

**OPTIMAL NUMERICAL METHODS FOR INVERSE HEAT CONDUCTION AND
INVERSE DESIGN SOLIDIFICATION PROBLEMS**

By

KEI OKAMOTO

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

WASHINGTON STATE UNIVERSITY
School of Mechanical and Materials Engineering

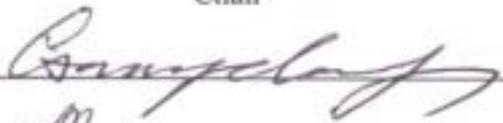
DECEMBER 2005

To the Faculty of Washington State University:

The members of the Committee appointed to examine the dissertation of KEI OKAMOTO find it satisfactory and recommend that it be accepted.



Chair





ACKNOWLEDGMENT

I am grateful for the efforts of Professor Ben Li, whose direction, support and encouragement have made this possible. I would like to acknowledge the valuable comments and advice by Professor Ariyawansa and Professor Cheng. The vital work for the experiment of spray cooling was conducted by Mr. G. Sovar and Isothermal Research Systems (ISR). Special thanks are due to Dr. Shu for the picture (Figure 2.3) of the solidification process and the assistance of the numerical calculations of the solidification problems. I would like to thank my coworkers at Washington State University for their helpful discussion and encouragements. Finally, many thanks to my parents who have continued to be enthusiastic about what I do.

This work was sponsored by NASA (Grant #: NAG8-1697).

OPTIMAL NUMERICAL METHODS FOR INVERSE HEAT TRANSFER AND INVERSE DESIGN SOLIDIFICATION PROBLEMS

Abstract

by Kei Okamoto, Ph.D.
Washington State University
DECEMBER 2005

Chair: Ben Q. Li

The optimization methods for the inverse heat conduction and solidification problems are discussed. Three different methods, the Tikhonov regularization method, the singular value decomposition (SVD) method, and the Levenberg-Marquardt method, are discussed and their performance is assessed comparatively in the inverse heat conduction problems. Several schemes for choosing the optimal regularization parameters are also discussed. These schemes include the maximum likelihood method (ML), the ordinary cross-validation method (OCV), the generalized cross-validation method (GCV), the discrepancy principle (DP), and the L-curve method. 2-D steady-state heat conduction problems are used for the case studies. Parameter estimation and function estimation for the optimal solution are also discussed and compared using 1-D transient heat conduction problems. In the inverse design solidification problems, on the other hand, the regularization method along with the L-curve method is discussed. The design algorithm is applied to determine the appropriate boundary heat flux distribution to obtain prescribed solid-liquid interfaces in a 2-D cavity. A new finite difference scheme for determining the sensitivity coefficients is proposed in the inverse steady-state solidification problems. A sequential method and a whole time-domain method are used and evaluated in the inverse design of solidification processes.

Based on the cases studied for the inverse heat conduction problems, the regularization method and the SVD method are comparative with the Levenberg-Marquardt method with a trust-region strategy. It is also found that the discrepancy principle (DP) gives the best estimate of choosing the regularization parameters. As for the inverse design solidification problems, the L-curve based regularization method is reasonably accurate to control the solid-liquid interfaces. The proposed finite difference scheme is found to be superior to the conventional finite difference scheme in determination of the sensitivity coefficients for the inverse steady-state solidification problems. The results of the inverse design of solidification processes show that the sequential method is comparative to the whole time-domain method if the diffusion time of the heat flux is carefully considered.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iii
ABSTRACT.....	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
NOMENCLATURE	xvii
CHAPTER	
1. INTRODUCTION	1
1.1 BACKGROUND.....	1
1.2 LITERATURE REVIEW.....	3
1.3 OBJECTIVES.....	6
1.4 OVERVIEW OF THE THESIS.....	7
2. DIRECT PROBLEMS.....	8
2.1 HEAT CONDUCTION PROBLEMS.....	8
2.1.1 Problem formulation.....	8
2.1.2 Problem solution.....	9
2.2 WELL-POSED AND ILL-POSED PROBLEMS.....	9
2.2.1 Well-posedness.....	10
2.2.2 Ill-posedness.....	11
2.2.3 Inverse problems.....	12

2.3 SOLIDIFICATION PROBLEMS.....	13
2.3.1 Problem formulation.....	13
2.3.2 Problem solution.....	14
2.3.3 Validation of the direct solidification problem.....	17
3. INVERSE PROBLEMS	20
3.1 INTRODUCTION.....	20
3.2 INVERSE HEAT CONDUCTION PROBLEMS.....	21
3.2.1 The regularization method.....	22
3.2.1.1 Maximum likelihood method.....	24
3.2.1.2 Ordinary cross-validation (OCV) and generalized cross-validation methods (GCV).....	26
3.2.1.3 The L-curve method.....	28
3.2.1.4 Discrepancy principle (DP).....	28
3.2.1.5 Yagola's discussion.....	29
3.2.2 Singular value decomposition (SVD).....	30
3.2.3 Levenberg-Marquardt method.....	32
3.2.3.1 Ozisik's implementation.....	32
3.2.3.2 More's implementation.....	33
3.3 INVERSE SOLIDIFICATION PROBLEMS.....	36
3.3.1 The regularization method.....	36
3.3.1.1 Steady-state problems.....	36
3.3.1.2 Transient problems.....	38

3.3.2 Finite difference approximation for determining sensitivity coefficients.....	40
3.3.3 The L-curve method.....	43
3.3.4 A Sequential method and a whole domain method.....	43
3.3.5 Scaling.....	44
3.3.6 Piecewise functions.....	45
3.4 SUMMARY OF THE NUMERICAL ALGORITHM.....	46
3.4.1 Inverse heat conduction problems.....	46
3.4.2 Inverse solidification problems.....	48
4. RESULTS AND DISCUSSION.....	52
4.1 INVERSE HEAT TRANSFER PROBLEMS.....	52
4.1.1 Steady-state problems.....	54
4.1.1.1 2-D heat conduction.....	54
4.1.1.2 Axisymmetric over-specified problems.....	55
4.1.2 Transient problems.....	58
4.1.2.1 Example problem 1.....	58
4.1.2.2 Example problem 2.....	66
4.1.3 Heat flux for 3-D spray cooling of electronic components.....	68
4.2 INVERSE DESIGN SOLIDIFICATION PROBLEMS.....	69
4.2.1 Steady-state solidification problems.....	70
4.2.2 Transient solidification problems.....	75
5. CONCLUSIONS.....	135

5.1 INVERSE HEAT CONDUCTION PROBLEMS.....	133
5.2 INVERSE DESIGN SOLIDIFICATION PROBLEMS.....	134
BIBLIOGRAPHY.....	136
APPENDIX.....	141
A. Derivation of Maximum likelihood method	141
B. Derivation of OCV and GCV	144

LIST OF TABLES

Table 4.1.1 Input heat generation and heat flux at $x=1m$	126
Table 4.1.2 The regularization parameters chosen in each method and the corresponding least square error for the 2-D problem.....	126
Table 4.1.3 Behavior of the Levenberg-Marquardt method using the More's implementation for axisymmetric over-specified problems: $m=36, \sigma=0.05$	127
Table 4.1.4 The regularization parameters chosen in each method and the corresponding least square error for the axisymmetric problem.....	128
Table 4.1.5 Sum of square errors in calculated heat flux values in $\sigma=0$ for example problem 1.....	129
Table 4.1.6 Sum of square errors in calculated heat flux values in $\sigma=0.01$ for example problem1.....	129
Table 4.1.7 Sum of square errors in calculated heat flux values in $\sigma=0$ for example problem2.....	130
Table 4.1.8 Sum of square errors in calculated heat flux values in $\sigma=0.01$ for example problem2.....	130
Table 4.2.1 Thermal properties used in the example problem.....	131
Table 4.2.2 Thermal properties of aluminum.....	132

LIST OF FIGURES

Figure 2.1. Illustration of (a) direct and (b) inverse heat conduction problems.....	18
Figure 2.2. Schematic of solidification in a 2-D cavity.....	19
Figure 2.3. Velocity fields at t=40 minutes of (a) experimental measurements and (b) FEM calculation for solidification of succinonitrile (SCN).....	19
Figure 3.1. Illustration of determining optimal α using the L-curve method in the case of (a) the inverse heat conduction problems and (b) the inverse solidification problems.....	46
Figure 4.1.1. Illustration of the 2-D inverse heat conduction problem.....	84
Figure 4.1.2. Inverse calculations for the 2-D heat conduction problems: (a) error indicators vs. optimal regularization parameter, (b) L-curve method for optimal regularization parameter and (c) discrepancy principle determination of the singularity threshold for truncated SVD method.....	86
Figure 4.1.3. Dependence of inversely calculated heat flux distribution on the distribution of sensor locations: (a) randomly spread sensors, and (b) specially organized sensors.....	87
Figure 4.1.4. Schematic of concentric problem with over-specified boundary conditions.....	88
Figure 4.1.5. Inverse calculations for over-specified heat conduction in a concentric geometry: (a) error indicator vs. optimal regularization parameter,	

(b) L-Curve method, and (c) discrepancy principle determination of the singularity threshold for the truncated SVD method.....	89
Figure 4.1.6. Comparison of the inverse calculations by the regularization and truncated SVD methods: (a) percent errors for the Levenberg-Marquardt method, (b) percent errors for the regularization method with an optimal parameter, and (c) Percent errors for the truncated SVD.....	91
Figure 4.1.7. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0$) using the regularization method: (a) Heat flux solution. (b) Percentage errors of the heat flux.....	92
Figure 4.1.8. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0$) using the SVD: (a) Heat flux solution. (b) Percentage errors of the heat flux.....	93
Figure 4.1.9. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0.01$) using the regularization method: (a) Heat flux solution. (b) Percentage errors of the heat flux.....	94
Figure 4.1.10. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0.01$) using the SVD: (a) Heat flux solution. (b) Percentage errors of the heat flux.....	95
Figure 4.1.11. Error indicators for the 200 unknowns.....	96
Figure 4.1.12. Discrepancy principle for the SVD for the 20 unknowns.....	96
Figure 4.1.13. Heat flux solution for the regularization method (Example problem 2, $\sigma=0$).....	97

Figure 4.1.14. Schematic of spray cooling of electronics.....	97
Figure 4.1.15. The FEM mesh used for the analysis of the spray cooling.....	98
Figure 4.1.16. Discrepancy principle determination of the singularity threshold for the truncated SVD method, which is used for the inverse prediction of heat flux distribution in spray cooling of electronics.....	98
Figure 4.1.17. Inverse calculations of heat flux during spray cooling of electronics for different spray angles: (a) 3-D view of the heat flux on the cooling surface of the die and (b) Comparison of calculated and measured temperatures	99
Figure 4.2.1. The deforming FEM quadrilateral linear elements used in the example problem.....	100
Figure 4.2.2. Direct solution of solidification problem: (a) Velocity distribution. (b) Temperature distribution.....	101
Figure 4.2.3. Norm of error of the solidification distance versus the number of iteration for the regularization method with respect to the regularization parameters.....	102
Figure 4.2.4. L-curve plots for the case2.....	102
Figure 4.2.5. Inverse solution of solidification problem for case2: (a) Inversely predicted heat flux distribution. (b) Percentage errors in the interface positions. (c) Velocity distribution and (d) Temperature distribution.....	104
Figure 4.2.6. Prescribed solid-liquid interfaces for case3 and case4.....	105

Figure 4.2.7. Norm of error of the solidification distance versus the number of iteration with respect to the regularization parameters (a) The conventional finite difference scheme (b) The proposed finite difference scheme.....	106
Figure 4.2.8. L-curve for case 3 using the proposed finite difference scheme for the sensitivity coefficient.....	106
Figure 4.2.9. Inverse solution of solidification problem for case3. The interface position is prescribed as a sine curve (Figure 4.24 case3) (a) Inversely predicted heat flux distribution. (b) Percentage errors in interface position. (c) Velocity distribution and (d) Temperature distribution.....	108
Figure 4.2.10. Norm of error of the solidification distance versus the number of iteration with respect to the regularization parameters.....	109
Figure 4.2.11. Temperature distribution by using the heat flux solution at the second iteration of $\alpha=1E-18$	109
Figure 4.2.12. Sensitivity coefficients $J_{i,20}$ at the third iteration of $\alpha=1E-17, 1E-18$	110
Figure 4.2.13. L-curve plots for the case 3 using the proposed finite difference scheme for the sensitivity coefficient.....	110
Figure 4.2.14. Inverse solution of the solidification problem for the case 4. The interface position is prescribed as a sharp sine curve. (a) Inversely predicted heat flux distribution. (b) Error in interface position. (c) Velocity distribution	

and (d) Temperature distribution.....	112
Figure 4.2.15. Prescribed locations for the solid-liquid interface in the case 1 and case2.....	113
Figure 4.2.16. Temperature distributions for the inverse steady solidification problem for (a) the case1 and (b) the case2.....	114
Figure 4.2.17. Velocity distribution at $t=0.6s$ for the direct solidification problem.....	114
Figure 4.2.18. Convergence rate for the whole domain with 9 unknowns in the case 1	115
Figure 4.2.19. L-curve plots for the whole domain with 9 unknowns.	115
Figure 4.2.20. Heat flux solution using the whole domain method with 9 unknowns....	116
Figure 4.2.21. Standard deviation for the whole domain method with 9 unknowns.....	116
Figure 4.2.22. Percent errors of the solidification distance at $t=0.6$ for the whole time-domain with 9 unknowns.....	117
Figure 4.2.23. Convergence rate for the first time domain in the case 1 for the sequential method.....	117
Figure 4.2.24. Convergence rate for the case 1 for the whole domain method.....	118
Figure 4.2.25. L-curve plots for the first time domain in the case 1 for the sequential method.....	118
Figure 4.2.26. L-curve method for the case 1 for the whole time-domain method.....	119
Figure 4.2.27. Optimal heat flux solutions in the case 1 for (a) the sequential method and (b) the whole time-domain method with the piecewise polynomial function.....	120

Figure 4.2.28. Optimal heat flux solution in the case 2 for (a) the sequential method and (b) the whole time-domain method with the piecewise polynomial function.....	121
Figure 4.2.29. Standard deviations σ for the error of the solidification distance for (a) the case 1 and (b) the case 2.....	122
Figure 4.2.30. Percent errors of the solid-liquid distance at $t=0.6$ for the sequential method for the case1.....	123
Figure 4.2.31. Percent errors of the solid-liquid distance at $t=0.6$ for the whole time-domain method for the case2.....	123
Figure 4.2.32. Velocity (left) and temperature (right) distributions at times $t=0.2, 0.4, 0.6$ for the case1.....	124
Figure 4.2.33. Velocity (left) and temperature (right) distributions at times $t=0.2, 0.4, 0.6$ for the case2.....	125

NOMENCLATURE

Mathematical symbols

[] matrix

{ } vector

$\| \|$ norm of a matrix or vector

$| |$ absolute value

tr [] trace of matrix

det [] determinant of matrix

Latin

C=specific heat

d_{ij} =solidification interface distance

\bar{d}_{ij} = ideal solidification interface distance

DP(α)=variance for discrepancy principle

f =trial function

{g}=gravity force

H=latent heat

\hat{i}, \hat{j} =unit vectors of $i^{\text{th}}, j^{\text{th}}$ components

[J]=sensitivity matrix

J_{ij} =sensitivity coefficients

[K] =stiffness matrix

l =number of nodes in the azimuthal direction

M =number of sensors

ML_{variance}(α)=variance of deviation by ML

m =number of nodes in the radial direction

N =number of unknowns

n =normal direction
 P =piecewise function
 p =probability distribution
 $\{q\}$ =heat flux vector
 $\{r\}$ =interior point
 S =area
 $S(\alpha)$ =variance of heat fluxes
 s =singular value
 T = temperature
 \bar{T} = measured temperature
 $T(q_0)$ =temperature at $q=0$
 t =time
 u =non-dimensional temperature
 $\{u\}$ =velocity
 V =volume
 $V_o(\alpha)$ =ordinary cross validation function
 $V(\alpha)$ =generalized cross-validation function
 x,y,z =coordinates
 z =uniformly distributed number

Greek

α =regularization parameter
 β =thermal expansion coefficient
 χ =normally distributed random number
 δ =norm of input error
 ε =input error (measurement error)
 Φ =variable of temperature
 ϕ =circumferential angle
 γ = positive constant

η = positive constant
 κ = thermal conductivity
 λ =perturbation parameter
 μ =damping parameter
 $\{\theta\}$ =shape function
 ρ =non-dimensional radii
 σ =standard deviation
 Ω =domain
 τ =singularity threshold
 Ψ =monotonically increasing function
 ξ = normally distributed random number

Subscripts

1=Dirichlet boundary condition
2=Neumann boundary condition
u=unknowns
analytic=analytic value
computed=computed value
a=inner circular boundary
b=outer circular boundary
trans=transformation
max=maximum value

Superscripts

k=iteration number
T=transpose

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

Inverse problems are encountered in various branches of science and engineering. Engineers, mathematicians, statisticians, and specialists of many other fields are all interested in inverse problems, each with different applications in mind. The applications of the inverse problem include biomedical imaging [Horacek *et al.*, 1997], thermotherapy, metabolism modeling, spline smoothing [Wahba, 1985], electrocardiography [Johnson *et al.*, 1998], and image reconstruction [Demoment, 1989; Galatsanos *et al.*, 1992].

In particular, inverse heat transfer problems arise from thermal system design considerations. These problems are concerned with the estimation of the heat flux and/or temperature distribution on a boundary using the temperature measurements at some points in the interior or with both temperature and flux specified over the same portion of the boundary [Martin *et al.*, 1996]. For example, the direct measurement of heat flux at the inside surface of a wall subjected to fire, at the outer surface of a reentry vehicle, or at the inside surface of combustion chamber is extremely difficult. These problems need to be solved using inverse computational algorithms. While inverse problems are physically feasible, they are mathematically ill-posed, which is manifested by the fact that noises present in the measured data can cause instabilities in the estimated heat fluxes. Proper numerical treatments are needed to overcome these instability problems. This paper discusses three popular methods for the inverse heat transfer calculations: the regularization method, the singular value decomposition method, and the Levenberg-

Marquardt method, as well as the schemes for estimating the optimal parameters for the first two methods.

Solidification is a well-known process, which contains a phase transformation of liquid and solid. Transient heat transfer problems involving solidification are important in many engineering applications. These applications include the making of ice, the freezing of food, and crystal growth. Solidification is also a widely used manufacturing process that is a very economical way to form components. The solution of the solidification process is inherently difficult because the interface between the solid and liquid phases is moving as the latent heat is absorbed or released at the interface. However, knowledge of liquid-solid interface morphology during solidification processing is of paramount importance to the microstructure formation in solidified materials. Because often the solid-liquid interface position is unknown *a priori*, the problem of finding the interface is classified as moving boundary problems. The widespread use of solidification principle in materials processing systems has resulted in both theoretical and experimental studies on the subject. A wide variety of numerical models have been developed for virtually every kind of solidification processing systems. Both the fixed grid and moving grid methods have been used to model the solidification phenomena. The fixed grid method involves the use of enthalpy-based formulation in which the latent heat is factored into an effective heat capacity. The moving grid method, however, tracks the solidification front, that is, the solid-liquid interface continuously by deforming the grids or elements. There are alternatives that involve a level set field variable to mark the solid-liquid interface. Model developed using these techniques have been applied mainly to answer the question concerning the interface position and morphological development for given operating

conditions and specified geometric constraints. However, to assure the quality and reliability of a casting, the ability to control the solid-liquid interface morphology is important. Thus, inverse computational algorithms are used to solve the design solidification processing systems. In practice, a desired solidification microstructure in the final products dictates a certain type of solid-liquid interface front morphology. Thus, question often arises of how the boundary heat flux distribution needs to be specified in order to obtain the desired solid-liquid interface during solidification processing. For this purpose, the regularization method along with the L-curve method is used to solve the inverse design of solidification processes in this paper.

1.2 LITERATURE REVIEW

Various inverse algorithms including the regularization schemes for heat transfer calculations have been summarized in two monographs [Sawaf *et al.*, 1995; Trujillo *et al.*, 1989]. The subject has received continuous interest owing to its wide range of applications including the estimation of thermophysical properties and the estimation of the unknown source functions. The Levenberg-Marquardt method is discussed in detail to solve inverse problems [Sawaf *et al.*, 1995; Kim *et al.*, 2003]. More (1977) proposed a robust and efficient implementation of the Levenberg-Marquardt method using a trust-region framework. Martin and Dulikravich (1996) used the regularization method combined with the boundary element method (BEM), but the procedure to choose the regularization parameter was limited to special cases. Trujillo and Busby (1989) used the regularization method and the generalized cross-validation (GCV) method in their study of inverse heat transfer problems. There are also other methods, such as the conjugate

gradient method (CGM), which are either less general or less easy to implement and are not included in the present study.

One of the crucial aspects associated with the regularization method is the appropriate selection of optimal regularization parameters for inverse estimation. Several approaches have been reported. The maximum likelihood (ML) method was presented by Wahba *et al.* (1985) from the standpoint of the statistical inference [Fitzpatrick *et al.*, 1991]. The ML method was found to yield a regularization parameter somewhat smaller than the optimal regularization parameter [Galatsanos *et al.*, 1992]. The ordinary cross-validation (OCV) method was proposed as another approach to obtain optimal regularization parameters [Allen, 1974; Golub *et al.*, 1979]. In order for the OCV method to remain invariant after transformation, the generalized cross-validation (GCV) method is needed and used in other engineering fields [Yoon *et al.*, 2000]. It has been reported that the GCV method gave poor accuracy in some cases. The L-curve method proposed by Hansen and O'Leary (1993) can be applied to the linear and non-linear inverse problems [Reginska, 1996; Tautenhahn *et al.*, 2003] and was used in the BEM-based inverse elasticity problems [Martin *et al.*, 2003]. One possible limitation of the L-curve method is the difficulty of accurately determining the corner point. The discrepancy principle (DP) was proposed to relate input errors (measurement errors) to the computed parameter [Morozov *et al.*, 1984]. Hollingsworth and Johns (2003) showed that the DP method was reliable when measurement errors were observed correctly. Some of these techniques have been applied in transient inverse heat transfer analyses; their performance, however, for steady-state inverse heat transfer analyses, is not well appreciated. The truncated SVD method has been considered a very useful technique for

inverse calculations, and yet there has not been a systematic approach to select the singularity threshold value, which is critical for the accurate solution.

It is, however, conceivable that inverse free boundary problems can be more complex in general than inverse heat conduction problems. One of the applications of the inverse problems for solidification processing systems is the determination of the boundary condition by utilizing either experimental measurements (inverse solidification problems) or prescribed conditions (optimal design solidification problems). The inverse solidification problems have been attempted in literature. Krishnan and Sharma (1996) found casting/mold interfacial heat transfer coefficients using experimental temperature measurements for casting solidification problems. They used the finite difference method (FDM) combined with the Beck's method for their inverse algorithm. O'Mahoney and Browne (2000) combined their inverse algorithm with the integral-derivative method to find interfacial heat transfer coefficients using temperature measurements. Xu and Naterer (2001) found temperature distribution using the prescribed solid-liquid interface location and heat fluxes. Hale *et al.* (2000) used the Global Time Method (GTM) for their inverse algorithm to find heat flux distribution in the boundary of both liquid and solid phase using the prescribed temperature and heat flux in the liquid-solid interface. In their approach, the solid and liquid regions are treated as two distinct inverse heat transfer problems. Dulikravich *et al.* (2003) found optimal magnetic fields on the boundary by specifying desired magnetic field lines and temperature distribution. Zabaras *et al.* (1990, 1993, 1995) have studied the inverse solidification problems both with and without fluid flow being considered. They reported various algorithms including the Beck's method, the steepest descent method (SDM), and the conjugate gradient method (CGM). Most of

the work up to date has been limited to conduction mode only and few have considered the fluid flow effects. In addition, the solid-liquid interfaces to control in the inverse solidification problems are limited to the simple shape such as a straight line.

1.3 OBJECTIVES

The objectives of this research consist of two parts. One is solving inverse heat transfer problems. The other one is solving inverse solidification problems.

The regularization method, the singular value decomposition method (SVD), and the Levenberg-Marquardt method are used and evaluated in the inverse heat transfer problems. We also consider the schemes for estimating the optimal parameters for the regularization method and the SVD. The discrepancy principle is used for determining the truncated threshold for the SVD method. The selection of optimal parameters is considered for Tikhonov regularization solutions of inverse heat transfer problems. For this purpose, the five popular methods including the maximum likelihood (ML), the ordinary cross-validation (OCV), the generalized cross-validation (GCV), and the L-curve methods, as well as the discrepancy principle (DP) are evaluated in detail. The testing cases include 1-D and 2-D inverse steady-state heat conduction problems where analytic solutions are available. Both the Tikhonov regularization method and the truncated SVD method, with optimal parameters determined, are compared with the Levenberg-Marquardt method. In addition, a parameter estimation and a function estimation approach are also used and assessed by using 1-D inverse transient heat conduction problems. As a last example of the inverse heat conduction problems, the

inverse algorithm is applied to estimate the heat flux experienced due to the spray cooling of a 3-D microelectronic component with an embedded heating source.

In the inverse solidification problems, on the other hand, we use the regularization method along with the L-curve method. The algorithm is applied to determine the boundary heat flux distribution for prescribed solid-liquid interfaces in a 2-D cavity for both steady and transient problems. We specify not only a straight line but also sine functions for the solid-liquid interfaces. We also discuss the finite difference scheme for computing sensitivity coefficients by using the inverse steady-state solidification problem. In addition, a sequential method and a whole domain method are used and evaluated for the transient problems.

1.4 OVERVIEW OF THE THESIS

In this thesis, the direct and inverse algorithms of heat conduction problems and solidification problems are presented.

In Chapter 2, the direct problems of the heat transfer and solidification processes are formulated and discussed. The concept of the well-posed and ill-posed problems is also discussed. Chapter 3 describes the inverse calculation methods of the heat transfer and solidification problems. In Chapter 4, the formulation of the cases studied is presented and the results are discussed. Finally, the results of the present study are summarized in Chapter 5.

CHAPTER TWO

DIRECT PROBLEMS

In this chapter, we first formulate heat conduction problems and address the solution to facilitate subsequent discussions. We then revise and discuss the concept of direct and inverse heat conduction problems in the context of the well-posedness and ill-posedness of problems. We formulate solidification (phase-change) problems and address the solution of the problems. Validation of our FEM code is also discussed using the experimental measurements, which were previously conducted.

2.1 HEAT CONDUCTION PROBLEMS

Heat conduction is the mode of heat transfer in which energy exchange takes place in solids or fluids from the region of high temperature to the region of low temperature. The direct heat conduction problem is mainly concerned with the determination of the temperature distribution within solids. In this section, we present the problem formulation and the solution of the heat conduction problems.

2.1.1 Problem formulation

For *direct* heat conduction problems, the temperature inside a domain Ω is sought using conditions prescribed along the *entire* boundary (see Figure 2.1a). Mathematically, a direct heat conduction problem is formulated as follows:

$$\rho C \frac{\partial T}{\partial t} = \kappa \nabla^2 T + Q \quad \in \Omega \quad (2.1.1)$$

In the above equation, ρ is the density, C is the specific heat, T is the temperature, t is time, κ is the thermal conductivity, Q is the heat generation rate. The above equation is subject to the following boundary conditions:

$$T = T_1 \in \partial\Omega_1 \quad (2.1.2)$$

$$-\{\mathbf{n}\} \cdot \kappa \nabla T = q_2 \in \partial\Omega_2 \quad (2.1.3)$$

Here it is important to note that the temperature or flux can be prescribed on either of the boundaries but not both.

2.1.2 Problem solution

Many existing techniques can be employed to solve the direct heat conduction problems. In this paper, the Galerkin finite element method is applied, which discretizes the governing equations in a matrix form:

$$[\mathbf{N}_T][\dot{\mathbf{T}}] + [\mathbf{K}][\mathbf{T}] = [\mathbf{F}] \quad (2.1.4)$$

where the matrix elements are calculated by

$$[\mathbf{N}_T] = \int_{\Omega} \rho C \{\theta\} \{\theta\}^T dV, \quad [\mathbf{K}] = \int_{\Omega} \kappa \nabla \{\theta\} \cdot \nabla \{\theta\}^T dV$$

$$[\mathbf{F}] = \int_{\partial\Omega} \{\mathbf{n}\} \cdot \kappa \nabla T \{\theta\} dS + \int_{\Omega} Q \{\theta\} dV \quad (2.1.5)$$

2.2 WELL-POSED AND ILL-POSED PROBLEMS

The concept of well-posed and ill-posed problems is crucial to formulate and solve field problems. In this section, the well-posed and ill-posed problems are revised using heat conduction problems.

