + i \times (y\text{-component of } \mathbf{r}_k \times \mathbf{p}_k)$. See Appendix C for the derivation of (21). Substituting (20) and (21) into (18) gives

$$\begin{aligned} & \sum_{k=1}^N \left([\mathbf{r}_k \times \mathbf{p}_k]_{x+iy} (x_k + iy_k)^{n-1} + (n-1) \right. \\ & \quad \times \left(\left(i(q_{xx,k} - q_{yy,k}) - 2q_{xy,k} \right) z_k \right. \\ & \quad \left. + (q_{yz,k} - iq_{xz,k})(x_k + iy_k) \right) (x_k + iy_k)^{n-2} \\ & = \frac{2n + 1}{\mu_0 n} \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(x + iy)^n dS. \end{aligned} \quad (22)$$

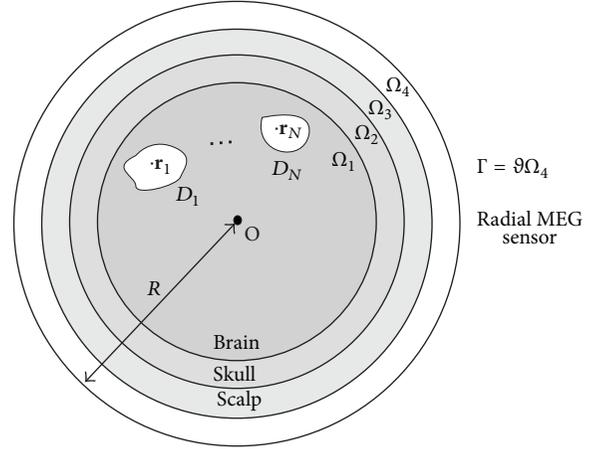


FIGURE 1: Head and sensor model.

We now put

$$\begin{aligned} \zeta_k &= x_k + iy_k, \\ \mu_k &= [\mathbf{r}_k \times \mathbf{p}_k]_{x+iy}, \\ \mathbf{v}_k &= \left(i(q_{xx,k} - q_{yy,k}) - 2q_{xy,k} \right) z_k \\ & \quad + (q_{yz,k} - iq_{xz,k})(x_k + iy_k), \\ c_n &= \frac{2n + 3}{\mu_0 (n + 1)} \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(x + iy)^{n+1} dS \end{aligned} \quad (23)$$

and rewrite (22) as

$$\sum_{k=1}^N \mu_k \zeta_k^n + n \sum_{k=1}^N \mathbf{v}_k \zeta_k^{n-1} = c_n \quad (n \in \mathbb{N}^*). \quad (24)$$

Comparing with the special case where the dipoles were distributed in a plane parallel to the xy -plane in [20], one finds that \mathbf{v}_k has an extra term: $(q_{yz,k} - iq_{xz,k})(x_k + iy_k)$.

Equation (24) has the same form as that of (6) in [17], and consequently the dipole-quadrupole position projected on the xy -plane, ζ_k ($k = 1, 2, \dots, N$), can be algebraically reconstructed from c_n ($n = 0, 1, \dots, 2N - 1$), which is shown as follows. In Section 3.1, we assume that N is known *a priori* and construct an algorithm to estimate ζ_k in (24). A method to estimate N as well as μ_k and \mathbf{v}_k is then described in Section 3.2.

Let us define $\sigma_1, \sigma_2, \dots, \sigma_N$ by the coefficients of

$$\prod_{k=1}^N (\zeta - \zeta_k) = \zeta^N + \sigma_1 \zeta^{N-1} + \dots + \sigma_N. \quad (25)$$

It was proved in [17] that (24) can be transformed into the second-degree equations for $\sigma_1, \sigma_2, \dots, \sigma_N$:

$$\boldsymbol{\sigma}^T H_{N,m} \boldsymbol{\sigma} = 0 \quad (m \in \mathbb{N}^*), \quad (26)$$

where $\boldsymbol{\sigma} = (\sigma_N, \dots, \sigma_1, 1)^T$ and $H_{N,m}$ is the $(N+1) \times (N+1)$ Hankel matrix given by

$$H_{N,m} = \begin{pmatrix} c_m & c_{m+1} & \cdots & c_{m+N} \\ c_{m+1} & c_{m+2} & \cdots & c_{m+N+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m+N} & c_{m+N+1} & \cdots & c_{m+2N} \end{pmatrix}. \quad (27)$$

Moreover, it has been shown that the $2N$ equations (26) for $m = 0, 1, \dots, 2N-1$ can be turned into linear equations as follows. First, from those $2N$ equations, we make $2N-1$ equations

$$\boldsymbol{\sigma}^T H_{N,m}^{(1)} \boldsymbol{\sigma} = 0 \quad (m = 0, 1, \dots, 2N-2), \quad (28)$$

where

$$H_{N,m}^{(1)} \equiv c_{m+1} H_{N,m} - c_m H_{N,m+1}. \quad (29)$$

By this definition, $H_{N,m}^{(1)}$ are the Hankel matrices whose $(1, 1)$ -components are zero. We denote their components by

$$H_{N,m}^{(1)} = \begin{pmatrix} 0 & c_{m+1}^{(1)} & \cdots & c_{m+N}^{(1)} \\ c_{m+1}^{(1)} & c_{m+2}^{(1)} & \cdots & c_{m+N+1}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m+N}^{(1)} & c_{m+N+1}^{(1)} & \cdots & c_{m+2N}^{(1)} \end{pmatrix}. \quad (30)$$

Next, from those $2N-1$ equations, we make $2N-2$ equations:

$$\boldsymbol{\sigma}^T H_{N,m}^{(2)} \boldsymbol{\sigma} = 0 \quad (m = 0, 1, \dots, 2N-3), \quad (31)$$

where

$$H_{N,m}^{(2)} \equiv c_{m+2}^{(1)} H_{N,m}^{(1)} - c_{m+1}^{(1)} H_{N,m+1}^{(1)}. \quad (32)$$

By definition, $H_{N,m}^{(2)}$ are the Hankel matrices expressed by

$$H_{N,m}^{(2)} = \begin{pmatrix} 0 & 0 & c_{m+2}^{(2)} & \cdots & c_{m+N}^{(2)} \\ 0 & c_{m+2}^{(2)} & \cdots & \cdots & c_{m+N+1}^{(2)} \\ c_{m+2}^{(2)} & \cdots & \cdots & \cdots & c_{m+N+1}^{(2)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{m+N}^{(2)} & c_{m+N+1}^{(2)} & \cdots & \cdots & c_{m+2N}^{(2)} \end{pmatrix}. \quad (33)$$

Note that the components along the first and second lines in the anti-diagonal direction become zero. Repeating this procedure to eliminate the components on the j th line in the anti-diagonal direction for $j = 3, \dots, 2N-1$, we obtain

$$\boldsymbol{\sigma}^T H_{N,m}^{(j)} \boldsymbol{\sigma} = 0 \quad (m = 0, 1, \dots, 2N-1-j), \quad (34)$$

where $H_{N,m}^{(j)}$ are the Hankel matrices whose components along the first through j th lines in the anti-diagonal direction are zero. Here, (34) for $j = 2N-1$ is a linear equation for σ_N . Solving it and substituting the solution into (34) for $j = 2N-2$ gives a linear equation for σ_{N-1} . Repeating the backward

substitution gives $\sigma_{N-2}, \dots, \sigma_1$, successively. The conditions for the unique solvability of these linear equations are $\zeta_i \neq \zeta_j$ for $i \neq j$ and $v_k \neq 0$. See [17] for more details. Once $\sigma_1, \dots, \sigma_N$ are obtained, we have ζ_k as the roots of $\zeta^N + \sigma_1 \zeta^{N-1} + \cdots + \sigma_N = 0$.

In a practical situation when data includes noise, the above procedure proposed in [17] to transform the second-degree equations (26) into linear equations (34) to obtain $\sigma_N, \dots, \sigma_1$ may be sensitive to the noise due to the cancellation. To reduce the cancellation error, we propose the following algorithm. First, we solve the second-degree equations (26) for $m = 0, \dots, N-1$ using a well-established method for solving simultaneous algebraic equations, for example, by means of the Gröbner bases. Then, out of the obtained 2^N solutions, $(\sigma_{1,j}, \dots, \sigma_{N,j})$ for $j = 1, \dots, 2^N$, we choose one which minimizes the sum of the absolute values of the left-hand sides of (26) for $m = N, \dots, 2N-1$ given by

$$\sum_{m=N}^{2N-1} |\boldsymbol{\sigma}^T H_{N,m} \boldsymbol{\sigma}|. \quad (35)$$

In other words, to obtain the theoretically unique solution of (26) for $m = 0, \dots, 2N-1$, first we find 2^N candidates from (26) for $m = 0, \dots, N-1$ and then choose the one which best explains the remaining equations (26) for $m = N, \dots, 2N-1$.

3.2. Reconstruction of μ_k , v_k , and N . To estimate N , following the method for the dipole source model [10], we assume that there are $N' (> N)$ dipole-quadrupole sources and then estimate ζ_k for $k = 1, 2, \dots, N'$ using the algorithm in Section 3.1. Once they are obtained, μ_k and v_k for $k = 1, 2, \dots, N'$ are linearly solved using (24) for $n = 0, 1, \dots, 2N'-1$. Then, we compute $|\mu_{n+1}/\mu_n|$ and $|v_{n+1}/v_n|$ for $n = 1, 2, \dots, N'-1$, which are expected to be sufficiently small when $n = N$. Practically, we estimate N such that these ratios become smaller than some threshold set *a priori*. The thresholds should be determined by the ratios of the noise level contained in the data to the dipole and quadrupole strength which can be regarded as a physiologically meaningful source. As for the dipole source model in the 2D problem, the threshold is theoretically evaluated in the context of the Padé approximation [24]. A similar theory for the dipole-quadrupole source model, although greatly required, is left for further research; in this paper we show only numerical simulations in Section 4.

3.3. Algebraic Reconstruction of the z -Coordinates. To determine the z -coordinates of the dipole-quadrupole positions, after we determine the xy -coordinates of the source, we use the method proposed in [7]. We put $l = n+1$ and $m = n$ in (16) and (17), where $n \in \mathbb{N}^*$. From the identity $r^{n+1} \hat{Y}_{n+1,n} = (2n+1)!(x+iy)^n z$ [23], we have

$$\begin{aligned} & \sum_{k=1}^N \left(\nabla(x+iy)^n z \Big|_{\mathbf{r}=\mathbf{r}_k} \cdot (\mathbf{r}_k \times \mathbf{p}_k) \right. \\ & \quad \left. + Q_k : \left(\nabla \left(\nabla(x+iy)^n z \times \mathbf{r} \right) \Big|_{\mathbf{r}=\mathbf{r}_k} \right) \right) \\ & = \frac{2n+3}{\mu_0} \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x+iy)^n z dS. \end{aligned} \quad (36)$$

Now, let

$$\phi_j(t) \equiv (t - \zeta_j)^2 \prod_{l=1, l \neq j}^N (t - \zeta_l)^3 \equiv \sum_{n=0}^{3N-1} \alpha_n t^n. \quad (37)$$

Then, from $\sum_{n=0}^{3N-1} \alpha_n \times (18)$ and $\sum_{n=0}^{3N-1} \alpha_n \times (36)$, we obtain

$$\begin{aligned} & \sum_{k=1}^N \left(\nabla \phi_j(x + iy) \Big|_{\mathbf{r}=\mathbf{r}_k} \cdot (\mathbf{r}_k \times \mathbf{p}_k) \right. \\ & \left. + Q_k : \left(\nabla \left(\nabla \phi_j(x + iy) \times \mathbf{r} \right) \Big|_{\mathbf{r}=\mathbf{r}_k} \right)^T \right) \\ & = \frac{1}{\mu_0} \sum_{n=0}^{3N-1} (2n+1) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n dS, \end{aligned} \quad (38)$$

$$\begin{aligned} & \sum_{k=1}^N \left(\nabla \phi_j(x + iy) z \Big|_{\mathbf{r}=\mathbf{r}_k} \cdot (\mathbf{r}_k \times \mathbf{p}_k) \right. \\ & \left. + Q_k : \left(\nabla \left(\nabla \phi_j(x + iy) z \times \mathbf{r} \right) \Big|_{\mathbf{r}=\mathbf{r}_k} \right)^T \right) \\ & = \frac{1}{\mu_0} \sum_{n=0}^{3N-1} (2n+3) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n z dS, \end{aligned} \quad (39)$$

respectively. In the left-hand sides of (38) and (39),

$$\nabla \phi_j(x + iy) \Big|_{\mathbf{r}=\mathbf{r}_k} = \begin{pmatrix} \phi_j'(\zeta_k) \\ i\phi_j'(\zeta_k) \\ 0 \end{pmatrix} = \mathbf{0}, \quad (40)$$

$$\nabla \phi_j(x + iy) z \Big|_{\mathbf{r}=\mathbf{r}_k} = \begin{pmatrix} z\phi_j'(\zeta_k) \\ iz\phi_j'(\zeta_k) \\ \phi_j(\zeta_k) \end{pmatrix} = \mathbf{0} \quad (41)$$

since from (37) it holds that $\phi_j(\zeta_k) = \phi_j'(\zeta_k) = 0$ for all $k = 1, 2, \dots, N$ irrespective of whether or not k is equal to j . Also, we have

$$\nabla \left(\nabla \phi_j(x + iy) \times \mathbf{r} \right) \Big|_{\mathbf{r}=\mathbf{r}_k} = \delta_{jk} \Phi_j, \quad (42)$$

$$\nabla \left(\nabla \phi_j(x + iy) z \times \mathbf{r} \right) \Big|_{\mathbf{r}=\mathbf{r}_k} = z_j \delta_{jk} \Phi_j, \quad (43)$$

where δ_{jk} is the Kronecker delta, and

$$\Phi_j = 2 \prod_{l=1, l \neq j}^N (\zeta_j - \zeta_l)^3 \begin{pmatrix} iz_j & -z_j & -i\zeta_j \\ -z_j & -iz_j & \zeta_j \\ 0 & 0 & 0 \end{pmatrix}. \quad (44)$$

See Appendix D for the derivation of (42) and (43).

Substituting (40) through (43) into (38) and (39) gives

$$\begin{aligned} Q_j : \Phi_j^T &= \frac{1}{\mu_0} \sum_{n=0}^{3N-1} (2n+1) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n dS, \\ Q_j : z_j \Phi_j^T &= \frac{1}{\mu_0} \sum_{n=0}^{3N-1} (2n+3) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n z dS, \end{aligned} \quad (45)$$

from which we obtain

$$z_j = \frac{\sum_{n=0}^{3N-1} (2n+3) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n z dS}{\sum_{n=0}^{3N-1} (2n+1) \alpha_n \int_{\Gamma} \mathbf{n} \cdot \mathbf{B}(\mathbf{r}) (x + iy)^n dS}. \quad (46)$$

This is the reconstruction formula of the z -coordinates: once ζ_k for $k = 1, 2, \dots, N$ are obtained, then z_j are determined using (46) for each $j = 1, 2, \dots, N$ with α_n defined by (37).

3.4. Algorithm. Our algorithm is summarized as follows.

Step 1 (estimate the number of the dipole-quadrupole sources). (1) Assume that there are $N' (> N)$ dipole-quadrupole sources. Using the algorithm in Step 2(1)–(5) below, where N is replaced by N' , obtain μ_k and ν_k for $k = 1, 2, \dots, N'$.

(2) Compute $|\mu_{n+1}/\mu_n|$ and $|\nu_{n+1}/\nu_n|$ for $n = 1, 2, \dots, N' - 1$. When $|\mu_{\widehat{N}+1}/\mu_{\widehat{N}}| < \epsilon_\mu$ and $|\nu_{\widehat{N}+1}/\nu_{\widehat{N}}| < \epsilon_\nu$ for the thresholds ϵ_μ and ϵ_ν , which are appropriately set *a priori*, we estimate N to be \widehat{N} .

Step 2 (estimate the position parameters of N dipole-quadrupole sources). (1) Given $\mathbf{n} \cdot \mathbf{B}$ on Γ , compute c_n in (23) for $n = 0, 1, \dots, 4N - 1$ and construct the Hankel matrices (27) for $m = 0, 1, \dots, 2N - 1$.

(2) Solve the simultaneous second-degree equation (26) for $m = 0, 1, \dots, N - 1$, using the Gröbner bases to get 2^N solutions.

(3) Out of the 2^N solutions, choose one solution $\sigma_1, \dots, \sigma_N$ that minimizes (35).

(4) Solve $z^N + \sigma_1 z^{N-1} + \dots + \sigma_N = 0$ to obtain ζ_1, \dots, ζ_N .

(5) Solve (24) for μ_k and ν_k , where $k = 1, \dots, N$.

(6) Using the obtained ζ_1, \dots, ζ_N , compute α_n in (37) for $n = 0, 1, \dots, 3N - 1$ and determine z_1, \dots, z_N using (46).

4. Numerical Simulations

In this section, our algorithm is verified numerically. Γ was assumed to be a sphere with $R = 0.12$ m on which 361 measurement points were distributed uniformly using the spherical t-design [25]. To validate our algorithm for the dipole-quadrupole model, we assumed that the data was available on the whole sphere which enclosed the source. Identification using data on a part of Γ is an important aspect of future studies, for which the method proposed by Taulu et al. [22] would be useful.

4.1. Demonstration of Our Method. In this subsection, We examine the case where there are $N = 2$ dipole-quadrupole sources whose parameters are given in Table 1. The theoretical data was computed using (11) to which Gaussian noise with $\sigma = 10^{-2}$ was added where the noise level σ is defined by the ratio of the standard deviation of the noise to the root mean squares of the data. 100 data sets with the different noise added were used for reconstruction.

First, to determine the number of the sources N , we assume that there are $N' = 3$ dipole-quadrupole sources.

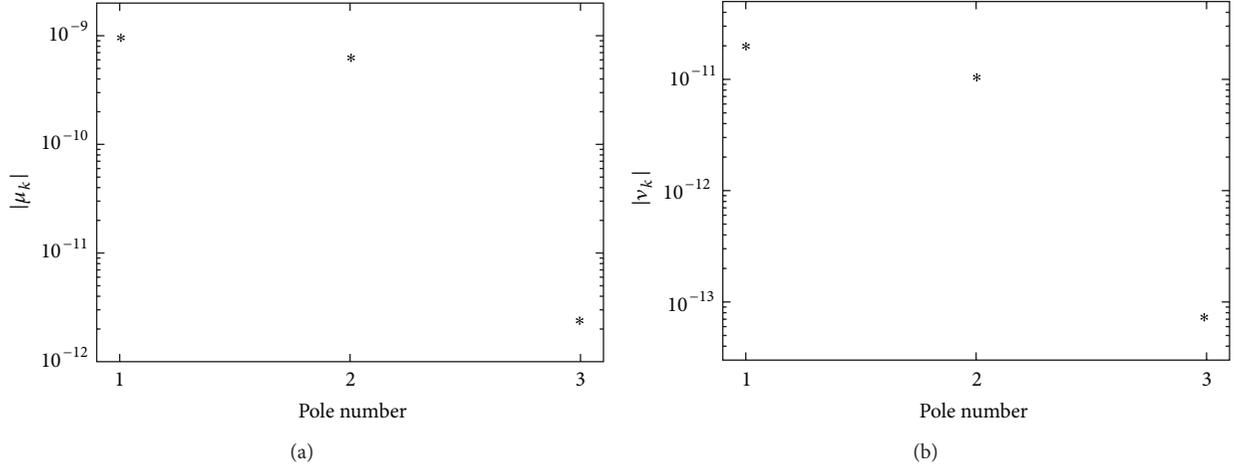


FIGURE 2: Estimation of N : the obtained $|\mu_k|$ (a) and $|\nu_k|$ (b) for $k = 1, 2, 3$ when assuming $N^t = 3$. Since $|\mu_3|$ and $|\nu_3|$ are much smaller than $|\mu_k|$ and $|\nu_k|$ for $k = 1, 2$, respectively ($|\mu_3/\mu_2| = 3.8 \times 10^{-3}$ and $|\nu_3/\nu_2| = 7.2 \times 10^{-3}$), we can judge that there are substantially $N = 2$ dipole-quadrupole sources.

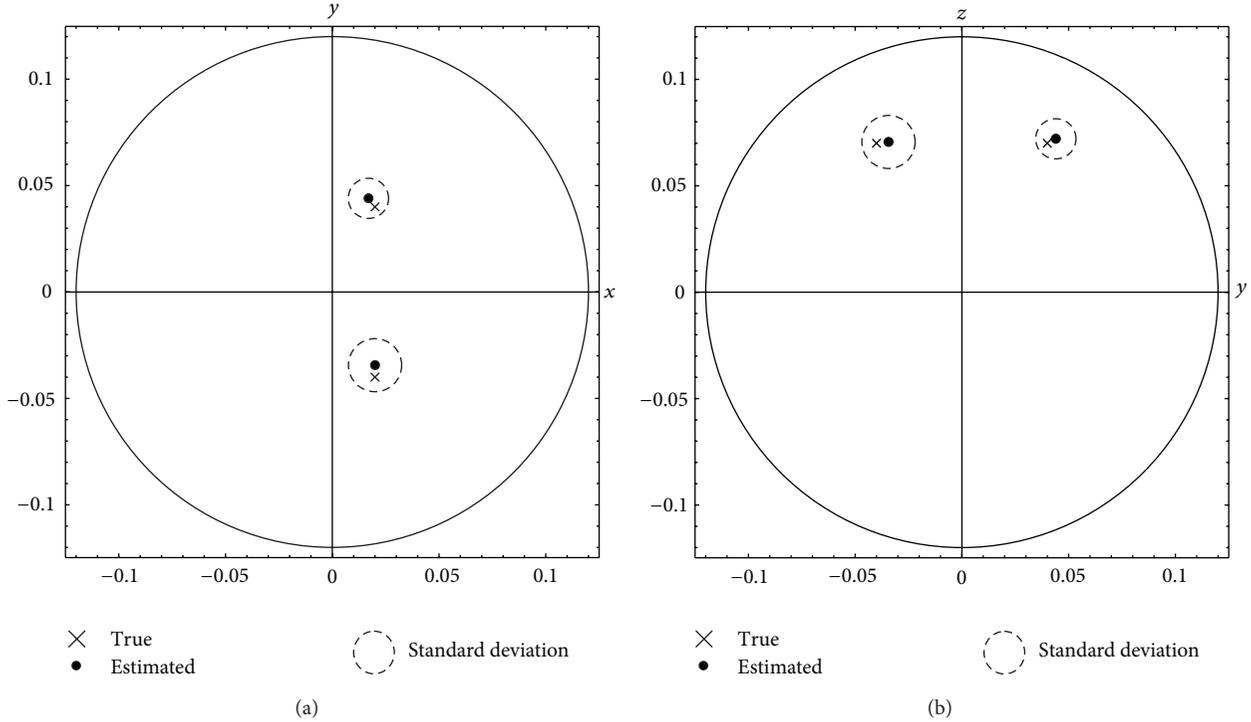


FIGURE 3: Localization result when $N = 2$ where the noise level $\sigma = 10^{-2}$.

Figure 2 shows the obtained $|\mu_k|$ and $|\nu_k|$ for $k = 1, 2, 3$. One observes that $|\mu_3|$ and $|\nu_3|$ are much smaller than $|\mu_k|$ and $|\nu_k|$ for $k = 1, 2$, respectively. In fact, the geometric means of $|\mu_3/\mu_2|$ and $|\nu_3/\nu_2|$ for 100 data sets were 3.8×10^{-3} and 7.2×10^{-3} , respectively. From this, we can judge that there are substantially $N = 2$ poles.

Figure 3 shows the localization result projected on the xy - and yz -planes. We find that, under this noise level, two dipole-quadrupole sources were stably reconstructed. The average and standard deviations of the computational time to

identify the parameters of the two dipole-quadrupole sources were 0.69 sec and 0.01 sec, respectively, using an Intel Core i7 CPU 870 (2.93 GHz) with 8 GB RAM, showing that our algebraic algorithm can determine the source parameters with low computational cost.

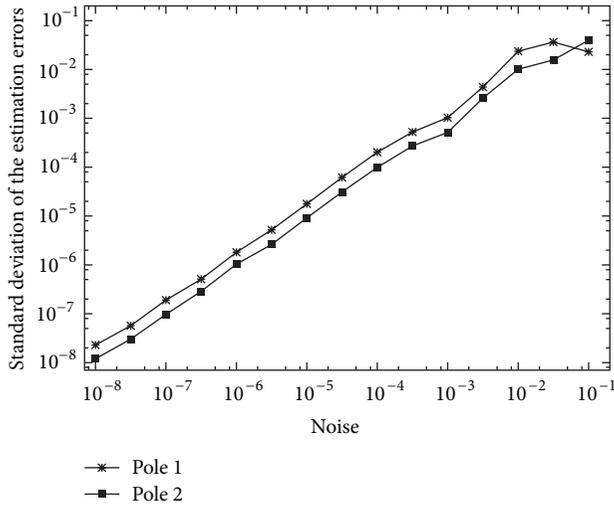
4.2. *Effect of Noise on Localization Accuracy.* To examine the stability of our algorithm, the noise level σ was varied in the wide range $10^{-8} \leq \sigma \leq 10^{-1}$. The source and sensor configuration were the same as in Section 4.1. Figure 4 shows

TABLE 1: Dipole-quadrupole parameters when $N = 2$.

k (pole no.)	\mathbf{r}_k	\mathbf{p}_k	$(q_{xx,k}, q_{xy,k}, q_{xz,k}, q_{yy,k}, q_{yz,k}, q_{zz,k})$
1	(0.02, 0.04, 0.07) m	(10, 0, 10) nAm	$(1, 1, 0, 2, 1, 0) \times 10^{-1} \text{ nAm}^2$
2	(0.02, -0.04, 0.07) m	(10, 0, -10) nAm	$(1, 1, 0, -2, 1, 0) \times 10^{-1} \text{ nAm}^2$

TABLE 2: dipole-quadrupole parameters when $N = 3$.

k (pole no.)	\mathbf{r}_k	\mathbf{p}_k	$(q_{xx,k}, q_{xy,k}, q_{xz,k}, q_{yy,k}, q_{yz,k}, q_{zz,k})$
1	(0.02, -0.04, 0.07) mm	(10, 0, 10) nAm	$(1, 1, 0, 2, 1, 0) \times 10^{-1} \text{ nAm}^2$
2	(-0.04, 0, 0.07) mm	(10, 0, -10) nAm	$(1, 1, 0, -2, 1, 0) \times 10^{-1} \text{ nAm}^2$
3	(0, 0.05, 0.07) mm	(0, 10, 10) nAm	$(2, 1, 0, 1, 1, 0) \times 10^{-1} \text{ nAm}^2$

FIGURE 4: Standard deviation of the localization errors with respect to the noise level σ .

the standard deviation (s.d.) of the localization errors with respect to the noise level σ for the two poles. It is observed that s.d. increases as σ increases. For this source in the sphere with the radius 0.12 m, to realize s.d. less than 0.01 m, one finds that the noise level should be less than $\sigma = 10^{-2}$.

4.3. Effect of the Distance between Two Sources on Localization Accuracy. We varied the distance between the two dipole-quadrupole sources: the two dipole-quadrupole source were set at $(0.02, \pm d, 0.07)$ where $2d$ was changed from 0.01 through 0.09. Figure 5 shows the relative localization error (error divided by the radius of Γ , 0.12). It is observed that the distance between the two dipole-quadrupole sources should be larger than about 0.06, that is the half of the radius of Γ , in order to obtain the relative error less than 10^{-2} .

4.4. Example When $N = 3$. We examine the case when there are $N = 3$ dipole-quadrupole sources whose parameters are given in Table 2. The noise level was set at $\sigma = 10^{-2}$. In this subsection, we assumed that N is known *a priori*. Figure 6 shows the localization result. Comparing with the result when $N = 2$ in Figure 3, the biases as well as the standard deviations of the localization errors become

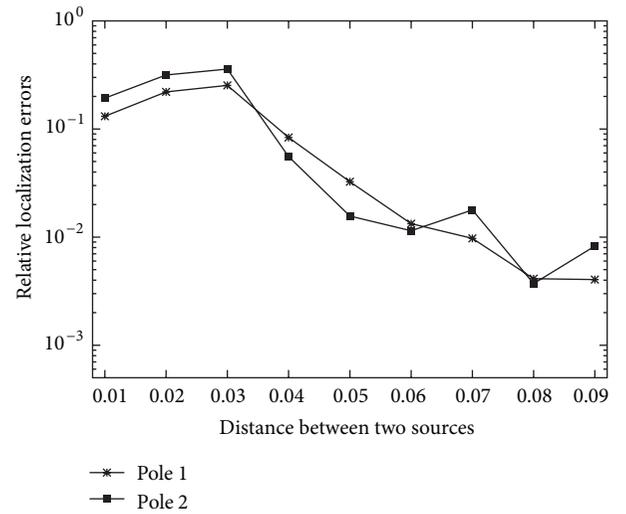


FIGURE 5: Relative localization error with respect to the distance between the two dipole-quadrupole sources.

larger when $N = 3$. It is numerically shown in [26] that, when using dipole-only source model, stability gets worse when N increases and it is difficult to estimate the dipoles greater than about $N = 5$. The example in this subsection shows that it is more difficult to estimate the large number ($N \geq 3$) of dipole-quadrupole sources. This is considered due to fast spatial decay of the quadrupolar field. However, the dipole-quadrupole source even with $N = 1$ or 2 has an advantage that it can well estimate a spatially distributed sources, especially when they are modeled with two, close, oppositely directed dipoles, which is a weak point of the conventional dipole model. This is shown in the next subsection.

4.5. Identification of Dipoles Distributed on Curved Surfaces. In this subsection, we compare our algebraic method assuming the dipole-quadrupole model (DQM) with a conventional algebraic method assuming the dipole model (DM) for estimating the spatially distributed dipole sources. To model dipoles on cerebral convolutions, we assume that dipoles are placed on a mesh on a half-cylinder with a radius of $r = 5$ mm and a height of $h = 5$ mm, as shown in Figure 7. There are six dipoles in the circumferential direction by five in

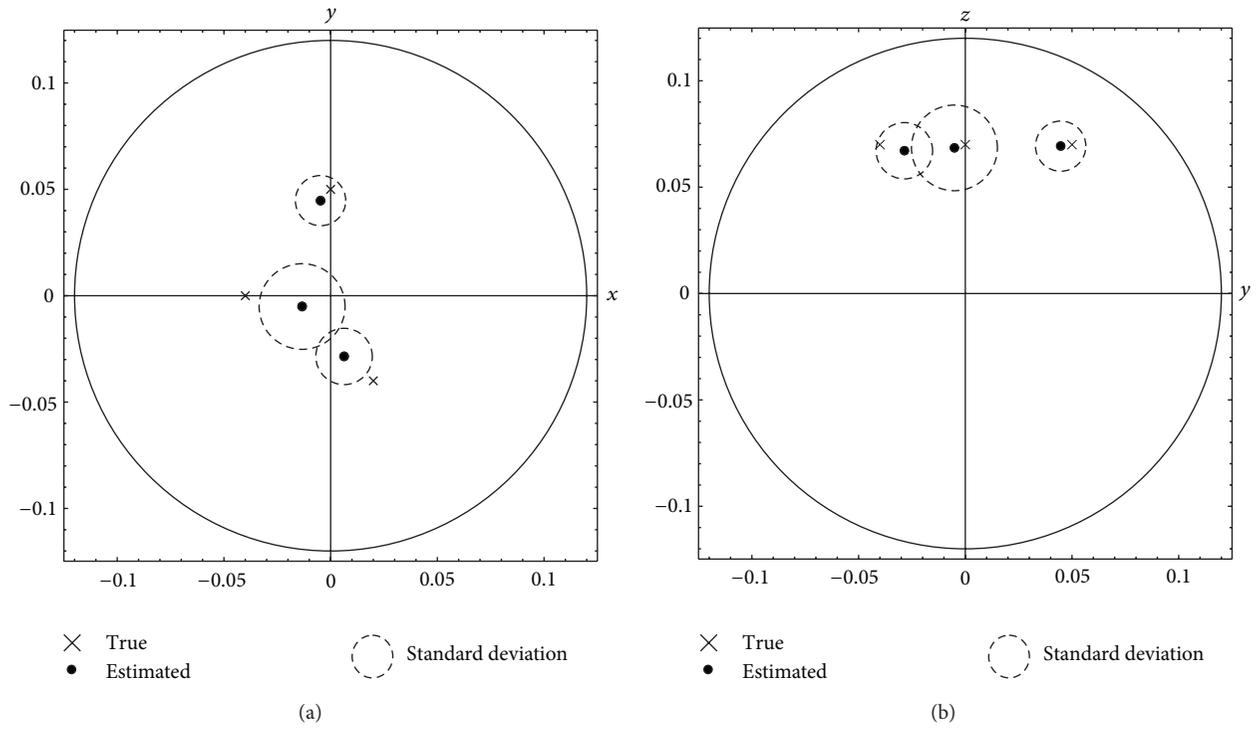


FIGURE 6: Localization result when $N = 3$.

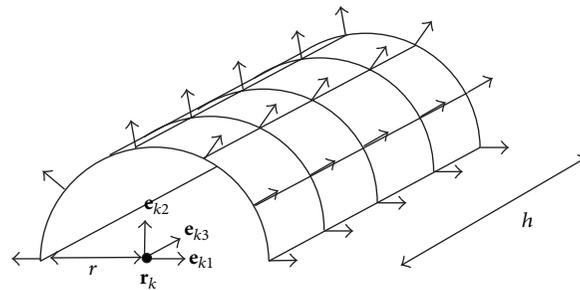


FIGURE 7: Source of distributed dipoles on a half-cylinder.

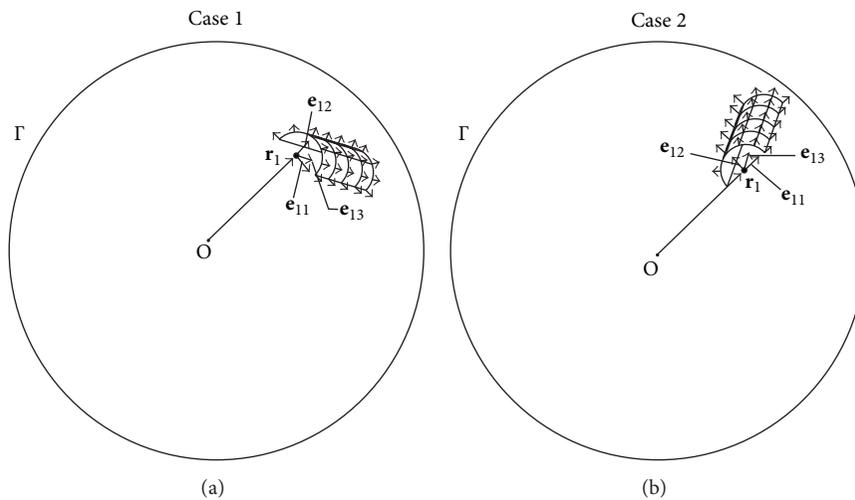


FIGURE 8: Cases 1 and 2 of the cylindrically distributed dipoles.

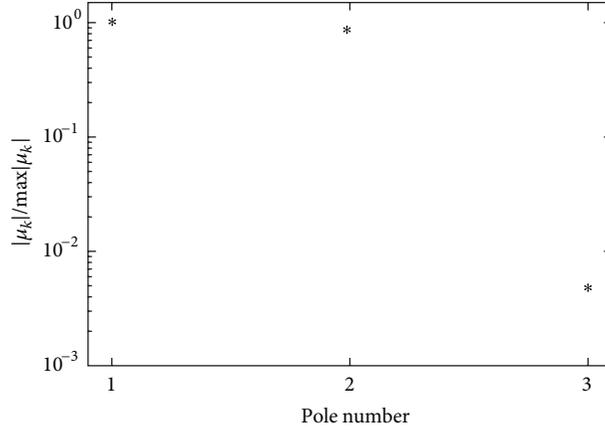


FIGURE 9: Estimation of N in Case 1 using the dipole model (DM) when assuming $N' = 3$. From $|\mu_3/\mu_2| = 0.005$, DM estimates $N = 2$.

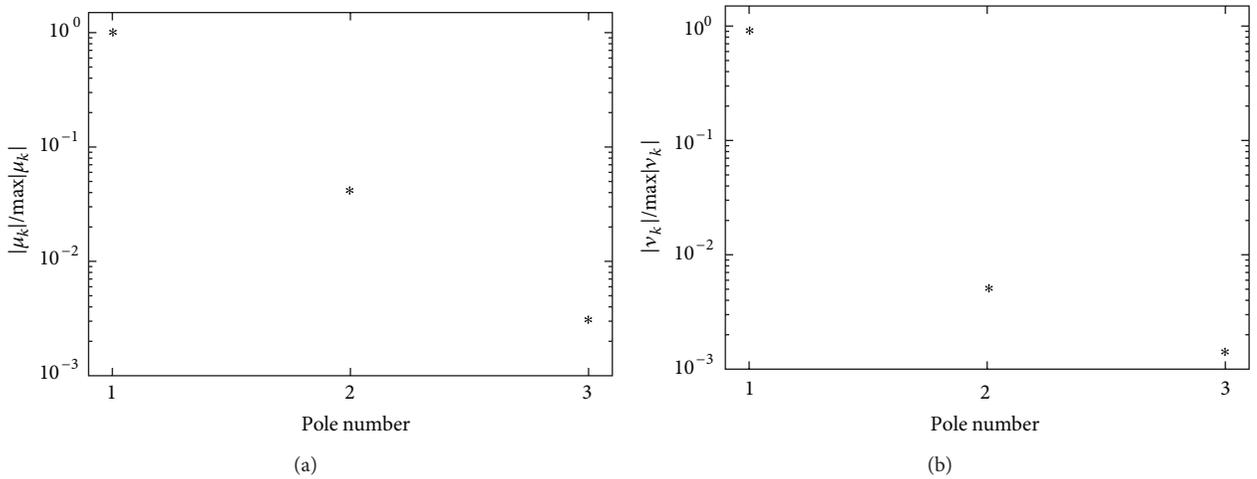


FIGURE 10: Estimation of N in Case 1 using the dipole-quadrupole model (DQM) when assuming $N' = 3$. From $|\mu_3/\mu_2| = 0.041$ and $|\nu_3/\nu_2| = 0.004$, DQM estimates $N = 1$.

the longitudinal direction, and hence a total of 30 dipoles on the half-cylinder. All the dipoles are aligned perpendicular to the surface of the cylinder to model the fact that the dipoles are perpendicular to the cerebral surface. We examined the following two cases.

Case 1. A single half-cylinder source at $\mathbf{r}_1 = (20, 50, 20)$ mm. The vectors which determine the posture of the cylinder, \mathbf{e}_{12} and \mathbf{e}_{11} , are set to be \mathbf{e}_r and \mathbf{e}_θ , respectively, where \mathbf{e}_r and \mathbf{e}_θ are the unit vectors in the r and θ directions at \mathbf{r}_1 . (see Figure 8(a)). In this case, the total dipole moment \mathbf{p}_1 is almost parallel to \mathbf{r}_1 ; the angle between them is 2.4 degrees. Since a radial dipole is a silent source for the radial component of the magnetic field [1], this cylindrical source is regarded as being almost quadrupolar.

Case 2. The same half-cylinder at \mathbf{r}_1 as in Case 1, but with $\mathbf{e}_{12} = -\mathbf{e}_\theta$ and $\mathbf{e}_{11} = \mathbf{e}_r$ (Figure 8(b)). In this case, the angle between \mathbf{r}_1 and \mathbf{p}_1 is 87 degrees, so that the source has a detectable equivalent dipole moment as well as the equivalent quadrupole moment.

We computed the forward solution generated by the 30 elemental dipoles using (1). Note that (11) was not used to compute the theoretical data. $N' = 3$ was assumed to estimate N .

4.5.1. Case 1. Figure 9 shows the estimated $|\mu_k|$ normalized by $\max\{\mu_k\}$ using DM. From the fact that $|\mu_3/\mu_2| = 0.005$, N is estimated to be two using DM. The reason why not a single equivalent dipole but two dipoles are estimated for a single cylindrical source is that the sum of the dipoles on the half-cylindrical surface is directed to the radial direction which is “silent” to the radial magnetic field outside the head. In contrast, Figures 10(a) and 10(b) show the estimated $|\mu_k|$ normalized by $\max\{\mu_k\}$ and $|\nu_k|$ normalized by $\max\{\nu_k\}$, respectively, using DQM. From $|\mu_3/\mu_2| = 0.041$ and $|\nu_3/\nu_2| = 0.004$, N is estimated to be one.

Figure 11 shows the localization result using DM and DQM. It is observed that the two positions estimated with DM were far apart from the cylindrical surface, whereas DQM well estimated the center of the cylinder. This is a great

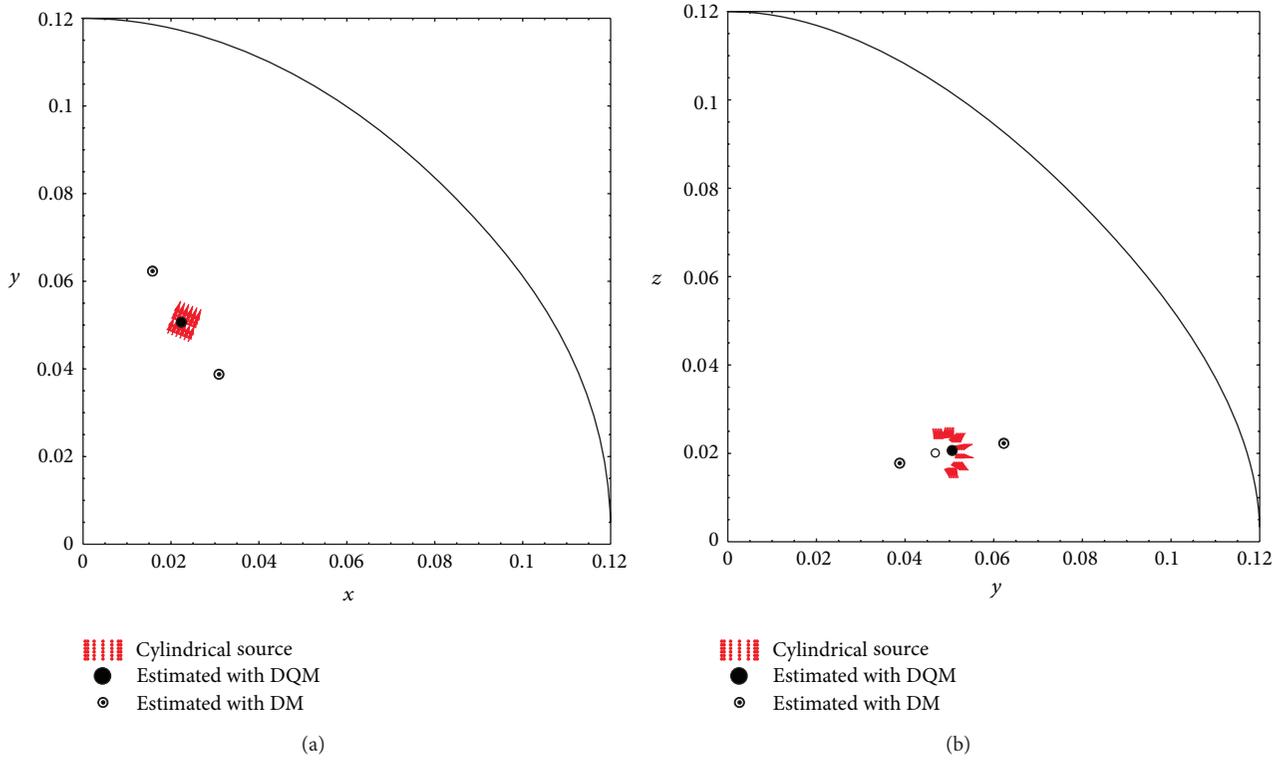


FIGURE 11: Localization result of the cylindrically distributed dipole sources using dipole model (DM) and dipole-quadrupole model (DQM).

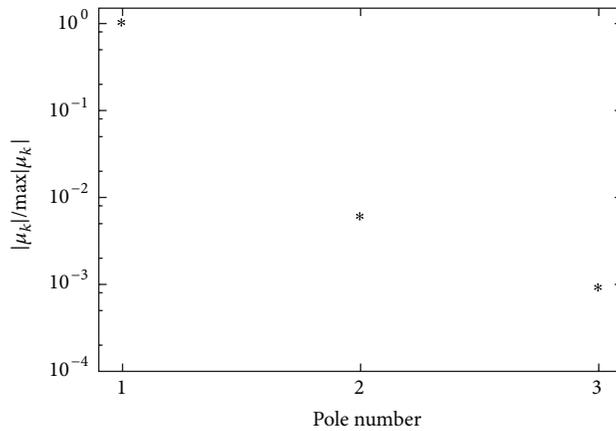


FIGURE 12: Estimation of N in Case 1 using the dipole model (DM) when assuming $N^t = 3$. From $|\mu_3/\mu_2| = 0.006$, DM estimates $N = 1$.

advantage of DQM; although it is difficult for DM to estimate stably two, close, oppositely directed dipoles which are tangential to the spherical surface, it is easy for DQM.

4.5.2. Case 2. Figures 12, 13(a), and 13(b) show the estimated $|\mu_k|$ using DM and $|\mu_k|$ and $|\nu_k|$ using DQM, respectively. In this Case, from $|\mu_2/\mu_1| = 0.006$ in Figure 12 using DM and $|\mu_2/\mu_1| = 0.006$ and $|\nu_2/\nu_1| = 0.073$ in Figure 13 using DQM, both model estimate $N = 1$. This is because the total of the elemental dipoles on the cylinder is almost perpendicular to the radial direction in Case 2 so that this cylindrical source can be regarded as a single equivalent current dipole.

Figure 14 shows the localization result. Both DM and DQM can well estimate the center position of the cylindrical source in this case.

5. Conclusion

We considered an inverse source problem of the Poisson equation for the radial component of the magnetic flux density when the current source is equivalently represented by multiple dipoles and quadrupoles in layered, spherical domains. By expressing the magnetic field in terms of the spherical harmonic expansion, we showed that the sectorial

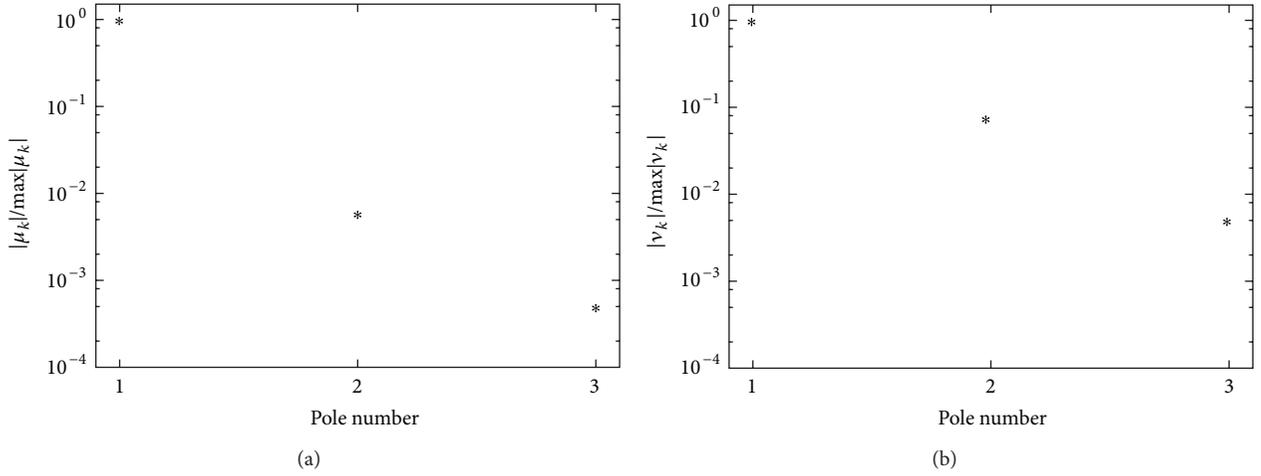


FIGURE 13: Estimation of N in Case 1 using the dipole-quadrupole model (DQM) when assuming $N' = 3$. From $|\mu_3/\mu_2| = 0.006$ and $|\nu_3/\nu_2| = 0.072$, DQM estimates $N = 1$.

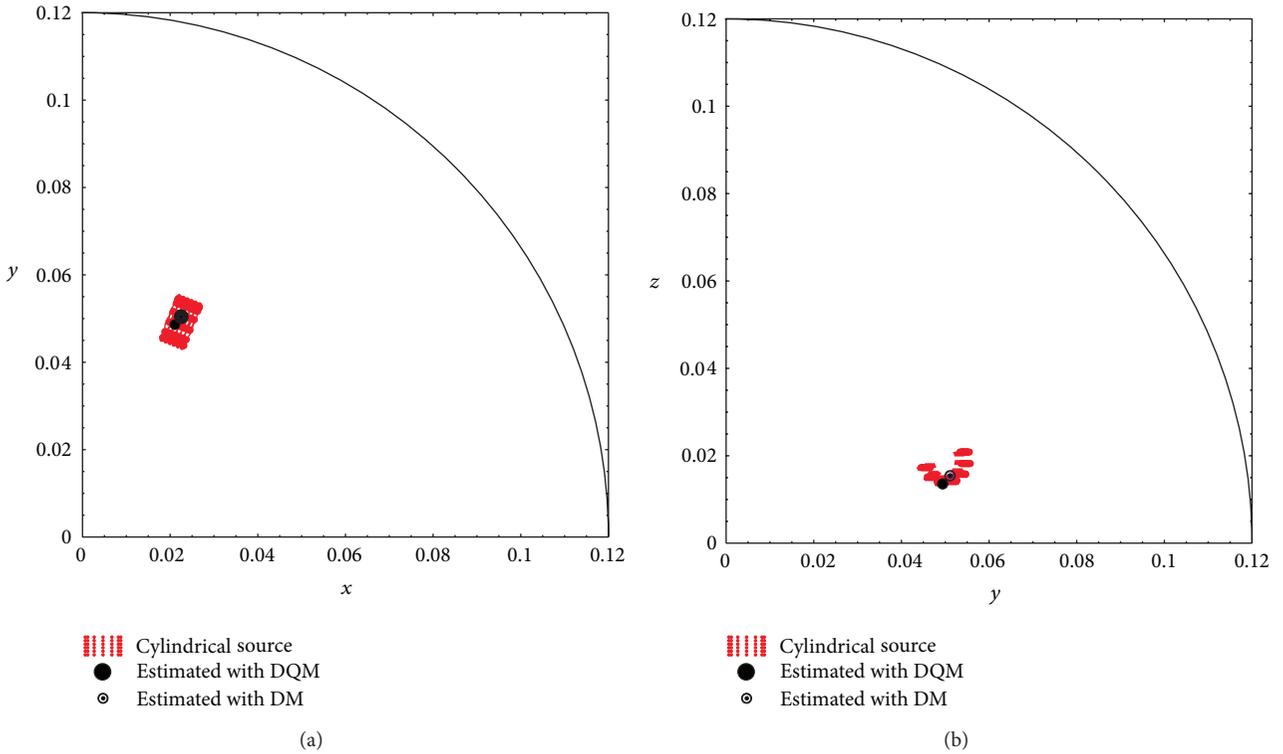


FIGURE 14: Localization result of the cylindrically distributed dipole sources using dipole model (DM) and dipole-quadrupole model (DQM).

harmonics gave the simultaneous algebraic equations relating the dipole and quadrupole parameters to the boundary data. The equations had the same form as derived in the simple- and double-pole reconstruction problem in the 2D case [17], so that the source parameters can be algebraically reconstructed. We proposed a method to obtain the x - and y -coordinates of the dipole-quadrupole source by solving the simultaneous second-degree equations, with the result of which the z -coordinates are also determined. Numerical simulations show that the proposed algorithm estimates the

number and centroid positions of the spatially extended dipoles well, especially when they include elemental dipoles oriented in opposite directions and their equivalent dipole moment is almost parallel to the radial direction of the spherical domain. A theoretical analysis to determine the thresholds depending on the noise in the algorithm for the estimation of the number of the dipole-quadrupole sources as well as an extension of the method to the case when the data is available only for part of the spherical boundary is left for further research.

Appendices

A. Derivation of (6)

Substituting the Taylor expansion of $\nabla'(1/|\mathbf{r} - \mathbf{r}'|)$ into the right-hand side of (3), we have

$$\begin{aligned} \mathbf{r} \cdot \mathbf{B}(\mathbf{r}) &= \frac{\mu_0}{4\pi} \sum_{k=1}^N \int_{D_k} \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}'=\mathbf{r}_k} \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &+ \frac{\mu_0}{4\pi} \sum_{k=1}^N \int_{D_k} \left(\nabla' \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \\ &\cdot \mathbf{j}_p(\mathbf{r}') dv' + \dots \end{aligned} \quad (\text{A.1})$$

The first term of (A.1) is rewritten as

$$\begin{aligned} &\int_{D_k} \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}'=\mathbf{r}_k} \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\frac{\mathbf{r} - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\frac{\mathbf{r} - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} \times \mathbf{r}_k \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \left(\mathbf{r}_k \times \int_{D_k} \mathbf{j}_p(\mathbf{r}') dv' \right) \cdot \frac{\mathbf{r} - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3}. \end{aligned} \quad (\text{A.2})$$

The second term of (A.1) is rewritten as

$$\begin{aligned} &\int_{D_k} \left(\nabla' \nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\nabla' \frac{\mathbf{r} - \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} \right) (\mathbf{r} - \mathbf{r}') \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \\ &\cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &+ \int_{D_k} \left(\frac{\nabla' (\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\frac{3(\mathbf{r} - \mathbf{r}')(\mathbf{r} - \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^5} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \end{aligned}$$

$$\begin{aligned} &+ \int_{D_k} \left(\frac{-I}{|\mathbf{r} - \mathbf{r}'|^3} \Big|_{\mathbf{r}'=\mathbf{r}_k} (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \\ &\cdot \mathbf{j}_p(\mathbf{r}') dv' \quad (I : 3 \times 3 \text{ identity matrix}) \\ &= \int_{D_k} \left(\frac{3(\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^5} (\mathbf{r} - \mathbf{r}_k) \cdot (\mathbf{r}' - \mathbf{r}_k) \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &- \int_{D_k} \left(\frac{\mathbf{r}' - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} \times \mathbf{r} \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &= \int_{D_k} \left(\frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^5} (\mathbf{r}' - \mathbf{r}_k) \right) \cdot \mathbf{j}_p(\mathbf{r}') dv' \\ &+ \int_{D_k} X_r \frac{\mathbf{r}' - \mathbf{r}_k}{|\mathbf{r} - \mathbf{r}_k|^3} \cdot \mathbf{j}_p(\mathbf{r}') dv'. \end{aligned} \quad (\text{A.3})$$

Here, for a 3×3 matrix A whose row vectors are \mathbf{a}_1^T , \mathbf{a}_2^T , and \mathbf{a}_3^T , it holds that

$$(\mathbf{A}\mathbf{x}) \cdot \mathbf{y} = \begin{pmatrix} \mathbf{a}_1^T \mathbf{x} \\ \mathbf{a}_2^T \mathbf{x} \\ \mathbf{a}_3^T \mathbf{x} \end{pmatrix} \cdot \mathbf{y} = \mathbf{a}_1^T \mathbf{x} y_1 + \mathbf{a}_2^T \mathbf{x} y_2 + \mathbf{a}_3^T \mathbf{x} y_3 = A : \mathbf{y}\mathbf{x}. \quad (\text{A.4})$$

Using this property of the tensor contraction, (A.3) is further rewritten as

$$\begin{aligned} &\frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^5} : \int_{D_k} \mathbf{j}_p(\mathbf{r}') (\mathbf{r}' - \mathbf{r}_k) dv' + \frac{X_r}{|\mathbf{r} - \mathbf{r}_k|^3} \\ &: \int_{D_k} \mathbf{j}_p(\mathbf{r}') (\mathbf{r}' - \mathbf{r}_k) dv' \\ &= \frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k) + |\mathbf{r} - \mathbf{r}_k|^2 X_r}{|\mathbf{r} - \mathbf{r}_k|^5} \\ &: \int_{D_k} \mathbf{j}_p(\mathbf{r}') (\mathbf{r}' - \mathbf{r}_k) dv'. \end{aligned} \quad (\text{A.5})$$

Therefore, substituting (A.2) and (A.5) into (3) gives (6).

B.

This can be shown as follows. It is easy to obtain the dipole terms. For the quadrupole terms, we use the identity:

$$\int_{\Omega_1} \mathbf{a}(\mathbf{r}') \cdot Q_k \nabla' \delta(\mathbf{r}' - \mathbf{r}_k) dv' = Q : (\nabla' \mathbf{a}(\mathbf{r}'))^T \Big|_{\mathbf{r}'=\mathbf{r}_k} \quad (\text{B.1})$$

for an arbitrary vector field $\mathbf{a}(\mathbf{r}') = (a_x(\mathbf{r}'), a_y(\mathbf{r}'), a_z(\mathbf{r}'))^T$, where T represents the transpose. When inserting

the quadrupole terms in (12) into (1), we have from (B.1)

$$\begin{aligned} & \int_{\Omega_1} \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r}' \right) \cdot \mathbf{Q}_k \nabla' \delta(\mathbf{r}' - \mathbf{r}_k) d\mathbf{v}' \\ &= \mathbf{Q}_k : \left(\nabla' \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r}' \right) \right)^T \Big|_{\mathbf{r}' = \mathbf{r}_k}. \end{aligned} \quad (\text{B.2})$$

Here, it holds that

$$\begin{aligned} & \nabla' \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r}' \right) \\ &= \nabla' \frac{\mathbf{r} \times \mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|^3} = \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} \right) (\mathbf{r} \times \mathbf{r}') + \frac{\nabla' (\mathbf{r} \times \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^3} \\ &= \frac{3(\mathbf{r} - \mathbf{r}')(\mathbf{r} \times \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|^5} + \frac{1}{|\mathbf{r} - \mathbf{r}'|^3} \begin{pmatrix} 0 & z & -y \\ -z & 0 & x \\ y & -x & 0 \end{pmatrix}, \end{aligned} \quad (\text{B.3})$$

and hence

$$\begin{aligned} & \left(\nabla' \left(\nabla' \frac{1}{|\mathbf{r} - \mathbf{r}'|} \times \mathbf{r}' \right) \right)^T \Big|_{\mathbf{r}' = \mathbf{r}_k} \\ &= \frac{3(\mathbf{r} \times \mathbf{r}_k)(\mathbf{r} - \mathbf{r}_k)}{|\mathbf{r} - \mathbf{r}_k|^5} + \frac{X_r}{|\mathbf{r} - \mathbf{r}_k|^3}. \end{aligned} \quad (\text{B.4})$$

Thus we have (11).

C. Derivation of (21)

It holds that

$$\begin{aligned} & \mathbf{Q}_k : \left(\nabla (\nabla(x + iy)^n \times \mathbf{r}) \right)^T \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= \begin{pmatrix} q_{xx,k} & q_{xy,k} & q_{xz,k} \\ q_{xy,k} & q_{yy,k} & q_{yz,k} \\ q_{xz,k} & q_{yz,k} & q_{zz,k} \end{pmatrix} \\ &: n \begin{pmatrix} i(n-1)z(x+iy)^{n-2} & -(n-1)z(x+iy)^{n-2} & i(x+iy)^{n-1} \\ -(n-1)z(x+iy)^{n-2} & -i(n-1)z(x+iy)^{n-2} & -(x+iy)^{n-1} \\ -in(x+iy)^{n-1} & n(x+iy)^{n-1} & 0 \end{pmatrix} \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= n \left(i(n-1)q_{xx,k}z_k(x_k + iy_k)^{n-2} - (n-1)q_{xy,k}z_k(x_k + iy_k)^{n-2} \right. \\ &\quad \left. + iq_{xz,k}(x_k + iy_k)^{n-1} - (n-1)q_{xy,k}z_k(x_k + iy_k)^{n-2} \right. \\ &\quad \left. - i(n-1)q_{yy,k}z_k(x_k + iy_k)^{n-2} - q_{yz,k}(x_k + iy_k)^{n-1} \right. \\ &\quad \left. - inq_{xz,k}(x_k + iy_k)^{n-1} + nq_{yz,k}(x_k + iy_k)^{n-1} \right) \\ &= n(n-1) \left((iq_{xx,k} - q_{yy,k}) - 2q_{xy,k} \right) z_k \\ &\quad + (q_{yz,k} - iq_{xz,k})(x_k + iy_k)(x_k + iy_k)^{n-2}, \end{aligned} \quad (\text{C.1})$$

which gives (21).

D. Derivation of (42) and (43)

It is easy to see that

$$\begin{aligned} & \nabla (\nabla \phi_j(x + iy) \times \mathbf{r}) \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= \begin{pmatrix} iz\phi_j'' & -z\phi_j'' & -i(x+iy)\phi_j'' - i\phi_j' \\ -z\phi_j'' & -iz\phi_j'' & (x+iy)\phi_j'' + \phi_j' \\ i\phi_j' & -\phi_j' & 0 \end{pmatrix} \Big|_{\mathbf{r} = \mathbf{r}_k}, \end{aligned} \quad (\text{D.1})$$

where ϕ_j , ϕ_j' , and ϕ_j'' represent $\phi_j(x+iy)$, $\phi_j'(x+iy)$, and $\phi_j''(x+iy)$, respectively. From the definition of ϕ_j in (37), one easily finds that $\phi_j(\zeta_k) = \phi_j'(\zeta_k) = 0$ for all $k = 1, 2, \dots, N$, and hence

$$\begin{aligned} & \nabla (\nabla \phi_j(x + iy) \times \mathbf{r}) \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= \begin{pmatrix} iz_k\phi_j''(\zeta_k) & -z_k\phi_j''(\zeta_k) & -i\zeta_k\phi_j''(\zeta_k) \\ -z_k\phi_j''(\zeta_k) & -iz_k\phi_j''(\zeta_k) & \zeta_k\phi_j''(\zeta_k) \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (\text{D.2})$$

Also, from the definition of ϕ_j , we have

$$\phi_j''(\zeta_k) = 2\delta_{jk} \prod_{l=1, l \neq j}^N (\zeta_j - \zeta_l)^3, \quad (\text{D.3})$$

where δ_{jk} is Kronecker's delta, which gives (42). In the same way,

$$\begin{aligned} & \nabla (\nabla \phi_j(x + iy) z \times \mathbf{r}) \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= \begin{pmatrix} -y\phi_j' + iz^2\phi_j'' & \phi_j + x\phi_j' - z^2\phi_j'' & -i(x+iy)z\phi_j'' - iz\phi_j' \\ -\phi_j - iy\phi_j' - z^2\phi_j'' & ix\phi_j' - iz^2\phi_j'' & (x+iy)z\phi_j'' + z\phi_j' \\ 2iz\phi_j' & -2z\phi_j' & -i(x+iy)\phi_j' \end{pmatrix} \Big|_{\mathbf{r} = \mathbf{r}_k} \\ &= z_k \begin{pmatrix} iz_k\phi_j''(\zeta_k) & -z_k\phi_j''(\zeta_k) & -i\zeta_k\phi_j''(\zeta_k) \\ -z_k\phi_j''(\zeta_k) & -iz_k\phi_j''(\zeta_k) & \zeta_k\phi_j''(\zeta_k) \\ 0 & 0 & 0 \end{pmatrix}, \end{aligned} \quad (\text{D.4})$$

which gives (43).

Acknowledgments

This research has been partially supported by the Moritani Scholarship Foundation and the 2012 Inamori Research Grants Program.