2.2.1 Well-posedness

According to the theory of partial differential equations (PDE), a problem is well-posed if and only if all of these three conditions are satisfied: (1) its solution exists, (2) the equation has one and only one solution, and (3) a small change in the data (such as prescribed boundary conditions, source strengths, coefficients in the PDE, etc.) produces only a small change in the solution [Nagle *et al.*, 1993]. The second condition requires that the solution, if it exists, is unique; the third requires that the solution is stable. The existence of the solution can be easily proved by noticing that a Green's function is constructed, and the solution can be obtained by the superimposition principle. To see the uniqueness of the solution, we consider a steady-state heat conduction problem. We construct two solutions, T^1 and T^2 , which each satisfies the above equation and the boundary conditions. With the new variable, $\Phi = T^1 - T^2$, the governing equations becomes,

$$0 = \nabla \cdot \kappa \nabla \Phi \quad (2.2.1)$$

$$\Phi = 0 \quad \in \partial\Omega_1 \quad (2.2.2)$$

$$-\{\mathbf{n}\} \cdot \kappa \nabla \Phi = 0 \quad \in \partial\Omega_2 \quad (2.2.3)$$

Integration of Eq. (2.2.1) with respect to Φ , followed by the use of Green's theorem, yields,

$$\begin{aligned} 0 &= \int_{\Omega} \Phi \nabla \cdot \kappa \nabla \Phi dV = \int_{\Omega} \kappa \nabla \cdot (\Phi \nabla \Phi) - \kappa (\nabla \Phi)^2 dV \\ &= \int_{\partial\Omega} \kappa \{\mathbf{n}\} \cdot (\Phi \nabla \Phi) dS - \int_{\partial\Omega} \kappa (\nabla \Phi)^2 dV \end{aligned} \quad (2.2.4)$$

Thus, along with Eqs. (2.2.2) and (2.2.3), immediately leads to the following result:

$$\int_{\partial\Omega} \kappa (\nabla \Phi)^2 dV = \int_{\partial\Omega_1} \kappa \{\mathbf{n}\} \cdot (\Phi \nabla \Phi) dS + \int_{\partial\Omega_2} \kappa \{\mathbf{n}\} \cdot (\Phi \nabla \Phi) dS = 0 \quad (2.2.5)$$

Thus, $\Phi=0$ or $T^1 = T^2$, that is, the above problem has a unique solution. Also, any small change in conditions (source, boundary conditions, etc.) will not affect the above result. Therefore, the direct heat conduction problem is well-posed. Note that if the whole boundary is prescribed with flux, the solution is also unique, albeit an arbitrary additive constant. The constant may be determined by requiring that the integration of boundary flux be zero. The problem is still well-posed.

2.2.2 Ill-posedness

An ill-posed problem violates any or all conditions required for well-posedness. Consider the above problem but with both temperature and fluxes chosen on part of the boundary, say $\partial\Omega_1$, but no conditions on $\partial\Omega_2$. In this case, the second integration involving $\partial\Omega_2$ in Eq. (2.2.5) is not necessarily zero, and thus the solution may not be unique. If a solution exists, it may not be continuously dependent on the boundary data; otherwise the stability condition is violated. To illustrate this point, we consider the following heat conduction example with constant κ :

$$0 = \nabla \cdot \nabla T \quad \in \Omega = [0, 1] \times [0, 1] \quad (2.2.6)$$

$$T(x,0) = 0 \quad \text{and} \quad -\mathbf{n}_y \cdot \nabla T(x,0) = \delta \sin(x/\delta) \quad (2.2.7)$$

By the principle of separation of variables, the following solution may be obtained, which satisfies the above two equations:

$$T(x, y) = \delta^2 \sin(x/\delta) \sinh(y/\delta) \quad (2.2.8)$$

With $\delta = 0$ or finite, $T(x, y)$ is well-behaved. However, if $\delta \rightarrow 0^\pm$, the solution is unbounded,

$$\lim_{\delta \rightarrow 0^{\pm}} \delta^2 \sin(x/\delta) \sinh(y/\delta) = \pm\infty \quad (2.2.9)$$

Thus, a very small change in $\partial T(x,0)/\partial y$ leads to an arbitrarily large change in the solution. That is, the solution is unstable. This example illustrates the nature of an ill-posed problem.

2.2.3 Inverse problems

In *inverse* heat transfer problems, temperatures at some interior points are known, while some of the heat fluxes are unknown and need to be part of the solution (See Figure 2.1b). A statement of inverse heat transfer problems involving the need to determine the boundary heat flux distribution may be made as follows:

$$0 = \nabla \cdot \kappa \nabla T + Q \quad \in \Omega \quad (2.2.10)$$

subject to the following boundary conditions:

$$T = T_1 \quad \in \partial\Omega_1 \quad (2.2.11)$$

$$-\{n\} \cdot \kappa \nabla T = q_2 \quad \in \partial\Omega_2 \quad (2.2.12)$$

and also to the known temperature information at the discrete interior points,

$$T(\{r\}) = \bar{T}_i \quad \{r\} \in \partial\Omega_2 \quad (2.2.13)$$

where q_2 is the unknown heat flux to be determined. The inverse heat conduction problem is ill-posed. To see this, we consider Eq. (2.2.8) again, with $T(x,0) = 0$, $\partial T(x,0) = 2$ and heat flux is to be sought on other boundaries. If further a value of $T(0.5,0.5) = 2$ is measured, then we can show that $\partial T(0,y) = \partial T(1,y) = \text{constant}$ would be a solution, and thus the solution is not unique. Beck *et al.* (1985) also show that the solution to inverse heat conduction problems is unstable. The solution to an inverse heat

conduction problem is thus the most probable and most stable solution among all the solutions available. Any inverse algorithm in general would need to incorporate stabilization schemes to obtain the solution to inverse heat conduction problems.

2.3 SOLIDIFICATION PROBLEMS

Solidification problems, also called moving boundary problems, are important in many engineering applications including crystal growth and the solidification of metals in casting. The solution of such problems is inherently difficult because the interface between the solid and liquid phases is moving and its location is not known *a priori*. Many solidification problems have appeared in the literature, but the exact solutions are limited to a number of idealized situations. When exact solutions are not available, numerical methods can be used to solve the solidification problems. In this section, we address the problem formulation of the solidification problems. Then we present the numerical methods of the solidification problems used in the present study.

2.3.1 Problem formulation

Figure 2.2 shows a 2-D cavity for the solidification problem under consideration. The top and bottom walls are thermally insulated. The temperature on the left wall is fixed at a constant temperature above the melting point, while the right wall is subjected to cooling. The melt, which is initially above the melting temperature, starts to solidify as a result of cooling at the right wall. The fluid flow and heat transfer in the system are governed by the continuity equation, the Navier-Stokes equations, and the energy equation. For the melt flow, the standard Boussinesq approximation,

$\rho = \rho_0[1 - \beta(T - T_m)]$, is used. The governing equations for the problem are given as follows:

$$\nabla \cdot \{\mathbf{u}\} = 0 \quad (2.3.1)$$

$$\rho \frac{\partial \{\mathbf{u}\}}{\partial t} + \rho \{\mathbf{u}\} \cdot \nabla \{\mathbf{u}\} = -\nabla p + \mu \nabla^2 \{\mathbf{u}\} - \{\mathbf{g}\} \rho_0 \beta (T - T_m) \quad (2.3.2)$$

$$\rho C \frac{\partial T}{\partial t} + \rho C \{\mathbf{u}\} \cdot \nabla T = k \nabla^2 T \quad (2.3.3)$$

$$\{n\} \cdot k \nabla T|_l - \{n\} \cdot k \nabla T|_s = \rho H \frac{\partial s}{\partial t} \quad (2.3.4)$$

The no slip condition is specified at the walls. The boundary conditions are as follows:

$$\{\mathbf{u}\} = 0 \quad \text{at all boundaries} \quad (2.3.5)$$

$$\frac{\partial T}{\partial y} = 0 \quad \text{at } y=0 \text{ and } y=h \quad (2.3.6)$$

$$T = T_o \quad T_o > T_m \text{ at } x=0 \quad (2.3.7)$$

$$T = T_L \text{ or } q = q_0 \quad T_m > T_L \text{ at } x=l \quad (2.3.8)$$

The solid-liquid interface is to be obtained for the direct problem or the well-posed problem. For the inverse problem, however, the solid-liquid interface shape is specified, and the cooling condition on the right wall is to be obtained.

2.3.2 Problem solution

The governing equations described above along with the boundary conditions are solved using the deforming Galerkin finite element method. The stiffness matrix is obtained by using the Galerkin's method of Weighted Residuals. The formulations and

relevant benchmark tests were detailed in a series of papers published earlier [Shu *et al.*, 2002; Li *et al.*, 2003; Song *et al.*, 2002]. Thus, only a brief summary is given here. The governing equations are recast in integral forms, and the field variables are interpolated using shape functions over the computational domain. With appropriate algebraic manipulations, the following set of equations are obtained:

$$\left(\int_{\Omega} \psi \hat{i} \cdot \nabla \phi^T dV \right) \mathbf{U}_i = -\varepsilon \left(\int_{\Omega} \psi \psi^T dV \right) \mathbf{P} \quad (2.3.9)$$

$$\begin{aligned} & \left(\int_{\Omega} \rho \phi \phi^T dV \right) \frac{d\mathbf{U}_i}{dt} + \left(\int_{\Omega} \phi \rho \mathbf{u} \cdot \nabla \phi^T dV \right) \mathbf{U}_i - \left(\int_{\Omega} \hat{i} \cdot \nabla \phi \psi^T dV \right) \mathbf{P} \\ & + \left(\int_{\Omega} \mu \nabla \phi \cdot \nabla \phi^T dV \right) \mathbf{U}_i + \left(\int_{\Omega} \mu (\hat{i} \cdot \nabla \phi) (\hat{j} \cdot \nabla \phi^T) dV \right) \mathbf{U}_j \\ & + \left(\int_{\Omega} \phi \theta^T \rho \beta g dV \right) (T - T_m) = \int_{\partial\Omega} \mathbf{n} \cdot \boldsymbol{\tau} \cdot \hat{i} \phi dS \end{aligned} \quad (2.3.10)$$

$$\left(\int_{\Omega} \rho C \theta \theta^T dV \right) \frac{d\mathbf{T}}{dt} + \left(\int_{\Omega} \rho C \theta \mathbf{u} \cdot \nabla \theta^T dV \right) \mathbf{T} + \left(\int_{\Omega} k \nabla \theta \cdot \nabla \theta^T dV \right) \mathbf{T} = - \int_{\partial\Omega} q_T \theta dS \quad (2.3.11)$$

Once the form of shape functions ϕ , θ , and ψ for velocity, pressure, and scalars are specified, the integrals defined in the above equations can be expressed in matrix form. Combining the momentum and energy equations into a single matrix equation gives rise to the following element stiffness matrix equation:

$$\begin{bmatrix} \mathbf{M} & 0 \\ 0 & \mathbf{N}_T \end{bmatrix} \begin{bmatrix} \dot{\mathbf{U}} \\ \dot{\mathbf{T}} \end{bmatrix} + \begin{bmatrix} \mathbf{A}(\mathbf{U}) + \mathbf{K} + \frac{1}{\varepsilon} \mathbf{E} \mathbf{M}_p^{-1} \mathbf{E}^T & \mathbf{B}_T \\ 0 & \mathbf{D}_T(\mathbf{U}) + \mathbf{L}_T \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{T} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{G}_T \end{bmatrix} \quad (2.3.12)$$

Note that in constructing the above element matrix equation, the penalty formulation is applied, and \mathbf{P} in the momentum equation is substituted by $\frac{1}{\varepsilon} \mathbf{M}_p^{-1} \mathbf{E}^T \mathbf{U}$. The assembled

global matrix equations are stored in the skyline form and solved using the Gaussian elimination method. The coefficient matrices of Eq. (2.3.12) above are calculated by

$$\begin{aligned}
\mathbf{M}_p &= \int_{\Omega} \psi \psi^T dV, \quad \mathbf{N}_T = \int_{\Omega} \rho C \theta \theta^T dV, \quad \mathbf{M} = \int_{\Omega} k \theta \theta^T dV, \quad \mathbf{E}_i = \int_{\Omega} \hat{i} \cdot \nabla \phi \psi^T dV, \\
\mathbf{L}_T &= \int_{\Omega} \nabla \theta \cdot \nabla \theta^T dV, \quad \mathbf{A}(\mathbf{U}) = \int_{\Omega} \rho \phi \mathbf{u} \cdot \nabla \theta^T dV, \quad \mathbf{D}_T(\mathbf{U}) = \int_{\Omega} \rho C \theta \mathbf{u} \cdot \nabla \theta^T dV, \\
\mathbf{B}_T &= \int_{\Omega} \rho \beta \mathbf{g} \phi \theta^T dV, \quad \mathbf{G}_T = - \int_{\partial \Omega} q_T \theta d\Gamma, \quad \mathbf{F} = \int_{\partial \Omega} \mathbf{n} \cdot \tau \phi d\Gamma, \\
\mathbf{K}_{ij} &= \left(\int_{\Omega} \mu \nabla \phi \cdot \nabla \phi^T dV \right) \delta_{ij} + \int_{\Omega} \mu (\hat{i} \cdot \nabla \phi) (\hat{j} \cdot \nabla \phi^T) dV
\end{aligned}$$

To implement the deforming finite elements to model the dynamic change of the moving interface, i.e., solidification front between the liquid and solid, a quasi-Lagrangian description is adopted. By this method, a region that covers the solidifying liquid and solid is defined, and the nodes within the region are allowed to move in accordance with the interface movement. These additional velocities that result from the mesh movement are added to the velocity field as given in the above equations. The energy balance equation describing the latent heat release and interface change is directly integrated within the context of weighted residuals,

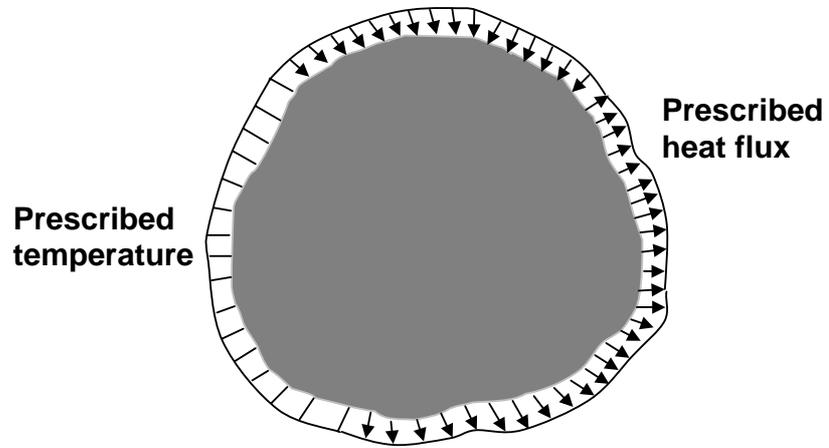
$$\left(\int_{\Omega} k \theta \hat{n} \cdot \nabla \theta^T dV \right) \mathbf{T}_l - \left(\int_{\Omega} k \theta \hat{n} \cdot \nabla \theta^T dV \right) \mathbf{T}_s = - \int_{\partial \Omega} \rho H \frac{\partial s}{\partial t} \theta d\Gamma \quad (2.3.13)$$

which is applied as a surface energy source to the total thermal energy balance equation, and added to the right-hand side of Eq. (2.3.12). The above equation systems are solved iteratively. Separation of the moving interface boundary coordinates from the global finite element solutions for field variables, however, requires the convergence of both moving interface coordinates and field variables in two related loops. The interface

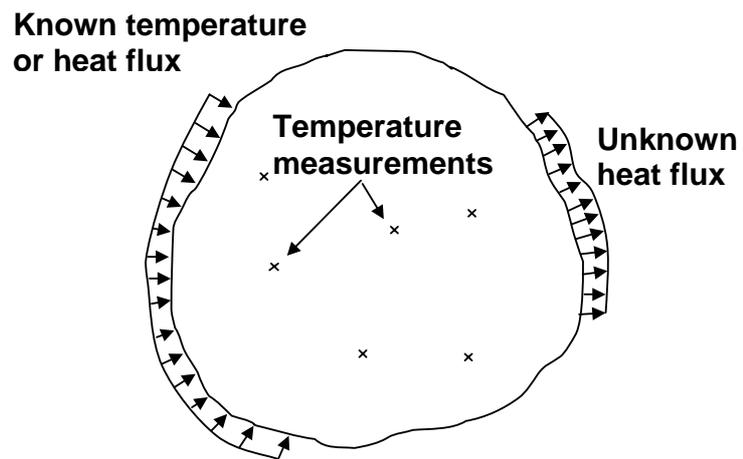
tracking strategy used in the present study involves an iterative procedure that entails applying the energy balance equation along with the interface as a surface source and searching for the interface position coordinates based on each converged field calculations. The updated interface positions are then fed back to the field calculations until both the interface position coordinates and field variables are converged within a preset criterion, which is set at 1×10^{-4} (relative error) for the results presented below.

2.3.3 Validation of the direct solidification problem

Shu (2003) conducted an experimental and numerical study of solidification in a square mold. Figure 2.3 shows the velocity fields at $t=40$ minutes of (a) the experimental measurements and (b) the FEM calculations. As seen in figures, they found good agreements of the experimental measurements and the numerical simulations. Their FEM code is utilized for the design solidification problems. Thus, the FEM calculations of the solidification processes in the present study are considered to be accurate.



(a)



(b)

Figure 2.1. Illustration of (a) direct and (b) inverse heat conduction problems.

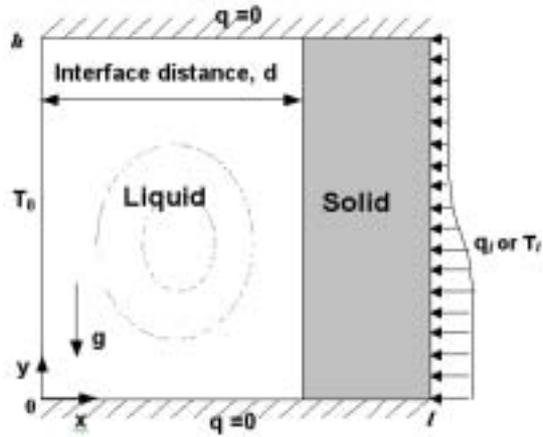


Figure 2.2. Schematic of solidification in a 2-D cavity

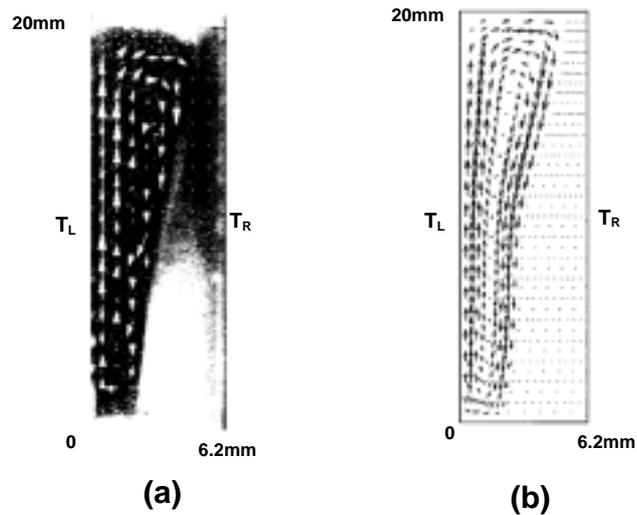


Figure 2.3. Velocity fields at $t=40$ minutes of (a) experimental measurements and (b) FEM calculation for solidification of succinonitrile (SCN). The boundary conditions are $T_L=334.12\text{K}$, $T_R=328.61\text{K}$, $T_m=331.23\text{K}$, and the upper and bottom wall is kept adiabatic. Figures courtesy of Y. Shu (2003)

CHAPTER THREE

INVERSE PROBLEMS

In this chapter, the inverse problems are discussed. We formulate the inverse heat conduction problems and inverse design solidification problems. Various methods are presented to solve the inverse problems. The numerical algorithm is then summarized.

3.1 INTRODUCTION

A common approach to formulating the inverse problems is to seek the distribution of the unknown parameters such that the difference between the observations and calculated parameters at given points is minimized. The least-squares criterion is solidly grounded in statistics, as we show below. Let the discrepancies between model and observation in a general data-fitting problem be denoted by ε , that is,

$$\varepsilon_j = y_j - \phi(x, t_j) \quad (3.1.1)$$

It is often reasonable to assume that the ε_j 's are independent and identically distributed with a certain variance σ^2 and probability density function g_σ . Under this assumption, the likelihood of a particular set of observations $y_j, j=1,2,\dots,m$, given that the actual parameter vector is x , is given by the function,

$$p(y/\sigma) = \prod_{j=1}^m g_\sigma(\varepsilon_j) = \prod_{j=1}^m g_\sigma(y_j - \phi(x, t_j)) \quad (3.1.2)$$

Since we know the value of the observations y_1, y_2, \dots, y_m , the “most likely” value of x is obtained by maximizing $p(y/\sigma)$ with respect to x . The resulting value of x is called the

maximum likelihood estimate of the parameters. When we assume that the discrepancies follow a normal distribution, we have

$$g_{\sigma}(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right) \quad (3.1.3)$$

Substitution in Eq. (3.1.2) yields

$$p(y/x, \sigma) = (2\pi\sigma^2)^{-m/2} \exp\left(-\frac{1}{2} \sum_{j=1}^m \frac{[y_j - \phi(x, t_i)]^2}{\sigma^2}\right) \quad (3.1.4)$$

For any fixed value of the variance σ^2 , it is obvious that p is maximized when the sum of squares is minimized. Thus, the minimization problem may be mathematically stated in terms of the objective function:

$$F(x, t) = \min \left\{ \sum_{j=1}^m [y_j - \phi(x, t)]^2 \right\} \quad (3.1.5)$$

3.2 INVERSE HEAT CONDUCTION PROBLEMS

Using the concept described above, the inverse heat conduction problem may be mathematically stated in terms of the objective function:

$$F(\{q_u\}) = \min \left\{ \sum_{i=1}^M [\bar{T}_i - T_i(\{q_u\})]^2 \right\} \quad (3.2.1)$$

where $T_i(\{q_u\})$ is the calculated temperature at $\{r\}$ and is a function of unknown heat fluxes $\{q_u\}$. To find the unknowns, one resorts to the least square procedures by differentiating the above equation with respect to $\{q_u\}$,

$$\frac{\partial F(\{q_u\})}{\partial \{q_u\}} = \sum_{i=1}^M [T_i(\{q_u\}) - \bar{T}_i] \frac{\partial T_i(\{q_u\})}{\partial \{q_u\}} = 0 \quad (3.2.2a)$$

or

$$[J]\{T(\{q_u\})\} = [J]\{\bar{T}(\{q_u\})\} \quad (3.2.2b)$$

where $\partial T_i(\{q_u\})/\partial\{q_u\}$ is the element of the sensitivity coefficient matrix $[J]$ and is in general a rectangular matrix and singular,

$$J_{ij} = \left[\frac{\partial T_i^T(\{q_u\})}{\partial q_j} \right]^T \quad (3.2.3)$$

If a forward difference is used, the sensitivity coefficient matrix with respect to q_j is approximated by

$$J_{ij} \cong \frac{T_i(q_1, q_2, \dots, q_j + \lambda q_j, \dots, q_N) - T_i(q_1, q_2, \dots, q_j, \dots, q_N)}{\lambda q_j} \quad (3.2.4)$$

where λ is a small number.

3.2.1 The regularized inverse method

For ill-posed heat transfer problems given by Eq. (3.2.1), direct use of either the Levenberg-Marquardt method or the SVD method sometimes does not give satisfactory answers because of their insufficient control of the instability associated with the problems. In these cases, the regularization method offers a viable alternative. The popular approach is Tikhonov regularization, by which the minimization function is augmented as follows:

$$F(\{q_u\}) = \min \left\{ \sum_{i=1}^M [\bar{T}_i - T_i(\{q_u\})]^2 + \alpha \sum_{j=1}^N q_j^2 \right\} \quad (3.2.5)$$

where α is the regularization parameter. The above function represents a trade-off between two optimization processes: (1) agreement between the data and solution and (2)

smoothness or stability of the solution. The parameter α regulates these two processes.

Using the index notation, one has

$$F(\{q_u\}) = (\bar{T}_i - T_i)(\bar{T}_i - T_i) + \alpha q_j q_j \quad (3.2.6)$$

Again the differentiation with respect to q_j will give rise to the following matrix equation:

$$\frac{\partial F(\{q_u\})}{\partial q_j} = 2 \left[-\frac{\partial T_i^T(\{q_u\})}{\partial q_j} \right] [\bar{T}_i - T_i(\{q_u\})] + 2\alpha q_j = 0 \quad (3.2.7)$$

In terms of the sensitivity coefficients J_{ij} , Eq. (3.2.7) becomes

$$J_{ij}^T [\bar{T}_i - T_i(\{q_u\})] = \alpha q_j \quad (3.2.8)$$

For linear inverse problems, the sensitivity matrix is not a function of unknown parameters. Thus, we have

$$T_i(\{q_u\}) = \left[\frac{\partial T_i^T(\{q_u\})}{\partial q_j} \right]^T q_j + T_i(\{q_0\}) = J_{ij} q_j + T_i(\{q_0\}) \quad (3.2.9)$$

where $T_i(\{q_0\})$ is the estimated temperature when $\{q\}=0$. With this relation substituted in, Eq. (3.2.8) becomes

$$J_{ij}^T [\bar{T}_i - J_{ij} q_j - T_i(\{q_0\})] = \alpha q_j \quad (3.2.10)$$

or in matrix notation,

$$\{q_u\} = ([J]^T [J] + \alpha [I])^{-1} [J]^T [\{\bar{T}\} - \{T(\{q_0\})\}] \quad (3.2.11)$$

where $[J]$ is the sensitivity coefficient matrix and $[I]$ is the identity matrix. For the above method to be successful, an appropriate choice of the regularization parameter α needs to be chosen. While there has been much research on choosing the best regularization parameter in an *a priori* manner, most schemes depend in some way on the data and/or constraints placed on the solution, and thus are problem dependent. In the next sections,

several different techniques for selecting α for inverse heat transfer calculations are evaluated.

3.2.1.1 Maximum likelihood method

As stated earlier, a solution to an inverse problem is the most probable solution. The maximum likelihood method is used to determine the optimal regularization parameter such that the solution is most probable from the given set of experimentally measured temperature data. The method is based on the use of *a priori* information on the parameters to be estimated, which is added to the information provided by the data. In a statistical context, the *a priori* information on $\{q\}$ is expressed in the form of *a priori* probability distribution $p(\{q\})$. Bayes' rule allows us to combine it with the information contained in the data to obtain the *a posteriori* distribution [Demoment, 1989],

$$p(\{q\}/\{\bar{T}\}) = p(\{\bar{T}\}/\{q\})p(\{q\})/p(\{\bar{T}\}) \quad (3.2.12)$$

Here $p(\{\bar{T}\}/\{q\})$ denotes the probability distribution of the data based on the real solution $\{q\}$. The solution is completely determined by the knowledge of models and noise distribution. Since $p(\{\bar{T}\})$ is constant, the above equation can be expressed as

$$p(\{q\}/\{\bar{T}\}) \propto p(\{\bar{T}\}/\{q\})p(\{q\}) \quad (3.2.13)$$

The regularization method is then written as

$$\{q\} = \arg \text{Min}(\left[\{\bar{T}\} - \{T(\{q\})\}\right]^T \left[\{\bar{T}\} - \{T(\{q\})\}\right] + \alpha \{q\}^T \{q\}) \quad (3.2.14)$$

The above equation is tantamount to maximizing the *a posteriori* distribution, $p(\{q\}/\{\bar{T}\})$,

$$\{q\} = \arg \text{Max } p(\{q\}/\{\bar{T}\}) \quad (3.2.15)$$

where

$$p(\{q\}/\{\bar{T}\}) = \exp\left(-\left\{\left\{\bar{T}\right\} - \{T(\{q\})\}\right\}^T \left[\left\{\bar{T}\right\} - \{T(\{q\})\}\right] + \alpha \{q\}^T \{q\}\right) \quad (3.2.16)$$

The measured temperature data $\{\bar{T}\}$ is of the form

$$\{\bar{T}\} = [J]\{q\} + \{T(\{q_0\})\} + \{\varepsilon\} \quad (3.2.17)$$

Assuming that $\{\varepsilon\}$ follows a normal distribution $N(0, \sigma^2)$, we obtain the following probability function [Bartoszynski, 1996]:

$$p(\{\bar{T}\}/\{q\}, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{M/2} \cdot \exp\left[\left(\frac{-1}{2\sigma^2}\right)\left[\left\{\bar{T}\right\} - [J]\{q\} - \{T(\{q_0\})\}\right]^T \left[\left\{\bar{T}\right\} - [J]\{q\} - \{T(\{q_0\})\}\right]\right] \quad (3.2.18)$$

where M is the number of thermal sensors. According to the Bayesian theory, the *prior* distribution of $\{q\}$ should follow a normal distribution. Thus, combining equations (3.2.13), (3.2.16) and (3.2.18) yields the following expression for the distribution:

$$p(\{q\}/\alpha, \sigma^2) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \det[\alpha[I]]^{1/2} \exp\left[\left(\frac{-\alpha}{2\sigma^2}\right)\{q\}^T \{q\}\right] \quad (3.2.19)$$

where N is the number of unknown components of $\{q\}$. The most commonly used maximum likelihood approach consists of maximizing a marginal likelihood, which is obtained by integrating the object out of the problem [Galatsanos *et al.*, 1992],

$$L_{\bar{T}}(\alpha, \sigma^2) = \int_{-\infty}^{+\infty} p(\{q\}, \{\bar{T}\}) dq = \int_{-\infty}^{+\infty} p(\{\bar{T}\}/\{q\}, \sigma^2) p(\{q\}/\alpha, \sigma^2) dq \quad (3.2.20)$$

which means the probability of $\{T\}$ falling in a short interval of length Δ near $\{\bar{T}\}$ [Pitman, 1993]. This approach can be considered as a special case of the expectation-maximum likelihood (EM). The analytical expression for $L_{\bar{T}}(\alpha, \sigma^2)$ is given in Appendix A. Taking the derivative of $L_{\bar{T}}(\alpha, \sigma^2)$ with respect to σ^2 and setting it equal to zero yields

$$\sigma^2 = \left(\frac{1}{M}\right) \cdot R(\alpha) \quad (3.2.21)$$

where $R(\alpha) = [\{\bar{T}\} - \{T(\{q_0\})\}]^T [[I] - [J][A(\alpha)]] [\{\bar{T}\} - \{T(\{q_0\})\}]$. Substituting the above equation into Eq. (3.2.20), one has the error indicator as a function of α ,

$$ML(\alpha) = \frac{[\{\bar{T}\} - \{T(\{q_0\})\}]^T ([I] - [J][A(\alpha)]) [\{\bar{T}\} - \{T(\{q_0\})\}]}{(\det[[I] - [J][A(\alpha)]])^{1/M}} \quad (3.2.22)$$

where $[A(\alpha)] = ([J]^T [J] + \alpha [I])^{-1} [J]^T$. Thus, the optimal α corresponds to the minimum of the error indicator.

3.2.1.2 Ordinary cross-validation (OCV) and generalized cross-validation (GCV) methods

The OCV and GCV methods are widely recognized methods in the field of statistical data analysis. The essence of the method, when viewed in the current context, is to first find the vector of $S_o(\alpha, k)$ that minimizes the cost function,

$$S_o(\alpha, k) = \sum_{\substack{i=1 \\ i \neq k}}^M [\bar{T}_i - T_i]^2 + \alpha \sum_{i=1}^N q_i^2 \quad (3.2.23)$$

with the k^{th} values of measured and computed temperatures (\bar{T}_k and T_k) omitted from the calculation. In this sense the ordinary cross-validation technique is also referred to as “the leaving-one-out method.”

Having arrived at $S_o(\alpha, k)$ that minimizes the cost function, we evaluate the effectiveness of this vector in predicting the value of \bar{T}_k that was “left out” of the

calculation of the cost function. We denote the predicted value of \bar{T}_k by $T_k(\alpha, k)$. The ordinary cross-validation function $V_o(\alpha)$ is then defined in order to measure the success of this prediction when the "leaving-one-out" process is repeated for all the available data points \bar{T}_k (i.e., for $k=1$ to M). The ordinary cross-validation error indicator is thus defined by

$$V_o(\alpha) = \frac{1}{M} \sum_{k=1}^M [\bar{T}_k - T_k(\alpha, k)]^2 \quad (3.2.24)$$

For linear problems, the following equation is derived (See Appendix B for the detailed derivation):

$$\begin{aligned} V_o(\alpha) &= \frac{1}{M} \sum_{k=1}^M \left[\frac{(\bar{T}_k - T_k(\alpha))}{(1 - b_{kk})} \right]^2 \\ &= \frac{1}{M} \left\| [C]([I] - [B(\alpha)])(\{\bar{T}\} - \{T(\{q_o\})\}) \right\|^2 \end{aligned} \quad (3.2.25)$$

where $[B(\alpha)] = [J]([J]^T[J] + \alpha[I])^{-1} [J]^T$

The ordinary cross-validation technique described above may fail in cases where the matrix $[B(\alpha)]$ is close to diagonal. It is clear that if $[B(\alpha)]$ is diagonal, then the cross-validation function $V_o(\alpha)$ given above reduces simply to $(1/M) \|\{\bar{T}\}\|^2$ which is entirely independent of the choice of α . Therefore, the use of the GCV method is based on the argument that any good choice of α should be "invariant under rotation of the measurement co-ordinate system". The error indicator for the GCV method can be written as,

$$V(\alpha) = \frac{(1/M) \left\| ([I] - [B]) \{\bar{T}\} \right\|^2}{\left[(1/M) \text{Tr}([I] - [B]) \right]^2} \quad (3.2.26)$$

The detailed derivation is shown in Appendix B.

3.2.1.3 The L-curve method

As discussed above, the method of regularization represents a balance between the data-solution agreement and smoothness of solution. Thus, an appropriate choice of α should give an optimal balance. This idea has led to the development of the L-curve method (Fig.3.1), which was first proposed by Hansen (1993). This method locates the ‘corner’ on a plot of the function of the norm of computed heat fluxes, $\| \{q\} \|$, versus the norm of the difference between sensor temperatures and computed temperatures, $\| \{\bar{T}\} - \{T_{computed}\} \|$. Let Ψ be a monotonically increasing function and define the curve,

$$L = \left\{ \left(\Psi(\| \{q\} \|^2), \Psi(\| \{\bar{T}\} - \{T_{computed}\} \|^2) \right) : \alpha > 0 \right\} \quad (3.2.27)$$

Often Ψ is chosen to be one of the functions,

$$\Psi(t) = t, \quad \Psi(t) = \sqrt{t}, \quad \text{or} \quad \Psi(t) = \frac{1}{2} \log_{10} t, t > 0 \quad (3.2.28)$$

3.2.1.4 Discrepancy principle (DP)

The method for *a priori* parameter selection, such as the GCV, the ML and the L-curve methods, requires no knowledge of the random errors (noise) in the experimental data. A *posteriori* method such as the discrepancy principle does require an estimate of sensor errors [Morozov, 1984]. The discrepancy principle demands that the problem is solved so that the residual norm is the same as the norm of errors δ in the measurements.

That is,

$$\| \{\bar{T}\} - \{T_{computed}\} \| = \delta \quad (3.2.29)$$

The value of α is selected as the optimal regularization parameter. This method demands a good estimate of the experimental errors.

3.2.1.5 Yagola's discussion

Yagola *et al.* (2002) mentioned that L-curve method was mathematically inapplicable for the solution of an ill-posed problem. In fact, they pointed out that the error must be known prior to the calculations to solve the ill-posed problem. Therefore, in their theory, GCV is also inapplicable because GCV doesn't use the prior information of the error. Their algorithm is as follows:

Consider

$$\left. \begin{array}{l} Az = \bar{u}, z \in Z, \bar{u} \in U \\ \|A_h - A\| \leq h, \|u_\delta - \bar{u}\| \leq \delta \end{array} \right\} \text{(I)}$$

To find the approximate solution means to be able to construct a regularizing algorithm.

The Regularizing algorithm implies

1. brings an element $z_{h\delta} = R(h, \delta, A_h, u_\delta)$ into correspondence with any data $(h, \delta, A_h, u_\delta)$,

$h \geq 0, \delta \geq 0, A_h \in L(Z, U), u_\delta \in U$ of the problem (I);

2. has the convergence property $z_{h\delta} \rightarrow \bar{z} = A^+ \bar{u}$ as $h, \delta \rightarrow 0, \bar{u} \in R(A) \oplus R(A)^\perp$

Theorem

Let $R(A_h, u_\delta)$ be a map of the set $L(Z, U) \otimes U$ into Z . If $R(A_h, u_\delta)$ is a regularizing algorithm (not depending explicitly on h, δ), then the map $P(A, \bar{u}) = A^+ \bar{u}$ is continuous on its domain $L(Z, U) \otimes (R(A) \oplus R(A)^\perp)$

The theorem above shows that $R(A_h, u_\delta)$ is for well-posed problem because $R(A_h, u_\delta)$ is continuous (stable) on the domain. However, the theorem above is based on the regularizing algorithm. In our cases, the second condition, which is convergence as $h, \delta \rightarrow 0$, is not guaranteed. Therefore, the theorem does not prove that the L-curve is not applicable for the ill-posed problem of our algorithm. It is worth using and evaluating the L-curve method or GCV for the estimation of the regularization parameter in the computational optimization of the inverse heat conduction problems.

3.2.2 Singular value decomposition method (SVD)

One popular method to solve Eq. (3.2.2b) is to use the singular value decomposition (SVD) method. Using this method, Eq. (3.2.2b) is solved in the following manner. First the temperature is linearized,

$$\{T(\{q\})\} - \{T(\{q_0\})\} = [J]\{q\} \quad (3.2.30)$$

with $T(q_0)$ being the temperature field at $\{q\}=0$. Substituting into Eq. (3.2.1), we have the following expression:

$$F(\{q_u\}) = \min \left\{ \sum_{i=1}^M \left(\bar{T}_i - T_i(\{q_0\}) - \sum_{j=1}^N J_{ij} q_j \right)^2 \right\} \quad (3.2.31)$$

The SVD procedure to solve the above equation is to choose a solution vector $\{q\}$ such that it minimizes the above error function. The procedure gives the following expression:

$$[A]\{q\} = \{g\} \quad (3.2.32)$$

with $[A] = [J]^T [J]$ and $\{g\} = [J]^T [\{T\} - \{T(\{q_0\})\}]$. Note that $F(\{q_u\})$ measures the distance from the point $\{g\}$ to the point $[A]\{q\}$ in the column space. In the case of matrix $[A]$ being singular or degenerate, as is often the case in inverse heat transfer problems, the solution to Eq. (3.2.32) chosen by the SVD process has the smallest distance, that is, $\{q\}^T \{q\} = \text{minimum}$. According to the theory of linear algebra, an $M \times N$ matrix $[A]$ can be written as the product of an $M \times N$ column-orthogonal matrix, $[U]$, an $N \times N$ diagonal matrix $[S]$ with positive singular values, and the transpose of an $N \times N$ orthogonal matrix $[V]$. Therefore,

$$[A] = [U][S][V]^T \quad (3.2.33)$$

The singular values, s_1, s_2, \dots, s_N , are the squares of the eigenvalues of the matrix $[A]^T [A]$. Also, $[U][U]^T = [U][U]^{-1} = [I]$ and $[V][V]^T = [V][V]^{-1} = [I]$, with $[I]$ being the identity matrix. Thus, the matrix $[A]$ can be readily inverted, $[A]^{-1} = [V]^{-T} [S]^{-1} [U]^{-1} = [V][S]^{-1} [U]^T$ and the unknown flux vector may be calculated as follows:

$$\{q\} = [V][S]^{-1} ([U]^T \cdot \{g\}) \quad (3.2.34)$$

In a well-conditioned matrix $[A]$, the singular values are roughly of the same order of magnitude. As the matrix becomes ill-conditioned, that is, singular, these values become more dispersed. Formally, the condition number of a matrix is defined as the base 10 logarithm of the ratio between the largest to the smallest of the s_j 's. For ill-conditioned $[A]$, the solution vector, $\{q\}$, when pre-multiplied by the inverse matrix of

[A], results in a very poor approximation to the force vector $\{g\}$. This can be particularly true when some of the singular values are zero or very small. One remedy to the problem is to use the truncated SVD by which a judicial decision is made to throw away the singular values smaller than a singularity threshold value, that is, to set $1/s_j = 0$ if $s_j / s_{\max} < \tau$ ($j=1,2,\dots,N$).

From the experience of selecting the regularization parameters, we propose the use of the discrepancy principle to choose the singularity threshold value. Thus, we have

$$\|\{\bar{T}\} - \{T(\tau)\}\| = \delta \quad (3.2.35)$$

This approach eliminates the *ad hoc* guess for the threshold value τ . The testing cases below show that this method produces a consistently good estimate.

3.2.3 Levenberg-Marquardt method

3.2.3.1 Ozisik's implementation

The minimization problem can be solved using the Levenberg-Marquardt method, which in essence is a modification of the Gauss-Newton algorithm that more dynamically mixes the Gauss-Newton algorithm and gradient-descent iterations. The method is iterative in nature [Sawaf *et al.*, 1995]. This method uses the second order derivative of F with respect to $\{q\}$, that is, the Hessian matrix or the curvature matrix, which is given by $[J]^T[J]$. Thus, Eq. (3.2.2b) may be solved using Newton's method,

$$0 = [J]^T [T] - [J]^T \{\bar{T}\} + [J]^T [J] \{\Delta q\} \quad (3.2.36)$$

The Levenberg-Marquardt algorithm for the solution of the above equation stabilizes the Hessian matrix by adding to an adjustable parameter term $\mu^k \text{diag}(([J]^k)^T [J]^k)$ and the resultant iterative procedure is thus as follows:

$$\begin{aligned} \{q_u\}^{k+1} = \{q_u\}^k + & \left[([J]^k)^T [J]^k + \mu^k \text{diag}(([J]^k)^T [J]^k) \right]^{-1} \\ & \cdot ([J]^k)^T \left[\{\bar{T}\} - \{T(\{q_u\}^k)\} \right] \end{aligned} \quad (3.2.37)$$

where *diag* means taking the diagonal terms of the matrix in the bracket and μ is a positive scalar damping parameter whose magnitude is dynamically adjusted to condition the iterative process. The superscript k denotes the iteration number. The calculation starts with an initial guess for $\{q_u\}$, and the temperature distribution is predicted and compared with the measured data. If the error increases during updating, then μ is increased, often by a factor of 10. On the other hand, if the error reduces, then μ is decreased, often by a factor of 10 [Marquardt, 1963]. The iteration convergence and the objective function $F(\{q_u\})$ are both used as stopping criteria [Ozisik *et al.*, 2000],

$$\| \{q_u\}^{k+1} - \{q_u\}^k \| \leq \gamma_1 \quad (3.2.38)$$

$$F(\{q_u\}^{k+1}) \leq \gamma_2 \quad (3.2.39)$$

$$\left\| ([J]^k)^T \left[\{\bar{T}\} - \{T(\{q_u\}^k)\} \right] \right\| \leq \gamma_3 \quad (3.2.40)$$

where $\gamma_1, \gamma_2, \gamma_3$ are pre-chosen constants. The Levenberg-Marquardt algorithm has proved to be an effective and popular way to solve nonlinear least squares problems. It can also be used to solve linear problems.

3.2.3.2 More's implementation (trust-region strategy)

The Levenberg-Marquardt method can be described and analyzed using the trust-region framework. More proposed a robust and efficient implementation of the Levenberg-Marquardt method using a trust-region strategy [More, 1977]. The solution

for the estimation of the N unknown parameters $P_j, j=1, \dots, N$, is based on the minimization of the ordinary least squares norm given by

$$S(\{P\}) = [\{\bar{T}\} - \{T(\{x\} + \{P\})\}]^T [\{\bar{T}\} - \{T(\{x\} + \{P\})\}] \quad (3.2.41)$$

We utilize a trust-region strategy for our calculation. That is, the objective function above has a constraint of

$$\|\{P\}\| \leq \Delta \quad (3.2.42)$$

where Δ is a trust-region radius. Using the Lagrange function

$$\Psi = [\{\bar{T}\} - \{T(\{x\} + \{P\})\}]^T [\{\bar{T}\} - \{T(\{x\} + \{P\})\}] - \mu(\Delta^2 - \{P\}^T \{P\}) \quad (3.2.43)$$

In the above equation, μ is a Lagrange multiplier. To minimize the Lagrange function given by equation (3.2.43), we need to equate to zero the derivatives of $\Psi(\{P\})$ with respect to each of the unknown parameters $[P_1, P_2, \dots, P_N]$, that is,

$$\nabla \Psi(\{P\}) = -[J]^T(\{P\})[\{\bar{T}\} - \{T(\{x\} + \{P\})\}] + \mu\{P\} = 0 \quad (3.2.44)$$

where $[J(\{P\})] = \left[\frac{\partial \{T^T(\{x\} + \{P\})\}}{\partial \{P\}} \right]^T$. In the linear inverse problem, the following relation holds:

$$\{T(\{x\} + \{P\})\} = \{T(\{P\})\} + [J]\{P\} \quad (3.2.45)$$

The resulting expression is rearranged to yield the following procedure to obtain the vector of unknown parameters $\{P\}$:

$$\{P\} = [[J]^T [J] + \mu[I]]^{-1} [J]^T [\{\bar{T}\} - \{T(\{P\})\}] \quad (3.2.46)$$

The complementarity condition of this problem is

$$\mu(\|\{P\}\| - \Delta) = 0 \quad (3.2.47)$$

The Lagrange multiplier μ is called a damping parameter for the Levenberg-Marquard method. To find an optimal damping parameter for the calculation, the following function is utilized:

$$\phi(\mu) = \|\{P\}\| - \Delta = \left\| \left[[J]^T [J] + \mu [I] \right]^{-1} [J]^T [\{\bar{T}\} - \{T(\{P\})\}] \right\| - \Delta \quad (3.2.48)$$

If $\phi(0) \leq 0$, then $\mu=0$ is the required parameter. Otherwise, the following procedure is utilized. Let $[J]=[U][S][V]^T$ be the singular value decomposition of $[J]$, then

$$\begin{aligned} \|\{P\}\|^2 &= \{P\}^T \{P\} \\ &= [\{\bar{T}\} - \{T(\{P\})\}]^T [U][S] \left[(S^2 + \mu [I])^{-1} \right]^2 [S][U] [\{\bar{T}\} - \{T(\{P\})\}] \\ &= \left[\sum_{i=1}^n \frac{s_i^2 u_i^2 (\bar{T}_i - T_i)^2}{(s_i^2 + \mu)^2} \right] \end{aligned} \quad (3.2.49)$$

where s_1, \dots, s_n are the singular values of $[J]$. Equation (3.2.49) is

$$\phi(\mu) = \left[\sum_{i=1}^n \frac{s_i^2 u_i^2 (\bar{T}_i - T_i)^2}{(s_i^2 + \mu)^2} \right]^{\frac{1}{2}} - \Delta \quad (3.2.50)$$

Notice that the above equation is expressed as

$$\phi(\mu) = \frac{a}{b + \mu} - \Delta \quad (3.2.51)$$

Therefore, $\phi(\mu_{k+1})=0$ if

$$\mu_{k+1} = \mu_k - \left(\frac{\phi(\mu_k) + \Delta}{\Delta} \right) \left[\frac{\phi(\mu_k)}{\phi'(\mu_k)} \right] \quad (3.2.52)$$

The function $\phi'(\mu)$ is derived in the following form:

$$\phi'(\mu) = \frac{\partial \phi(\mu)}{\partial \mu} = \frac{\partial \|\{P(\mu)\}\|}{\partial \mu} = \frac{1}{2} \|\{P(\mu)\}\|^{-1} \frac{\partial \|\{P(\mu)\}\|^2}{\partial \mu}$$

$$= -\frac{1}{\|\{P(\mu)\}\|} \sum_{j=1}^n \frac{s_i^2 u_i^2 (\bar{T}_i - T_i)^2}{(s_i^2 + \mu)^3} = -\frac{\{P\}^T ([J]^T [J] + \mu [I])^{-1} \{P\}}{\|\{P(\mu)\}\|} \quad (3.2.53)$$

In the above equation, the following equations are utilized:

$$\frac{\partial \|\{P(\mu)\}\|^2}{\partial \mu} = -2 \sum_{j=1}^n \frac{s_i^2 u_i^2 (\bar{T}_i - T_i)^2}{(s_i^2 + \mu)^3} \quad (3.2.54)$$

$$\{P\}^T ([J]^T [J] + \mu [I])^{-1} \{P\} = \sum_{j=1}^n \frac{s_i^2 u_i^2 (\bar{T}_i - T_i)^2}{(s_i^2 + \mu)^3} \quad (3.2.55)$$

3.3 INVERSE SOLIDIFICATION PROBLEMS

The inverse solidification problem seeks the heat flux distribution $\{q\}$ that leads to a given solidification interface distance $\{\bar{d}\}$. The inverse solidification problem belongs to the class of a nonlinear inverse problem. Therefore, the SVD method, which is used for the inverse heat transfer problem, is not applicable in the nonlinear problem. In addition, the Levenberg-Marquardt algorithm is not effective based on the cases studied in the inverse heat transfer problems. The regularization method seems like a viable choice for nonlinear inverse problems. In this section, we present the regularization method for the inverse solidification problem. Then, a new method for computing sensitivity coefficients is addressed. The L-curve method, sequential and whole domain methods, scaling, and piecewise polynomial functions are also presented.

3.3.1 The regularization method

3.3.1.1 Steady-state problems

The inverse solidification problem seeks the heat flux distribution $\{q\}$ that leads to a given solid-liquid interface shape $\{\bar{d}\}$. Note that solidification interface distance,

$\{d\}$, is defined as the length between the left wall and the solid-liquid interface (See Fig. 2.2). For the purpose of numerical analysis, the unknown $\{q\}$ is discretized into N different values, $q_j, j=1, \dots, N$. Thus, the inverse solution is to estimate the N unknown parameters $q_j (j=1, \dots, N)$ by *minimizing* the following objective function:

$$S(\{q\}) = \sum_{i=1}^M [\bar{d}_i - d_i(\{q\})]^2 + \alpha \sum_{i=1}^N q_i^2 \quad (3.3.1)$$

where $\{q\}^T = [q_1, q_2, \dots, q_N]$ is the vector of unknown parameters, $d_i(\{q\})$ is the estimated solidification interface position, \bar{d}_i is the ideal solidification interface position, N is the total number of unknown parameters, and M is the total number of controlled solidification interface position, where $M \geq N$. In the above equation α is the regularization parameter. Equation (3.3.1) can be written in matrix form,

$$S(\{q\}) = [\{\bar{d}\} - \{d(\{q\})\}]^T [\{\bar{d}\} - \{d(\{q\})\}] + \alpha \{q\}^T \{q\} \quad (3.3.2)$$

To minimize the least squares norm given by Equation (3.3.2), the derivatives of $S(\{q\})$ with respect to each of the unknown parameters $[q_1, q_2, \dots, q_N]$ are set to zero, that is,

$$\frac{\partial S(\{q\})}{\partial q_1} = \frac{\partial S(\{q\})}{\partial q_2} = \dots = \frac{\partial S(\{q\})}{\partial q_N} = 0 \quad (3.3.3)$$

This necessary condition for the minimization of $S(\{q\})$ can be expressed in matrix notation by setting the gradient of $S(\{q\})$ to zero, that is,

$$\nabla S(\{q\}) = 2 \left[-\frac{\partial \{d^T(\{q\})\}}{\partial \{q\}} \right] [\{\bar{d}\} - \{d(\{q\})\}] + 2\alpha \{q\}^T = 0 \quad (3.3.4)$$

Defining the sensitivity coefficient matrix,

$$[J(\{q\})] = \left[\frac{\partial \{d^T(\{q\})\}}{\partial \{q\}} \right]^T \quad (3.3.5)$$

To approximate the sensitivity coefficients, the finite difference scheme is used (See the next section). Equation (3.3.4) becomes

$$[J(\{q\})]^T [\{\bar{d}\} - \{d(\{q\})\}] = \alpha \{q\} \quad (3.3.6)$$

For non-linear problems, when the sensitivity coefficients depend on the vector of unknown parameters $\{q\}$, the problem is referred to as non-linear inverse problems. By the Taylor series expansion,

$$\{d(\{q\}^{k+1})\} = \{d(\{q\}^k)\} + [J(\{q\}^k)]^T (\{q\}^{k+1} - \{q\}^k) \quad (3.3.7)$$

Equation (3.3.6) becomes

$$[J^T(\{q\}^{k+1})]^T [\{\bar{d}\} - \{d(\{q\}^{k+1})\}] = \alpha \{q\}^{k+1} \quad (3.3.8)$$

Using Equation (3.3.7) for the non-linear equation, Equation (3.3.8) becomes

$$[J(\{q\}^{k+1})]^T [\{\bar{d}\} - \{d(\{q\}^k)\} - [J]^k (\{q\}^{k+1} - \{q\}^k)] = \alpha \{q\}^{k+1} \quad (3.3.9)$$

and then

$$\{q\}^{k+1} = \left[([J]^{k+1})^T [J]^k + \alpha [I] \right]^{-1} ([J]^{k+1})^T [\{\bar{d}\} - \{d(\{q\}^k)\} + [J]^k \{q\}^k] \quad (3.3.10)$$

Further assuming $[J\{q\}^{k+1}] \approx [J\{q\}^k]$, we have the following estimation for the unknown heat flux distribution:

$$\{q\}^{k+1} = \left[([J]^k)^T [J]^k + \alpha [I] \right]^{-1} ([J]^k)^T [\{\bar{d}\} - \{d(\{q\}^k)\} + [J]^k \{q\}^k] \quad (3.3.11)$$

3.3.1.2 Transient problems

The inverse transient solidification problem seeks the heat flux distribution $q(y,t)$ that leads to a given solidification interface distance \bar{d}_{ij} for time and space. For the purpose of a numerical analysis, we consider the unknown function $q(y,t)$ to be parameterized in the following form:

$$q(y,t) = \sum_{j=1}^N q_j f_j(y,t) \quad (3.3.12)$$

where $f_j(y,t)$ ($j=1, \dots, N$) are known trial functions. This approach is categorized to the parameter estimation approach. The inverse solution is to estimate the N unknown parameters q_j ($j=1, \dots, N$) by minimizing the following objective function:

$$S(\{q\}) = \sum_{i=1}^L \sum_{j=1}^M [\bar{d}_{ij} - d_{ij}(\{q\})]^2 + \alpha \sum_{j=1}^N q_j^2 \quad (3.3.13)$$

where q_j is the unknown parameters, \bar{d}_{ij} is the ideal solidification interface distance at time and space, $d_{ij}(\{q\})$ is the estimated solidification interface distance corresponding to the same time and space as \bar{d}_{ij} , L is the number of controlled solidification distance for time, M is the number of controlled solidification distance for space, and N is the total number of unknown parameters. In the above equation, α is the regularization parameter. Equation (3.3.13) can be written in matrix form,

$$S(\{q\}) = [\{\bar{d}\} - \{d(\{q\})\}]^T [\{\bar{d}\} - \{d(\{q\})\}] + \alpha \{q\}^T \{q\} \quad (3.3.14)$$

where $\{q\}^T = [q_1, q_2, \dots, q_i, \dots, q_N]$,

$$\{\bar{d}\} = [\{\bar{d}_1\} \{\bar{d}_2\} \dots \{\bar{d}_i\} \dots \{\bar{d}_L\}]^T = [\bar{d}_{11} \bar{d}_{12} \dots \bar{d}_{1M} \bar{d}_{21} \dots \bar{d}_{ij} \dots \bar{d}_{LM}]^T,$$

$$\{d(\{q\})\} = [\{d_1\} \{d_2\} \dots \{d_i\} \dots \{d_L\}]^T = [d_{11} d_{12} \dots d_{1M} d_{21} \dots d_{ij} \dots d_{LM}]^T.$$

Notice that equation (3.3.14) is identical to equation (3.3.2). Thus, we can solve the inverse transient problem using the same procedure as the steady-state calculations described above. We use equation (3.3.11) to estimate the heat flux solutions. In addition, we use the following equation for computing the sensitivity coefficients:

$$J_{ij} = \frac{d_i(q_1, q_2, \dots, q_j + \lambda q_j, \dots, q_N) - d_i(q_1, q_2, \dots, q_j, \dots, q_N)}{\lambda q_j} \quad (3.3.15)$$

3.3.2 Finite difference approximation for determining sensitivity coefficients

It is common knowledge in the inverse community that the following finite-difference approximation for determining sensitivity coefficients (Eq.3.3.5) is used in the

inverse problem. If a conventional forward difference approximation $\left(\frac{\partial d_i}{\partial q_j}\right)_{FW}$ is used,

the sensitivity coefficients with respect to q_j is approximated by

$$J_{ij} = \frac{\partial d_i}{\partial q_j} \cong \left(\frac{\partial d_i}{\partial q_j}\right)_{FW} = \frac{d_i(q_1, q_2, \dots, q_j + \lambda q_j, \dots, q_N) - d_i(q_1, q_2, \dots, q_j, \dots, q_N)}{\lambda q_j} \quad (3.3.16)$$

where λ is a small number. Different values of λ were selected by authors. In the inverse heat transfer problem, Ozisik and Orlande (2000) used $\lambda=10^{-5}$ or 10^{-6} , and Beck and Blackwell (1988) used $\lambda=10^{-4}$. As for the inverse solidification problem, Ruan (1990) used $\lambda=10^{-3}$ in the regularization method, and Zabararas (1990) used $\lambda=10^{-3}$ in the steepest descent method. In this conventional scheme, the perturbed values, λq_j , with constant λ depend on unknown parameters q_i . Therefore, the perturbed values vary in the each calculated sensitivity coefficient. This finite difference scheme works well if the unknown parameters, $\{q\}$, do not distribute in a wide range. However, if the unknown parameters are in a wide range, especially, in the case where the values are crossing zero, the calculations to find the optimal solution may be unstable. More specifically, we

consider truncation error. The truncation error results from the neglected terms in the Taylor series expansion of the perturbed function. The Taylor series expansion of $d_i(q_1, q_2, \dots, q_j + \Delta q_j, \dots, q_N)$ can be written as

$$d_i(q_1, q_2, \dots, q_j + \Delta q_j, \dots, q_N) = d_i(q_1, q_2, \dots, q_j, \dots, q_N) + \Delta q_j \frac{\partial d_i}{\partial q_j} + \frac{(\Delta q_j)^2}{2} \frac{\partial^2 d_i}{\partial q_j^2} + \dots \quad (3.3.17)$$

The truncation error of the leading term for the forward-difference approximation is

$$Err_{Trun}(\Delta q_j) = \left| \left(\frac{\partial d_i}{\partial q_j} \right)_{FW} - \left(\frac{\partial d_i}{\partial q_j} \right) \right| = \frac{\Delta q_j}{2} \left| \frac{\partial^2 d_i}{\partial q_j^2} \right| \quad (3.3.18)$$

Any arithmetic operation among floating numbers introduces an additional fractional error. This type of error is called *round-off error*. Round-off errors accumulate with increasing amounts of calculation. The total error (Err_{total}) of the forward-difference approximation is the combination of the truncation and the round-off errors, which is given as

$$Err_{total} = Err_{Trun}(\Delta q_j) + Err_{round}(\Delta q_j) \quad (3.3.19)$$

If the conventional forward-difference scheme is used, the perturbed values, λq_j , vary with the unknown parameters. This leads to the disparity of the truncation error in the sensitivity coefficient. That is, if the unknown parameter is large, the truncation error also becomes large. Thus, the ratio of the truncation error to the total error is large. On the other hand, if the unknown parameter is small, the ratio of the round-off error to the total error is large. This discrepancy in each sensitivity coefficient causes the unstable and inaccurate solution of the inverse problems.