References

- [1] J. Sarvas, "Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem," *Physics in Medicine and Biology*, vol. 32, no. 1, article 4, pp. 11–22, 1987.
- [2] S. Baillet, J. C. Mosher, and R. M. Leahy, "Electromagnetic brain mapping," *IEEE Signal Processing Magazine*, vol. 18, no. 6, pp. 14–30, 2001.
- [3] G. Dassios and D. Hadjiloizi, "On the non-uniqueness of the inverse problem associated with electroencephalography," *Inverse Problems*, vol. 25, no. 11, Article ID 115012, 18 pages, 2009.
- [4] A. S. Fokas, Y. Kurylev, and V. Marinakis, "The unique determination of neuronal currents in the brain via magnetoencephalography," *Inverse Problems*, vol. 20, no. 4, pp. 1067–1082, 2004.
- [5] M. Huang, C. J. Aine, S. Supek, E. Best, D. Ranken, and E. R. Flynn, "Multi-start downhill simplex method for spatio-temporal source localization in magnetoencephalography," *Electroencephalography and Clinical Neurophysiology*, vol. 108, no. 1, pp. 32–44, 1998.
- [6] T. Ohe and K. Ohnaka, "A precise estimation method for locations in an inverse logarithmic potential problem for point mass models," *Applied Mathematical Modelling*, vol. 18, no. 8, pp. 446–452, 1994.
- [7] A. El Badia and T. Ha-Duong, "An inverse source problem in potential analysis," *Inverse Problems*, vol. 16, no. 3, pp. 651–663, 2000.
- [8] K. Yamatani, T. Ohe, and K. Ohnaka, "An identification method of electric current dipoles in spherically symmetric conductor," *Journal of Computational and Applied Mathematics*, vol. 143, no. 2, pp. 189–200, 2002.
- [9] T. Nara and S. Ando, "A projective method for an inverse source problem of the Poisson equation," *Inverse Problems*, vol. 19, no. 2, pp. 355–369, 2003.
- [10] T. Nara, J. Oohama, M. Hashimoto, T. Takeda, and S. Ando, "Direct reconstruction algorithm of current dipoles for vector magnetoencephalography and electroencephalography," *Physics in Medicine and Biology*, vol. 52, no. 13, article 14, pp. 3859–3879, 2007.
- [11] F. Darvas, D. Pantazis, E. Kucukaltun-Yildirim, and R. M. Leahy, "Mapping human brain function with MEG and EEG: methods and validation," *NeuroImage*, vol. 23, no. 1, pp. S289–S299, 2004.
- [12] K. Matsuura and Y. Okabe, "Selective minimum-norm solution of the biomagnetic inverse problem," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 6, pp. 608–615, 1995.
- [13] G. Nolte and G. Curio, "On the calculation of magnetic fields based on multipole modeling of focal biological current sources," *Biophysical Journal*, vol. 73, no. 3, pp. 1253–1262, 1997.
- [14] G. Nolte and G. Curio, "Current multipole expansion to estimate lateral extent of neuronal activity: a theoretical analysis," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 10, pp. 1347–1355, 2000.
- [15] K. Jerbi, J. C. Mosher, S. Baillet, and R. M. Leahy, "On MEG forward modelling using multipolar expansions," *Physics in Medicine and Biology*, vol. 47, no. 4, pp. 523–555, 2002.
- [16] K. Jerbi, S. Baillet, J. C. Mosher, G. Nolte, L. Garnero, and R. M. Leahy, "Localization of realistic cortical activity in MEG using current multipoles," *NeuroImage*, vol. 22, no. 2, pp. 779–793, 2004.
- [17] T. Nara, "An algebraic method for identification of dipoles and quadrupoles," *Inverse Problems*, vol. 24, no. 2, Article ID 025010, 19 pages, 2008.
- [18] A. Irimia, K. R. Swinney, and J. P. Wikswo, "Partial independence of bioelectric and biomagnetic fields and its implications for encephalography and cardiography," *Physical Review E*, vol. 79, no. 5, Article ID 051908, 2009.
- [19] T. Nara, "Reconstruction of the number and positions of dipoles and quadrupoles using an algebraic method," *Journal of Physics: Conference Series*, vol. 135, no. 1, Article ID 012076, 2008.
- [20] T. Nara, "An explicit method for inverse reconstruction of equivalent current dipoles and quadrupoles," in *Magnetoencephalography*, E. W. Pang, Ed., Chapter 6, pp. 113–128, InTech, 2011.
- [21] R. E. Alvarez, "Filter functions for computing multipole moments from the magnetic field normal to a plane," *IEEE Transactions on Medical Imaging*, vol. 10, no. 3, pp. 375–381, 1991.
- [22] S. Taulu, J. Simola, and M. Kajola, "Applications of the signal space separation method," *IEEE Transactions on Signal Processing*, vol. 53, no. 9, pp. 3359–3372, 2005.
- [23] E. W. Hobson, *The Theory of Spherical and Ellipsoidal Harmonics*, Cambridge University Press, 1931.
- [24] P. Barone and R. March, "Some properties of the asymptotic location of poles of Padé approximants to noisy rational functions, relevant for modal analysis," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2448–2457, 1998.
- [25] E. B. Saff and A. B. J. Kuijlaars, "Distributing many points on a sphere," *The Mathematical Intelligencer*, vol. 19, no. 1, pp. 5–11, 1997.
- [26] T. Nara and S. Ando, "Direct localization of poles of a meromorphic function from measurements on an incomplete boundary," *Inverse Problems*, vol. 26, no. 1, Article ID 015011, 2010.

Research Article

Estimation of Open Boundary Conditions for an Internal Tidal Model with Adjoint Method: A Comparative Study on Optimization Methods

Haibo Chen,^{1,2} Chunbao Miao,¹ and Xianqing Lv¹

¹Laboratory of Physical Oceanography, Ocean University of China, Qingdao 266100, China

²Research and Development Center for Offshore Oil Safety and Environmental Technology, China Offshore Environmental Services Ltd., Qingdao 266061, China

Correspondence should be addressed to Xianqing Lv; xqinglv@ouc.edu.cn

Received 16 October 2012; Revised 7 January 2013; Accepted 10 January 2013

Academic Editor: Fatih Yaman

Copyright © 2013 Haibo Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Based on an internal tidal model, the practical performances of the limited-memory BFGS (L-BFGS) method and two gradient descent (GD) methods (the normal one with Wolfe's line search and the simplified one) are investigated computationally through a series of ideal experiments in which the open boundary conditions (OBCs) are inverted by assimilating the interior observations with the adjoint method. In the case that the observations closer to the unknown boundary are included for assimilation, the L-BFGS method performs the best. As compared with the simplified GD method, the normal one really uses less iteration to reach a satisfactory solution, but its advantage over the simplified one is much smaller than expected. In the case that only the observations that are further from the unknown boundary are assimilated, the simplified GD method performs the best instead, whereas the performances of the other two methods are not satisfactory. The advanced L-BFGS algorithm and Wolfe's line search still need to be improved when applied to the practical cases. The simplified GD method, which is controllable and easy to implement, should be regarded seriously as a choice, especially when the classical advanced optimization techniques fail or perform poorly.

1. Introduction

In the recent years, with the rapid advances in ocean observing systems, increasingly more oceanic observation data have become available. At the same time, mainframe supercomputers have become more powerful, with improvements in memory size, the introduction of multiprocessor capabilities as well as enhanced processor speed. These advances have provided us with an opportunity to improve the precision of numerical simulations of the ocean with these available observation data, for both operational and research purposes. Indeed, this is an active topic of a fast-growing research field, data assimilation, which is an effective method for marine research, and it has become widely used in meteorological and oceanographic predictions in recent years. Among all data assimilation methods, four-dimensional variational (4D-Var) data assimilation is often considered as one of the most effective and powerful approaches developed over

the past two decades. It is an advanced data assimilation method that involves the adjoint technique. This method has the advantage of directly assimilating various observations distributed in time and space into the numerical model and can maintain dynamical and physical consistency with the model at the same time. Thus, 4D-Var has been widely applied to meteorological and oceanographic data assimilation, sensitivity studies, and parameter estimation. The issue of 4D-Var applied to the shallow water equations model has been the subject of numerous works (see, e.g., [1–29] and references therein).

In this paper, we are interested in solving a large-scale optimization problem minimizing a cost functional

$$\min_{\mathbf{p}} J(\mathbf{p}) = \min_{\mathbf{p}} \left\{ \frac{1}{2} \int_{t_i}^{t_f} [C(\mathbf{x}) - \hat{\mathbf{x}}]^T K_{\mathbf{x}} [C(\mathbf{x}) - \hat{\mathbf{x}}] dt \right\}, \quad (1)$$

where \mathbf{p} represents the control variables and is defined by the open boundary conditions (OBCs) for a regional internal tidal model, which are denoted by a series of Fourier coefficients in the present work, t is the time, $[t_i, t_f]$ denotes the assimilation window where t_i and t_f are the initial and final time, respectively, and $\mathbf{x}(t)$ and $\hat{\mathbf{x}}(t)$ are the time-dependent state variable in an n -dimensional Euclidean space \mathbf{R}^n and the time-dependent observation variable in an m -dimensional Euclidean space \mathbf{R}^m , respectively. The operator $C : \mathbf{R}^n \rightarrow \mathbf{R}^m$ represents a projection from the n -dimensional space (\mathbf{R}^n) of the model solution \mathbf{x} to the m -dimensional space (\mathbf{R}^m) of observations. Superscript T indicates transpose. The components of \mathbf{x} are values of the various model fields (elevation, velocity, temperature, salinity, etc., alone or in combination) at each of the model grid locations. The number of components of \mathbf{x} is denoted by n , and the number m represents the number of observations at any given time. In general, n is larger than m . In fact, the objective function J is the weighted sum of squares of distance between the model solution and available observations distributed in space and time. $K_{\mathbf{x}}$ is the weighting matrix and theoretically should be the inverse of observation error covariance matrix [1, 2]. However, determining the “exact” form for $K_{\mathbf{x}}$ is far from easy [3]. Since this work represents a preliminary study in the application of an adjoint assimilation model, we proceed under the best and even unrealistic scenario by assuming that there are no observation errors and that the model solutions will have a perfect fit to the observations such that $\min_{\mathbf{p}} J(\mathbf{p}) = 0$. Therefore, $K_{\mathbf{x}}$ is simplified to be unit matrix [30].

The control variables \mathbf{p} and the state vector \mathbf{x} satisfy the time-dependent model equations of the form

$$\frac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}, \mathbf{p}), \quad (2)$$

where $F : \mathbf{R}^n \rightarrow \mathbf{R}^n$ is model vector function of \mathbf{x} . In general, the model equations (2) are nonlinear and are assumed to be twice continuously differentiable.

The objective of the adjoint assimilation method is to find model control variables \mathbf{p} that minimize the cost functional J . Note that J is only a function of the control variables \mathbf{p} because \mathbf{x} is uniquely defined by the model equations (2) for any prescribed \mathbf{p} . The control variables can be initial conditions, boundary conditions or parameters to be estimated, or combination of them. In a regional ocean model, the open boundary conditions, which must be prescribed to complete the model description at non-land boundaries, are very important and have a critical impact on the modeling results. Therefore, for simplicity, only OBCs are taken as control variables in this paper. A number of studies have been concerned about the OBC estimation with adjoint method. For example, the work of Lardner et al. [4] discussed the optimal control of OBC in the channel using a 2D adjoint tidal model. Seiler [5] performed a series of identical twin assimilation experiments using the adjoint method to estimate lateral open boundary values of stream function and relative vorticity for a quasigeostrophic open ocean model. Zou et al. [2] also developed a sequential open boundary

control scheme augmenting radiation conditions and applied it to idealized barotropic wind-driven ocean simulations. ten Brummelhuis et al. [6] used the data assimilation technique to estimate the OBCs for a shallow sea model of the entire European continental shelf. Heemink et al. [7] used the adjoint approach for a 3D shallow water flow system (TRIWAQ) to estimate the harmonic constants in the OBCs, the friction parameter, viscosity parameter, and water depth by assimilating tide gauge data as well as altimeter data. Shulman [8] proposed a data assimilation approach to specify open boundary conditions, and their method was then tested by Shulman et al. [9] with respect to the improvement of the model prediction skills of the M_2 tidal amplitudes and phases in the Yellow Sea. Taillandier et al. [10] used a four-dimensional variational data assimilation method to control boundary conditions in a nonstratified open coastal model. In the work of Zhang et al. [11], lateral tidal OBCs that forced tides in the internal region were estimated by assimilating predicted coastal tidal elevations into a 2D POM with the adjoint method. Gejadze and Copeland [12] studied the open boundary control problem for free-surface barotropic Navier-Stokes equations with adjoint data assimilation technology. By assimilating the tidal harmonic constants derived from T/P altimeter data with a 3D numerical barotropic adjoint tidal model constructed in [13], Zhang and Lu [14] optimized the OBCs and simulated the M_2 tide and tidal current in the Bohai and Yellow Seas. Based on a 2D tidal model, Cao et al. [15] and Guo et al. [16] studied the inversion of OBCs by assimilating the T/P altimeter data with adjoint method. Kazantsev [17] used a variational data assimilation technique to control boundary conditions on rigid boundaries for a shallow-water model. Recently, in the work of Chen et al. [18], an adjoint assimilation model for the simulation of internal tides was constructed and simply tested with a series of ideal experiments in which several prescribed spatial distributions of OBCs were successfully inverted by assimilating the model-generated pseudoobservations.

The previous minimization problem (1) requires large-scale optimization techniques. This problem attempts to find a solution of the model equations (2) that best fits, in the generalized least-squares sense, the observations distributed in space and time. In meteorology and oceanography, the fluid dynamics can be enforced through the use of a Lagrangian function constructed by appending the model equations to the cost function as constraints. To solve problem (1), many feasible large-scale optimization methods can be found in [31]. However, the number of studies discussing the performances of various optimization methods in the meteorological and oceanographic application is still relatively small. Wang et al. [19] proposed an adjoint truncated Newton algorithm for the large-scale unconstrained minimization in the meteorology application and applied this algorithm to a limited-area shallow-water equation model with model-generated data where the initial conditions served as control variables. Subsequently, Wang et al. [20] presented a new adjoint Newton algorithm for carrying out large-scale unconstrained optimization required in variational data assimilation and applied this algorithm to three 1D models and to one 2D limited-area shallow-water equation model

with both model-generated and First Global Geophysical Experiment data. Zou et al. [21] compared the performances of several limited-memory quasi-Newton and truncated Newton methods for unconstrained nonlinear optimization on some large-scale problems in oceanography and meteorology. Their results showed that among the tested limited-memory quasi-Newton methods, the limited-memory BFGS (L-BFGS) method of Liu and Nocedal [32] has the best overall performance for the problems examined, and the numerical performance of two truncated Newton methods, differing in the inner-loop solution for the search vector, is competitive with that of L-BFGS. More recently, Alekseev et al. [33] compared the performance of several advanced large-scale minimization algorithms, including the conjugate gradient method, quasi-Newton (BFGS), the limited-memory quasi-Newton (L-BFGS), Hessian Free Newton method, and a new hybrid algorithm, applied to the minimization of the cost functional in the solution of ill-posed inverse computational fluid dynamics problems related to parameter estimation. In the work of [14], the comparison of a simple gradient descent (GD) method with the L-BFGS algorithm was conducted within a 3D barotropic tidal model, and the results demonstrate that, compared with the steepest descent method, the L-BFGS really uses much fewer steps to reach a minimum, while the values of the minimized cost function and inverted Fourier coefficient are almost the same. In the work of [18], the L-BFGS version of Liu and Nocedal [32], combined with the adjoint assimilation technique, was applied to the optimization of the OBCs for an internal tidal model and was proved to be very efficient.

The main objective of this paper is to make a computational investigation on the performances of the L-BFGS method and two versions of GD method. All these methods do not require any evaluations of the Hessian matrices but gradient vectors and thus are computationally feasible. In the present work, the number of the control variables (OBCs) is often large. In this case, the limited-memory methods are suggested, and one of the most commonly used limited-memory methods is the L-BFGS method [31, 32]. The L-BFGS method is an adaptation of the standard BFGS method to large-scale optimization problems. This method has been proved to be globally convergent for convex objective functions and provides a fast rate of linear convergence as well as requiring minimal storage. As the competitor, we choose the GD method. The GD method is a fundamental and probably the earliest method for nonlinear optimization so that it attracts many researchers' attention since its proposition. A recent and interesting report on the GD method can be found in [31, 34]. The GD method has the negative gradient direction as the search direction and is one of the simplest minimization algorithms that require calculation of derivatives. The GD method is particularly useful when the dimension of the problem is very large [35]. Two versions of GD method are considered here, differing in the formulation of the step length. One normally uses the step length satisfying the strong Wolfe conditions [31, 36, 37] which have been widely used in line search (see [38, 39] and references therein). And the other is a simplified GD method without the line search which has been used in a relatively

small number of studies [13, 14, 22, 23]. In this paper, based on the model constructed by Chen et al. [18], the performances of the L-BFGS method and the two versions of GD method are compared through a series of ideal experiments in which the OBCs for a numerical internal tidal model are estimated by assimilating the interior observations with the adjoint method.

This paper is organized as follows. We start with a brief introduction of the two-layered internal tidal model in Section 2. For the details of the model, we refer interested readers to [18]. The optimization methods for comparison, including the L-BFGS method and two GD methods, are described briefly in Section 3. In Section 4, as illustrative numerical examples, a series of ideal experiments are carried out to invert the prescribed OBCs with the different methods examined here, and the experimental results are presented. Finally, we make a summary and draw some conclusions in Section 5.

2. Internal Tidal Model

2.1. Forward Model. A two-layer version of the numerical internal tidal model described in [18] is considered in this paper. Assuming that the potential density ρ_k ($k = 1, 2$) in each layer is constant, the layer-averaged, nonlinear, time-dependent continuity and momentum equations of each layer subject to the hydrostatic approximations are derived from the primitive 3D governing equations. Here, the subscripts 1 and 2 refer to the upper and lower layers, respectively. Using spherical coordinates in the horizontal direction and isopycnic coordinates in the vertical one, we obtain the internal mode equations as follows:
upper layer:

$$\frac{\partial q_1}{\partial t} + \frac{1}{R \cos \phi} \frac{\partial (q_1 u_1)}{\partial \lambda} + \frac{1}{R \cos \phi} \frac{\partial (q_1 v_1 \cos \phi)}{\partial \phi} = 0, \quad (3a)$$

$$\begin{aligned} \frac{\partial u_1}{\partial t} + \frac{u_1}{R \cos \phi} \frac{\partial u_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial u_1}{\partial \phi} - \frac{u_1 v_1 \tan \phi}{R} \\ - f v_1 + \frac{g}{R \cos \phi} \frac{\partial}{\partial \lambda} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) \\ - A_h \Delta u_1 + \frac{A_v \bar{\rho}}{q_1 \bar{h}} (u_1 - u_2) = 0, \end{aligned} \quad (3b)$$

$$\begin{aligned} \frac{\partial v_1}{\partial t} + \frac{u_1}{R \cos \phi} \frac{\partial v_1}{\partial \lambda} + \frac{v_1}{R} \frac{\partial v_1}{\partial \phi} + \frac{u_1^2 \tan \phi}{R} \\ + f u_1 + \frac{g}{R} \frac{\partial}{\partial \phi} \sum_{m=1}^l \left(\frac{q_m}{\rho_m} - h_m \right) \\ - A_h \Delta v_1 + \frac{A_v \bar{\rho}}{q_1 \bar{h}} (v_1 - v_2) = 0; \end{aligned} \quad (3c)$$

lower layer:

$$\frac{\partial q_2}{\partial t} + \frac{1}{R \cos \phi} \frac{\partial (q_2 u_2)}{\partial \lambda} + \frac{1}{R \cos \phi} \frac{\partial (q_2 v_2 \cos \phi)}{\partial \phi} = 0, \quad (4a)$$

$$\begin{aligned} & \frac{\partial u_2}{\partial t} + \frac{u_2}{R \cos \phi} \frac{\partial u_2}{\partial \lambda} + \frac{v_2}{R} \frac{\partial u_2}{\partial \phi} - \frac{u_2 v_2 \tan \phi}{R} \\ & - f v_2 + \frac{g}{R \cos \phi} \frac{\partial}{\partial \lambda} \sum_{k=1}^2 \left(\frac{q_k}{\rho_2} - h_k \right) \end{aligned} \quad (4b)$$

$$\begin{aligned} & - A_h \Delta u_2 - \frac{A_v \bar{\rho}}{q_2 h} (u_1 - u_2) \\ & + \frac{\kappa \rho_2 u_2}{q_2} \sqrt{(u_2)^2 + (v_2)^2} = 0, \end{aligned}$$

$$\begin{aligned} & \frac{\partial v_2}{\partial t} + \frac{u_2}{R \cos \phi} \frac{\partial v_2}{\partial \lambda} + \frac{v_2}{R} \frac{\partial v_2}{\partial \phi} + \frac{u_2^2 \tan \phi}{R} \\ & + f u_2 + \frac{g}{R} \frac{\partial}{\partial \phi} \sum_{k=1}^2 \left(\frac{q_k}{\rho_2} - h_k \right) \end{aligned} \quad (4c)$$

$$\begin{aligned} & - A_h \Delta v_2 - \frac{A_v \bar{\rho}}{q_2 h} (v_1 - v_2) \\ & + \frac{\kappa \rho_2 v_2}{q_2} \sqrt{(u_2)^2 + (v_2)^2} = 0. \end{aligned}$$

Here, t is the time, λ and ϕ are the eastern longitude and northern latitude, respectively, $u(\lambda, \phi, t)$ and $v(\lambda, \phi, t)$ are horizontal velocities in λ and ϕ , respectively, $q(\lambda, \phi, t)$ is the time-varying layer mass, and

$$q_1 = \rho_1 (h_1 + \eta_1 - \eta_2), \quad q_2 = \rho_2 (h_2 + \eta_2), \quad (5)$$

where h_k and η_k are the undisturbed layer thickness and interface (surface for $k = 1$) elevation above the undisturbed level, respectively. R is the radius of the earth, g the gravitational acceleration, f the Coriolis parameter, and $f = 2\Omega \sin \phi$, where Ω represents the angular speed of Earth's rotation, A_h the horizontal eddy viscosity coefficient, Δ the Laplace operator, and

$$\Delta(u, v) = \frac{1}{R^2 \cos^2 \phi} \frac{\partial^2 (u, v)}{\partial \lambda^2} + \frac{1}{R^2 \cos \phi} \frac{\partial}{\partial \phi} \left[\cos \phi \frac{\partial (u, v)}{\partial \phi} \right]. \quad (6)$$

A_v and κ are the interface and bottom friction coefficient, respectively, $\bar{\rho} = (\rho_1 + \rho_2)/2$, and $\bar{h} = (h_1 + h_2)/2$. In the forward model q , u and v are the main outputs and called the state variables in this paper.

By defining the barotropic currents

$$\begin{aligned} U &= \frac{(q_1 u_1 + q_2 u_2)}{Q}, \\ V &= \frac{(q_1 v_1 + q_2 v_2)}{Q}, \quad \text{where } Q = q_1 + q_2, \end{aligned} \quad (7)$$

we can derive a 2D external mode from the internal mode equations. For further details about the external mode, we refer the interested reader to the previous work of [18].

Usually, the model domain Σ is defined within a certain range of space and time, enclosed by the initial and boundary conditions. In the present model, a zero field is assigned to the model state variables as the initial conditions. The model is run for several tidal cycles, and the simulation results in the last cycle are taken as the model output. The boundary located in the wet grid is treated as the open boundary, and the boundary located in the dry grid (e.g., the island and the land) is treated as the closed boundary. The details of the boundary conditions are described as follows.

The Flather condition which was originally proposed by Flather [40] yielded good results [18, 41] when applied to the open boundaries in this model. In this paper, an adaptation of the Flather condition used by [18] is installed in the external mode of the forward model, which is given as follows:

$$\begin{aligned} \eta_1 &= \tilde{\eta} \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right) \frac{H}{g}} U, \\ \eta_1 &= \tilde{\eta} \pm \sqrt{\left(1 - \frac{f^2}{\omega^2}\right) \frac{H}{g}} V, \end{aligned} \quad (8)$$

where $\tilde{\eta}$ is external data beyond the model boundary representing the clamped surface elevation relating to the boundary barotropic tidal force and $H = h + \eta_1$ is the time-varying total water depth. The sign in (8) depends on the boundary location (positive for eastern and northern boundaries; negative for western and southern boundaries).

For the internal mode, the relaxation conditions, which have been found to have a robust performance in a variety of situations [42, 43], are applied at open boundaries. In this condition, a relaxation region of several grids close to the boundary is defined. Within this region, total tidal values including the interface elevation and fluid velocity of both barotropic and baroclinic tides are first calculated using the standard discretization and then relaxed towards the barotropic values as follows:

$$x(i, j) = x_{Bt}(i, j) + [1 - \mu(i, j)] \cdot x_{Bc}(i, j), \quad (9)$$

to give the final values x . Here, x represents the tidal states η_k , u_k , or v_k , x_{Bt} and x_{Bc} represent the barotropic and baroclinic tides, respectively, and $\mu(i, j) \in [0, 1]$ is the relaxation factor in grid (i, j) . In our model, μ is chosen to be increased linearly from 0 to 1 while the grid is getting close to the open boundary.

Closed boundary conditions for both the internal and the external modes are zero flow normal to the coast; that is, $\tilde{u} \cdot \tilde{n} = 0$ for the grid points at closed boundary, where \tilde{n} is the outward unit vector and $\tilde{u} = (u, v)$ is the velocity vector.

2.2. Adjoint Model. The adjoint method is a powerful tool for parameter estimation. The basic idea of the adjoint method is quite simple: a model is defined by an algorithm and its control variables including initial conditions, boundary conditions, and empirical parameters. The cost function that

measures the data misfit between the model output and observations is minimized through optimizing the control variables. In detail, the cost function decreases along a certain search direction which can be calculated with a certain optimization algorithm according to the gradient of cost function with respect to the control variables, and this gradient is calculated by what has been known as the adjoint model. Based on the governing equations (3a)–(4c) of the forward model, its adjoint model can be constructed as follows. The details of the adjoint model in a generalized form (not limited to the two-layer case) can be found in the work of [18].

Concretely, the cost function is defined by

$$J(q, u, v; \mathbf{p}) = \frac{1}{2} \int_{\Sigma} \left[\sum_{k=1}^2 (u_k - \hat{u}_k)^2 + \sum_{k=1}^2 (v_k - \hat{v}_k)^2 \right] d\sigma, \quad (10)$$

where k is layer index, Σ denotes the model domain of both time and space, \mathbf{p} represents the generalized control variables and is defined by the open boundary conditions (OBCs) for a regional internal tidal model, which are denoted by a series of Fourier coefficients in the present work, u_k and v_k are the models simulated, and \hat{u}_k and \hat{v}_k are the observations. In (10), the variables u_k and v_k can be uniquely determined by the control \mathbf{p} via the model equations (3a)–(4c). Hence, the functional J is implicitly dependent on \mathbf{p} .

The Lagrangian function is defined by

$$\begin{aligned} L(q, u, v; q_a, u_a, v_a; \mathbf{p}) &= J(q, u, v; \mathbf{p}) \\ &+ \int_{\Sigma} [q_{a1} \cdot F_{(3a)} + u_{a1} q_1 \cdot F_{(3b)} \\ &\quad + v_{a1} q_1 \cdot F_{(3c)}] d\sigma \\ &+ \int_{\Sigma} [q_{a2} \cdot F_{(4a)} + u_{a2} q_2 \cdot F_{(4b)} \\ &\quad + v_{a2} q_2 \cdot F_{(4a)}] d\sigma, \end{aligned} \quad (11)$$

where q_{ak} , u_{ak} , and v_{ak} are called adjoint variables of the state variables q_k , u_k , and v_k , respectively. $F_{(3a)}$, $F_{(3b)}$, \dots , $F_{(4c)}$ denote the left-hand side of (3a), (3b), \dots , (4c), respectively.

According to the typical theory of Lagrangian multiplier method, we have the following first-order derivatives of Lagrangian function with respect to all the variables and parameters:

$$\frac{\partial L}{\partial q_{ak}} = 0, \quad \frac{\partial L}{\partial u_{ak}} = 0, \quad \frac{\partial L}{\partial v_{ak}} = 0; \quad (12)$$

$$\frac{\partial L}{\partial q_k} = 0, \quad \frac{\partial L}{\partial u_k} = 0, \quad \frac{\partial L}{\partial v_k} = 0; \quad (13)$$

$$\frac{\partial L}{\partial \mathbf{p}} = 0. \quad (14)$$

Equations (12) return the governing equations (3a)–(4c). The adjoint equations can be derived using (13). From (14), we

can obtain the gradients of the cost function with respect to control variables.

Similar to the forward model, the adjoint model also consists of the internal and external modes. Actually, the equations derived from (13) are considered as the internal mode, and the external mode can be derived from the internal mode in a similar way the external mode of the forward model is derived. The details of both the internal and external modes of the adjoint model can be found in [18].

2.3. Discretization. Several numerical methods have been widely used in the discretization of time-dependent 3D primitive equations. The time integration schemes of these methods can be fully explicit [44], semi-implicit [45, 46] or fully implicit [47]. For large-scale oceanic problems, the applications of 3D models are becoming a reality with the aid of modern computers. The fully explicit finite difference method is relatively simple to implement, except that its time step is strictly restricted by the Courant-Friedrich-Lewy (CFL) stability criterion [48]. At present, many existing ocean models are based on an Alternating Direction Implicit (ADI) method which was proposed for the approximate solution of the shallow-water equations in [49]. ADI method results in computational efficiency superior to fully explicit methods because their improved stability allows large time steps to be employed, and it is also easy for implementation (see [13, 14]). Since the model must simulate fields of both velocity and elevations in each isopycnic layer, a technique known as external-internal mode splitting has been used in several ocean models by Simons [50]. Complete details on the model discretization were given in [18].

2.4. Gradients. Among all the control variables in this model, the OBCs are the most important and have critical impacts for a regional tidal simulation. Solutions in model interior are uniquely determined by the tidal OBCs once the initial conditions and other parameters have been determined. The Flather condition (8) can be determined once the external data $\bar{\eta}$ is described. Hence, the determination of the open boundary conditions for the present model is equivalent to the estimation of the tidal force $\bar{\eta}$. Assume that at a certain open boundary grid point (i_B, j_B) , the M_2 tidal force $\bar{\eta}$ at the n th time step is subject to

$$\bar{\eta}_{i_B, j_B}^n = a_{i_B, j_B} \cos(\omega n \Delta t) + b_{i_B, j_B} \sin(\omega n \Delta t), \quad (15)$$

where ω denotes the frequency of M_2 constituent, Δt is the time step length, and a_{i_B, j_B} and b_{i_B, j_B} are the Fourier coefficients as well as tunable parameters in the model. The gradients of cost function with respect to a_{i_B, j_B} and b_{i_B, j_B} can be deduced from (14) which yields

$$\frac{\partial J}{\partial a_{i_B, j_B}} + \sum_n T_{i_B, j_B}^n \cos(\omega n \Delta t) = 0, \quad (16)$$

$$\frac{\partial J}{\partial b_{i_B, j_B}} + \sum_n T_{i_B, j_B}^n \sin(\omega n \Delta t) = 0.$$

Details on computing T_{i_B, j_B}^n were given in [18].

Having determined the adjoint variables (q_a , u_a and v_a) with the adjoint model, the gradients of cost function with respect to the OBCs (i.e., the Fourier coefficients a and b) can thus be calculated. Then, the OBCs are optimized with some minimization algorithms which will be described in the following section.

3. Optimization Methods

In general, numerical methods solving the minimization problem (1) have the following iterative formula:

$$\mathbf{p}_{k+1} = \mathbf{p}_k + \alpha_k \mathbf{d}_k, \quad (17)$$

where \mathbf{p}_k , α_k , and \mathbf{d}_k are current iterative point, a positive step length, and a search direction, respectively. For simplicity, $J(\mathbf{p}_k)$ and $\nabla J(\mathbf{p}_k)$ are denoted by J_k and \mathbf{g}_k , respectively. There are many formulas to define search direction \mathbf{d}_k in formula (17) [31]. In this paper, three optimization methods are compared when applied to the assimilation model. They are different in the selection of α_k or \mathbf{d}_k . The three methods are given as follows.

3.1. L-BFGS Method. The L-BFGS method is a well-known limited-memory quasi-Newton method for solving large-scale problems whose Hessian matrices cannot be computed at a reasonable cost or are not sparse [31]. It requires the search direction to be

$$\mathbf{d}_k = -\mathbf{H}_k \mathbf{g}_k, \quad (18)$$

where

$$\begin{aligned} \mathbf{H}_k = & (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m}^T) \mathbf{H}_k^0 (\mathbf{V}_{k-m} \cdots \mathbf{V}_{k-1}) \\ & + \delta_{k-m} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+1}^T) \mathbf{s}_{k-m} \mathbf{s}_{k-m}^T (\mathbf{V}_{k-m+1} \cdots \mathbf{V}_{k-1}) \\ & + \delta_{k-m+1} (\mathbf{V}_{k-1}^T \cdots \mathbf{V}_{k-m+2}^T) \\ & \times \mathbf{s}_{k-m+1} \mathbf{s}_{k-m+1}^T (\mathbf{V}_{k-m+2} \cdots \mathbf{V}_{k-1}) \\ & + \cdots \\ & + \delta_{k-1} \mathbf{s}_{k-1} \mathbf{s}_{k-1}^T. \end{aligned} \quad (19)$$

Here, $\delta_k = 1/\mathbf{y}_k^T \mathbf{s}_k$, $\mathbf{V}_k = \mathbf{I} - \delta_k \mathbf{y}_k \mathbf{s}_k^T$, $\mathbf{s}_k = \mathbf{p}_{k+1} - \mathbf{p}_k = \alpha_k \mathbf{d}_k$, $\mathbf{y}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, \mathbf{I} is the identity matrix, and \mathbf{H}_k^0 is the initial Hessian approximation. Many previous studies have shown that typically $3 \leq m \leq 7$, where $m > 7$ does not improve the performance of L-BFGS. In this paper, the L-BFGS version of Liu and Nocedal [32] is employed, and the Fortran codes are authored by Nocedal [51], while and the number of corrections m used in the L-BFGS update is taken as 5 [33].

3.2. Gradient Descent Method with Wolfe's Line Search (GDM-W). This method requires the search direction to satisfy

$$\mathbf{d}_k = -\mathbf{g}_k, \quad (20)$$

while the step length α_k satisfies the following strong Wolfe conditions [36, 37]:

$$J(\mathbf{p}_k + \alpha_k \mathbf{d}_k) \leq J_k + c_1 \alpha_k \mathbf{d}_k^T \mathbf{g}_k, \quad (21a)$$

$$\left| \nabla J(\mathbf{p}_k + \alpha_k \mathbf{d}_k)^T \mathbf{d}_k \right| \leq c_2 \left| \mathbf{g}_k^T \mathbf{d}_k \right| \quad (21b)$$

with $0 < c_1 < c_2 < 1$. The line search routine enforcing the strong Wolfe conditions can also be found in [31].

3.3. Simplified Gradient Descent Method (GDM-S). This method requires the search direction to satisfy

$$\mathbf{d}_k = -\mathbf{g}_k, \quad (22)$$

while the step length α_k is simply defined by

$$\alpha_k = \frac{\alpha_0}{\|\mathbf{g}_k\|_2}, \quad (23)$$

where the denominator $\|\mathbf{g}_k\|_2$ denotes the L_2 norm of the gradient \mathbf{g}_k and α_0 is an empirical constant. This method has been used in a relatively small number of studies (see [13, 14, 23, 52]) showing that this method is indeed feasible and effective in practical application.

4. Numerical Experiments

4.1. Numerical Implementation. The performances of L-BFGS, GDM-W, and GDM-S are compared by a set of ideal experiments. The computing domain has 51×41 horizontal grids (with the horizontal resolution $10' \times 10'$) and two vertical isopycnic layers (Figures 1(a) and 1(b)). The maximum undisturbed water depth is 6000 m, and the undisturbed interface is placed at the depth of 2000 m. The 2D bottom topography is constructed by the superposition of an inclined plane and a Gaussian surface (Figure 1(b)). The angular frequency of M_2 tide is $1.4050789025 \times 10^{-4} \text{ s}^{-1}$, and the whole-time step is 496.863 s (1/90 of the period of M_2 tide) for both external and internal modes. The horizontal eddy viscosity coefficient is chosen to be $A_h = 1000$. The coefficients of bottom and interface frictions are taken as $\kappa = 0.0025$ and $A_v = 0.003$, respectively.

As shown in Figure 1(a), all four boundaries are open and installed with the passive Flather condition allowing the outgoing wave to radiate out of the computing domain through the open boundaries. In this work, for simplicity and no loss of generality, only the eastern boundary is treated as the source of tidal force driving the forward model (meaning that both Fourier coefficients a and b are equal to zero on the other three boundaries), and only the estimation of Fourier coefficient a is studied. To be specific, on the eastern boundary, the Fourier coefficient b is always treated as the known parameter (equal to 0 in this work), and the Fourier coefficient a to be estimated here is assumed to be spatially

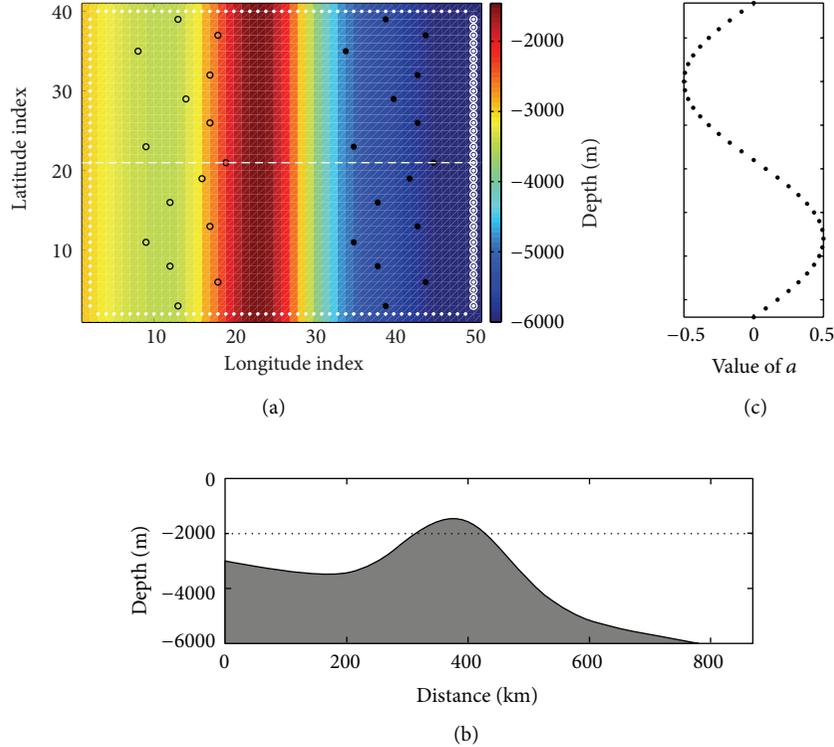


FIGURE 1: Experiment design. (a) Plane view of the computing domain and observation locations. The black dots and the black open circles represent the locations of the observations O1 and O2, respectively. The white dots and the white open circles denote the open boundaries and the open boundary points where the tidal force is installed, respectively. The white dashed line indicates the position of the longitudinal section which will be shown in (b). (b) Bottom topography along the dashed line shown in (a) and the location of the undisturbed interface between the two layers shown as the dashed line. (c) Prescribed spatial distribution of OBCs (Fourier coefficient a specially here) along the eastern boundary.

distributed and is prescribed with a spatial distribution which is constructed by the trigonometric function (see Figure 1(c)).

Some grids are randomly picked out from the 51×41 horizontal grids and treated as the observation positions (see Figure 1(a)). The forward model is run with the prescribed OBCs, and the model-generated results of the surface currents (i.e., currents in the upper layer) at these observation points are taken as the pseudo-observations (referred to as observations for brevity hereafter). Then, the OBCs can be optimized with the following steps.

Step 1. An initial value is chosen empirically and assigned to the unknown control variables.

Step 2. Perform the simulation by running the forward model, and the simulation results, mainly the state variables especially including the current velocity at observation points, are obtained. The value of the cost function is worked out in this step.

Step 3. The difference between simulation results and observations serves as the external force driving the adjoint model. Also, the adjoint variables in a period of M_2 tide are calculated through backward integration of the adjoint equations.

Step 4. Having known both the state and adjoint variables from Steps 2 and 3, respectively, the gradients with respect to control variables to be inverted are calculated by formula (16).

Step 5. Update the unknown control variables with a certain optimization algorithm.

Step 6. If some stopping criterion is satisfied, then stop and return the optimized control variables, otherwise; replace the initial value with the new control variables and go to Step 2.

With the procedure previous repeated, the OBCs will be optimized continuously, and the difference between simulated values and observations will be diminished. Meanwhile, the difference between the prescribed and the inverted OBCs will also be decreased. The initial value of a is taken as zero in each experiment. For all methods examined here, the total number of iterations is allowed to be 100 at most, and we report the cost function values, the gradient L_2 norms, and the inversion results obtained by each method. Note that the L-BFGS iteration may stop before it reaches the maximum number due to the termination criterion installed in the L-BFGS routine authored by Nocedal [51]. The number of corrections m used in the L-BFGS update is taken as 5. For the

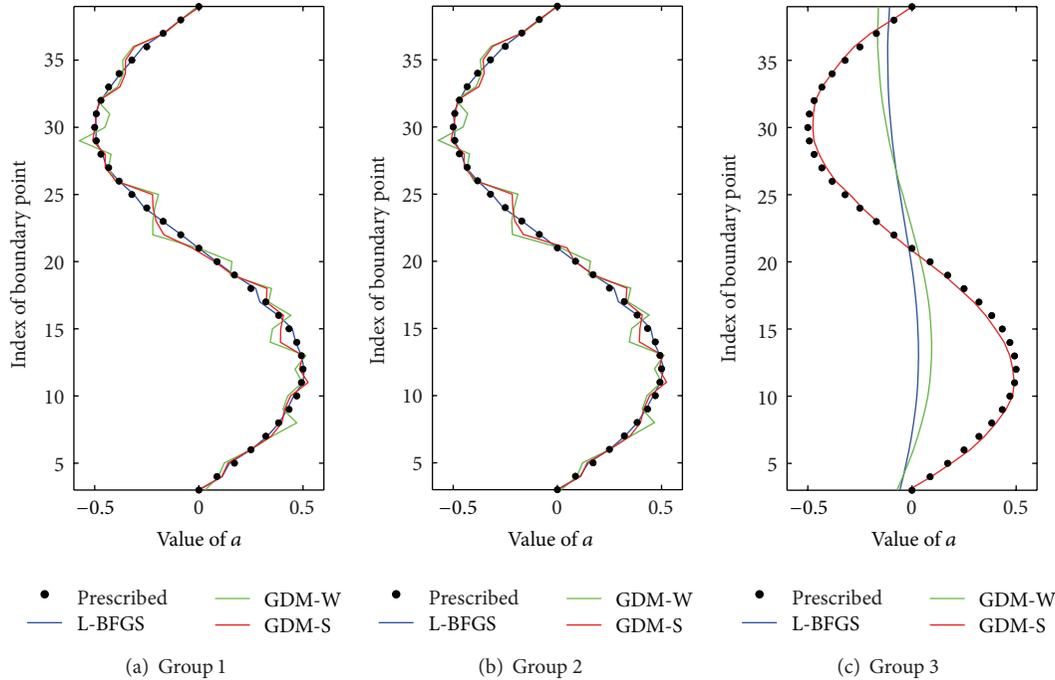


FIGURE 2: Inverted spatial distribution of OBCs (Fourier coefficient a specially here) compared with the prescribed one.

GDM-W, Wolfe's line search is carried out with parameters $c_1 = 10^{-4}$ and $c_2 = 0.9$. For the GDM-S, the constant α_0 is experimentally chosen to be 0.1.

4.2. Results. As shown in Figure 1(a), all observations can be divided into two groups: O1 which are closer to the eastern boundary and O2 which are further from the eastern boundary. Accordingly, the experiments are divided into three groups: Group 1 (both O1 and O2 are used), Group 2 (only O1 is used), Group 3 (only O2 is used). After assimilation, the inversion error (in the root mean square sense) and correlation coefficient between the prescribed and inverted OBCs are reported in Table 1, and the inverted OBCs are compared with the prescribed ones in Figure 2. All experiment results will be examined group by group as follows.

First, let us examine the performances of the three methods in the experiments of Group 1 in which observations of both O1 and O2 are assimilated. The results of Group 1 shown in Figure 2(a) and Table 1 indicate that with sufficient observations assimilated, the inversion results obtained by using all three methods are satisfactory in terms of the inverted OBC curve as well as the inversion error and the correlation coefficient. And especially, the solution obtained with the L-BFGS method is the closest to the exact solution (the prescribed OBCs). The variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 3(a), 3(b), and 3(c), respectively, as a function of the iteration number. The optimization procedure clearly shows that among the three methods examined, the L-BFGS performs the best

and its convergence rate is much faster than those of the GD methods, which is consistent with the classical theories about the convergence rate of the quasi-Newton methods and GD method [31]. Now, let us pay more attention to the results obtained with GDM-W and GDM-S. The comparison between the two versions of GD method is rather interesting. At the beginning stage of the iteration, the GDM-W has a faster convergence rate than the GDM-S. After more than 20 iteration steps, the result of GDM-S begins to be better than that of the GDM-W. Then, after a neighborhood of the exact solution is reached, a slow convergence of GDM-S occurs in the following iterations, also with oscillations in both of the cost function and gradient norm which may be caused by the constant α_0 making the step length relatively overlarge when the solution is close to the exact one. In contrast, the GDM-W keeps the cost function and the gradient norm declining on the whole due to its step length satisfying the strong Wolfe conditions (21a) and (21b). It is worth noting that in the latter part of the assimilation procedure (after about 30 iteration steps when the oscillations in the cost function and in the gradient norm for the GDM-S appear (see Figures 3(a) and 3(b)), the convergence rate of the GDM-W is a little faster than that of the GDM-S. This leads to the final values of both the cost function and gradient norm obtained with the GDM-W slightly smaller than those obtained with the GDM-S. On the other hand, however, the comparison between the inversion errors for the GDM-W and GDM-S presents a different pattern. As shown in Figure 3(c), during the beginning several iterations, the descent rate for both GD methods is very rapid and the inversion error for the GDM-W has a faster descent rate than that for the GDM-S. This state continues until the 25th iteration step when the inversion

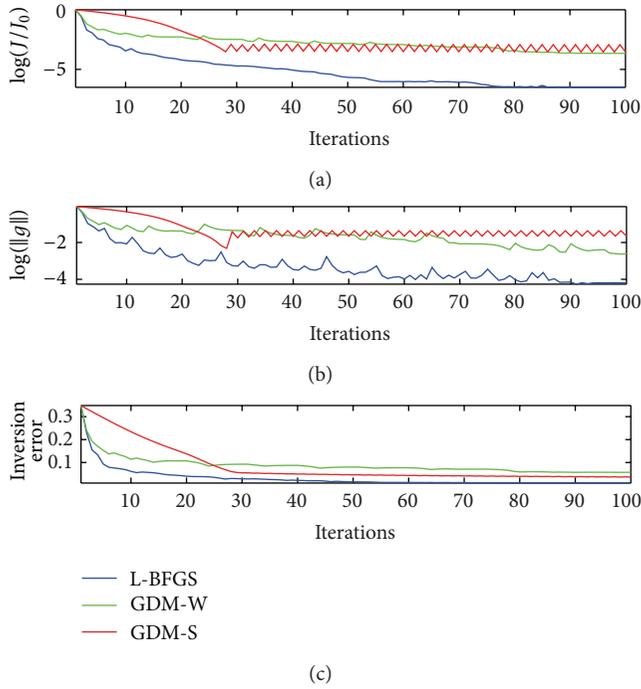


FIGURE 3: Optimization history for experiments of Group 1, about (a) the cost function normalized by its initial value J_0 , (b) the L_2 norm of the gradient of the cost function with respect to the OBCs, and (c) the inversion error (in the root mean square sense) between the inverted and prescribed OBCs.

error for the GDM-W begins to exceed that for the GDM-S, and these two methods almost begin to have the same slow descent rate of the inversion error. Finally, the inversion result obtained with the GDM-S is better than that obtained with the GDM-W (also see Figure 2(a) and Table 1).

For all experiments in Group 2, only the observations of O1 are assimilated. For this group, the variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 4(a), 4(b), and 4(c), respectively, as a function of the iteration number. As we can see, the assimilation procedure of Group 2 shown in Figure 4 follows a similar pattern as that of Group 1 shown in Figure 3, although the number of observations for Group 2 is only half of that for Group 1. This suggests that removing the observations that are far from the unknown boundary may have very little effect upon the experiment results, whereas the observations that are closer to the eastern open boundary play a much more important role in the estimation of the OBCs in our model. Probably, this may be the main reason why we can obtain the satisfactory results in Group 1.

To sum up, from the previous results, we have seen that the L-BFGS method can hold its advantage over the other two versions of GD method in terms of both the optimization procedure and the final results providing the observations that are closer to the unknown boundary are included for assimilation, and the L-BFGS method is recommended in this

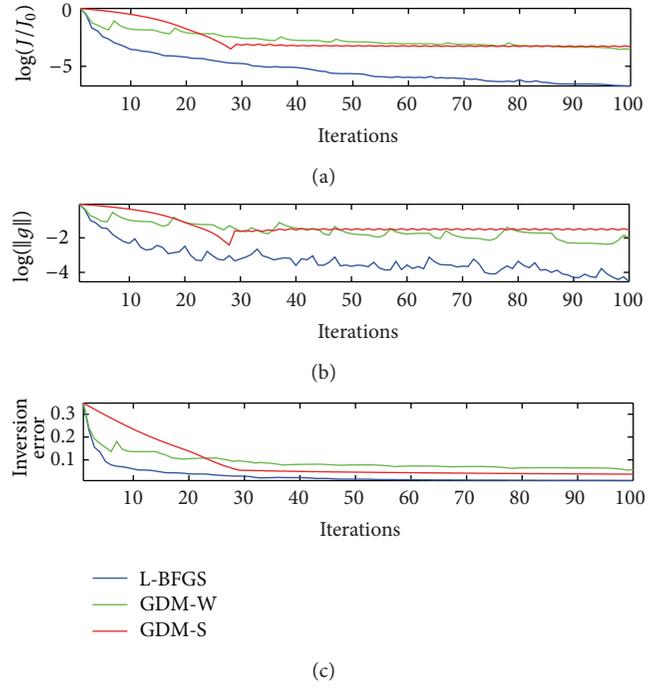


FIGURE 4: As in Figure 3 but for the experiments of Group 2.

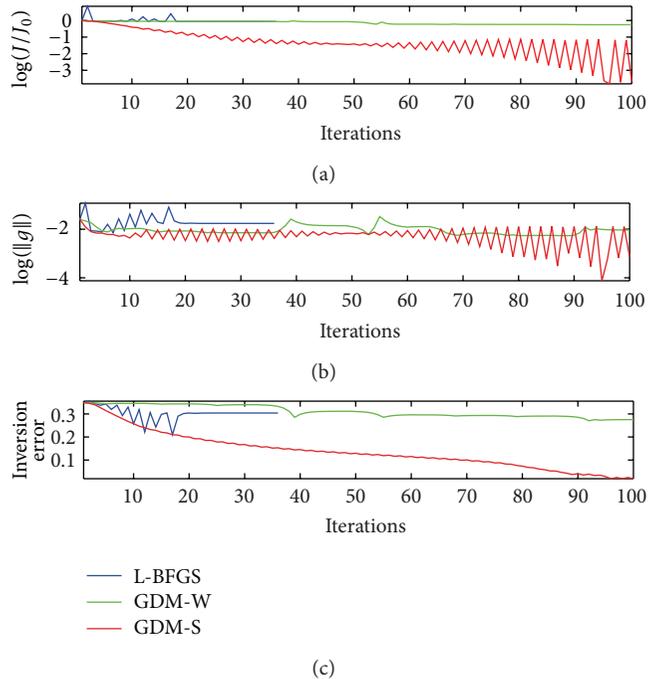


FIGURE 5: As in Figure 3 but for the experiments of Group 3.

case. Meanwhile, both results obtained with the two versions of GD method are also satisfactory. The comparison between the results obtained with the GDM-W and GDM-S indicates that the Wolfe's line search is effective in finding a reasonable step length that can achieve adequate reductions in the cost function and can guarantee a rapid rate of convergence at

TABLE 1: Inversion error (in the root mean square sense) and correlation coefficient between the inverted and prescribed OBCs.

Group	Inversion error			Correlation coefficient		
	L-BFGS	GDM-W	GDM-S	L-BFGS	GDM-W	GDM-S
1	0.0105	0.0571	0.0361	0.9995	0.9865	0.9947
2	0.0094	0.0566	0.0372	0.9996	0.9868	0.9943
3	0.3038	0.2753	0.0193	0.8935	0.8482	0.9986

the beginning of iteration. This allows the GDM-W to gain an advantage over the GDM-S in terms of the final values of the cost function and the gradient norm. Nevertheless, their difference is smaller than expected, and finally, the GDM-S excels the GDM-W in the inversion result which is an essential criterion testing whether the whole model is valid. Therefore, the GD method simplified with a constant step length is competitive to that with the Wolfe's line search in this case.

It is very interesting to find that when we come to the results of Group 3 which are dramatically different from those of Group 2 although the same number of observations is used in these two groups. In Group 3, only the observations O2 that are further from the unknown boundary are assimilated. For this group, the variation of the cost function normalized by its initial value, that of the L_2 norm of the gradient of the cost function with respect to the OBCs, and that of the inversion error are plotted in Figures 5(a), 5(b) and 5(c), respectively, as a function of the iteration number. From the optimization procedure shown in Figure 5, we can clearly see that, in this case, the L-BFGS method stops at the 36th iteration step and fails to converge, and the GDM-W has an unacceptable slow convergence rate, both leading to a relatively large discrepancy between the inverted and prescribed OBCs (see Figure 2(c) and Table 1), whereas the simple GDM-S performs the best and, more importantly, can still yield satisfactory inversion result (see Figure 2(c) and Table 1). The performance of GDM-S shown in Figure 5 is more interesting. After about 10 steps of iteration in the beginning, the cost function and the gradient norm begin to vary with oscillations which become increasingly larger as the iteration goes on, but meanwhile, their minimum values are getting smaller. At last when the maximum of 100 iterations is reached, the cost function and the gradient norm are reduced by more than 3 orders of magnitude and more than 1.5 orders of magnitude, respectively, compared with their initial values, which are comparable to the results of GDM-S in Groups 1 and 2. On the other hand, the variation of the inversion error for GDM-S also contains oscillations, but compared with the oscillations in the cost function and gradient norm (see Figures 5(a) and 5(b)), the ones in the inversion error are much smaller, and therefore the variation curve looks much smoother (see Figure 5(c)). This may be caused by the high complexity of the cost function in the control variable space if the observations to be assimilated are far from the unknown boundary, and this high complexity may also be a possible reason for the failure of the L-BFGS method as well as for the inefficiency of Wolfe's line search in Group 3. The results of Group 3 indicate that the far distance between the observations and the unknown boundary has almost

little effect on the validity and feasibility of the GDM-S, and therefore the simplified GD method is recommended in this case, instead of the L-BFGS method.

5. Summary and Conclusions

In this paper, based on the numerical internal tidal model constructed by Chen et al. [18], the practical performances of the L-BFGS method (given by Liu and Nocedal [32]) and two versions of GD method are compared and investigated computationally through a series of ideal experiments in which the OBCs are estimated by assimilating the interior observations with the adjoint method. The two GD methods include the normal one with Wolfe's line search for the step length and the simplified one with a fixed step length which should be predetermined by modeler. The cost function, the gradient norm, and the inversion result are reported and examined for each experiment.

According to the locations of the observations assimilated, all observations can be divided into two groups. Accordingly, all experiments are divided into three groups. In both of the first two groups, the observations that are closer to the unknown boundary are included for assimilation. In this case, as expected, the L-BFGS method has a remarkably better performance than the GD methods, not only in terms of the convergence rate but also in terms of the final result. This is consistent with the classical theory about the convergence properties of the L-BFGS method and the GD method. On the other hand, compared with the simplified GD method, the normal one with Wolfe's line search can really use fewer iteration steps to reach a satisfactory solution, but the values of the minimized cost function and the gradient norm are almost the same. Although great computational efforts have been made to implement Wolfe's line search to optimize the step length, the effect is much smaller than expected, and even the normal GD method finally yields a worse inversion result than the simplified one. In the third group of experiments, only the observations that are further from the unknown boundary are assimilated. In this case, the simplified GD method remains effective and yields satisfactory results, whereas the L-BFGS method fails to converge and the GD method with Wolfe's line search presents an unacceptable slow convergence rate. This demonstrates that in the practical application, the simplified GD method, which is controllable and easy to implement, is competitive to both the classical L-BFGS method and the normal GD method with Wolfe's line search. This suggests that when applied to the practical cases, the advanced L-BFGS algorithm and Wolfe's line search may still need to be improved, while how to take sufficient advantage of some simple methods is still

worth investigating. Nonetheless, the simplified GD method should not be ignored and should be regarded seriously as a choice, especially when the classical advanced optimization techniques fail or perform poorly. Furthermore, how to take sufficient advantage of some other simplified methods is also worth investigating.

Acknowledgment

The authors would like to deeply thank the editor and two anonymous reviewers for their constructive suggestions whereby their paper has been improved greatly. Partial support for this research was provided by the National Natural Science Foundation of China through Grant Nos. 41076006 and 41206001, the State Ministry of Science and Technology of China under Contract nos. 2008AA09A402, 2013AA091201, and 2013AA091202, the Ministry of Education's III Project through Grant B07036, and the Fundamental Research Funds for the Central Universities 201261006 and 201262007.

References

- [1] I. M. Navon, X. Zou, J. Derber, and J. Sela, "Variational data assimilation with an adiabatic version of the NMC spectral model," *Monthly Weather Review*, vol. 120, no. 7, pp. 1433–1446, 1992.
- [2] J. Zou, W. H. Hsieh, and I. M. Navon, "Sequential open-boundary control by data assimilation in a limited-area model," *Monthly Weather Review*, vol. 123, pp. 2899–2909, 1995.
- [3] D. L. T. Anderson, J. Sheinbaum, and K. Haines, "Data assimilation in ocean models," *Reports on Progress in Physics*, vol. 59, no. 10, pp. 1209–1266, 1996.
- [4] R. W. Lardner, A. H. Al-Rabeh, and N. Gunay, "Optimal estimation of parameters for a two-dimensional hydro-dynamical model of the Arabian gulf," *Journal of Geophysical Research*, vol. 98, no. C10, pp. 18229–18242, 1993.
- [5] U. Seiler, "Estimation of open boundary conditions with the adjoint method," *Journal of Geophysical Research*, vol. 98, no. C12, pp. 22855–22870, 1993.
- [6] P. G. J. ten Brummelhuis, A. W. Heemink, and H. F. P. van den Boogaard, "Identification of shallow sea models," *International Journal for Numerical Methods in Fluids*, vol. 17, no. 8, pp. 637–665, 1993.
- [7] A. W. Heemink, E. E. A. Moutaan, M. R. T. Roest, E. A. H. Vollebregt, K. B. Robaczewska, and M. Verlaan, "Inverse 3D shallow water flow modelling of the continental shelf," *Continental Shelf Research*, vol. 22, no. 3, pp. 465–484, 2002.
- [8] I. Shulman, "Local data assimilation in specification of open boundary conditions," *Journal of Atmospheric and Oceanic Technology*, vol. 14, no. 6, pp. 1409–1419, 1997.
- [9] I. Shulman, J. K. Lewis, A. F. Blumberg, and B. N. Kim, "Optimized boundary conditions and data assimilation with application to the M_2 tide in the yellow sea," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 4, pp. 1066–1071, 1998.
- [10] V. Taillandier, V. Echevin, L. Mortier, and J. L. Devenon, "Controlling boundary conditions with a four-dimensional variational data-assimilation method in a non-stratified open coastal model," *Ocean Dynamics*, vol. 54, no. 2, pp. 284–298, 2004.
- [11] A. J. Zhang, E. Wei, and B. B. Parker, "Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique," *Continental Shelf Research*, vol. 23, no. 11–13, pp. 1055–1070, 2003.
- [12] I. Y. Gejadze and G. J. M. Copeland, "Open boundary control problem for Navier-Stokes equations including a free surface: adjoint sensitivity analysis," *Computers & Mathematics with Applications*, vol. 52, no. 8–9, pp. 1243–1268, 2006.
- [13] J. C. Zhang and X. Q. Lu, "Parameter estimation for a three-dimensional numerical barotropic tidal model with adjoint method," *International Journal for Numerical Methods in Fluids*, vol. 57, no. 1, pp. 47–92, 2008.
- [14] J. C. Zhang and X. Q. Lu, "Inversion of three-dimensional tidal currents in marginal seas by assimilating satellite altimetry," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 49–52, pp. 3125–3136, 2010.
- [15] A. Z. Cao, Z. Guo, and X. Q. Lu, "Inversion of two-dimensional tidal open boundary conditions of M_2 constituent in the Bohai and Yellow Seas," *Chinese Journal of Oceanology and Limnology*, vol. 30, no. 5, pp. 868–875, 2012.
- [16] Z. Guo, A. Z. Cao, and X. Q. Lv, "Inverse estimation of open boundary conditions in the Bohai Sea," *Mathematical Problems in Engineering*, vol. 2012, Article ID 628061, 9 pages, 2012.
- [17] E. Kazantsev, "Boundary conditions control for a shallow-water model," *International Journal for Numerical Methods in Fluids*, vol. 68, no. 5, pp. 625–641, 2012.
- [18] H. B. Chen, C. B. Miao, and X. Q. Lv, "A three-dimensional numerical internal tidal model involving adjoint method," *International Journal for Numerical Methods in Fluids*, vol. 69, no. 10, pp. 1584–1613, 2012.
- [19] Z. Wang, I. M. Navon, X. Zou, and F. X. Le Dimet, "A truncated Newton optimization algorithm in meteorology applications with analytic Hessian/vector products," *Computational Optimization and Applications*, vol. 4, no. 3, pp. 241–262, 1995.
- [20] Z. Wang, K. Droegemeier, and L. White, "The adjoint Newton algorithm for large-scale unconstrained optimization in meteorology applications," *Computational Optimization and Applications*, vol. 10, no. 3, pp. 283–320, 1998.
- [21] X. Zou, I. M. Navon, M. Berger, K. H. Phua, T. Schlick, and F. X. Le Dimet, "Numerical experience with limited-memory quasi-Newton and truncated Newton methods," *SIAM Journal on Optimization*, vol. 3, no. 3, pp. 582–608, 1993.
- [22] X. Q. Lu and J. C. Zhang, "Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method," *Continental Shelf Research*, vol. 26, no. 16, pp. 1905–1923, 2006.
- [23] J. C. Zhang, X. Q. Lu, P. Wang, and Y. P. Wang, "Study on linear and nonlinear bottom friction parameterizations for regional tidal models using data assimilation," *Continental Shelf Research*, vol. 31, no. 6, pp. 555–573, 2011.
- [24] A. F. Bennett and P. C. McIntosh, "Open ocean modeling as an inverse problem: tidal theory," *Journal of Physical Oceanography*, vol. 12, no. 10, pp. 1004–1018, 1982.
- [25] F. X. Le Dimet and O. Talagrand, "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects," *Tellus A*, vol. 38, no. 2, pp. 97–110, 1986.
- [26] E. Courtier and O. Talagrand, "Variational assimilation of meteorological observations with the adjoint equations—part 2: numerical results," *Quarterly Journal of the Royal Meteorological Society*, vol. 113, no. 478, pp. 1329–1347, 1987.

- [27] X. Zou, I. M. Navon, and F. X. Le Dimet, "Incomplete observations and control of gravity waves in variational data assimilation," *Tellus A*, vol. 44, no. 4, pp. 273–296, 1992.
- [28] M. Ghil and P. Malanotte-Rizzoli, "Data assimilation in meteorology and oceanography," in *Advances in Geophysics*, B. Saltzman, Ed., vol. 33, pp. 141–266, Academic Press, New York, NY, USA, 1991.
- [29] I. M. Navon, "Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography," *Dynamics of Atmospheres and Oceans*, vol. 27, no. 1–4, pp. 55–79, 1998.
- [30] L. S. Yu and J. J. O'Brien, "On the initial condition parameter estimation," *Journal of Physical Oceanography*, vol. 22, pp. 1361–1364, 1992.
- [31] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer Series in Operations Research and Financial Engineering, Springer, New York, NY, USA, 2nd edition, 2006.
- [32] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, vol. 45, no. 3, pp. 503–528, 1989.
- [33] A. K. Alekseev, I. M. Navon, and J. L. Steward, "Comparison of advanced large-scale minimization algorithms for the solution of inverse ill-posed problems," *Optimization Methods & Software*, vol. 24, no. 1, pp. 63–87, 2009.
- [34] J. Nocedal, A. Sartenaer, and C. Zhu, "On the behavior of the gradient norm in the steepest descent method," *Computational Optimization and Applications*, vol. 22, no. 1, pp. 5–35, 2002.
- [35] M. N. Vrahatis, G. S. Androulakis, J. N. Lambrinos, and G. D. Magoulas, "A class of gradient unconstrained minimization algorithms with adaptive stepsize," *Journal of Computational and Applied Mathematics*, vol. 114, no. 2, pp. 367–386, 2000.
- [36] P. Wolfe, "Convergence conditions for ascent methods," *SIAM Review*, vol. 11, pp. 226–235, 1969.
- [37] P. Wolfe, "Convergence conditions for ascent methods. II. Some corrections," *SIAM Review*, vol. 13, pp. 185–188, 1971.
- [38] G. Zhou, "A descent algorithm without line search for unconstrained optimization," *Applied Mathematics and Computation*, vol. 215, no. 7, pp. 2528–2533, 2009.
- [39] Z.-J. Shi and J. Shen, "Convergence of descent method without line search," *Applied Mathematics and Computation*, vol. 167, no. 1, pp. 94–107, 2005.
- [40] R. A. Flather, "A tidal model of the north-west European continental shelf," *Memoires de la Societe Royale des Sciences de Liege*, vol. 6, no. 10, pp. 141–164, 1976.
- [41] C. B. Miao, H. B. Chen, and X. Q. Lu, "An isopycnic-coordinate internal tide model and its application to the South China Sea," *Chinese Journal of Oceanology and Limnology*, vol. 29, no. 6, pp. 1339–1356, 2011.
- [42] E. D. Palma and R. P. Matano, "On the implementation of passive open boundary conditions for a general circulation model: the barotropic mode," *Journal of Geophysical Research C*, vol. 103, no. 1, pp. 1319–1341, 1998.
- [43] J. Nycander and K. Ds, "Open boundary conditions for barotropic waves," *Journal of Geophysical Research*, vol. 108, no. C5, 2002.
- [44] J. J. Leendertse, R. C. Alexander, and S. K. Liu, "A three dimensional model for estuaries and coastal seas: volume I, principles of computations," Report R-1417-OWRR, Rand Corporation, Santa Monica, Calif, USA, 1973.
- [45] J. O. Backhaus, "A three-dimensional model for the simulation of shelf sea dynamics," *Deutsche Hydrographische Zeitschrift*, vol. 38, no. 4, pp. 165–187, 1985.
- [46] V. Casulli, "A semi-implicit finite difference method for the non-hydrostatic free-surface flow," *International Journal for Numerical Methods in Fluids*, vol. 30, no. 4, pp. 425–440, 1999.
- [47] H. L. Yuan and C. H. Wu, "An implicit three-dimensional fully non-hydrostatic model for free-surface flows," *International Journal for Numerical Methods in Fluids*, vol. 46, no. 7, pp. 709–733, 2004.
- [48] P. J. Roache, *Fundamentals of Computational Fluid Dynamics*, Hermosa Publishers Press, Albuquerque, NM, USA, 1982.
- [49] G. Fairweather and I. M. Navon, "A linear ADI method for the shallow-water equations," *Journal of Computational Physics*, vol. 37, no. 1, pp. 1–18, 1980.
- [50] T. J. Simons, "Verification of numerical models of Lake Ontario—part I: circulation in spring and early summer," *Journal of Physical Oceanography*, vol. 4, no. 4, pp. 507–523, 1974.
- [51] L. J. Nocedal, "BFGS subroutine, software for large-scale unconstrained optimization," <http://www.ece.northwestern.edu/~nocedal/lbfgs.html>.
- [52] H. B. Chen, X. Q. Lv, and Y. S. Qiao, "Application of gradient descent method to the sedimentary grain-size distribution fitting," *Journal of Computational and Applied Mathematics*, vol. 233, no. 4, pp. 1128–1138, 2009.