As an example, we consider the case where unknown parameters are ranging $-1,000 \leq q \leq 1,000$, and we have $q_1=0.001$ and $q_2=1,000$. If the conventional finite difference scheme (Eq.3.3.16) is used, the perturbed value for the sensitivity coefficient J_{1j} equals to $q_1 \cdot \lambda = 0.001 \times 0.001 = 1 \times 10^{-6}$. Similarly, the perturbed value for J_{2j} is $q_2 \cdot \lambda = 1,000 \times 0.001 = 1$. In the former case, the perturbed value is very small, which is 1×10^{-6} . Thus, the ratio of the truncation error to the total error is very small. On the other hand, the perturbed value for J_{2j} , which is 1, is large compared to the last case. Thus, the ratio of the truncation error to the total error is smaller than that in the last case. This discrepancy of the error in the sensitivity coefficients causes unstable and slow convergence. To overcome this drawback, the following forward difference scheme is proposed:

$$J_{ij} \cong \frac{d_i(q_1, q_2, \dots, q_j + \varepsilon \cdot \text{Max}(q_k), \dots, q_N) - d_i(q_1, \dots, q_j, \dots, q_N)}{\varepsilon \cdot \text{Max}(q_k)} \quad (3.3.20)$$

where ε is a small number, and $\text{Max}(q_k)$ is a maximum value of $\{q\}$. In this proposed method, the same perturbed value, $\varepsilon \cdot \text{Max}(q_k)$, is used for all of the sensitivity coefficients no matter what values we have for the unknown parameters. This scheme removes the drawback of the conventional forward difference method since the ratio of the truncation error has the same level in each sensitivity coefficient. This modification results in the stable and fast convergence.

3.3.3 L-curve method

It is well known that the success of the regularized minimization method described above depends on an appropriate choice of the regularization parameter. Note that since the problem is non-linear, the Ordinary Cross-Validation (OCV), the Generalized Cross-Validation (GCV), and Maximum likelihood method (ML) methods, which were discussed in Chapter 3.2 for the inverse heat conduction problems, are in general ineffective to find the optimal regularization parameter. In addition, the discrepancy principle based on the measurement error is not utilized because error may not be easily specified. The L-curve method is considered a viable choice for this purpose. The L-curve method is based on an algorithm that locates the ‘corner’ of a plot of the function of norm of computed heat fluxes, $\|\{q\}\|$, versus norm of the difference between computed solidification distance and prescribed solidification distance, $\|\{\bar{d}\} - \{d\}\|$ (Figure 3.1). Let Φ be a monotonically increasing function, we may define the curve,

$$L = \left\{ (\Phi(\|\{q\}\|), \Phi(\|\{\bar{d}\} - \{d\}\|)) : \alpha > 0 \right\} \quad (3.3.21)$$

where Φ is chosen to be one of the following functions:

$$\Phi(t) = t, \quad \Phi(t) = \sqrt{t}, \quad \text{or} \quad \Phi(t) = \frac{1}{2} \log_{10} t, \quad t > 0 \quad (3.3.22)$$

3.3.4 A sequential method and a whole domain method

The time domain in which the inverse problem is calculated may be another way to classify the methods of solution. Several schemes have been proposed for the time-dependent inverse problem. A whole domain method and a sequential method are useful methods for a parameter specification inverse problem. The whole domain method

utilizes the whole part of the time domain to compute unknown parameters in the inverse algorithm. On the other hand, the time domain in the sequential method is split into some parts of the time domain. Then unknown variables in each time domain are calculated separately and independently. The final conditions that are calculated in each time domain are carried over to the next time domains for the initial conditions. In the regularization method, the most of the computation time is spent for computing the sensitivity coefficient matrix. All of the time steps must be considered to form the sensitivity coefficient matrix for the whole domain method. Therefore, the whole time domain method is computationally expensive and time-consuming. On the other hand, each time domain has the smaller number of the time steps in the sequential method. As a result, the smaller number of the sensitivity coefficients is calculated. Therefore, the sequential method allows us to store small space of a computer memory. In addition, the computation time of the sequential method is faster than that of the whole domain method. The whole domain method, however, may be suitable for the problems where variables in the early time domain affect the controlled parameters in the late time. These problems include the solidification process in which the effect of diffusion and latent heat is considered. We use both the whole domain and sequential methods in the present research.

3.3.5 Scaling

We use a polynomial approximation for the trial functions (Eq. 3.3.12) to find optimal heat flux solutions. When variables of the polynomial function are either in a small or large range, *scaling* needs to be performed for the variables of the polynomial

functions. *Scaling* is a transformation of the polynomial function in order to obtain an optimal solution of the inverse calculations. The poorly scaled functions make the solution unbalanced. For example, when a second order polynomial function, $q(x)=a_1x^2+a_2x+a_3$ ($0 \leq x \leq 0.01$), is used for approximating an optimal solution, the coefficient a_1 for x^2 has very small effect compared to a_2 and a_3 because x^2 is a very small value in $0 \leq x \leq 0.01$. In other words, $q(x)$ is less sensitive to small changes in a_1 than those in a_2 . Thus, the sensitivity coefficients J_{i1} (Eq. 3.3.15) for a_1 become very small. When these sensitivity coefficients are used to find the optimal solution, we obtain the solution in which the value of a_1x^2 is small and negligible. That is, the solution using this polynomial function has a strong linear relationship between x and q . For overcoming this drawback, the polynomial function should be changed to

$$q(x) = a_1 \cdot (10^{-4} x^2) + a_2 \cdot (10^{-2} x) + a_3 \quad (3.3.23)$$

Since the variables ($10^{-4} x^2$, $10^{-2} x$, and 1) are linearly independent and generate a second order polynomial function, no loss of generality is incurred by the use of the scaled polynomial function. In addition, the scaled function makes the solution more balanced.

3.3.6 Piecewise functions

In the whole domain method, we use the piecewise approximation for an optimal heat flux solution. The final time t_f is given in the calculation. We split the time domain into $(n+1)$ domains as follows:

$$0 = t_0 < t_1 < t_2 < \dots < t_n = t_f \quad (3.3.24)$$

Piecewise functions $P_i(t)$ are defined as

$$P_i(t) \quad t_i < t < t_{i+1} \quad i = 0, 1, \dots, n-1 \quad (3.3.25)$$

3.4 SUMMARY OF THE NUMERICAL ALGORITHM

In this section, the computational algorithm for the inverse heat conduction problems and the inverse solidification problems described above are summarized.

3.4.1 Inverse heat conduction problems

Suppose that the temperature measurements $\{\bar{T}\} = (\bar{T}_1 \cdots \bar{T}_M)$, the norm of measurement error δ , and the standard deviation of measurement error σ are given. Unknown heat flux $\{q\}$ is computed as follows.

3.4.1.1 The regularization method and the SVD

Step 1. Set a perturbation value $\lambda = \delta q = 0.01$.

Step 2. Set $\{q\} = \{q_0\} = (q_1, q_2, \dots, q_N) = (0, 0, \dots, 0)$.

Step 3. Solve the direct problem given by equations (2.1.1), (2.1.2), (2.1.3) in order to compute $\{T(\{q_0\})\} = (T'_1 \cdots T'_i \cdots T'_M)$ using $\{q_0\} = 0$.

For $j=1, N$

Step 4. Set $q_j = \delta q$. $q_1 = q_2 = \dots = q_{j-1} = q_{j+1} = \dots = q_N = 0$. That is, $\{q''\} = (q_1, q_2, \dots, q_j, \dots, q_N) = (0, \dots, 0, \delta q, 0, \dots, 0)$

Step 5. Solve the direct problem given by equations (2.1.1), (2.1.2), (2.1.3) in order to

compute $\{T''(\{q''\})\} = (T_1'' \cdots T_i'' \cdots T_M'')$ using $\{q''\}$.

For $i=1, M$

Step 6. Compute the sensitivity coefficients, $J_{ij} = \frac{T_i'' - T_i'}{\delta q_j}$.

End for i

End for j

Knowing $[J]$, $\{T(\{q_0\})\}$, $\{\bar{T}\}$, the regularization method and the SVD are computed as

follows:

i) The regularization method

Step 7. Set $\alpha=1$

Step 8. Compute $\{q\}$ using Eq. (3.2.11).

Step 9. Compute the value of ML, OCV, GCV, DP, L-curve using Eq. (3.2.21), (3.2.22), (3.2.25), (3.2.26), (3.2.27), (3.2.29)

Step 10. Replace α by 0.9α , and return to Step 8 until $\alpha < 10^{-20}$

ii) SVD

Step 7. $\tau=1$

Step 8. Compute $\{q\}$ using Eq. (3.2.34).

Step 9. Compute the value of DP using Eq. (3.2.35)

Step 10. Replace τ by 0.9τ , and return to Step 8 until $\tau < 10^{-20}$

3.4.1.2 The Levenberg-Marquardt method

(a) Ozisik implementation

Step 1. Set $k=1$, $\lambda=0.01$, $\mu=0.001$.

Step 2. Take an initial guess for the heat flux, $\{q^0\}=(q_1^0, \dots, q_i^0, \dots, q_N^0)$.

Step 3. Solve the direct problem given by equations (2.1.1), (2.1.2), (2.1.3) in order to

$$\text{compute } \{T(\{q\}^{k-1})\} = (T'_1 \dots T'_i \dots T'_M) \text{ using } \{q\}^{k-1} = (q_1^{k-1}, \dots, q_j^{k-1}, \dots, q_N^{k-1})$$

For $j=1, N$

Step 4. Set $\{q''\} = (q_1^{k-1}, \dots, q_j^{k-1} + \lambda q_j^{k-1}, \dots, q_N^{k-1})$.

Step 5. Solve the direct problem given by equations (2.1.1), (2.1.2), (2.1.3) in order to

$$\text{compute } \{T''(\{q''\})\} = (T''_1 \dots T''_i \dots T''_M).$$

For $i=1, M$

Step 6. Compute the sensitivity coefficients, $J_{ij} = \frac{T''_i - T'_i}{\lambda q_j^{k-1}}$.

End for i

End for j

Step 7. Compute $\{q\}^{k+1}$ from Eq. (3.2.37).

Step 8. Compute $F(\{q\}^{k+1})$.

If $F(\{q\}^{k+1}) \geq F(\{q\}^k)$, replace μ by 10μ and return to Step 7.

If $F(\{q\}^{k+1}) < F(\{q\}^k)$, accept the new estimate $\{q\}^{k+1}$ and replace μ by 0.1μ .

Step 9. Check the stopping criteria given by Eq.(3.2.38-40). If any of the criteria is satisfied, stop the calculations; Otherwise, replace k by $k+1$ and return to Step 3.

(b) More's implementation

We consider a linear inverse heat transfer problem in this algorithm.

Step 1. Set $\mu_0, \Delta^0, \{q^0\}, \lambda, m=0$

Step 2. Compute $\{T^0\}, [J]$ using the direct problem given by equations (2.1.1), (2.1.2), (2.1.3).

Step 3. $k=0$

Step 4. Compute $\{P(\mu_k)\}$ using the equation(3.2.46).

Step 5. Compute $\{T(\{P\})\}=\{T\}^m+[J]\{P(\mu_k)\}$

Step 6. Evaluate $\phi(\mu_k)$ and $\phi'(\mu_k)$ using the equations (3.2.48) and (3.2.53)

Step 7. If $\phi(\mu_k) \approx 0$, then accept the step $\{P\}$. Set $\{q\}^{m+1}=\{P\}+\{q\}^m, \{T\}^{m+1}=\{T(\{P\})\}, \Delta^{m+1}=\lambda\Delta^m$, and go to step 10. Otherwise go to step 8.

Step 8. Obtain μ_{k+1} using equation (3.2.52).

Step 9. If $\mu_{k+1} < 0$, then set $\mu_0 = \mu_0 \times 0.01$ and go to step 4. Otherwise set $k=k+1$ and go to step 4.

Step 10. If Δ^{m+1} is small enough, stop the calculation. Otherwise replace m by $m+1$ and go to step 3.

3.4.2 Inverse solidification problems

Step 1. Set $k=1, \lambda=0.01, \alpha=1$

Step 2. Set the ideal solidification distance, $\{\bar{d}\} = (\bar{d}_1, \dots, \bar{d}_M)$.

Step 3. Set the initial guess of heat flux, $\{q^0\} = (q_1^0, \dots, q_j^0, \dots, q_N^0)$.

For $j=1, N$

Step 4. Solve the direct problem given by equations (2.3.1) to (2.3.8) in order to compute

the solidification distance $\{d\} = (d_1^{k-1}, \dots, d_i^{k-1}, \dots, d_M^{k-1})$ using

$$\{q\}^{k-1} = (q_1^{k-1}, \dots, q_j^{k-1}, \dots, q_N^{k-1})$$

Step 5. Set $\{q''\} = (q_1^{k-1}, \dots, q_j^{k-1} + \lambda q_j^{k-1}, \dots, q_N^{k-1})$.

Step 6. Solve the direct problem given by equations (2.3.1) to (2.3.8) in order to compute

the solidification distance $\{d''\} = (d_1'', \dots, d_i'', \dots, d_M'')$ using $\{q''\}$.

For $i=1, M$

Step 7. Compute the sensitivity coefficients, $J_{ij} = \frac{d_i'' - d_i^{k-1}}{\lambda q_j^{k-1}}$.

End for i

End for j

Step 8. Compute $\{q\}$ using Eq. (3.3.11).

Step 9. If $\|\{q\}^{k+1} - \{q\}^k\| \leq \omega$ is satisfied, replace α by 0.8α and go to Step 3 until $\alpha < 10^{-30}$;

Otherwise, replace k by $k+1$ and return to Step 4.

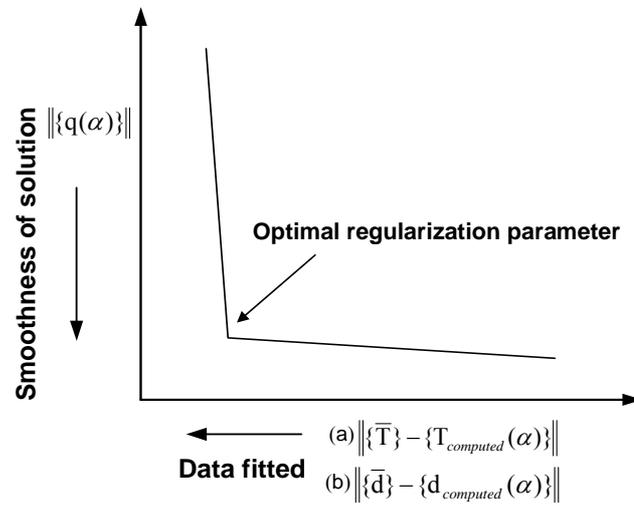


Figure 3.1. Illustration of determining optimal α using the L-curve method in the case of (a) the inverse heat conduction problems and (b) the inverse solidification problems

CHAPTER FOUR

RESULTS AND DISCUSSION

We address and discuss the formulations and results of the inverse heat conduction problems and the inverse solidification problems in this chapter.

4.1 INVERSE HEAT CONDUCTION PROBLEMS

In this section, various schemes presented for the selection of optimal regularization parameter α are discussed in a comparative fashion for the inverse heat transfer calculations. The results are then compared with those obtained using the SVD and Levenberg-Marquardt methods. The examples used steady-state 2-D problems. A parameter estimation and a function estimation approach are also evaluated using transient 1-D problems. The direct problems are calculated using the Galerkin finite element method. The detailed formulation and accuracy of the calculation were given in early publications [Song *et al.*, 2002]. As an application, the heat flux distribution calculated using the inverse heat transfer algorithm described above is also presented.

In general, error of the temperature measurement is assumed to follow a normal distribution. Thus, in the example problems, the input error is the normally distributed random number generated in the following scheme. Let z_1 and z_2 be independent uniformly distributed random number in $[0,1]$. Then the random variables

$$\xi = \sqrt{-2 \ln z_1} \cos(2\pi z_2) \quad (4.1.1)$$

are independent and distributed according to the standard normal distribution, $N(0,1)$. If the temperature distribution follows $N(T, \sigma)$, then the temperatures are determined in the following forms:

$$T_{sensor} = T_{analytic} + \sigma\xi \quad (4.1.2)$$

Regularization parameters for the analytic solution are determined by finding the minimum points of the following sum of square errors in the computed heat flux $\{q(\alpha)\}$:

$$S(\alpha) = \sum_{j=1}^M [q_j(\alpha) - q_{j,analytic}]^2 \quad (4.1.3)$$

In the above equation, $q_{j,analytic}$ is the heat flux for the analytic solution, and $q_j(\alpha)$ is the computed heat flux. The value of $S(\alpha)$ indicates the accuracy of the heat flux using the inverse algorithm. Equations for the ML, OCV, and GCV methods are shown in equations (3.2.22), (3.2.25), (3.2.26), respectively. The minimum points of these functions indicate the optimal regularization parameter. The regularization parameters by the L-curve method were found in the corner points. Since the variance $\sigma^2(\alpha)$ found using the maximum likelihood method is found in equation (3.2.21), the variance is compared with the input error variance σ^2 . The regularization parameters were then found by minimizing

$$ML_{variance}(\alpha) = |\sigma^2(\alpha) - \sigma^2_{analytic}| \quad (4.1.4)$$

Similarly, utilizing equation (3.2.29) for the discrepancy principle, the regularization parameters were found by utilizing

$$DP(\alpha) = |\delta - \|\{\bar{T}\} - \{T_{computed}(\alpha)\}\| | \quad (4.1.5)$$

where δ is the norm of input noise in the sensors, $\{\bar{T}\}$ is the sensor temperature, and $\{T_{computed}\}$ is the computed temperature depending on α . The same equation is used to find the optimal singularity threshold for SVD. These functions are examined for a range of α and τ varying between 10^{-15} and 1.

4.1.1 Steady-state problems

4.1.1.1 2-D heat conduction with a rectangular region

First, we consider a 2-D inverse heat conduction problem over a rectangular region ($0 \leq x \leq 1m$, $0 \leq y \leq 1m$) with and without heat generation. The boundaries at $y=0$ and $y=1m$ were insulated. The boundary at $x=0$ was kept at 10K, and the heat fluxes at $x=1m$ are to be determined using the inverse algorithm. The problem is schematically shown in Figure 4.1.1. Three different cases are considered, which are given in Table 4.1.1. The analytic solution is obtained using the prescribed temperature of 110K. The “measured” interior sensor temperatures are generated by tempering the analytic solution with different noise levels. The inverse calculations used the regularization parameters selected by the maximum likelihood (ML), the ordinary cross-validation (OCV), the generalized cross-validation (GCV), and the L-curve methods, as well as the discrepancy principle (DP). As a comparison, the problems are also solved using the SVD.

The results for the case where 6 sensor points with $\sigma=0.9$ are plotted in Figure 4.1.2. Table 4.1.2 shows that the regularization parameter chosen in each method and its corresponding least square error. From Figure 4.1.2a and Table 4.1.2, we can see that the DP gives the best estimate of α , and the OCV and GCV give the next best estimate, though not as good as the DP. As shown in Figure 4.1.2b, the L-curve plots form the L-curve shape, but the regularization parameter at the convex point, which is taken as an optimal parameter, is smaller than in the analytic solution. The ML and ML-variance do not perform very well. Once again, from Figure 4.1.2c, the DP gives a good estimate of the singularity threshold value in comparison with $S(\text{analytic})$.

For other cases with the different number of sensors and error level, the DP is consistently the best estimate for choosing the regularization parameter. For some cases of the OCV and GCV, two local minimum points are shown. But overall, the OCV and GCV are the second and third best methods to find α . In all of the cases, the ML, ML-variance, and the L-curve find the smaller value than in the analytical solution.

Calculations with sensors located at different positions were also made. Two extreme cases are presented in Figure 4.1.3. In Figure 4.1.3a, the sensors are distributed evenly around the center of the square, and the SVD method outperforms the regularization method. In the case of the regularization method, the total net heat balance still holds, though the heat flux is not distributed symmetrically. In Figure 4.1.3b, on the other hand, the sensors are placed at the middle plane horizontally, and the regularization method performs much better than the SVD method.

4.1.1.2 Axisymmetric over-specified problems

This problem is concerned with an annular, homogeneous, isotropic, planar region between two concentric circles (see Fig.4.1.4) with nondimensionalized radii, where $\rho_a = 0.5$ and $\rho_b = 1.2$ is considered. This problem is chosen to test the algorithms for the problems with over-specified boundary conditions on the outer boundary. The discretization uses linear elements with $l=10$ along the azimuthal direction and $m=24, 36, 48$ in the radial direction.

The analytic solution for the well-posed problem, where $u_a=0.5$ and $u_b=1$, gives the results of $q_a = -1.142$, and $q_b = 0.4759$ [Martin *et al.*, 1996]. For the numerical inverse solution, the outer boundary was over-specified with both constant temperature and flux

boundary conditions taken from the analytic solution of the well-posed problem, while nothing was specified on the inner circular boundary. We use the case with $m=36$ and $\sigma=0.05$ in the following results of the Levenberg-Marquardt method, the regularization method, and the SVD.

Results of the Levenberg-Marquardt method

We use both the Ozisik's implementation (Sec. 3.2.3.1) and the More's implementation (Sec. 3.2.3.2) for the Levenberg-Marquardt method. The Ozisik's implementation diverges and fails every time. On the other hand, the More's implementation, which is used for the first time in the inverse heat transfer problems, is likely to converge to the optimal solution.

The behavior of our algorithm is shown in Table 4.1.3. We chose $\Delta^0=3.3$ and $\lambda=0.5$. The initial estimate of the heat flux is set at $q^0=0$. Selecting the appropriate Δ^0 and λ is critical to obtain an optimal heat flux solution. If Δ^0 is too small or λ is too large, the algorithm may miss a chance to reach an optimal solution. On the other hand, if Δ^0 is too large or λ is too small, the solution may be far from the optimal solution. Table 4.1.3 shows that the complementarity condition (Eq. (3.2.47)) holds at every iteration. The damping parameter, μ , increases as the trust-region radius, Δ , is gradually reduced. This result mathematically makes sense because $\|P(\mu)\|$ is a decreasing function. In the Ozisik's implementation for choosing the damping parameter, the parameter is gradually reduced as the iteration procedure advances to the optimal solution. This decreasing damping parameter is considered one of the reasons for the divergence of the calculation, which occurs likely in the Ozisik's implementation, because $\|P(\mu \rightarrow 0)\| = +\infty$ in the ill-

posed problems. However, the Ozisik's implementation may work well if the problem is well-posed. In this example problem, we find that the problem is ill-posed because $s_{\min} = 0$. Table 4.1.3 also shows that our calculation is efficiently performed to converge to the solution.

Results for the regularization method and the SVD

Table 4.1.4 shows the regularization parameters chosen in each method and the corresponding least square error. The error indicators for the regularization parameter are shown in Figure 4.1.5a. It is clear that the regularization parameter chosen by the DP method is close to that chosen by the analytical solution $S(\alpha)$. The parameter chosen by OCV and GCV is closed to the optimal parameter, and its corresponding least square error is reasonably small. The regularization parameters selected by $ML(\alpha)$, and its variance, $ML_{variance}(\alpha)$, are smaller than the optimal regularization parameter selected by $S(\alpha)$. The regularization parameter corresponding to the corner points in the L-curve (Fig. 4.1.5b) is $\alpha=3.64 \times 10^{-3}$, and that is far from that given by $S(\alpha)$ or $DP(\alpha)$. Similar results were found for $m=24, 36$, and 48 and other conditions. Figure 4.1.5c shows the error indicator for the singularity threshold. For this problem, the minimum points of the discrepancy principle (DP) match well with the analytic solution.

Comparison of the accuracy of the three methods

Figures 4.1.6a, 4.1.6b, and 4.1.7c plot the errors associated with the heat fluxes and the inner temperatures with $m=36$ and $\sigma=0.05$ calculated in the Levenberg-Marquardt method with the More's implementation, the regularization method and the SVD method

using the optimal α and τ , determined using the DP. The SVD method outperforms the regularization method and the Levenberg-Marquardt method in both the inner temperatures and inner heat flux solutions. The solution of the regularization method has the same level of accuracy of the Levenberg-Marquardt method. We note that the parameters chosen in each method play a central role of the comparison of the accuracy.

4.1.2 Transient problems

In this section, the inverse transient heat transfer calculations are discussed. We first use an example problem where the analytic solution is available. Then we consider another case where heat flux solution contains discontinuities and sharp corners. Toward this end, both function estimation and parameter estimation are used for the inverse calculations. We also consider the case where the sensor temperatures have measurement errors. The accuracy of the computational solution is then discussed.