Research Article

A Mollification Regularization Method for a Fractional-Diffusion Inverse Heat Conduction Problem

Zhi-Liang Deng,¹ Xiao-Mei Yang,² and Xiao-Li Feng³

¹ School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China

² School of Mathematics, Southwest Jiaotong University, Chengdu 610031, China

³ Department of Mathematical Sciences, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Zhi-Liang Deng; zhiliangdeng@gmail.com

Received 12 June 2012; Revised 19 December 2012; Accepted 20 December 2012

Academic Editor: Fatih Yaman

Copyright © 2013 Zhi-Liang Deng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The ill-posed problem of attempting to recover the temperature functions from one measured transient data temperature at some interior point of a one-dimensional semi-infinite conductor when the governing linear diffusion equation is of fractional type is discussed. A simple regularization method based on Dirichlet kernel mollification techniques is introduced. We also propose *a priori* and *a posteriori* parameter choice rules and get the corresponding error estimate between the exact solution and its regularized approximation. Moreover, a numerical example is provided to verify our theoretical results.

1. Introduction

It is well known that the classical diffusion equation requires second spatial derivatives and first time derivatives. However, people are shifting their partial focus to fractional-order differential equations with the realization that the use of fractional-order derivatives and integrals leads to formulas of certain physical processes which is more economical and useful than the classical approach in terms of Fick's laws of diffusion [1–4]. Some time fractional diffusion equations involving only a first-order spatial derivative and a half-order time derivative are successfully used for modeling some anomalous diffusion physical phenomena (see, e.g., [5–7] and the references therein). One can also find some applications of such fractional diffusion equations in several scenarios in [2, 7], such as relaxation to equilibrium in systems (such as polymers chains and membranes) with long temporal memory, anomalous transport in disordered systems, diffusion on fractals, and to model non-Markovian dynamical processes in protein folding. Fractional diffusion equations have been extensively investigated both in theory itself and

numerical computation based on the broad applications in many application fields, especially in describing phenomena related to anomalous diffusion processes. The following is a partial list of articles which contain theoretical results and numerical tests. Some fundamental solutions and Green functions of fractional differential equations are given in [8, 9]. Finite difference techniques [10–12] and finite element methods [13, 14] provide some efficiency numerical results for several kinds of fractional differential equations. In [15], the decomposition method is used to construct analytical approximate solutions of time-fractional wave equation subject to specified boundary conditions. An analytical solution of a fractional diffusion equation by Adomian decomposition method is presented in [16]. One can also refer [17–19] to the regularity of the solution of fractional diffusion equations and some *a priori* estimates.

In this paper, we consider the following one-dimensional fractional diffusion problem on a semi-infinite slab: suppose the temperature $f(t)$ at some interior point $x = x_0 > 0$ is approximately measurable. For convenience, we set $x_0 = 1$. The temperature $u(x, t)$ at the $0 < x < 1$ is desired and

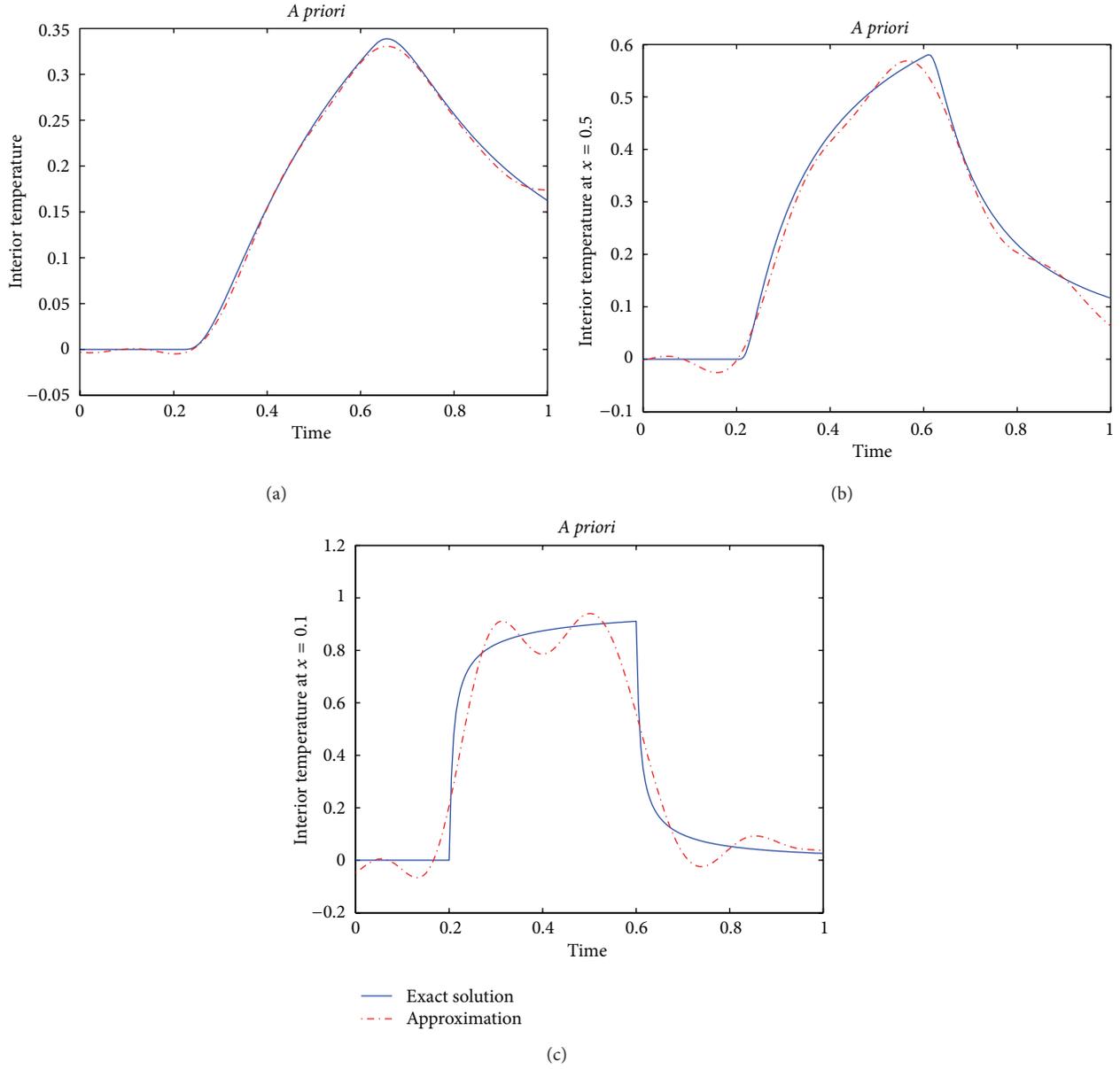


FIGURE 1: The comparisons of exact solution and regularization solution at (a) $x = 0.9$, (b) $x = 0.5$, (c) $x = 0.1$, $\nu = 34.3956$, $\delta = 0.01$.

unknown. The mathematical description of the fractional diffusion problem is listed in the following. The unknown temperature $u(x, t)$ satisfies

$$\begin{aligned} \frac{\partial}{\partial x} u(x, t) &= -\frac{1}{\sqrt{\kappa}} \frac{\partial^{1/2}}{\partial t^{1/2}} u(x, t) + \frac{u_0}{\sqrt{\pi \kappa t}}, \quad x > 0, \quad t \geq 0, \\ u(1, t) &= f(t), \end{aligned} \quad (1)$$

where κ is the constant diffusivity coefficient, $u_0 = u(x, 0)$, a constant. The half-time differentiation $(\partial^{1/2}/\partial t^{1/2})u(x, t)$

indicates the Riemann-Liouville fractional derivative with order $\alpha = 1/2$ which is given by the convolution integral

$$\begin{aligned} (\partial^\alpha u)(t) &= \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t u(s) (t-s)^{-\alpha} ds, \quad 0 \leq t \leq T, \quad 0 < \alpha < 1, \end{aligned} \quad (2)$$

where $\Gamma(\cdot)$ is the Gamma function. More detailed information on fractional derivatives and a general historical perspective may be found in [7, 20, 21]. The situations we will treat are those in which the system is initially at equilibrium, so that

$$u(x, t) = u_0, \quad t < 0, \quad x \geq 0. \quad (3)$$

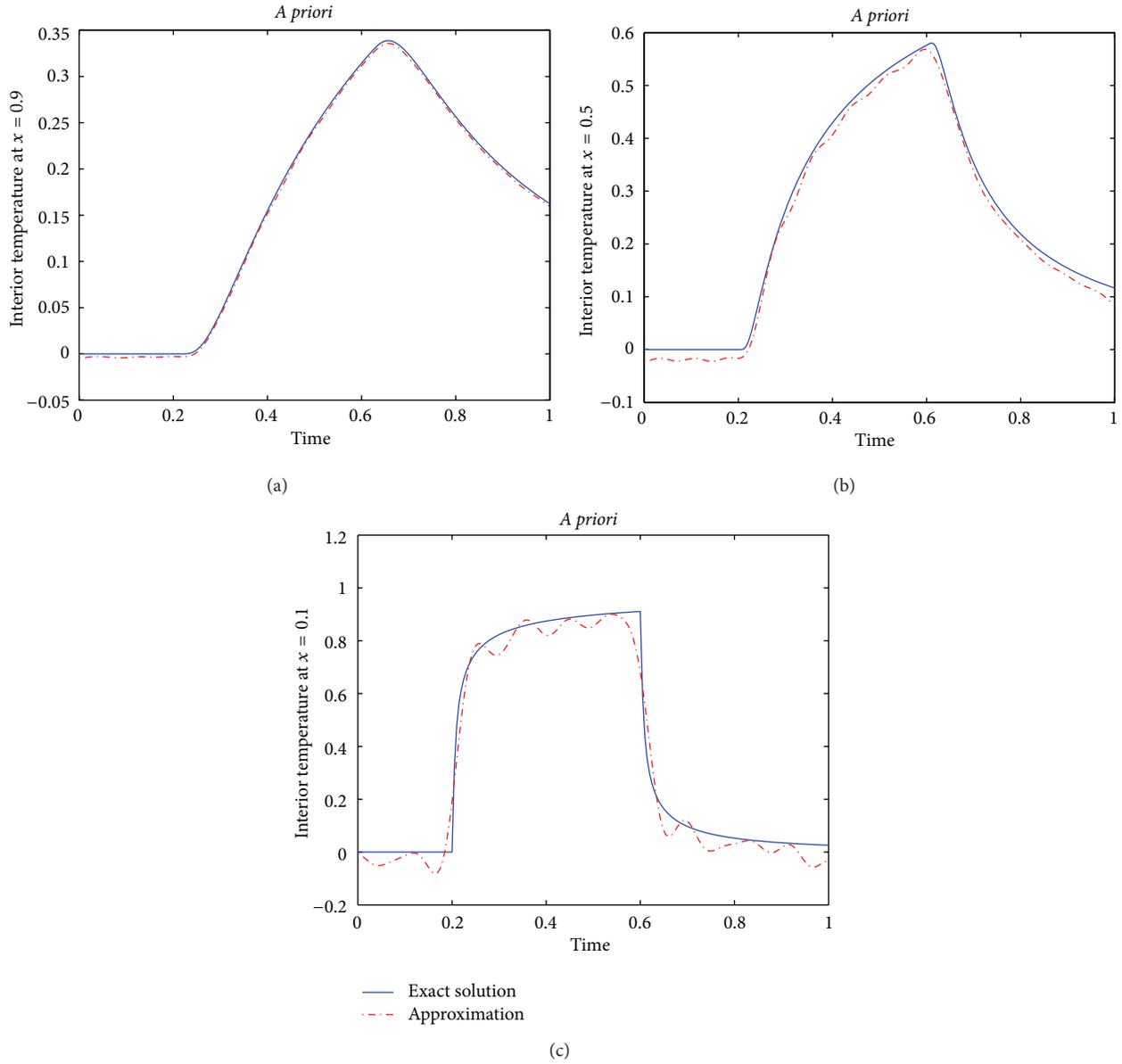


FIGURE 2: The comparisons of exact solution and regularization solution at (a) $x = 0.9$, (b) $x = 0.5$, (c) $x = 0.1$, $\nu = 83.1949$, $\delta = 0.001$.

At $t = 0$ a perturbation of the system commences by some unspecified process occurring at the boundary. During times of interest T , this perturbation does not affect regions remote from the $x = 0$ boundary, so that the relationship

$$\lim_{x \rightarrow +\infty} u(x, t) = u_0, \quad t \leq T \quad (4)$$

applies. In [22], Murio discussed the similar problems using mollification method with Gauss kernel. The idea used in this current work is a development of the ideas in [22]. However, there is only formal stability and discrete error discussion in the cited paper. Here, we give some analysis of error estimates under *a priori* and *a posteriori* regularization parameter, and a comparison of those two choice methods in our regularization method. In the following, the regularization method is outlined.

In order to simplify the Fourier analysis of fractional diffusion problem (1), and in the rest of the paper, we assume without loss of generality, $\kappa = 1$, $u_0 = 0$. We also assume that all the functions involved are $L^2(\mathbb{R})$ and use the corresponding L^2 norm, as defined in the following, to measure errors:

$$\|y\| = \left(\int_{\mathbb{R}} |y(t)|^2 dt \right)^{1/2}. \quad (5)$$

If the Fourier transform of a function $y(t)$ is written as

$$\hat{y}(\xi) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y(t) e^{-i\xi t} dt, \quad i = \sqrt{-1}, \quad (6)$$

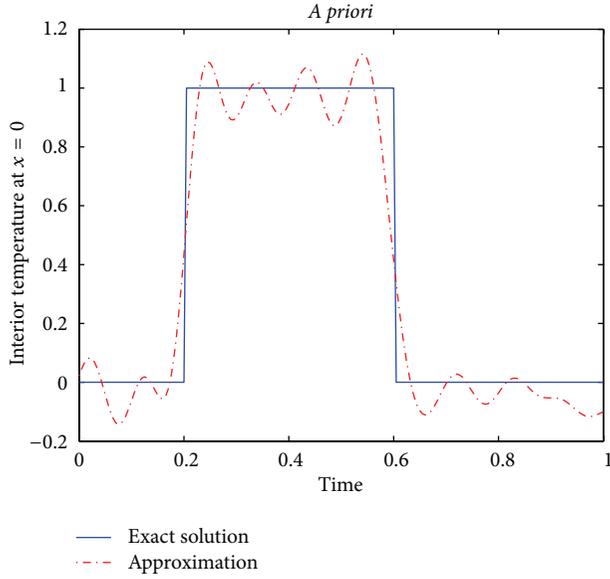


FIGURE 3: The comparisons of exact solution and regularization solution at $x = 0$, $\nu = 83.1949$, $\delta = 0.001$.

then the Fourier inversion formula reads

$$y(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \hat{y}(\xi) e^{i\xi t} d\xi. \quad (7)$$

Fourier transforming (1) with respect to variable t , it follows that

$$\frac{\partial}{\partial x} \hat{u}(x, \xi) = -\sqrt{i\xi} \hat{u}(x, \xi), \quad x > 0, \xi \in \mathbb{R} \quad (8)$$

according to the fact that [7]

$$\frac{\partial^{1/2}}{\partial t^{1/2}} h(t) = \sqrt{i\xi} \hat{h}(\xi). \quad (9)$$

The first-order ordinary differential equation (8) has the general solution

$$\hat{u}(x, \xi) = \hat{u}(1, \xi) e^{-\sqrt{i\xi}(1-x)} = \hat{f}(\xi) e^{-\sqrt{i\xi}(1-x)}, \quad (10)$$

or equivalently,

$$e^{-\sqrt{i\xi}(1-x)} \hat{u}(x, \xi) = \hat{f}(\xi). \quad (11)$$

Now, we can write the problem (1) as

$$\widehat{A}(\xi, x) \hat{u}(x, \xi) = \hat{f}(\xi), \quad (12)$$

where $\widehat{A}(\xi, x) := e^{-\sqrt{i\xi}(1-x)}$ is a multiplication operator.

Since we require $\hat{u}(x, \cdot) \in L^2(\mathbb{R})$ and $u_0 = 0$, we see from (4) and (10) that the real part of $\sqrt{i\xi}$ is positive. Thereby, the multiplication operator $\widehat{A}(\xi, x)$ is a bounded linear operator for $0 \leq x < 1$. Then its inverse operator is unbounded. Therefore problem (12) is linear ill posed. For the general theory of linear ill posed problems, we can

refer to [23]. It is worth pointing out that ill posed problems of a large number of diffusion equations, both fractional-order as well as integral order, have been discussed by many authors. Yang et al. [24–28] discuss the identification of source terms for some integral-order diffusion equations using some regularization strategies. Hon et al. [29, 30] apply some meshless methods to the ill posed problems of heat conduction equations. In [17, 22, 29, 31, 32], some uniqueness results and numerical methods are given for some fractional diffusion ill posed problems. Here, we apply a simple stabilizing method, namely, the mollification method with Dirichlet kernel [33], to stabilize the problem (1). Suppose that the measured data function $f_\delta(t)$ satisfies

$$\|f - f_\delta\| \leq \delta, \quad (13)$$

where δ is noise level. Take the Dirichlet function

$$D_\nu(t) := \sqrt{\frac{2}{\pi}} \frac{\sin \nu t}{t} \quad (14)$$

as the mollifier kernel, where ν is a positive constant. Define operator M_ν as

$$M_\nu u(x, t) := \int_{-\infty}^{+\infty} D_\nu(\tau) u(x, t - \tau) d\tau. \quad (15)$$

Then, we have the following associated problem: for some $\delta > 0$, find $M_\nu u_\delta(x, t)$ that satisfies

$$\frac{\partial}{\partial x} M_\nu u_\delta(x, t) = -\frac{\partial^{1/2}}{\partial t^{1/2}} M_\nu u_\delta(x, t), \quad x > 0, t \geq 0, \quad (16)$$

$$M_\nu u_\delta(1, t) = M_\nu f_\delta(t).$$

In fact, ν plays the role of regularization parameter.

In Section 2 an *a priori* parameter choice rule and the corresponding error estimate are discussed. In Section 3, we propose an *a posteriori* parameter choice rule and get the error bound. Numerical tests are given in Section 4.

2. The Error Estimate with *a Priori* Parameter Choice

In this section, the error estimate of the mollification regularization method will be derived under the *a priori* parameter choice rule. Suppose that the following source condition holds:

$$\|u(0, \cdot)\| \leq E. \quad (17)$$

The source condition (17) is indispensable, otherwise, there can be no uniform convergence rates for any regularization method, that is, the convergence is arbitrarily slow

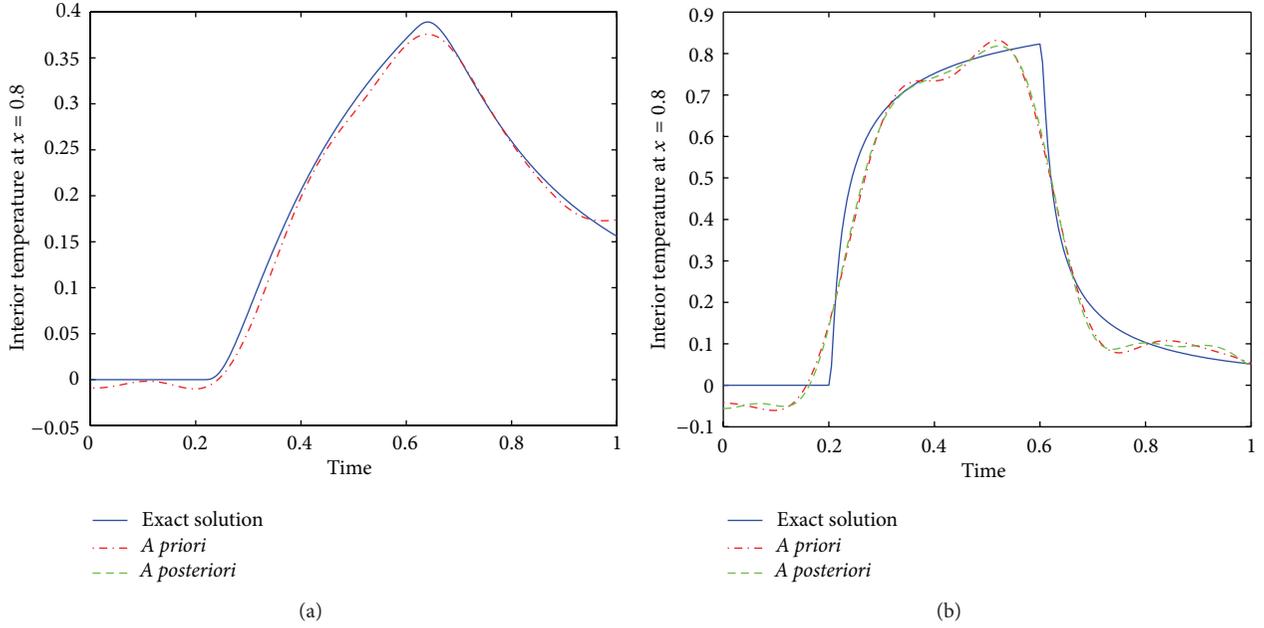


FIGURE 4: The comparisons of exact solution and regularization solution at (a) $x = 0.8$, (b) $x = 0.2$ with $\delta = 0.01$.

(see Proposition 3.11 in [23]). For $0 < x < 1$, by Parseval formula and triangle inequality, we know that

$$\begin{aligned}
 & \|M_\nu u_\delta(x, \cdot) - u(x, \cdot)\| \\
 &= \|\widehat{M_\nu u_\delta}(x, \cdot) - \widehat{u}(x, \cdot)\| \\
 &= \|e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f_\delta}(\xi) - e^{(1-x)\sqrt{i\xi}} \widehat{f}(\xi)\| \\
 &\leq \|e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f_\delta}(\xi) - e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f}(\xi)\| \\
 &\quad + \|e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f}(\xi) - e^{(1-x)\sqrt{i\xi}} \widehat{f}(\xi)\| \\
 &:= I_1 + I_2.
 \end{aligned} \tag{18}$$

Since $|e^{\sqrt{i\xi}}| = e^{\sqrt{|\xi|/2}}$, for I_1 , we get by (13)

$$\begin{aligned}
 I_1 &= \left\| e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f_\delta}(\xi) - e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f}(\xi) \right\| \\
 &= \left(\int_{-\nu}^{\nu} \left| e^{(1-x)\sqrt{i\xi}} (\widehat{f_\delta} - \widehat{f}) \right|^2 d\xi \right)^{1/2} \\
 &\leq e^{(1-x)\sqrt{\nu/2}} \delta.
 \end{aligned} \tag{19}$$

For I_2 , we use the source condition (17) and obtain

$$\begin{aligned}
 I_2 &= \left\| e^{(1-x)\sqrt{i\xi}} \widehat{D_\nu f}(\xi) - e^{(1-x)\sqrt{i\xi}} \widehat{f}(\xi) \right\| \\
 &= \left(\int_{|\xi| \geq \nu} \left| e^{(1-x)\sqrt{i\xi}} \widehat{f}(\xi) \right|^2 d\xi \right)^{1/2} \\
 &= \left(\int_{|\xi| \geq \nu} \left| e^{(1-x)\sqrt{i\xi}} e^{-\sqrt{i\xi}} \widehat{u}(0, \xi) \right|^2 d\xi \right)^{1/2} \\
 &\leq e^{-\sqrt{\nu/2}x} E.
 \end{aligned} \tag{20}$$

Therefore,

$$\|M_\nu u_\delta(x, \cdot) - u(x, \cdot)\| \leq e^{(1-x)\sqrt{\nu/2}} \delta + e^{-\sqrt{\nu/2}x} E. \tag{21}$$

Minimizing the right-hand side of (21), we set $e^{(1-x)\sqrt{\nu/2}} \delta = e^{-\sqrt{\nu/2}x} E$ and get

$$\nu = 2 \left(\log \frac{E}{\delta} \right)^2, \tag{22}$$

which also implies

$$\|M_\nu u_\delta(x, \cdot) - u(x, \cdot)\| \leq 2E^{1-x} \delta^x, \quad \text{for } 0 < x < 1. \tag{23}$$

Therefore, we get the following theorem.

Theorem 1. Assume conditions (13), (17) hold. If the regularization parameter ν is taken by (22), then for $0 < x < 1$, there holds the error estimate (23).

Remark 2. The error estimate (23) is order optimal in the sense of Tautenhahn [34]. In our application $\|u(0, \cdot)\|$ is

usually not known, therefore we have no exact *a priori* bound E and cannot choose the parameter ν according to (22). However, if selecting $\nu = 2(\log(1/\delta))^2$, we can obtain the convergence rate

$$\|M_\nu u_\delta(x, \cdot) - u(x, \cdot)\| \leq 2\delta^x, \quad \text{for } 0 < x < 1. \quad (24)$$

Theorem 1 provides no information about the convergence and convergence rates of $M_\nu u_\delta(x, \cdot)$ at $x = 0$. The question is settled by our next result. We now give the error estimate at $x = 0$ under a stronger *a priori* assumption

$$\|u(0, \cdot)\|_p \leq E, \quad p > 0, \quad (25)$$

where $\|\cdot\|_p$ denotes the norm on Sobolev space $H_p(\mathbb{R})$ defined by

$$\|f\|_p^2 := \int_{-\infty}^{\infty} (1 + \xi^2)^p |\widehat{f}(\xi)|^2 d\xi \quad (26)$$

for $f \in H^p(\mathbb{R})$. We only need to reestimate I_2 for $x = 0$. Under the stronger *a priori* bound (25), it is not hard to get

$$\begin{aligned} I_2 &= \left(\int_{|\xi| \geq \nu} |\widehat{u}(0, \xi)|^2 d\xi \right)^{1/2} \\ &= \left(\int_{|\xi| \geq \nu} (1 + \xi^2)^{-p} (1 + \xi^2)^p |\widehat{u}(0, \xi)|^2 d\xi \right)^{1/2} \\ &\leq \frac{1}{(1 + \nu^2)^p} E \leq \frac{E}{\nu^{2p}}. \end{aligned} \quad (27)$$

Then, (18), (19), and (27) lead to the following error bound:

$$\|M_\nu u_\delta(0, \cdot) - u(0, \cdot)\| \leq e^{\sqrt{\nu/2}} \delta + \frac{E}{\nu^{2p}}. \quad (28)$$

The error bound (28) does not provide the convergence as $\delta \rightarrow 0$ obviously. Hence, we need to choose a proper parameter ν . For this reason, some proper lower bound of the right-hand side of (28) should be given. In order to minimize the right-hand side of (28), it is necessary to introduce the following lemma and its proof can be found in [34].

Lemma 3 (see [34]). *Let the function $f(\lambda) : (0, a] \rightarrow \mathbb{R}$ be given by*

$$f(\lambda) = \lambda^b \left[d \log \frac{1}{\lambda} \right]^{-c} \quad (29)$$

with a constant $c \in \mathbb{R}$ and positive constants $a < 1$, b and d , then for the inverse function $f^{-1}(\lambda)$ one has

$$f^{-1}(\lambda) = \lambda^{1/b} \left[\frac{d}{b} \log \frac{1}{\lambda} \right]^{c/b} (1 + o(1)) \quad \text{for } \lambda \rightarrow 0. \quad (30)$$

Minimizing the right-hand side of (28), we let $e^{\sqrt{\nu/2}} \delta = E/\nu^{2p}$. Denote $e^{-\sqrt{\nu/2}} := \lambda$. Simple computation shows $\lambda(\log(1/\lambda))^{-4p} = \delta^{2p}/E$. By using (30) in Lemma 3, we obtain

$$\lambda = \frac{\delta^{2p}}{E} \left(\log \frac{E}{\delta^{2p}} \right)^{4p} (1 + o(1)), \quad \text{for } \delta \rightarrow 0, \quad (31)$$

which also implies that

$$\nu = 2 \log^2 \left(\frac{\delta^{2p}}{E} \left(\log \frac{E}{\delta^{2p}} \right)^{4p} \right) (1 + o(1)), \quad \text{for } \delta \rightarrow 0. \quad (32)$$

Therefore, the following estimate holds

$$\begin{aligned} &\|M_\nu u_\delta(0, \cdot) - u(0, \cdot)\| \\ &\leq CE \left(2 \log^2 \left(\frac{\delta^{2p}}{E} \left(\log \frac{E}{\delta^{2p}} \right)^{4p} \right) \right)^{(-2p)}, \end{aligned} \quad (33)$$

where C is a constant, which also shows that the convergence rate at $x = 0$ is logarithmic.

Theorem 4. *Assume conditions (13), (25) hold. If the regularization parameter ν is taken as given by (32), then the error estimate (33) holds.*

Remark 5. From the error estimate (23), as $x \rightarrow 0$, we see that the accuracy of regularization solution becomes progressively lower and even cannot get convergence for $x = 0$. This is common in the theory of ill posed problems. Nevertheless, if a stronger *a priori* assumption (25) is imposed, the regularization solution converges to the exact solution at $x = 0$, but only in a slower way.

3. The Error Estimate with a Posteriori Parameter Choice

In this section, we consider the *a posteriori* regularization parameter choice rule. Choose the regularization parameter ν as the solution of the equation

$$d(\nu) := \|\widehat{D}_\nu \widehat{f}_\delta(\cdot) - \widehat{f}_\delta(\cdot)\| = \tau \delta, \quad (34)$$

where $\tau > 1$ is a constant. To establish existence and uniqueness of solution of (34), we need the following lemma.

Lemma 6. *If $\delta > 0$, then there hold the following:*

- (a) $d(\nu)$ is a continuous function;
- (b) $\lim_{\nu \rightarrow 0} d(\nu) = \|\widehat{f}_\delta\|$;
- (c) $\lim_{\nu \rightarrow +\infty} d(\nu) = 0$;
- (d) $d(\nu)$ is a strictly decreasing function.

The proof is very easy and we omit it here. Denote

$$\zeta(x, t) := M_\nu u_\delta(x, t) - u(x, t). \quad (35)$$

We give the main result of this section as follows.

Theorem 7. *Assume the conditions (11) and (17) hold and $\tau > 1$. Take the solution ν of (34) as the regularization parameter, then there holds the error estimate for $0 < x < 1$:*

$$\|M_\nu u_\delta(x, t) - u(x, t)\| \leq C(E, \tau) E^{1-x} \delta^x, \quad (36)$$

where $C(E, \tau)$ is a constant depending on E, τ .

TABLE 1: Errors with *a priori* rule.

x	$\delta = 0.01$		$\delta = 0.001$	
	e_r	e_i	e_r	e_i
0.1	0.1590	1.2154	0.1000	0.7643
0.5	0.0596	0.2729	0.0517	0.2370
0.9	0.0173	0.0497	0.0154	0.0443

Proof. By (35), the Parseval formula, triangle inequality, and (17), we have

$$\begin{aligned}
 \|\zeta(0, \cdot)\| &= \|M_\nu u_\delta(0, \cdot) - u(0, \cdot)\| = \|\widehat{M_\nu u_\delta}(0, \cdot) - \widehat{u}(0, \cdot)\| \\
 &= \|\widehat{D_\nu \widehat{u}_\delta}(0, \cdot) - \widehat{u}(0, \cdot)\| = \|\widehat{D_\nu} e^{\sqrt{i\xi}} \widehat{f}_\delta - e^{\sqrt{i\xi}} \widehat{f}\| \\
 &= \|\widehat{D_\nu} e^{\sqrt{i\xi}} (\widehat{f}_\delta - \widehat{f}) + (\widehat{D_\nu} - 1) e^{\sqrt{i\xi}} \widehat{f}\| \\
 &\leq \|\widehat{D_\nu} e^{\sqrt{i\xi}} (\widehat{f}_\delta - \widehat{f})\| + E \\
 &\leq e^{\sqrt{\nu/2}} \delta + E.
 \end{aligned} \tag{37}$$

By virtue of (34), we know that

$$\begin{aligned}
 \tau \delta &= \|\widehat{D_\nu} \widehat{f}_\delta - \widehat{f}_\delta\| = \|(\widehat{D_\nu} - 1)(\widehat{f}_\delta - \widehat{f} + \widehat{f})\| \\
 &\leq \|(\widehat{D_\nu} - 1)(\widehat{f}_\delta - \widehat{f})\| + \|(\widehat{D_\nu} - 1)\widehat{f}\| \\
 &\leq \delta + \|(\widehat{D_\nu} - 1) e^{-\sqrt{i\xi}} \widehat{u}(0, \xi)\| \\
 &\leq \delta + e^{-\sqrt{\nu/2}} E,
 \end{aligned} \tag{38}$$

which means

$$(\tau - 1) \delta \leq e^{-\sqrt{\nu/2}} E. \tag{39}$$

Thus it is obvious that

$$e^{\sqrt{\nu/2}} \delta \leq \frac{E}{\tau - 1}. \tag{40}$$

Moreover, inserting (40) into (37), we get the following inequality:

$$\|\zeta(0, \cdot)\| \leq \frac{\tau}{\tau - 1} E. \tag{41}$$

In addition,

$$\begin{aligned}
 \|\zeta(1, \cdot)\| &= \|M_\nu u_\delta(1, \cdot) - u(1, \cdot)\| = \|\widehat{M_\nu u_\delta}(1, \cdot) - \widehat{u}(1, \cdot)\| \\
 &\leq \|\widehat{M_\nu u_\delta}(1, \cdot) - \widehat{u}_\delta(1, \cdot)\| + \|\widehat{u}_\delta(1, \cdot) - \widehat{u}(1, \cdot)\| \\
 &\leq \tau \delta + \delta = (\tau + 1) \delta.
 \end{aligned} \tag{42}$$

It is easy to see

$$\begin{aligned}
 \|\zeta(x, \cdot)\|^2 &= \int_{-\infty}^{+\infty} |\widehat{D_\nu \widehat{u}_\delta}(x, \xi) - \widehat{u}(x, \xi)|^2 d\xi \\
 &= \int_{-\infty}^{+\infty} |\widehat{D_\nu} e^{\sqrt{i\xi(1-x)} \widehat{f}_\delta(\xi) - e^{\sqrt{i\xi(1-x)} \widehat{f}(\xi)}|^2 d\xi \\
 &= \int_{-\infty}^{+\infty} |e^{\sqrt{i\xi(1-x)} (\widehat{D_\nu} \widehat{f}_\delta(\xi) - \widehat{f}(\xi))}|^2 d\xi.
 \end{aligned} \tag{43}$$

Denote $\theta_f = \widehat{D_\nu} \widehat{f}_\delta(\xi) - \widehat{f}(\xi)$. By using the Hölder inequality, we know that

$$\begin{aligned}
 \|\zeta(x, \cdot)\|^2 &= \int_{-\infty}^{+\infty} |e^{\sqrt{i\xi(1-x)} \theta_f}|^2 |\theta_f|^{2x} |\theta_f|^{2(1-x)} d\xi \\
 &\leq \left[\int_{-\infty}^{+\infty} \left(|e^{\sqrt{i\xi} \theta_f} \right)^{2(1-x)} d\xi \right]^{1/(1-x)} \left[\int_{-\infty}^{+\infty} (\theta_f^{2x})^{1/x} d\xi \right]^x \\
 &= \left[\int_{-\infty}^{+\infty} \left(|e^{\sqrt{i\xi} \theta_f} \right)^2 d\xi \right]^{1-x} \left[\int_{-\infty}^{+\infty} (\theta_f^2) d\xi \right]^x.
 \end{aligned} \tag{44}$$

Therefore, we obtain

$$\|\zeta(x, \cdot)\| \leq \|\zeta(0, \cdot)\|^{1-x} \|\zeta(1, \cdot)\|^x, \tag{45}$$

which also implies that

$$\begin{aligned}
 \|M_\nu u_\delta(x, \cdot) - u(x, \cdot)\| &\leq \|\zeta(0, \cdot)\|^{1-x} \|\zeta(1, \cdot)\|^x \\
 &\leq \left(\frac{\tau}{\tau - 1} E \right)^{1-x} ((\tau + 1) \delta)^x.
 \end{aligned} \tag{46}$$

Therefore, we complete the proof. \square

4. Numerical Examples

For linear heat diffusion, analytic solutions for the temperature distribution $u(x, t)$, $x > 0$, $t > 0$ in a semi-infinite solid with zero initial temperature, and $u(0, t) := H(t)$ at the surface, are obtained using the integral equation

$$u(x, t) = \left(H * \frac{\partial}{\partial t} K \right) (x, t) = \int_0^t H(s) \frac{\partial}{\partial t} K(x, t-s) ds, \tag{47}$$

TABLE 2: Errors with *a priori* rule at $x = 0$.

x	$\delta = 0.001$	
	e_r	e_i
0	0.1873	1.6748

TABLE 3: The comparison between *a priori* and *a posteriori* rules with $\delta = 0.01$.

x	<i>A priori</i>		<i>A posteriori</i>	
0.8	$e_r = 0.0321$	$e_i = 0.1063$	$e_r = 0.0313$	$e_i = 0.1010$
0.2	$e_r = 0.1175$	$e_i = 0.7817$	$e_r = 0.1085$	$e_i = 0.7214$

where the kernel function

$$K(x, t) = \operatorname{erfc}\left(\frac{x}{2\sqrt{t}}\right) \quad (48)$$

is the temperature distribution corresponding to a unit step boundary temperature, $H(t) = 1, t > 0$. Here, $\operatorname{erfc}(\cdot)$ denotes the complementary error function defined by

$$\operatorname{erfc}(z) := \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-t^2} dt. \quad (49)$$

We take the example from [22]. As an interesting and challenging test for the numerical method, we proposed the surface temperature function $H(t)$, which is 1 between 0.2 and 0.6 and zero otherwise. The exact solutions for the FICHP, in this example, are the functions

$$u(x, t) = \begin{cases} 0, & 0 < t \leq 0.2, \\ \operatorname{erfc}\left(\frac{x}{2\sqrt{t-0.2}}\right), & 0.2 < t \leq 0.6, \\ \operatorname{erfc}\left(\frac{x}{2\sqrt{t-0.2}}\right) & 0.6 < t < \infty. \\ -\operatorname{erfc}\left(\frac{x}{2\sqrt{t-0.6}}\right), & \end{cases} \quad (50)$$

Consequently, the exact interior data temperature is given by

$$f(t) = \begin{cases} 0, & 0 < t \leq 0.2, \\ \operatorname{erfc}\left(\frac{1}{2\sqrt{t-0.2}}\right), & 0.2 < t \leq 0.6, \\ \operatorname{erfc}\left(\frac{1}{2\sqrt{t-0.2}}\right) & 0.6 < t < \infty. \\ -\operatorname{erfc}\left(\frac{1}{2\sqrt{t-0.6}}\right), & \end{cases} \quad (51)$$

Suppose the vector F represents samples from the function $f(t)$. The noisy discrete data function is generated by adding random errors to the exact data function, at every grid point, that is,

$$F_{\delta}(n) = F(n) + \epsilon_n, \quad |\epsilon_n| \leq \delta, \quad n = 0, 1, \dots, N_t, \quad (52)$$

where the (ϵ_n) 's are independent random Gaussian variables with variance $\sigma^2 = \epsilon^2$. The absolute and relative weighted l^2 errors for the recovered interior temperatures are calculated as

$$e_i := \left[\frac{1}{N_t + 1} \sum_{n=0}^{N_t} |M_v u_{\delta}(x, n) - u(x, n)| \right]^{1/2}, \quad (53)$$

$$e_r := \frac{\left[(1/(N_t + 1)) \sum_{n=0}^{N_t} |M_v u_{\delta}(x, n) - u(x, n)| \right]^{1/2}}{\left[(1/(N_t + 1)) \sum_{n=0}^{N_t} |u(x, n)|^2 \right]^{1/2}}, \quad (54)$$

respectively.

It is easy to implement the algorithm described in (16). Using an available fast Fourier transform (FFT) subroutine, a simple program was written to test the algorithm for the above example in Matlab. Firstly, some tests are implemented to verify the effectiveness of the regularization method under *a priori* parameter choice rule at several interior points. In addition, we would like to compare the *a posteriori* parameter choice rule (34) with the *a priori* parameter choice rule (22). The *a priori* bound E in (22) can be calculated easily in this example. The Newton's bisection is used to solve (34), where we choose $\tau = 1.1$.

Figures 1 and 2 provide the comparisons between the exact solutions and regularization solutions with error bound $\delta = 0.01$ and $\delta = 0.001$ at interior point $x = 0.1, 0.5, 0.9$ using *a priori* parameter choice rule (22), respectively. Figure 3 gives the comparison between the exact solutions and regularization solutions with error bound $\delta = 0.001$ at boundary point $x = 0$ using *a priori* parameter choice rule (32). And intuitively, it seems that there are better numerical effects for closer distance from $x = 1$. We list the error in Tables 1 and 2 to verify our result. Figure 4 demonstrates the comparisons between *a priori* and *a posteriori* parameter choice rules and Table 3 tells us that it has better effect using *a posteriori* parameter choice rule than *a priori* parameter choice rule. We also see that the difference of the numerical results between *a priori* choice rule and *a posteriori* choice rule is slight, which agrees with our theoretical results.

Acknowledgments

The authors are indebted to the referees of this paper for their most helpful comments and suggestions, which helped to improve the presentation greatly. This work was supported by

the Fundamental Research Funds for the Central Universities ZYGX2011J104 and SWJTU11BR078 and the NSF of China (Nos. 11126102, 11226040, and 11126187).

References

- [1] V. V. Kulish and J. L. Lage, "Fractional-diffusion solution for transient local temperature and heat flux," *Transactions of ASME*, vol. 122, pp. 372–376, 2000.
- [2] K. B. Oldham and J. Spanier, *The Fractional Calculus: Theory and Application of Differential and Integration to Arbitrary Order*, Academic Press, 1974.
- [3] K. B. Oldham and J. Spanier, "A general solution of the diffusion equation for semiinfinite geometries," *Journal of Mathematical Analysis and Applications*, vol. 39, pp. 655–669, 1972.
- [4] K. B. Oldham and J. Spanier, "The replacement of Fick's law by a formulation involving semidifferentiation," *Journal of Electroanalytical Chemistry*, vol. 26, pp. 331–341, 1970.
- [5] R. Gorenflo, F. Mainardi, D. Moretti, and P. Paradisi, "Time fractional diffusion: a discrete random walk approach," *Nonlinear Dynamics*, vol. 29, no. 1–4, pp. 129–143, 2002.
- [6] V. V. Anh and N. N. Leonenko, "Non-Gaussian scenarios for the heat equation with singular initial conditions," *Stochastic Processes and their Applications*, vol. 84, no. 1, pp. 91–114, 1999.
- [7] I. Podlubny, *Fractional Differential Equations*, Academic Press, 1999.
- [8] M. M. El-Borai, "The fundamental solutions for fractional evolution equations of parabolic type," *Boletín de la Asociación Matemática Venezolana*, vol. 6, no. 1, pp. 29–43, 2004.
- [9] F. Mainardi, Y. Luchko, and G. Pagnini, "The fundamental solution of the space-time fractional diffusion equation," *Fractional Calculus & Applied Analysis*, vol. 4, no. 2, pp. 153–192, 2001.
- [10] F. Liu, V. Anh, and I. Turner, "Numerical solution of the space fractional Fokker-Planck equation," *Journal of Computational and Applied Mathematics*, vol. 166, pp. 209–219, 2004.
- [11] M. M. Meerschaert and C. Tadjeran, "Finite difference approximations for fractional advection-dispersion flow equations," *Journal of Computational and Applied Mathematics*, vol. 172, no. 1, pp. 65–77, 2004.
- [12] S. B. Yuste, "Weighted average finite difference methods for fractional diffusion equations," *Journal of Computational Physics*, vol. 216, no. 1, pp. 264–274, 2006.
- [13] G. J. Fix and J. P. Roop, "Least squares finite-element solution of a fractional order two-point boundary value problem," *Computers & Mathematics with Applications*, vol. 48, no. 7–8, pp. 1017–1033, 2004.
- [14] J. P. Roop, "Computational aspects of FEM approximation of fractional advection dispersion equations on bounded domains in R^2 ," *Journal of Computational and Applied Mathematics*, vol. 193, no. 1, pp. 243–268, 2006.
- [15] Z. M. Odibat and S. Momani, "Approximate solutions for boundary value problems of time-fractional wave equation," *Applied Mathematics and Computation*, vol. 181, no. 1, pp. 767–774, 2006.
- [16] S. S. Ray and R. K. Bera, "Analytical solution of a fractional diffusion equation by Adomian decomposition method," *Applied Mathematics and Computation*, vol. 174, no. 1, pp. 329–336, 2006.
- [17] J. J. Liu and M. Yamamoto, "A backward problem for the time-fractional diffusion equation," *Applicable Analysis*, vol. 89, no. 11, pp. 1769–1788, 2010.
- [18] W. McLean, "Regularity of solutions to a time-fractional diffusion equation," *The ANZIAM Journal*, vol. 52, no. 2, pp. 123–138, 2010.
- [19] J. Prüss, *Evolutionary Integral Equations and Applications*, vol. 87, Birkhäuser, Basel, Switzerland, 1993.
- [20] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and Applications of Fractional Differential Equations*, Elsevier, 2006.
- [21] S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Fractional Integrals and Derivatives*, Gordon and Breach Science, Yverdon, Switzerland, 1993.
- [22] D. A. Murio, "Stable numerical solution of a fractional-diffusion inverse heat conduction problem," *Computers & Mathematics with Applications*, vol. 53, no. 10, pp. 1492–1501, 2007.
- [23] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic, 1996.
- [24] L. Yang, Z.-C. Deng, J.-N. Yu, and G.-W. Luo, "Two regularization strategies for an evolutionary type inverse heat source problem," *Journal of Physics A*, vol. 42, no. 36, Article ID 365203, 16 pages, 2009.
- [25] L. Yang, Z.-C. Deng, J.-N. Yu, and G.-W. Luo, "Optimization method for the inverse problem of reconstructing the source term in a parabolic equation," *Mathematics and Computers in Simulation*, vol. 80, no. 2, pp. 314–326, 2009.
- [26] L. Yang, M. Dehghan, J.-N. Yu, and G.-W. Luo, "Inverse problem of time-dependent heat sources numerical reconstruction," *Mathematics and Computers in Simulation*, vol. 81, no. 8, pp. 1656–1672, 2011.
- [27] L. Yang, J. N. Yu, G. W. Luo, and Z. C. Deng, "Reconstruction of a space and time dependent heat source from finite measurement data," *International Journal of Heat and Mass Transfer*, vol. 55, pp. 6573–6581, 2012.
- [28] L. Yang, J. N. Yu, G. W. Luo, and Z. C. Deng, "Numerical identification of source terms for a two dimensional heat conduction problem in polar coordinate system," *Applied Mathematical Modelling*, vol. 37, pp. 939–957, 2013.
- [29] F. F. Dou and Y. C. Hon, "Kernel-based approximation for Cauchy problem of the time-fractional diffusion equation," *Engineering Analysis with Boundary Elements*, vol. 36, no. 9, pp. 1344–1352, 2012.
- [30] Y. C. Hon and T. Wei, "A fundamental solution method for inverse heat conduction problem," *Engineering Analysis With Boundary Elements*, vol. 28, no. 5, pp. 489–495, 2004.
- [31] J. Cheng, J. Nakagawa, M. Yamamoto, and T. Yamazaki, "Uniqueness in an inverse problem for a one-dimensional fractional diffusion equation," *Inverse Problems*, vol. 25, no. 11, pp. 1–16, 2009.
- [32] B. Jin and W. Rundell, "An inverse problem for a one-dimensional time-fractional diffusion problem," *Inverse Problems*, vol. 28, no. 7, Article ID 075010, 19 pages, 2012.
- [33] D. N. Hào, "A mollification method for ill-posed problems," *Numerische Mathematik*, vol. 68, no. 4, pp. 469–506, 1994.
- [34] U. Tautenhahn, "Optimality for ill-posed problems under general source conditions," *Numerical Functional Analysis and Optimization*, vol. 19, no. 3–4, pp. 377–398, 1998.

Research Article

Stochastic Inverse Identification of Nonlinear Roll Damping Moment of a Ship Moving at Nonzero-Forward Speeds

S. L. Han and Takeshi Kinoshita

Institute of Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Correspondence should be addressed to Takeshi Kinoshita, kinoshita@iis.u-tokyo.ac.jp

Received 1 August 2012; Revised 14 November 2012; Accepted 26 November 2012

Academic Editor: Fatih Yaman

Copyright © 2012 S. L. Han and T. Kinoshita. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The nonlinear responses of ship rolling motion characterized by a roll damping moment are of great interest to naval architects and ocean engineers. Modeling and identification of the nonlinear damping moment are essential to incorporate the inherent nonlinearity in design, analysis, and control of a ship. A stochastic nonparametric approach for identification of nonlinear damping in the general mechanical system has been presented in the literature (Han and Kinoshita 2012). The method has been also applied to identification of the nonlinear damping moment of a ship at zero-forward speed (Han and Kinoshita 2013). In the presence of forward speed, however, the characteristic of roll damping moment of a ship is significantly changed due to the lift effect. In this paper, the stochastic inverse method is applied to identification of the nonlinear damping moment of a ship moving at nonzero-forward speed. The workability and validity of the method are verified with laboratory tests under controlled conditions. In experimental trials, two different types of ship rolling motion are considered: time-dependent transient motion and frequency-dependent periodic motion. It is shown that this method enables the inherent nonlinearity in damping moment to be estimated, including its reliability analysis.

1. Introduction

Ship motions are defined by the six degrees of freedom that a ship can experience at sea. Among them, rolling motion, which is defined by a rotational motion around the longitudinal axis of a ship, has been attracting considerable research attention over the years because large roll motions might be a serious threat to the safety of a ship such as ship capsizing or structure failure. Therefore, an appropriate model to describe the rolling motion is crucial for accurate predictions of ship's roll response in a given sea state.

There are considerable studies on the modeling of the rolling motion [1–20]. The rolling motion of a ship can be characterized by analyzing components of moments such as moment of inertia, damping and restoring moments. Among them, the roll damping moment is considered as the most important component due to difficulties in identifying its value. The difficulties result from its nonlinearity due to the fluid viscosity and its dependence on the forward speed. The usual practice of identifying roll damping moment of a ship is to use the parametric method [8–14], which assumes the form of nonlinearity a priori. The coefficients of the prescribed form of nonlinearity are then estimated by relating the loss in potential energy over each cycle of the roll decay test with the energy dissipation as the damping. This method gives quite accurate result but sometimes fails when one adopt inaccurate nonlinear model because different types of nonlinear model produce different motion responses [15].

A large number of studies on this subject have been concentrated on the parametric identification mentioned above. In contrast, there are few studies involving nonparametric identification where the prior knowledge for the nonlinearities of damping moments is not necessary. For example, Haddara and Hinchey [16] proposed a method, based on the combination of the neural networks technique with the standard parametric identification, for modeling nonlinear damping moment from free-roll decay curve. They also applied the method to random roll responses measured during the sea trials to investigate roll characteristics of full-scale ships at sea [17]. Roberts and Vasta [18] presented a stochastic method for estimating the damping moment and the excitation spectrum using the Markov model for the energy envelope of the response.

In recent years, inverse methods have been presented [1, 2, 19, 20] for the nonparametric identification of damping moment of a ship. Conceptually, these methods are based on inverse formalism originated from the transformation of an original nonlinear motion equation for ship rolling motion in either a deterministic [19, 20] or stochastic manner [1, 2]. The roll damping moment of a ship at zero-forward speed was identified from the measurement of the responses of the system based on the inverse problem formulation [2, 20], which is ill-posed in the sense that small variations in the input data can result in erroneous inverse solution. Compared with the deterministic inverse method [19, 20], the stochastic inverse method [1, 2] is robust and reliable in the sense that it can also provide a reliability analysis of the identified results.

In the literature [2], the study on the application of the stochastic inverse method was presented for the case of zero-forward speed. The inherent nonlinearity in damping moment of a test ship at zero-forward speed was successfully identified, including its credible intervals as an indicator of reliability of identification. However, there is no guarantee that the method can be applied to the case of nonzero-forward speed since the nonlinear characteristic of roll damping has changed significantly due to the lift effect in the presence of the forward speed. It is well known that the roll damping moment of a ship is also strongly dependent on the forward speed [14].

In this paper, attention is focused on the real practical application of the stochastic inverse method [1] to identification of nonlinear damping moment of a ship moving with “non-zero” forward speeds. For this purpose, we first derive a stochastic inverse model by defining nonlinear damping moment as a nondeterministic parameter, which is multivariate random variable. This stochastic inverse modeling method is then applied to the laboratory tests to assess workability and practicability. Two different motions, transient motion and forced periodic motion, are considered with various trial conditions. The unique features of the method can be summarized in the following sense: firstly, it is nonparametric, that is, it does not require a prescribed form of nonlinearity unlike the conventional parametric

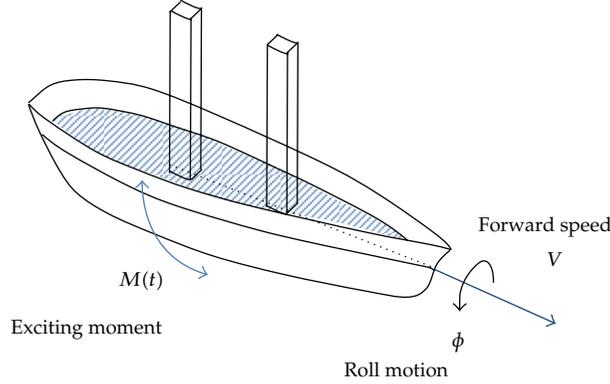


Figure 1: The schematic of the experiment.

methods. Secondly, it also offers a way of quantification of confidence level of identified solution given noisy data since the method is based on a stochastic inverse model.

The outline of this paper is as follows. The stochastic inverse model for nonlinear damping of a ship is derived in Section 2. Experimental setups are explained in Section 3. Section 4 presents analysis results of experimental data. Concluding remarks are made in Section 5.

2. Problem Formulation

2.1. Governing Equation for Ship Rolling Motion

It is assumed that the rolling motion ϕ of a ship moving at a forward speed V , illustrated in Figure 1, is governed by the following nonlinear differential equation of motion:

$$(I + \Delta I)\ddot{\phi}(t) + B(\phi, \dot{\phi}; t) + K(\phi; t) = M(t), \quad (2.1)$$

where I is the actual mass moment of inertia, ΔI is the added mass moment of inertia, $B(\phi, \dot{\phi})$ is the nonlinear roll damping moment, $K(\phi)$ is the restoring moment, and $M(t)$ is the roll-excitation moment. In the above equation, an overdot denotes a differentiation with respect to time.

The roll damping moment is expressed as a positive nonlinear function of the roll angle and angular velocity. The roll damping moment $B(\phi, \dot{\phi})$ is also affected by the forward speed of the ship since a lift effect occurs due to the presence of the forward speed. The restoring moment $K(\phi)$, which is expressed as an antisymmetric function of the roll angle, is induced by hydrostatic pressure exerted by a fluid at equilibrium due to the force of gravity. If the amplitude of roll angle is sufficiently small, then it is often expressed as the multiplication of the displacement of a ship and the distance between metacentric height and center of gravity GM , that is, $K(\phi) = \rho g \nabla \cdot GM\phi$. In this study, our attention is restricted to the small amplitude motion.

2.2. Inverse Setting for the Nonlinear Damping Moment

From the mathematical manipulation based on the concept of variation of parameters, the following relationship is obtained [1, 2, 19, 20]:

$$g(\phi; t) = \int_0^t \frac{x_1(\tau)x_2(t) - x_1(t)x_2(\tau)}{(I + \Delta I)W(\tau)} \{B(\phi, \dot{\phi}; t)\} d\tau, \quad (2.2)$$

where $W \equiv |x_1\dot{x}_2 - x_2\dot{x}_1|$ and x_1, x_2 are chosen to satisfy the following equations, respectively:

$$\begin{aligned} (I + \Delta I)\dot{x}_1 + \rho g \nabla \cdot \text{GM}x_1 &= 0, & x_1(0) &= 1, & \dot{x}_1(0) &= 0, \\ (I + \Delta I)\dot{x}_2 + \rho g \nabla \cdot \text{GM}x_2 &= 0, & x_2(0) &= 0, & \dot{x}_2(0) &= 1. \end{aligned} \quad (2.3)$$

The left hand side of (2.2) is given by

$$g(\phi; t) = \phi(t) - \alpha x_1(t) - \beta x_2(t) - \int_0^t \frac{x_1(\tau)x_2(t) - x_1(t)x_2(\tau)}{(I + \Delta I)W(\tau)} M(\tau) d\tau, \quad (2.4)$$

where $\alpha = \phi(0)$ and $\beta = \dot{\phi}(0)$.

The purpose of the study is to inversely identify the nonlinear damping moment given a measured response data ϕ . For a set of measured roll response data $\phi_i = \phi(t_i)$ during $0 < t < T$, the following system is given for the unknown nonlinear damping $B_j = B(\tau_j)$, $j = 1, \dots, n$:

$$g_i = \sum_{\tau_j < t_i} H_{ij} B_j, \quad i = 1, \dots, m, \quad (2.5)$$

where

$$H_{ij} = \Delta \tau \frac{x_1(\tau_j)x_2(t_i) - x_1(t_i)x_2(\tau_j)}{(I + \Delta I)W(\tau_j)}, \quad (2.6)$$

$$g_i = \phi(t_i) - \alpha x_1(t_i) - \beta x_2(t_i) - \sum_{\tau_j < t_i} \Delta \tau \frac{x_1(\tau_j)x_2(t_i) - x_1(t_i)x_2(\tau_j)}{(I + \Delta I)W(\tau_j)} M(\tau_j). \quad (2.7)$$

The identification of the nonlinear damping can be achieved by inverting the matrix system (2.5) from the observable parameter g_i . This identification procedure can be considered as the inverse problem [21–23] since its aim is to find the cause from the effect of the physically observable data.

It is worth noting that the system (2.5) is given by discretizing the first-kind integral operator (2.2). According to the inverse problem theory [21–23], the inverse of such systems is very sensitive to small changes in the data. The data-driven nature of the inverse problem makes the inverse analysis more difficult because one should use the measured data containing random errors arising from various sources of noise in the inverse setting. For such systems, which are often called ill-posed, the usual inverse procedure yields erroneous solutions which are physically meaningless.

2.3. Stochastic Inversion

To ensure a stable solution procedure, the unknown damping moment $B(\phi, \dot{\phi}; t)$ will here be modeled as a sequence of random variable $U(t; \xi)$, $U : \Omega \rightarrow \mathbb{R}^n$, where ξ is every possible outcome of an arbitrary nonempty sample space Ω . Then, the random variable U can be related to the directly observable quantity g by the Bayesian formula [23]:

$$p(U | g) \propto p(g | U)p(U), \quad (2.8)$$

where $p(g | U)$ is the likelihood, $p(U)$ is the prior probability density function, and $p(U | g)$ is the posterior probability density function.

Using adequate probabilistic models, the probabilistic expression (2.8) can be specified more clearly. In the case where the measurement error is independent additive Gaussian random noise with zero mean and standard deviation σ , the likelihood $p(g | U)$ has the form:

$$p(g | U) \propto (\sigma^2)^{-m/2} \exp\left(-\frac{\|HU - g\|_2^2}{2\sigma^2}\right), \quad (2.9)$$

where $\|\cdot\|_2$ refers to Euclidean norm, and m is the number of measurements. Using a pairwise Markov random field model [23–25], the prior $p(U)$ can be written as

$$p(U) \propto \lambda^{n/2} \exp\left(-\frac{\lambda}{2}U^T W U\right), \quad (2.10)$$

where the matrix $W \in \mathbb{R}^{n \times n}$ is defined by

$$W_{ij} = \begin{cases} n_i, & i = j, \\ -1, & i \sim j, \\ 0, & \text{otherwise,} \end{cases} \quad (2.11)$$

where n_i is the number of neighbors for the point i , and $i \sim j$ means that i and j are adjacent. Thus, (2.8) can be specified as

$$p(Ug) \propto (\sigma^2)^{-m/2} \exp\left(-\frac{\|HU - g\|_2^2}{2\sigma^2}\right) \lambda^{n/2} \exp\left(-\frac{\lambda}{2}U^T W U\right). \quad (2.12)$$

The parameters σ and λ for the probabilistic model are also nondeterministic and difficult to be known a priori. In the stochastic inversion, these uncertainties are naturally resolved by expanding into the hierarchical model [23–25]:

$$p(U, \lambda, \sigma | g) \propto (\sigma^2)^{-(m/2 + \alpha_2 + 1)} \exp\left(-\frac{\|HU - g\|_2^2 + 2\beta_2}{2\sigma^2}\right) \lambda^{(n/2 + \alpha_1 - 1)} \exp\left(-\frac{\lambda(U^T W U + \beta_1)}{2}\right), \quad (2.13)$$

where (α_1, β_1) is a pair of gamma distribution for λ , and (α_2, β_2) is a pair of inverse gamma distribution for the prior of σ . Based on the hierarchical model (2.13), it is possible to determine, at the same time, the stable inverse solution and the hyperparameters σ and λ through a numerical sampling technique such as Markov chain Monte Carlo [1, 23–25].

2.4. MCMC Simulation

To extract information on the damping moment from the constructed stochastic inverse model, it is necessary to employ the simulation technique such as Markov chain Monte Carlo, whose aim is to draw an identical independent distributed set of samples from a target density. The hierarchical model (2.13) with the given observable quantities can be explored by the following hybrid algorithm which is designed by mixing Metropolis-Hastings steps in the Gibbs sampler [1, 23–26].

- (i) Initialize $U^{(0)} = \{U_1^{(0)}, U_2^{(0)}, \dots, U_n^{(0)}\}$, $\lambda^{(0)}$, and $\sigma^{(0)}$.
- (ii) For $i = 0 : N_{\text{MCMC}} - 1$
 - Sample $U_1^{(i+1)}$ from $p(U_1 | U_2^{(i)}, U_3^{(i)}, \dots, U_n^{(i)}, \lambda^{(i)}, \sigma^{(i)})$
 - Sample $U_2^{(i+1)}$ from $p(U_2 | U_1^{(i+1)}, U_3^{(i)}, \dots, U_n^{(i)}, \lambda^{(i)}, \sigma^{(i)})$
 - \vdots
 - Sample $U_n^{(i+1)}$ from $p(U_n | U_1^{(i+1)}, U_2^{(i+1)}, \dots, U_{n-1}^{(i+1)}, \lambda^{(i)}, \sigma^{(i)})$
 - Sample u_1 from the uniform distribution on $[0, 1]$
 - Sample $\lambda^{(*)}$ from $q(\lambda^{(*)} | \lambda^{(i)})$
 - if $u_1 < \min\{1, p(\lambda^{(*)} | U^{(i+1)}, \sigma^{(i)})q(\lambda^{(i)} | \lambda^{(*)})/p(\lambda^{(i)} | U^{(i+1)}, \sigma^{(i)})q(\lambda^{(*)} | \lambda^{(i)})\}$, $\lambda^{(i+1)} = \lambda^{(*)}$
 - else $\lambda^{(i+1)} = \lambda^{(i)}$
 - Sample u_2 from the uniform distribution on $[0, 1]$
 - Sample $\sigma^{(*)}$ from $q(\sigma^{(*)} | \sigma^{(i)})$
 - if $u_2 < \min\{1, p(\sigma^{(*)} | U^{(i+1)}, \lambda^{(i+1)})q(\sigma^{(i)} | \sigma^{(*)})/p(\sigma^{(i)} | U^{(i+1)}, \lambda^{(i+1)})q(\sigma^{(*)} | \sigma^{(i)})\}$, $\sigma^{(i+1)} = \sigma^{(*)}$
 - else $\sigma^{(i+1)} = \sigma^{(i)}$.

The sampled set of realizations $\{U^{(i)}\}$ can then be used to approximate the statistics of the target densities $p(U, \lambda, \sigma | g)$ by

$$p(U, \lambda, \sigma | g) \approx \frac{1}{N_{\text{MCMC}}} \sum_{k=1}^{N_{\text{MCMC}}} \delta(U - U^{(k)}), \quad (2.14)$$

where N_{MCMC} is the number of Monte Carlo simulation and δ is Dirac delta function.

In summary, the damping moment of a ship moving with nonzero-forward speed can be identified through the following steps. (1) Derive the inverse model (2.5) relating the desired damping moment with the observable parameter, which is a function of the roll response; (2) measure the dynamic roll response of a ship moving at a nonzero-forward speed; (3) formulate stochastic inverse model (2.8) given the measured roll response; (4) use the designed MCMC algorithm to identify the damping moment.

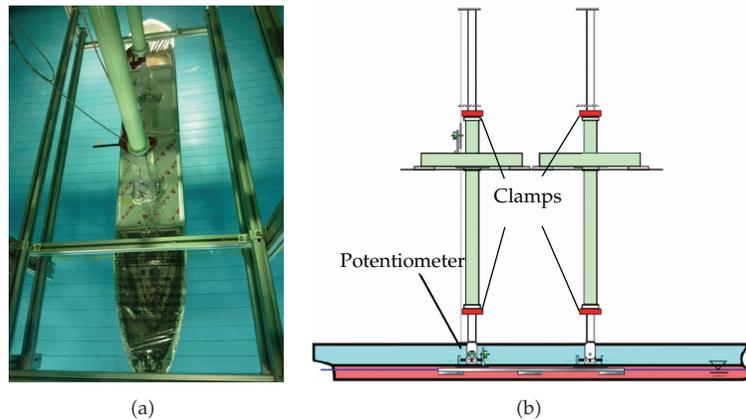


Figure 2: Overview of (a) the test model and (b) experimental setups.

Table 1: Particulars of the test model.

Length L_{pp}	2.500 m
Breadth B	0.387 m
Draft D	0.132 m
Displacement volume ∇	0.110 m ³
The distance between metacentric height and center of gravity GM	0.074 m
Vertical position of the center of buoyancy KB	0.071 m
Vertical position of the transverse metacenter above the keel line KM	0.179 m
Natural frequency ω_n	6.905 rad/s (without BK) 6.136 rad/s (with BK)

3. Experimental Setup

The workability and accuracy of the present method were verified by damping moment identification of the test model. The related experiments were conducted at the Ocean Engineering (OE) Basin of the University of Tokyo. The basin, which is often called as the towing tank, is a physical water tank to perform hydrodynamic tests with ship models. Figure 2 shows the test model used for the experimental analysis and an overview of experimental setups. The plan of the test model is shown in Figure 3. Table 1 summarizes particulars of the test model. During experiments, the test model was rigidly clamped in all degrees of freedom, excepting roll motion to minimize effects from the other modes of motions.

Figure 4 shows the layout of the OE Basin, whose length, breadth, and depth are 50 m, 10 m, and 5 m, respectively. The basin is equipped with a towing carriage that runs on two rails on either side of the basin at a speed from zero to 2 m/s. The carriage is also equipped with computers and devices to measure or control the speed of the towing carriage. The test model was initially positioned in one side of the basin and then towed by the carriage to the opposite side ranging from zero to 2 m/s. The roll motions are recorded by the potentiometer attached to the center of gravity of the test model.

The model was first tested without any appendages such as bilge-keels to consider the hull damping characteristics. After that, for the purpose of assessing the workability, two

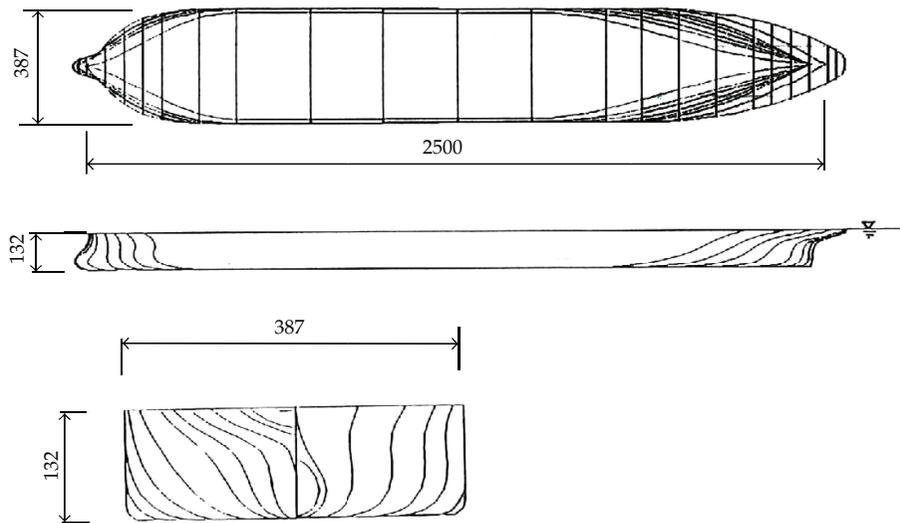


Figure 3: Body plan of the test model.



Figure 4: Layout of ocean engineering basin at the University of Tokyo.

bilge-keels (BK) with about 1 m long are attached to both sides of the test model at the turn of the bilge as in Figure 5. The installation of BK generates totally different roll characteristic since BK increases hydrodynamic resistance to rolling motion and makes the ship roll less. The effect of BK can easily observed even at zero-forward speed and becomes larger with the presence of the nonzero-forward speed.

For the experimental application, two different types of motions, time-dependent transient motion and frequency-dependent periodic motion, are considered. Firstly, the transient motion caused by an initial roll angle is considered. An external moment is first applied to the test model by static means to give an initial roll angle. Then the moment is eliminated and the decaying roll motion is measured. Secondly, the forced motion caused by a periodic excitation is considered. Monofrequency sinusoidal roll motion is first imposed by the vertical force generated by the force oscillating device in Figure 6. The resulting vertical force is recorded by the load cell and converted to the exciting moment by multiplying the moment arm.

It should be noted that a linear approximation for the restoring moment is valid only for sufficiently small roll angle. As pointed out earlier, in this study, we restrict our attention to the small amplitude motion so that the restoring moment can be expressed by the linear form.

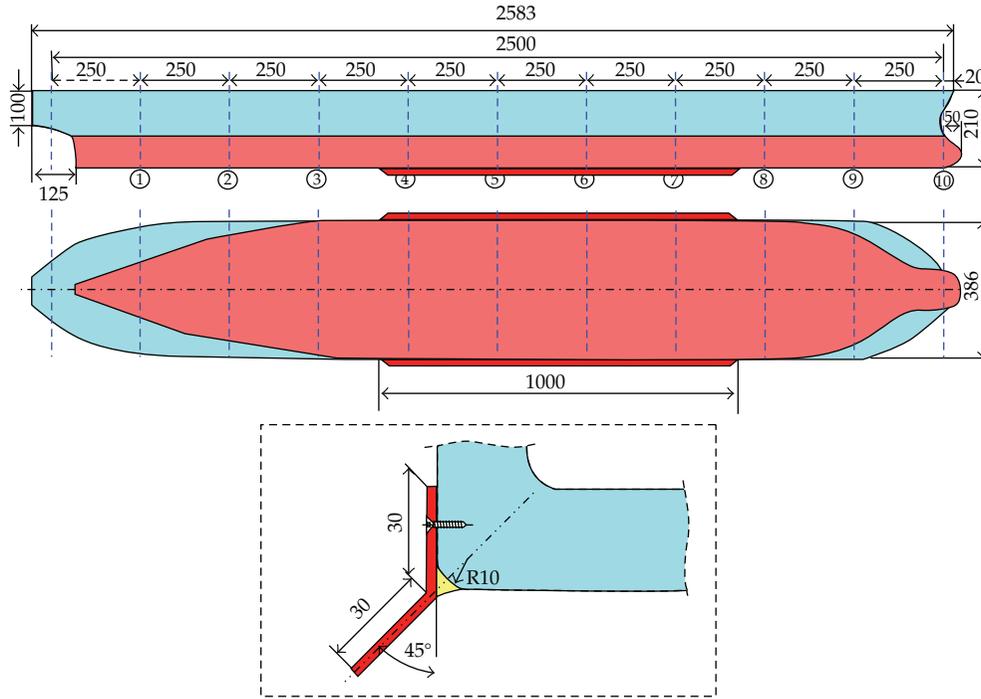


Figure 5: Attachment of bilge-keels to the both sides of the test model.

4. Analysis of Experimental Data

4.1. Transient Motion: Free-Decay Rolling

As a first application, a transient motion induced by an initial roll angle is considered. This motion is referred to as the free-decay rolling motion. For the trial, an initial roll angle is first given as $\phi(0) = \alpha$ while the test model is being towed by the carriage with a forward speed V . The static moment is then eliminated and the resultant response is recorded by measuring devices. This free-decay rolling motion is governed by the following initial value problem from (2.1) since $M(t) = 0$:

$$(I + \Delta I)\ddot{\phi} + B(\phi, \dot{\phi}) + K\phi = 0, \quad \phi(0) = \alpha, \quad \dot{\phi}(0) = 0. \quad (4.1)$$

Dividing both sides of (4.1) by $I + \Delta I$, we can obtain

$$\ddot{\phi} + \tilde{B}(\phi, \dot{\phi}) + \omega_n^2 \phi = 0, \quad (4.2)$$

where $\tilde{B} = B/(I + \Delta I)$ and $\omega_n = \sqrt{K/(I + \Delta I)}$.

Recorded roll responses are presented in Figure 7 for the test model without and with BK. Here, the Froude number (Fr) is defined by $V/\sqrt{L_{pp}g}$, where V is the velocity of the test model, g is the gravitational acceleration, and L_{pp} is the length of the ship. It can be observed that responses of the test model with BK decay to zero faster than the test model without BK.

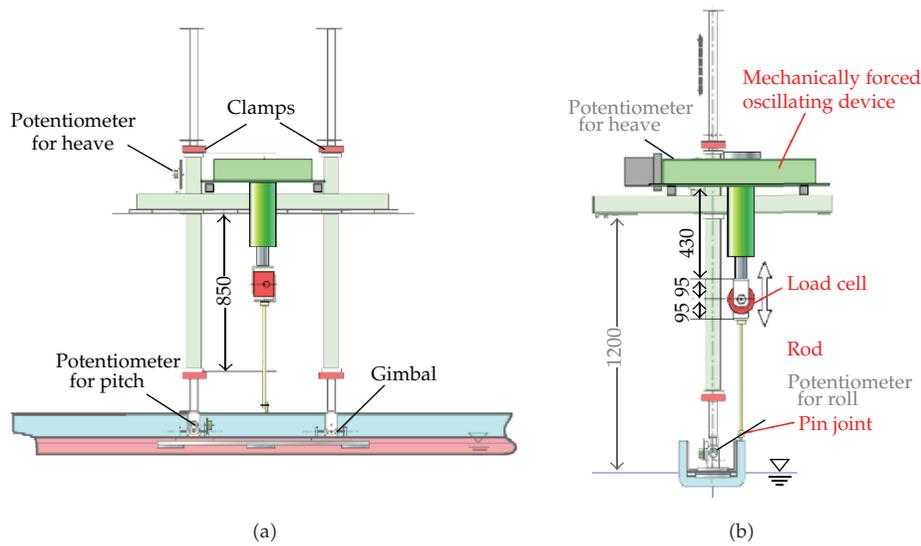


Figure 6: Forced oscillating devices: (a) side view. (b) front view.

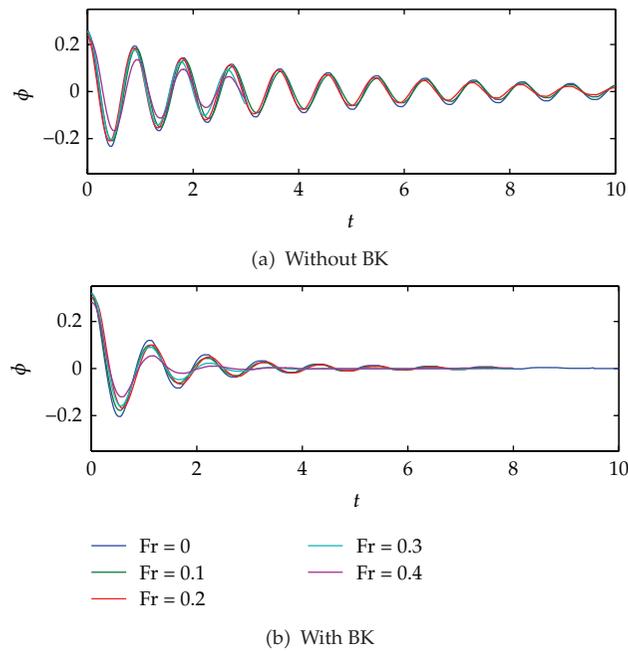


Figure 7: Roll decay curve of the test model for different forward speeds.

Moreover, the rate of decrease of roll angle becomes larger when the forward speed increases for both cases of the model without and with BK. This clearly shows that the damping is dependent on the forward speed. This fact is well known and described in various papers [14, 27].

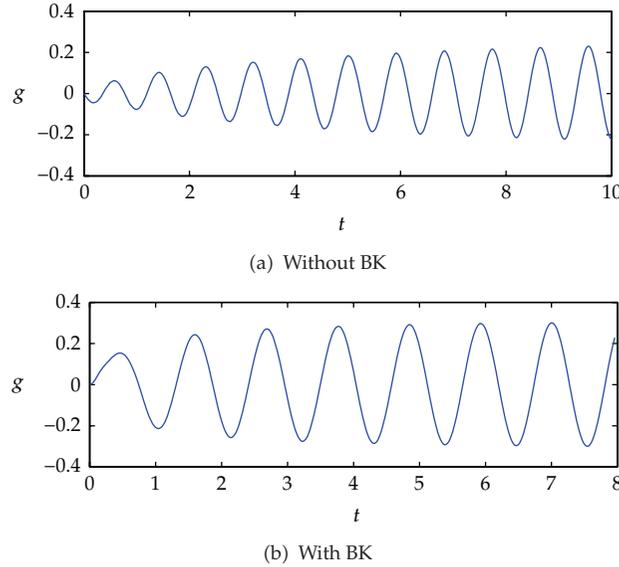


Figure 8: The converted quantity g from the measured roll-angle data when $Fr = 0.2$.

It is worthwhile to note that the physical coefficients are not necessary for the case of free-decay rolling. The only thing that is required for applying the present method is the natural angular frequency. The natural frequency for the test model is shown in Table 1. The variation of the roll natural frequency due to the addition of bilge keels is extremely large. In this study, for simplicity, the frequency of the largest Fourier component is considered as the natural frequency for the BK model.

To illustrate the method, through application to experimental data in Figure 7, a particular case of the roll-angle data with $Fr = 0.2$ was chosen. As a first stage of the identification, the roll-angle data set ϕ_i ($i = 1, 2, \dots, m$) is converted to a set of the observable quantity g_i using the relationship given by (2.7). This is clearly straightforward if the roll angle is recorded since the excitation moment $M(t)$ is zero for the case of the free-decay rolling motions. Consequently, the observable quantity can be computed by $g_i = \phi_i - \alpha x_1(t_i)$ and the results are shown in Figure 8. Using this quantity, we can construct a stochastic inverse model $p(U, \lambda, \sigma | g)$ for both cases of experiments.

Before illustrating results from MCMC simulation, the least-squares estimation for the inverse solution of (2.5) is first presented in Figure 9. The results clearly show the difficulties with the standard inverse solution to (2.5), that is, lack of stability in solution. The instability, frequent sign changes, is often encountered in solving data-driven inverse problems since the inevitable errors, no matter how small, incurred in taking and analyzing measured signal, are the main reason of this instability.

We consider now the stochastic inverse model $p(U, \lambda, \sigma | g)$ and its MCMC simulation for a stable and reliable solution. For the purpose of the simulation, $\lambda^{(0)} = 1$, $\sigma^{(0)} = 0.1$, and zero vectors for $U^{(0)}$ are used for the initial values of the algorithm. A small value 10^{-1} is chosen for the pairs of parameters (α_1, β_1) and (α_2, β_2) . For MCMC simulation, 50,000 samples were generated and the last 25,000 samples are used to estimate statistics such as mean and standard deviations for U . The MCMC results are illustrated in Figure 10.

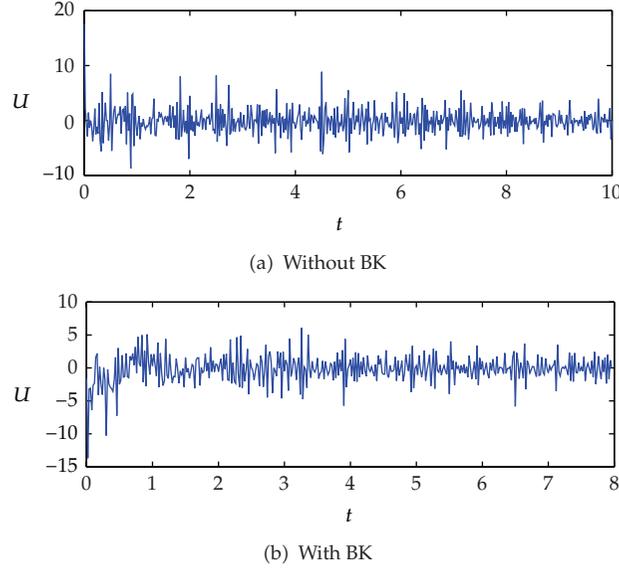


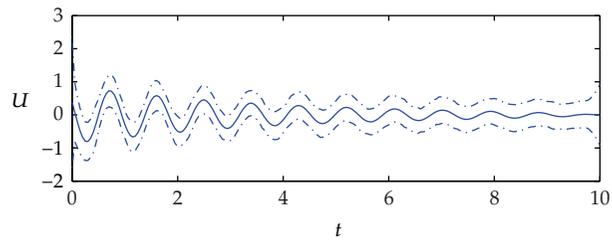
Figure 9: Inverse solutions through least-squares estimation.

The upper and lower dotted lines denote the 95% credible interval, quantifying the degree of uncertainties in the solution given the measurements. The mean estimate of the posterior density is fairly stable compared with the results in Figure 9. The result in Figure 10 describes the relative likelihood for the multivariate random variable U , which includes information of the unknown nonlinear damping moments. The posterior mean represents the most probable value of the stochastic inverse model $p(U, \lambda, \sigma | g)$. The confidence interval can be interpreted as an indicator of the reliability of the estimated value conditional on the measured data.

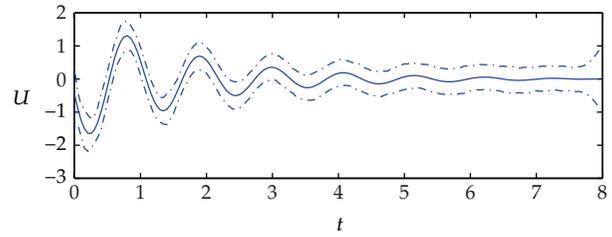
It is worth to note that MCMC simulation takes a while to properly sample the target distribution. In this study, we used the evolution of components to check if the chain works properly. Figure 11 shows an example of trace plots for the marginal distribution of the MCMC result for the trial with BK. Trace plots of each component are clearly in the burn-in phase, that is, stationary states. This implies that the posterior density $p(U, \lambda, \sigma | g)$ is successfully explored with the designed MCMC algorithm.

Once the probability density function $p(U, \lambda, \sigma | g)$ is estimated, the nonlinear damping moment can then be identified by $E[p(U(t; \xi) | g)] = -\tilde{B}(\dot{\phi}(t))$. The identified nonlinear roll damping moments are shown in Figure 12. It is worthwhile to note that the present method is fully nonparametric. Thus, the information of nonlinear damping moment is first revealed by a set of data based on the most probable value of the estimated function U . The set of scattered points is then used to specify the nonlinear model for damping moment. Here, the form $\tilde{B}(\dot{\phi}(t)) = B_1\dot{\phi} + B_3\dot{\phi}^3$ is used to fit the identified data. The units of the coefficients are $[s^{-1}]$ for B_1 and $[s]$ for B_3 .

It is also important to check the accuracy of the identified model. For this purpose, the roll motion is resimulated by using the identified damping moment. Figure 13 shows the comparison between the re-simulated and the measured roll responses for both cases. It is confirmed that the both trials are in well coincidence with the measured roll response.



(a) Without BK



(b) With BK

— Posterior mean
 - - - 95% credible interval

Figure 10: Solutions from stochastic inverse modeling with MCMC simulation.

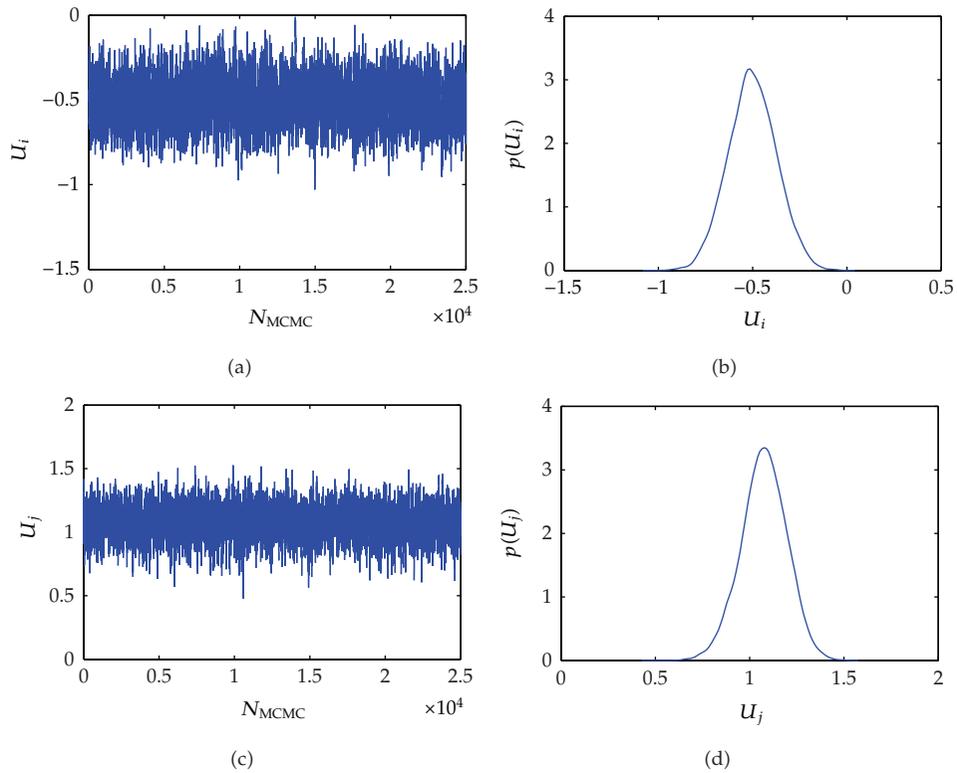


Figure 11: Typical examples of the trace plots for two components.

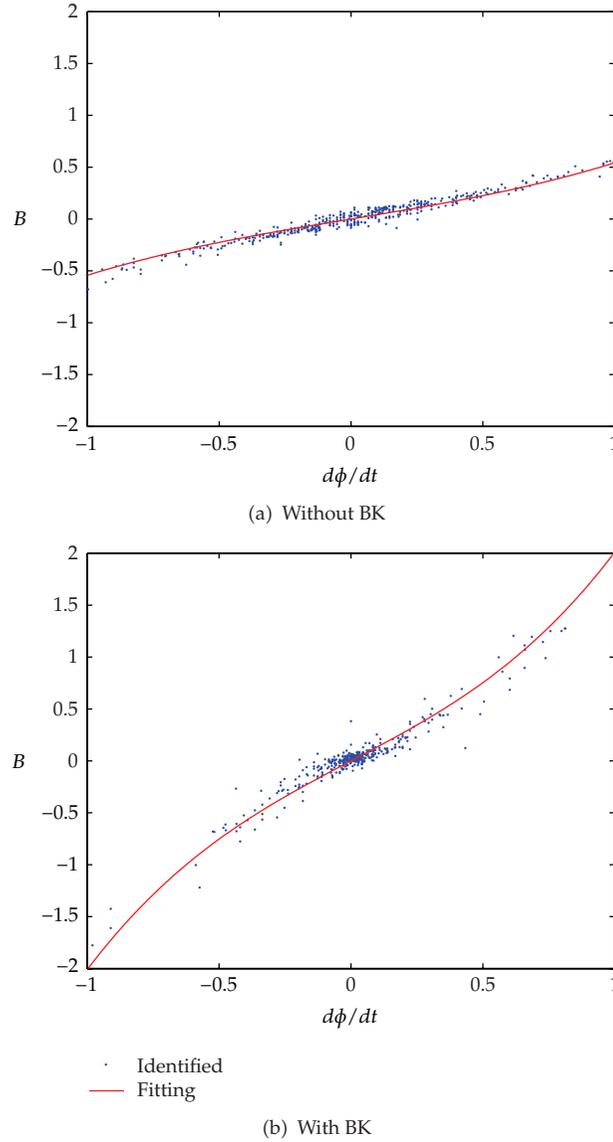


Figure 12: Identified nonlinear roll damping moments.

Finally, the preceding procedures are applied to all other experimental data for the case of free-decay rolling motion. The results are summarized in Table 2. The results clearly illustrate the effect of increasing forward speed on the roll damping moments. Furthermore, the damping moment obtained with the trials with BK is much larger than those of trials without BK.

It can be naturally concluded, based on the identified results in Table 2, that it would be possible to develop an empirical formula for the nonlinear damping moments as the form:

$$\tilde{B}(\dot{\phi}(t)) = B_1\dot{\phi} + B_3\dot{\phi}^3 = B_0\left(1 + \kappa(\text{Fr})^2\right)\dot{\phi} + B_3\dot{\phi}^3, \quad (4.3)$$

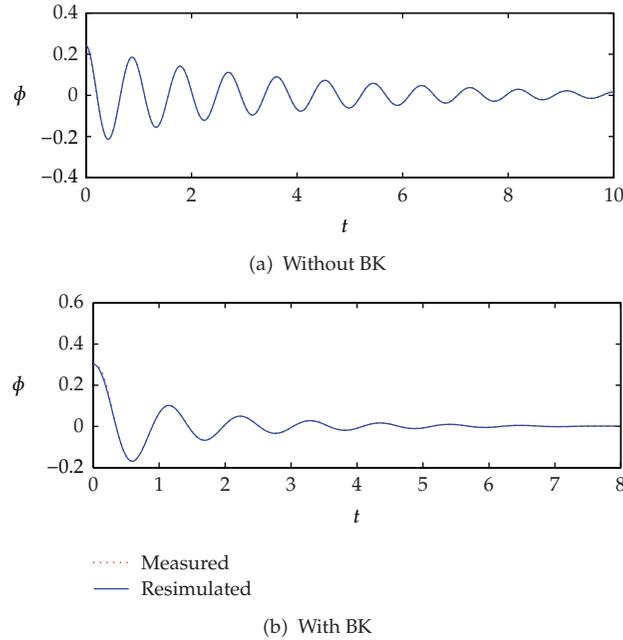


Figure 13: Resimulated roll response with the identified damping moment.

Table 2: Identified nonlinear damping for the case of free-decay rolling.

Fr	Trials with BK		Trials without BK	
	B_1	B_3	B_1	B_3
0	0.3796	0.1436	1.2408	0.6682
0.1	0.5070	0.1016	1.2475	0.4852
0.2	0.4203	0.1219	1.3262	0.6725
0.3	0.7208	0.2027	2.5105	0.4689
0.4	1.1007	0.1829	3.1812	0.6552

where B_0 is the linear contribution for zero-forward speed. The value of κ , which can be determined by fitting the identified results in Table 2, was turned out to be $\kappa = 4.7536$ for the non-BK model and $\kappa = 13.097$ for BK model, see Figure 14. It can be found that the effect of forward speed is much greater in the case of the BK model.

4.2. Periodic Motion: Forced Rolling Motion

As a second application, a periodic forced motion induced by periodic excitation is considered. The monofrequency periodic motion is imposed while ship is moving with a forward speed V :

$$\phi(t) = A_1 \cos(\omega t). \tag{4.4}$$

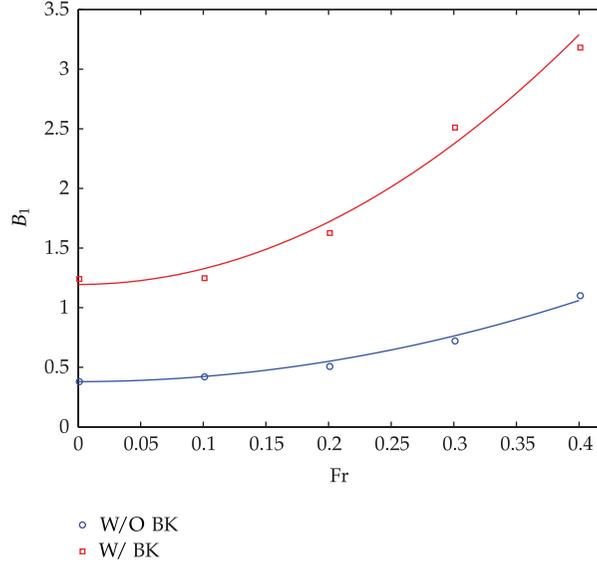


Figure 14: Determination of the coefficient κ for empirical formula.

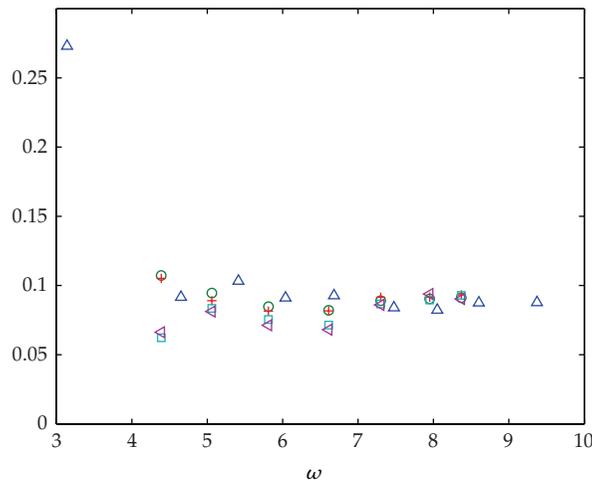
The corresponding roll exciting moment $M(t)$ is then measured with the load cell as explained in Section 3. It is worth to point out that, in applying the proposed method to forced rolling motion, it is first necessary to know roll moment of inertia unlike free-roll decay test. The values were determined a priori through the related tests.

For the forced oscillation test, the frequency-dependent coefficients of the roll moment of inertia are generally obtained by Fourier analysis:

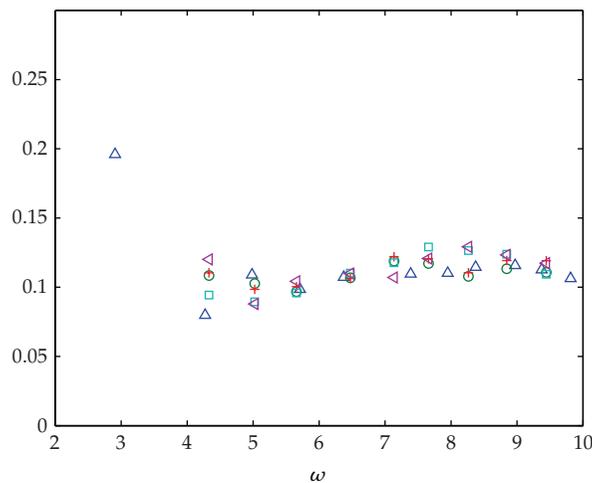
$$(I + \Delta I) = \frac{KA_1 - M_a \cos(\varepsilon)}{\omega^2 A_1}, \quad (4.5)$$

where M_a is the amplitude of the exciting moment, ε is the phase difference between the roll angle and the exciting moment, and ω is the exciting frequency. The obtained coefficients are illustrated in Figure 15.

Now we are ready to apply the present method to the measured data. For illustrating purposes, a particular case of experimental data was chosen as identification examples for the forced rolling motion. Figure 16 shows the measured roll response and the exciting moment for the test model with BK when $Fr = 1.0$ and exciting frequency $\omega = 5.12$. Figure 17 shows the step-by-step results of identifying roll damping moment from the measured data. The measured data was first converted to the observable parameter g through (2.7). The converted quantity, illustrated in Figure 17(a), leads to the construction of stochastic inverse model as in (2.13). Without loss of generality, the same conditions were used to explore the constructed stochastic inverse model for the MCMC algorithm as explained in Section 4.1. The MCMC results were shown in Figure 17(b). Next, the roll response is resimulated with the identified inverse solution. The result is shown in Figure 17(c). It is observed that the resimulated response is in good agreement with the measured roll response. This proves the accuracy of the identified roll damping moment.



(a) Without BK



(b) With BK

Figure 15: Nondimensional moment of inertia of the test model: $(I + \Delta I)/mB^2$.

It is worth noting here that, for the case of free-decay rolling motion, the effects on roll amplitude in a roll-decay are mainly due to the damping [28], but the same cannot be said in the case of forced rolling motion. It could be expected that the nonlinearity in the restoring plays a role and the estimated damping moment is not dependent only on the roll angular velocity. Consequently, it is difficult to build an analytical model of the identified nonlinear damping moment unlike the case of free-decay rolling motion as described in Section 4.1.

Instead, we performed here a quantitative analysis of the results obtained by applying the preceding procedure to all other experimental data of the forced rolling motion. The rationale behind this is that the relating results can be considered to be periodic with the

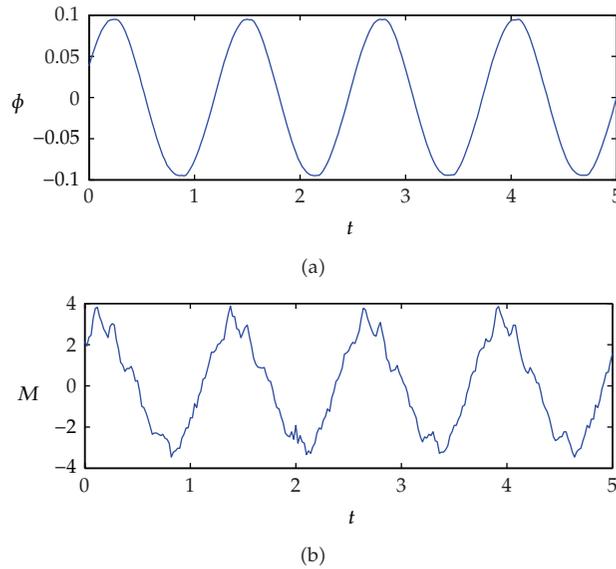


Figure 16: Measured roll response and exciting moment.

exciting frequency since the monofrequency periodic motion was imposed. It is convenient to illustrate and compare the results in terms of frequency for this case. Figures 18 and 19 show the peak values of the identified solution from the present stochastic identification procedure. For all cases, the same amplitude of the motion is imposed. It should be noted that the function U includes the information on the damping moment. That is, the results in Figures 18 and 19 can be explained by the fact that the damping moment is dependent on the forward speed and the exciting frequency.

5. Conclusions

In this paper, a stochastic inverse method has been investigated for the identification of roll characteristics of a ship moving at nonzero-forward speeds. The rolling motion has been treated as a single-degree-of-freedom nonlinear equation of motion, uncoupled from other motions. On this basis, the stochastic inverse model for the nonlinear damping moment was derived as a probabilistic expression in terms of the observable parameter which is a function of the measurements of the roll angle and excitation. The stochastic inverse model contains the information of the nonlinear damping contribution as the multivariate random variables. Given measured data, the nonlinear damping moments were identified through the designed Markov chain Monte Carlo algorithm.

To ensure applicability, the proposed method has been applied to the experimental data for the two different mechanical phenomena regarding ship roll motions, that is, the transient motion and forced periodic motion. In nonlinear system identification, it is difficult to define the quality of the identified results because it depends on its purposes. The aim of the present study is to find the nonlinear system model which can reproduce the measured system response. In this sense, it can be concluded that the proposed method

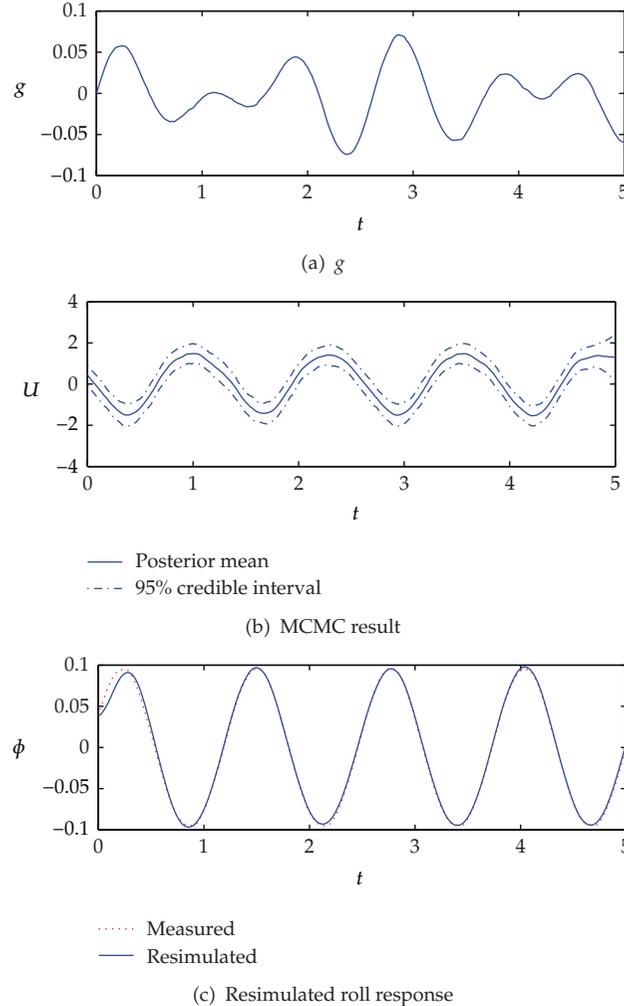
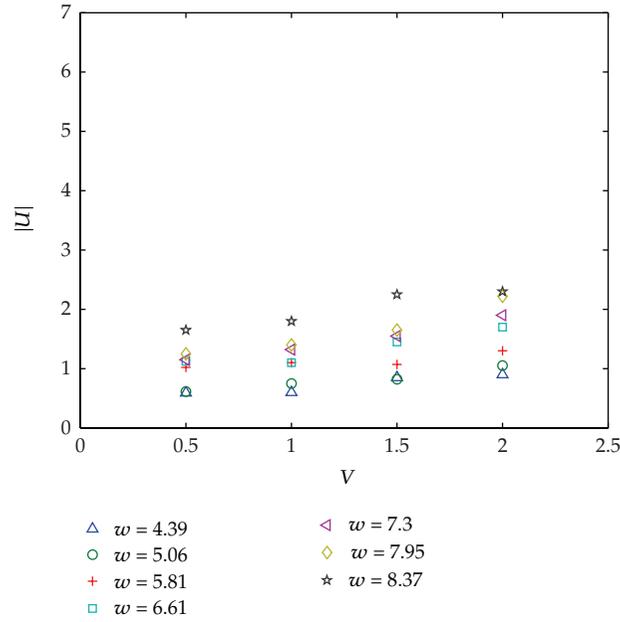


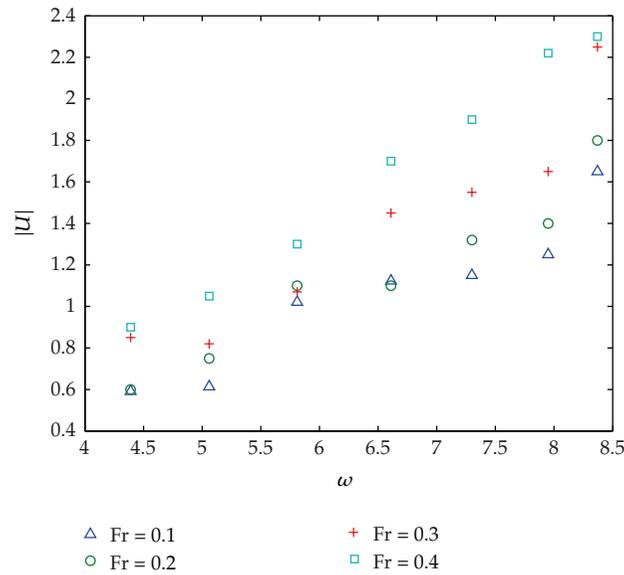
Figure 17: Step-by-step result for the identification of roll damping moment.

can accurately identify the nonlinearity in damping of a ship moving at nonzero-forward speeds.

The preset stochastic inverse method has the following limitations. Firstly, the method is derived based on the assumption of small amplitude of the rolling motion so that the restoring moment can be approximated by the linear form. In reality, the restoring moment is also nonlinear. The nonlinear contribution cannot be neglected. The extension to a system with the nonlinear restoring is not difficult mathematically. However, the new formulation needs additional information on the restoring nonlinearity for the identification purpose. Secondly, the experimental setups do not strictly reflect the actual motion of the ship. Rolling motion is generally coupled with other motions, such as sway and heave. In reality, the coupling effects should also be taken into account. However, for simplicity of the experimental application, in this paper, the test model was restrained in all degrees of motion except the roll motion while the model was moving at a constant speed.



(a) Identified solution versus forward speed



(b) Identified solution versus exciting frequency

Figure 18: Illustration of the peak value of the identified result for the trial without BK.

Acknowledgments

The authors would like to thank the editor and anonymous reviewers for their valuable comments, suggestions, and constructive criticism, which were very helpful in improving the paper substantially.

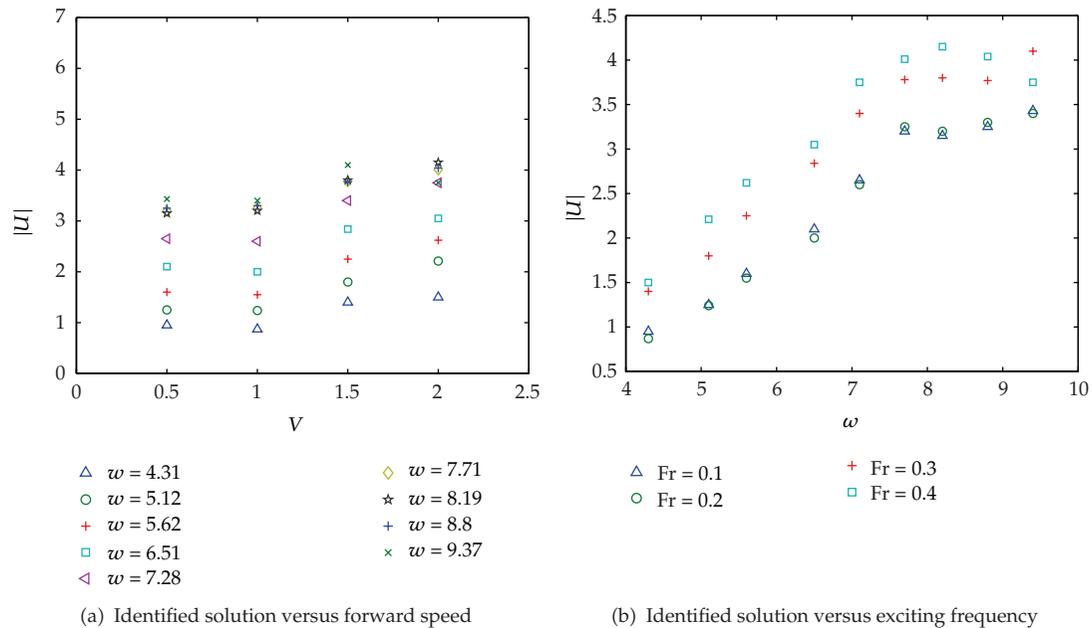


Figure 19: Illustration of the peak value of the identified result for the trial with BK.

References

- [1] S. L. Han and T. Kinoshita, "Nonlinear damping identification in nonlinear dynamic system based on stochastic inverse approach," *Mathematical Problems in Engineering*, vol. 2012, Article ID 574291, 2012.
- [2] S. L. Han and T. Kinoshita, "Stochastic inverse modeling of nonlinear roll damping moment of a ship," *Applied Ocean Research*, vol. 39, pp. 11–19, 2013.
- [3] R. Bhattacharyya, *Dynamics of Marine Vehicles*, John Wiley and Sons, New York, NY, USA, 1978.
- [4] S. N. Blagoveshchensky, *Theory of Ship Motions, vols. I and II*, Dover Publications, New York, NY, USA, 1962.
- [5] J. O. De Kat and J. R. Paulling, "The simulation of ship motions and capsizing in severe seas," *Transactions of SNAME*, vol. 117, pp. 127–135, 1989.
- [6] A. Zborowski and M. Taylan, "Evaluation of small vessels' roll motion stability reserve for resonance conditions," in *Proceedings of the 14th Ship Technology and Research STAR Symposium 1989 in conjunction with the SNAME Spring Meeting*, New Orleans, La, USA, April 1989.
- [7] J. A. Witz, C. B. Ablett, and J. H. Harrison, "Roll response of semisubmersibles with non-linear restoring moment characteristics," *Applied Ocean Research*, vol. 11, no. 3, pp. 153–166, 1989.
- [8] J. F. Dalzell, "A note on the form of ship roll damping," *Journal of Ship Research*, vol. 22, no. 3, pp. 178–185, 1978.
- [9] J. B. Roberts, "Estimation of nonlinear ship roll damping from free-decay data," *Journal of Ship Research*, vol. 29, no. 2, pp. 127–138, 1985.
- [10] J. B. Mathisen and W. G. Price, "Estimation of ship roll damping coefficients," *RINA Transaction*, vol. 127, pp. 169–186, 1985.
- [11] D. W. Bass and M. R. Haddara, "Nonlinear models of ship roll damping," *International Shipbuilding Progress*, vol. 35, no. 401, pp. 5–24, 1988.
- [12] J. R. Spouge, "Non-linear analysis of large-amplitude rolling experiments," *International Shipbuilding Progress*, vol. 35, no. 403, pp. 271–320, 1988.
- [13] H. S. Y. Chan, Z. Xu, and W. L. Huang, "Estimation of nonlinear damping coefficients from large-amplitude ship rolling motions," *Applied Ocean Research*, vol. 17, no. 4, pp. 217–224, 1995.
- [14] Y. Himeno, "Prediction of ship roll damping—a state of the art," Tech. Rep. 239, Department of Naval Architecture and Marine Engineering, University of Michigan, Ann Arbor, Mich, USA, 1981.

- [15] M. Taylan, "The effect of nonlinear damping and restoring in ship rolling," *Ocean Engineering*, vol. 27, no. 9, pp. 921–932, 2000.
- [16] M. R. Haddara and M. Hinchey, "On the use of neural network techniques in the analysis of free roll decay curves," *International Shipbuilding Progress*, vol. 42, pp. 166–178, 1995.
- [17] M. R. Haddara and M. Wishahy, "An investigation of roll characteristics of two full scale ships at sea," *Ocean Engineering*, vol. 29, no. 6, pp. 651–666, 2002.
- [18] J. B. Roberts and M. Vasta, "Markov modelling and stochastic identification for nonlinear ship rolling in random waves," *Philosophical Transactions of the Royal Society A*, vol. 358, no. 1771, pp. 1917–1941, 2000.
- [19] T. S. Jang, H. S. Choi, and S. L. Han, "A new method for detecting non-linear damping and restoring forces in non-linear oscillation systems from transient data," *International Journal of Non-Linear Mechanics*, vol. 44, no. 7, pp. 801–808, 2009.
- [20] T. S. Jang, J. W. Son, S. L. Han, H. G. Sung, S. K. Lee, and S. C. Shin, "A numerical investigation on nonparametric identification of nonlinear roll damping moment of a ship from transient response," *The Open Ocean Engineering Journal*, vol. 3, pp. 100–107, 2010.
- [21] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, vol. 120, Springer, New York, NY, USA, 1996.
- [22] C. W. Groetsch, *Inverse Problems in the Mathematical Sciences*, Informatica International, 1993.
- [23] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, New York, NY, USA, 2005.
- [24] J. Wang and N. Zabaras, "Hierarchical Bayesian models for inverse problems in heat conduction," *Inverse Problems*, vol. 21, no. 1, pp. 183–206, 2005.
- [25] J. Wang and N. Zabaras, "A Markov random field model of contamination source identification in porous media flow," *International Journal of Heat and Mass Transfer*, vol. 49, no. 5-6, pp. 939–950, 2006.
- [26] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to MCMC for machine learning," *Machine Learning*, vol. 50, no. 1-2, pp. 5–43, 2003.
- [27] M. R. Haddara and S. Zhang, "Effect of forward speed on the roll damping of three small fishing vessels," *Journal of Offshore Mechanics and Arctic Engineering*, vol. 116, no. 2, pp. 102–108, 1994.
- [28] G. Bulian, "Estimation of nonlinear roll decay parameters using an analytical approximate solution of the decay time history," *International Shipbuilding Progress*, vol. 51, pp. 5–32, 2004.

Research Article

A Review of Soft Techniques for Electromagnetic Assessment of Concrete Condition

**F. A. A. Queiroz,¹ D. A. G. Vieira,¹ X. L. Travassos,²
and M. F. Pantoja³**

¹ Parque Tecnológico de Belo Horizonte, Rua Professor José Vieira de Mendonça N 770, Salas 406 e 407, Engenho Nogueira, 31310-260 Belo Horizonte, MG, Brazil

² Departamento Regional da Bahia, Centro Integrado de Manufatura e Tecnologia, SENAI, Avenida Orlando Gomes N 1845, Piatã, 41650-010 Salvador, BA, Brazil

³ Departamento de Electromagnetismo y Física de la Materia Facultad de Ciencias, Universidad de Granada Campus de Fuentenueva, Severo Ochoa s/n, 18071 Granada, Spain

Correspondence should be addressed to F. A. A. Queiroz, francisco.queiroz@enacom.com.br

Received 26 July 2012; Revised 5 November 2012; Accepted 14 November 2012

Academic Editor: Fatih Yaman

Copyright © 2012 F. A. A. Queiroz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper describes some efforts to design more efficient analysis for the GPR raw-data interpretation. Such analysis requires algorithms by which problems having complex scattering properties can be solved as accurately and as quickly as possible. This specification is difficult to achieve when dealing with iteratively solved algorithms characterized by a forward solver as part of the loop, which often makes the solution process computationally prohibitive for large problems. The inverse problem is solved by an alternative approach which uses model-free methods based on example data. Measurements with GPR equipment were performed to validate the algorithms using concrete slabs.

1. Introduction

Concrete is a highly versatile construction material well suited for many applications. Strength, durability and many other factors depend on the relative amounts and properties of the individual components. It is almost impossible to predict its final electrical characteristics at the job site. It is electrically nonconductive but usually contains significant amounts of steel reinforcement. While concrete is noted for its durability, it is susceptible to a range of environmental degradation factors, which can limit its service life.

Throughout the literature, radar technology proved to be the cheapest and easiest method for mapping reinforcements but neither characterization of flaws by dimension and material nor crack detection could be demonstrated [1–5]. The concept of using Ground

Penetrating Radar (GPR) for the nondestructive testing (NDT) of concrete structures has grown in the past years, involving researchers from many different disciplines. The interest in this technique can be explained, basically, by its advantages compared to other techniques such as portability, relative low cost and weight, the ease of application to different surfaces. However, the imaging of typical field data may be difficult due to problems like limited coverage, noisy data, or nonlinear relations between observed quantities and physical parameters to be reconstructed. Several algorithms for reconstructing an unknown object have been implemented, most based on Newton-type iteration schemes what requires significant computational burden and a high level of complexity.

Nevertheless, radar has significant potential for development by way of software for signal and image processing to improve resolution solving an inverse problem.

The inverse problem consists in using the results of actual observations to infer the values of the parameters characterizing the system under investigation. Estimating inclusions characteristics using measured signals obtained by scanning antennas is known as an ill-posed inverse problem. Electromagnetic wave propagation inverse problems are typically ill-posed, as opposed to the well-posed problems more typical when modeling physical situations where the model parameters or material properties are known.

Some different reference techniques of non-destructive evaluation of reinforced concrete structures were collected in the review book [6]. One of those works is about migration algorithms [7], which can be used in simulated and real GPR data to focus the information contained in hyperbola, facilitating the visual location and assessment of inclusions in concrete.

Kirchhoff migration and Synthetic Aperture Focusing Technique (SAFT) [8] help to identify inclusions in the host medium raw data. The latter method tries to enhance the spatial resolution of inclusions whereas the first method concentrates in one point the consecutive hyperbolic echoes of a cylindrical inclusion.

For example, in [9, 10], a migration algorithm was used to define the number of inclusions and locate their centers inside nonhomogeneous host medium. Although that type of medium may affect the results of these algorithms, the obtained information was then used in the subsequent phase wherein the geometry and the electrical characteristics of inclusions were estimated by other method.

The general problem of testing concrete using subsurface radar including assisting with intelligent systems is addressed in [11]. Emphasizing the interest in developing methods to solve one of these inverse problems, many works including Computational Intelligence were reviewed in [12] with the specific objective of rebar detection in concrete using GPR, which is the same topic treated in [13] using Neural Networks. A Multilayer Perceptron model was used in [14] as a simplified hyperbolic shape detector in GPR images to identify steel reinforcing bars in concrete. As plastic pipes can be found inside concrete structures too, their identification in GPR images is a similar problem whose solution is proposed in [15] using multi-agent systems.

In [16], three works using Neural Networks in GPR data sets to inspect highways, bridge decks, and pavements are listed and can find applications in these structures sometimes made of concrete. This is the case considered in [17] to be solved by using Independent Component Analysis and fractals to detect defects in bridge decks. A Fuzzy System was developed in [18] to merge data from GPR and infrared and digital inspection system in order to detect problems in mountain tunnel lining inspected. In [19], water and chloride contents were predicted by Neural Networks from radar data to evaluate physical

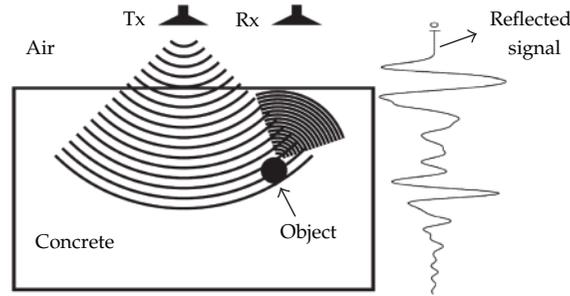


Figure 1: GPR inverse problem description.

condition of concrete. Thus, much effort has been done to improve automated concrete inspection by developing solutions for this inverse problem.

The aim of this paper is to expose efficient inverse methods to GPR raw data of concrete. The analysis steps of raw radargrams involve a priori information, subjective criteria, and indefinite sequences of data processing steps due to specific set of scan conditions. That makes it hard to completely characterize the inclusions in concrete using simple models. Therefore, this work describes alternative ways of solving the inverse problem of retrieving buried inclusions in dielectrics with electrical characteristics similar to concrete slabs.

The information extracted from the inverse models is used to estimate the electrical parameters of the probed materials inside concrete. It may help the development of software to aid analysis of GPR data from non-destructive evaluation of concrete structures.

The inverse problem considering the GPR assessment of concrete is outlined. The concept of inverse problem is defined with respect to the inclusion reconstruction in concrete slabs. Three reconstruction algorithms are implemented and experimental tests are performed to investigate the performance of the proposed algorithms.

2. Problem Settings

The objective of this work is the inverse problem of finding the characteristics of a target buried in a dielectric material given the reflected field measured by the antenna (Figure 1). In practice, the reflected signal is a collection of discrete observations in time. For the radar assessment of concrete, the objective is to determine a finite number of parameters. The parameters needed to characterize inclusions in a dielectric slab are found by identifying electrical (permittivity and conductivity) and geometrical (depth and radius) properties.

Such a problem is called discrete inverse problem or parameter estimation problem. In general, this is a difficult problem because the obtained information is not sufficient for estimation, which requires some a priori information about the inclusions. On the other hand, a large number of data can bring the problem to instability.

One possible technique to overcome these shortcomings is the use of parametric algorithms which are based on optimization algorithms. This kind of algorithms update the parameters iteratively to minimize (or maximize) a certain evaluation function. The computation of an approximate electric or magnetic field is done by the finite-difference time-domain (FDTD) method as a forward solver.

The optimization process can be carried out by using a variety of algorithms such as the Newton method, conjugate gradient method [20], or evolutionary algorithms. One major

limitation of using parametric algorithms is the time of calculation that can be prohibitive for 3D problems. Nonparametric algorithms can also be used. Usually, they are complex to implement but they can solve inverse problems faster than parametric algorithms because it is not necessary to iteratively evaluate the function.

Those approaches are less efficient in the sense of using all the available raw data without extract relevant prior information. So, they do not contribute to the development of more efficient models that are closer to white boxes. An alternative way is to use feature selection algorithms, such as Simba and Relief [21], to select characteristics from a scattered GPR wave. These are based on the concept of margin to define the relevant features of the data.

3. Parameter Estimation Approaches

3.1. Model Fitting

Model fitting method is one of the parametric imaging algorithms which utilize the evaluation function and optimization algorithms. In the model fitting method, target characteristics and location are expressed as parameters. The parameters are updated iteratively to minimize the difference between the observed data and the estimated data.

An incident wave and a scattered wave can be used to characterize the scattering object. Usually, in real world problems, the incident and scattered waves are known and it is desired to identify the scattering object. This problem can be written as an optimization problem involving the scattered wave of the unknown object $E(\theta_0)$, the reference object, and the scattered wave of a test object $E(\theta)$ [9].

Thus, θ^* , the optimum θ , is the argument that minimizes the error of the reference object scattered wave $E(\theta_0)$ relative to the test object scattered wave $E(\theta)$. Mathematically:

$$\theta^* = \arg \min f(\theta) = \sum_{i=1}^{ns} (E(\theta_0) - E(\theta))^2, \quad (3.1)$$

where ns is the number of sample points where the scattered wave is measured. Note that $E(\theta_0)$ is known even though θ_0 is unknown. The scattered field $E(\theta)$ is then generated assuming one tests θ , and the optimization procedure aims at minimizing the error between $E(\theta_0)$ and $E(\theta)$ in such a way as to identify the scatter object θ_0 .

The problem described in (3.1) is usually multimodal, as shown in Figure 2, where the unknowns are the radii of two inclusions given their physical properties. This multi-modal characteristic motivates the use of a stochastic approach instead of a deterministic one. This problem was solved in this paper using the Particle Swarm Optimization (PSO), which is described next.

The Particle Swarm Optimization (PSO) algorithm is similar to genetic algorithms (GAs) due to the random initialization. The first difference is that each potential solution is called particles, instead of individuals, and they “fly” on the search space. To each particle of the swarm during the iterations, the position of the best solution found to a given particle, called *pbest* (particle best), is saved.

The best value found considering all the particle is also saved, and is called *gbest* (global best). At each iteration, the PSO is based on the change in the particle’s velocity in the direction of its *pbest* and *gbest*, weighted by random terms.

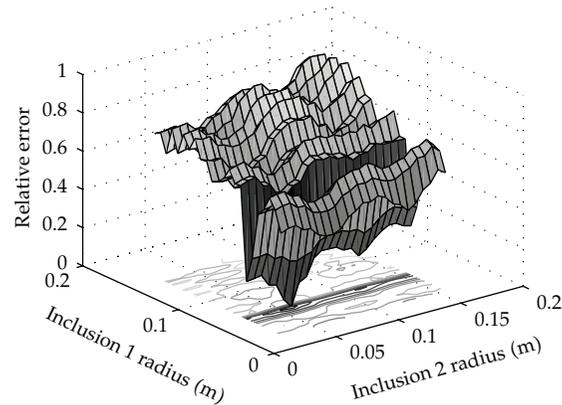


Figure 2: Surface plot of (3.1) representing the problem of finding the radii of two inclusions, given that their location and physical properties are known. Even in this simple example the surface is clearly multi-modal, thus, stochastic methods must be employed in the parametric approach.

Table 1: Inverse scattering problem definitions.

	Depth (mm)	Radii (mm)
Min	50	25
Max	250	100
Ref. Object	100	45

In our radar problem, the unknown object was obtained by experiment at depth of the inclusion equal to 100 mm, the radius equal to 45 mm and the standard deviation of the nonhomogeneous medium, sd , equal to 0.15. The target considered was water. The definition of the reference object as well as the range of the variables for the optimization process are summarized in Table 1.

The PSO was initialized using 50 particles over 50 iterations. The *gbest* evolution throughout the iterations is shown in Table 2. In the 50th iteration, the algorithm was capable of finding a solution very close to the desired one.

Nonetheless, good approximations were already available in the 5th iteration; thus, if a very accurate outcome is not necessary the algorithm would need only about 5 iterations to converge. It was studied in [10] an algorithm to mix migration algorithms with PSO in order to minimize the computational burden of this kind of investigation.

Although this methodology for model fitting may get better results by exploring a wide range of values for the independent parameters, better results are not guaranteed and the computer resources (hardware and time) needed to optimize the problem may not be available.

3.2. Artificial Neural Networks

Several imaging or inversion techniques have been developed to refocus the scattered signals back to their true spatial location and then decrease the interpretation time for effective maintenance and/or repair. Among them the Artificial Neural Networks (ANN) has been

Table 2: PSO results.

Iteration	Depth (mm)	Radii (mm)
1	87	32
5	97	44
10	97	47
50	100	44.4
Ref. Object	100	45

proven to be a promising technique in the solution of inverse electromagnetic field problems [22].

In this section it is addressed a 2D problem where a cylinder of unknown characteristics is buried in a nonhomogeneous dielectric. The incident and scattered wave are simulated using FDTD to train the ANN. The dielectric medium uses the electrical characteristics of concrete [1] with a mean relative electrical permittivity value of 6 and standard deviation 0.15, that is, a non-homogenous medium [10].

The investigation domain is illuminated by a differentiated Gaussian pulse with a center frequency of $f = 900$ MHz and bandwidth between 0.3 and 2 GHz. In order to control the numerical dispersion and provide good discretization for the inclusions the spatial steps were chosen as $\Delta x = \Delta y = 6$ mm. The aim of this problem is, given an incident wave, W_i , and scattered wave, W_s , to determine the radius and depth of the inclusion.

To train the ANN a set of different inclusion examples, say $S(\cdot)$, for different radii, depths, ϵ_r , and σ were generated.

The ANN has been trained with a set of different inclusion examples, constructed by varying the radius in the range [0.02–0.1] m according to the rule $\text{radius} = 0.02 + i \times 0.001$, $i = 0, \dots, 80$, with ϵ_r in the range [1–10], according to the rule $\epsilon_r = 1 + i \times 1$, $i = 0, \dots, 9$, σ in the range [0–4000] S/m according to the rule $\sigma = 0 + i \times 500$, $i = 0, \dots, 8$ and depth in the range [0.05–0.25] m according to the rule $\text{depth} = 0.05 + i \times 0.025$, $i = 0, \dots, 9$ summing a total of 1640 examples.

The W_s is the only information available in the real cases; therefore, it has to be used to characterize the inclusion. In this paper 1200 time steps were considered; thus, a problem in \mathbb{R}^{1200} must be solved. The ANNs suffer from a phenomenon called the curse of dimensionality [23], that is, the learning process becomes slower and less effective.

The dimensions of data features have great impact in samples density in input dimension and in the number of parameters in the structure of a chosen model. So, the samples with more features get more distant from each other and even if numerous samples are available in information extraction process, the number of features may require great computer resources to represent and compute model structure during parameters tuning in training.

To reduce the dimensionality of the problem Principal Component Analysis (PCA) was used. The main advantage of PCA is that once these patterns in the data are found, the data can be compressed, that is, by reducing the number of dimensions, without much loss of information. This technique is commonly used in image compression.

The input vectors are first normalized so that they have zero mean and unity variance. For PCA to work properly, one has to subtract the mean from each of the data dimensions. The PCA uses a linear mapping of a given set of samples $S_q = \{x_1, \dots, x_T \mid x_i \in \mathbb{R}^p\}$ to construct a new data set $S_p = \{y_1, \dots, y_T \mid y_i \in \mathbb{R}^q\}$, where $q \leq p$.

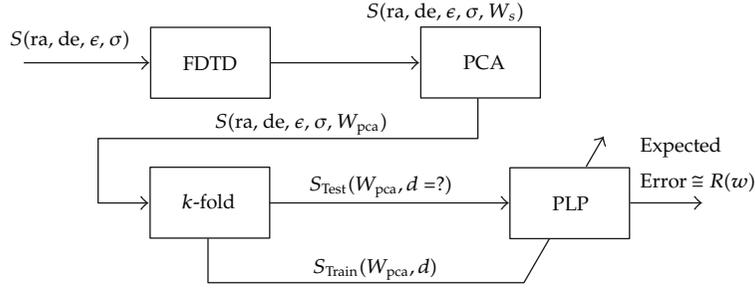


Figure 3: This figure shows an overview of the detection system employed using the ANN. Firstly, given a set S of radii (ra), depths (de), ϵ_r , and σ a scattered wave ($W_s \in \mathbb{R}^p$) is calculated using FDTD. Afterwards, a dimensional reduction is applied. In the next step, the expected error is estimated to adapt the Parallel Layer Perceptron network parameters. In this work a 10-fold cross-validation was employed.

Another interpretation of the PCA is the construction of directions that maximize the variance. The transformation V_q generates a projection space in which the covariance matrix is diagonal. The diagonal covariance matrix implies that the variance of a variable with itself is maximized while it is minimized with respect to any other variable. Thus, the q variables with higher variance in the new space should be kept. The principal components of a set of data in \mathbb{R}^p provide a sequence of best linear approximations to that data, of all ranks $q \leq p$.