4.1.2.1 Example problem 1

Direct problem

Solutions to the diffusion equation (the Fourier equation)

$$\frac{\partial T}{\partial t} - \alpha \frac{\partial^2 T}{\partial x^2} = 0 \quad (4.1.6)$$

are obtained in the spatial interval $0.1 \leq x \leq 1.0$, with boundary conditions

$$\frac{\partial T}{\partial x} = c = 2 - 2\pi \sin(0.05\pi) \exp\left[-\alpha \left(\frac{\pi}{2}\right)^2 t\right] \text{ at } x=0.1 \quad (4.1.7a)$$

and

$$T = 2 \quad \text{at } x=1 \quad (4.1.7b)$$

Initial conditions are chosen to be

$$T = 2x + 4 \cos(0.5\pi x) \quad (4.1.8)$$

This problem has the following exact analytic solution [Fletcher, 1991]:

$$T = 2x + 4 \cos(0.5\pi x) \exp\left[-\alpha \left(\frac{\pi}{2}\right)^2 t\right]. \quad (4.1.9)$$

First, the direct problem of the above example is solved. For this purpose, the Galerkin finite element method is used to discretize with linear elements. The FEM calculations are then compared to the exact analytic solution (Eq.4.1.9). Note that assessing the accuracy of the FEM calculations is not the objective of the present research. However, we use the FEM calculations for computing the sensitivity coefficients for both the regularization method and the SVD. The truncation error and the round-off error of the FEM calculations may cause a significant error to the inverse calculations. Thus, the FEM calculations should have a good agreement with the exact analytic solutions for the accurate calculations of the inverse problem. In the FEM calculations, we use $\Delta t=0.01$, $\Delta x=0.1$, and $\alpha =1$ to compute the temperature distribution in the above example problem. We utilize the following RMS error to assess the accuracy of the FEM calculations:

$$RMS_{error} = \left[\left(\sum_{i=1}^N (T_{i,analytic} - T_{i,computed})^2 \right) / N \right]^{1/2} \quad (4.1.10)$$

where N is the number of nodes, $T_{i,analytic}$ is the temperature for the analytic solution (Eq.4.1.9), and $T_{i,computed}$ is the computed temperature. At $t=2s$, the RMS error is 4.368×10^{-4} . Thus, the FEM calculations with $\Delta t=0.01$ and $\Delta x=0.1$ are considered to be accurate enough to be utilized for the inverse calculations.

Inverse problem

We utilize the above direct problem for formulating the inverse calculations and assessing the accuracy. First, we assume that the boundary condition at $x=1$ is not specified. However, the temperatures (sensor temperatures) in the interior domain are specified using the analytic solution (Eq.4.1.9). Then, the heat flux at $x=1$ is found using the inverse calculations.

The mesh discretization ($\Delta t=0.01$ and $\Delta x=0.1$) is the same as the one used in the direct problem. Sensor temperatures are located at $x=0.1, 0.4,$ and 0.7 , which are specified at every 0.1 seconds. The sensor temperatures are computed using the analytic solution (Eq.4.1.9). The final time is set at $t=2s$. Therefore, the total number of sensor temperatures in the inverse calculations is $60 (20 \times 3)$ points. We also examine the case where the sensor temperatures have measurement errors. We assume that the measurement errors of the sensor temperatures follow a normal distribution $N(T, \sigma)$. Thus, the sensor temperatures are determined using Eq.(4.1.2). The standard deviation σ is set at $\sigma=0.01$ in the inverse calculations. The heat flux at $x=1$ is computed using the temperature (Eq. 4.1.7b) at $x=1$, the initial temperatures (Eq.4.1.8), and the sensor temperatures located at $x=0.1, 0.4, 0.7$. To this end, we use the regularization method and the SVD along with the function estimations and the parameter estimations. The accuracy of the heat flux solutions are then assessed using the analytic solution shown in Eq.(4.1.7a).

Function estimation and parameter estimation

There are two methods for estimating the inverse solutions: a *parameter estimation* and a *function estimation* approach. The function estimation is the method that has no use of functions. In other words, the function estimation is an estimation approach in an infinite dimensional space of functions. Therefore, the function estimation is time-consuming and computationally expensive, but often accurate. On the other hand, in the parameter estimation, functions such as polynomial functions are used for approximating the solution. In general, the parameter estimation is expressed in the following linear form:

$$g(t) = \sum_{j=1}^N P_j C_j(t) \quad (4.1.11)$$

where $C_j(t), j=1, \dots, N$ are known trial functions. The N unknown parameters $P_j, j=1, \dots, N$, are estimated in the inverse calculations. In the parameter estimation, appropriately chosen trial functions are required to obtain an accurate solution; however, the calculation time and computer's memory usage is drastically reduced.

We use four approaches to estimate the heat flux solutions in the above example problem.

Approach 1. Function estimation (200unknowns)

No function is assumed prior to the calculations. Therefore, each time step has a different heat flux. Since the number of time step is 200, we find 200 unknown heat fluxes in the following manner:

$$\begin{array}{ll} q_1 & \text{for } 0 \leq t \leq 0.01 \\ q_2 & \text{for } 0.01 \leq t \leq 0.02 \\ \vdots & \end{array}$$

$$q_{200} \quad \text{for } 1.99 \leq t \leq 2.0$$

Approach 2. Function estimation (20 unknowns)

To reduce the calculation time and also stabilize the inverse calculations, the number of unknown parameters in approach 1 is reduced to 20 using the following assumption:

$$\begin{aligned} q_1 = q_2 = q_3 = \dots = q_{10} & \quad \text{for } 0 \leq t \leq 0.1 \\ q_{11} = q_{12} = q_{13} = \dots = q_{20} & \quad \text{for } 0.1 \leq t \leq 0.2 \\ & \quad \vdots \\ q_{191} = q_{192} = q_{193} = \dots = q_{200} & \quad \text{for } 1.9 \leq t \leq 2.0 \end{aligned}$$

Approach 3. Parameter estimation (Cubic polynomial approximation)

Polynomial functions are well-known trial functions to approximate solutions for any engineering problems. Ozisik (1993) used the cubic polynomial function to approximate the temperature distribution for the integral methods of the direct heat conduction problem. His experience has shown that there is no significant improvement in the accuracy of the solution to choose a polynomial greater than the fourth degree. However, when the solution contains discontinuities and sharp corners, the polynomial functions may not be a proper choice for approximating the solution. The reason is obvious. We use the following cubic polynomial representation to approximate the heat flux solutions:

$$q = q_1 + q_2 \times t + q_3 \times t^2 + q_4 \times t^3 \quad (4.1.12)$$

where t is time. The parameters, q_1, q_2, q_3, q_4 , are found using the inverse calculations.

Approach 4. Parameter estimation (Prior knowledge of the function)

This time, we assume that the form of the function for the analytic solution is known prior to the inverse calculations. Since the analytic solution is available in Eq. (4.1.7a), we use the exponential equation for an approximate solution. Thus, the parameters (q_1, q_2, q_3) are estimated in the following equation:

$$q = q_1 + q_2 \exp(q_3 \times t) \quad (4.1.13)$$

This equation is not a linear form of the parameter estimation shown in Eq. (4.1.11) because two unknown parameters coexist in the second term in the right hand side of Eq. (4.1.13). This second term causes the nonlinearity of the inverse calculations. Note that based on the definition of the nonlinear inverse problem, the sensitivity coefficients have some functional dependence on unknown parameters q . The sensitivity coefficients are computed as

$$J_{ij} = \left[\frac{\partial T_i^T(q)}{\partial q_j} \right]^T = \left[\frac{\Delta T_i(q)}{\Delta q_j} \right]^T$$

In the case of Eq. (4.1.13), to compute $J_{i2} = \left[\frac{\Delta T_i(q)}{\Delta q_2} \right]^T$, the change of q with respect to the change of q_2 varies with the value of q_3 . Thus, the sensitivity coefficient J_{i2} is dependant on the value of q_3 . Similarly, the sensitivity coefficient J_{i3} is dependant on the value of q_2 . Hence, the parameter estimation with Eq. (4.1.13) will be a nonlinear inverse problem. In the nonlinear problem, the SVD is not applicable since the SVD is utilized only for the linear inverse problem.

Results for example problem 1

Results with no input error ($\sigma=0$)

First, we consider the case where the sensor temperatures are the same as the exact analytic solutions (Eq.4.1.9) ($\sigma=0$). Figure 4.1.7a shows the heat flux solution using the regularization method for the function estimations (Approach 1 and Approach 2) and the parameter estimations (Approach 3 and Approach 4). Figure 4.1.7b shows the corresponding percentage errors to the analytic solution. The function estimation with 200 unknowns (Approach 1) is less accurate than the one with 20 unknowns (Approach 2). Too many unknown parameters are considered to make the calculations unstable. We also find that the parameter estimation with the exponential function (Approach 4) has the best accuracy, and the cubic polynomial function has moderate accuracy (Approach 3).

Figure 4.1.8a shows the heat flux solution using the SVD. Figure 4.1.8b shows the corresponding percentage errors to the analytic solution. Note that the SVD with the exponential function (Eq.4.1.13) is not applicable (See section of Approach 4). Similar to the results of the regularization method, the parameter estimation with the cubic polynomial function is better than the function estimations.

Table 4.1.5 shows the sum of the square errors for each approach of solution with respect to the regularization method and the SVD. The sum of the square errors is calculated as

$$S(\alpha) = \sum_{j=1}^M [q_j(\alpha) - q_{j,analytic}]^2 \quad (4.1.14)$$

where M is the number of the time step, $q_{j,analytic}$ is the analytic heat flux, and $q_j(\alpha)$ is the computed heat flux. As seen in Table 4.1.5, the accuracy of the heat flux solutions for the

SVD is largely the same as the regularization method. Thus, the SVD is comparative to the regularization method when there is no measurement error at the sensors.

Results with input data noise $\sigma=0.01$

In this case, the sensor temperatures have measurement errors of $\sigma=0.01$. Figure 4.1.9a shows the heat flux solution using the regularization method. Figure 4.1.9b shows the corresponding percentage errors to the analytic solution. Figure 4.1.10a shows the heat flux solution using the SVD. Figure 4.1.10b shows the corresponding percentage errors to the analytic solution. Table 4.1.6 shows the sum of the square errors for each approach of solution with respect to the regularization method and the SVD. As seen from Fig. 4.1.10a, the heat flux solution of the function estimation with 200 unknowns (Approach 1) in the SVD exhibits oscillatory behavior. The SVD is comparative to the regularization method when the parameter estimation with the cubic polynomial function (Approach 3) and the function estimation with 20 unknowns (Approach 2) are used. The function estimation with 20 unknowns (Approach 2) works better than 200 unknowns (Approach 1) in both the regularization method and the SVD. In addition, the parameter estimations are more accurate than the function estimations.

Figure 4.1.11 shows the error indicators that show the optimal regularization parameter. Figure 4.1.12 shows the discrepancy principle (DP) for the SVD to choose the optimal singularity threshold. Note here that the vertical axis means the error indicator, but not the true errors. The minimum point of each curve shows the optimal regularization parameter or the optimal singularity threshold chosen by the each method. The accuracy of the each calculation is assessed and compared using the analytic

solution. As seen from Fig.4.1.11, the regularization parameters at the minimum points of discrepancy principle (DP) and the OCV are located close to that of the analytic solution. Thus, the discrepancy principle (DP), the OCV, and the GCV are the best methods to choose the optimal regularization parameter. The minimum point of the variance based on the maximum likelihood method indicates that $\alpha=1 \times 10^{-5}$ is an optimal parameter. However, this value is much smaller than that of the analytic solution. The regularization parameters selected by $ML(\alpha)$, and its variance, $ML_{variance}(\alpha)$, are smaller than the optimal regularization parameter. As for the singularity threshold (Fig. 4.1.12), the minimum points of the DP are located at $\tau < 2.5 \times 10^{-1}$, which match well with the analytic solution. Similar results are found for other cases for the regularization parameters and the singularity threshold values.

4.1.2.2 Example problem 2

We then consider the case where the heat flux solution contains discontinuities and sharp corners. We use the same geometry as in the example problem 1. This time, however, no exact analytic solution is available. We consider the diffusion equation (Eq.4.1.6) in the spatial interval $0.1 \leq x \leq 1.0$ with the following boundary conditions:

$$\frac{\partial T}{\partial x} = \begin{cases} 1 & \text{for } 0 \leq t < 0.7 \\ 2 & \text{for } 0.7 \leq t \leq 1.3 \\ 1 & \text{for } 1.3 < t \leq 2 \end{cases} \quad \text{at } x=0.1 \quad (4.1.15a)$$

and

$$T = 2 \quad \text{at } x=1 \quad (4.1.15b)$$

Initial conditions are

$$T = 2 \tag{4.1.16}$$

First, we solve the above direct problem using the FEM calculations. Because the exact analytic solution of this problem is not available, we cannot assess the accuracy of the FEM calculations. But we verified that our FEM calculations were accurate in the example problem 1. In addition, we choose the same discretization ($\Delta t=0.01$, $\Delta x=0.1$, and $\alpha=1$) as in the example problem 1. Therefore, the temperature distribution obtained by the FEM calculations is considered accurate.

In the inverse calculations, the problem formulation is the same as the example problem 1. That is, we assume that heat flux distribution at $x=0.1$ is not specified prior to the calculations. However, the sensor temperatures at $x=0.1$, 0.4 , 0.7 are specified every 0.1 seconds. In this case, the sensor temperatures are found by the FEM calculations since the exact analytic solution is not available. We use the temperature (Eq.4.1.15b) at $x=1$, the initial temperatures (Eq.4.1.16), and the sensor temperatures for our inverse calculations. Then the heat flux at $x=0.1$ is found by the regularization method and the SVD. We use the same approaches for estimating the solution as in the example problem 1, which are Approach 1 and 2 for function estimations and Approach 3 for the parameter estimation.

Results for example problem 2

Figure 4.1.13 shows the heat flux solution for the regularization method with no input error ($\sigma=0$). The figure shows that the function estimations (Approach 1 and 2) are superior to the parameter estimation (Approach 3). Since the analytic heat flux solution (Eq. 4.1.15a) contains discontinuities and sharp corners at $t=0.7$ and 1.3 , it is

mathematically impossible to obtain an accurate solution if the cubic polynomial function is utilized. Tables 4.1.7 and 4.1.8 show the sum of the square errors for the each approach in $\sigma=0$ and $\sigma=0.01$ respectively. The regularization method is comparative to the SVD in all of the solution approaches except for the function estimation with 200 unknowns with the noisy sensor temperatures ($\sigma=0.01$). This result holds true for the example problem 1. We also examine the error indicators for the optimal regularization parameter and the discrepancy principle (DP) for the singularity threshold. The similar results to the example problem 1 are found.

4.1.3 Heat flux for 3-D spray cooling of electronic components

With the knowledge gained above and the regularization inverse algorithm, we have studied the inverse heat transfer that occurs during the cooling of a microchip by a liquid droplet stream sprayed from a nozzle above the microprocessor. Microthermal sensors were embedded in the microprocessor package at various locations, and readings are recorded (Figure 4.1.14). The details of the design and experimental setup are given in a recent paper [Schwarzkopf *et al.*, 2004]. The boundary conditions are as follows:

$$\begin{aligned} \frac{\partial T}{\partial x} &= 0 && \text{at } x=0, 21\text{mm} \\ \frac{\partial T}{\partial y} &= 0 && \text{at } y=0, 21\text{mm} \\ \frac{\partial T}{\partial z} &= 0 && \text{at } z=0 \\ h=500\text{w/m} &&& \text{in } \begin{cases} 0 < x < 4.2\text{mm}, 16.8\text{mm} < x < 21\text{mm} \\ 0 < y < 4.2\text{mm}, 16.8\text{mm} < y < 21\text{mm} \\ z = 0.15\text{mm} \end{cases} \end{aligned}$$

where h is a heat transfer coefficient. With the measurements and locations known, the above inverse heat transfer scheme is then applied to determine the heat flux distribution over the surface ($4.2\text{mm} < x, y < 16.8\text{mm}$, $z = 0.15\text{mm}$) upon which the droplet sprays are impinged. The inverse heat transfer model is 3-D. The FEM mesh used in the calculation is shown in Figure 4.1.15. Approximately 30 temperature points are measured in each case. The heater is located in $8.4\text{mm} < x, y < 12.6\text{mm}$, at $z = 0\text{mm}$. Note that the number of sensors and the rate of the heat capacity of the heater vary with the cases. One hundred unknown parameters for the heat flux are computed in inverse calculations. For this practical problem, the Levenberg-Marquardt method is not applicable because the proper values of Δ^0 and λ cannot be estimated. The regularization method and the SVD method are utilized to find the heat flux distribution on the cooling surface. Results from the regularization method and the SVD method, with both the regularization parameter α and the singularity threshold τ determined using the discrepancy principle (DP), give the same heat flux distribution. Figure 4.1.16, for example, shows the scheme to choose the singularity threshold. The singularity threshold corresponding to the minimum point is chosen as an optimal parameter. One typical heat flux distribution calculated by this algorithm is shown in Figure 4.1.17a. Figure 4.1.17b shows a comparison of the experimentally measured and inversely computed results for temperature distribution at the centerline.

4.2 INVERSE DESIGN SOLIDIFICATION PROBLEMS

The regularization method is implemented for the inverse design of solidification processes. The design problem is set up to find optimal heat flux solutions for controlling

the solid-liquid interface. The inverse steady-state solidification problems are first discussed and solved using the function estimation analysis. Then, the inverse transient solidification problems are solved with the parameter estimation approach.

4.2.1 Steady-state solidification problems

The direct and inverse algorithms described above enable the prediction of solid-liquid interface shape for a set of given boundary conditions and of the heat flux distribution along the boundaries for a prescribed interface movement. We consider a steady-state solidification problem with natural convection in a square cavity. We test four different cases using the material properties shown in Table 4.2.1. The cavity geometry and the mesh discretization are shown in Fig. 4.2.1. In addition, gravity force, 9.8 m/s^2 , is applied downward.

Case 1. Direct Problems

Let us first consider a simple case where solidification occurs in a square cavity ($0 \leq x \leq 0.02\text{m}, 0 \leq y \leq 0.02\text{m}$) under the prescribed conditions along the wall. Quadrilateral linear elements, 22×44 , are used in each solid region and liquid region (Figure 4.2.1). The cavity is thermally insulated at the top and bottom. The left wall is fixed at a constant temperature ($T=2280\text{K}$) above the melting point ($T_m=2243.15\text{K}$), while the right side wall is fixed at $T=2220\text{K}$. Figures 4.2.2 illustrate the quasi-steady state velocity and temperature fields in the cavity. For this problem, it is shown that the moving interface is strongly affected by the convection in the cavity. If the natural convection were not present, a vertical solidification front would have been achieved.

Case 2. Inverse calculation: prescribed vertical interface

In this case, the location of the solid-liquid interface is specified as a vertical straight line. The interest is in the prediction of heat flux distribution required to achieve this interface. Boundary conditions are the same as those in Case 1 except for the heat flux at the right side. Theoretically, this means that heat flux distribution is designed and tuned such that the effect of natural convection on solid-liquid interface is eliminated. The inverse computational model described above is used to obtain the desired heat flux distribution at the right side wall. More specifically, we use the regularization method along with the appropriately selected regularization parameter using the L-curve method. For this problem, the interface position is set at $x= 0.01\text{m}$ from the left wall. We use the proposed finite difference scheme (Eq. 3.3.20) for computing the sensitivity coefficients in this case. Since this problem is a nonlinear regularization problem, convergence rates of the calculations are investigated, which are shown in Figure 4.2.3. We use the following norm of error of the solidification distance as a convergence criterion:

$$S_l = \left[\sum_{i=1}^M (\bar{d}_i - d_i)^2 \right]^{1/2} \quad (4.2.1)$$

where S_l is the norm of error of the solidification distance, d_i is the calculated solidification interface distance, \bar{d}_i is the ideal solidification interface distance, and M is the total number of controlled solidification distance. The figure shows that the smaller the regularization parameters are, the slower but more accurate the convergence is. The figure also shows that after the third iteration, the reduction in norm of error is slower. Thus, we consider that the inverse calculation reaches its convergence in the first three

iterations. To determine the appropriate parameter α for this problem, simulations are made to generate the L-curve plots following the procedures described in section 3.3.3. Figure 4.2.4 plots the L-curve, which clearly shows a turning point in the error chart. Corresponding to the turning point is the optimal α for this problem. Thus, the regularization parameter α of 5×10^{-20} is chosen for an optimal regularization parameter. Figures 4.2.5a and 4.2.5b show the inversely calculated heat flux distribution and the percentage error distribution of solidification distance for this problem. Heat fluxes found in the regularization parameter are considered to be accurate, because the percentage errors of solidification distance are very small, which are within $\pm 0.2\%$. The predicted heat flux distribution varies strongly from the bottom to the top. The temperature distribution and fluid flow field calculated by this heat flux distribution are plotted in Figures 4.2.5c and 4.2.5d. It is shown that a vertical solid-liquid interface is achievable and the natural convection effect is balanced by the inversely determined heat flux distribution.

Case 3: Inverse calculation: prescribed sine curve

As another example, a solid-liquid interface is specified as a sine curve shown in Figure 4.2.6. We first compare the proposed finite difference scheme (Eq. 3.3.20) to the conventional finite difference scheme (Eq. 3.3.16) for computing the sensitivity coefficients. Figures 4.2.7a and 4.2.7b show the convergence rate for two methods: (a) the conventional finite difference scheme with $\lambda=0.01$ and (b) the proposed finite difference scheme with $\varepsilon=0.01$. The figures show that the convergence rate for the conventional method is unstable and less accurate than the one for the proposed method.

The regularization parameter α of 5×10^{-20} is chosen as an optimal regularization parameter using the L-curve method for the proposed finite difference scheme (Fig. 4.2.8). The calculated heat flux distribution along the right side wall is depicted in Figure 4.2.9a. The heat flux is distributed in a wide range, and there are four points at which the heat flux is crossing zero. This heat flux distribution is considered causing the unstable and inaccurate convergence for the conventional finite difference scheme (Figure 4.2.7a). Figure 4.2.9b shows the percentage errors between the calculated and prescribed solidification distance. It is shown that the percentage error between the prescribed solidification distance and the calculated solidification distance is within $\pm 0.15\%$. This percentage error is largely the same as in the case2. The calculated fluid flow and temperature distribution using the inversely determined heat flux distribution are given in Figures 4.2.9c and 4.2.9d. These figures show that the solid and liquid regions are separated with the prescribed sinuous interface, and our calculated heat flux distribution is feasible to control the solid-liquid interface.

Case 4: Inverse calculation: prescribed sharp sine curve

We prescribed another sinuous solid-liquid interface (Figure 4.2.6). This time, the frequency of the sine curve is double of that of the case 3. That is, this curve has sharper bending curve than that in the case3. First, we study the convergence rate (Figure 4.2.10) with respect to different values of the regularization parameter using the proposed finite difference scheme. When the regularization parameter is less than 1×10^{-18} , the convergence of the inverse calculation becomes unstable. One of the reasons of the unstable convergence stems from the infeasible heat flux solution. Figure 4.2.11 shows

the temperature distribution obtained by using the heat flux solution at the second iteration of $\alpha=1\times 10^{-18}$. This figure shows that the temperature distribution above the melting point ($T_m=2234.15\text{K}$) is observed in the solid region (right side). Since we specified in the FEM code that the right side is a solid region, this temperature distribution is not acceptable. That is, the heat flux solution in the calculation is not practical. This unpractical heat flux distribution is used in the next iteration (third iteration) for calculating the sensitivity coefficients. The calculation results in discontinuous sensitivity coefficients. Figure 4.2.12 shows that the sensitivity coefficients $J_{i,20}$ in the second iteration of $\alpha=1\times 10^{-17}$ and 1×10^{-18} . Note that the sensitivity coefficient $J_{i,20}$ is derived as $J_{i,20} = \frac{\partial d_i}{\partial q_{20}}$ where q_{20} is heat flux located at $y=8.864\times 10^{-3}\text{m}$ in the right wall, and d_i is controlled parameter (solidification distance). There are the discontinuous sensitivity coefficients for $\alpha=1\times 10^{-18}$, whereas the sensitivity coefficients for $\alpha=1\times 10^{-17}$ are continuous. The same holds true for other sensitivity coefficients. The unpractical heat flux solution leads to the discontinuous values, and ultimately to the unstable convergence for $\alpha=1\times 10^{-18}$. Figure 4.2.13 shows the L-curve plots. These plots do not form the L-curve shape. Thus, we cannot use the L-curve method to select the optimal regularization parameter. We select the regularization parameter α of 5×10^{-18} , where the sum of the square errors of the solidification distance is minimum. The calculated heat flux distribution along the right wall is depicted in Figure 4.2.14a. Figure 4.2.14b shows the percentage errors between the calculated and prescribed solidification interface distance. The maximum percentage error, 1.5%, in this case is ten times larger than in case3. This shows that the sharp bending curve for the prescribed solid-liquid interface is

more difficult for the inverse algorithm to control than the case where the slowly bending curve is specified. The calculated fluid flow and temperature distribution using the inversely determined heat flux distribution are given in Figure 4.2.14c and 4.2.14d. These figures show that the heat flux solution is feasible to obtain the prescribed sinuous solid-liquid interface.

4.2.2 Transient solidification problems

In this section, an inverse design of solidification processes with natural convection is solved to find the optimal heat flux solution for the solid-liquid interface specified to move at a constant velocity. The regularization method along with the L-curve method is implemented for the inverse design solidification problem.

We address the direct transient solidification problem with convection and briefly discuss the effects of convection on the solid-liquid interface (Sec. 4.2.2.2). Finally, we consider the inverse design of solidification processes (Sec. 4.2.2.3). In particular, the initial conditions of the transient solidification problem is obtained using the inverse steady-state solidification problem without convection (Sec.4.2.2.1)

In the direct and inverse problems, we consider a solidification process for pure aluminum confined in a square mold ($0 \leq x \leq 0.02m, 0 \leq y \leq 0.02m$). The properties of aluminum are shown in Table 4.2.2. Quadrilateral linear elements, 22×44 , are used in the solid and liquid phases (Figure 4.2.1).

4.2.2.1 The inverse steady-state solidification problem without convection

First, the inverse steady-state solidification problem without natural convection is solved. The cavity is thermally insulated at the top and bottom. The left wall is fixed at $T=1220\text{K}$ above the melting temperature ($T_m=933\text{K}$). The gravity force is not applied. The locations of the desired solid-liquid interface for the case 1 and case 2 are specified in Figure 4.2.15. We find the optimal heat flux solution on the right wall required to achieve the prescribed solid-liquid interfaces. The direct problem is then solved to find the temperature distributions (Figure 4.2.16). These temperature distributions will be used for the initial conditions for the transient solidification calculations.

4.2.2.2 The direct problem for transient solidification processes with convection

The direct problems for transient solidification processes were studied extensively in [Shu *et al.*, 2002; Li *et al.*, 2003; Song *et al.*, 2002]. Our objective to present the direct problem in this paper is to see how much the initial vertical straight solid-liquid interface moves and deforms at a certain time by the effect of natural convection. The initial temperatures are given in Figure 4.2.16a. Thus, the solid-liquid interface for the initial state is located at $x=0.015$ vertically. The top and bottom walls are kept adiabatic. The temperature of the left wall is fixed at 1220K , which is above the freezing temperature ($T_m=933\text{K}$). The right side wall is imposed with heat extraction ($q=-4\times 10^6 \text{ W/m}^2$), which will cause the solidification to occur. The gravity force, 9.8m/s^2 , is applied downward. A constant time step $\Delta t=0.02\text{s}$ is selected with the final time of 0.6s . Figure 4.2.17 shows the velocity distribution at $t=0.6\text{s}$. For this problem, it is seen that the solid-liquid interface at $t=0.6\text{s}$ is strongly affected by the convection in the cavity. If the natural

convection were not present, a vertical solidification front would have been achieved. We also see that solidification occurs, and the solid-liquid interface moves leftward.

4.2.2.3 The inverse design of solidification processes

We tested two cases for the inverse design of solidification processes. In the case 1, the initial solid-liquid interface is a vertical straight line located at $x=1.5E-2$. The case 2 has a sinuous solid-liquid interface in the initial state shown in Figure 4.2.15. Initial temperatures in the case 1 and case 2 are given in Figure 4.2.16a and 4.2.16b, respectively. The top and bottom walls are kept adiabatic. The temperature of the left wall is fixed at 1220K. In addition to the above initial and boundary conditions, we also desire that solidification occurs with the prescribed growth conditions where a desired solid-liquid interface moves leftward with a prescribed constant velocity, $1.17 \times 10^{-2} \text{m/s}$. To this end, we specify the every node points at the solid-liquid interface in every 0.04 seconds for our inverse calculations. Note that the interfacial velocity must be constant for a non-uniform microstructure in the final casting. The gravity force, 9.8m/s^2 , is applied downward. A constant time step $\Delta t=0.02\text{s}$ is selected with the final time of 0.6s. Our interest is to find the optimal heat flux distribution on the right wall in order to obtain a prescribed solid-liquid interface. To evaluate the accuracy of our inverse algorithm, the standard deviation is utilized in the following equation:

$$\sigma_i = \left(\frac{\sum_{j=1}^M (\bar{d}_{ij} - d_{ij})^2}{M} \right)^{\frac{1}{2}} \quad (4.2.2)$$

where M is the number of controlled solidification distance for space. First, we use the parameter estimation with fourth order polynomial functions for time, t , and space, y . We also use the technique of scaling (See section 3.3.5). Thus, the approximate heat flux solution is

$$q = q_1(10^8 x^4) + q_2(10^6 x^3) + q_3(10^4 x^2) + q_4(10^2 x) + q_5(10^4 t^4) + q_6(10^3 t^3) + q_7(10^2 t^2) + q_8(10t) + q_9 \quad (4.2.3)$$

The parameters, $q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9$, are found in the regularization method with the L-curve method. If this inverse problem were solved in the function estimation (without using the function approximation), the number of unknowns would be 1,320 because there would be 44 unknowns for space and 30 unknowns for time. Thus, the parameter estimation with the polynomial function results in the significant reduction of the calculation time and computer memory.

We first examine the convergence for the non-linear regularization method. Figure 4.2.20 shows the convergence rates with respect to different values of the regularization parameter. Two iterations are enough for the regularization method to converge the calculations. Figure 4.2.19 shows the L-curve plots for this problem. There is a convex point located at $\alpha=1 \times 10^{-22}$. Thus, we select 1×10^{-22} as an optimal regularization parameter. Figure 4.2.20 shows that the heat flux solution using $\alpha=1 \times 10^{-22}$. Figure 4.2.21 is the standard deviation σ (Eq. 4.2.2) in each time step by using the heat flux solution. As seen from the figure, the error reaches the lowest point at $t=0.28$ and then shoots up until final time. Figure 4.2.22 shows the percent errors of the solidification distance at $t=0.6s$. The maximum percent error at $t=0.6s$ is 5 %. Since only nine unknown

parameters for the polynomial functions are solved, the heat flux distribution (Figure 4.2.20) found by the regularization method is a smooth and simple curve. To control the solid-liquid interface accurately, more complex curve of the heat flux distribution needs to be obtained. To this end, we use the sequential method and the whole time-domain method with the piecewise polynomial functions. These two methods make the inverse calculations more accurate, but time-consuming.

4.2.2.4 Sequential method

Due to the diffusion time between the boundary and the interface, it is critical to choose appropriate time domain for the sequential method. After doing numerical experiments, we found that 0.2 seconds was enough time to calculate the optimal heat flux solution. Thus, we divide the time domain into three parts. The first time domain ranges from 0 to 0.2s followed by the second time domain ($0.2s \leq t \leq 0.4s$) and the third time domain ($0.4s \leq t \leq 0.6s$). Each domain is calculated independently by using the regularization method. Thus, three calculations of the regularization methods are solved in a sequential manner. At the boundaries of the time domains ($t=0.2s$ and $0.4s$), the temperature and velocity distributions are carried over to the next time domains. That is, the temperature and velocity distributions at $t=0.2s$ calculated in the first time domain are used for the initial conditions in the second domain. The same holds true for the conditions at $t=0.4s$ for the second and third time domains. The following fifth-order scaled polynomial functions for time and space are used to find the optimal heat flux distribution:

$$\begin{aligned}
q &= q_1(10^8 x^4) + q_2(10^6 x^3) + q_3(10^4 x^2) + q_4(10^2 x) + \\
& q_5(10^4 t^4) + q_6(10^3 t^3) + q_7(10^2 t^2) + q_8(10t) + q_9
\end{aligned} \tag{4.2.4}$$

where $0 \leq x \leq 0.02, 0 \leq t \leq 0.2$. Nine unknown parameters, $q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8, q_9$, are computed in the inverse calculations for the each time domain. Thus, totally 36 unknown parameters are solved in the sequential method.

4.2.2.5 Whole time-domain method with piecewise polynomial functions

The piecewise polynomial functions are used to find the optimal heat flux solution. In the whole time-domain method, the following scaled piecewise polynomial functions are used:

$$\begin{aligned}
q &= q_i(10^8 x^4) + q_{i+1}(10^6 x^3) + q_{i+2}(10^4 x^2) + q_{i+3}(10^2 x) + \\
& q_{i+4}[10^4(t-a)^4] + q_{i+5}[10^3(t-a)^3] + q_{i+6}[10^2(t-a)^2] + q_{i+7}[10(t-a)] + q_{i+8}
\end{aligned} \tag{4.2.5}$$

$$\text{where } 0 \leq x \leq 0.02, \begin{cases} i = 1, a = 0, 0 \leq t \leq 0.2 \\ i = 10, a = 0.2, 0.2 \leq t \leq 0.4 \\ i = 19, a = 0.4, 0.4 \leq t \leq 0.6 \end{cases}$$

In the above equation, 36 coefficients of the polynomial function, q_1, q_2, \dots, q_{36} , are solved by the regularization method. Note that we use the same-scaled function as in the sequential method using Eq. (4.2.4). Thus, when the whole time-domain method is compared to the sequential method, the effect of scaling is neglected.

4.2.2.6 Convergence rate

We first examine the convergence for the sequential method and the whole time-domain method of the calculations of the regularization method. Figures 4.2.23 and 4.2.24 show the convergence rates for the first time domain in the case 1 for the sequential method and for the whole time-domain method with the piecewise polynomial function, respectively. In the sequential method, the calculations of the regularization method reach the convergence in the second iteration. The whole time-domain method needs as little as three iterations to reach the convergence.

4.2.2.7 L-curve method

To find an optimal regularization parameter, the L-curve method is used for our calculation. Figure 4.2.25 shows the L-curve plots for the first time domain of the sequential method in the case 1. Figure 4.2.26 shows the L-curve plots for the whole time-domain method with the piecewise polynomial function in the case1. These figures show the turning points, which are considered the optimal regularization parameters. The regularization parameters, $\alpha=1 \times 10^{-24}$ and 1×10^{-25} , are selected in the sequential method and the whole time-domain method, respectively. For other cases, the L-curve plots are also taken, and the optimal regularization parameters are chosen in the same manner.

4.2.2.8 Optimal heat flux solutions

Using the optimal regularization parameter found in the L-curve method, the optimal heat flux solutions are found for the case 1 and case 2 in both the sequential method and the whole time-domain method with the piecewise polynomial function.

Figures 4.2.27a and 4.2.27b show the optimal heat flux solutions in the case 1 for the sequential method and the whole time-domain method. These heat flux solutions are almost identical. Figures 4.2.28a and 4.2.28b show the optimal heat flux solution in the case 2 for the sequential method and the whole time-domain method. As seen from the figures, the heat flux solution for the whole time-domain method is almost identical to the heat flux solution for the sequential method. Since we divide the time domain for the heat flux solution into three parts, we have non-smooth functions with kinks for the heat flux distribution at the boundaries of the time domain ($t=0.2s, 0.4s$).

4.2.2.9 Validation of the inverse design solutions and comparison of the sequential method and the whole time-domain method

Figures 4.2.29 (a) and (b) show time history of the standard deviation σ (Eq. 4.2.2) for the error of the solidification distance in the case 1 and case 2, respectively. The standard deviation in the earlier time stage of each time domain ($t=0.2,0.4s$) is found to be less accurate in the sequential method. This is because the heat flux on the right wall diffuses toward the solid-liquid interface, and the diffusion time is required for a thermal front to penetrate the solid domain and reach the solid-liquid interface. Thus, it is physically difficult to control the solid-liquid interface immediately after the initial time. The error of the whole time-domain method with the piecewise polynomial functions is more damped and slightly accurate than that of the sequential method. Overall, the standard deviations for the error of distance in both the sequential and the whole time-domain method are small enough to consider that our inverse algorithm is reasonably accurate to control the solid-liquid interface. Figures 4.2.30 is the percent errors of the

solid-liquid distance at $t=0.6$ for the sequential method for case1. The percent errors of the sequential method are ten times smaller than in the whole time-domain method with nine unknowns (Figure 4.2.22). Figure 4.2.31 shows percent errors of the solidification distance at $t=0.6$ for the whole time-domain method in the case 2. The percent errors are less than 1%. Figures 4.2.32 and 4.2.33 show the velocity and temperature distributions at $t=0.2, 0.4, 0.6$ in the case 1 and case 2 for the sequential method by using the optimal heat flux distributions. The cold temperature distributions are shown at the upper right corner to eliminate the effect of natural convection. We also find that the solid-liquid interface moves leftward at a constant velocity as we specified in our calculation.

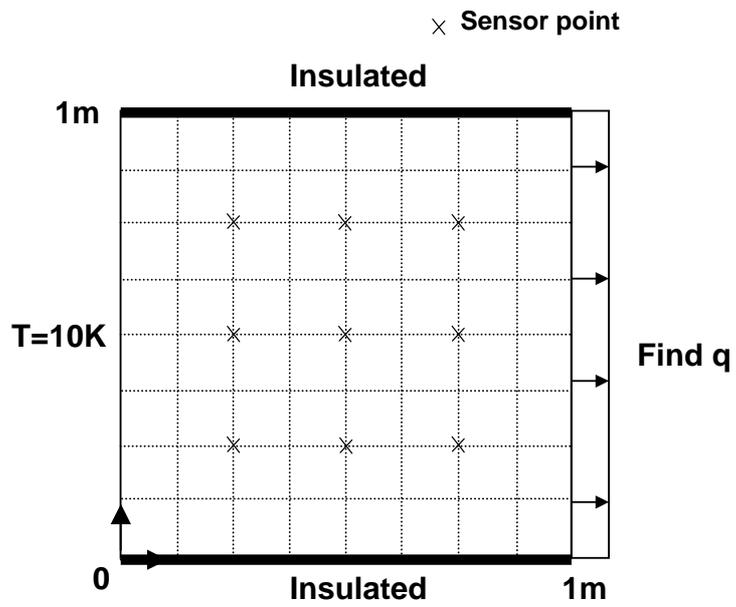
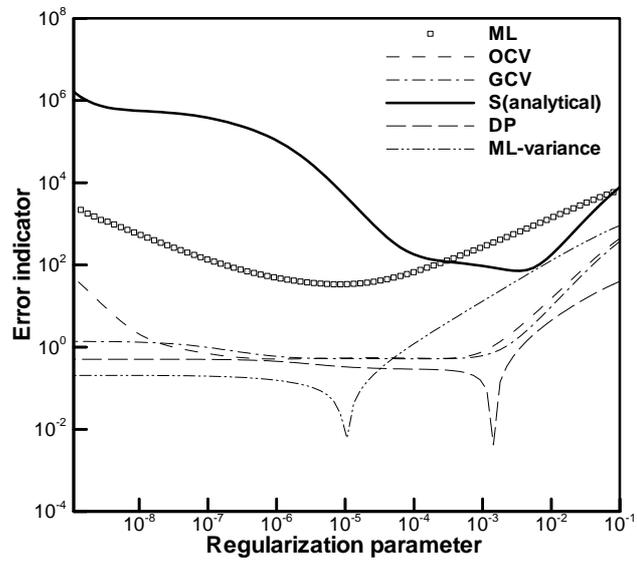
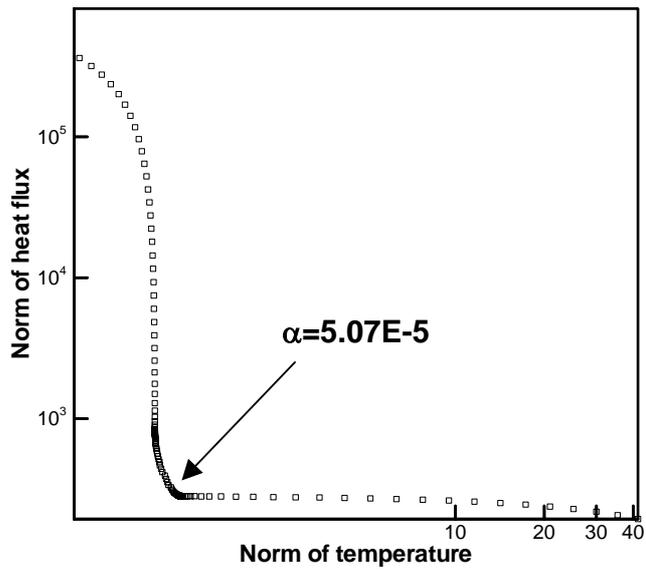


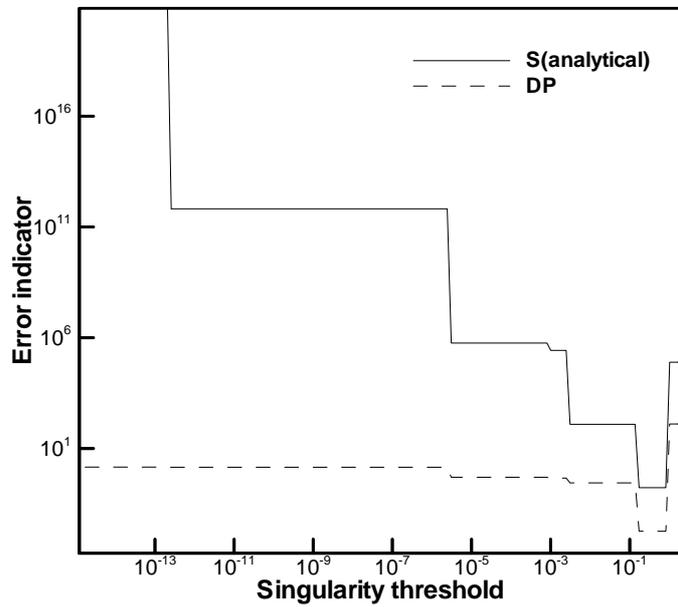
Figure 4.1.1. Illustration of the 2-D inverse heat conduction problem.



(a)

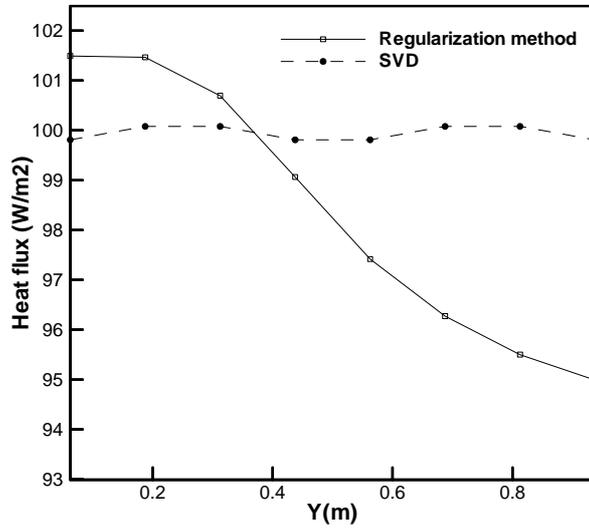


(b)

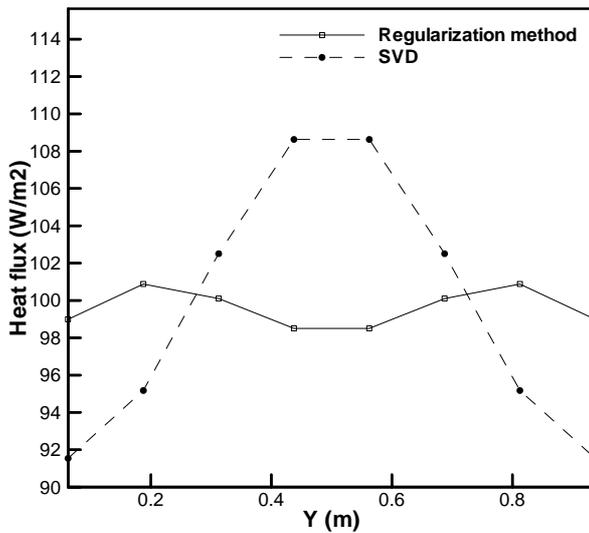


(c)

Figure 4.1.2. Inverse calculations for the 2-D heat conduction problems: (a) Error indicators vs. optimal regularization parameter, (b) L-curve method for optimal regularization parameter and (c) discrepancy principle determination of the singularity threshold for truncated SVD method.



(a)



(b)

Figure 4.1.3. Dependence of inversely calculated heat flux distribution on the distribution of sensor locations: (a) randomly spread sensors, and (b) specially organized sensors

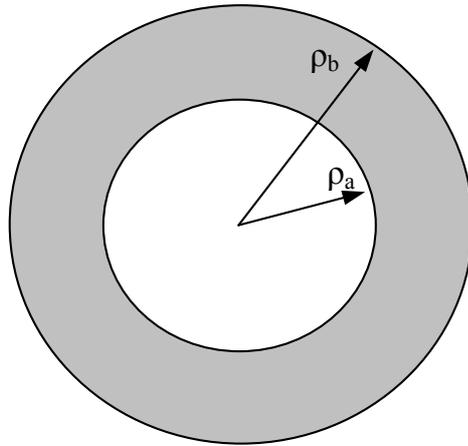
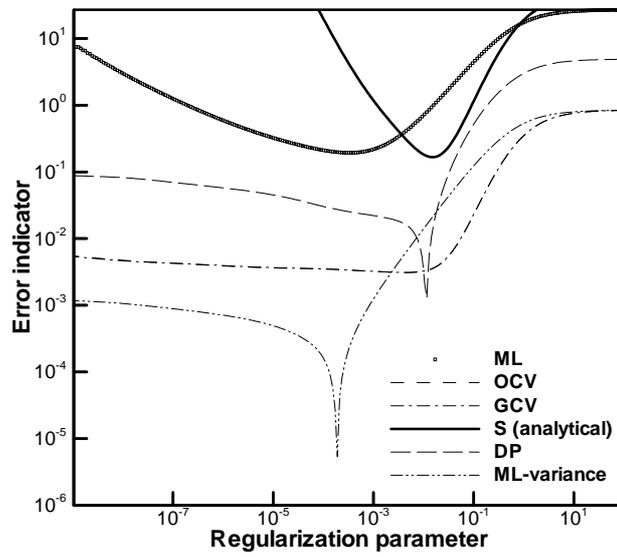
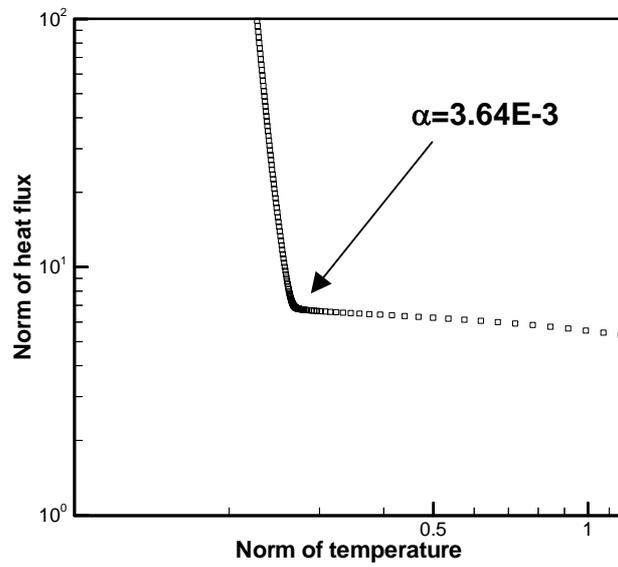


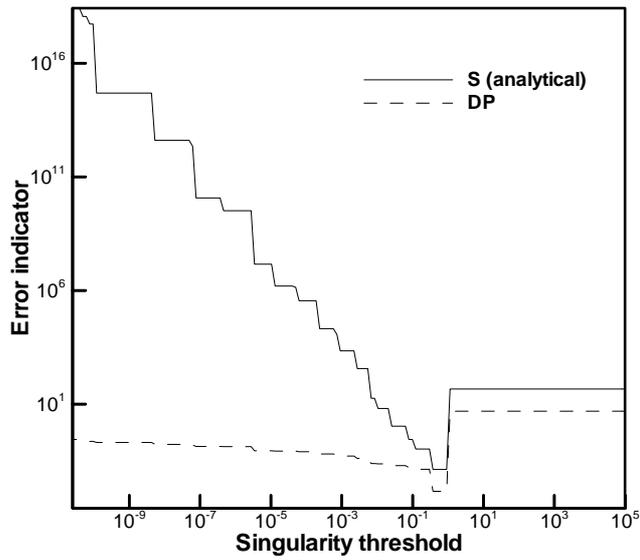
Figure 4.1.4. Schematic of concentric problem with over-specified boundary conditions.



(a)

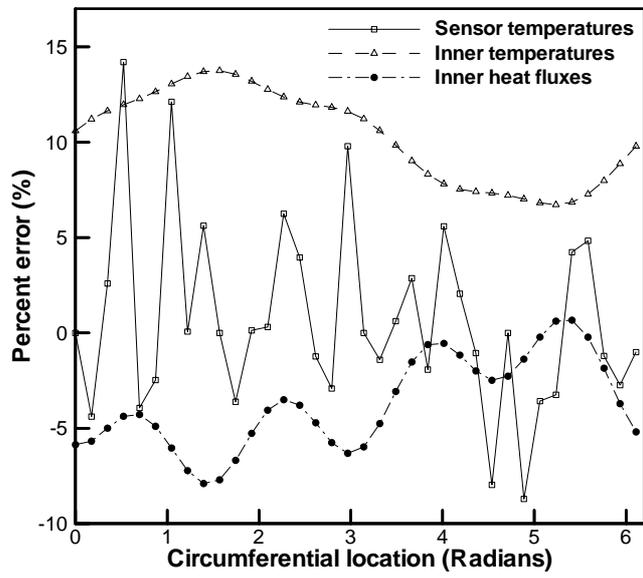


(b)

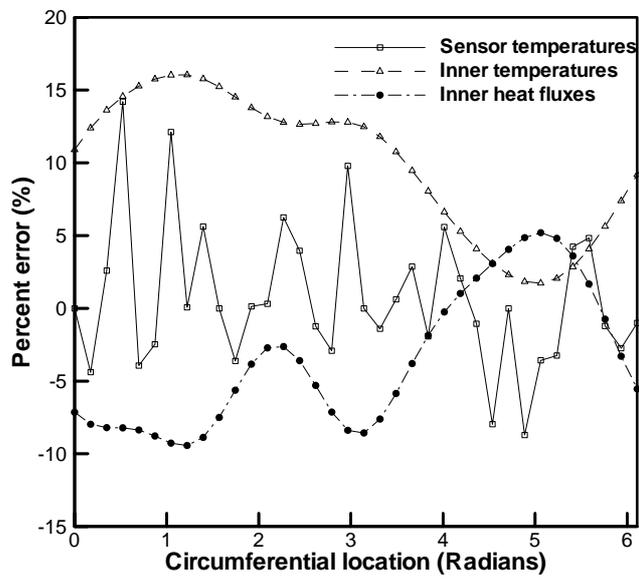


(c)

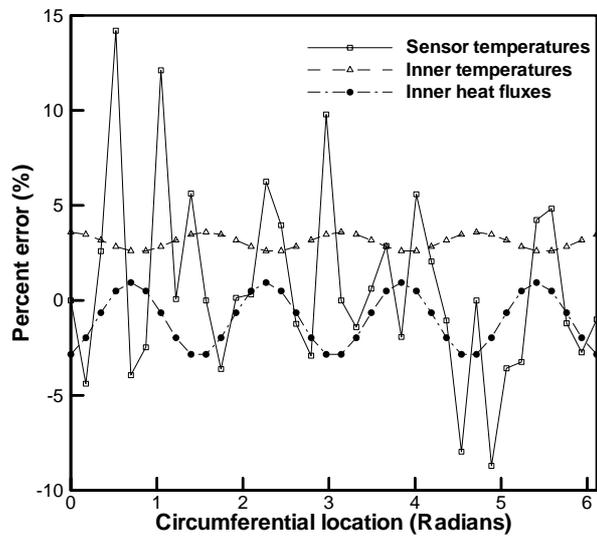
Figure 4.1.5. Inverse calculations for over-specified heat conduction in a concentric geometry: (a) Error indicator vs. optimal regularization parameter (b) L-Curve method and (c) Discrepancy principle determination of the singularity threshold for the truncated SVD method.



(a)

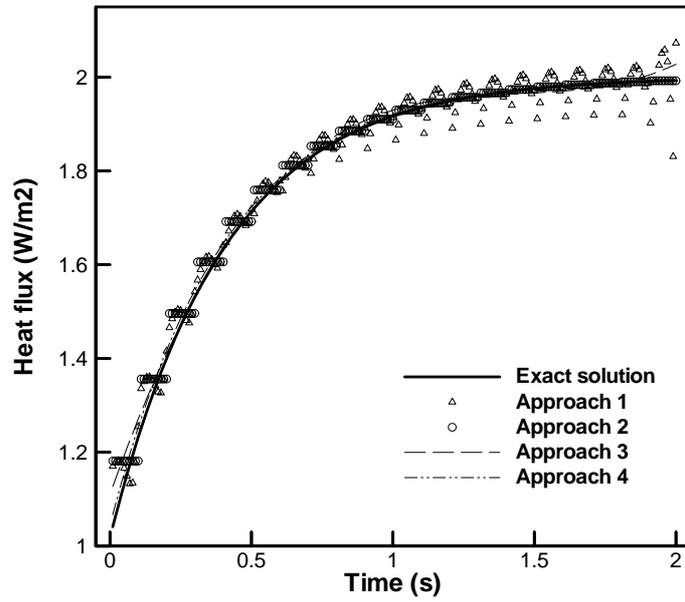


(b)

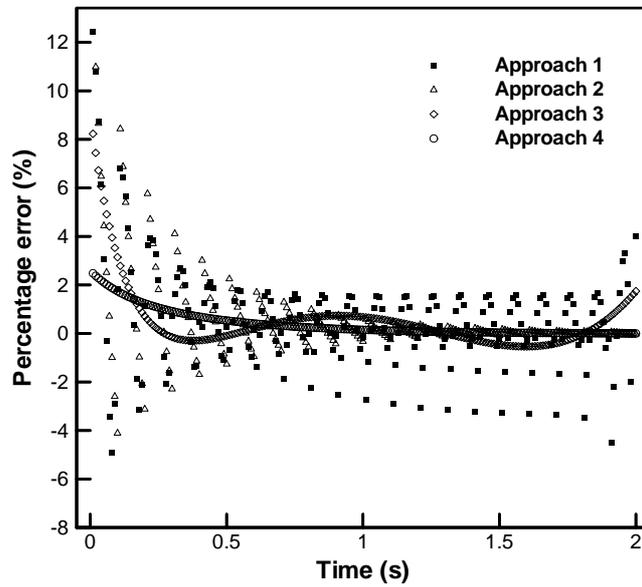


(c)

Figure 4.1.6. Comparison of the inverse calculations by the regularization and truncated SVD methods: (a) Percent errors for the Levenberg-Marquardt method (b) Percent errors for the regularization method and (c) Percent errors for the truncated SVD.