The problem of dimension reduction and ANN training has its scheme shown in Figure 3. Considering the input space initially in \mathbb{R}^{1200} , it can be projected in a \mathbb{R}^{286} without any loss of information, that is, 100% of the data variance was kept. Considering 99.99% of the variance, the variables can be projected in \mathbb{R}^{139} and in \mathbb{R}^{51} when 99% of the original variance is kept. These are remarkable reductions that help in reducing the curse of dimensionality.

To evaluate the performance of the studied techniques the following error (loss) figure is used:

$$L(d_r) = \frac{|d_t - d_r|}{d_r}, \quad (3.2)$$

where d is the unknown variable (depth or radius), the subscript t indicates the real value of the variable, and the subscript r indicates the value reconstructed by the neural network. This measures the percentage deviation of the reconstructed object from the real one (desired object).

In Table 4, it is presented the results considering the Relative Error in the prediction of the depth given 286 dimensions, varying the number of neurons. It is clear that the error does not vary much as the number of neurons changes. The optimum is achieved for 9 neurons. The network starts overfitting afterwards. The number of neurons was determined using cross-validation. However, the results given different number of neurons are also acceptable results, which can be used in a real-world problem.

The ANNs have the advantages of having a numerous class of models, reaching high levels of accuracy due to their capacity of generalization, having universal approximation models for regression and classification models, can be fast computed due to their natural parallel structure (real time applications), being tolerant to structure failures (knowledge distributed in structure), having well known strategies to control their complexity,

having satisfactory models easy to understand and implement even for non-experts, may be retrained or modified once having new samples.

On the other hand, the drawbacks of ANNs are the necessity of having many representative pre-processed (normalized) samples to guarantee satisfactory understanding of domains of problems being modeled (the accuracy is conditioned to samples representing the problem), cannot always be easily understood or translated to a set of symbolic rules or equations due to their structure, cannot easily deal with many different problems, inputs, and outputs with different codification, noise levels, and approximation and from different areas, the learning process may need great computational resources and time.

3.3. Feature Selection

Feature selection (also known as subset selection) is a process commonly used in machine learning, wherein a subset of the features available from the data is selected for application of a learning algorithm. The best subset contains the least number of dimensions that most contribute to accuracy, discarding the remaining, unimportant dimensions. This is an important stage of pre-processing and is one of two ways of avoiding the curse of dimensionality (the other is feature extraction). Feature selection is, therefore, the task of choosing a small subset of features which is sufficiently to predict the target labels. This step is fundamental to build reliable classifiers.

Another phenomenon that takes place in high dimensional problems is the rise of the computational effort. Suppose that one wants to approximate the scattered wave for the 1640 samples. The numerous methods for modeling the nonlinear behavior are classically divided into three categories: physical, empirical, and table-based ones. Some may be difficult to categorize in this way, however, and therefore models are divided here simply into physical and empirical ones, and into black-box and circuit-level ones.

Feature selection aims at choosing for a given data set a subset which can capture the relevant information. The choice of features is important to avoid the curse of dimensionality and, therefore, guarantee a good convergence of the learning machines. It can also provide some understanding concerning the nature of the problem, as it indicates the main physical properties to classify an underground target.

In literature three types of Feature Selection algorithms are distinguished, the so-called wrapper, filter, and embedded methods [21–23]. Wrapper methods estimate the usefulness of a subset of features by a given predictor or learning machine. The Wrapper methods try to directly optimize the performance of a given predictor, this is done estimating the generalization performance. Even though it sounds a natural approach it can be unfeasible given its computational burden, as many classifier must be cross-validate.

Filter or Variable Ranking methods compute relevance scores for each single feature and choose the most relevant ones according to those scores. This can be done using usually ad hoc, evaluation function aiming at searching the set of features that maximize the information. Some of the commonly used evaluation functions are: the mutual information, the margin, dependence measures, among others.

The main drawback of such simple filter methods is that they are not able to detect inter-feature-dependencies; one example of inter-feature-dependency in our GPR problem is relation between the clutter and reflected echoes from near targets that can occur at the same time. Neither the first nor the second dimension alone helps to determine from which class an example is stemming, only both dimensions together contain enough information about the class membership.

Another drawback of those methods is its computational deficiency. But there are methods, which may be categorized as variable ranking methods and are also able to reveal such feature dependencies; for example, Relief [21] is one of those. This paper will consider feature selection algorithms based on the filter model. In this case the feature selection is a type of pre-processing, using some predefined cost function, as the classes separation margin.

The algorithms considered here are Relief and Simba as in [24]. The Relief algorithm is based on the statistical relevance of the features while Simba is based on the concept of the separation hypothesis-margin between two classes.

From the reflected wave, 15 features were computed to be selected.

- (i) Delay of the first reflection ($Delay$) [25],
- (ii) maximum amplitude ($\max(W_s)$) [25],
- (iii) reflected wave mean ($\text{mean}(W_s)$),
- (iv) reflected wave standard deviation ($\text{std}(W_s)$),
- (v) mean of the wave derivative ($\text{mean}(dW_s)$),
- (vi) standard deviation of the wave derivative ($\text{std}(dW_s)$),
- (vii) maximum amplitude of the Fourier transform (mFFT),
- (viii) energy in six different bands (B_1, B_2, \dots, B_6) of the Fourier transform,
- (ix) frequency of maximum amplitude of the Fourier transform (fmFFT),
- (x) signal energy ($\text{sen}(W_s)$).

The six most selected features in [26] were: $\text{mean}(W_s)$, mFFT, fmFFT, $\text{sen}(W_s)$, $\text{std}(dW_s)$, and $Delay$.

As already explained, the feature reduction avoids large structures to data set modeling in both regression and classification problems. In addition, a smaller set of features may expose models to less noisy data, and hence, it may improve classification accuracy. Other advantage is the reduction of computational resources needed to store data and calculated metrics from samples. And models with fewer features tend to facilitate obtaining more explainable models because the resultant features can be justified by known information which relates samples and their labels.

To do so, the feature selection methods must have the ability to deal with large data sets with many samples and features. Depending on the selected strategy to feature extraction, which replaces the original data by some of their metrics for evaluation, the new representation of each sample should not vary so much due to noise. But if it is not the case, the feature selection process may be affected by metrics intolerance to noise. The final subset of features may greatly vary besides requiring a consensus response and a more complex data pre-processing step. So, the feature selection methods must be applied more than once. Another problem of filter methods for feature selection is the requirement of having all classes well represented in training set. As the methods addressed in this work are based on hypothesis-margin, the samples near the margin should be presented in the set considered to rank features. The metric and strategy to compare the combinations of features have considerable impact on the quality of results and the number of evaluations. As a consequence, different filter methods can reach distinct results requiring almost similar time for evaluation. Then, by the problems exposed above, the final selected features are not guaranteed to be the best subset.

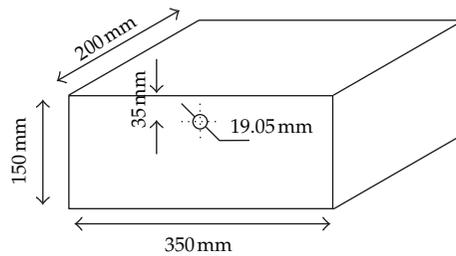


Figure 4: Concrete slab dimensions.

4. Real Data Processing

The concrete has been the most commonly used building material because it is durable, cheap, and readily available and can be cast into almost any shape. The main components of concrete are the hydrated cement paste, aggregates, water, and the transition zone.

The main constituent of concrete is cement, which is made up of calcium, silica, alumina, and iron oxide. The aggregates occupy 60 to 80% of the concrete's volume. They are usually stronger than cement and therefore play a role as a filler [27].

Concrete has many voids that are usually filled with water. The water can be classified depending on the type of void and degree of firmness with which the water is held in these voids. The water, which is responsible for the electrical conductivity of concrete, is called free water and is present in voids greater than 50 nm in size.

The transition zone, the region that exists between the hydrated cement paste and the aggregate, is very thin, with a thickness on the order of 10 to 50 μm . The transition zone is important in that it is the weakest zone in the concrete and thus influences its stiffness and durability.

The concrete used in the experiments is close to the one found at the Federal University of Minas Gerais. In addition, the dimensions of the four slabs were specified to meet the characteristics found in [28, 29]. In this way the mixture has the following characteristics: (1) cement Portland type I; (2) water/cement ratio 0.60; (3) cement/sand ratio 1:2.25. The inclusions have depth 35 mm from the surface of the concrete slab and have diameters of 19.05 mm as illustrated in Figure 4.

In order to get experimental results as close as possible to the numerical simulations it was prepared a GPR survey in four concrete slabs with the following characteristics: (1) without inclusion; (2) with metal inclusion; (3) with PVC inclusion; (4) water filled PVC inclusion. The survey was performed in a semi-anechoic chamber to avoid noise in the experiment according to the test setup illustrated in Figure 5.

The GPR equipment used in this work has its minimum configuration consisting of the following device grouped and used together:

- (a) 1.6 GHz or 2.3 GHz shielded antennas;
- (b) ProEx Control unity;
- (c) XV-11 Monitor with 1.2 m cable;
- (d) Cable X3M 4 m;
- (e) Battery Li-Ion 11.1 V/6.6 Ah.

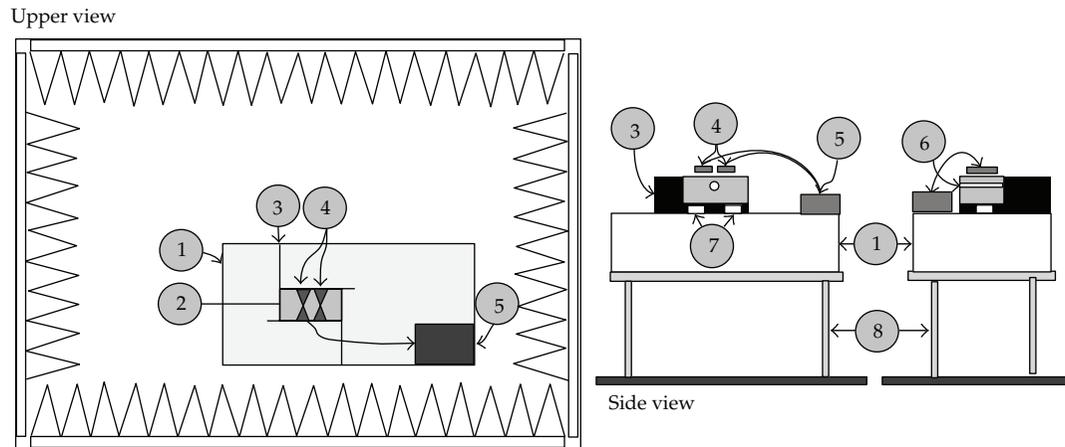


Figure 5: Schematic and photo of the assembly held in semi-anechoic chamber. (1) Polystyrene layer, (2) concrete block, (3) absorber plate, (4) bowtie antennas, (5) GPR control unit, (6) inclusion, (7) polystyrene blocks, and (8) wood table.

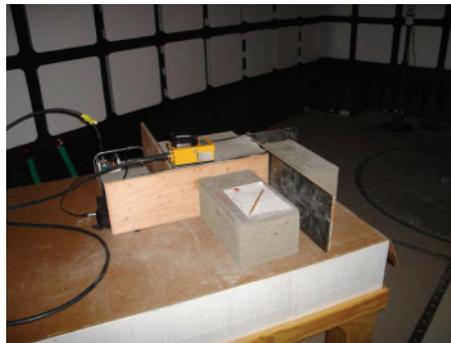


Figure 6: Experiment performed in a semi-anechoic chamber.

Prior to the GPR assessment of concrete in the chamber, test setup ambient levels (i.e., all equipment energized except the GPR equipment) were performed to verify 6 dB or more below the level required. The ambient measurements were performed using vertical polarization of a log periodic antenna.

The GPR equipment and its cable harness along with the concrete slabs lie on an insulated support 1000 mm (± 50 mm) above the floor of the test chamber. The dielectric constant of the insulated support was less than 1.4. The ground plane was not used for this experiment. The battery was located under the test bench. The equipment was at least 1000 mm from the chamber walls. No part of the GPR equipment was closer than 250 mm to the floor.

The GPR operated under typical loading and other conditions as in a regular survey to detect cracks and inclusions. The cables provided by the GPR equipment to perform the tests were shielded. Some cables that had excess lengths and were not bulk were bundled at the approximate center of the cable with the bundles 300 mm to 400 mm in length.

Figure 6 illustrates the experiment in the semi-anechoic chamber.

The climatic test conditions are defined in Table 5.

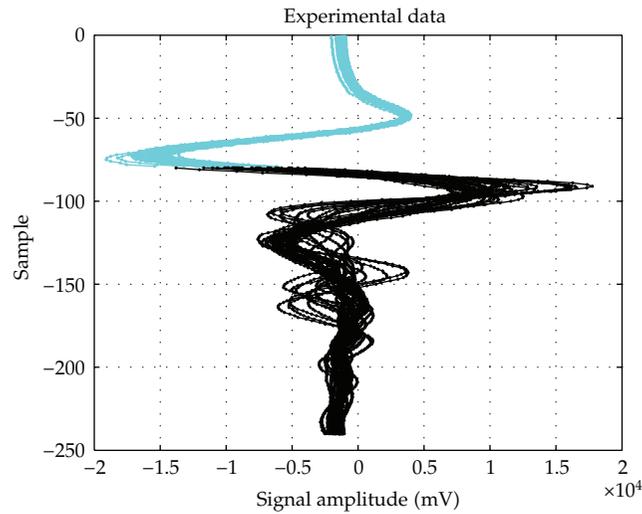


Figure 7: Experimental data of 1.6 GHz antenna: discarded samples in cyan and considered samples in black.

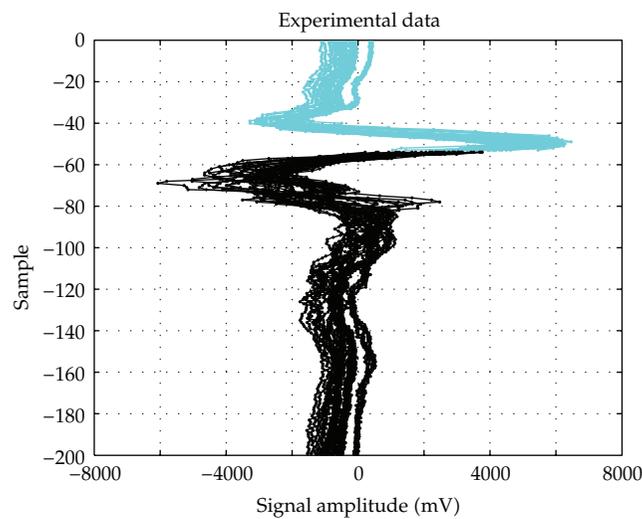


Figure 8: Experimental data of 2.3 GHz antenna: discarded samples in cyan and considered samples in black.

Based on simulated results and for higher number of samples, each concrete block was probed in its up and down sides. Therefore, 324 samples were collected with each antenna, 108 samples for each block inclusion (air, conductor, or PVC + water). As 12 samples of each block was taken over the inclusion, only these samples were considered to depth regression with half of them in top side and the other half in the underside of the blocks. The radius estimation was not experimented with real data set since all the inclusions in the concrete blocks had the same diameter (The experimental data set used can be downloaded in <http://www.enacom.com.br/doc/gprenacomsenai.zip>).

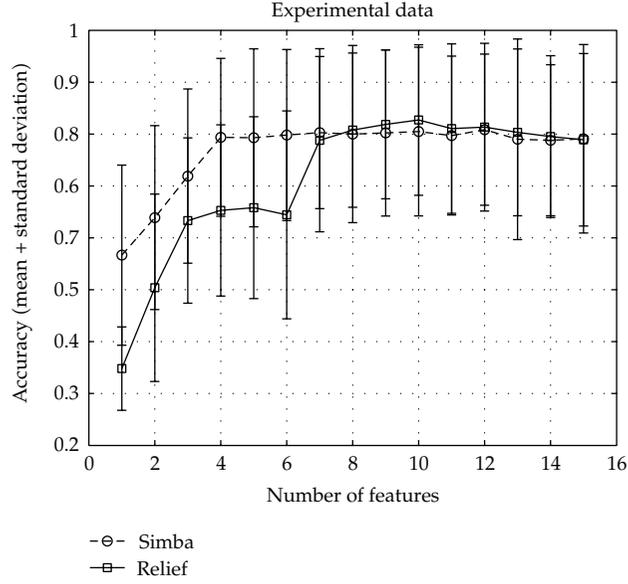


Figure 9: Accuracy results for k -NN with features extracted from experimental data of 1.6 GHz antenna to classify types of inclusions.

Each A-scan from both antennas had first echo excluded from data for models experiments. For 1.6 GHz antenna, just samples from 80th to 240th position were considered, as showed in Figure 7. And for 2.3 GHz, samples were selected from 54th to 200th position, as registered in Figure 8.

Before submitting these samples to classification and regression models, each data set was submitted to feature extraction as in [26] and after to Simba and Relief to rank features by their capacity of classes discrimination. The ranking results (from most to least important) for each antenna are as follows.

(i) 1.6 GHz antenna

(a) Simba: $\text{mean}(dW_s)$, fmFFT , $\text{std}(dW_s)$, mFFT , B_3 , SignalEnergy , B_2 , B_1 , $\text{std}(W_s)$, B_4 , $\text{max}(W_s)$, B_6 , B_5 , $\text{mean}(W_s)$, Delay .

(b) Relief: fmFFT , mFFT , $\text{std}(dW_s)$, SignalEnergy , $\text{std}(W_s)$, B_1 , $\text{mean}(dW_s)$, B_3 , B_6 , B_4 , B_2 , $\text{max}(W_s)$, Delay , B_5 , $\text{mean}(W_s)$.

(ii) 2.3 GHz antenna

(a) Simba: fmFFT , $\text{max}(W_s)$, $\text{mean}(dW_s)$, $\text{std}(dW_s)$, mFFT , B_6 , SignalEnergy , B_2 , $\text{std}(W_s)$, B_3 , B_5 , B_1 , B_4 , $\text{mean}(W_s)$, Delay .

(b) Relief: fmFFT , $\text{max}(W_s)$, $\text{std}(dW_s)$, B_3 , B_5 , B_2 , mFFT , SignalEnergy , $\text{mean}(dW_s)$, $\text{std}(W_s)$, Delay , B_6 , B_1 , $\text{mean}(W_s)$, B_4 .

Each of these features ranking was important to generate results to be compared to simulated data sets. As in the latter, no pre-processing of data was done before sample submission to models.

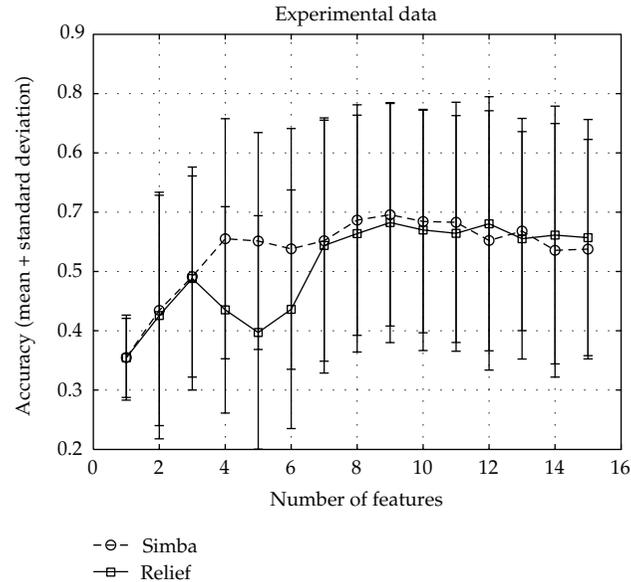


Figure 10: Accuracy results for k -NN with features extracted from experimental data of 2.3 GHz antenna to classify types of inclusions.

The $k = 3$ -NN for inclusion classification model used Squared Euclidean distance. And the accuracy was computed adopting $n = 50$ -fold cross-validation method adopting *cvpartition* of MATLAB with its default values. The correspondent results are registered in Figure 9 for 1.6 GHz antenna and in Figure 10 for 2.3 GHz antenna.

Higher accuracies with fewer features were attained using the sequences obtained by Simba. But in both data sets the accuracies were smaller than expected. The results seen to be affected by real data problems like noise, clutter, and equipment instabilities as signal shifting. Then, more care should be taken in pre-processing steps addition to improve the quality of signals submitted to modeling step.

The PLP model was used with Log-sigmoid activation function as the regression model with two neurons used in internal layer. The percent error was computed using $n = 25$ -fold cross-validation method taking for each model with Hybrid Levenberg Marquardt allowing 10 epoch and 0.001 as maximum mean squared error (MSE). The expected results were rescaled to the allowed output range (between 0 and 1). The results for both antennas are shown, respectively, in Figures 11 and 12.

The features values extracted from both data sets permitted stable results despite the apparent higher standard deviation in the second figure because of the magnification done to visualization improvement. The features addition to this model indicated the existence of a small group of 6 features which should be considered to achieve smaller percent errors.

On the other hand, in both cases, the *Delay* feature, which obtained constant values for all samples, had significant impact in the chosen model, causing bad conditioned matrix of weights. The first models without considering this feature had significant percent error for depth estimation. But the subsequent models with *Delay* and other new features could not be trained satisfactory. So, these features should be disregarded until a new function for first delay arrival can effectively extract the corresponding information for samples discrimination.

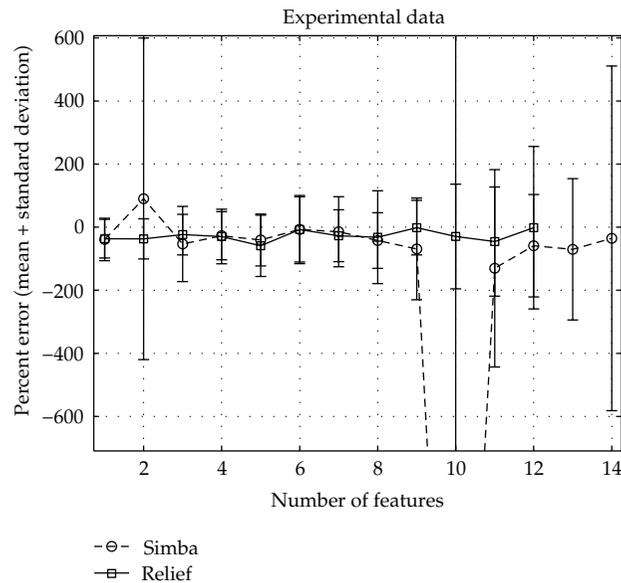


Figure 11: Percent Error main results for PLP with features extracted from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

To complete the set of earlier results to be compared with the ones with simulated data sets, the same experimental data set and its extracted features were submitted to PCA and PLP-MGM of [30–32] for regression modeling. Using $n = 10$ -fold cross-validation method, PLP-MGM model with 3 internal neurons and saddle activation function was trained using Hybrid Levenberg Marquardt up to 10 epochs or until maximum MSE reaches below 0.001 level. For samples from A-scans, only components which contribute with more than 0.005% were considered in absolute percent error evaluation in models incrementing the number of components.

Before analyzing the results, it should be emphasized that the number of available samples for training and testing are much smaller than the ones simulated. Therefore, higher standard deviations are expected in the results considering the same number of folds in cross-validation method.

For 1.6 GHz antenna, 28 components were selected whereas for 2.3 GHz antenna, 35 components passed the permitted threshold. The respective results are exposed in Figures 13 and 14. The Pareto plots in each of these figures showed fewer significant components to explain the variance of the whole set. For both sets, the absolute percent mean error indicates small groups of no more than 10 principal components to achieve better estimation models for depth.

Using the extracted features, which were only 15, all of them were submitted to PCA and submitted to PLP-MGM models taking the same allowed parameters. These final results, which are in Figures 15 and 16, showed better estimation considering the same number of features was almost the same used for principal components, respectively. Based on these graphics, there really should exist a small set of features similar to the ones extracted which may obtain less complex models with satisfactory estimation even if nonrobust characteristics were considered.

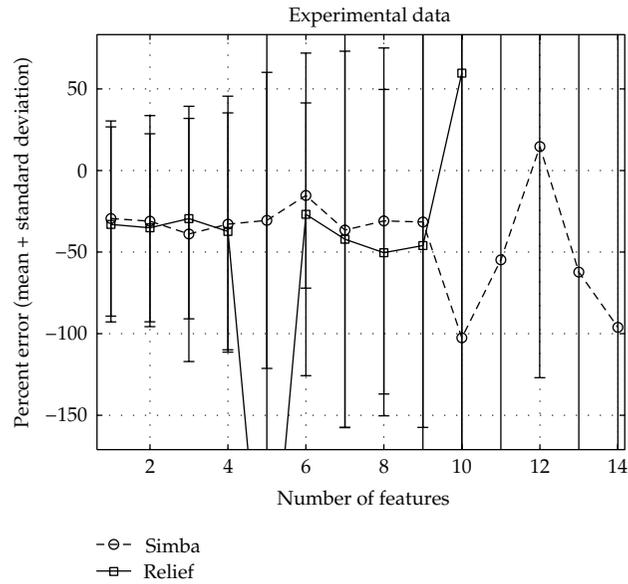
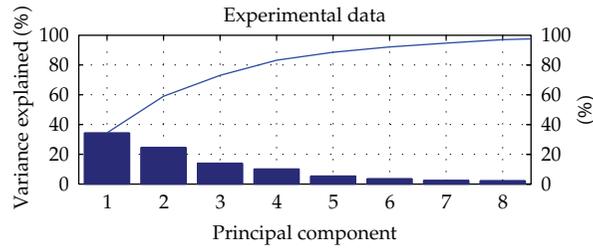
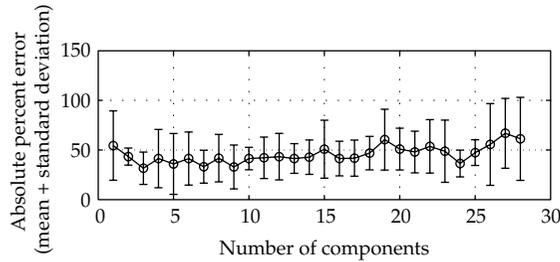


Figure 12: Percent Error main results for PLP with features extracted from experimental data of 2.3GHz antenna to approximate depth of inclusions.



(a)



(b)

Figure 13: Absolute Percent Error main results for PLP-MGM with PCA components of samples from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

Table 3: Results of the relative error considering the depth prediction.

Configuration	Max(L)	Train(s)	Test(s)
PLP (286)	0.04%	8.44	0.016
PLP (139)	0.15%	4.48	0.01
PLP (51)	20%	6.2	0.011

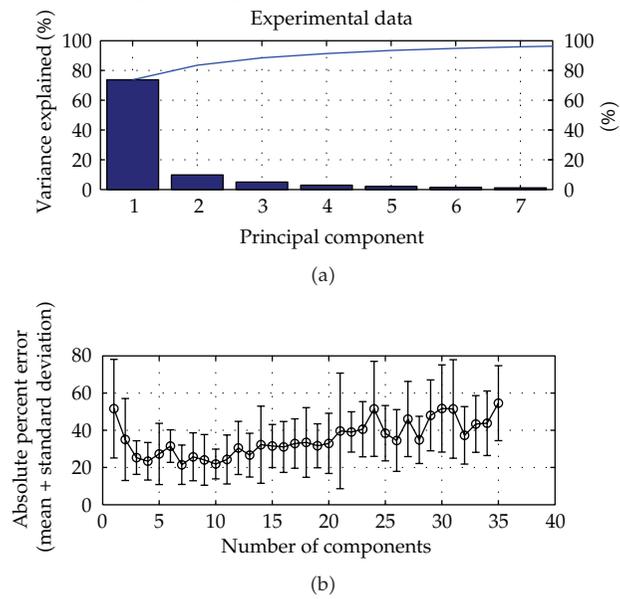


Figure 14: Absolute Error main results for PLP-MGM with PCA components of samples from experimental data of 2.3 GHz antenna to approximate depth of inclusions.

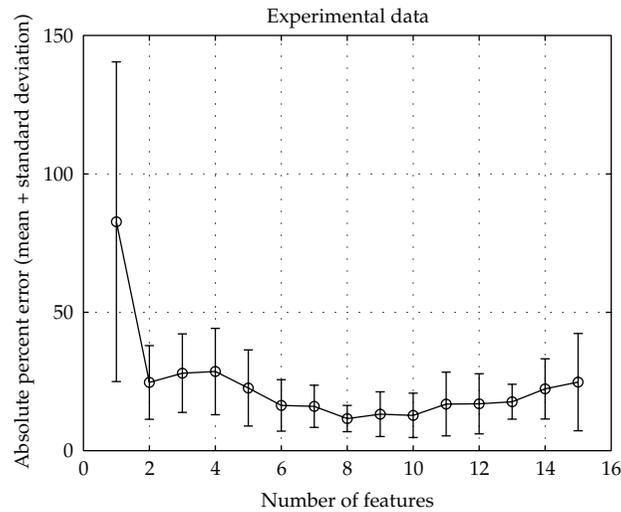


Figure 15: Absolute Percent Error main results for PLP-MGM with PCA components of features extracted from experimental data of 1.6 GHz antenna to approximate depth of inclusions.

However, models with extracted and selected features are preferable. Besides their better results, these latter models input components are more reliable since they consist of functions which behavior can be explained by the GPR operation and the related physics.

5. Conclusion

A number of important issues have been discussed in this paper. First the definition of estimation schemes for microwave imaging was outlined. The concept of inverse problem

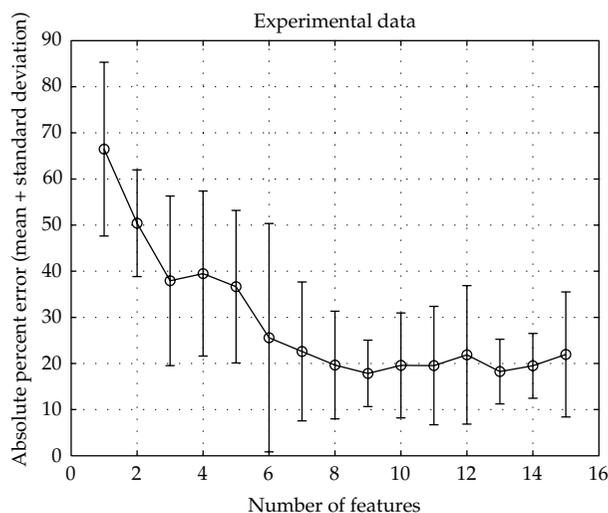


Figure 16: Absolute Error main results for PLP-MGM with PCA components of features extracted from experimental data of 2.3GHz antenna to approximate depth of inclusions.

Table 4: Results of the Relative Error considering the depth prediction given 286 dimensions, varying the number of neurons.

Neurons	Mean(L)	Max(L)
6	0.009%	0.27%
7	0.008%	0.16%
8	0.003%	0.05%
9	0.001%	0.04%
10	0.004%	0.09%
11	0.004%	0.09%
12	0.005%	0.10%
13	0.006%	0.15%
14	0.005%	0.16%
15	0.008%	0.10%

for the GPR assessment of concrete was defined with respect to the inclusion reconstruction. Three 2D reconstruction algorithms were implemented and new configurations were proposed using Feature Selection. Numerical simulations with different types of inclusions were presented to assess the accuracy and efficiency of these algorithms for image reconstruction in real-world settings. Experiments were performed in a controlled environment to validate the results. Both data sets were submitted to PCA, feature extraction, and feature selection methods not only to improve the results, but also to obtain efficient explainable models.

As could be seen, a few extracted features are not robust to noise contamination and need some pre-processing to be reasonably evaluated. Some models were accordingly developed to reduce signal contamination, reduce the data formatting steps, or in some way overcome the large number of features and samples from A-scans. But the feature extraction process still has to be improved to avoid constant values, to gain a small general set of characteristics, and to reduce the pre-processing steps needed. In other words, it has to be developed a set of features to efficiently capture the information in which is expected to find

Table 5: Environmental test conditions.

Temperature	23 ± 5 degrees C
Humidity	20 to 80% relative humidity

the elements capable of discriminate materials and explaining the phenomena involved in this inverse problem.

For most of the cases, the targets were satisfactorily reconstructed with the 2D algorithms and the convergence was assessed in terms of the relative error (see Tables 3 and 4). In conclusion, significant algorithmic flexibilities utilizing 2D based algorithm were demonstrated in terms of accommodating various forms of inclusions in single and iterative reconstruction. More informative and less complex models were obtained by the use of specific models to each inverse problem and trying to include some A-scans features that visually distinguish the materials and indicate inclusions.

Acknowledgment

This work was supported by Coordination for the Improvement of Higher Level—or Education—Personnel (CAPES), National Council for Scientific and Technological Development (CNPq), and Foundation for Research Support in Minas Gerais State (FAPEMIG), Brazil.

References

- [1] D. J. Daniels, *Ground Penetrating Radar*, The institute of Electrical Engineers, London, UK, 2004.
- [2] H. M. Jol, *Ground Penetrating Radar: Theory and Applications*, Elsevier Science, Oxford, UK, 2009.
- [3] L. B. Conyers, "Ground-penetrating radar techniques to discover and map historic graves," *Historical Archaeology*, vol. 40, no. 3, pp. 64–73, 2006.
- [4] M. M. Poulton, *Computational Neural Net-Works for Geophysical Data Processing*, Pergamon, Mysia, Turkey, 2001.
- [5] A. G. Davis, M. K. Lim, and C. G. Petersen, "Rapid and economical evaluation of concrete tunnel linings with impulse response and impulse radar non-destructive methods," *NDT and E International*, vol. 38, no. 3, pp. 181–186, 2005.
- [6] C. Maierhofer, H. W. Reinhardt, and G. Dobmann, Eds., *Non-Destructive Evaluation of Reinforced Concrete Structures: Non Destructive Testing Methods*, vol. 2, CRC Press, Boca Raton, Fla, USA, 2010.
- [7] K. J. Langenberg, K. Mayer, and R. Marklein, "Nondestructive testing of concrete with electromagnetic and elastic waves: modeling and imaging," *Cement and Concrete Composites*, vol. 28, no. 4, pp. 370–383, 2006.
- [8] S. R. Doctor, T. E. Hall, and L. D. Reid, "SAFT—the evolution of a signal processing technology for ultrasonic testing," *NDT International*, vol. 19, no. 3, pp. 163–167, 1986.
- [9] X. L. Travassos, D. A. G. Vieira, N. Ida, C. Vollaire, and A. Nicolas, "Inverse algorithms for the GPR assessment of concrete structures," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 994–997, 2008.
- [10] X. L. Travassos, D. A. G. Vieira, N. Ida, and A. Nicolas, "In the use of parametric and nonparametric algorithms for the nondestructive evaluation of concrete structures," *Research in Nondestructive Evaluation*, vol. 20, no. 2, pp. 71–93, 2009.
- [11] J. H. Bungey, "Sub-surface radar testing of concrete: a review," *Construction and Building Materials*, vol. 18, no. 1, pp. 1–8, 2004.
- [12] X.-Q. He, Z. Q. Zhu, Q. Y. Liu, and G. Y. Lu, "Review of GPR rebar detection," in *Proceedings of the Progress in Electromagnetics Research Symposium (PIERS '09)*, pp. 804–813, Beijing, China, March 2009.
- [13] T. C. K. Molyneaux, S. G. Millard, J. H. Bungey, and J. Q. Zhou, "Radar assessment of structural concrete using neural networks," *NDT and E International*, vol. 28, no. 5, pp. 281–288, 1995.

- [14] M. R. Shaw, S. G. Millard, T. C. K. Molyneaux, M. J. Taylor, and J. H. Bungey, "Location of steel reinforcement in concrete using ground penetrating radar and neural networks," *NDT and E International*, vol. 38, no. 3, pp. 203–212, 2005.
- [15] D. Ayala-Cabrera, M. Herrera, J. Izquierdo, and R. Pérez-García, "Location of buried plastic pipes using multi-agent support based on GPR images," *Journal of Applied Geophysics*, vol. 75, no. 4, pp. 679–686, 2011.
- [16] M. Heiler, S. McNeil, and J. Garrett, "Ground-penetrating radar for highway and bridge deck condition assessment and inventory," in *Nondestructive Evaluation of Aging Bridges and Highways*, pp. 195–206, June 1995.
- [17] I. Abdel-Qader, F. Abu-Amara, and O. Abu-dayyeh, "Fractals and independent component analysis for defect detection in bridge decks," *Advances in Civil Engineering*, vol. 2011, Article ID 506464, 14 pages, 2011.
- [18] J. Wang, X. Xie, and H. Huang, "A fuzzy comprehensive evaluation system of mountain tunnel lining based on the fast nondestructive inspection," in *Proceedings of the International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE '11)*, pp. 2832–2834, 2011.
- [19] Z. M. Sbartai, S. Laurens, K. Viriyametantont, J. P. Balayssac, and G. Arliguie, "Non-destructive evaluation of concrete physical condition using radar and artificial neural networks," *Construction and Building Materials*, vol. 23, no. 2, pp. 837–845, 2009.
- [20] S. Bazarara, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley-Interscience, New York, NY, USA, 2006.
- [21] K. Kira and L. Rendell, "A practical approach to feature selection," in *Proceedings of the 9th International Workshop on Machine Learning*, pp. 249–256, 1992.
- [22] S. R. H. Hoole, "Artificial neural networks in the solution of inverse electromagnetic field problems," *IEEE Transactions on Magnetics*, vol. 29, no. 2, pp. 1931–1934, 1993.
- [23] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, *Feature Extraction-Foundations and Applications, Studies in Fuzziness and Soft Computing*, vol. 207, Springer, New York, NY, USA, 2006.
- [24] R. Gilad-Bachrach, A. Navot, and N. Tishby, "Margin based feature selection-theory and algorithms," in *Proceedings of the 21st International Conference on Machine Learning (ICML '04)*, pp. 337–344, July 2004.
- [25] S. Caorsi and G. Cevini, "An electromagnetic approach based on neural networks for the GPR investigation of buried cylinders," *IEEE Geoscience and Remote Sensing Letters*, vol. 2, no. 1, pp. 3–7, 2005.
- [26] F. A. A. Queiroz, D. A. G. Vieira, and X. L. Travassos, "Analyzing the relevant features of GPR scatteredwaves in time and frequency-domain," *Research in Nondestructive Evaluation*. In press.
- [27] P. K. Mehta, *Concrete: Structure, Properties and Materials*, Prentice Hall, Englewood Cliffs, NJ, USA, 1986.
- [28] O. Buyukozturk, T.-Y. Yu, and J. A. Ortega, "A methodology for determining complex permittivity of construction materials based on transmission-only coherent, wide-bandwidth free-space measurements," *Cement and Concrete Composites*, vol. 28, no. 4, pp. 349–359, 2006.
- [29] I. L. Al-Qadi, O. A. Hazim, and S. M. W. Suriad S.M., "Dielectric properties of Portland cement concrete at low radio frequencies," *Journal of Materials in Civil Engineering*, vol. 7, no. 3, pp. 192–198, 1995.
- [30] D. A. G. Vieira, L. Travassos, R. R. Saldanha, and V. Palade, "Signal denoising in engineering problems through the minimum gradient method," *Neurocomputing*, vol. 72, no. 10–12, pp. 2270–2275, 2009.
- [31] X. L. Travassos, D. A. G. Vieira, N. Ida, C. Vollaire, and A. Nicolas, "Characterization of inclusions in a nonhomogeneous GPR problem by artificial neural networks," *IEEE Transactions on Magnetics*, vol. 44, no. 6, pp. 1630–1633, 2008.
- [32] X. L. Travassos, D. A. G. Vieira, V. Palade, and A. Nicolas, "Noise reduction in a non-homogenous ground penetrating radar problem by multiobjective neural networks," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1454–1457, 2009.

Research Article

Decomposition of the Time Reversal Operator for Target Detection

**Chun-xiao Li,^{1,2} Huan-cai Lu,^{1,2} Ming-fei Guo,¹
Hang-fang Zhao,³ and Jiang-ming Jin^{1,2}**

¹ MOE Key Laboratory of Mechanical Manufacture and Automation, Zhejiang University of Technology, Hangzhou 310014, China

² Zhejiang Key Laboratory of Signal Processing, Zhejiang University of Technology, Hangzhou 310014, China

³ Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China

Correspondence should be addressed to Chun-xiao Li, chunxiaoli@zju.edu.cn

Received 8 September 2012; Revised 6 November 2012; Accepted 12 November 2012

Academic Editor: Fatih Yaman

Copyright © 2012 Chun-xiao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A thorough theory of detection problem using active time reversal has been investigated in several recent papers. Although active time reversal method is theoretically superior to the others, its practical implementation for target detection is far more difficult. This paper investigates the detection problem using passive decomposition of the time reversal operator (DORT) method. Provided that the signal components can be modeled as a linear combination of basis vectors with an unknown signal subspace, the generalized likelihood ratio test (GLRT) is derived based on Neyman-Person lemma with the unknown signal subspace replaced by its maximum likelihood estimation. The test statistics is one of the dominant eigenvalues of the time reversal operator for a point-like scatterer. Finally, the performance of the DORT detector is investigated with acoustic data collected from a waveguide tank. The experimental results show that the DORT detector can provide, respectively, 1.4 dB, 1.1 dB, and 0.8 dB performance gains over the energy detector given false alarms rate of 0.0001, 0.001, and 0.01.

1. Introduction

It is acoustically difficult to detect a target in shallow water environment, for instance, the performances of the traditional detectors such as the matched filter are severely degraded because the echo of active sonar undergoes distortion during two-way propagation and reflection from the target. One of the approaches to improve the detection performance is the model-based matched filter that takes advantage of the physical model to enhance signal

processing [1]. Another compromising solution is the one to model the channel distortion, such as the segmented replica correlation detector and the replica correlation integration detector [2, 3].

Time-reversal approaches provide an alternative approach to this kind of problem [4–7]. The unique feature of time reversal is that it provides a robust set of waves without any signal analysis, which compensates for the distortion induced by spatial inhomogeneities in the propagation medium [4]. Two time-reversal approaches, the iterative time-reversal, and the decomposition of the time reversal operator (DORT) can be employed for detection and localization of target [6–9].

Time reversal has been intensively studied in ultrasound acoustics and electromagnetics. There have been several recent studies devoted to the detection problem from the statistic signal processing perspective [10–15]. Reference [12] investigates the detection problem using the active time reversal of DORT method and the iterative time reversal method in the context of acoustic wave processing in ocean, while [10, 11] studies the active iterative time reversal method in the context of electromagnetic wave processing in air.

Although active time reversal method is theoretically superior to the others, its practical implementation for target detection is far more difficult. As an alternative, this paper investigates the detection problem using passive DORT method in acoustic wave processing in ocean. Compared to electromagnetic wave propagation in air, sound propagation in ocean is more environmentally involved. The terms “active” and “passive” refer to the source-receiver arrays (SRA): “active” array works when a SRA attempts to focus on a target by retransmitting an environmentally dependent excitation vector, while “passive” array functions when a SRA simply transmits signals from individual source elements followed by processing of the returned time series [16]. Moreover, the performances of the derived detectors in [12] are only evaluated by numerical simulations. This paper tests the detector with real acoustic data from waveguide experiments.

The rest of the paper is organized as follows. The mathematical model of binary hypothesis is derived, and the generalized likelihood ratio test (GLRT) of the DORT approach is developed in Section 2. The proposed detectors are validated by laboratory waveguide experiments in Section 3. The conclusions and summary are drawn in Section 4.

2. Theory and Method

2.1. Notations

We denote scalars by lowercase letters, vector quantities by boldface lowercase letters, and matrices by uppercase boldface. $(\cdot)^H$ stands for conjugate transpose; $R(A)$ and $I(A)$ are the real and the imaginary parts of matrix A , respectively; $\text{tr}(\mathbf{A})$ is the sum of the diagonal elements of \mathbf{A} ; \mathbf{I}_N is the identity matrix of size N .

2.2. DORT Measurement

According to the theory of DORT, a SRA of N sensors collects data and then generates a response matrix with $N \times N$ dimension. The matrix is formed by sequentially transmitting signals from individual source elements of the SRA or orthogonal beams and recording the backscattered echo on each receiver element [8, 9, 17].

Provided that a normalized narrow-band signal is transmitted (the angular frequency ω is thus omitted in the formulas below), the received signals in time are first transformed into frequency domain. They can be modeled as a sum of interferences and possible signals whose presence we are trying to detect as follows:

$$\mathbf{y}_j = \mathbf{s}_j + \mathbf{v}_j, \quad j = 1, 2, \dots, N, \quad (2.1)$$

where \mathbf{s}_j is the signal term and \mathbf{v}_j is the interference-plus-noise term for j th transmission.

The signal vector \mathbf{s}_j is assumed to be deterministic but unknown. We model them as $\mathbf{s}_j = \mathbf{H}\boldsymbol{\theta}_j$, which are a linear combination of basis vectors with an unknown signal subspace \mathbf{H} with known rank- p (corresponding to p propagation modes of a scatterer). Considering that the target is not definitely a point-like scatterer for the application of DORT method, the signal subspace has dimension p . In practical implementation, \mathbf{H} is a $N \times p$ matrix, and each column \mathbf{h}_i is the Green's function of the channel between the SRA and one aspect of the scatterer.

Although the interference \mathbf{v}_j is the sum of the receiver noise and the reverberation signal, the DORT method has the ability to separate the echo of a target from reverberation [18]. We assume that both the receiver noise and the residual reverberation signal can be modeled as zero-mean white Gaussian random processes [2], that is,

$$\mathbf{v}_j \sim CN(0, \sigma_v^2 \mathbf{I}_N). \quad (2.2)$$

The real and imaginary components of \mathbf{v}_j are $R(\mathbf{v}_j) \sim N(0, (\sigma_v^2/2)\mathbf{I}_N)$ and $I(\mathbf{v}_j) \sim N(0, (\sigma_v^2/2)\mathbf{I}_N)$, respectively. Although the assumption of Gaussian reverberation may not always be true, it is often made to facilitate mathematical analysis.

After all the transmission are finished, the total received data can be arranged as a matrix by [19]

$$\mathbf{Y} = \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_j & \cdots & \mathbf{y}_N \end{bmatrix}. \quad (2.3)$$

Similarly, the signal matrix can be written as $\mathbf{S} = [\mathbf{s}_1 \cdots \mathbf{s}_j \cdots \mathbf{s}_N]$.

The detection problem then becomes to choose one of the following two hypotheses:

$$H_0 : \mathbf{Y} = \mathbf{V} \quad \text{versus} \quad H_1 : \mathbf{Y} = \mathbf{S} + \mathbf{V}. \quad (2.4)$$

2.3. The DORT Detector

As shown in [10, 12], the suboptimal detector using GLRT in the Neyman-Pearson sense is derived as an energy detector in conventional approaches when the Green's function of unknown channel is replaced by its maximum likelihood estimation. In this paper a suboptimal detector using the GLRT approach is derived for the time reversal approach, which is named the DORT detector.

The probability density functions (PDF) $p(\mathbf{Y} | H_1)$ and $p(\mathbf{Y} | H_0)$ under condition of H_1 and H_0 are, respectively,

$$\begin{aligned} p(\mathbf{Y} | H_1) &= \frac{1}{\pi^{NN} (\sigma_v^2)^{NN}} e^{\text{tr}(-[(\mathbf{Y}-\mathbf{S})(\mathbf{Y}-\mathbf{S})^H]/\sigma_v^2)}, \\ p(\mathbf{Y} | H_0) &= \frac{1}{\pi^{NN} (\sigma_v^2)^{NN}} e^{\text{tr}(-[\mathbf{Y}\mathbf{Y}^H]/\sigma_v^2)}. \end{aligned} \quad (2.5)$$

Recall that the generalized likelihood ratio is defined by the ratio of the likelihood functions under each hypothesis, but it is maximized over the unknown nuisance parameter space [20], that is,

$$\max_{\Theta} \frac{p(\mathbf{Y} | H_1, \Theta)}{p(\mathbf{Y} | H_0)}, \quad (2.6)$$

where Θ denotes the nuisance parameter set, for example, the unknown signal subspace \mathbf{H} .

It is assumed that the statistical characterization of the noise is completely known a priori in this paper. Then the detection problem becomes finding a maximum likelihood estimate of the signal subspace \mathbf{H} , or equivalently, its orthogonal subspace $\mathbf{A} = [\mathbf{a}_{p+1} \cdots \mathbf{a}_N]$ under condition of H_1 [21]. The log-likelihood function of \mathbf{Y} under H_1 is (ignoring constants)

$$L = \ln p(\mathbf{Y} | H_1) = -\frac{1}{\sigma_v^2} \text{tr}[(\mathbf{Y} - \mathbf{S})^H (\mathbf{Y} - \mathbf{S})]. \quad (2.7)$$

A maximum likelihood principle to identify the orthogonal subspace \mathbf{A} can form a Lagrangian for minimizing $-2\sigma_v^2 L$ in (2.7) under the constraints $\mathbf{a}_i^H \mathbf{S} = [0 \cdots 0]$:

$$\ell = \text{tr}[(\mathbf{Y} - \mathbf{S})^H (\mathbf{Y} - \mathbf{S})] + 2 \text{tr}(\mathbf{A}^H \mathbf{S} \Xi), \quad (2.8)$$

where Ξ is an $(N - p) \times N$ matrix, it contains the Lagrangians ξ_{ij} .

Firstly, we can estimate the signal matrix \mathbf{S} assuming that the orthogonal subspace \mathbf{A} is known, and then we further maximize likelihood with respect to \mathbf{A} (cf. [21] for detailed derivations). The estimate of signal matrix is

$$\hat{\mathbf{S}} = (\mathbf{I} - P_{\mathbf{A}})\mathbf{Y}, \quad (2.9)$$

where $P_{\mathbf{A}}$ is the projection $P_{\mathbf{A}} = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$.

Substitute (2.9) into (2.7), we obtain

$$L = -\frac{1}{\sigma_v^2} \text{tr}[P_{\mathbf{A}} \mathbf{Y} \mathbf{Y}^H]. \quad (2.10)$$

In practice, \mathbf{Y} is the measurement of the transfer matrix, and $\mathbf{R} = \mathbf{Y}\mathbf{Y}^H$ is the time reversal operator. As the subspace \mathbf{A} is unknown, we must further maximize likelihood with respect to it.

Let the time reversal operator have the orthogonal decomposition below

$$\begin{aligned}\mathbf{R} &= \mathbf{U}\Lambda^2\mathbf{U}^T, \\ \mathbf{U} &= [\mathbf{u}_1 \cdots \mathbf{u}_{p+1} \cdots \mathbf{u}_N], \\ \Lambda^2 &= \text{diag}\{\lambda_1^2 \cdots \lambda_{p+1}^2 \cdots \lambda_N^2\},\end{aligned}\tag{2.11}$$

where the eigenvalue matrix Λ^2 is sorted in descending order, and the matrix \mathbf{U} is composed of the corresponding eigenvectors. It is then straightforward proved that L is bounded as follows:

$$L \leq -\frac{1}{\sigma_v^2} \sum_{i=p+1}^N \lambda_i^2\tag{2.12}$$

for any rank- $N - p$ projector $P_{\mathbf{A}}$.

If $\hat{\mathbf{A}} = [\mathbf{u}_{p+1} \cdots \mathbf{u}_N]$, the bound is achieved for a projector $P_{\mathbf{A}}$ onto the subspace $\langle \hat{\mathbf{A}} \rangle$:

$$P_{\mathbf{A}} = \hat{\mathbf{A}}\hat{\mathbf{A}}^H.\tag{2.13}$$

Clearly, the dominant p eigenvectors form a rank- p signal subspace, and the rest of the eigenvectors builds a rank- $N - p$ orthogonal subspace \mathbf{A} . The estimate of the signal subspace is

$$\hat{\mathbf{H}} = [\mathbf{u}_1 \cdots \mathbf{u}_p].\tag{2.14}$$

Substituting (2.5), (2.9), and (2.13) into (2.6), and taking the logarithm operation with constants neglected yields

$$\begin{aligned}L_{\text{DORT}} &= -\frac{1}{\sigma_v^2} \text{tr}[(\mathbf{Y} - \mathbf{S})^H(\mathbf{Y} - \mathbf{S})] + \frac{1}{\sigma_v^2} \text{tr}[\mathbf{Y}^H\mathbf{Y}] \\ &= -\frac{1}{\sigma_v^2} \sum_{i=p+1}^N \lambda_i^2 + \frac{1}{\sigma_v^2} \sum_{i=1}^N \lambda_i^2 = \frac{1}{\sigma_v^2} \sum_{i=1}^p \lambda_i^2.\end{aligned}\tag{2.15}$$

Recall that only a single target is considered, that is, $p = 1$. That means

$$L_{\text{DORT}} = \frac{1}{\sigma_v^2} \lambda_1^2.\tag{2.16}$$

Therefore the test statistic is one of the dominant eigenvalues of the time reversal operator if the signal components are assumed deterministic unknown and modeled as

a linear combination of basis vectors with an unknown signal subspace. The test statistic with minimal signal analysis is similar as that obtained in [19], which is $1 + \lambda_1^2$. We call it DORT detector. Due to its simple procedure of signal processing and no knowledge about the environment and the array configuration required, the DORT detector can be used to coarsely search the suspicious objects as a preliminary detector. In next section, we compare the detection performance of DORT detector with energy detector that is based on the same detection theory.

3. Experimental Results of Detector Performance

Since it is difficult to theoretically analyze the receiver-operating characteristics (ROC) of the DORT detector, we study the ROC by Monte Carlo simulation. This section presents the study on the detection performance of the DORT detector and the performance gain of DORT detector over the energy detector with a mix of real acoustic data from waveguide experiments and simulated noise.

3.1. Experimental Setup

The experimental setup is shown in Figure 1. The experiment was performed in a waveguide tank using a vertical SRA of 32 elements equally spaced at 0.04 m apart. The waver tank can simulate a stationary shallow water environment with multipath propagations, mainly including direct path, sea surface-reflected path, and sea bottom-reflected path. Each of the 32 elements is individually controlled and amplified during transmission and reception. The waveguide tank is 14 m in length, 1.2 m in width, and 1.4 m in height. Its three of four vertical walls are covered with anechoic tiles, the other one is a steel sheet located 12 m away from the SRA. The bottom is a sandstone basement covered with sand of 0.22 m thickness.

The transmitted signal is a 0.5 ms PCW signal centered at 18 kHz. The wavelength is about 0.08 m. The first source transducer is excited by the transmitted signal. The backscattered waves received on the N channels of the SRA are stored. This operation is repeated for all the transducers. Finally a data matrix \mathbf{Y} is generated. The target was an air-filled steel cylinder of 0.21 m (2.6λ) in diameter and 0.51 m (6.2λ) in length suspended at a range of 8.2 m and a depth of 0.84 m. This means that the range of target to SRA is about 100λ and suspended at depth of 10.5λ .

The water column has a sound of equal speed profile. The bottom is modeled as a 0.22 m sand sediment layer above a sandstone basement. The environmental parameters are following: in the water column, wave speed $c_1 = 1480$ m/s; in the bottom layer, sand density $\rho_2 = 1800$ kg/m³, sand speed $c_2 = 1650$ m/s, and attenuation $\alpha_2 = 0.67$ dB per wavelength; in the basement layer, density $\rho_3 = 1800$ kg/m³, speed $c_3 = 1580$ m/s, and attenuation $\alpha_3 = 0.8$ dB per wavelength.

3.2. Signal Processing

In order to obtain high SNR, the time reversal operator is built from short time windows. That is, the matrix $\mathbf{Y}(\omega, r)$ is the Fourier transform $\mathbf{Y}(t)$ (the measurement of transfer matrix $\mathbf{K}(t)$ between time t and $t + \Delta t$, where t is related to the distance r through the equation $t = 2r/c$; c is the sound speed; Δt is the window length [18]). Especially, r_0 corresponds to the distance

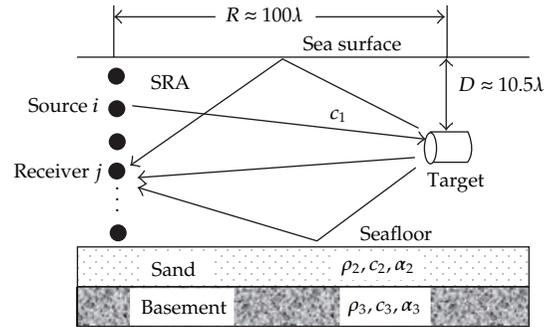


Figure 1: Experimental setup and model for the acoustic environment.

of the target. In order to obtain the test statistic in (2.16) using DORT observation matrix $\mathbf{Y}^H \mathbf{Y}$, the range of the target is firstly estimated. The matrix containing the target echo is then transformed into frequency domain. Finally the test statistic is got by the SVD decomposition of the matrix. A detailed illustration is shown in this section.

To estimate the distance of the target, a $\Delta t = 0.5$ ms time windows is shifted by steps of 0.25 ms. The selection of the windows length will be illustrated later on. For each distance, the matrices $\mathbf{Y}(r)$ are calculated at center frequency 18 kHz. After SVD decomposition, the first eigenvalues are normalized. In the approach of the energy detector, the average energy of the array at center frequency for each transmission is firstly calculated and then is accumulated to obtain the total average energy for the all transmissions. There is another simple method to calculate the energy is to summarize all eigenvalues after SVD decomposition.

Figure 2 depicts the first normalized eigenvalues and the normalized energies as a function of distance from 0 to 20 m. It can be seen that a peak value appears, respectively, at the distance of 8.3 m for DORT detector and 8.4 m for the energy detector, which tells that both detectors can correctly detect the target (the other peak at the distance of 12 m is caused by the steel sheet).

It is worth noting that the target cannot be simply treated as a point-like scatterer compared to the wavelength. It has been shown that the diagonalization of the time-reversal operator permits the various elastic components of the scattered field to be extracted, and more than one eigenvalue is associated to the scatterer [7]. However, it has been also shown the strong specular echo can be separated from the Lamb wave by selecting temporally with a proper time window in the same reference.

Since the distance of the target has been determined, the corresponding time window $\mathbf{Y}(\omega, r_0)$ which contains the target echo can also be determined. This time-windowed segment of echo data is used for the following signal analysis and detection performance investigation. Figure 3 shows eigenvalues of the time-reversal operator at the central frequency for the windows of increasing duration of $\Delta t = 0.5, 1, 1.5,$ and 2 ms. With the window duration increasing, all eigenvalues except the first one which is normalized are becoming larger and larger. It is possible that other propagation modes from the target or reverberation from other objects contribute to the signals in the time windows for long durations. Hence by choosing $\Delta t = 0.5$ ms as the window duration and selecting the time window associated to the strong specular echo, we still use the first eigenvalue as the test statistic for the DORT detector.

The backpropagation of the first eigenvector is calculated for the center frequency using KRAKEN [22]. The code KRAKEN is a normal mode propagation model developed

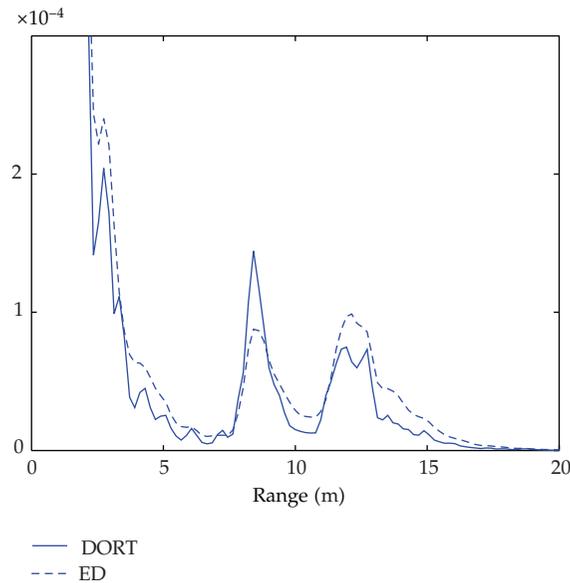


Figure 2: Target at 8.2 m: the first normalized eigenvalues (solid line) and the normalized energies (dashed line) as a function of distance from 0 m to 20 m.

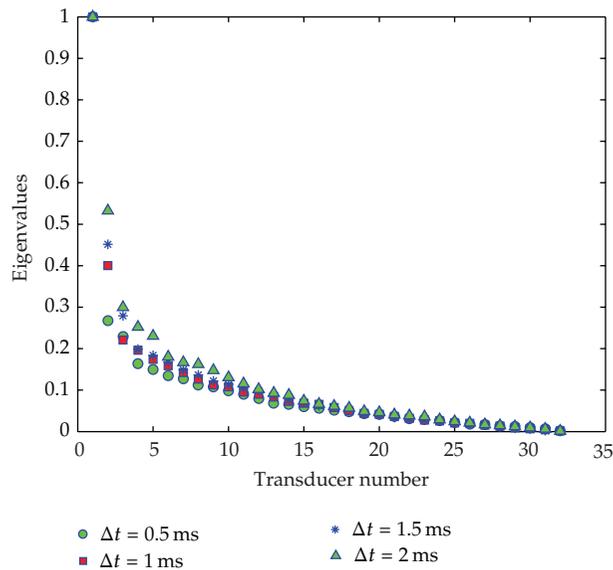


Figure 3: The experimental eigenvalues of of the time-reversal operator at the central frequency and the window duration is $\Delta t = 0.5$ ms(circle), $\Delta t = 1$ ms (square), $\Delta t = 1.5$ ms (star), and $\Delta t = 2$ ms (triangle).

at SACLANT Undersea Research Centre (now Nato Undersea Research Centre). It has been widely used for modeling ocean environments that are range independent, range dependent, or fully 3-dimensional. The image is displayed in the range from 1 m to 12 m on the whole height of the waveguide. As shown in Figure 4, the peak position in the ambiguity surface is at a range of 8.47 m and a depth of 0.85 m. The image provides information about the location

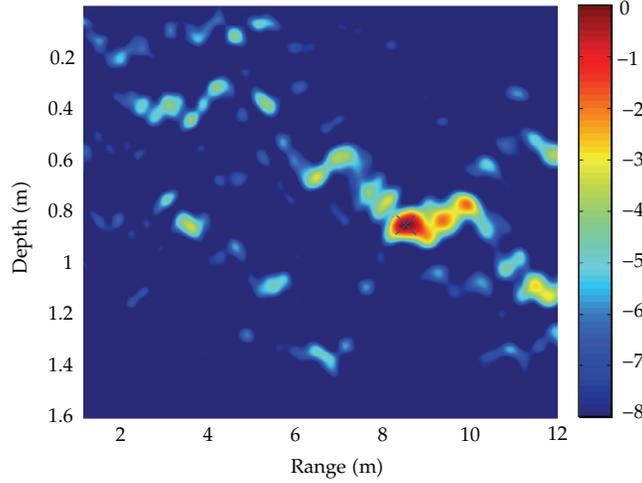


Figure 4: Numerical backpropagation of the first eigenvector using KRAKEN.

of the target, in almost agreement with the true position (8.2 m in range and 0.84 m in depth). The slight discrepancy is possibly caused by inexact environmental parameters in Figure 1.

3.3. Detection Performance

As mentioned in [23, 24], no analytical expression is available for the probability distribution of the first eigenvalue. Therefore, it is difficult to analytically express detection probability of DORT detector. The detection performance of DORT detector is numerically evaluated using Monte Carlo method.

The initial SNR is calculated by the ratio of the signal energy of the target cell to the energy of the time window near the target cell, which is about 8.2 dB. Note that the initial SNR is high, we considered that the corresponding time window $\mathbf{Y}(\omega, r_0)$ only contains the target echo; that is, the signal matrix is replaced by $\mathbf{Y}(\omega, r_0)$ containing noise in practice. To obtain noisy backscatter at different SNR, we added numerically generated zero-mean white Gaussian noise to the real data of acoustic backscatter [10]. The SNR is defined as

$$\text{SNR} = \frac{E\{\|\mathbf{S}\|^2\}}{E\{\|\mathbf{V}\|^2\}} = \frac{\sum_{j=1}^N \|\mathbf{s}_j\|^2}{N\sigma_v^2}. \quad (3.1)$$

For the practical signal processing, we set $\sigma_v = 1$ and scale the total signal energy to meet different SNR level. We then further modified the SNR levels in (3.1) according to the initial SNR to obtain values close to the actual ones.

To study the performance, we evaluated the detection probability P_D as a function of the SNR for a fixed probability of false alarm P_{FA} . In order to obtain the detection threshold, we, respectively, generated 10000 independent noise realizations and computed the test statistics given by (2.16) for DORT detector; the resulting 10000 test statistics are, respectively, sorted in ascending order; the thresholds are then selected so that $P_{FA} = 0.0001$, $P_{FA} = 0.001$, or $P_{FA} = 0.01$. To compute P_D , a new set of 5000 independent noise is generated and added

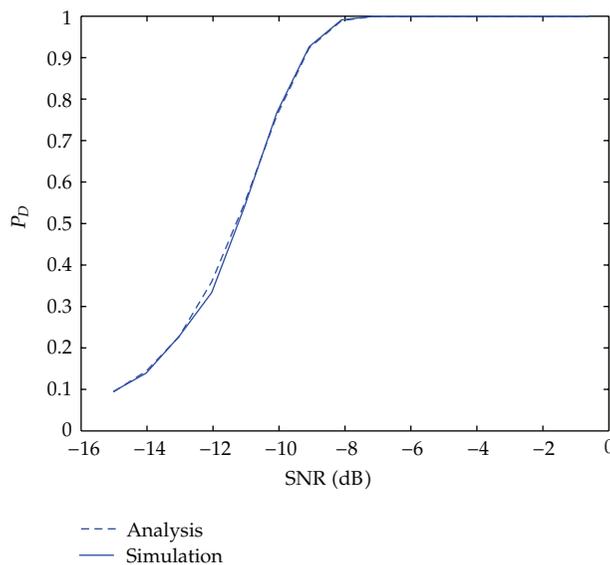


Figure 5: Detection probability versus SNR for $P_{FA} = 0.01$ using the energy detector with a mix of real acoustic data from laboratory waveguide experiment and simulated noise; analysis (dashed line) versus simulation (solid line).

to the actual acoustic data. The test statistics are, respectively, computed and compared to the corresponding thresholds. The percentage of the number of times that the test statistic exceeds the threshold is used as an estimate of the detection probability.