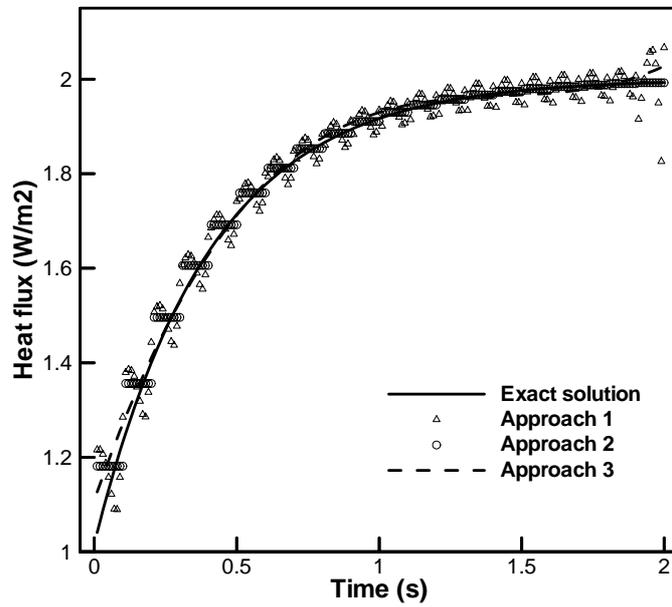


(a)

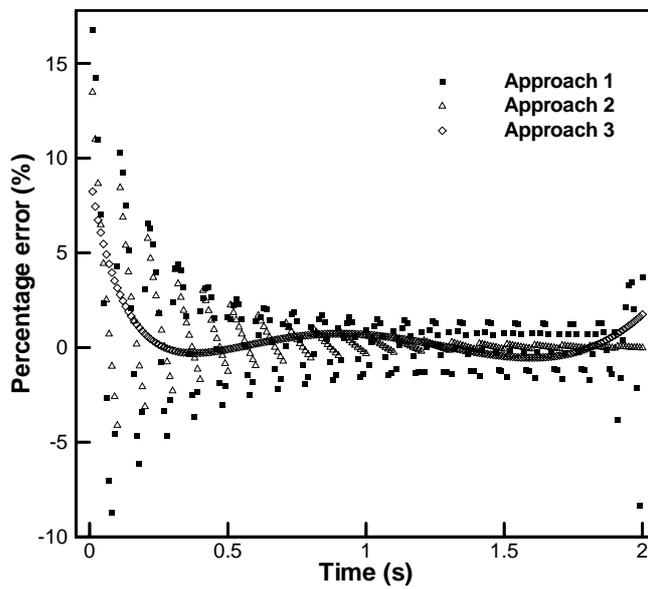


(b)

Figure 4.1.7. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0$) using the regularization method: (a) Heat flux solution. (b) Percentage errors of the heat flux

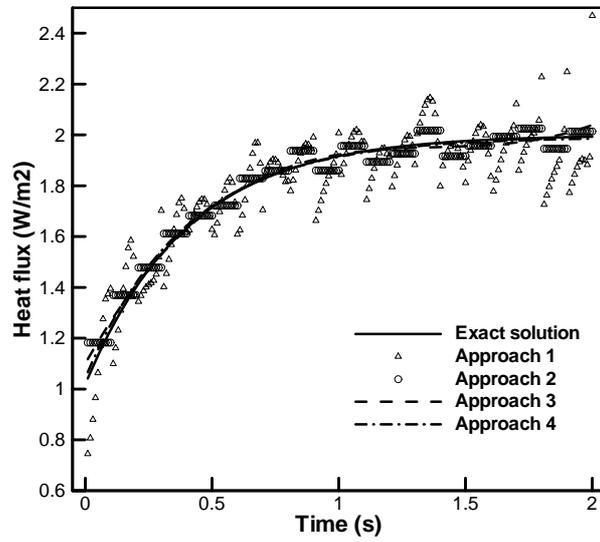


(a)

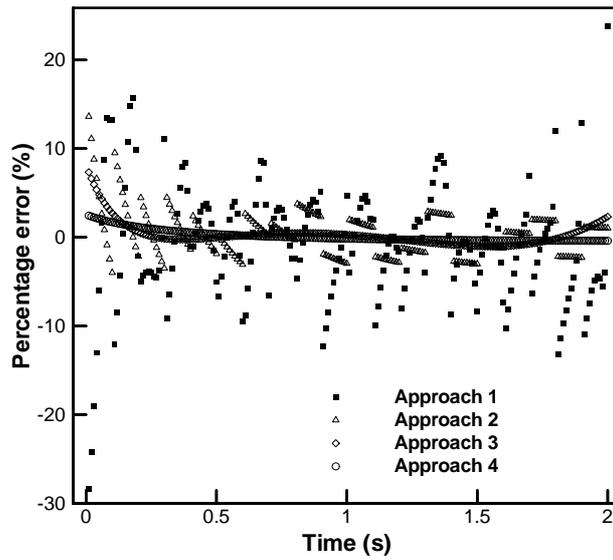


(b)

Figure 4.1.8. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0$) using the SVD: (a) Heat flux solution. (b) Percentage errors of the heat flux

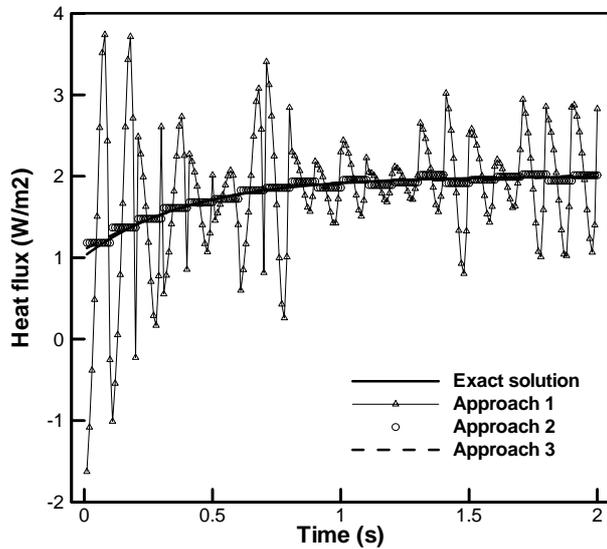


(a)

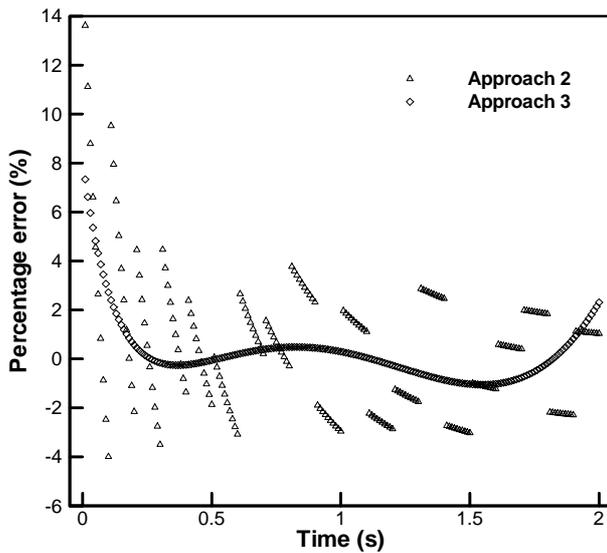


(b)

Figure 4.1.9. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0.01$) using the regularization method: (a) Heat flux solution. (b) Percentage errors of the heat flux



(a)



(b)

Figure 4.1.10. Inverse calculations for transient heat conduction (Example problem1 with $\sigma=0.01$) using the SVD: (a) Heat flux solution. (b) Percentage errors of the heat flux

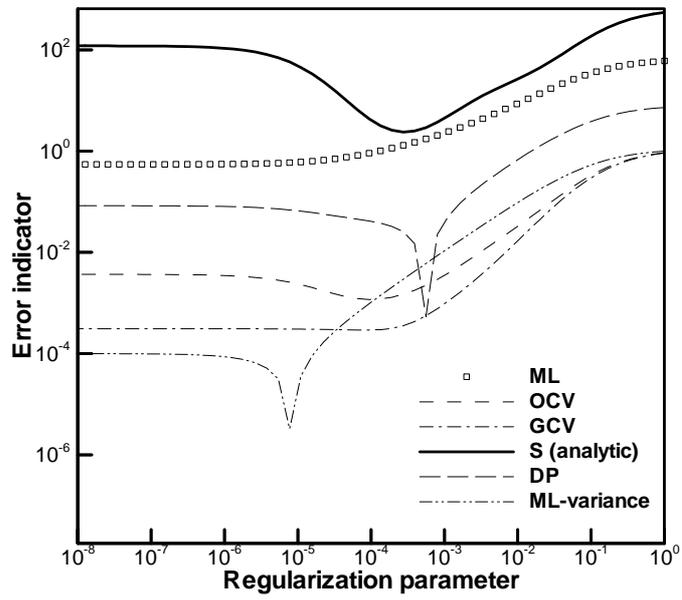


Figure 4.1.11. Error indicators for the 200 unknown

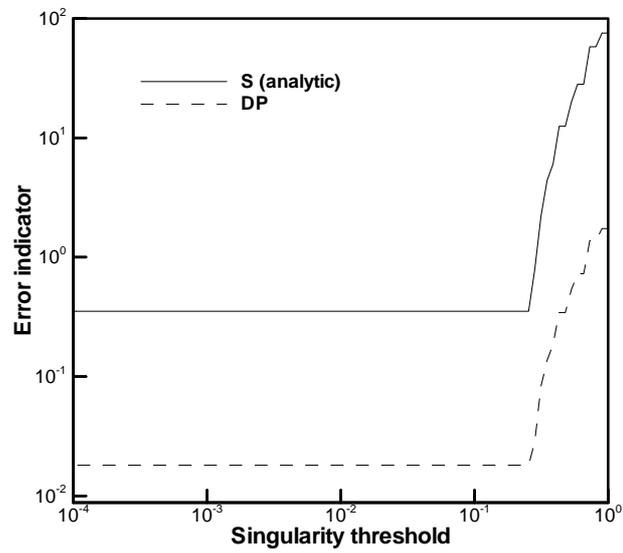


Figure 4.1.12. Discrepancy principle (DP) for the SVD for the 20 unknowns

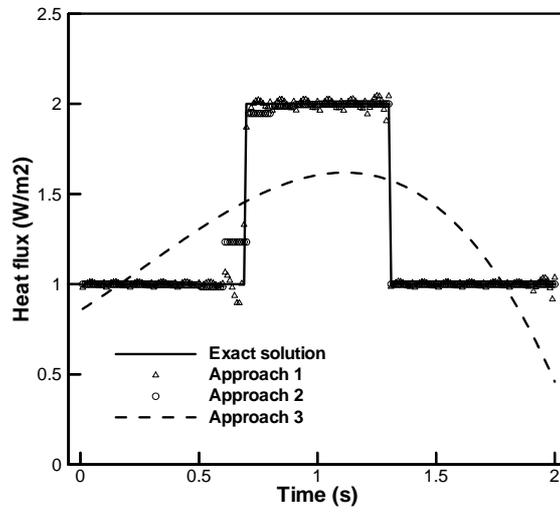


Figure 4.1.13. Heat flux solution for the regularization method (Example problem 2, $\sigma=0$)

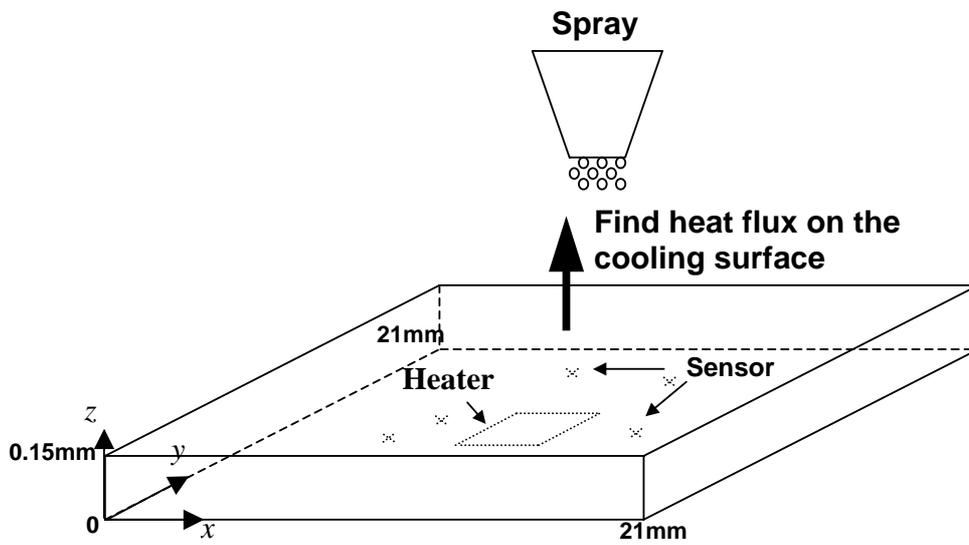


Figure 4.1.14. Schematic of spray cooling of electronics.

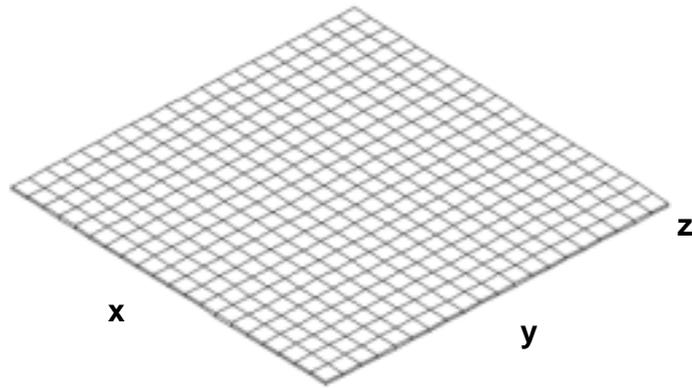


Figure 4.1.15. The FEM mesh used for the analysis of the spray cooling.

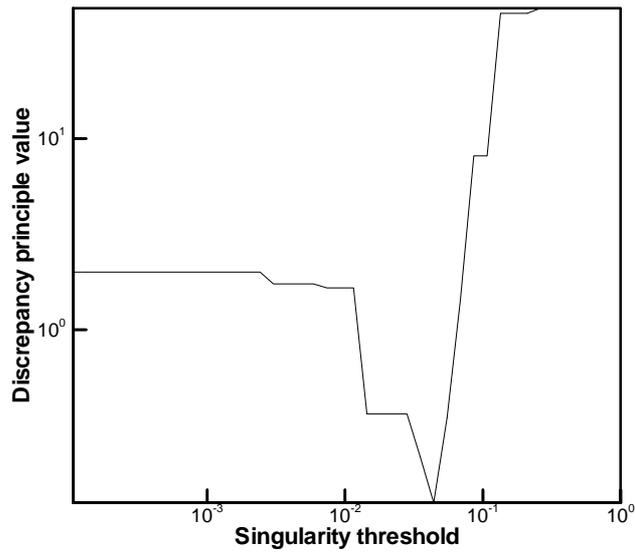


Figure 4.1.16. Discrepancy principle determination of the singularity threshold for the truncated SVD method, which is used for the inverse prediction of heat flux distribution in spray cooling of electronics.

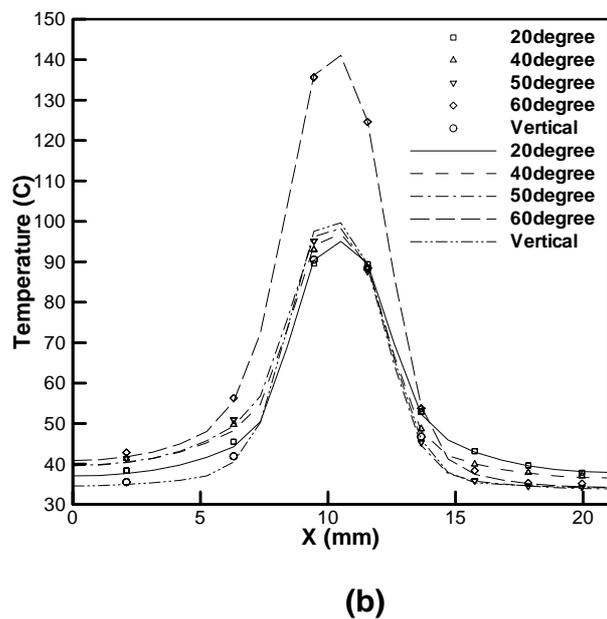
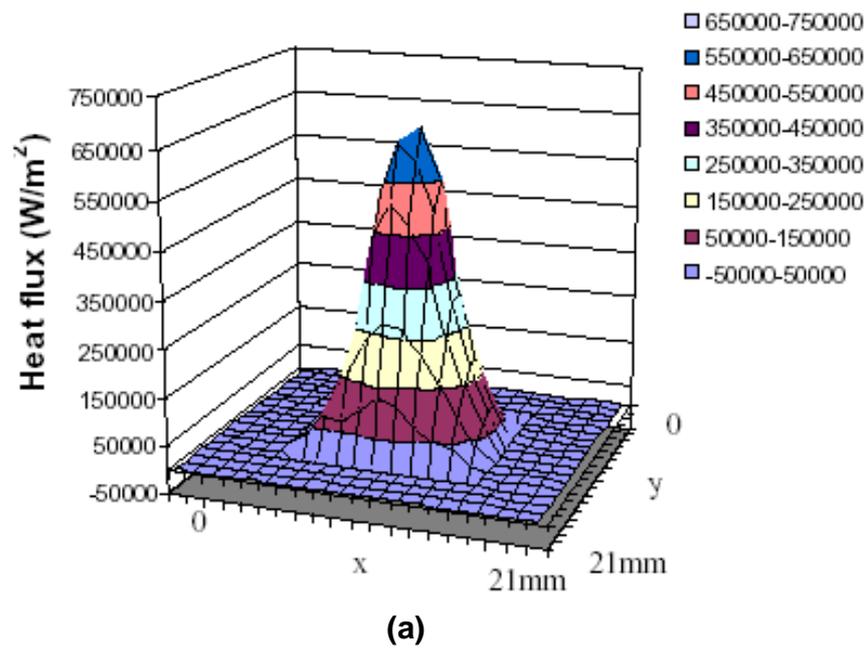


Figure 4.1.17. Inverse calculations of heat flux during spray cooling of electronics for different spray angles: (a) 3-D view of the heat flux on the cooling surface of the die and (b) Comparison of calculated and measured temperatures – dots: measurements and lines: inverse calculations.

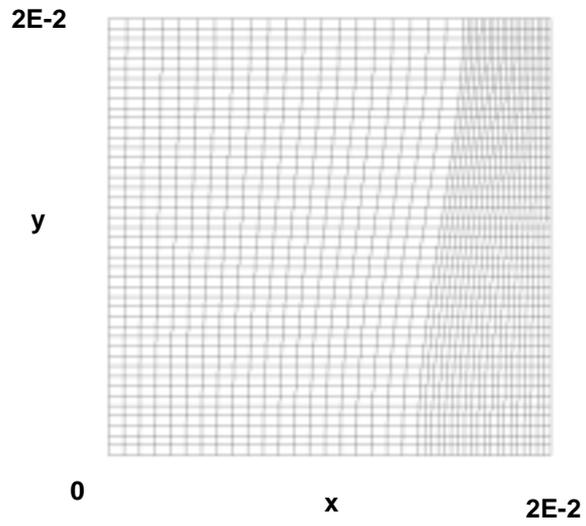
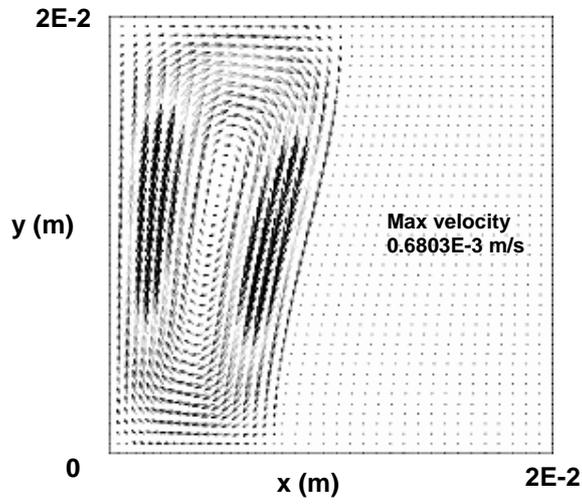
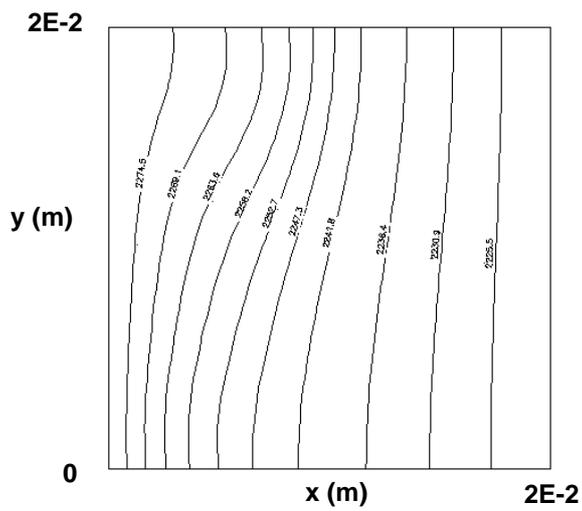


Figure 4.2.1. The deforming FEM quadrilateral linear elements used in the example problem.



(a)



(b)

Figure 4.2.2. Direct solution of the solidification problem: (a) Velocity distribution. (b) Temperature distribution. The natural convection has a strong effect on the solidification interface shape. Boundary condition used: 2280 K at the left wall and 2220K at the right wall. The melting point is 2243.15 K.

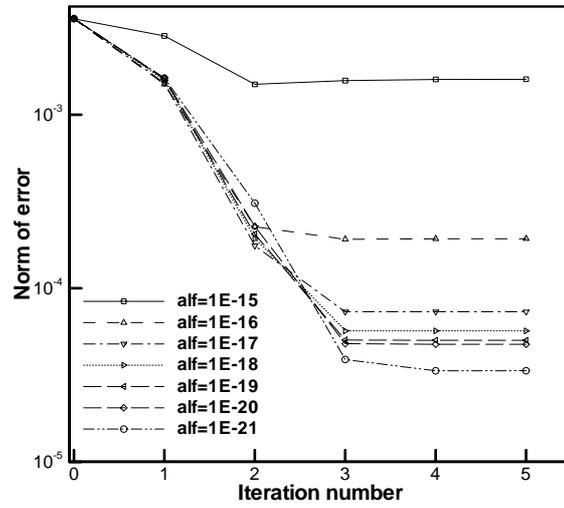


Figure 4.2.3. Norm of error of the solidification distance versus the number of iteration for the regularization method with respect to the regularization parameters

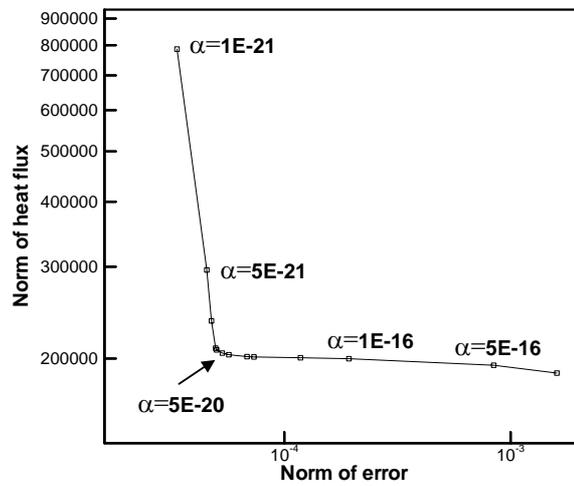
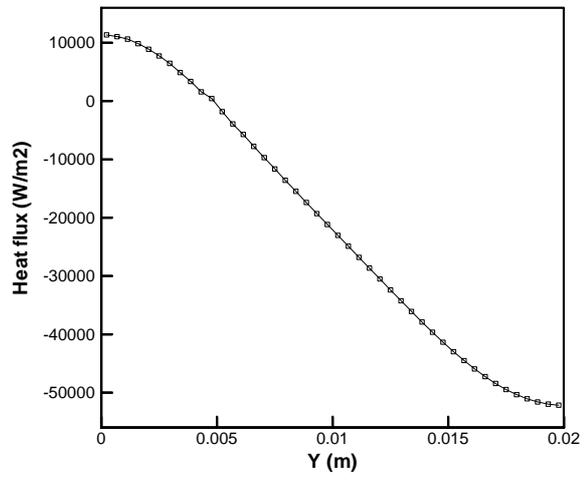
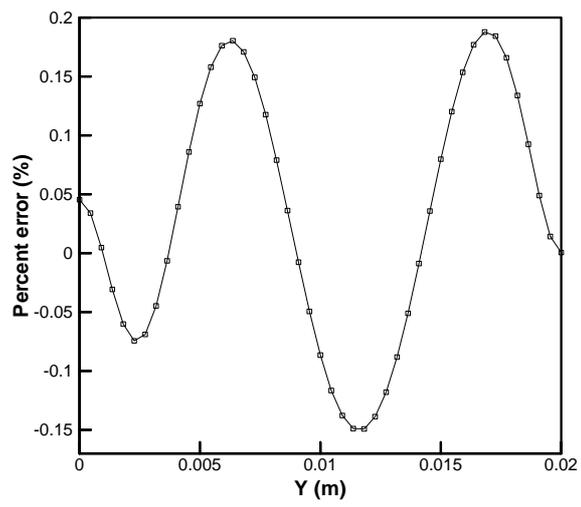


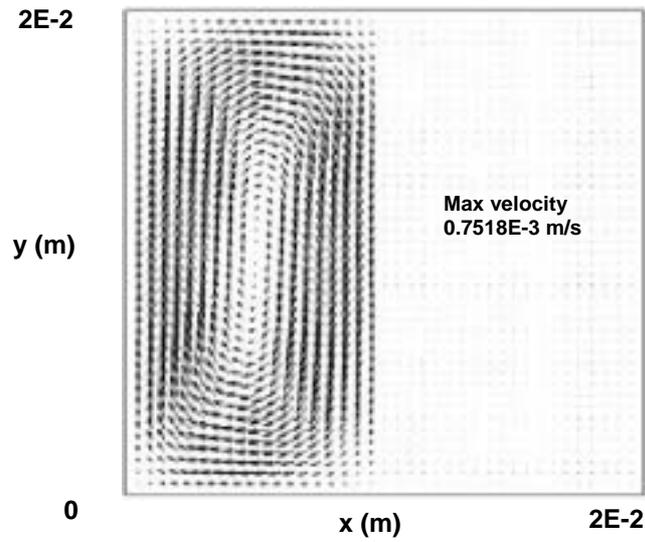
Figure 4.2.4. L-curve plots for case2. The regularization parameter α of 5×10^{-20} is selected.



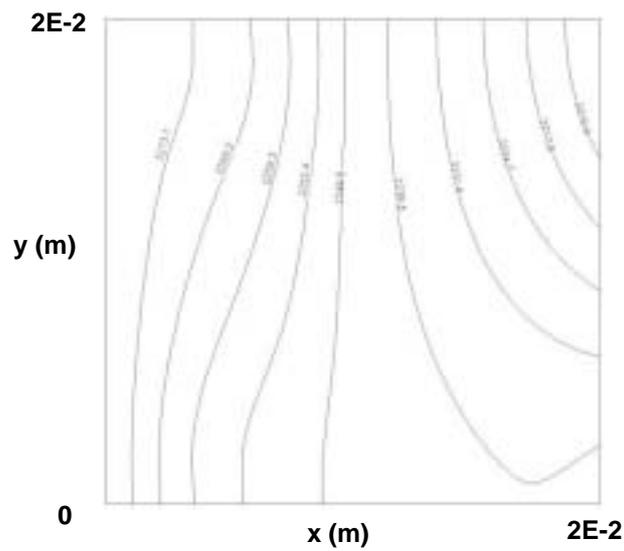
(a)



(b)



(c)



(d)

Figure 4.2.5. Inverse solution of solidification problem for case2. The interface position is prescribed as a vertical line at $x=0.01$. Other conditions are the same as in Figure 4.2.2. (a) Inversely predicted heat flux distribution. (b) Percentage errors in interface positions. (c) Velocity distribution and (d) Temperature distribution.

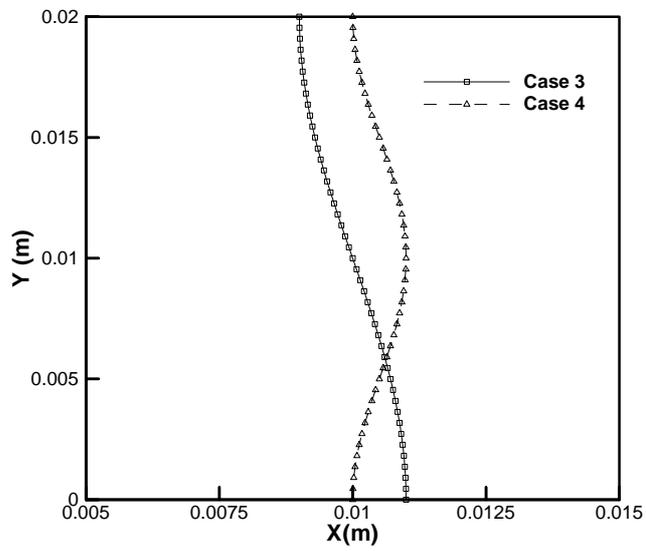
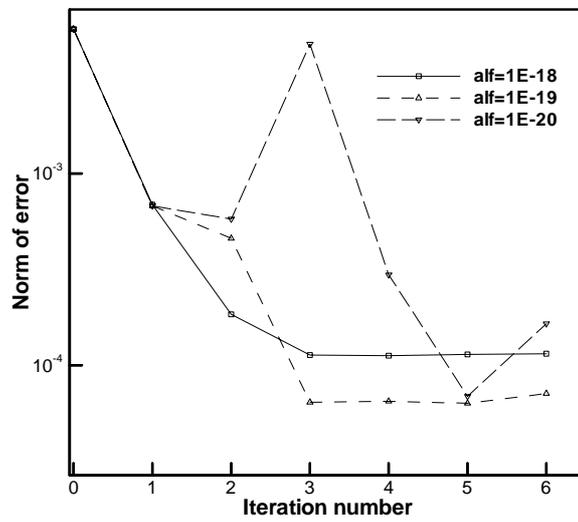
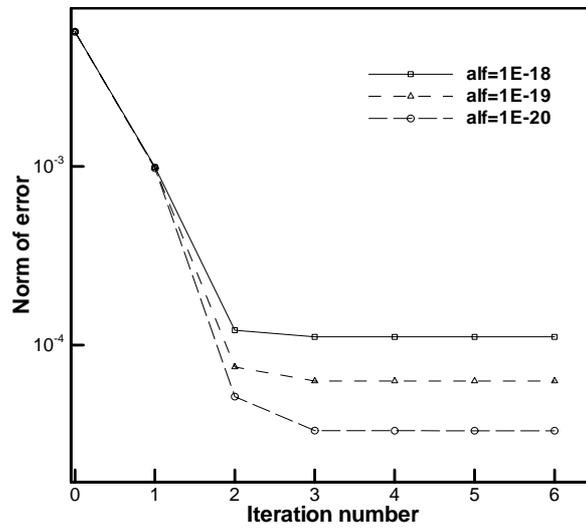


Figure 4.2.6. Prescribed solid-liquid interfaces for case3 ($x=-0.001\cos(50\pi(y-0.02))+0.01$) and case4 ($x=0.0005\cos(100\pi(y-0.01))+0.0105$).