Figure 5 shows the results of simulation versus the ones of the energy detector for $P_{FA} = 0.01$ developed in [12]. Figure 5 indicates that the simulation result of the energy detector is in excellent agreement with the theoretical analysis, thus it validated the proposed design of simulation experiment with a mix of real acoustic data from laboratory waveguide experiment and simulated noise. Figure 6 shows the detection performance of DORT detector and that of the energy detector. It can be seen that DORT detector provides detection gain with respect to the energy detector, except that the energy detector performs better in the case that the detection probabilities for both detectors is lower than 0.2 for extremely low SNR. For instance, we measured the performance gains by the relative SNR required to achieve detection probabilities of 0.5 given false alarm rates of 0.0001, 0.001, and 0.01. The performance gains of the DORT detector with respect to the energy detector are respectively about, 1.4 dB, 1.1 dB, and 0.8 dB. It should be pointed out that no knowledge and modeling of the propagation medium are needed for the design of DORT detector.

The reason that the DORT detector has performance gain over energy detector can be explained as follows. In fact, the test statistic of the energy detector can also be expressed as the summarization of all eigenvalues $\lambda_1^2 + \lambda_2^2 + \dots + \lambda_N^2$ after SVD decomposition, instead of λ_1^2 for the DORT detector. The increment of the first eigenvalue is relatively larger than that of the others with SNR increasing, which can be implied in Figure 3. With the window duration increasing, the others are being more and more small compared to the first eigenvalue due to more noises are contained in the observation matrix.

4. Conclusion

In this paper, the target detection using passive DORT method is presented and discussed. The mathematical models of detection problem are developed using a sequence of collected

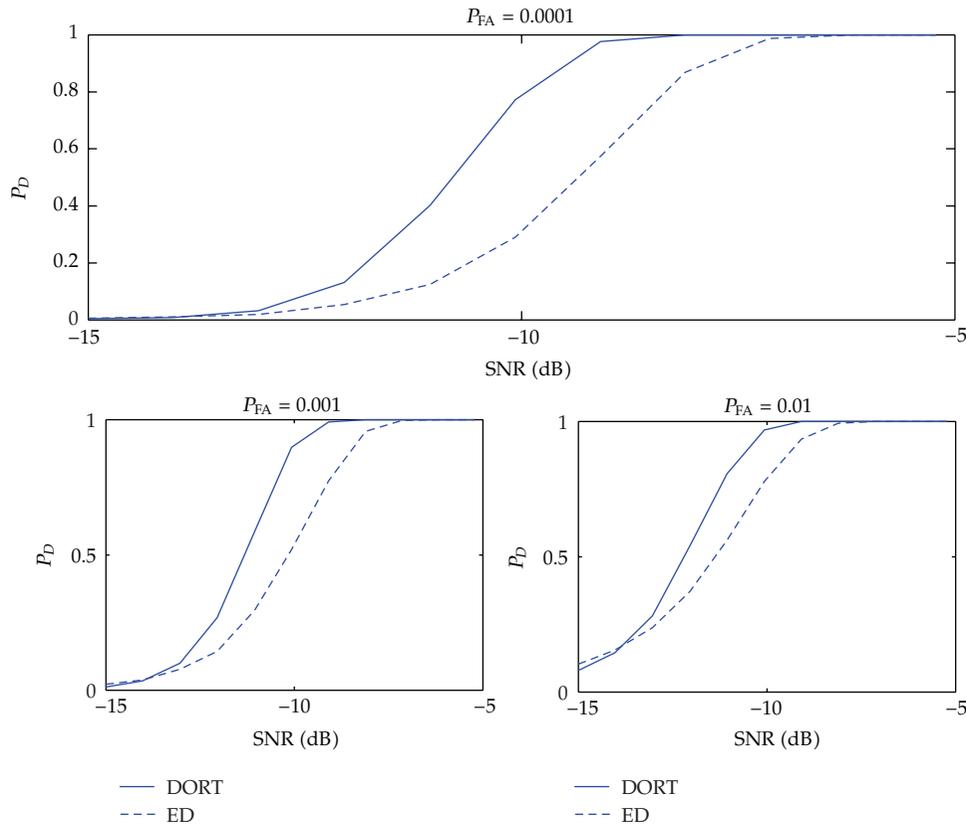


Figure 6: Detection probability versus SNR using a mix of real acoustic data from laboratory waveguide experiment and simulated noise; DORT detector (solid line) versus energy detector (dashed line).

data from typical DORT transmission. It is proved that the DORT detector can be derived with the unknown signal subspace replaced by its maximum likelihood estimate, if the signal components are assumed deterministic unknown and modeled as a linear combination of basis vectors with an unknown signal subspace; in addition, one of the dominant eigenvalue of the time reversal operator is the test statistic for point-like scatterer. Finally, the DORT detector and the energy detector are tested with the real acoustic data collected in the laboratory waveguide experiment.

The detection experiments using DORT method are conducted with the data measured from a 32 elements vertical SRA. A target has been detected and correctly located within the water depth. The experimental results show that the DORT detector can provide, respectively, 1.4 dB, 1.1 dB, and 0.8 dB performance gains over the energy detector given false alarms rate of 0.0001, 0.001, and 0.01. It should be pointed out that this paper only focuses on the passive time reversal method in which the SRA simply transmits conventional signals followed by processing of the returned time data series. Therefore, the detection gain is mainly provided by the decomposition that can separate the target echo from other contributions, but not by the adaption, say, the transmitted wavefront matched to the channel as the active time reversal methods. For future work, the detection problems for p independent point-like scatterers or in unstationary ocean environments using DORT method should be further investigated.

Acknowledgments

The authors gratefully acknowledge the support for the waveguide experiment from Hangzhou Applied Acoustics Research Institute. This work was supported by the Natural Science Foundation of China (Project no. 61101231), the Zhejiang Provincial Natural Science Foundation of China (Project no. Y1100629 and Y1100883), and the Foundation of Qianjiang Scholars (Project no. 56710203012).

References

- [1] J. P. Hermand and W. I. Roderick, "Acoustic model-based matched filter processing for fading time-dispersive ocean channels: theory and experiment," *IEEE Journal of Oceanic Engineering*, vol. 18, no. 4, pp. 447–465, 1993.
- [2] B. Friedlander and A. Zeira, "Detection of broadband signals in frequency and time dispersive channels," *IEEE Transactions on Signal Processing*, vol. 44, no. 7, pp. 1613–1622, 1996.
- [3] P. M. Baggenstoss, "On detecting linear frequency-modulated waveforms in frequency- and time-dispersive channels: alternatives to segmented replica correlation," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 4, pp. 591–598, 1994.
- [4] P. Roux and M. Fink, "Time reversal in a waveguide: study of the temporal and spatial focusing," *Journal of the Acoustical Society of America*, vol. 107, no. 5, pp. 2418–2429, 2000.
- [5] C. Prada, F. Wu, and M. Fink, "The iterative time reversal mirror: a solution to self-focusing in the pulse echo mode," *Journal of the Acoustical Society of America*, vol. 90, no. 2, pp. 1119–1129, 1991.
- [6] N. Mordant, C. Prada, and M. Fink, "Highly resolved detection and selective focusing in a waveguide using the D.O.R.T. method," *Journal of the Acoustical Society of America*, vol. 105, no. 5, pp. 2634–2642, 1999.
- [7] C. Prada and M. Fink, "Separation of interfering acoustic scattered signals using the invariants of the time—reversal operator: application to Lamb waves characterization," *Journal of the Acoustical Society of America*, vol. 104, no. 2, pp. 801–807, 1998.
- [8] C. Prada, J. L. Thomas, and M. Fink, "The iterative time reversal process: analysis of the convergence," *Journal of the Acoustical Society of America*, vol. 97, no. 1, pp. 62–71, 1995.
- [9] J. L. Li, H. F. Zhao, and W. Fang, "Experimental investigation of selective localisation by decomposition of the time reversal operator and subspace-based technique," *IET Radar, Sonar and Navigation*, vol. 2, no. 6, pp. 426–434, 2008.
- [10] J. M. F. Moura and Y. Jin, "Detection by time reversal: single antenna," *IEEE Transactions on Signal Processing*, vol. 55, no. 1, pp. 187–201, 2007.
- [11] Y. Jin and J. M. F. Moura, "Time-reversal detection using antenna arrays," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1396–1414, 2009.
- [12] C. X. Li, W. Xu, J. L. Li, and X. Y. Gong, "Time-reversal detection of multidimensional signals in underwater acoustics," *IEEE Journal of Oceanic Engineering*, vol. 36, no. 1, pp. 60–70, 2011.
- [13] S. Y. Chen, H. Tong, and C. Cattani, "Markov models for image labeling," *Mathematical Problems in Engineering*, vol. 2012, Article ID 814356, 18 pages, 2012.
- [14] C. Cattani, S. Chen, and G. Aldashev, "Information and modeling in complexity," *Mathematical Problems in Engineering*, vol. 2012, Article ID 868413, 4 pages, 2012.
- [15] S. Y. Chen, "Kalman filter for robot vision: a survey," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4409–4420, 2012.
- [16] H. C. Song, W. S. Hodgkiss, W. A. Kuperman, K. G. Sabra, T. Akal, and M. Stevenson, "Passive reverberation nulling for target enhancement," *Journal of the Acoustical Society of America*, vol. 122, no. 6, pp. 3296–3303, 2007.
- [17] J. F. Lingeitch, H. C. Song, and W. A. Kuperman, "Time reversed reverberation focusing in a waveguide," *Journal of the Acoustical Society of America*, vol. 111, no. 6, pp. 2609–2614, 2002.
- [18] C. Prada, J. De Rosny, D. Clorennec et al., "Experimental detection and focusing in shallow water by decomposition of the time reversal operator," *Journal of the Acoustical Society of America*, vol. 122, no. 2, pp. 761–768, 2007.
- [19] S. Bose and A. O. Steinhardt, "Adaptive array detection of uncertain rank one waveforms," *IEEE Transactions on Signal Processing*, vol. 44, no. 11, pp. 2801–2809, 1996.

- [20] S. M. Kay, *Fundamentals of Statistical Signal Processing, Detection Theory*, Prentice Hall, PTR Press, Upper Saddle River, NJ, USA, 1998.
- [21] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley Press, New York, NY, USA, 1991.
- [22] M. B. Porter, "The KRAKEN normal mode program," Memorandum SM-245, 1991.
- [23] A. Aubry and A. Derode, "Detection and imaging in a random medium: a matrix method to overcome multiple scattering and aberration," *Journal of Applied Physics*, vol. 106, no. 4, Article ID 044903, 2009.
- [24] A. Aubry and A. Derode, "Singular value distribution of the propagation matrix in random scattering media," *Waves in Random and Complex Media*, vol. 20, pp. 333–363, 2010.

Research Article

A Direct Solution Approach to the Inverse Shallow-Water Problem

Aleign Gessese^{1,2} and Mathieu Sellier¹

¹ Department of Mechanical Engineering, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

² School of Mechanical and Industrial Engineering, Bahir Dar University, P.O. Box 26, Bahir Dar, Ethiopia

Correspondence should be addressed to Mathieu Sellier, mathieu.sellier@canterbury.ac.nz

Received 26 July 2012; Revised 28 October 2012; Accepted 29 October 2012

Academic Editor: Fatih Yaman

Copyright © 2012 A. Gessese and M. Sellier. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The study of open channel flow modelling often requires an accurate representation of the channel bed topography to accurately predict the flow hydrodynamics. Experimental techniques are the most widely used approaches to measure the bed topographic elevation of open channels. However, they are usually cost and time consuming. Free surface measurement is, on the other hand, relatively easy to obtain using airborne photographic techniques. We present in this work an easy to implement and fast to solve numerical technique to identify the underlying bedrock topography from given free surface elevation data in shallow open channel flows. The main underlying idea is to derive explicit partial differential equations which govern this inverse reconstruction problem. The technique described here is a “one-shot technique” in the sense that the solution of the partial differential equation provides the solution to the inverse problem directly. The idea is tested on a set of artificial data obtained by first solving the forward problem governed by the shallow-water equations. Numerical results show that the channel bed topographic elevation can be reconstructed with a level of accuracy less than 3%. The method is also shown to be robust when noise is present in the input data.

1. Introduction

The study of open channel flows requires an understanding of the hydrodynamics of the flow in order to accurately capture its main characteristics. Designing and studying engineering structures on rivers, understanding the hydraulic phenomena of the watercourse for water quality control, and prediction of exceptional natural events associated with water flow constitute the three main objectives of the study of open channel flow simulation, [1]. All modelling studies of such applications require well-established governing equations along with the predefined parameters, boundary conditions, and initial conditions. Previous

studies show that there has been tremendous progress in the field of hydraulic modelling of open channel flows. However, the representation of accurate channel bed topography is still a challenge, [2–4]. Parameters such as bedrock topographic elevation and roughness coefficients are required prior to the simulation of open channel flows. The level of discrepancy in these data will have a direct effect on the quality of the simulation results. Thus in order to accurately predict the hydrodynamics of open channel flows, precise bedrock topographic elevation data and other hydraulic parameters are required. In [5] and references therein, it is argued that some hydrological models contain conceptual parameters that usually cannot be accessed by in-situ data measuring techniques. On the contrary, some models only use measurable parameters from field surveys. But the procedure is often long and expensive because it is difficult to determine parametric values at each computational grid points.

There have been significant developments in experimental techniques to measure river bathymetry and flow depths. Examples of these techniques for river bathymetry include interferometric synthetic aperture radar (SAR) digital photogrammetry [4, 6] used for the North Ashburton River, New Zealand; airborne laser altimetry (LiDAR); ground survey [2] used for the Yakima and Trinity River Basins in the USA to assess its ability to map bathymetry. However, most experimental techniques to identify the bed rock elevation are time consuming [4, 7] and expensive [2, 5] especially for wide, gravel bed, and long reach rivers.

On the other hand, many numerical techniques have been developed to simulate open channel flows. The characteristics of the numerical methods differ depending on the particular problem it addresses: methods suitable for studying hydraulic jumps; methods suitable to study dam break problems; methods to study flood plain flows; methods to study steady flows and methods suitable for high Froude numbers are relevant scenario worthy of mention. A detailed review of the relevant numerical techniques on the forward problem is beyond the scope of this study.

Comparatively, very little has been studied on the numerical reconstruction of the bedrock elevation data of open channel flows. Previous studies have shown that there are mainly two approaches that can be implemented to solve such an inverse problem: the direct numerical approach and optimization-based iterative methods. The former method was previously implemented by the authors of this work, [8–10] to infer the bedrock topography from a known free surface data in one-dimensional shallow water flows and by [11–13] to infer the substrate topography and the corresponding flow field [14] from free surface data in thin film flows. This paper extends to three-dimensional shallow-water flows the methodology presented in [8–10] which only dealt with planar flows only dealt with planar flows. In [15], a numerical bedrock reconstruction approach from known free surface elevation data is presented for zero-inertia two-dimensional shallow-water flows, while the optimization-based reconstruction approach has been implemented by many authors who tried to minimize a cost functional iteratively, see [16, 17].

In the following the forward problem refers to the simulation of open channel flows for the prediction of the water surface elevation, the velocity field, and flood coverage from known hydraulic parameters such as bedrock elevation distribution and the roughness coefficient. Conversely, we call the inverse problem when the bedrock elevation data is inferred from the given free surface elevation data and other hydraulic parameters. Section 2 introduces the shallow water equations and the respective discretization technique for the forward problem. Section 3 presents a set of forward problem benchmark test cases. Following Section 3, the inverse problem governing equations and the corresponding

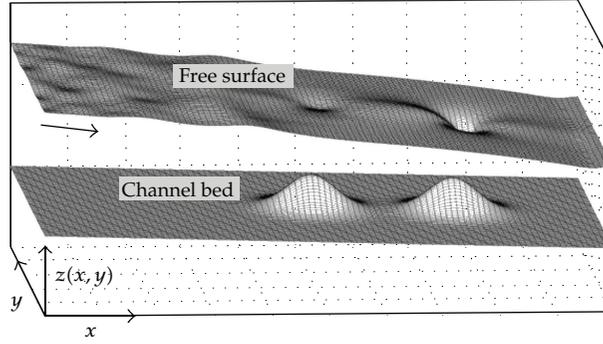


Figure 1: Two-dimensional shallow-water flow orientation.

discretization technique and the respective inverse problem test cases are presented in Section 4. In Section 5, we present the concluding remarks.

2. Governing Equations

Shallow-water flows are three dimensional in nature, but the three-dimensional Navier-Stokes equation can be simplified to a set of two-dimensional equations by considering vertically averaged quantities. This is done by integrating the Navier-Stokes equations over the flow depth for an incompressible fluid. While depth averaging, the two-dimensional shallow-water equations are derived with the following assumptions. (a) the pressure distribution is hydrostatic over the flow depth, (b) the angle of inclination of the channel is small such that flow depths measured along the vertical and normal to the channel bed directions are approximately the same, (c) the length in the vertical direction of the flow is smaller than that in the main flow direction, and (d) the Coriolis and wind stress effects are neglected, [18]. The hydrostatic pressure distribution assumption is valid for the case of long and shallow waves, (wave length is much larger than the depth of the flow), in which the vertical acceleration of fluid elements during the wave passage stays small. The solution of shallow water equations may not accurately represent open channel flows with short or high waves, because the hydrostatic pressure assumption will be violated for such flows, [18]. A general two-dimensional flow orientation is shown in Figure 1.

Thus, the two-dimensional governing equations can be written in conservative form [18] as

$$\begin{aligned}
 & H_t + E_x + F_y + S = 0, \\
 & H = \begin{pmatrix} h \\ uh \\ vh \end{pmatrix} \quad E = \begin{pmatrix} uh \\ u^2h + \frac{1}{2}gh^2 \\ uvh \end{pmatrix} \quad F = \begin{pmatrix} uh \\ uvh \\ v^2h + \frac{1}{2}gh^2 \end{pmatrix} \quad S = \begin{pmatrix} 0 \\ -gh(S_{ox} - S_{fx}) \\ -gh(S_{oy} - S_{fy}) \end{pmatrix}, \\
 & S_{fx} = \frac{n^2 u \sqrt{u^2 + v^2}}{C_o h^{4/3}} \quad S_{fy} = \frac{n^2 v \sqrt{u^2 + v^2}}{C_o h^{4/3}}, \\
 & S_{ox} = -\frac{\partial z}{\partial x} \quad S_{oy} = -\frac{\partial z}{\partial y},
 \end{aligned} \tag{2.1}$$

where $u(x, y, t)$ is the velocity component in x direction; $v(x, y, t)$ is the velocity component in y direction; $h(x, y, t)$ is the depth of the flow; C_o is dimensional constant: $C_o = 1$ for SI unit system; n is the Manning's roughness coefficient; $z(x, y)$ is bed topography; g is the acceleration due to gravity. The free surface function, $\psi(x, y, t) = h(x, y, t) + z(x, y)$, is the measure of the water surface elevation, and the corresponding depth of the flow is the difference between the water surface and the channel bed elevations.

2.1. Discretization of the Governing Equations of the Forward Problem

The discretization of the governing equations is one of the most important steps in the numerical solution of partial differential equations in engineering and science. Many discretization techniques have been developed to suit specific problems. In the context of this work, the MacCormack explicit numerical scheme is implemented for the forward problem. It has been chosen because it is a standard technique which is reasonably easy to implement and which produces reliable results. This discretization technique has a predictor and a corrector stage coupled with sequential solution methodology. Backward differencing numerical technique is implemented in the predictor stage, while forward differencing is used in the corrector stage. The flow variables are known at time step k , and their values at $k + 1$ are to be determined, [18]. Then, for grid point (i, j) , the approximate equation can be written as the following.

Predictor Stage:

$$H_{i,j}^* = H_{i,j}^k - \frac{\Delta t}{\Delta x} \nabla_x E_{i,j}^k - \frac{\Delta t}{\Delta y} \nabla_y F_{i,j}^k - \Delta t S_{i,j}^k \quad \begin{cases} 2 \leq i \leq N, \\ 2 \leq j \leq M. \end{cases} \quad (2.2)$$

Corrector Stage:

$$H_{i,j}^{**} = H_{i,j}^k - \frac{\Delta t}{\Delta x} \Delta_x E_{i,j}^* - \frac{\Delta t}{\Delta y} \Delta_y F_{i,j}^* - \Delta t S_{i,j}^* \quad \begin{cases} 1 \leq i \leq N - 1, \\ 1 \leq j \leq M - 1, \end{cases} \quad (2.3)$$

in which H^* and H^{**} are intermediate values of H . The new values of the vector H at time $k + 1$ are obtained from

$$H_{i,j}^{k+1} = \frac{1}{2} (H_{i,j}^* + H_{i,j}^{**}). \quad (2.4)$$

As mentioned above the scheme first uses backward space differencing (∇_x and ∇_y) to predict an intermediate solution from the known information at time step k . Then the forward

differences (Δ_x and Δ_y) are used in the corrector stage to correct the predicted values. The finite difference operators (∇) and (Δ) are defined as

$$\begin{aligned}\nabla_x H_{i,j} &= H_{i,j} - H_{i-1,j}, \\ \Delta_x H_{i,j} &= H_{i+1,j} - H_{i,j}.\end{aligned}\quad (2.5)$$

The subscripts in the operators show the direction of differencing. Substituting the flux terms as $U = uh$ and $V = vh$, the final form of the discretized equations can be written as the following.

Predictor Stage:

$$\begin{aligned}h_{i,j}^* &= h_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i,j}^k - U_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (V_{i,j}^k - V_{i,j-1}^k), \\ U_{i,j}^* &= U_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i,j}^k u_{i,j}^k - U_{i-1,j}^k u_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (U_{i,j}^k v_{i,j}^k - U_{i,j-1}^k v_{i,j-1}^k) - \frac{1}{2} g \frac{\Delta t}{\Delta x} \\ &\quad \times \left[(h_{i,j}^k)^2 - (h_{i-1,j}^k)^2 \right] + g h_{i,j}^k \Delta t (S_{ox} - S_{fx})_{i,j}, \\ V_{i,j}^* &= V_{i,j}^k - \frac{\Delta t}{\Delta x} (V_{i,j}^k u_{i,j}^k - V_{i-1,j}^k u_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (V_{i,j}^k v_{i,j}^k - V_{i,j-1}^k v_{i,j-1}^k) - \frac{1}{2} g \frac{\Delta t}{\Delta y} \\ &\quad \times \left[(h_{i,j}^k)^2 - (h_{i,j-1}^k)^2 \right] + g h_{i,j}^k \Delta t (S_{oy} - S_{fy})_{i,j}.\end{aligned}\quad (2.6)$$

Corrector Stage:

$$\begin{aligned}h_{i,j}^{**} &= h_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i+1,j}^* - U_{i,j}^*) - \frac{\Delta t}{\Delta y} (V_{i,j+1}^* - V_{i,j}^*), \\ U_{i,j}^{**} &= U_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i+1,j}^* u_{i+1,j}^* - U_{i,j}^* u_{i,j}^*) - \frac{\Delta t}{\Delta y} (U_{i,j+1}^* v_{i,j+1}^* - U_{i,j}^* v_{i,j}^*) - \frac{1}{2} g \frac{\Delta t}{\Delta x} \left[(h_{i+1,j}^*)^2 - (h_{i,j}^*)^2 \right] \\ &\quad + g h_{i,j}^* \Delta t (S_{ox}^* - S_{fx}^*)_{i,j}, \\ V_{i,j}^{**} &= V_{i,j}^k - \frac{\Delta t}{\Delta x} (V_{i+1,j}^* u_{i+1,j}^* - V_{i,j}^* u_{i,j}^*) - \frac{\Delta t}{\Delta y} (V_{i,j+1}^* v_{i,j+1}^* - V_{i,j}^* v_{i,j}^*) - \frac{1}{2} g \frac{\Delta t}{\Delta y} \left[(h_{i,j+1}^*)^2 - (h_{i,j}^*)^2 \right] \\ &\quad + g h_{i,j}^* \Delta t (S_{oy}^* - S_{fy}^*)_{i,j}.\end{aligned}\quad (2.7)$$

The new values at time step $k + 1$ are then calculated from the intermediate values which are determined from the predictor and corrector steps:

$$h_{i,j}^{k+1} = \frac{1}{2}(h_{i,j}^* + h_{i,j}^{**}); \quad U_{i,j}^{k+1} = \frac{1}{2}(U_{i,j}^* + U_{i,j}^{**}); \quad V_{i,j}^{k+1} = \frac{1}{2}(V_{i,j}^* + V_{i,j}^{**}), \quad (2.8)$$

whereas the primitive variables will be determined from the computed values of U and V in each time step:

$$h_{i,j}^{k+1} = h_{i,j}^{k+1}; \quad u_{i,j}^{k+1} = \frac{U_{i,j}^{k+1}}{h_{i,j}^{k+1}}; \quad v_{i,j}^{k+1} = \frac{V_{i,j}^{k+1}}{h_{i,j}^{k+1}}. \quad (2.9)$$

Reflective wall boundary is implemented at the side walls. This can be implemented by replacing the fictitious point in the solid wall by its mirror point in the flow domain while changing the sign of the normal component of velocity.

However, the inlet and the outlet boundary conditions are handled depending on the type of the flow. For subcritical flow the depth of the flow at the outlet boundary condition is defined with the flow rate at the upstream boundary. On the contrary, the depth and the flow rate at the upstream boundary are defined for supercritical flows. However, supercritical and transcritical flow test cases are not included in this study because of the limitation of the proposed approach. On the other side of the defined boundary conditions, it is assumed that the gradient of the parameters along the flow direction is zero. The transverse velocity component at the inlet is defined also as zero.

3. The Forward Problem

The governing equations of the forward problem along with its discretization are presented in the above section. The parameters that will be determined from the forward problem are the free surface elevation, the velocity components of the flow, and the depth of the flow. For this purpose some hydraulic parameters are required prior to the computation. These include bed topography elevation, roughness coefficient, flow rate, and the depth of the flow at the boundary depending on the flow regime. In the following, the results of the forward problem test cases are presented, and these results will be used as input data for the validation and assessment of the inverse problem methodology.

3.1. Test Case I: One-Dimensional Steady Subcritical Flow over a Frictionless Channel

A 1 m wide 25 m long channel is considered to test subcritical flow over a bump. The channel bed is assumed to have a rectangular cross section and frictionless with a bump. The bed topography is defined by

$$z = \begin{cases} 0 & x < 8, \quad x > 12, \\ 0.2 - 0.05(x - 10)^2 & 8 \leq x \leq 12. \end{cases} \quad (3.1)$$

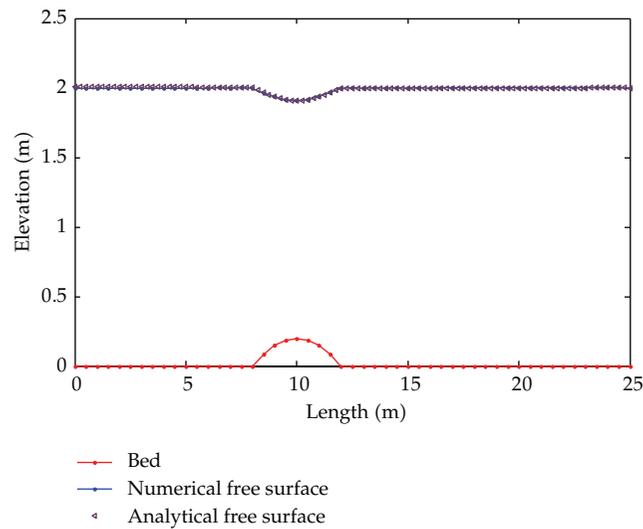


Figure 2: Free surface and bedrock topographies.

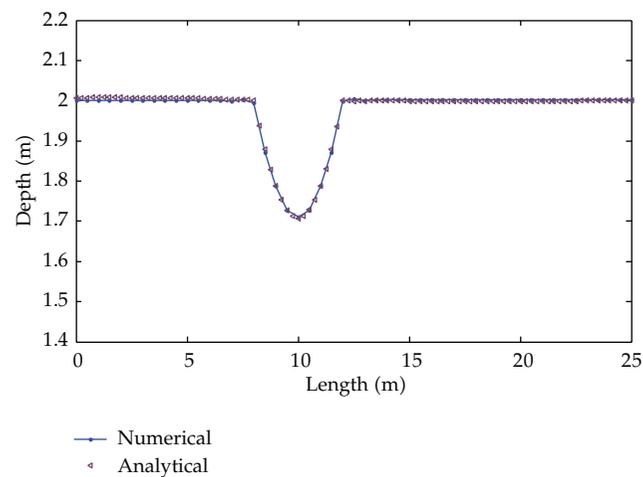


Figure 3: Depth of the flow.

The depth at downstream boundary condition of the flow is set as $h = 2$ m, and the water inflow condition $Q = 4.42 \text{ m}^3/\text{s}$ is imposed for the computation of results. This test case is the most common canonical test case used by different authors, [19–21] for example, in order to validate their numerical techniques for the convergence of the solution towards steady state and the conservation of discharge along the channel. The Froude numbers for this test case range from 0.496–0.635 showing that the flow is subcritical in the entire domain.

The results are in a good agreement with the results presented in the above references and with the respective analytical solution presented in [22]. As can be seen in Figures 2 and 3, the existence of the obstacle creates significant change on the free surface and depth profiles.

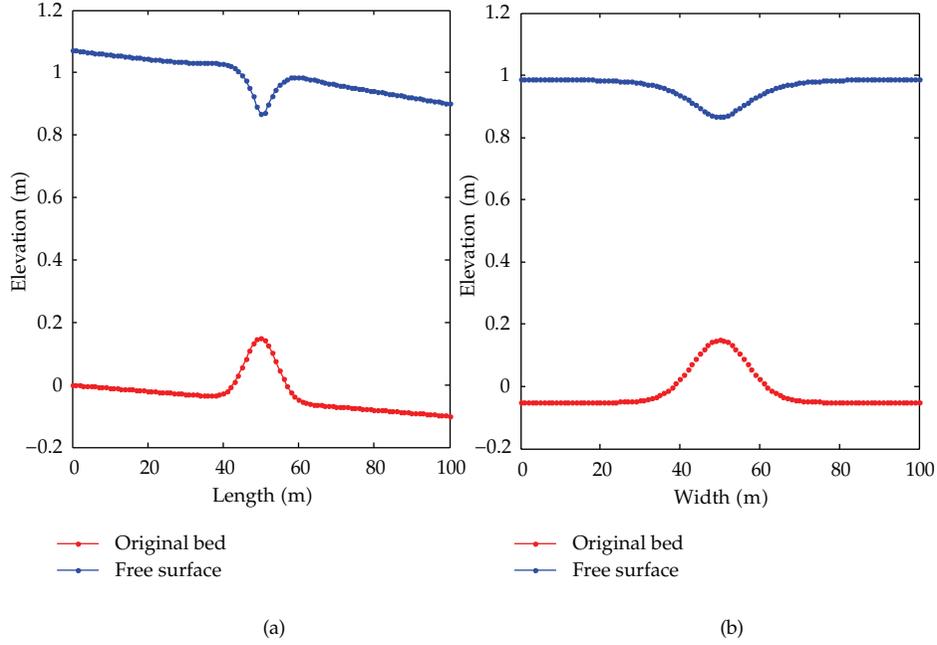


Figure 4: Centreline bed and free surface elevation variation (a) along the length and (b) along the transverse direction.

3.2. Test Case II: Three-Dimensional Flow over a Three-Dimensional Bump

In this test case, an artificial bed topography with a bump having three-dimensional characteristics is considered to simulate three-dimensional shallow water flows. Thus variation of the bed along the longitudinal and the transverse direction is considered in a computational domain of size 100 m by 100 m. The bump is defined by

$$z = 0.2 \exp\left(-\left(x - \frac{x_{\max}}{2}\right)^2 \frac{1}{\alpha_1^2} - \left(y - \frac{y_{\max}}{2}\right)^2 \frac{1}{\alpha_2^2}\right) - 0.001x, \quad (3.2)$$

where x_{\max} and y_{\max} are the length scales of the domain in x and y directions. In this test case $x_{\max} = y_{\max} = 100$ m and $\alpha_1 = 6$ $\alpha_2 = 10$ are constants. A spatial grid size $\Delta x = 1$ m and the temporal grid size $\Delta t = 0.1$ s are implemented. The flow rate $Q = 2$ m³/s per unit width is defined at the inlet boundary. A Manning's roughness value $n = 0.02$ is chosen to account for the effect friction. A subcritical flow is considered such that at the outlet a depth is 1.0 m is imposed along the transverse direction. Steady solution is generated from the transient governing equation and steady boundary conditions. The Froude numbers for this test case range from 0.56–0.92 showing that the flow is subcritical in the entire domain. In the following, plots of the bedrock elevation along the centre lines and the respective depth of the flow for forward problem are presented.

From the given bedrock topographic elevation, steady flow rate, roughness coefficient, and boundary conditions, the steady flow results are generated. The free surface variation along the centreline is given in Figure 4.

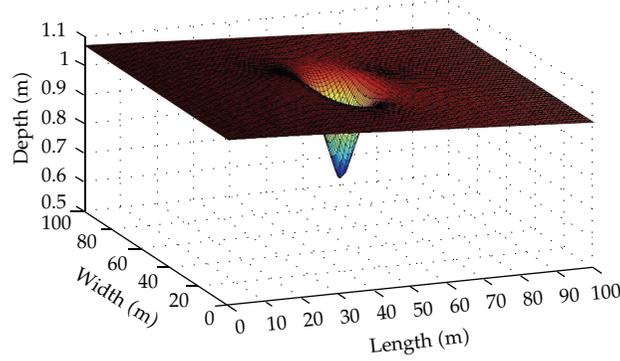


Figure 5: Depth variation.

As can be seen from Figures 4 and 5, the existence of the bump at the centre of flow domain creates a three-dimensional variation in the free surface elevation. The underlying hypothesis of this work is that the shape of this free surface with known boundary conditions contains sufficient information to reconstruct the underlying bedrock. This will be discussed in detail in the inverse problem analysis section. The free surface elevation data of this test case will be used as a known parameter along with other boundary conditions for the inverse problem analysis.

3.3. Test Case III: Three-Dimensional Flow over Number of Three-Dimensional Bumps

In order to investigate the effect of a more complex bedrock topography on the free surface, four bumps are considered in the computational domain. In this test case $\alpha_1 = 6$ $\alpha_2 = 10$ values in the bed topography function are considered. The flow rate $Q = 4.42 \text{ m}^3/\text{s}$ per unit width defined at the inlet boundary along with a depth of 2.0 m at the outlet boundary is imposed along the transverse direction. Grid sizes and Manning's roughness coefficient values considered here are similar to that of test case II. The free surface elevation is generated for the inverse problem analysis by solving numerically (2.6), (2.7), (2.8), and (2.9). The bedrock topography is given by

$$\begin{aligned}
 z = & 0.2 \exp\left(-\left(x - \frac{x_{\max}}{3}\right)^2 \frac{1}{\alpha_1^2} - \left(y - \frac{y_{\max}}{2}\right)^2 \frac{1}{\alpha_2^2}\right) \\
 & + 0.2 \exp\left(-\left(x - \frac{2}{3}x_{\max}\right)^2 \frac{1}{\alpha_1^2} - \left(y - \frac{y_{\max}}{2}\right)^2 \frac{1}{\alpha_2^2}\right) \\
 & + 0.2 \exp\left(-\left(x - \frac{x_{\max}}{2}\right)^2 \frac{1}{\alpha_1^2} - \left(y - \frac{y_{\max}}{3}\right)^2 \frac{1}{\alpha_2^2}\right) \\
 & + 0.2 \exp\left(-\left(x - \frac{x_{\max}}{2}\right)^2 \frac{1}{\alpha_1^2} - \left(y - \frac{2}{3}y_{\max}\right)^2 \frac{1}{\alpha_2^2}\right) - 0.001x.
 \end{aligned} \tag{3.3}$$

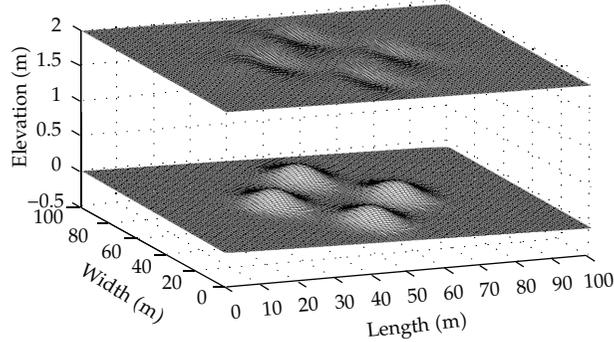


Figure 6: Water level on an inclined bed with four bumps.

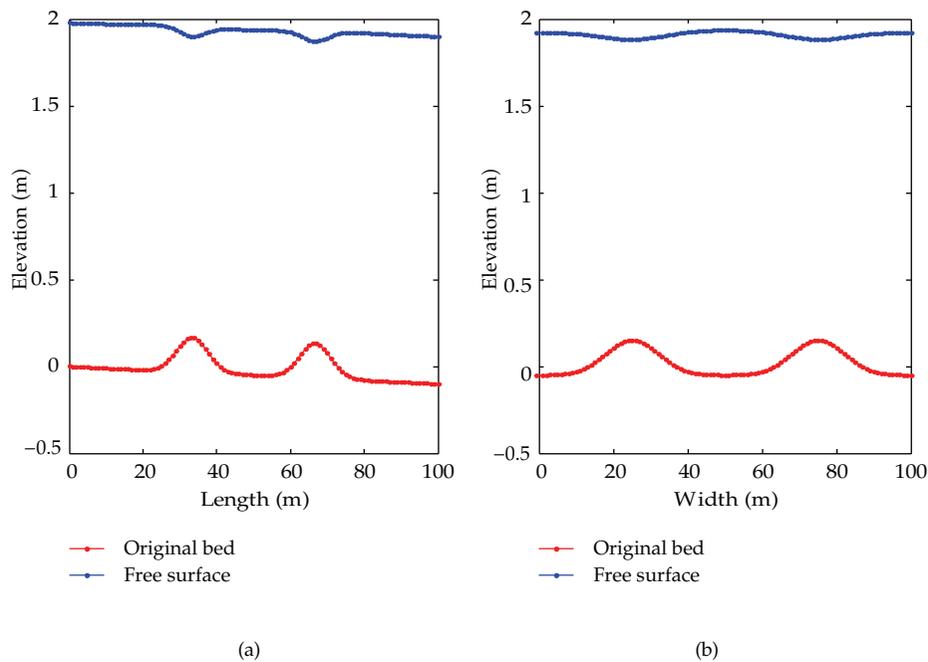


Figure 7: Bed and free surface along the centreline (a) longitudinal (b) transverse variation.

From Figures 6 and 7, it can be seen that each bump introduced on the bed has its own effect on the free surface. The undulations on the free surface are the manifestations of the bump on the bedrock, and these free surface undulations, if accurately captured, can provide sufficient information of the bedrock and other parameters. The Froude numbers for this test case range from 0.49–0.6 showing that the flow is subcritical in the entire domain. The free surface generated in this test case will also be used as input data for the respective inverse problem test case.

4. The Inverse Problem

Unlike the forward problem the inverse problem analysis will be dealing with identifying hydraulic parameters from the known fields and parameters. The shallow-water equations will be rearranged and used for the inverse problem analysis. However, there is a difference between the known and unknown parameters. The governing equation of the inverse problem can be rewritten after substituting a free surface function by $\psi(x, y) = z(x, y) + h(x, y)$

$$\begin{aligned}
 H_t + E_x + F_y + S &= 0, \\
 H &= \begin{pmatrix} h \\ uh \\ vh \end{pmatrix} \quad E = \begin{pmatrix} uh \\ u^2h \\ uvh \end{pmatrix} \quad F = \begin{pmatrix} vh \\ uvh \\ v^2h \end{pmatrix} \quad S = \begin{pmatrix} 0 \\ -gh(S_{\psi x} - S_{fx}) \\ -gh(S_{\psi y} - S_{fy}) \end{pmatrix}, \\
 S_{fx} &= \frac{n^2 u \sqrt{u^2 + v^2}}{C_o h^{4/3}} \quad S_{fy} = \frac{n^2 v \sqrt{u^2 + v^2}}{C_o h^{4/3}}, \\
 S_{\psi x} &= -\frac{\partial \psi}{\partial x} \quad S_{\psi y} = -\frac{\partial \psi}{\partial y},
 \end{aligned} \tag{4.1}$$

where $\psi(x, y)$ is the free surface elevation function.

The discretization of the inverse problem governing equations differs from the forward problem in such a way that only backward differencing explicit numerical scheme is implemented. This approach has stability restriction but is easy to implement and fast to provide a solution:

$$\begin{aligned}
 h_{i,j} &= h_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i,j}^k - U_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (V_{i,j}^k - V_{i,j-1}^k), \\
 U_{i,j} &= U_{i,j}^k - \frac{\Delta t}{\Delta x} (U_{i,j}^k u_{i,j}^k - U_{i-1,j}^k u_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (U_{i,j}^k v_{i,j}^k - U_{i,j-1}^k v_{i,j-1}^k) \\
 &\quad + gh_{i,j}^k \Delta t (S_{\psi x} - S_{fx})_{i,j}, \\
 V_{i,j} &= V_{i,j}^k - \frac{\Delta t}{\Delta x} (V_{i,j}^k u_{i,j}^k - V_{i-1,j}^k u_{i-1,j}^k) - \frac{\Delta t}{\Delta y} (V_{i,j}^k v_{i,j}^k - V_{i,j-1}^k v_{i,j-1}^k) \\
 &\quad + gh_{i,j}^k \Delta t (S_{\psi y} - S_{fy})_{i,j}.
 \end{aligned} \tag{4.2}$$

The boundary condition is implemented in a similar way as that of the forward problem. The new values at time step $k + 1$ are calculated from the previous time step values.

In this analysis, the expression used to find the new values differs from the forward problem approach. This is because, in the inverse problem, unlike to the forward problem analysis, no information is transferred from downstream to upstream direction making the backward difference approach a suitable choice. As the backward differencing scheme is implemented, the definition of the outlet boundary condition does not affect the solution of the problem in the upwind direction as there is no upstream wave propagation. Thus, inlet boundary condition is implemented. The primitive variables will be determined from the

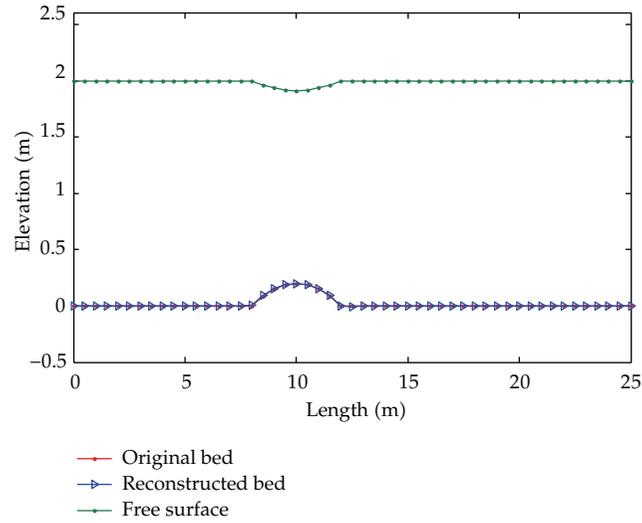


Figure 8: Reconstructed bed: subcritical test case.

computed values of U and V in each time step. Once the values of the primitive variables are determined the bed topography elevation will be evaluated by a simple subtraction for the depth of the flow from the free surface elevation function:

$$z_{i,j}^{k+1} = \psi_{i,j}^{k+1} - h_{i,j}^{k+1}. \quad (4.3)$$

The results of the inverse problem will be counterchecked with the forward problem results in order to identify the capability of the algorithm in reconstructing the bed topographic elevation. Thus similar test cases will be considered for the numerical experiment of the algorithm.

4.1. Test Case I: Steady One-Dimensional Subcritical Flow over a Bump in a Frictionless Channel

The free surface elevation obtained from the forward problem analysis is used as an input parameter for the inverse problem. This includes the free surface profile and the boundary and initial conditions along with the steady inflow rate. Unlike the depth at the downstream boundary in the case of the forward problem, the inverse problem requires boundary conditions at the inlet boundary. A depth at the upstream boundary $h = 2$ m and the water inflow rate $Q = 4.42$ m³/s are imposed. A time step of 0.05 sec and spatial grid size of 0.5 m are used in the computation.

Figure 8 shows the comparison of the reconstructed and the actual bed forms. In Figure 9 the reconstructed and the actual depth of the flow are compared. The results show that there is a perfect agreement between the reconstructed and the forward problem parameters. Relative to the results presented in [16], this approach has a capability to overcome the constant shift of the reconstructed bed topography. This confirms that the algorithm can be used for bedrock reconstructions of open channel flows of one-dimensional nature.

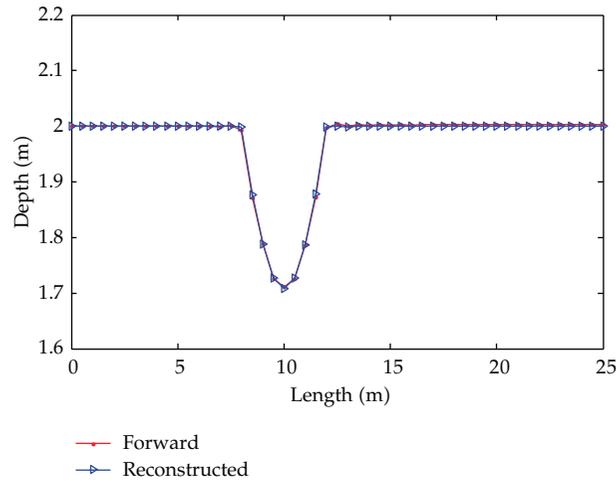


Figure 9: Comparison of the reconstructed and forward problem depth variation for the case of subcritical flow.

4.2. Test Case II: Steady Three-Dimensional Flows over a Bump in a Rectangular Channel

This test case is used to test the algorithm for its capability to reconstruct the unknown bed topographic elevation from known free surface elevation data which was generated in the forward test case II. In the respective forward test case, the effect of bed slope, friction coefficient, and local bed elevation was implemented to generate the free surface profile which we use as input parameter. Along with the free surface profile, depth variation and flow rate at the inlet boundary are used. For the inverse problem analysis, a spatial grid size $\Delta x = 1$ m and the time step $\Delta t = 0.05$ sec are implemented. The flow rate of $Q = 2 \text{ m}^3/\text{s}$ per unit width and depth at the upstream boundary which is imported from the forward problem are used. A Manning's roughness value $n = 0.02$ is used as in the case of the forward problem. A total computation time of 100 seconds was sufficient to generate steady reconstructed parameters.

As can be seen from Figure 10, the bedrock is successfully reconstructed with a very good agreement with the actual bed used in the forward problem. The bump is well reconstructed along with the bed slope showing that the methodology used can use the signal from the free surface to successfully identify its cause. Quantitatively the bed topography is reconstructed with 3% maximum deviation. The sensitivity of the algorithm to the introduction of noise in the free surface data was also tested on ranges of noise magnitude with respect to the signal on the free surface. Noise magnitudes ranging from 1–5% of the “peak-to-peak” free surface deviations are considered. In Figure 11, a reconstructed bed topography from a 1% noise is shown, and it is evident that the noise is not amplified in the reconstructed bed.

4.3. Test Case III: The Reconstruction of Bed Topography Used in Test Case III of the Forward Problem

Like the above inverse problem test cases, the free surface topographic data is used as input parameter to reconstruct the corresponding bed topographic elevation. Additionally a flow

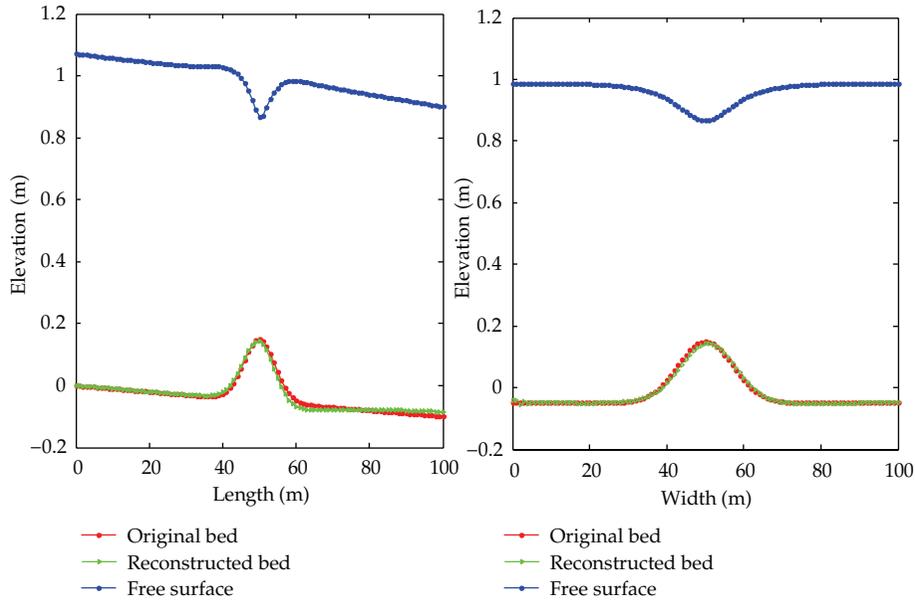


Figure 10: Comparison of original and reconstructed bed forms.

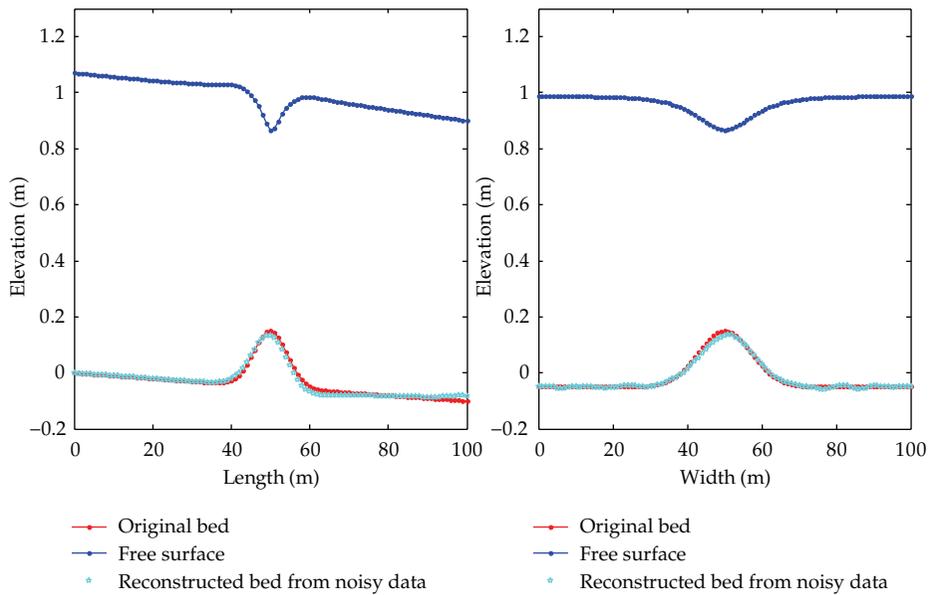


Figure 11: Reconstructed bed form from noisy free surface data.

rate of $Q = 4.42 \text{ m}^3/\text{s}$ per unit width and depth ($h = 2 \text{ m}$) at the upstream boundary are imported from the corresponding forward problem. The reconstructed bed along with the free surface profile is shown in Figure 12.

As can be seen from Figure 12, the four bumps in the original bump are successfully reconstructed. However, there is an apparent but small difference, downstream of the bumps,

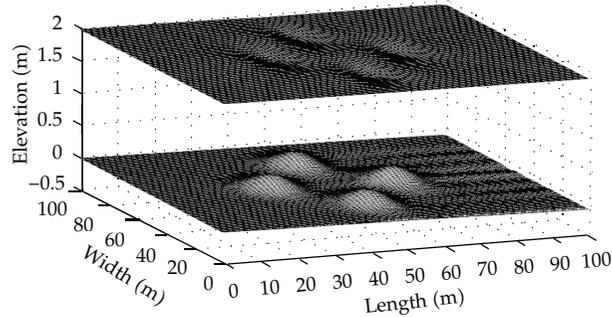


Figure 12: Reconstructed bed topographic elevation.

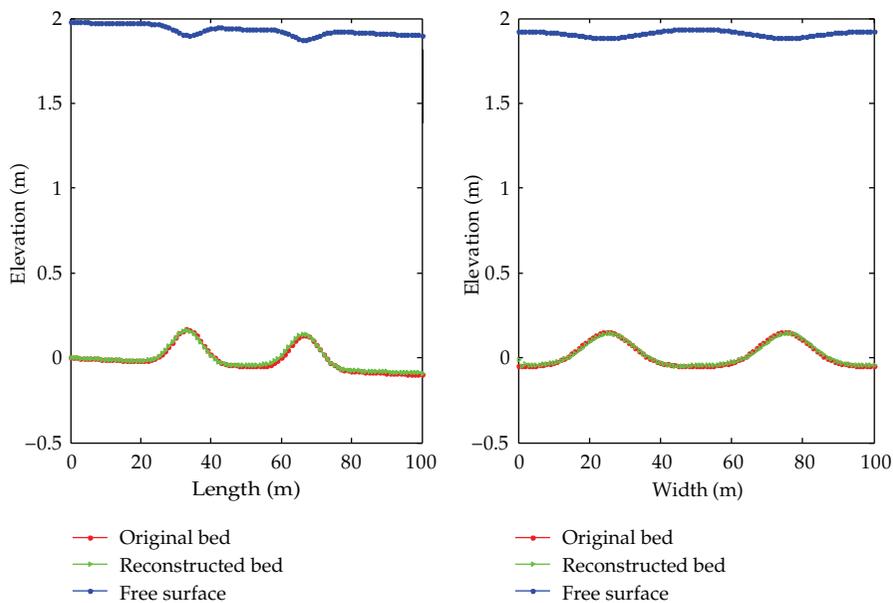


Figure 13: Comparison of the reconstructed and original bed forms.

between the reconstructed and the original bed. The comparison of the bed topography along the centreline is presented in Figure 13.

As can be seen in Figure 13, the reconstructed and original bed forms are in good agreement with each other. The difference between the original and the reconstructed bed lies below 0.5%.

Like test case II, the reconstruction algorithm is tested for its applicability to infer the channel bed topography from noisy free surface elevation data. The magnitude of the noise introduced is 1% of the signal of the free surface variation, and the results (Figure 14) show that the bed topography is reconstructed without the amplification of the noise in the reconstructed bed elevation.

The above test cases are all based on numerical results with smooth free surfaces which are used as input variable. However, in practice it is difficult to get a smooth free surface from field measurements. Often the measured free surface elevation includes noise. Thus,

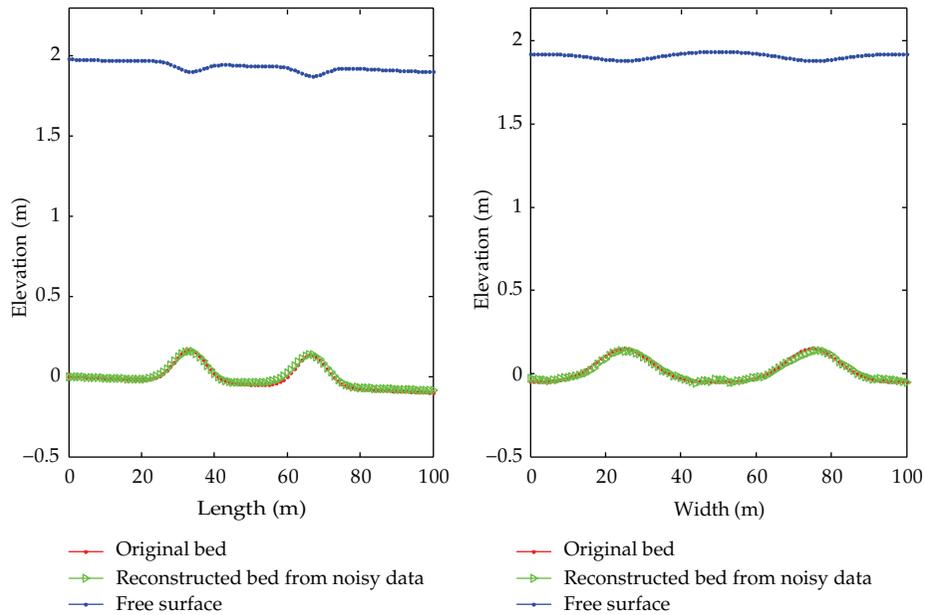


Figure 14: Comparison of the original and bed reconstructed forms from noisy free surface data.

to identify the capability of the algorithm to handle noisy input data, we introduce 1, 2.5, and 5% noise in the free surface data of all of the previous test cases. The noise is based on the signal on the free surface. A Savitzky-Golay smoothing filter is used in MATLAB to smooth the 2D noisy free surface data before the computation. The reconstruction of the bed topographies from the noisy data revealed that the level of discrepancy in the reconstructed bedrock elevation is in the range of 1%, 5%, and 9.5%, respectively.

The algorithm used for the inverse reconstruction of the bedrock topography from the free surface elevation data is fast and easy to implement. It has the ability to reconstruct the bed from noisy free surface data with the help of smoothing techniques. The methodology requires the values of the steady flow rate, depth of the flow at inlet boundary, and the roughness coefficient in addition to the free surface data. The flow in all test cases is subcritical as it is shown by the calculated ranges of the Froude number. Similar analysis indicated that the proposed approach is not suitable for transcritical and supercritical flows because hydraulic jumps and surges are not accommodated by the solution approach; thus it requires further developments to perform bedrock reconstruction in the cases of such flows.

The presented methodology works well with steady flows of subcritical nature. The effect of a flux change in the flow will have local effect on the free surface which will send a wrong signal to the algorithm in the reconstruction process. Thus scope of this study is therefore limited to steady shallow open channel flows.

5. Conclusion

The study of open channel flow modelling calls for input parameters like bed topography and roughness coefficient in order to accurately predict the hydrodynamics of the flow. A methodology based on an explicit finite difference scheme is used to reconstruct the bedrock

elevation from the free surface data. The methodology requires a steady flow, the knowledge of the roughness coefficient, and the depth of the flow at the inlet boundary.

The algorithm is tested on a set of benchmark test cases, and encouraging results are found. The bed topography is well reconstructed with a deviation of 3% or less. The numerical methodology is easy to implement and fast to produce a solution but has a CFL restriction because of its explicit nature. This approach is suitable for steady open channel flows for which the shallow-water approximation holds where the signal due to the existence of the underlying bed topography is captured by the free surface measurement technique.

In practice, the measured free surface contains noise. The presented methodology is found to be capable of reconstructing the channel bed topography from noisy data. When tested on a set of noisy numerical data, the methodology was found to introduce no noise amplification in all test cases considered.

References

- [1] J. A. Cunge, F. M. Holly, and A. Verwey, *Practical Aspects of Computational River Hydraulics*, Pitam Publishing, London, UK, 1980.
- [2] R. Hilldale and D. Raff, "Assessing the ability of airborne LiDAR to map river bathymetry," *Earth Surface Processes and Landforms*, vol. 33, pp. 773–783, 2008.
- [3] K. Marks and P. Bates, "of high-resolution topographic data with floodplain flow models," *Integration Hydrological Processes*, vol. 14, pp. 2109–2122, 2000.
- [4] R. M. Westaway, S. N. Lane, and D. M. Hicks, "The development of an automated correction procedure for digital photogrammetry for the study of wide, shallow, gravel-bed rivers," *Earth Surface Processes and Landforms*, vol. 25, pp. 209–225, 2000.
- [5] H. Roux and D. Dartus, "Estimating hydraulic parameters and geometric characteristics of a river from remote sensing data using optimization methods," in *Proceedings of the 2nd International Conference on Fluvial Hydraulics*, A. A. BALKEMA, Napoly, Italy, 2004.
- [6] R. M. Westaway, S. N. Lane, and D. M. Hicks, "Remote sensing of clear-water, shallow, gravel-bed rivers using digital photogrammetry," *Photogrammetric Engineering & Remote Sensing*, vol. 67, no. 11, pp. 1271–1281, 2001.
- [7] Y. Hirose and Y. Imai, "Airborne remote sensing for river environmental assessment," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 1149–1152, 2008.
- [8] A. F. Gessese, M. Sellier, E. Van Houten, and G. Smart, "Inferring channel bed topography from known free surface data," in *Proceedings of the 34th World Congress of the International Association for Hydro-Environment Engineering and Research (IAHR '11)*, Brisbane, Australia, June 2011.
- [9] A. F. Gessese, M. Sellier, E. Van Houten, and G. Smart, "Reconstruction of river bed topography from free surface data using a direct numerical approach in one-dimensional shallow water flow," *Inverse Problems*, vol. 27, no. 2, 2011.
- [10] A. F. Gessese, G. Smart, C. Heiningd, and M. Sellier, "One-dimensional bathymetry based on velocity measurements," *Inverse Problems in Science and Engineering*, pp. 1–17, 2012.
- [11] M. Sellier, "Substrate design or reconstruction from free surface data for thin film flows," *Physics of Fluids*, vol. 20, no. 6, Article ID 062106, 2008.
- [12] C. Heining and N. Aksel, "Bottom reconstruction in thin-film flow over topography: steady solution and linear stability," *Physics of Fluids*, vol. 21, no. 8, Article ID 083605, 10 pages, 2009.
- [13] M. Sellier and S. Panda, "Beating capillarity in thin film flows," *International Journal for Numerical Methods in Fluids*, vol. 63, pp. 431–448, 2009.
- [14] C. Heining, M. Sellier, and N. Aksel, "The inverse problem in creeping film flows," *Acta Mechanica*, vol. 223, no. 4, pp. 841–847, 2012.
- [15] A. Gessese, K. M. Wa, and M. Sellier, "Bathymetry reconstruction based on the zero-inertia shallow water approximation," *Theoretical and Computational Fluid Dynamics*. In press.
- [16] W. Castaigns, D. Dartus, M. Honnorat, F.-X. Le Dimet, Y. Loukili, J. Monnier et al., "Automatic differentiation: a tool for variational data assimilation and adjoint sensitivity analysis for flood modelling," *Lecture Notes in Computational Science and Engineering*, vol. 50, pp. 249–262, 2006.

- [17] M. Honnorat, J. Monnier, and F.-X. Le Dimet, "Lagrangian data assimilation for river hydraulics simulations," *Computing and Visualization in Science*, vol. 12, no. 5, pp. 235–246, 2009.
- [18] M. H. Chaudhry, *Open-Channel Flow*, Springer, New York, NY, USA, 2nd edition, 2007.
- [19] X. Ying, A. A. Khan, and S. S. Y. Wang, "Upwind conservative scheme for the Saint Venant equations," *Journal of Hydraulic Engineering*, vol. 130, no. 10, pp. 977–987, 2004.
- [20] Q. Liang and A. G. L. Borthwick, "Adaptive quadtree simulation of shallow flows with wet-dry fronts over complex topography," *Computers & Fluids*, vol. 38, no. 2, pp. 221–234, 2009.
- [21] Y. Xing and C.-W. Shu, "High order finite difference WENO schemes with the exact conservation property for the shallow water equations," *Journal of Computational Physics*, vol. 208, no. 1, pp. 206–227, 2005.
- [22] F. M. Henderson, *Open Channel Flow*, Macmillan, New York, NY, USA, 1966.

Research Article

Exponential Convergence of an Observer Based on Partial Field Measurements for the Wave Equation

**Dominique Chapelle, Nicolae Cîndea,
Maya De Buhan, and Philippe Moireau**

MACS Project Team, Inria, France

Correspondence should be addressed to Philippe Moireau, philippe.moireau@inria.fr

Received 27 July 2012; Accepted 2 October 2012

Academic Editor: Valery Yakhno

Copyright © 2012 Dominique Chapelle et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We analyze an observer strategy based on partial—that is, in a subdomain—measurements of the solution of a wave equation, in order to compensate for uncertain initial conditions. We prove the exponential convergence of this observer under a nonstandard observability condition, whereas using measurements of the time derivative of the solution would lead to a standard observability condition arising in stabilization and exact controllability. Nevertheless, we directly relate our specific observability condition to the classical geometric control condition. Finally, we provide some numerical illustrations of the effectiveness of the approach.

1. Introduction

Observer theory has been established for decades [1], but only much more recently has it been considered for systems governed by evolution partial differential equations (PDEs), see [2–4] and references therein. In this realm challenges abound, as in particular observer convergence can no longer be mathematically analyzed solely by investigating the poles of the observer dynamics operator. Furthermore, observers are only meaningful when they provide an actual *computable* estimate of the system state—and possibly also of some uncertain parameters, see [2]—which implies that we must consider *discretization* issues both in their design and in their analysis, and this in turn introduces additional serious difficulties, as is already well known in the field of stabilization of PDE-governed systems, see for example [5].

In this paper we consider an observer strategy originally proposed in [6] under the name “Schur Displacement Feedback” (SDF) for elasticity-like formulations, and directly adapted here to the scalar wave equation. A major advantage of this observer is that it exploits measurements of the primary variable rather than of the time derivative of this

variable. Such direct measurements are more easily obtained in practice, in general, hence no detrimental time differentiation of the data is required in this case. When analyzing the estimation error, we find—as always with Luenberger observers—a closed-loop stabilized dynamics, albeit here of a rather uncommon—nonphysical—type. Nevertheless, we will show that exponential convergence is achieved under an observability condition closely related to that associated with more standard stabilization strategies, a condition which is itself equivalent to the classical geometric control condition [7].

2. Observer Design

Let Ω be a bounded domain of \mathbb{R}^n with a regular boundary $\partial\Omega$. Given a known source term $f \in L^1(\mathbb{R}^+, L^2(\Omega))$ with \dot{f} —namely, the time derivative of f —also in $L^1(\mathbb{R}^+, L^2(\Omega))$, we consider the following wave equation:

$$\begin{aligned} \ddot{u}(x, t) - \Delta u(x, t) &= f(x, t), & (x, t) \in \Omega \times (0, \infty), \\ u(x, t) &= 0, & (x, t) \in \partial\Omega \times (0, \infty), \\ u(x, 0) &= u_0(x), \quad \dot{u}(x, 0) = v_0(x), & x \in \Omega. \end{aligned} \tag{2.1}$$

Denoting $x(t) = \begin{pmatrix} u(t) \\ \dot{u}(t) \end{pmatrix}$, we can rewrite (2.1) as the first-order system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + R, & t > 0, \\ x(0) &= x_0, \end{aligned} \tag{2.2}$$

where $x_0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}$ and $A : \mathfrak{D}(A) \rightarrow \mathcal{X}$ with

$$\mathfrak{D}(A) = \mathfrak{D}(-\Delta) \times \mathfrak{D}(-\Delta^{1/2}), \quad \mathcal{X} = \mathfrak{D}(-\Delta^{1/2}) \times \mathcal{L}, \quad A = \begin{pmatrix} 0 & I \\ \Delta & 0 \end{pmatrix}, \quad R = \begin{pmatrix} 0 \\ f \end{pmatrix}, \tag{2.3}$$

for $\mathcal{L} = L^2(\Omega)$ here. Note that with the boundary conditions considered in (2.1), we have $\mathfrak{D}(-\Delta^{1/2}) = H_0^1(\Omega)$ and $\mathfrak{D}(-\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$. The operator A generates a group and for all $x_0 \in \mathcal{X}$ then $(u, \dot{u}) \in C((0, T); \mathcal{X})$ for any $T > 0$, see for example [8].

For this system, we consider that some measurements are available—assumed to be without noise in this paper—in $\omega \subset \Omega$ an open and nonempty subset of Ω and at every time t . These measurements can be either of the form

$$\forall t \geq 0, \quad \left| \begin{array}{l} H_0^1(\Omega) \quad \longrightarrow H^1(\omega) \\ u(\cdot, t) \quad \longmapsto z(t) = u(\cdot, t)|_\omega \end{array} \right. \tag{2.4}$$

or alternatively

$$\forall t \geq 0, \quad \left| \begin{array}{l} L^2(\Omega) \quad \longrightarrow L^2(\omega) \\ \dot{u}(\cdot, t) \quad \longmapsto z(t) = \dot{u}(\cdot, t)|_\omega. \end{array} \right. \tag{2.5}$$

Our paper will mainly concentrate on the first type of observation, but we will consider the second type as a matter of comparison. In each case, accordingly introducing the observation space $\mathcal{Z} = H^1(\omega)$ or $\mathcal{Z} = L^2(\omega)$, we can define an observation operator $H \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$ by

$$H = (H_0 \ 0), \quad \text{or} \quad H = (0 \ H_0), \quad (2.6)$$

respectively, with H_0 the restriction operator on ω .

The aim in observer design is to define a system \hat{x} , with modified dynamics compared to that followed by the original system x and based on using the discrepancy $z - H\hat{x}$, so that for any initial condition \hat{x}_0 , the state system \hat{x} tends to x in time. The dynamics of the observer \hat{x} read

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + R + G(z(t) - H\hat{x}(t)), \quad t > 0, \\ \hat{x}(0) &= \hat{x}_0, \end{aligned} \quad (2.7)$$

with G a gain operator to be defined in $\mathcal{L}(\mathcal{Z}, \mathcal{X})$ in our case.

In order to assess the efficiency of the observer, we can consider the dynamics followed by the error $\tilde{x} = x - \hat{x}$, namely

$$\begin{aligned} \dot{\tilde{x}} &= (A - GH)\tilde{x}(t), \quad t > 0, \\ \tilde{x}(0) &= x_0 - \hat{x}_0, \end{aligned} \quad (2.8)$$

and then G should stabilize \tilde{x} to 0. This type of approach can be categorized as a Luenberger observer [1]—also sometimes referred to as “nudging” [3]—applied on an evolution PDE.

In the case of time-derivative measurements (2.5), we can choose $G = \gamma \begin{pmatrix} 0 \\ \mathbb{1}_\omega \end{pmatrix}$, for any $\gamma > 0$. Then the observer system derived from a *Direct Velocity Feedback (DVF)* gives

$$\begin{aligned} \ddot{\hat{u}}(x, t) - \Delta \hat{u}(x, t) &= f + \gamma \mathbb{1}_\omega(x)(z(x, t) - \dot{\hat{u}}(x, t)), \quad (x, t) \in \Omega \times (0, \infty), \\ \hat{u}(x, t) &= 0, \quad (x, t) \in \partial\Omega \times (0, \infty), \\ \hat{u}(x, 0) &= \hat{u}_0(x), \quad \dot{\hat{u}}(x, 0) = \hat{v}_0(x), \quad x \in \Omega, \end{aligned} \quad (2.9)$$

and the error follows the dynamics

$$\begin{aligned} \ddot{\tilde{u}}(x, t) - \Delta \tilde{u}(x, t) + \gamma \mathbb{1}_\omega(x) \dot{\tilde{u}}(x, t) &= 0, \quad (x, t) \in \Omega \times (0, \infty), \\ \tilde{u}(x, t) &= 0, \quad (x, t) \in \partial\Omega \times (0, \infty), \\ \tilde{u}(x, 0) &= \tilde{u}_0(x), \quad \dot{\tilde{u}}(x, 0) = \tilde{v}_0(x), \quad x \in \Omega. \end{aligned} \quad (2.10)$$

This system is well known to be exponentially stable—see [7, 9] and references therein—if and only if the observability condition

$$\begin{aligned} \exists(T_0, C), \quad \forall T \geq T_0, \quad \forall(u_0, v_0) \in H_0^1(\Omega) \times L^2(\Omega), \\ \int_0^T \int_{\omega} |\dot{u}(x, t)|^2 d\Omega dt \geq C \left(\|u_0\|_{H^1(\Omega)}^2 + \|v_0\|_{L^2(\Omega)}^2 \right), \end{aligned} \quad (2.11)$$

is satisfied for arbitrary solutions of the system (2.1) with $f = 0$, which is equivalent to the geometric control condition (GCC) of [7].

In the first case of direct measurements of the field (2.4), we can also propose an observer without any time differentiation involved—typically to avoid amplifying the measurement errors which always arise in practice. To that purpose, we define the operator

$$\mathcal{L}_{\omega} : H^1(\omega) \longrightarrow H_0^1(\Omega), \quad \mathcal{L}_{\omega}\phi = \psi, \quad (2.12)$$

where ψ is the solution of the following elliptic equation:

$$\begin{aligned} \Delta\psi &= 0, \quad \text{in } \Omega \setminus \omega, \\ \psi &= 0, \quad \text{on } \partial\Omega, \\ \psi &= \phi, \quad \text{in } \bar{\omega}, \end{aligned} \quad (2.13)$$

namely, \mathcal{L}_{ω} is a harmonic lifting operator. We easily verify that \mathcal{L}_{ω} is bounded from $H^1(\omega)$ to $H_0^1(\Omega)$ using the trace and lifting properties to write

$$\|\mathcal{L}_{\omega}\phi\|_{H_0^1(\Omega)}^2 \leq \|\nabla\phi\|_{L^2(\omega)}^2 + C_1 \|\phi\|_{\partial\omega}^2 \leq C_2 \|\phi\|_{H^1(\omega)}^2. \quad (2.14)$$

As a candidate observer, we will consider $G = \gamma \binom{\mathcal{L}_{\omega}}{0}$ which gives the following first-order system:

$$\begin{aligned} \dot{\hat{u}}(x, t) &= \hat{v}(x, t) + \gamma \mathcal{L}_{\omega}(z(t) - \mathbb{1}_{\omega}(x)\hat{u}(x, t)|_{\omega}), \quad (x, t) \in \Omega \times (0, \infty), \\ \dot{\hat{v}}(x, t) - \Delta\hat{u}(x, t) &= 0, \quad (x, t) \in \Omega \times (0, \infty), \\ \hat{u}(x, 0) &= \hat{u}_0(x), \quad \hat{v}(x, 0) = \hat{v}_0(x), \quad x \in \Omega \end{aligned} \quad (2.15)$$

with $\gamma > 0$ a gain parameter. Note that this system amounts to a modification of the wave equation written in first-order form, with a correction term based on the discrepancy between the measurement and the observer primary variable. This strategy is the direct adaptation to the wave equation of the “Schur Displacement Feedback” (SDF) filtering methodology originally proposed in [6] for elasticity-like formulations.

Remark 2.1. The above choices of gain operator G fulfill the identity $G = \gamma H^*$ where the $*$ symbol denotes the adjoint via Riesz representation. This is obvious in the case of (2.5), whereas we will provide the detailed proof with Proposition 3.1 below for (2.4).

Remark 2.2. We consider in this paper a known source term, but the observers considered here can be extended to also estimate an unknown source term, following the strategy introduced in [2].

3. Exponential Convergence of the Observer System for Field Measurements

In order to establish that (2.15) is an adequate observer for (2.1), we should study the decay of the error satisfying for (2.15) the following system:

$$\begin{aligned}\dot{\tilde{u}}(x, t) &= \tilde{v}(x, t) - \gamma \mathcal{L}_\omega(\mathbb{1}_\omega(x)\tilde{u}(x, t)), & (x, t) \in \Omega \times (0, \infty), \\ \tilde{v}(x, t) - \Delta \tilde{u}(x, t) &= 0, & (x, t) \in \Omega \times (0, \infty), \\ \tilde{u}(x, t) &= 0, & (x, t) \in \partial\Omega \times (0, \infty), \\ \tilde{u}(x, 0) &= \tilde{u}_0(x), \quad \tilde{v}(x, 0) = \tilde{v}_0(x), & x \in \Omega\end{aligned}\tag{3.1}$$

and we will prove the exponential decay of the associated energy, namely,

$$\tilde{E}(t) = \frac{1}{2} \left(\int_{\Omega} |\nabla \tilde{u}(x, t)|^2 d\Omega + \int_{\Omega} |\tilde{v}(x, t)|^2 d\Omega \right).\tag{3.2}$$

Proposition 3.1. *Assume that we have the observability condition*

$$\int_0^{T_0} \|u(\cdot, t)\|_{H^1(\omega)}^2 dt \geq C \left(\|u_0\|_{H^1(\Omega)}^2 + \|v_0\|_{L^2(\Omega)}^2 \right), \quad \forall (u_0, v_0) \in H_0^1(\Omega) \times L^2(\Omega),\tag{3.3}$$

for some $T_0 > 0$, for arbitrary solutions of (2.1) with $f = 0$. Then, there exist strictly positive constants M and μ such that

$$\tilde{E}(t) \leq M e^{-\mu t} \tilde{E}(0), \quad \forall t > 0.\tag{3.4}$$

Proof. Let us consider the space H_ω^1 defined by $H^1(\omega)$ equipped with the norm

$$\|\phi\|_{H_\omega^1} = \|\nabla(\mathcal{L}_\omega\phi)\|_{L^2(\Omega)}. \quad (3.5)$$

It is straightforward to see that this norm is equivalent to the usual norm of $H^1(\omega)$, using first (2.14), and secondly

$$\begin{aligned} \|\phi\|_{H^1(\omega)}^2 &= \|\nabla\phi\|_{L^2(\omega)}^2 + \|\phi\|_{L^2(\omega)}^2 \\ &\leq \|\nabla(\mathcal{L}_\omega\phi)\|_{L^2(\Omega)}^2 + \|\mathcal{L}_\omega\phi\|_{L^2(\Omega)}^2 \leq (1+C)\|\nabla(\mathcal{L}_\omega\phi)\|_{L^2(\Omega)}^2, \end{aligned} \quad (3.6)$$

with C given by the Poincaré inequality. Considering now \mathcal{L}_ω as a bounded operator from H_ω^1 to $H_0^1(\Omega)$, we want to identify \mathcal{L}_ω^* . Using the orthogonality property

$$\forall \psi \in H_0^1(\Omega) \text{ s.t. } \psi|_\omega = 0, \quad \langle \nabla(\mathcal{L}_\omega\phi), \nabla\psi \rangle_{L^2(\Omega)} = 0, \quad (3.7)$$

directly inferred from the first equation of (2.13), we have that for all $\psi \in H_0^1(\Omega)$ and $\phi \in H_\omega^1$

$$\langle \nabla(\mathcal{L}_\omega\phi), \nabla\psi \rangle_{L^2(\Omega)} = \langle \nabla(\mathcal{L}_\omega\phi), \nabla(\mathcal{L}_\omega\psi|_\omega) \rangle_{L^2(\Omega)} = \langle \phi, \psi|_\omega \rangle_{H_\omega^1}, \quad (3.8)$$

showing that the adjoint via Riesz representation is then

$$\mathcal{L}_\omega^* : H_0^1(\Omega) \longrightarrow H_\omega^1, \quad \mathcal{L}_\omega^*\psi = \psi|_\omega. \quad (3.9)$$

Hence, we have proven that $G = \gamma H^*$, as claimed in Remark 2.1. Therefore, the error system (3.1) dynamics can be rewritten as

$$\dot{\tilde{x}} = (A - \gamma H^* H) \tilde{x}(t), \quad (3.10)$$

and the exponential decay of the energy (3.2) is then equivalent to the following observability inequality, see for example Theorem 2.3 in [9] with $B = H^*$,

$$\int_0^{T_0} \|Hx\|_{H_\omega^1}^2 dt \geq C\|x_0\|_{\mathcal{X}}^2, \quad \forall x_0 \in \mathcal{X}, \quad (3.11)$$

where x is the solution of the original wave equation in first-order form (2.2) with $f = 0$, and for some strictly positive constants T_0 and C , and of course this condition is directly equivalent to (3.3). \square

However, the observability condition (3.3) is somewhat nonstandard since we are using the H^1 -norm of the field in the observation space instead of the L^2 -norm of the time derivative. Nevertheless, the following proposition directly relates our observability condition with classical results.

Proposition 3.2. *Assume that the geometric control condition of Bardos et al. [7] is satisfied for some $\tilde{\omega}$ strict subset of ω such that $\text{dist}(\Omega \setminus \omega, \tilde{\omega}) > 0$, in the time interval $]0, T_0[$ for $T_0 > 0$. Then the observability condition (3.3) holds for the same time T_0 .*

Proof. Since the geometric control condition holds, we have the classical observability condition [7]

$$\int_0^{\tilde{T}} \int_{\tilde{\omega}} |\dot{u}(x, t)|^2 d\Omega dt \geq C \left(\|u_0\|_{H^1(\Omega)}^2 + \|v_0\|_{L^2(\Omega)}^2 \right), \quad \forall (u_0, v_0) \in H_0^1(\Omega) \times L^2(\Omega), \quad (3.12)$$

for some time $\tilde{T} = T_0 - \delta$ with δ sufficiently small, for arbitrary solutions of (2.1) with $f = 0$. We will show that this entails (3.3) by an argument inspired from [10]. Let $\psi \in C_c^\infty(\overline{\Omega})$ be a cutoff function satisfying

$$\psi(x) = \begin{cases} 0, & \text{if } x \in \Omega \setminus \omega \\ 1, & \text{if } x \in \tilde{\omega} \end{cases} \quad (3.13)$$

and $0 \leq \psi(x) \leq 1$ for every $x \in \overline{\Omega}$. Denote also $\phi(t) = t^2(\tilde{T} - t)^2$. Then, by repeated integrations by parts, we obtain

$$\begin{aligned} 0 &= \int_0^{\tilde{T}} \int_{\omega} \phi \psi (\ddot{u} - \Delta u) u \, d\Omega \, dt \\ &= \int_0^{\tilde{T}} \int_{\omega} \ddot{\phi} \psi \frac{|u|^2}{2} \, d\Omega \, dt - \int_0^{\tilde{T}} \int_{\omega} \phi \psi |\dot{u}|^2 \, d\Omega \, dt + \int_0^{\tilde{T}} \int_{\partial\omega} \phi \frac{\partial \psi}{\partial n} \frac{|u|^2}{2} \, d\Gamma \, dt \\ &\quad - \int_0^{\tilde{T}} \int_{\omega} \phi \Delta \psi \frac{|u|^2}{2} \, d\Omega \, dt + \int_0^{\tilde{T}} \int_{\omega} \phi \psi |\nabla u|^2 \, d\Omega \, dt. \end{aligned} \quad (3.14)$$

The definition of ψ entails $(\partial\psi/\partial n)|_{\partial\omega} = 0$, hence

$$\int_0^{\tilde{T}} \int_{\omega} \phi \psi |\dot{u}|^2 \, d\Omega \, dt = \int_0^{\tilde{T}} \int_{\omega} \phi \psi |\nabla u|^2 \, d\Omega \, dt + \int_0^{\tilde{T}} \int_{\omega} \ddot{\phi} \psi \frac{|u|^2}{2} \, d\Omega \, dt - \int_0^{\tilde{T}} \int_{\omega} \phi \Delta \psi \frac{|u|^2}{2} \, d\Omega \, dt. \quad (3.15)$$

This identity combined with the properties of the cutoff functions ϕ and ψ provides, for any strictly positive ε , the existence of a constant $C > 0$ such that

$$\int_{\varepsilon}^{T-\varepsilon} \int_{\tilde{\omega}} |\dot{u}|^2 d\Omega dt \leq C \left(\int_0^T \int_{\omega} |\nabla u|^2 d\Omega dt + \int_0^T \int_{\omega} |u|^2 d\Omega dt \right). \quad (3.16)$$

Substituting $\tilde{T} + 2\varepsilon$ for \tilde{T} in all the above computations gives

$$\int_{\varepsilon}^{\tilde{T}+\varepsilon} \int_{\tilde{\omega}} |\dot{u}|^2 d\Omega dt \leq C \int_0^{\tilde{T}+2\varepsilon} \|u(\cdot, t)\|_{H^1(\omega)}^2 dt. \quad (3.17)$$

We proceed by making the change of variable $\tau = t - \varepsilon$ in the left-hand side integral, yielding

$$\int_0^{\tilde{T}} \int_{\tilde{\omega}} |\dot{u}(x, \tau + \varepsilon)|^2 d\Omega d\tau \leq C \int_0^{\tilde{T}+2\varepsilon} \|u\|_{H^1(\omega)}^2 dt. \quad (3.18)$$

Noting that $u(x, t + \varepsilon)$ satisfies the wave equation with initial data $(u(x, \varepsilon), \dot{u}(x, \varepsilon))$ and applying (3.12) with this shifted solution, we obtain that there exists also C such that

$$\int_0^{\tilde{T}} \int_{\tilde{\omega}} |\dot{u}(x, t + \varepsilon)|^2 d\Omega dt \geq C \left(\|u(\varepsilon)\|_{H^1(\Omega)}^2 + \|\dot{u}(\varepsilon)\|_{L^2(\Omega)}^2 \right). \quad (3.19)$$

Combining (3.18), (3.19), and the the fact that the energy of the solution of the wave equation is exactly conserved over time, we have our observability inequality (3.3) upon choosing $\varepsilon = \delta/2$. \square

Remark 3.3. Note that this proof requires a geometric control condition slightly stronger than that directly associated with the measurement domain ω , since $\tilde{\omega}$ is a strict subset. However, it is likely that our desired observability condition (2.11) could also be established directly by adapting a microlocal analysis strategy used for the standard observability condition (3.3), for example, using defect measures [11]. In our case, the defect measure of interest would be in H^1 as considered more recently in [12]. This H^1 defect measure should be proven to be made of the same bicharacteristic set as the L^2 defect measure introduced in [13] for the classical GCC. In this context we would have the new observability condition with the same subdomain ω . Nevertheless, the advantage of our approach lies in its compactness, and moreover it demonstrates an interesting property of equirepartition (over time) of the total energy localized within the subdomain ω between the kinetic and potential contributions.

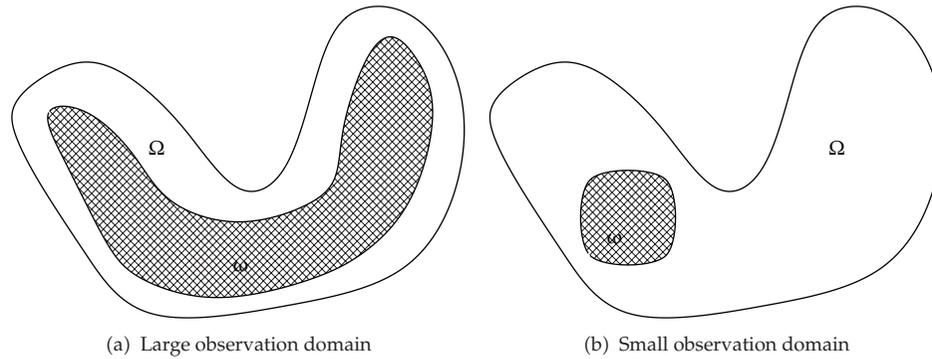


Figure 1: Geometry and two observation domains considered.

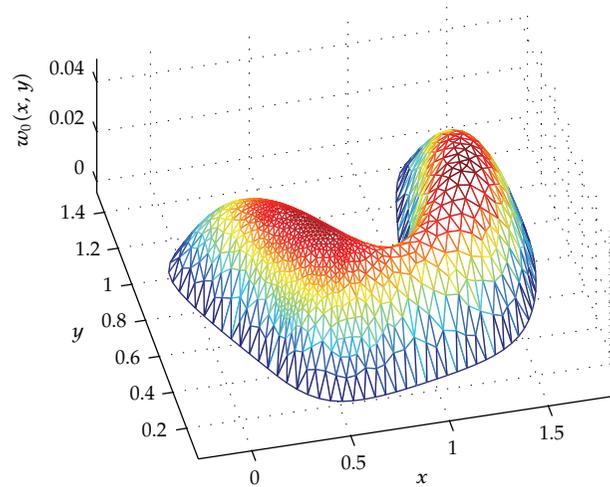


Figure 2: Initial condition in mesh used.

Remark 3.4. For both observer strategies, the potential noise in the measurements—disregarded in this paper—would simply entail an additional source term in the error equations, without amplification of this error term by time or space differentiation. The parameter γ should then be chosen to obtain the best decay rate without undue amplification of the noise [2, 6].

4. Numerical Illustrations

In order to illustrate the effectiveness of our Schur Displacement Feedback (SDF) observer approach—and compare it to the classical Direct Velocity Feedback (DVF) observer—we consider the two-dimensional domain shown in Figure 1 and, with the two choices of observation domains also displayed in the Figure. In order for the mesh to be conforming with the observation domain, we use one specific triangular mesh for each observation choice, with very similar mesh sizes—corresponding to about 950 vertices—for both cases, one of which being shown in Figure 2 with the initial condition considered in the simulations.

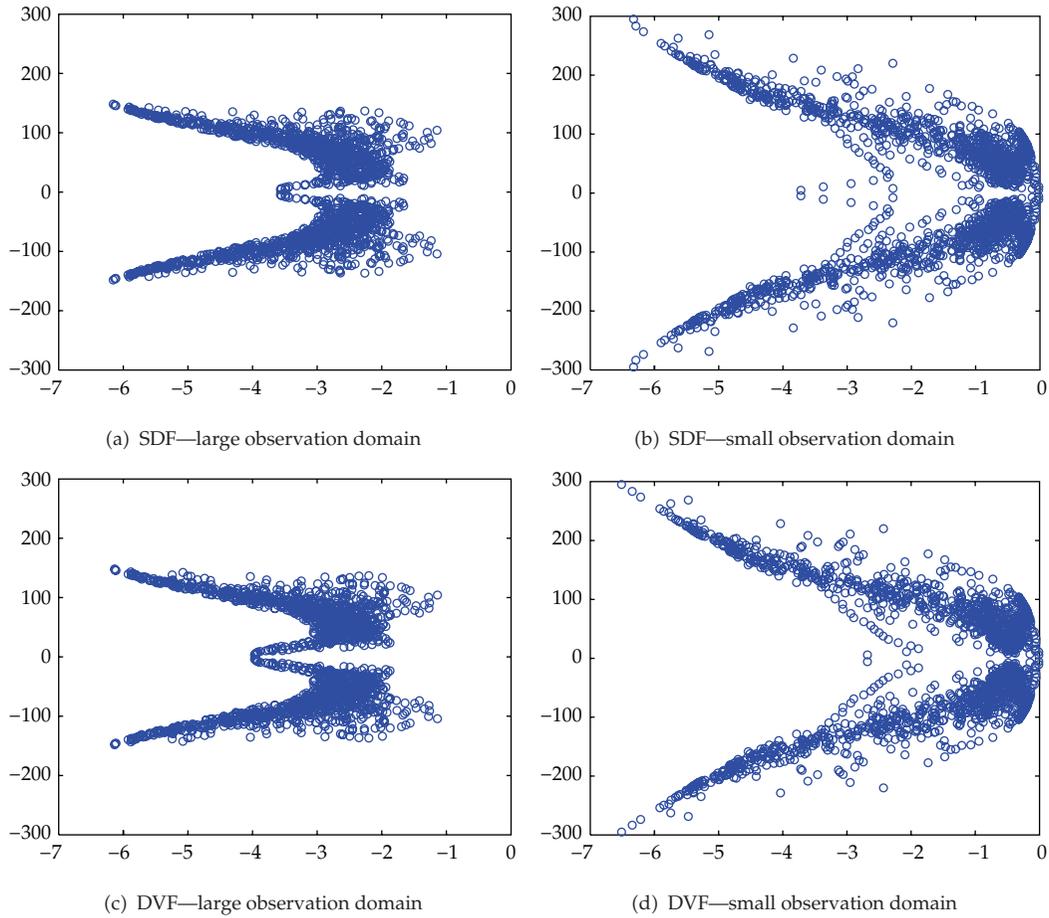


Figure 3: Poles for SDF and DVF stabilization (up and down, resp.) with large and small observation domains (left and right, resp.).

The variational form of the wave equation is then discretized using P_1 -Lagrange finite elements in space and a mid-point Newmark type time-discretization scheme [6]. As discretization issues in relation to observation and stabilization properties are not the topic of this paper—we instead refer to [6, 14] for detailed discretization considerations and analysis—we use sufficiently small values of the time step to obtain “converged” solutions in time, and we focus on the semidiscrete equation—namely, discretized space-wise only—when computing poles. Furthermore, as it is well known that spatial discretization itself induces undesirable numerical artifacts in the stabilization and control of our type of partial differential equation, with some numerical “high-frequency” poles featuring vanishing dissipation, we resort to incorporating a numerical viscosity term as advocated in [14, 15] to circumvent this difficulty.

We show in Figures 3(a) and 3(b) the numerical poles of the SDF-stabilized equation, for the two observation domains considered ($\gamma = 8$). We can see that the stabilization strategy is very effective for the large observation domain, as all real parts lie below -1 , and below -2 or so when excluding high-frequency poles. As a matter of comparison,

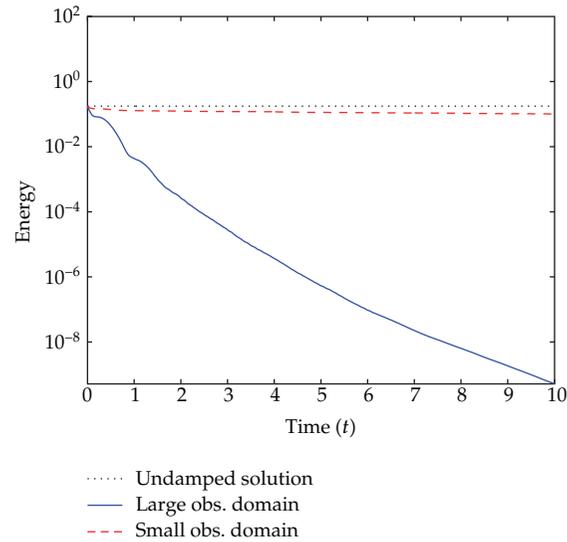


Figure 4: Energies of the undamped and SDF-stabilized solutions.

the imaginary part of the first undamped eigenmode is about 5.6. Of course, the effectiveness is significantly reduced in the case of the second observation domain, which does not satisfy the above geometric control condition, indeed. We also compare in Figures 3(c) and 3(d) the corresponding poles obtained with the DVF strategy ($\gamma = 8$), and we observe that the two approaches provide similar stabilization properties.

This is further illustrated in Figure 4, where we compare the energy norms of the SDF-stabilized numerical solutions in time with that of the wave equation itself—note that the time discretization scheme used is energy preserving. This confirms the excellent stabilization properties of the SDF approach, as expected provided that the controllability (or equivalent observability) condition is fulfilled. Of course, this directly translates into equivalent observation convergence properties following our above discussion.

References

- [1] D. G. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.
- [2] P. Moireau, D. Chapelle, and P. Le Tallec, "Joint state and parameter estimation for distributed mechanical systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 6-8, pp. 659–677, 2008.
- [3] D. Auroux and J. Blum, "A nudging-based data assimilation method: the back and forth nudging (BFN) algorithm," *Nonlinear Processes Geophysics*, vol. 15, no. 2, pp. 305–319, 2008.
- [4] K. Ramdani, M. Tucsnak, and G. Weiss, "Recovering the initial state of an infinite-dimensional system using observers," *Automatica*, vol. 46, no. 10, pp. 1616–1625, 2010.
- [5] L. R. Tcheugoué Tébou and E. Zuazua, "Uniform exponential long time decay for the space semi-discretization of a locally damped wave equation via an artificial numerical viscosity," *Numerische Mathematik*, vol. 95, no. 3, pp. 563–598, 2003.
- [6] P. Moireau, D. Chapelle, and P. Le Tallec, "Filtering for distributed mechanical systems using position measurements: perspectives in medical imaging," *Inverse Problems*, vol. 25, no. 3, Article ID 035010, 25 pages, 2009.

- [7] C. Bardos, G. Lebeau, and J. Rauch, "Un exemple d'utilisation des notions de propagation pour le contrôle et la stabilisation des problèmes hyperboliques," *Rendiconti del Seminario Matematico del Università Politecnico Torino, Fascicolo Speciale (Hyperbolic Equations)*, pp. 12–31, 1988.
- [8] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, vol. 44 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1983.
- [9] K. Liu, "Locally distributed control and damping for the conservative systems," *SIAM Journal on Control and Optimization*, vol. 35, no. 5, pp. 1574–1590, 1997.
- [10] X. Zhang, "Explicit observability estimate for the wave equation with potential and its application," *Proceedings: Mathematical, Physical and Engineering Sciences*, vol. 456, no. 1997, pp. 1101–1115, 2000.
- [11] P. Gérard, "Microlocal defect measures," *Communications in Partial Differential Equations*, vol. 16, no. 11, pp. 1761–1794, 1991.
- [12] N. Burq and G. Lebeau, "Mesures de défaut de compacité, application au système de lamé," *Annales Scientifiques de l'École Normale Supérieure*, vol. 34, no. 6, pp. 817–870, 2001.
- [13] N. Burq and P. Gérard, *Contrôle optimal des équations aux dérivées partielles*, Cours de l'École Polytechnique, 2002.
- [14] D. Chapelle, N. Cîrdea, and P. Moireau, "Improving convergence in numerical analysis using observers—the wave-like equation case," *Mathematical Models and Methods in Applied Sciences*. In press.
- [15] S. Ervedoza and E. Zuazua, "Uniformly exponentially stable approximations for a class of damped systems," *Journal de Mathématiques Pures et Appliquées*, vol. 91, no. 1, pp. 20–48, 2009.

Research Article

Design of an Annular Disc Subject to Thermomechanical Loading

Sergei Alexandrov,^{1,2} Elena Lyamina,¹ and Yeau-Ren Jeng²

¹ *A.Yu. Ishlinskii Institute for Problems in Mechanics, Russian Academy of Sciences,
101-1 Prospect Vernadskogo, Moscow 119526, Russia*

² *Department of Mechanical Engineering and Advanced Institute for Manufacturing with High-tech
Innovations, National Chung Cheng University, 168 University Road, Chia-Yi 62102, Taiwan*

Correspondence should be addressed to Yeau-Ren Jeng, imeyrj@ccu.edu.tw

Received 27 July 2012; Accepted 3 September 2012

Academic Editor: Valery Yakhno

Copyright © 2012 Sergei Alexandrov et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Two solutions to design a thin annular disc of variable thickness subject to thermomechanical loading are proposed. It is assumed that the thickness of the disc is everywhere sufficiently small for the stresses to be averaged through the thickness. The state of stress is plane. The initiation of plastic yielding is controlled by Mises yield criterion. The design criterion for one of the solutions proposed requires that the distribution of stresses is uniform over the entire disc. In this case there is a relation between optimal values of the loading parameters at the final stage. The specific shape of the disc corresponds to each pair of such parameters. The other solution is obtained under the additional requirement that the distribution of strains is uniform. This solution exists for the disc of constant thickness at specific values of the loading parameters.

1. Introduction

Thin annular discs subject to various loading conditions are a class of commonly used structures in mechanical engineering. Therefore, there is a vast amount of literature on this topic. These studies can be conveniently divided into two groups, namely, analysis and design. The present paper deals with the design of thin discs. Therefore, previous works solely related to the analysis of discs are not considered here. Reviews of some results on the design of thin discs are provided in [1, 2]. An analytical solution to design an annular disc of variable thickness under internal and external pressures has been proposed in [1]. It has been assumed that the initiation of plastic yielding is controlled by Tresca yield criterion. The goal of the design has been to find conditions under which the yield criterion is simultaneously satisfied at all points of the disc. A disc rotating in a temperature field has been considered

in [2]. The objective function in the problem is the disc weight. A numerical method has been used to solve the problem. The present paper concerns with an annular disc inserted into a rigid container and subject to thermomechanical loading. One of the loading parameters is uniform temperature and the other is pressure over the inner radius of the disc. Both vary with the time. Two design criteria are adopted. The requirement of one of the design criteria is that the distribution of stresses is uniform over the entire disc at the final stage. Using this criterion, a relation between optimal loading parameters is obtained. The solution is given in analytical form. The additional requirement of the other design criterion is that the state of strain is uniform. The same requirement is adopted to design Michell structures (see [3]). It is shown that in the case of the disc under consideration there is the unique solution for the second design problem. In particular, the thickness of the disc is constant.

2. Statement of the Problem

Consider a thin annular disc of outer radius b and inner radius a inserted into a rigid container of radius a . It is convenient to introduce a cylindrical coordinate system (r, θ, z) with its z -axis coinciding with the axis of symmetry of the disc. The initial thickness of the disc, h , is a function of r . The disc is subject to thermal loading by a uniform temperature field varying with the time. The disc has no stress at the initial temperature. Uniform pressure varying with the time is applied over the inner radius of the disc. The outer radius is fixed to the container. It is evident that the problem is axisymmetric. In particular, the solution is independent of θ . Moreover, the normal stresses in the cylindrical coordinates, σ_r , σ_θ , and σ_z are the principal stresses. It is also assumed that the state of stress is two-dimensional, $\sigma_z = 0$. The pressure applied, thermal expansion caused by a rise of temperature, and the constraints imposed on the disc affect the initial zero-stress state. It is assumed that the rise of temperature above the reference state, T , and the pressure over the inner radius are monotonically nondecreasing functions of the time, t . The boundary conditions are

$$\sigma_r = -\beta\sigma_0, \quad (2.1)$$

at $r = a$ and

$$u = 0, \quad (2.2)$$

at $r = b$. Here u is the radial displacement, β is a function of the time, and σ_0 is a constant introduced for further convenience. The circumferential displacement vanishes everywhere.

It is assumed that the thickness of the disc is everywhere sufficiently small for the stresses to be averaged through the thickness. In this case the only nontrivial equilibrium equation becomes

$$\frac{\partial}{\partial r}(hr\sigma_r) = h\sigma_\theta. \quad (2.3)$$

The total radial, ε_r , and circumferential, ε_θ , strains are defined by

$$\varepsilon_r = \varepsilon_r^T + \varepsilon_r^e + \varepsilon_r^p, \quad \varepsilon_\theta = \varepsilon_\theta^T + \varepsilon_\theta^e + \varepsilon_\theta^p, \quad (2.4)$$

where the superscript T denotes the thermal portions of the total strains, the superscript e the elastic portions of the total strains, and the superscript p the plastic portions of the total strains. It follows from Hooke's law that

$$\varepsilon_r^e = \frac{\sigma_r - \nu\sigma_\theta}{E}, \quad \varepsilon_\theta^e = \frac{\sigma_\theta - \nu\sigma_r}{E}, \quad (2.5)$$

where E is Young's modulus and ν is Poisson's ratio. The thermal portions of the total strains are given by

$$\varepsilon_r^T = \varepsilon_\theta^T = \alpha T, \quad (2.6)$$

where α is the thermal coefficient of linear expansion. In the plastic range, Mises yield criterion is adopted. For the problem under consideration this criterion reduces to

$$\sigma_r^2 + \sigma_\theta^2 - \sigma_\theta\sigma_r = \sigma_0^2, \quad (2.7)$$

where σ_0 is the yield stress in tension, a material constant for perfectly plastic materials. This quantity is also involved in (2.1). The associated flow rule is written in terms of the strain rate components. A consequence of this rule is

$$\frac{\dot{\varepsilon}_r^p}{\dot{\varepsilon}_\theta^p} = \frac{2\sigma_r - \sigma_\theta}{2\sigma_\theta - \sigma_r}, \quad (2.8)$$

where $\dot{\varepsilon}_r^p$ and $\dot{\varepsilon}_\theta^p$ are the plastic portions of the total radial and circumferential strain rates. Another essential equation following from the associated flow rule expresses plastic incompressibility, $\dot{\varepsilon}_r^p + \dot{\varepsilon}_\theta^p + \dot{\varepsilon}_z^p = 0$, where $\dot{\varepsilon}_z^p$ is the plastic portion of the total axial strain rate. This equation serves to determine $\dot{\varepsilon}_z^p$ and is not important for the present solution. At small strains,

$$\dot{\varepsilon}_r^p = \frac{\partial \varepsilon_r^p}{\partial t}, \quad \dot{\varepsilon}_\theta^p = \frac{\partial \varepsilon_\theta^p}{\partial t}. \quad (2.9)$$

According to the design criterion proposed by Michell (see [3]), all of the structural elements must be strained by exactly the same strain magnitude in either simple tension or pure compression. This criterion can be too restrictive for the structure under consideration. Therefore, in the present paper two design criteria are adopted. First, it is required that an equistressed state occurs in the entire disc. Then, the possibility to obtain a uniform distribution of strains is explored.

3. Restriction on Thickness Variation

The same magnitude of the elastic portion of strains can be obtained if and only if the distribution of stress components is uniform. Then, it follows from (2.3) that

$$\frac{\partial}{h\partial r}(hr) = n + 1, \quad (3.1)$$

where n is constant. Let h_0 be the thickness of the disc at $r = b$. Then, the solution of (3.1) satisfying this condition is

$$h = h_0 \left(\frac{r}{b} \right)^n. \quad (3.2)$$

Note that this function is often adopted in studies devoted to analysis of thin discs, for example [4–6]. The uniform distribution of stresses is only required in the final stage of loading. Using (3.2) the equation of equilibrium (2.3) for intermediate stages becomes

$$r \frac{\partial \sigma_r}{\partial r} + (1 + n)\sigma_r = \sigma_\theta. \quad (3.3)$$

4. Thermoelastic Solution

At the beginning of the process of loading the entire disc is elastic. At this stage,

$$\frac{\partial u}{\partial r} = \varepsilon_r^T + \varepsilon_r^e, \quad \frac{u}{r} = \varepsilon_\theta^T + \varepsilon_\theta^e. \quad (4.1)$$

Eliminating u between these two equations, using (2.5) and (2.6), and taking into account that T is independent of r yield

$$r \left(\frac{\partial \sigma_\theta}{\partial r} - \nu \frac{\partial \sigma_r}{\partial r} \right) + (1 + \nu)(\sigma_\theta - \sigma_r) = 0. \quad (4.2)$$

Eliminating the stress σ_θ in (4.2) by means of (3.3) gives

$$r^2 \frac{\partial^2 \sigma_r}{\partial r^2} + (3 + n)r \frac{\partial \sigma_r}{\partial r} + n(1 + \nu)\sigma_r = 0. \quad (4.3)$$

It is convenient to introduce the dimensionless radius ρ by $\rho = r/b$. Then, the general solution of (4.3) is

$$\frac{\sigma_r}{\sigma_0} = A\rho^{s_1} + B\rho^{s_2}, \quad \frac{\sigma_\theta}{\sigma_0} = A(1 + n + s_1)\rho^{s_1} + B(1 + n + s_2)\rho^{s_2} \quad (4.4)$$

where A and B are constants of integration and

$$\begin{aligned} s_1 &= -\left(1 + \frac{n}{2}\right) - \frac{1}{2}\sqrt{(2-n)^2 + 4n(1-\nu)}, \\ s_2 &= -\left(1 + \frac{n}{2}\right) + \frac{1}{2}\sqrt{(2-n)^2 + 4n(1-\nu)}. \end{aligned} \quad (4.5)$$

Substituting (4.4) into (2.5) determines ε_θ^e . Then, using this expression for ε_θ^e and (2.6) the radial displacement can be found from the equation $\varepsilon_\theta^e + \varepsilon_\theta^T = u/r$. As a result,

$$\frac{u}{rq} = A(1+n-\nu+s_1)\rho^{s_1} + B(1+n-\nu+s_2)\rho^{s_2} + \tau, \quad (4.6)$$

where $q = \sigma_0/E$ and $\tau = \alpha T/q$. Substituting the boundary conditions (2.1) and (2.2) into (4.4) and (4.6) leads to

$$\begin{aligned} A &= A^e, & A^e &= \frac{\beta(1+n-\nu+s_2) - \tau\omega^{s_2}}{(1+n-\nu+s_1)\omega^{s_2} - (1+n-\nu+s_2)\omega^{s_1}}, \\ B &= B^e, & B^e &= \frac{\tau\omega^{s_1} - \beta(1+n-\nu+s_1)}{(1+n-\nu+s_1)\omega^{s_2} - (1+n-\nu+s_2)\omega^{s_1}}, \end{aligned} \quad (4.7)$$

where $\omega = a/b$.

5. Thermoelastic-Plastic Solution for Design

The yield criterion (2.7) is satisfied by the following substitution:

$$\frac{\sigma_r}{\sigma_0} = -\frac{2 \sin \gamma}{\sqrt{3}}, \quad \frac{\sigma_\theta}{\sigma_0} = -\frac{(\sin \gamma + \sqrt{3} \cos \gamma)}{\sqrt{3}}, \quad (5.1)$$

where γ is a function of ρ and τ . Substituting (5.1) into (3.3) results in

$$\rho \frac{\partial \gamma}{\partial \rho} = \frac{\sqrt{3} \cos \gamma - (1+2n) \sin \gamma}{2 \cos \gamma}. \quad (5.2)$$

The zone where the yield criterion is satisfied should occupy the entire disc at the final stage. As it has been mentioned before, the design criterion chosen is satisfied if and only if the distribution of stress is uniform at this stage. Therefore, it should be uniform over the domain where (5.2) is valid. It follows from (5.1) that the condition that the distribution of the stresses in the plastic zone is uniform is equivalent to the condition that γ is independent of ρ . It is evident that the general solution of (5.2) does not satisfy this requirement. However, this equation has a special solution in the form $\gamma = \gamma_0$, where

$$\tan \gamma_0 = \frac{\sqrt{3}}{1+2n}. \quad (5.3)$$

It is seen from (2.1), (5.1) and (5.3) that this special solution takes place if and only if

$$\beta = \frac{2}{\sqrt{3}} \sin \gamma_0. \quad (5.4)$$

Let R be the dimensionless radius of the elastic/plastic boundary. The general solution (4.4) and (4.6) is valid in the elastic zone. However, A and B are not given by (4.7). The stresses σ_r and σ_θ are continuous across the elastic/plastic boundary. Therefore, it follows from (4.4) and (5.1) that

$$\begin{aligned} -\frac{2}{\sqrt{3}} \sin \gamma_0 &= AR^{s_1} + BR^{s_2}, \\ -\frac{(\sin \gamma_0 + \sqrt{3} \cos \gamma_0)}{\sqrt{3}} &= A(1+n+s_1)R^{s_1} + B(1+n+s_2)R^{s_2}. \end{aligned} \quad (5.5)$$

The boundary condition (2.2) combined with (4.6) gives

$$A(1+n-\nu+s_1) + B(1+n-\nu+s_2) + \tau = 0. \quad (5.6)$$

Solving (5.5) for AR^{s_1} and BR^{s_2} results in

$$\begin{aligned} AR^{s_1} = A_0 &= \frac{\sqrt{3} \cos \gamma_0 - (2n + 2s_2 + 1) \sin \gamma_0}{\sqrt{3}s}, \\ BR^{s_2} = B_0 &= \frac{(2n + 2s_1 + 1) \sin \gamma_0 - \sqrt{3} \cos \gamma_0}{\sqrt{3}s}. \end{aligned} \quad (5.7)$$

Thus, the quantities AR^{s_1} and BR^{s_2} are independent of τ . Eliminating A and B in (5.6) by means of (5.7) leads to

$$A_0 R^{-s_1} (1+n-\nu+s_1) + B_0 R^{-s_2} (1+n-\nu+s_2) + \tau = 0. \quad (5.8)$$

It is convenient to introduce the following quantities $\xi_r^p = \partial \varepsilon_r^p / \partial \tau$ and $\xi_\theta^p = \partial \varepsilon_\theta^p / \partial \tau$. Then, (2.8) becomes

$$\xi_r^p = \xi_\theta^p \left(\frac{2\sigma_r - \sigma_\theta}{2\sigma_\theta - \sigma_r} \right). \quad (5.9)$$

Since the stresses are constant in the plastic zone, the elastic strain rates vanish. Moreover, it follows from (2.6) that $\xi_r^T = \partial \varepsilon_r^T / \partial \tau = q$ and $\xi_\theta^T = \partial \varepsilon_\theta^T / \partial \tau = q$. Therefore, the total strain rates in the equation of compatibility can be replaced with their plastic portions. Then, using the definition for ξ_r^p and ξ_θ^p this equation is reduced to

$$\rho \frac{\partial \xi_\theta^p}{\partial \rho} + \xi_\theta^p - \xi_r^p = 0. \quad (5.10)$$

Substituting (5.1) at $\gamma = \gamma_0$ into (5.9) and then eliminating ξ_r^p in (5.10) yield

$$\rho \frac{\partial \xi_\theta^p}{\partial \rho} = -\frac{3n\xi_\theta^p}{1+2n}. \quad (5.11)$$

Here, (5.3) has been used to eliminate $\tan \gamma_0$. Equation (5.11) can be immediately integrated to give

$$\xi_\theta^p = q\xi_0(\tau)\rho^m, \quad m = -\frac{3n}{1+2n}, \quad (5.12)$$

where $\xi_0(\tau)$ is a function of integration. Introduce the notation $w = du/d\tau$. Note that w is proportional to the radial velocity. Using (5.12) the value of w on the plastic side of the elastic/plastic boundary is determined as

$$w_R^p = qR + q\xi_0(\tau)R^{m+1}. \quad (5.13)$$

Differentiating (4.6) with respect to τ yields

$$\frac{w}{rq} = \frac{dA}{d\tau}(1+n-\nu+s_1)\rho^{s_1} + \frac{dB}{d\tau}(1+n-\nu+s_2)\rho^{s_2} + 1. \quad (5.14)$$

Thus, the value of w on the elastic side of the elastic/plastic boundary is

$$w_R^e = \left[\frac{dA}{d\tau}(1+n-\nu+s_1)R^{s_1} + \frac{dB}{d\tau}(1+n-\nu+s_2)R^{s_2} + 1 \right] Rq. \quad (5.15)$$

Since $w_R^p = w_R^e$, it follows from (5.13) and (5.15) that

$$\xi_0(\tau) = \left[\frac{dA}{d\tau}(1+n-\nu+s_1)R^{s_1} + \frac{dB}{d\tau}(1+n-\nu+s_2)R^{s_2} \right] R^{-m}. \quad (5.16)$$

This equation can be rewritten in the following equivalent form:

$$\xi_0(\tau) = \left[\frac{dA}{dR}(1+n-\nu+s_1)R^{s_1} + \frac{dB}{dR}(1+n-\nu+s_2)R^{s_2} \right] \frac{dR}{d\tau} R^{-m}. \quad (5.17)$$

Eliminating here the derivatives dA/dR and dB/dR by means of (5.7) gives

$$\begin{aligned} \xi_0(\tau) = & - \left[s_1(1+n-\nu+s_1) \left[\sqrt{3} - (2n+2s_2+1) \tan \gamma_0 \right] \right. \\ & \left. + s_2(1+n-\nu+s_2) \left[(2n+2s_1+1) \tan \gamma_0 - \sqrt{3} \right] \right] \frac{dR}{d\tau} \frac{R^{-m-1}}{\sqrt{3}s} \cos \gamma_0. \end{aligned} \quad (5.18)$$

Eliminating here $\tan \gamma_0$ by means of (5.3) yields

$$\xi_0(\tau) = -2 \frac{dR}{d\tau} \frac{s_1 s_2 R^{-m-1}}{(1+2n)} \cos \gamma_0. \quad (5.19)$$

Substituting (5.19) into (5.12) gives

$$\xi_\theta^p = -\frac{2q s_1 s_2 \cos \gamma_0 \rho^m}{(1+2n)} \frac{R^{-m-1} dR}{d\tau}. \quad (5.20)$$

Integrating with respect to τ determines the circumferential plastic strain as

$$\varepsilon_\theta^p = \frac{2q s_1 s_2 \cos \gamma_0 \rho^m}{(1+2n)m} R^{-m} + \varepsilon_0(\rho), \quad (5.21)$$

where $\varepsilon_0(\rho)$ is an arbitrary function of ρ . This function should be found using the condition $\varepsilon_\theta^p = 0$ at the elastic/plastic boundary. Then, it follows from (5.21) that

$$\varepsilon_\theta^p = \frac{2q s_1 s_2 \cos \gamma_0}{(1+2n)m} \left(\frac{\rho^m}{R^m} - 1 \right). \quad (5.22)$$

Substituting (5.1) at $\gamma = \gamma_0$ into (5.9), eliminating $\tan \gamma_0$ by means of (5.3), and integrating with the respect to τ using the condition that $\varepsilon_\theta^p = 0$ when $\varepsilon_r^p = 0$ give

$$\varepsilon_r^p = \varepsilon_\theta^p \frac{(1-n)}{(1+2n)}. \quad (5.23)$$

6. Design of the Disc

The solution found can be used to search for two kinds of optimal conditions. In particular, it is possible to search for a uniform distribution of stresses at the final stage of loading. This kind of design requires that the plastic zone occupies the entire disc. The stresses at any point of the disc are given by (5.1), where γ should be replaced with γ_0 . Putting $R = 1$ in (5.8) determines the value of $\tau = \tau_p$ at which the entire disc becomes plastic

$$\tau_p = -A_0(1+n-\nu+s_1) - B_0(1+n-\nu+s_2). \quad (6.1)$$

Using (4.5), (5.3), (5.4), (5.7), and (6.1) it is possible to find a relation between the two optimal loading parameters at the final stage, τ_p and β , at a given value of n . The distribution of the elastic and thermal portions of the strain tensor is uniform at these values of the parameters. However, the plastic portion of the strain tensor varies with the radius according to (5.22) and (5.23). Therefore, not all of the requirements of Michell structures are satisfied. The following solution enables the total strain distribution to be uniform. It is evident from (5.12) that it is possible if and only if $\xi_0(\tau) = 0$. Then, it follows from (5.19) that $s_1 = 0$ or $s_2 = 0$. These conditions along with (4.5) provide two equations for n . The equation corresponding to $s_1 = 0$

has no solution. The other equation gives $n = 0$. In this case $s_1 = -2$. Thus, the distribution of strains in the disc of constant thickness is uniform if $\beta = 1$, as follows from (5.3) and (5.4). Moreover, it is seen from (5.3) and (5.7) that $A_0 = 0$ at $n = 0$ and $s_2 = 0$. Therefore, (5.8) does not provide any relation between R and τ . The physical meaning of this feature of the solution is that the plastic zone simultaneously occupies the entire disc of any size. Thus, this design satisfies the criterion adopted in [1]. The corresponding value of τ can be found from the thermoelastic solution. In particular, since β is given, A_e and B_e in (4.7) are solely dependent of τ . Therefore, replacing A and B in (4.4) with A_e and B_e , respectively, and putting $\rho = \omega$ determine the stresses as functions of τ . Finally, substituting these functions into the yield criterion (2.7) gives the equation for the value of τ at which the entire disc becomes plastic.

7. Conclusions

New solutions for design of a thin annular disc subject to thermomechanical loading have been proposed. Two design criteria have been adopted. One of the criteria requires that the state of stress is uniform at the final stage of loading. This criterion leads to a relation between optimal values of the loading parameters for each specific shape of the disc. A more restrictive criterion additionally requires, by analogy to Michell structures, that the state of strain is uniform at the final stage. Application of this criterion has shown that the state of stress and strain required appears in the disc of constant thickness at certain values of the loading parameters. This design also satisfies the requirements formulated in [1]. Possible developments of the approach proposed include plastic anisotropy, pressure-dependency of the yield criterion, and variation of some material properties along the radius.

Acknowledgment

The research described in this paper has been supported by the Grants NSC-99-2218-E-194-003-MY3, RFBR-11-01-93001, and GK No. 11.519.11.3015. A part of this work was done while the first author was with National Chung Cheng University (Taiwan) as a Research Scholar under the recruitment program supported by the National Science Council of Taiwan (contact 99-2811-E-194-009).

References

- [1] K. Hein and M. Heinloo, "The design of nonhomogeneous equi-strength annular discs of variable thickness under internal and external pressures," *International Journal of Solids and Structures*, vol. 26, no. 5-6, pp. 617-630, 1990.
- [2] B. Farshi, H. Jahed, and A. Mehrabian, "Optimum design of inhomogeneous non-uniform rotating discs," *Computers and Structures*, vol. 82, no. 9-10, pp. 773-779, 2004.
- [3] P. Dewhurst, "Analytical solutions and numerical procedures for minimum-weight Michell structures," *Journal of the Mechanics and Physics of Solids*, vol. 49, no. 3, pp. 445-467, 2001.
- [4] J. Chakrabarty, "Elastic-plastic expansion of a circular hole in an infinite plate of variable thickness," *International Journal of Mechanical Sciences*, vol. 13, no. 5, pp. 439-444, 1971.
- [5] J. W. Phillips and M. Schrock, "Note on shear stresses in accelerating disks of variable thickness," *International Journal of Mechanical Sciences*, vol. 13, no. 5, pp. 445-449, 1971.
- [6] A. N. Sherbourne and D. N. S. Murthy, "Stresses in disks with variable profile," *International Journal of Mechanical Sciences*, vol. 16, no. 7, pp. 449-459, 1974.

Research Article

CT Metal Artifact Reduction Method Based on Improved Image Segmentation and Sinogram In-Painting

**Yang Chen,^{1,2,3} Yinsheng Li,^{1,2,3} Hong Guo,⁴
Yining Hu,^{1,2,3} Limin Luo,^{1,2} Xindao Yin,⁵ Jianping Gu,⁵
and Christine Toumoulin^{2,3}**

¹ *Laboratory of Image Science and Technology, Southeast University, Nanjing 210096, China*

² *Centre de Recherche en Information Biomedicale Sino-Francais (LIA CRIBs), 35042 Rennes, France*

³ *Laboratoire Traitement du Signal et de l'Image (LTSI) INSERM U642, Université de Rennes I,
35042 Rennes Cedex, France*

⁴ *Department of Radiology, General Hospital of Tianjin Medical University, Tianjing 300052, China*

⁵ *Department of Radiology, Nanjing First Hospital Affiliated to Nanjing Medical University,
Nanjing 210029, China*

Correspondence should be addressed to Jianping Gu, gujianpingdoctor@gmail.com

Received 21 February 2012; Revised 25 June 2012; Accepted 27 June 2012

Academic Editor: Fatih Yaman

Copyright © 2012 Yang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The streak artifacts caused by metal implants degrade the image quality and limit the applications of CT imaging. The standard method used to reduce these metallic artifacts often consists of interpolating the missing projection data but the result is often a loss of image quality with additional artifacts in the whole image. This paper proposes a new strategy based on a three-stage process: (1) the application of a large-scale non local means filter (LS-NLM) to suppress the noise and enhance the original CT image, (2) the segmentation of metal artifacts and metallic objects using a mutual information maximized segmentation algorithm (MIMS), (3) a modified exemplar-based in-painting technique to restore the corrupted projection data in sinogram. The final corrected image is then obtained by merging the segmented metallic object image with the filtered back-projection (FBP) reconstructed image from the in-painted sinogram. Quantitative and qualitative experiments have been conducted on both a simulated phantom and clinical CT images and a comparative study has been led with Bal's algorithm that proposed a similar segmentation-based method.

1. Introduction

Metallic objects like dental implants, surgical clips, or steel-hip prostheses lead to severe shadow and streak artifacts in CT images that superimpose the structures of interest and

deteriorate image quality. The reason is that metallic objects have a very high density in the human body, which creates a barrier to the transmitted x-ray beam during CT examination. It results a lack of data in the projection data that lead to the production of streak artifacts in CT images [1, 2]. This photo deficiency caused by metallic object would become more severe under low dose scanning [2]. In the last decade, many approaches have been proposed to reduce these artifacts. These methods can be roughly classified into iterative and interpolation-based methods.

Iterative algorithms operate in a feed-back mode in both the image and projection data spaces [3–5]. However, three major difficulties can be pointed out for that methods: (1) the well-formatted original raw projection data are often unavailable because the leading manufacturers of CT imaging devices are often reluctant to provide it; (2) the involved high computational cost of iterative algorithms often requires an implementation on specialized processor units; (3) the iterative algorithm still need to be combined with sinogram correction method when the metal artifacts are rather severe.

Interpolation-based methods correct metal artifacts directly in sinogram space. Compared to iterative algorithms, these methods are less computationally expensive and can be implemented without the availability of the original raw projection data. They aim at identifying the corrupted segments in the sinogram and interpolating the data from noncorrupted neighboring projections. Some of these methods add other steps to improve the sinogram correction accuracy and design a four-stage process that consists of image enhancement, metallic object segmentation, image forward projection, and sinogram in-painting, final image reconstruction using FBP [6–12]. It is worth notable that the normalization operation suggested in [9, 10] has been considered to give a more accurate attenuation estimation.

We propose a new method to suppress the metal artifacts and improve the sinogram completeness that is based on the above described scheme. The major contribution states at the image enhancement, segmentation, and sinogram inpainting levels with, respectively, the application of a large scale nonlocal means filter (LS-NLM), a mutual information maximized segmentation (MIMS), and a modified exemplar-based in-painting technique. The description of this method is given in Section 2. Comparative experiments with the method proposed by Bal and Spies in [12] are then provided in Section 3. This method was chosen for comparison because it made use of a similar strategy to address the corrupted sinogram problem, that is, Tensor filtering, k-means clustering technique-based segmentation, and linear interpolation-based sinogram in-painting. It will be referred thereafter in this paper by “Bal’s algorithm.” We will show that, compared to this algorithm, our method provides a better sinogram correction. Visual and quantitative analysis are also reported to highlight this superiority. A CUDA parallelization technique has been applied to accelerate the calculations of the patch distances involved in the image enhancement and sinogram inpainting steps, respectively.

2. Method

The proposed sinogram completeness algorithm is divided into four major stages.

Step 1 (prefiltering). The original CT image including metal artifacts is enhanced with the edge-preserving LS-NLM filter.

Step 2 (image segmentation). The metallic artifacts and objects are respectively extracted using the MIMS algorithm with a partitioning of the image into different regions.

Step 3 (sinogram inpainting). Once the metallic objects and artifacts have been extracted, the segmented artifact image is forward projected to determine the projection data in the sinogram space which are affected by the artifacts. A subtraction is performed between the corrupted sinogram and the original one. The missing projection data in the subtracted sinogram are then restored using a modified exemplar-based in-painting technique.

Step 4 (backward projection of the in-painted sinogram and image correction). The artifact compensated image is then reconstructed from the in-painted sinogram using the FBP algorithm. Afterward, the final corrected image is obtained by inserting the previously segmented metal component into the reconstructed image.

Of all the four steps, Step 2 and Step 3 are the two key steps in which the damaged sinogram data are estimated and corrected. The above stages are detailed in the following subsections with the flowchart displayed in Figure 1.

2.1. Image Enhancement

This first stage aims at applying an edge-preserving filtering operation to smooth and denoise the streak artifacts in the original CT images. The LS-NLM filter has been proven to be efficient for image denoising with edge preservation. It was, for instance, applied with success to suppress mottled noise in low-dose abdominal CT images [13]. The principle is to replace the value of a pixel by the weighted average of pixels located in a neighbourhood window of size N . Each weight expresses the similarity between the central pixel in the window and each neighboring pixel and is given by the pair-wise difference between patches surrounding each pair of considered pixels [14, 15].

Let f^O and f^F , respectively, denote the images before and after filtering, the filter output is given by

$$f_i^F = \frac{\sum_{j \in N_i} w_{ij} f_j^O}{\sum_{j \in N_i} w_{ij}}, \quad (2.1)$$

$$w_{ij} = \exp\left(\frac{-\|s_i - s_j\|_{2,\alpha}^2}{h}\right), \quad (2.2)$$

where f_i^O is the pixel located at the center of the neighborhood N and f_j^O the pixels located in the neighborhood of f_i^O . w_{ij} denotes the weight between pixels f_i^O and f_j^O and is calculated as a similarity measure between the two patches s_i and s_j surrounding each pair of pixels i and j in the neighbourhood N_i , respectively. The decay parameter h acts as a filtering parameter. In (2.2), a Gaussian kernel of standard deviation α is used to take into account the distance between the central pixel and other pixels in the patch. The LS-NLM filter involves working with a large-size neighborhood N and a number of size patches s equal to the number of pixel in the neighborhood N , which implies high costs for calculating the distance between each patch pair in each neighborhood N . To accelerate the computation, a GPU parallelization using the CUDA framework was applied for the original pixel-wise processing based on [16–18].

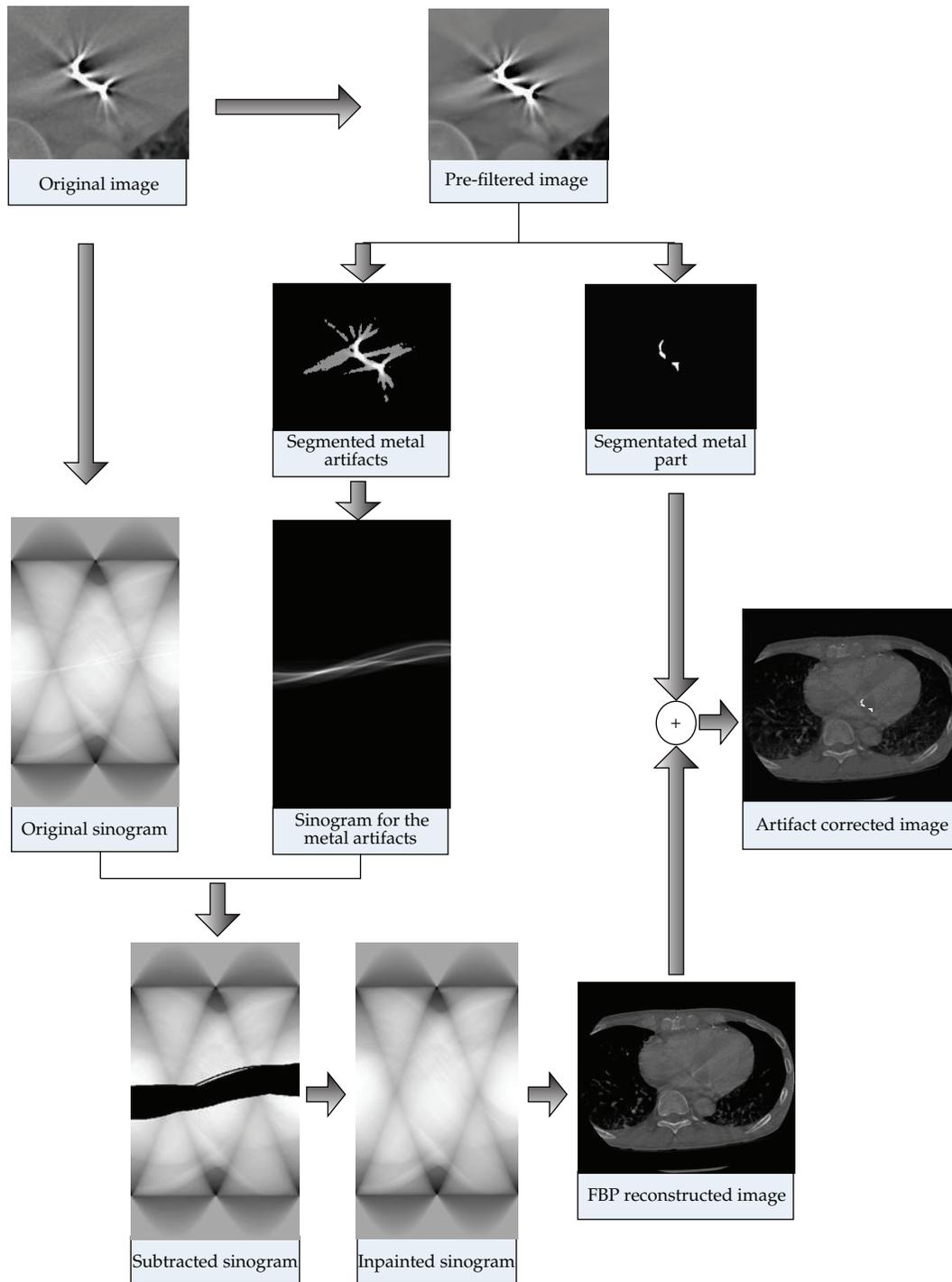


Figure 1: Flowchart of the proposed correction method.

2.2. Image Segmentation

The MIMS method is based on the maximum mutual information (MMI) and allows determining the class number n based on the difference of mutual information (DMI) [19]. The mutual information $MI(A, B)$ between image A and B is defined by

$$MI(A, B) = H(A) + H(B) - H(A, B), \quad (2.3)$$

where $H(A)$ and $H(B)$, respectively, denote the entropies of the images A and B , and $H(A, B)$ the mutual entropy. Based on [15, 16], the $MI(A, B)$ can be rewritten as the joint probability density distribution of images A and B :

$$MI(A, B) = \sum P(A, B) \log \frac{P(A, B)}{P(A)P(B)}, \quad (2.4)$$

$$DMI_n(f^F) = \frac{MI(f^F, S_n) - MI(f^F, S_{n-1})}{MI(f^F, f^F)}, \quad (2.5)$$

where $P(A)$, $P(B)$ are the probability density distribution of A and B , respectively, and $P(A, B)$ the joint probability density function. $P(A)$, $P(B)$, and $P(A, B)$ can be computed from the histogram quantization [20]. Based on [18], in (2.5), $MI(f^F, S_n)$ is the normalized mutual entropy between the image f^F and the current image S_n segmented into n classes, and $DMI_n(f^F)$ the normalized difference between the entropies $MI(f^F, S_n)$ and $MI(f^F, S_{n-1})$. When the class number increases, $DMI_n(f^F)$ decreases while $MI(f^F, S_n)$ converges towards $MI(f^F, f^F)$. This convergence is reached when DMI becomes smaller than a specified threshold ε . A local optimality can be obtained when DMI converges towards a local minimum while the mutual entropy synchronously reaches its maximum [19]. MIMS implementation only requires to set the maximum class numbers MCN (MCN_a and MCN_m for the segmentations of artifacts and metals, resp.) and the threshold ε . A description of the algorithm is given in Figure 2. The filtered image f^F is classified into n classes by using an intensity threshold vector G_n^k (k being the k th iteration and n referring to the threshold number). The thresholds in G_n^k and the class centers can be automatically computed within a simulated annealing- (SA-) based optimization process [21].

2.3. Sinogram In-Painting

The output of the MIMS algorithm provides the segmented metal artifact and object images. We forward projected the segmented metal artifact image into the sinogram domain, subtracted then the original sinogram (built from the original image that includes the metallic objects) with the metal artifact sinogram to delete the corrupted projection data, and applied an in-painting technique on the subtracted sinogram to restore the missing projection data. The proposed in-painting technique refers to a modified exemplar-based in-painting method to find the best matched sinogram patch for the restoration of the missing projection data [22, 23].

Let consider the following hold.

- (1) The splitting of the subtracted sinogram into two regions: Ω the region to be filled (in which data are missing) and θ the region where the information is complete.

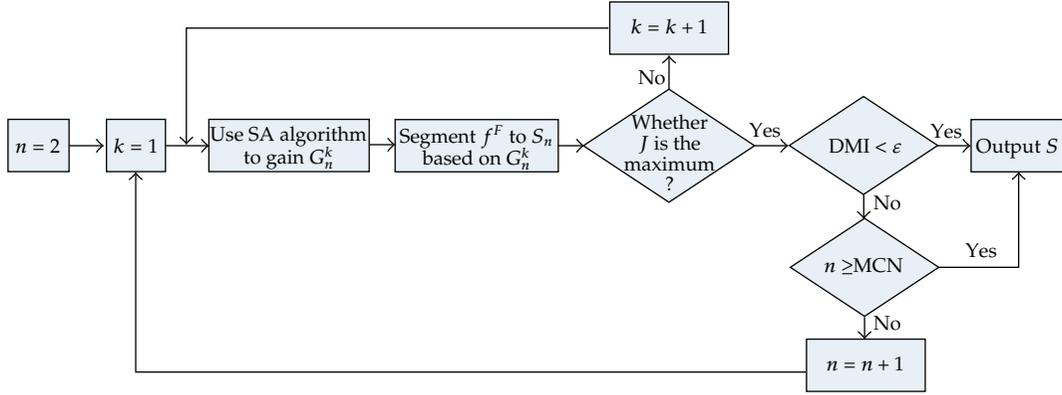


Figure 2: Flowchart of the MIMS algorithm: J corresponds to $MI_n(f^F, S_n)$ and G_n^k is the intensity threshold vector for the k th iteration with current class number n . f^F refers to the filtered image at the previous stage.

- (2) A patch P^p centered on each pixel p located in the region Ω at the border of the frontier between the two regions Ω and θ so that the patch P^p includes a certain number of pixels belonging to θ .
- (3) A set of patches P^q centered on each pixel q of the region θ .

Then the patch P^p is compared with each patch P^q using the following similarity metric:

$$d(p, q) = \|P^p - P^q\|, \quad (2.6)$$

where the distance $d(p, q)$ is calculated between the corresponding pixels in patches P^p and P^q that belong to the region θ . P^p and P^q are of the same sizes, and $\|\cdot\|$ denotes the Euclidean distance between them. The patch P^q that minimizes this distance is selected and its contents are copied into P^p to restore the missing pixels of P^p that are located in the region Ω . An automatic region filling process is conducted based on [22] that introduces a filling order of the pixels located in the region Ω . It goes through the iteration of a three-stage process for each pixel.

- (1) The computation of priority for all the patches P^p whose central pixels p are located in region Ω but just behind the border line of θ . The priority is computed with the following function $\text{Pr}(p)$ for each of the pixels p :

$$\text{Pr}(p) = \lambda C(p) + (1 - \lambda)D(p), \quad (2.7)$$

where $D(p)$ and $C(p)$ are, respectively, called data and fidelity term. The latter one is introduced to quantify the number of points, surrounding the target pixel p , that are known or have already been in-painted. This term tends to privilege those patches that have more pixels from the known region θ . $D(p)$ is given by the scalar product between the vector normal to the front and the maximum gradient orientation at point p . Its objective is to encourage linear structures to be processed

first against high curvature line. λ is a weighting coefficient ($0 < \lambda < 1$) that controls the balance between $C(p)$ and $D(p)$. Details of calculus of $C(p)$ and $D(p)$ can be found in [22].

- (2) The choice of the pixel p with the highest priority (calculated by (2.7)) and the search for the best match between the patch P^p and the set of patches P^q - determined by the similarity metric (2.6). Then, the pixels of patch P^p located in the region Ω are then restored with the corresponding known pixels of the patch P^q .
- (3) The updates of the fidelity and data terms for pixels of P^p that have been filled and are afterward located just behind the border line of θ .

In practice, to increase the accuracy in the in-painting process, the size of patch P^p for in-painting is in fact set smaller than when applying the search for the best match in (2.6).

3. Experiments

Experiments were conducted both on a simulated phantom and clinical CT images. All the images have a size of 512×512 . Clinical CT datasets were acquired from a multidetector row CT unit with 16 detector rows (Somatom Sensation 16; Siemens Medical Solutions). The scanning protocol was 100 mAs, 120 kVp, 5 mm slice thickness and the spatial resolution was 0.457 mm^2 . The images were reconstructed with a FBP algorithm using a convolution kernel "B40f." Here, convolution kernel is used to control the smoothing effect in CT images for Siemens Somatom Sensation 16 system, and B40f is the routine convolution kernel for brain CT reconstruction. Figures 3(a), 3(b), and 3(c) display the three original clinical CT images that include metallic artifacts. The first dataset depicts a chest image in which metallic artifacts come from a metallic suture material. The two other datasets refer to a brain image where the metallic artifacts originate from golden earrings. A phantom image including metallic artifacts was also simulated to allow quantitative comparisons. It consists of a cylindrical metallic insert incorporated in a cylindrical water container (Figure 4(a)). The phantom was simulated from an artifact-free phantom CT image (Figure 4(b)) applying the following intensity settings: air: 0; square container: 500; cylindrical water receptacle: 3000, water: 1000. Subsequently, we name the images in Figures 3(a), 3(b), 3(c), and 4(a), ClinicalImage1, ClinicalImage2, ClinicalImage3 and PhantomImage, respectively. We applied a parallel geometry for the forward and backward projection operations involved in the metallic artifact reduction algorithm. This algorithm has been written in C language, MATLAB (release R2006b), and NVIDIA CUDA libraries. It was then run on a PC with an Inter Core i5 processor, $2.68 \times 4 \text{ GHz}$, 6 G RAM, and GPU (NVIDIA GTX465).

We compared our method with Bal's algorithm in [12] which made use of a similar strategy to reduce the metal artifacts.

- (1) A linear structure tensor (LST) filtering proposed in [24] is applied to reduce noise and smooth the artifacts in the original image with artifacts. Three parameters needed to be sets: the mask size v , the scaling factor σ_0 , and the relationship between width and length of Gaussian filter.
- (2) A cluster-based K-means method to segment the metallic objects. This algorithm requires a suitable setting of the class numbers and the initial center of each cluster.

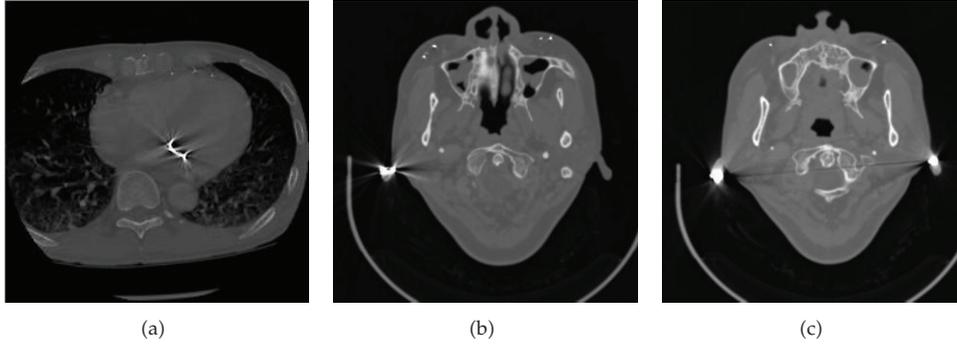


Figure 3: Clinical CT images including metal artifacts. (a) ClinicalImage1 is a chest image with artifacts caused by metallic suture material; (b) ClinicalImage2 is a brain image with artifacts caused by one golden earring; (c) ClinicalImage3 is a brain image with artifacts caused by two golden earrings. Note that the image quality is severely degraded by metal artifacts.

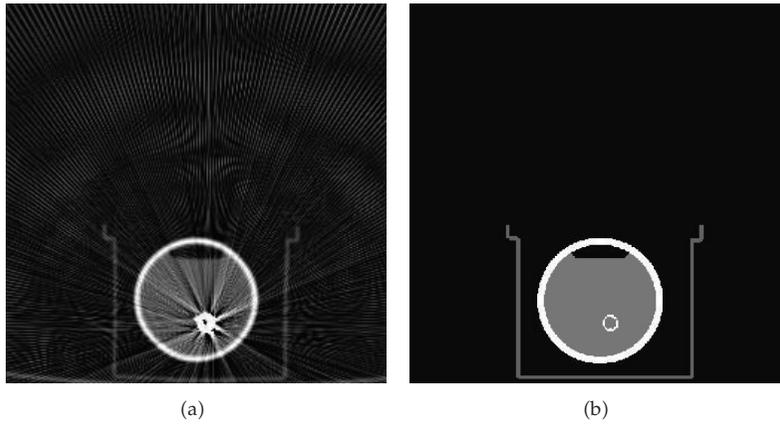


Figure 4: Simulated CT images including metal artifacts. (a) Phantom image including simulated metal artifact (PhantomImage); (b) the original artifact-free phantom image used to create PhantomImage. Note that the image quality is also severely degraded by metal artifacts.

- (3) An in-painting step in which the neighboring sinogram data (from the projection of above K-means segmentation) is used to complete the tagged metallic projection for the segmented metallic objects.

The proposed process involves also to specify a certain number of parameters: N , s , the decaying parameter h in the prefiltering step, the maximum class numbers MCN_a , MCN_m to, respectively, extract the metallic artifacts and object, the threshold ε for the segmentation step, the size of the patches P^p , and the value of λ for the in-painting step.

For both algorithms, the involved parameters were manually set to find the optimal parameter combination that led to the best qualitative results. This qualitative evaluation was carried out in collaboration with a radiologist (Xindao Yin 10 years clinical experience). These optimal parameters are listed in Table 1. It is notable that we used the same parameter setting in Bal's method because we found similar results can be obtained when using the same parameter settings.

Table 1: Parameter setting for Bal's algorithm and our proposed method. For our proposed correction, MCN_a and MCN_m depict the maximum class number (MCN) in the MIMS segmentation of artifacts and metal components, respectively.

	ClinicalImage1	ClinicalImage2	ClinicalImage3	PhantomImage
Bal algorithm in [7]				
Pre-filtering step		$v = 2$, the scaling factor σ_0 is set to 10,		
		the relationship between width and length of Gaussian filter is set to 2 ($\sigma_1 = 2\sigma_2$).		
Segmentation step		K-means segmentation using 5 classes with CT values: -950 (air), 0 (soft tissue), 200 (normal tissue), 750 (bone), 5000 (metal).		
Inpainting step		Linear interpolation using 5 points in each symmetric side		
Our proposed method				
Pre-filtering step	$41 \times 41N$, $11 \times 11s$, $h = 500$	$41 \times 41N$, $11 \times 11s$, $h = 450$	$41 \times 41N$, $11 \times 11s$, $h = 450$	$41 \times 41N$, $11 \times 11s$, $h = 350$
Segmentation step	$MCN_a = 16$, $MCN_m = 22$, $\epsilon = 600$	$MCN_a = 12$, $MCN_m = 18$, $\epsilon = 600$	$MCN_a = 12$, $MCN_m = 18$, $\epsilon = 600$	$MCN_a = 5$, $MCN_m = 7$, $\epsilon = 600$
In-painting step	$11 \times 11P^p$, $3 \times 3P^{iq}$, $\lambda = 0.65$	$11 \times 11P^p$, $3 \times 3P^{iq}$, $\lambda = 0.3$	$11 \times 11P^p$, $3 \times 3P^{iq}$, $\lambda = 0.3$	$11 \times 11P^p$, $3 \times 3P^{iq}$, $\lambda = 0.8$

3.1. Qualitative Study

Figures 5(a) and 5(b) and Figures 5(a) and 5(b) display for comparison each of the steps of each algorithm (Bal's algorithm and our proposed method). ClinicalImage1 is used for validation. From Figures 5(a1) and 5(b1), we can see that the LS-NLM filtering shows better properties of noise suppression and structure preservation than the ASF filtering used in Bal's algorithm. Also, only the metallic object was segmented with Bal's K-means algorithm (Figure 5(a2)) while our MIMS algorithm allowed both the artifacts and the metallic objects to be extracted (Figure 5(b2)). Figures 5(a3) and 5(b3) point up the differences between the two resulting projected corrupted sinogram. In particular, the corrupted sinogram obtained from the segmented metallic artifact component in Figure 5(b2) is larger than the one issues from the segmented metallic component in Figure 5(a2). Figures 5(a4) and 5(b4) provide the two resulting in-painted sinogram. In the one obtained with Bal's algorithm, we can see metallic shadows resulting from the metallic artifacts segmentation (see the red arrow in the zoomed region in Figure 5(a4)). These shadows are absent in Figure 5(b4), that is, the sinogram computed from our method. These metallic shadows create new artifacts in the reconstructed image in Figure 5(a5). Finally, this example illustrates the best performance of our method on Bal's algorithm.

Figures 6 and 7 illustrate these results on the other two clinical CT images ClinicalImage2 and ClinicalImage3, respectively. A region of interest (ROI) delineated by red lines is zoomed to emphasize the behavior of each algorithm. Severe streak artifacts can be observed in the original image that spread from the ears at the metallic component location. The resulting image analysis (Figures 6(c), 6(d), 7(c) and 7(d)) makes appear that each method provides a substantial reduction of the metallic artifacts. However, results depicted in (Figures 6(d) and 7(d)) appear smoother with less accentuated artifacts. Structures are better preserved while in (Figures 6(c) and 7(c)) some artifacts remain that deteriorate the quality of the image. These artifacts are more pronounced in the close metallic object surrounding.

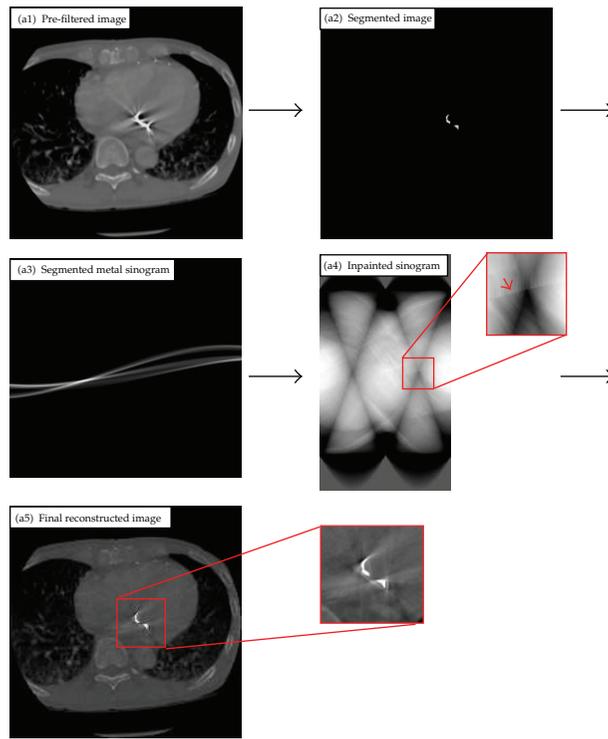
We can also see in Figures 6(d) and 7(d) that some new streak artifacts (pointed by black arrows in Figures 6(d) and 7(d)), though not strong, were introduced into the processed images. The introduced new artifacts might come from the errors of the segmentation and inpainting in the proposed approach.

3.2. Quantitative Study

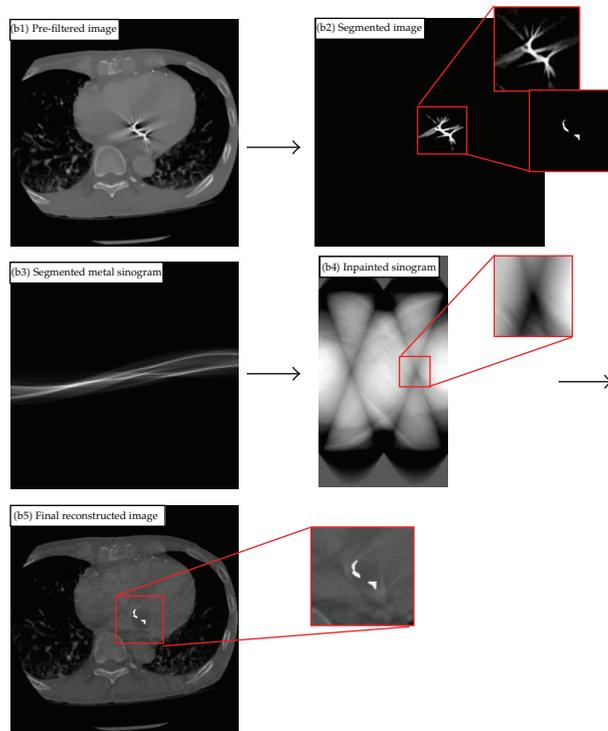
A quantitative study was then carried out on the simulated phantom of Figure 4(a) with respect to the original artifact-free phantom image in Figure 4(b). We displayed the intensity profiles along a specified horizontal line in the original and corrected images to highlight the differences in the behavior of each method according to the crossed structure properties. We also computed the mean square error (MSE) and the standard deviation (STD) in two ROI located in two different homogeneous regions that is, inside (region1) and outside (region2) the phantom:

$$\text{MSE} = \frac{1}{N_{\Omega}} \sqrt{\sum_{j \in \Omega} (f_j^c - f_j^o)^2}, \quad (3.1)$$

$$\text{STD} = \sqrt{\frac{1}{N_{\Omega}} \sum_{j \in \Omega} (f_j^c - \bar{f}_{\Omega}^c)^2}, \quad (3.2)$$



(a)



(b)

Figure 5: Artifact correction on ClinicalImage1 (Figure 3(a)) using: (a), Bal's algorithm [7]; (b), our proposed method.

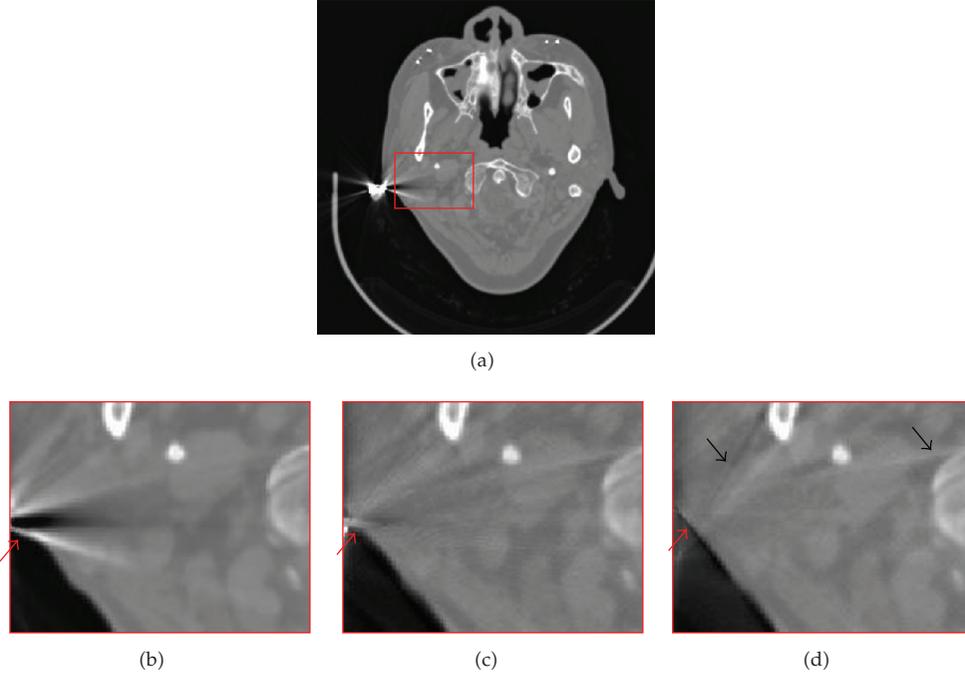


Figure 6: Artifact correction on ClinicalImage2. (a) Original CT image with artifacts; (b) zoomed ROI in the original image; (c) zoomed ROI in the result image obtained from Bal's algorithm; (d) zoomed ROI in the result image obtained from the proposed method.

where N_{Ω} denotes the pixel number in the chosen region Ω (Region1 or Region2). f_j^c and f_j^o are the intensity of the pixel within the region Ω in the corrected image f^c and the original reference image f^o , respectively. \bar{f}_{Ω}^c characterizes the mean intensity in the ROI, in the corrected images.

The phantom images after correction by Bal's algorithm and our proposed method are, respectively, displayed in Figures 8(a) and 8(b). The qualitative evaluation highlights the capacity of our algorithm to better suppress the metallic artifacts as to provide a superior consistency in the homogeneous region preservation. Figure 8(c) plots the intensity profiles along the same given horizontal line in the original (Figures 4(a) and 4(b)) and corrected (Figures 8(a) and 8(b)) images, respectively. The profiles confirm that our method brings a better quality correction in the homogeneous region. Table 2 provides the MSE and STD measures for each method and in each ROI. The figures also confirm the supremacy of our approach with an MSE and an STD that are the lowest on the image set.

Table 3 provides the total computation costs (in CPU seconds) for each step for each method. For our method, the computation cost is given without and with GPU acceleration. This method is rather expensive in computation time. The CUDA parallelization brings a substantial gain with an acceleration that can reach a rate of 10 to 30 depending on the complexity of the image. The parallelization makes then our method competitive in computation time with Bal's algorithm.

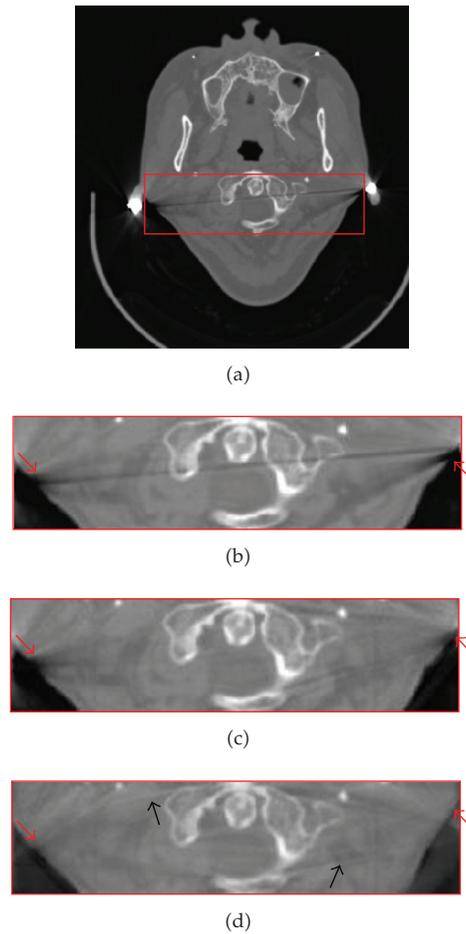


Figure 7: Artifact correction on ClinicalImage3. (a) Original CT image with artifacts; (b) zoomed ROI in the original image; (c) zoomed ROI in the result image obtained from Bal' algorithm; (d) zoomed ROI in the result image obtained from the proposed method.

4. Discussion

The proposed strategy adopted for reducing the metallic artifacts in the reconstructed image relies on a four-stage process that consists of image enhancement, metallic object segmentation, image forward projection, and sinogram in-painting, final image reconstruction using FBP. The image enhancement makes use of an LS-NLM filter. This filter exploits a patch similarity measure to smooth the image while preserving the edges. Its response is not very sensitive to the size of the neighborhood N and patch s . However, the decaying parameter h , that quantifies the smoothing rate and how fast the weights decay with increasing dissimilarity of respective patches, is sensitive to the noise ratio in the image. Its value is set as a function of the noise variance. The MIMS-based segmentation involves only to set the maximum class number MCN (MCN_a and MCN_m for the segmentations of artifacts and metals, resp.) and the threshold ε that is applied for the convergence of the difference mutual information (DMI). The choice of this threshold is not sensitive as we can see in the experiments. Its value is relatively stable on the set of the processed

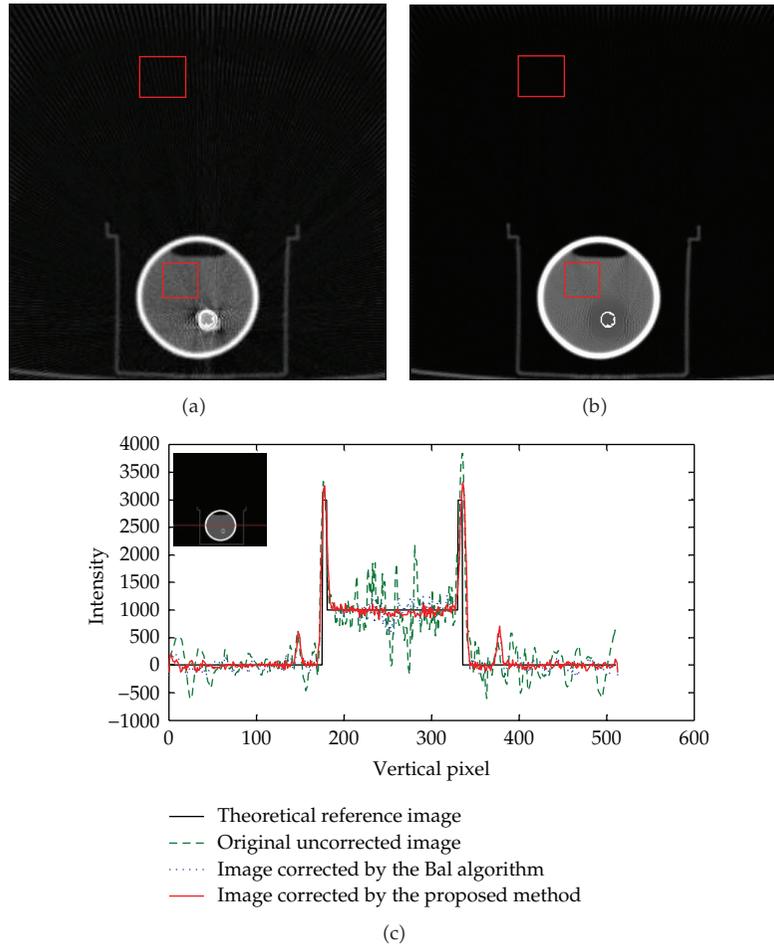


Figure 8: Line profile comparison for different artifact corrections. (a) corrected image using Bal's algorithm; (b) corrected image using our method; (c) line profiles along a given line in the original image and each corrected image. The location of the line is drawn in red in the medallion image that appears at the upper left part of the plot.

Table 2: MSE and STD on the original corrupted and corrected phantom images. We can see the calculated MSE and STD are lower for the images corrected with our method.

		Original uncorrected image (Figure 4(a))	Image corrected by the Bal algorithm (Figure 8(a))	Image corrected by the proposed method (Figure 8(a))
MSE of the whole images		1.37	1.31	1.26
STD	Region1	302.98	64.21	42.20
	Region2	293.52	99.04	56.81
MSE	Region1	15.15	3.27	2.16
	Region2	6.22	2.66	1.13

Table 3: Computation cost (in CPU seconds) for each steps in the Bal algorithm and the proposed correction method.

		ClinicalImage1	ClinicalImage2	ClinicalImage3	PhantomImage
Bal's Method in [7]	Pre-filtering step	0.98	0.97	0.96	0.97
	Segmentation step	67.23	23.44	24.21	14.13
	Inpainting step	1.32	1.57	1.22	0.85
	Total	69.53	25.98	26.39	15.95
Proposed correction (unparallelized)	Pre-filtering step	237.56	236.92	237.17	236.32
	Segmentation step (metal)	34.45	26.26	28.34	12.39
	Segmentation step (artifacts)	26.21	16.96	16.92	4.93
	Inpainting step	44.02	49.09	64.78	52.81
Total	342.24	329.23	347.21	306.45	
Proposed correction (CUDA parallelized)	Pre-filtering step	6.23	5.97	6.39	5.98
	Segmentation step (metal)	34.45	26.26	28.34	12.39
	Segmentation step (artifacts)	26.21	16.96	16.92	4.93
	Inpainting step	1.42	1.58	2.06	1.74
Total	33.86	24.51	25.37	12.65	

images (Table 3). The MMI (maximum mutual information) and DMI are computed within a simulated annealing optimization process. This allows the thresholds in G_n^k and the class centers to be automatically computed. This guarantees an optimal choice of these parameters. Moreover, the metal artifacts and objects to be extracted have a very high density and if other highly contrasted structures are not located in the close neighborhood, the algorithm run quite well and provides satisfactory results. Considering now the sinogram inpainting stage, the proposed exemplar-based in-painting technique considers a global similarity measure and relies on redundant information present in the image. This modified exemplar-based in-painting method is reasonable for the CT sinogram completion because it is observed there are lots of repetitive structures in CT sinogram. When a large scale is selected, the proposed exemplar-based in-painting can give an effective restoration of repetitive structures in sinogram space. Parameters to be set relate to the sizes of the patches (Table 1) and the weighting coefficient λ that is used to balance the fidelity and data term in (2.7). We preferred to consider small patches (3×3) for the inpainting process in order not to lose subtle details.

For comparison, Bal's algorithm used a similar strategy: (1) a linear structure tensor (LST) filtering was first applied. Its response is highly dependent on the parametric streak strength and orientation quantification and several parameters were empirically tested to ensure a satisfying quantification and efficient filtering. (2) A K-means clustering technique was considered for the metallic artifact segmentation. It required choosing a suitable number of classes and the initial center of each cluster. The K-means algorithm is based on the intensity clustering of one single image and appears finally less efficient than the MIMS method to accurately segment the metallic artifacts. (3) An interpolation technique was then considered to recover missing or metallic data from neighboring projections from non-corrupted segments.

The global process is carried out in a small region surrounding the metallic object. Artifacts around the metal objects can be removed so as to remove projection data inconsistencies with data consistent with similar neighborhood. However, image details may be noticeably altered especially and some remaining artifacts still appear in the regions closest to the metal objects. The presence of pathology has not yet been considered and is something to be further evaluated in collaboration with our medical expert. The reason might be the metallic artefacts can in some situation completely superimpose the structures and the presence of a pathology may be hidden. Thus, the density change due to the artefact removing process may be difficult to evaluate.

5. Conclusion

This paper proposed a new strategy for reducing metallic artifacts in CT images. The proposed method outperforms Bal's algorithm in each of the three steps: image prefiltering, image segmentation, and sinogram inpainting. Visual and quantitative analyses on phantom and clinical data show that the proposed correction method provides a substantial reduction of the metallic artifacts in the corrected images. The pixel-wise operations in the pre-filtering and sinogram inpainting steps are greatly accelerated by a CUDA parallelization that makes the algorithm also competitive in computation time.

Although this algorithm demonstrated a good potential for the reduction of metallic artifacts in CT images, some improvements have to be further considered. First, segmentation accuracy might be further increased by applying more dedicate method such as the segmentation method in [25]. Second, the exemplar-based in-painting procedure is expensive in computation time due to the search for the patch priorities, which needs to be updated after the in-painting of each point within the corrupted sinogram. Third, some intensity inconsistencies can be still observed around the regions of the metallic objects in the corrected images, and the sinusoid property of sinogram has not be exploited in inpainting the missing sinogram data [26]. At last, the presence of pathology in the surrounding of the metal object has not been considered in the evaluation of the algorithm. Thus, further work will be devoted to solve the set of problems as to perform an extensive evaluation on clinical data.

Abbreviation

CT:	Computed tomography
LST:	Linear structure tensor
ASF:	Adaptive Steering Filter
NAST:	Nonlinear anisotropic structure tensor
LS-NLM:	Large scale nonlocal means
MIMS:	Mutual information Maximized Segmentation
MMI:	Maximum mutual information
DMI:	Difference of mutual information.

Acknowledgment

This research was supported by National Basic Research Program of China under Grant (2010CB732503), National Natural Science Foundation under Grants (81000636), the Project supported by Natural Science Foundations of Jiangsu Province (BK2009012 and BK2011593).

References

- [1] W. A. Kalender, R. Hebel, and J. Ebersberger, "Reduction of CT artifacts caused by metallic implants," *Radiology*, vol. 164, no. 2, pp. 576–577, 1987.
- [2] M. Robert, G. Michael, and H. David, "Artifact analysis and reconstruction improvement in helical cardiac cone beam CT," *IEEE Transactions on Medical Imaging*, vol. 23, no. 9, pp. 1150–1164, 2004.
- [3] X. Zhang, J. Wang, and L. Xing, "Metal artifact reduction in x-ray computed tomography (CT) by constrained optimization," *Medical Physics*, vol. 38, no. 2, pp. 701–711, 2011.
- [4] C. Lemmens, D. Faul, and J. Nuyts, "Suppression of metal artifacts in CT using a reconstruction procedure that combines MAP and projection completion," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 250–260, 2009.
- [5] Y. Chen, D. Gao, C. Nie et al., "Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior," *Computerized Medical Imaging and Graphics*, vol. 33, no. 7, pp. 495–500, 2009.
- [6] J. W. Gu, L. Zhang, G. Q. Yu, Y. X. Xing, and Z. Q. Chen, "X-ray CT metal artifacts reduction through curvature based sinogram inpainting," *Journal of X-Ray Science and Technology*, vol. 14, no. 2, pp. 73–82, 2006.
- [7] W. J. H. Veldkamp, R. M. S. Joemai, A. J. Van Der Molen, and J. Geleijns, "Development and validation of segmentation and interpolation techniques in sinograms for metal artifact suppression in CT," *Medical Physics*, vol. 37, no. 2, pp. 620–628, 2010.
- [8] D. Prell, Y. Kyriakou, M. Beister, and W. A. Kalender, "A novel forward projection-based metal artifact reduction method for flat-detector computed tomography," *Physics in Medicine and Biology*, vol. 54, no. 21, pp. 6575–6591, 2009.
- [9] D. Prell, Y. Kyriakou, T. Struffert, A. Dörfler, and W. A. Kalender, "Metal artifact reduction for clipping and coiling in interventional C-Arm CT," *American Journal of Neuroradiology*, vol. 31, no. 4, pp. 634–639, 2010.
- [10] E. Meyer, R. Raupach, M. Lell, B. Schmidt, and M. Kachelrieß, "Normalized metal artifact reduction (NMAR) in computed tomography," *Medical Physics*, vol. 37, no. 10, pp. 5482–5493, 2010.
- [11] Y. Zhang, Y. F. Pu, J. R. Hu, Y. Liu, Q. L. Chen, and J. L. Zhou, "Efficient CT metal artifact reduction based on fractional-order curvature diffusion," *Computational and Mathematical Methods in Medicine*, vol. 2011, Article ID Article ID, 173748, 9 pages, 2011.
- [12] M. Bal and L. Spies, "Metal artifact reduction in CT using tissue-class modeling and adaptive prefiltering," *Medical Physics*, vol. 33, no. 8, pp. 2852–2859, 2006.
- [13] Y. Chen, W. Chen, X. Yin et al., "Improving low-dose abdominal CT images by Weighted Intensity Averaging over Large-scale Neighborhoods," *European Journal of Radiology*, 2010.
- [14] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2005.
- [15] Y. Chen, J. Ma, Q. Feng, L. Luo, P. Shi, and W. Chen, "Nonlocal prior Bayesian tomographic reconstruction," *Journal of Mathematical Imaging and Vision*, vol. 30, no. 2, pp. 133–146, 2008.
- [16] A. Kharlamov and V. Podlozhnyuk, "Image Denoising," Tech. Rep., NVIDIA, Santa Clara, Calif, USA, 2007.
- [17] Accelerating MATLAB with CUDA Using MEX Files, (White Paper), <http://developer.nvidia.com/object/matlab>.
- [18] Y. Chen, Y. S. Li, W. Yu, L. M. Luo, W. Chen, and C. Toumoulin, "Joint-map tomographic reconstruction with patch similarity based mixture prior model," *Multiscale Modeling and Simulation*, vol. 9, no. 4, pp. 1399–1419, 2011.
- [19] Q. W. Lu and W. F. Chen, "Image segmentation based on mutual information," *Chinese Journal of Computers*, vol. 29, no. 2, pp. 296–301, 2006.
- [20] J. Rigau, M. Feixas, M. Sbert, A. Bardera, and I. Boada, "Medical image segmentation based on mutual information maximization," in *Proceeding of the 7th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI '04)*, pp. 135–142, September 2004.
- [21] D. Bertsimas and J. Tsitsiklis, "Simulated annealing," *Statistical Science*, vol. 8, no. 1, pp. 10–15, 1993.
- [22] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.
- [23] A. Wong and J. Orchard, "A nonlocal-means approach to exemplar-based inpainting," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '08)*, pp. 2600–2603, October 2008.

- [24] W. Wei, G. Jinghuai, and L. Kang, "Structure-Adaptive anisotropic filter with local structure tensors," in *Proceedings of the 2nd International Symposium on Intelligent Information Technology Application (IITA '08)*, pp. 1005–1010, December 2008.
- [25] M. Meilinger, O. Schütz, C. Schmidgunst, and E. W. Lang, "Projective segmentation of metal implants in cone beam computed tomographic images," in *Proceedings of the International Symposium on Image and Signal Processing and Analysis*, Dubrovnik, Croatia, 2011.
- [26] Y. S. Li, Y. Chen, Y. N. Hu et al., "Strategy of computed tomography sinogram inpainting based on sinusoid-like curve decomposition and eigenvector-guided interpolation," *Journal of the Optical Society of America A*, vol. 29, no. 1, pp. 153–163, 2012.

Research Article

Inverse Estimation of Open Boundary Conditions in the Bohai Sea

Zheng Guo,^{1,2} Anzhou Cao,¹ and Xianqing Lv¹

¹ Laboratory of Physical Oceanography, Ocean University of China, Qingdao 266100, China

² National Marine Data and Information Service, Tianjin 300171, China

Correspondence should be addressed to Xianqing Lv, xqinglv@ouc.edu.cn

Received 26 April 2012; Revised 13 June 2012; Accepted 25 June 2012

Academic Editor: Fatih Yaman

Copyright © 2012 Zheng Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents an algorithm for the estimation of open boundary conditions (OBCs) which force tides in the interior region by an adjoint data assimilation approach. Assuming that OBCs are position dependent, OBCs can be approximated by linear interpolation among values at certain independent points (IPs). Twin experiments are performed to examine the sensitivity of the model to the IP distribution and interpolation radius. It is proved that the prescribed OBCs can be well recovered with appropriate number of IP and interpolation radius. In the Bohai Sea model domain with horizontal resolution of $10' \times 10'$, the appropriate number of IP is 3 and the interpolation radius is $60'$. In the practical experiment, the M_2 constituent in the Bohai Sea is simulated by assimilating the T/P data and tidal gauge data. The mean absolute errors in amplitude and phase are 5.0 cm and 5.7° , respectively, and the cochart obtained shows the character of M_2 constituent in the Bohai Sea.

1. Introduction

Open boundary conditions (OBCs) are crucial for the representation of tidal processes in the regional ocean model. Traditionally, they can be obtained from global tidal estimates, larger scale models, or extrapolation from available observations. However, experience is required when the methods mentioned above are used. In addition, OBCs obtained need to be adjusted by comparison between model results and observations, which is time consuming.

The adjoint method is a powerful tool for parameter estimation. Inverse estimation of OBCs in meteorology and oceanography dates back to Sasaki et al. [1]. Shulman et al. [2], Nguyen [3], Das and Lardner [4, 5], Lardner et al. [6], Seiler [7], and Zhang et al. [8, 9] made some studies on optimizing parameters or estimating OBCs in different models over the past few decades.

Efforts also have been made with regard to inverse estimation of OBCs in the China Seas. He et al. [10] investigated the shallow water tidal constituents in the Bohai and Yellow Sea by assimilating T/P altimeter data with the adjoint method. Han et al. [11] employed a 2-dimensional nonlinear numerical POM to describe tide in the China Seas, in which the OBCs and other unknown internal model parameters were inversed in the meanwhile. With the help of a 3D numerical barotropic adjoint tidal model, Zhang and Lu [12] studied the estimation of OBCs more deeply and simulated the 3D M_2 tides and tidal currents in the Bohai and Yellow Seas by assimilating the satellite altimetry data.

In the above-mentioned studies about inverse estimation of OBCs, all grid points on the open boundary are inverted independently. In order to reduce the number of control variables, amplitude and phase of each tidal constituent at a given open boundary grid were taken as a quadratic polynomial based on the results of Schwiderski's global tidal model, which decreased the control variables to six for each tidal constituent [9]. Similar to what had been done in Das and Lardner [4, 5], we assume several independent points (IPs) on the open boundary. Values at the IPs can be optimized by correction equation, and those at other grid points on the open boundary are determined by linearly interpolating the values at IPs.

2. Optimization of Independent Points in the OBCs and Interpolation Scheme

We use the same adjoint tidal model as in Lu and Zhang [13].

In this study, the amplitude and phase at IPs are transformed into Fourier coefficients. The correction of the independent Fourier coefficients is as follows:

$$\begin{aligned} K(a_{jj} - a'_{jj}) + \sum_j \left[W_{j,jj} \sum_{i=1}^I T^i \cos(\omega i \Delta t) \right] &= 0, \\ K(b_{jj} - b'_{jj}) + \sum_j \left[W_{j,jj} \sum_{i=1}^I T^i \sin(\omega i \Delta t) \right] &= 0, \end{aligned} \quad (2.1)$$

where jj is the index of IPs and j is the index of grids on the entire open boundary. a_{jj} and b_{jj} are optimized values of Fourier coefficients at IPs, while a'_{jj} and b'_{jj} are prior values. $W_{j,jj}$ is the weight of linear interpolation. K is the undetermined constant. T^i varies with specific conditions, that is,

$$-g\mu_{m,n} / \Delta x, ((m, n) \text{ is on the left of the computing domain});$$

$$g\mu_{m-1,n} / \Delta x, ((m, n) \text{ is on the right of the computing domain});$$

$$-g\mu_{m,n} / \Delta y, ((m, n) \text{ is below the computing domain});$$

$$g\mu_{m,n-1} / \Delta y, ((m, n) \text{ is above the computing domain}).$$

Fourier coefficients at other grid points obtained by interpolation are denoted by A and B

$$\begin{aligned} A_j &= \sum_{jj} W_{j,jj} a_{jj}, \\ B_j &= \sum_{jj} W_{j,jj} b_{jj}, \end{aligned} \quad (2.2)$$

where

$$W_{j,jj} = \frac{w_{j,jj}}{\sum_{jj} w_{j,jj}}, \quad (2.3)$$

where $w_{j,jj}$ is the weight coefficient in the Cressman form [14]

$$w_{j,jj} = \frac{R^2 - r_{j,jj}^2}{R^2 + r_{j,jj}^2}, \quad (2.4)$$

where R is the interpolation radius and $r_{j,jj}$ is the distance from grid point j to jj on the boundary.

3. Twin Experiments

In this study, twin experiments are implemented in the Bohai Sea (Figure 1) from 117.5°E to 122.5°E and from 37°N to 41°N. The horizontal resolution of this model is $10' \times 10'$. The open boundary is installed on the right edge. The bottom friction coefficients are taken as constant (1.2×10^{-3}). The iteration process of twin experiment (TE) is designed as follow: (1) Run the original dynamic forward model with prescribed OBCs. Take the simulated results at the grid points on T/P satellite tracks and tidal gauges as the “observations.” (2) A value is assigned to each IP (the initial guess is taken as zero here) and values of other grid points on the open boundary are obtained by interpolation. The forward simulation is performed under this condition, which gives the values of state variables. (3) The differences between simulated results and observations serve as the external force of the adjoint model. And the adjoint variables in a period of M_2 constituent are calculated through backward integration of the adjoint equations. (4) Update OBCs by correction equation and interpolation. Repeat steps (2)–(4) until some stopping criterion is satisfied. As the OBCs are increasingly optimized, the differences between simulated results and observations will decrease. Meanwhile, differences between the prescribed OBCs and inverted ones are usually decreased as well.

The iteration will be terminated once a stopping criterion is met. The criterion could be that the last two values of the cost function are sufficiently close, the magnitude of the gradient is sufficiently small, the discrepancy between the updated and old parameters is sufficiently small, or a combination of these. In this work, the criterion is that the number of iteration steps is equal to 100 exactly in both twin experiments and practical experiment.

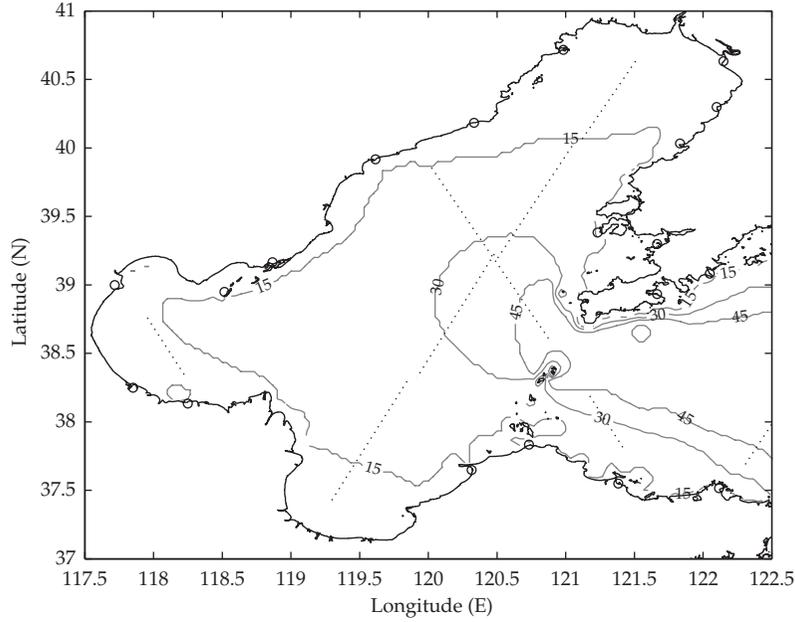


Figure 1: The bathymetry map of the Bohai Sea and positions of T/P altimeter tracks (“.”) and tidal gauges (“o”).

3.1. Influence of the Interpolation Radius on Inversion Results

The model is sensitive to interpolation radius when the algorithm presented in this paper is implemented. In this part, TE1~TE4 are conducted with prescribed OBCs (Figure 2), in which 3 IPs are uniformly distributed (Figure 4(a)) and the interpolation radius is varying, that is, $R = 60'$, $R = 80'$, $R = 100'$, and $R = 120'$. The results are shown as follow.

The decreasing process of cost function value is an important criterion to access the inversion. The ratio of the cost function and its initial value is used to describe this process. The inverted OBCs, cost function, and gradient are shown in Figures 2 and 3, respectively. From Figure 3, we can find that after 100 iteration steps, the cost function and gradient can reach 1% or even 0.1% and $10^{-3} \sim 10^{-5}$ of their initial value. From Figure 2, it can be concluded that by the use of IP scenarios, the prescribed OBCs can be inverted successfully by assimilating the “observations.” In addition, we can find that when R is larger than the maximum distance between adjacent IPs (the cases of $R = 80'$, $R = 100'$ and $R = 120'$), the inversion results may have strange values between adjacent IPs which is related to the interpolation in the Cressman form.

In this paper, mean absolute error (MAE) is used to test the simulated results. MAE is defined as follow:

$$\Delta A = \frac{\sum_{i=1}^n |A_{\text{obs}} - A_{\text{simu}}|}{n}, \quad (3.1)$$

where A is amplitude or phase, ΔA is the MAE of A , n is the total number of observations, and A_{obs} and A_{simu} represent the observation and simulated results at observation locations, respectively.

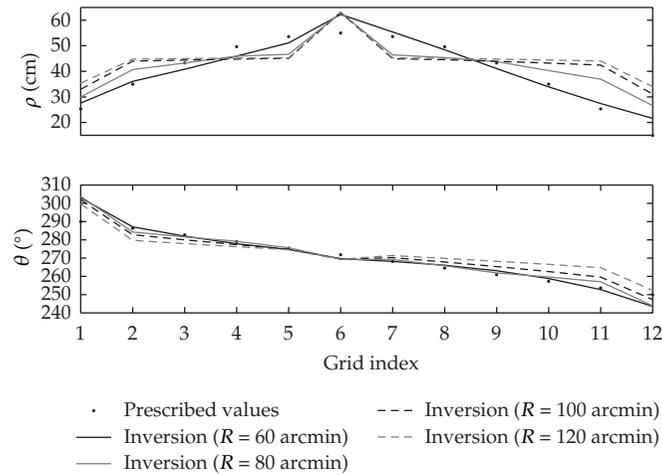


Figure 2: Prescribed OBCs and inverted ones corresponding to different interpolation radius.

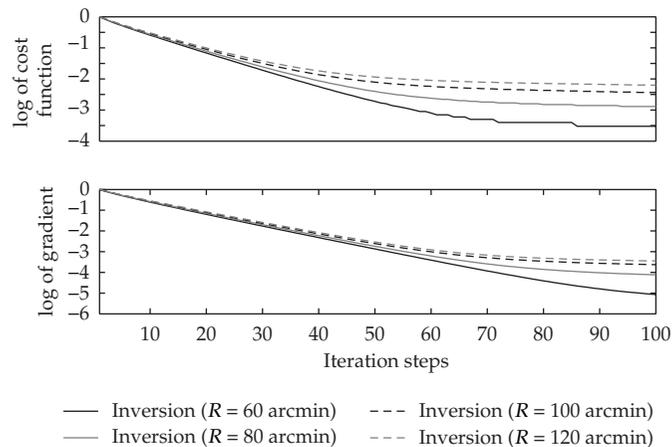


Figure 3: Cost function and gradient in TE1~TE4.

Considering the MAEs shown in Table 1, we can conclude that better inversion results can be obtained when the interpolation radius is equal to the maximum distance between adjacent IPs. Furthermore, the difference increases with the growth of interpolation radius.

3.2. Influence of IP Distribution on Inversion Results

The number and distribution of IPs also have an important influence on the inverse estimation. To explore the influence, three strategies are designed, that is, 3 IPs in strategy (A), 5 IPs in strategy (B), and 7 IPs in strategy (C), and applied to TE5~TE7 where R is equal to the maximum distance between adjacent IPs, respectively. Figure 5 displays the prescribed OBCs and inversion results of TE5~TE7.

Cost function and gradient are shown in Figure 6 from which we can find that after 100 iteration steps, the cost function and gradient can reach 1% or even 0.1% and $10^{-3} \sim 10^{-5}$ of

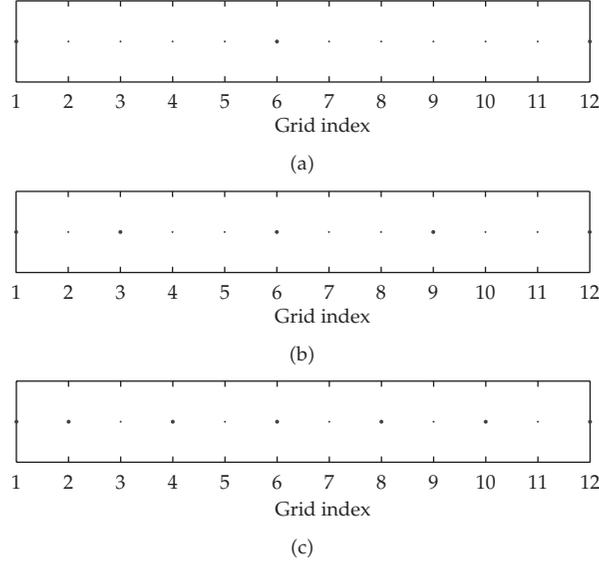


Figure 4: IP distribution of the 3 strategies (large and small dots denote IPs and the other grid points on the open boundary, resp).

Table 1: Differences between prescribed OBCs and inverted ones.

Interpolation radius	ΔH (cm)	Δg ($^{\circ}$)
$R = 60'$	2.9	2.6
$R = 80'$	5.8	2.9
$R = 100'$	7.9	3.9
$R = 120'$	8.7	5.5

their initial value, respectively. Considering the inverted OBCs shown in Figure 5 and MAEs between prescribed OBCs and inverted ones shown in Table 2, we can conclude that the OBCs can be inverted successfully by using IP scenarios, which demonstrates the reasonability and feasibility of the model.

Though the prescribed OBCs can be inverted by using all these 3 strategies, strategy (A) with 3 IPs and R equal to $60'$ leads to better results (smaller values of MAE, cost function, and gradient).

4. Practical Experiments

According to the conclusion of Sections 3.1 and 3.2, in the practical experiment, only strategy (A) is used to estimate the OBCs and simulate the M_2 constituent in the Bohai Sea. Tidal gauge data are used during the assimilation, while T/P data are used to test the simulated results. The bottom friction coefficients are similar to those in Lu and Zhang [13].

The MAEs of amplitude and phase are 5.0 cm and 5.7° , respectively. Cochart based on the model results shows the characteristic of M_2 constituent in the Bohai Sea (Figure 7). There are two amphidromic points in the Bohai Sea, of which one is near Qinhuangdao and the other is near the Yellow River delta.

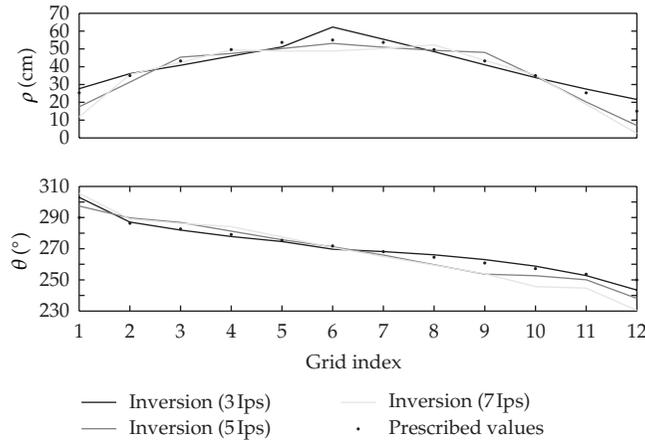


Figure 5: Prescribed OBCs and inverted ones corresponding to different strategies.

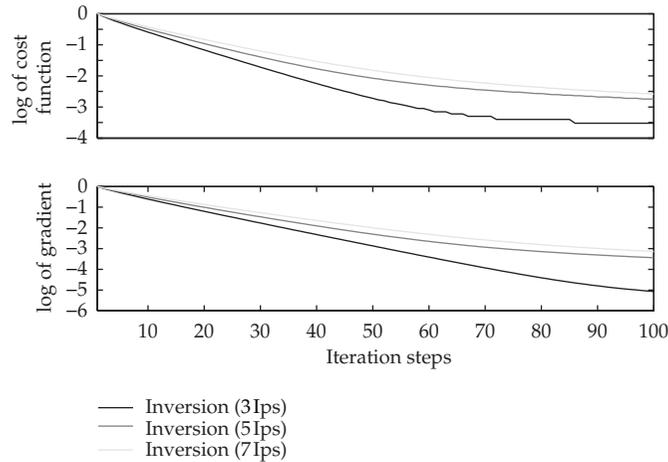


Figure 6: Cost function and gradient in TE5~TE7.

5. Summary

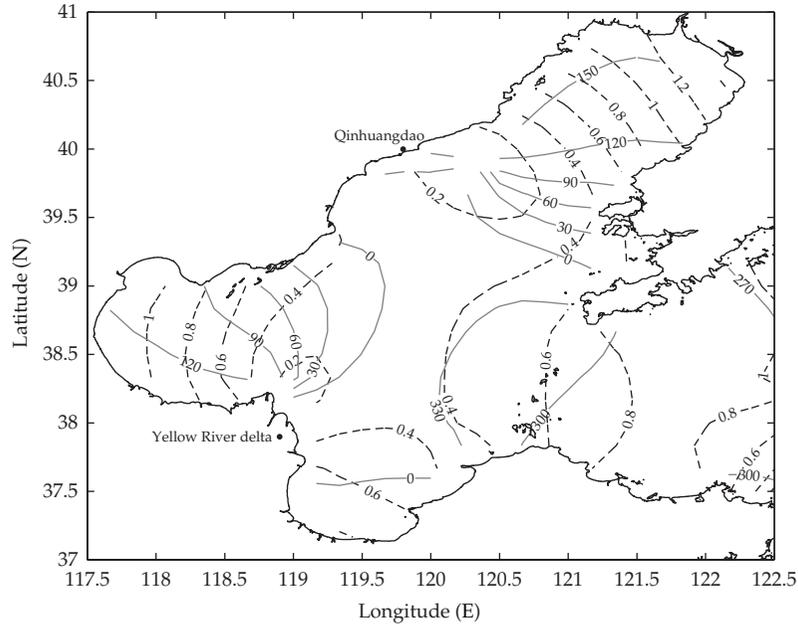
OBCs have crucial impacts on the representation of tidal processes in regional ocean models. They can be inverted by the adjoint method. In this study, OBCs are assumed to be position-dependent and therefore can be approximated by linear interpolation among certain nodal values. In this way, we reduce the number of control variables.

Twin experiments are performed to verify the feasibility of the method and examine the sensitivity of the model to the number of IPs and interpolation radius. It is proved that the prescribed OBCs can be well recovered with appropriate number of independent points and interpolation radius. In the Bohai Sea model domain with horizontal resolution of $10' \times 10'$, the appropriate number of IPs is 3 and the interpolation radius is $60'$.

Based on the strategy of appropriate number of IPs and interpolation radius obtained in the twin experiment, the M_2 constituent in the Bohai Sea is simulated by assimilating the T/P data and tidal gauge data. The MAEs of amplitude and phase are 5.0 cm and 5.7° ,

Table 2: Differences between prescribed OBCs and inverted ones.

Strategy	ΔH (cm)	Δg (°)
(A), 3 IPs, $R = 60'$	2.9	2.6
(B), 5 IPs, $R = 30'$	3.6	4.4
(C), 7 IPs, $R = 20'$	4.3	7.1

**Figure 7:** Cotidal chart obtained in the practical experiment using strategy (A) (the dashed line denotes cotidal line (meter) and the solid line denotes cophase line (degree)).

respectively, and the cochart obtained can describe the character of M_2 constituent in the Bohai Sea.

Used to more general and open domain, the IP strategy is still feasible and reasonable; however, the results will vary with specific conditions while two conclusions stay unchanged. On one hand, under the situation that IPs are distributed evenly, better inversion results can be obtained with interpolation radius equal to the distance between adjacent IPs. On the other hand, the IP number should be chosen carefully, avoiding a too large one. So when this method is used to a more general and open domain, only several experiments are needed in order to obtain the best simulated results.

Acknowledgments

This work is supported by the State Ministry of Science and Technology of China through Grant 2008AA09A402, the National Nature Science Foundation of China through Grant 41076006, and the Ministry of Education's 111 Project through Grant B07036.

References

- [1] Y. Sasaki, P. Gu, and L. Yan, "A fundamental study of the numerical prediction based on the variational principle," *Journal of the Meteorological Society of Japan*, vol. 33, pp. 262–275, 1955.
- [2] I. Shulman, J. K. Lewis, A. F. Blumberg, and B. N. Kim, "Optimized boundary conditions and data assimilation with application to the M2 tide in the yellow sea," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 4, pp. 1066–1071, 1998.
- [3] N. T. Nguyen, "Adjoint optimization of one-dimensional hyperbolic equations with constrained periodic boundary conditions," *Computer Methods in Applied Mechanics and Engineering*, vol. 197, no. 51-52, pp. 4683–4691, 2008.
- [4] S. K. Das and R. W. Lardner, "On the estimation of parameters of hydraulic models by assimilation of periodic tidal data," *Journal of Geophysical Research*, vol. 96, no. 8, pp. 15187–15196, 1991.
- [5] S. K. Das and R. W. Lardner, "Variational parameter estimation for a two-dimensional numerical tidal model," *International Journal for Numerical Methods in Fluids*, vol. 15, no. 3, pp. 313–327, 1992.
- [6] R. W. Lardner, A. H. Al-Rabeh, and N. Gunay, "Optimal estimation of parameters for a two-dimensional hydrodynamical model of the Arabian Gulf," *Journal of Geophysical Research*, vol. 98, no. 10, pp. 18229–18242, 1993.
- [7] U. Seiler, "Estimation of open boundary conditions with the adjoint method," *Journal of Geophysical Research*, vol. 98, no. 12, pp. 22855–22870, 1993.
- [8] A. Zhang, B. B. Parker, and E. Wei, "Assimilation of water level data into a coastal hydrodynamic model by an adjoint optimal technique," *Continental Shelf Research*, vol. 22, no. 14, pp. 1909–1934, 2002.
- [9] A. Zhang, E. Wei, and B. B. Parker, "Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique," *Continental Shelf Research*, vol. 23, no. 11–13, pp. 1055–1070, 2003.
- [10] Y. He, X. Lu, Z. Qiu, and J. Zhao, "Shallow water tidal constituents in the Bohai Sea and the Yellow Sea from a numerical adjoint model with TOPEX/POSEIDON altimeter data," *Continental Shelf Research*, vol. 24, no. 13-14, pp. 1521–1529, 2004.
- [11] G. Han, W. Li, Z. He, K. Liu, and J. Ma, "Assimilated tidal results of tide gauge and TOPEX/POSEIDON data over the China Seas using a variational adjoint approach with a nonlinear numerical model," *Advances in Atmospheric Sciences*, vol. 23, no. 3, pp. 449–460, 2006.
- [12] J. Zhang and X. Lu, "Inversion of three-dimensional tidal currents in marginal seas by assimilating satellite altimetry," *Computer Methods in Applied Mechanics and Engineering*, vol. 199, no. 49–52, pp. 3125–3136, 2010.
- [13] X. Lu and J. Zhang, "Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method," *Continental Shelf Research*, vol. 26, no. 16, pp. 1905–1923, 2006.
- [14] G. P. Cressman, "An operational objective analysis system," *Monthly Weather Review*, vol. 87, pp. 367–374, 1959.