(a)



(b)

Figure 4.2.7. Norm of error of the solidification distance versus the number of iteration with respect to the regularization parameters (a) The conventional finite difference scheme (b) The proposed finite difference scheme

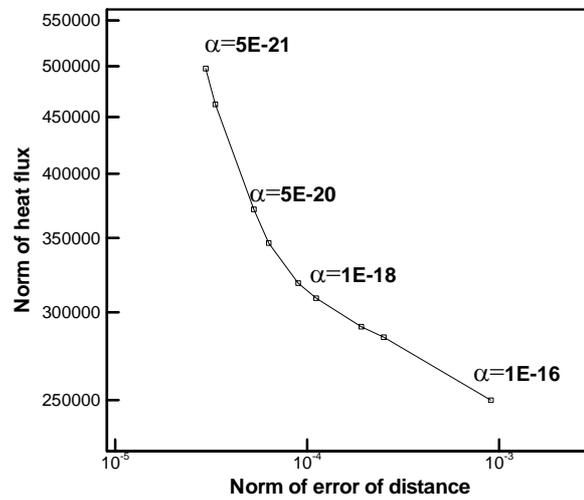
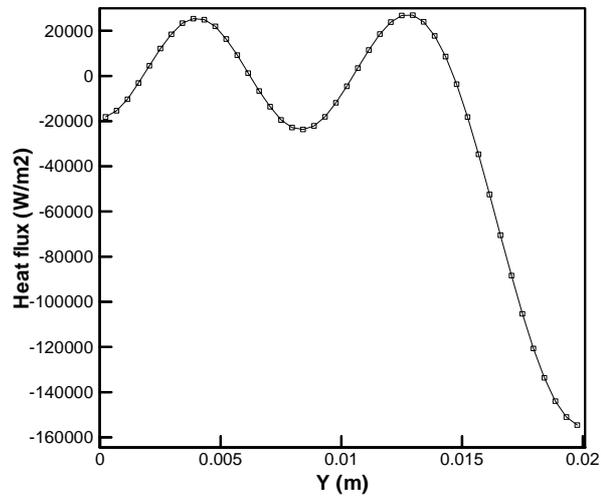
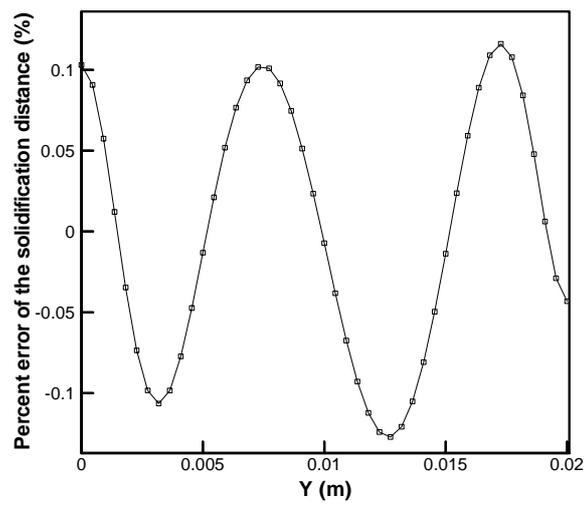


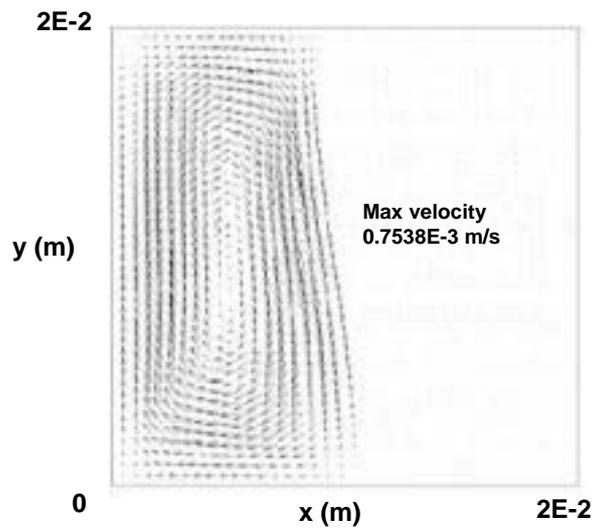
Figure 4.2.8. L-curve for case 3 using the proposed finite difference scheme for the sensitivity coefficient.



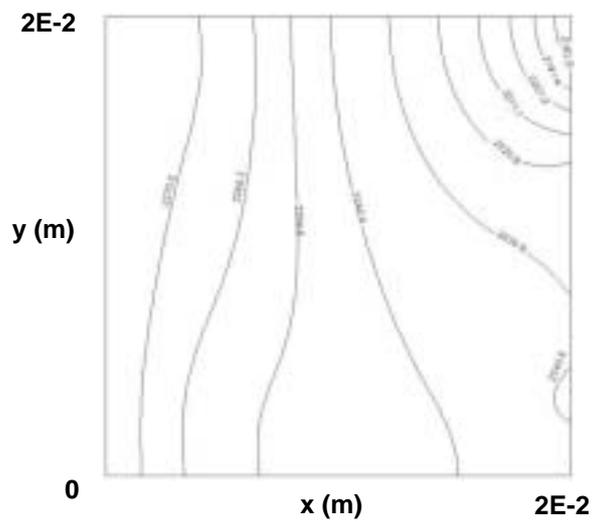
(a)



(b)



(c)



(d)

Figure 4.2.9. Inverse solution of solidification problem for case3. The interface position is prescribed as a sine curve (Figure 4.2.6 case3) Other conditions are the same as in Figure 4.2.2. (a) Inversely predicted heat flux distribution. (b) Percentage errors in interface position. (c) Velocity distribution and (d) Temperature distribution.

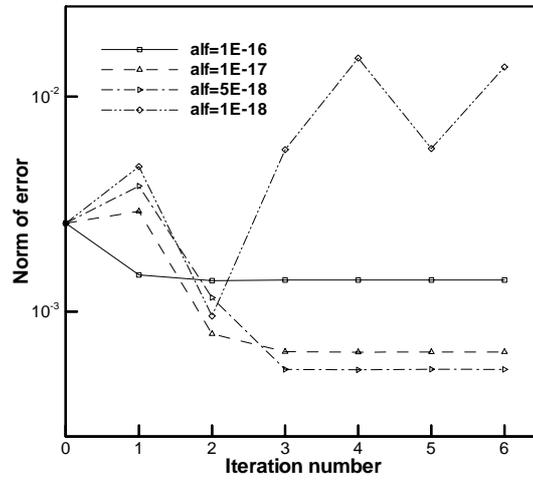


Figure 4.2.10. Norm of error of the solidification distance versus the number of iteration with respect to the regularization parameters.

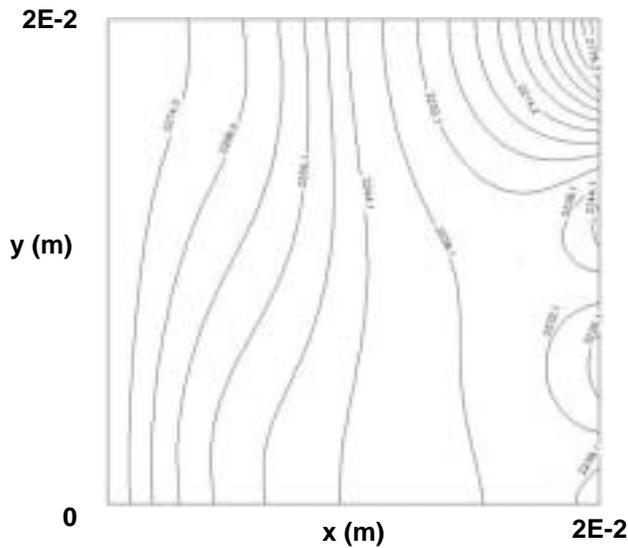


Figure 4.2.11. Temperature distribution by using the heat flux solution at the second iteration of $\alpha=1E-18$. The infeasible temperature distribution is observed in the right corner.

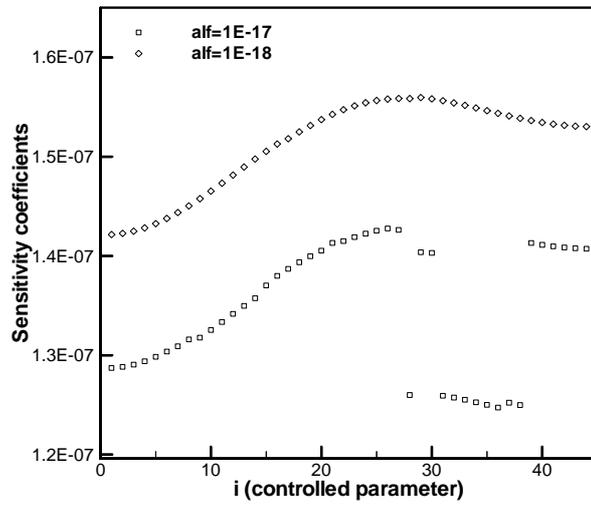


Figure 4.2.12. Sensitivity coefficients $J_{i,20}$ at the third iteration of $\alpha=1E-17, 1E-18$.

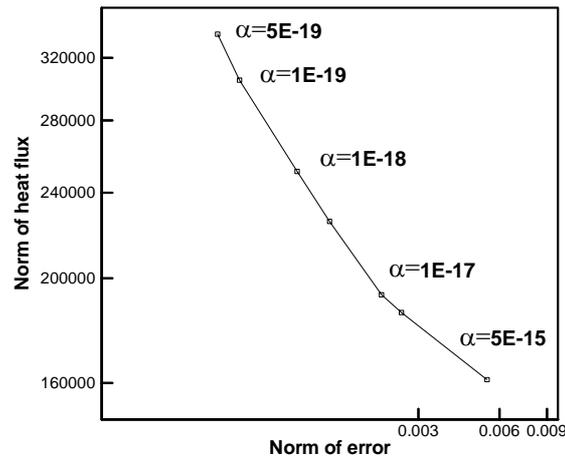
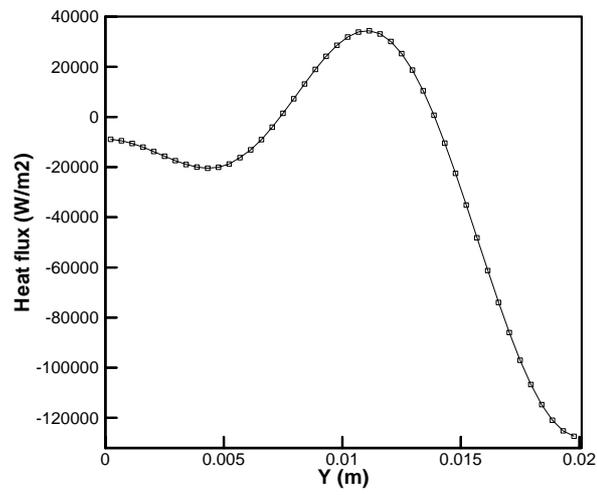
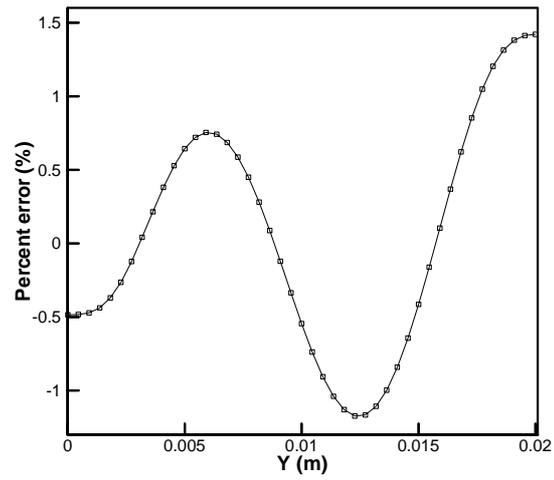


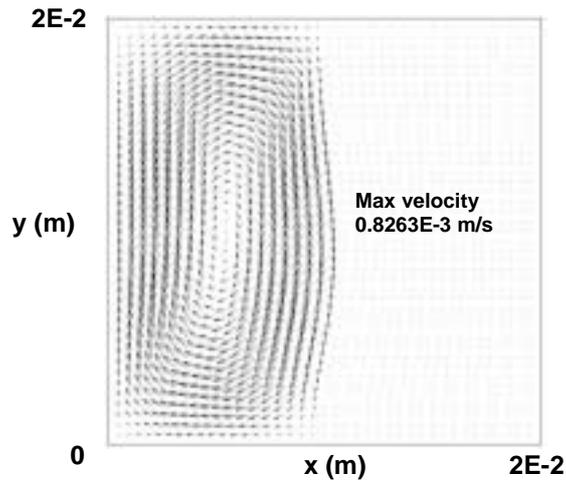
Figure 4.2.13. L-curve plots for the case 3 using the proposed finite difference scheme for the sensitivity coefficient.



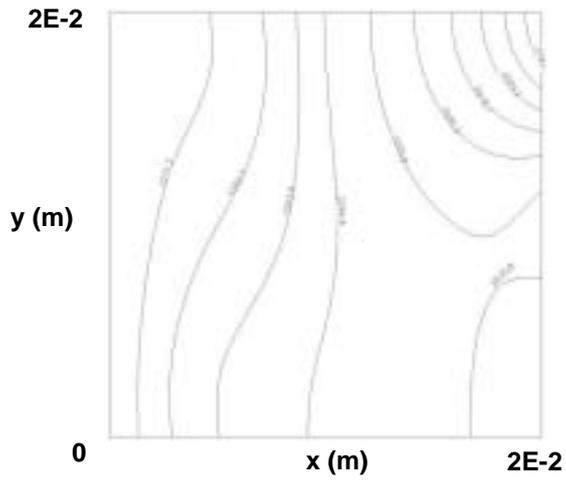
(a)



(b)



(c)



(d)

Figure 4.2.14 Inverse solution of the solidification problem for the case 4. The interface position is prescribed as a sharp sine curve. Other conditions are the same as in Figure 4.2.2. (a) Inversely predicted heat flux distribution. (b) Error in interface position. (c) Velocity distribution and (d) Temperature distribution.

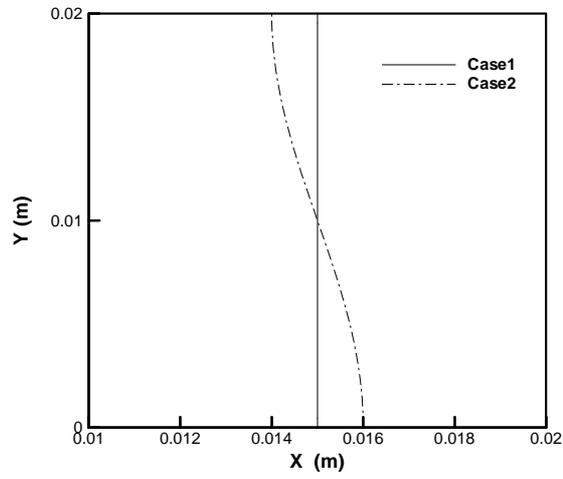
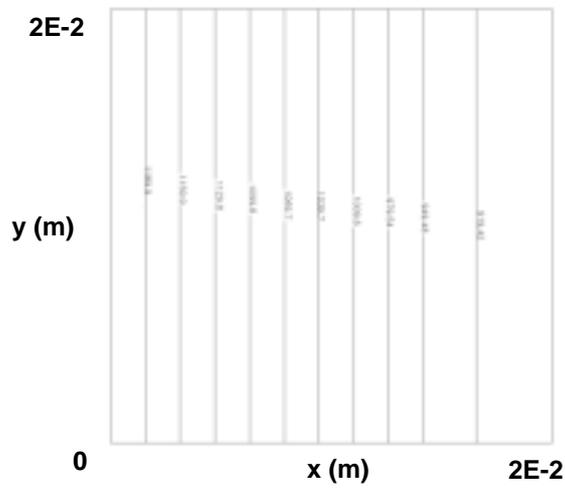


Figure 4.2.15. Prescribed locations for the solid-liquid interface in the case 1 and case2. Case1 is a vertical straight line ($x=0.015$). Case 2 is a sine curve ($x=-0.001\cos(50\pi(y-0.02))+0.015$).



(a)

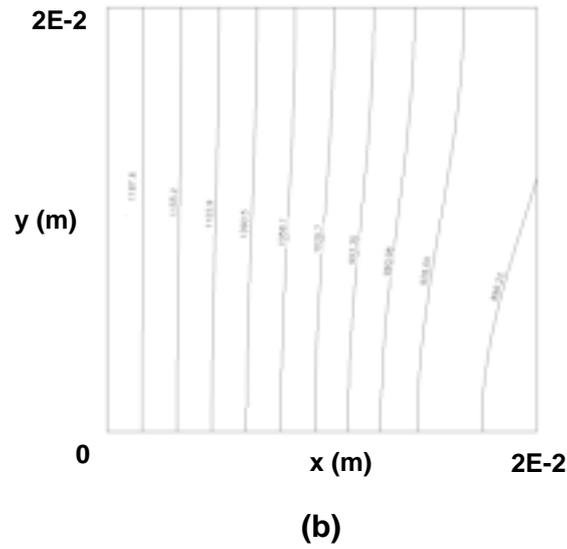


Figure 4.2.16. Temperature distributions for the inverse steady solidification problem for (a) the case1 and (b) the case2. These temperature distributions will be initial conditions for the transient calculations.

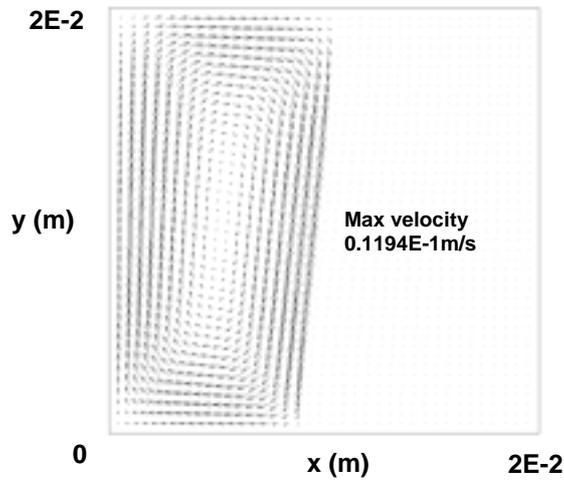


Figure 4.2.17. Velocity distribution at $t=0.6s$ for the direct solidification problem. The initial solid-liquid interface was located at $x=1.5E-2m$ vertically.

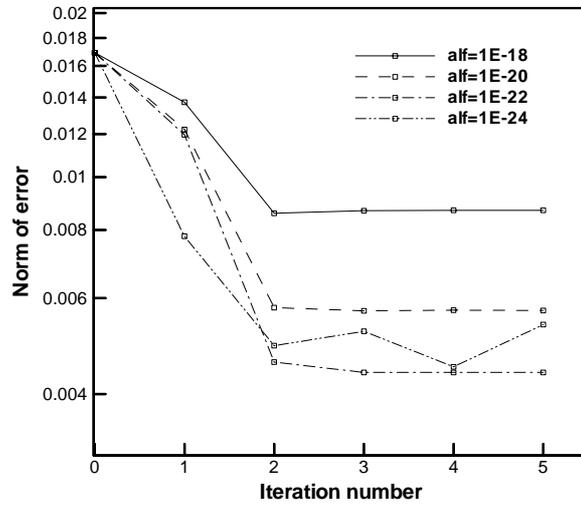


Figure 4.2.18. Convergence rates for the whole time-domain method with 9 unknowns in the case 1

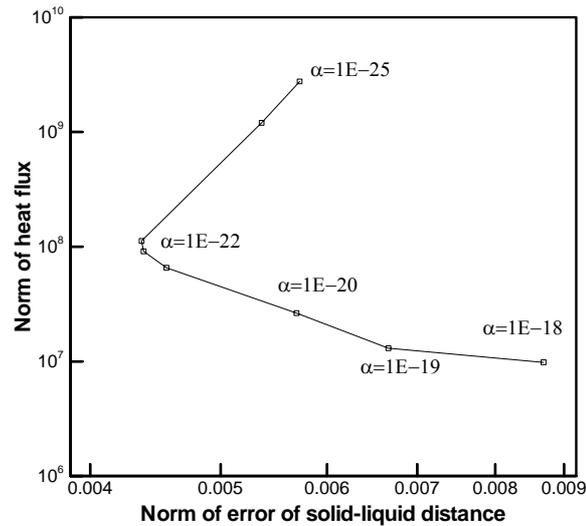


Figure 4.2.19. L-curve plots for the whole time-domain method with 9 unknowns. $\alpha=1E-22$ is selected for the optimal regularization parameter.

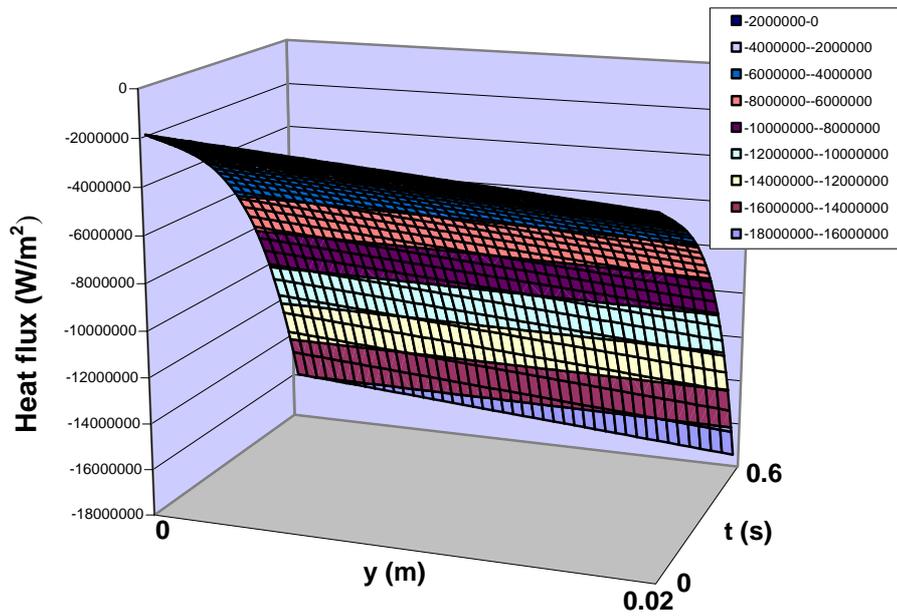


Figure 4.2.20. Heat flux solution using the whole time-domain method with 9 unknowns

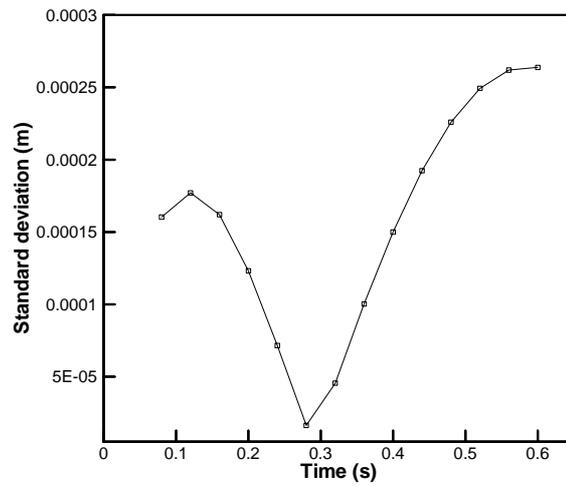


Figure 4.2.21. Standard deviation for the whole time-domain method with 9 unknowns

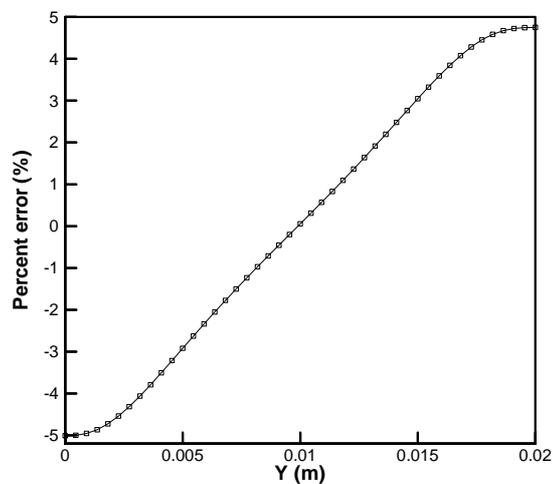


Figure 4.2.22. Percent errors of the solidification distance at $t=0.6$ for the whole time-domain with 9 unknowns

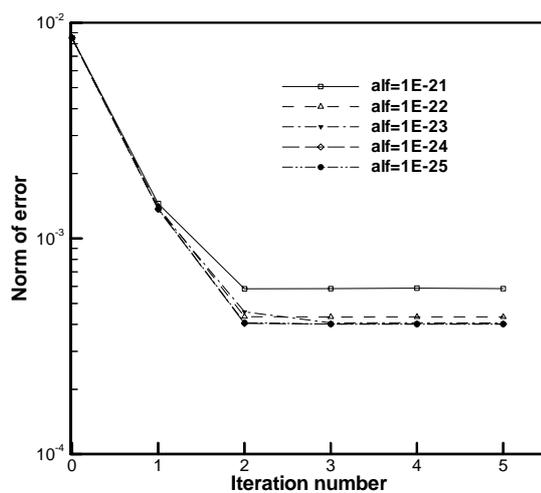


Figure 4.2.23. Convergence rate for the first time domain in the case 1 for the sequential method

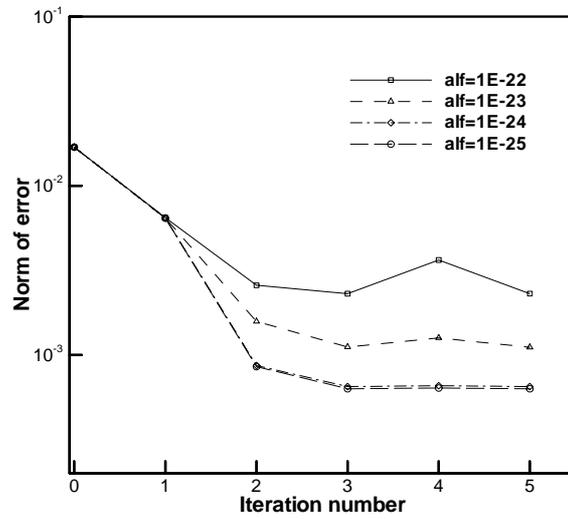


Figure 4.2.24. Convergence rates for the case 1 for the whole time-domain method

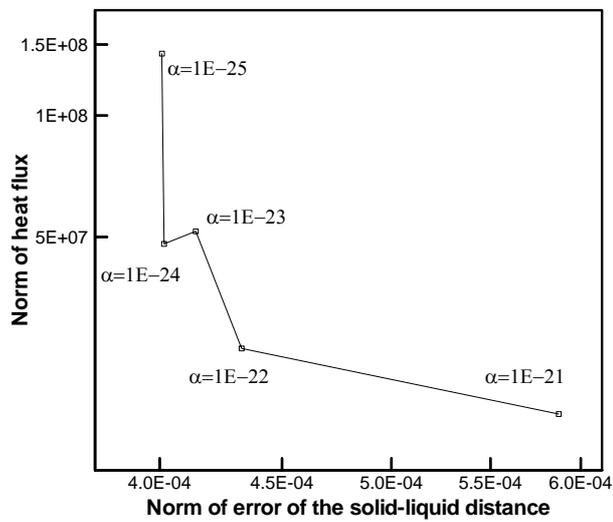


Figure 4.2.25. L-curve plots for the first time domain in the case 1 for the sequential method. The regularization parameter $\alpha=1E-24$ is selected for the optimal solution.

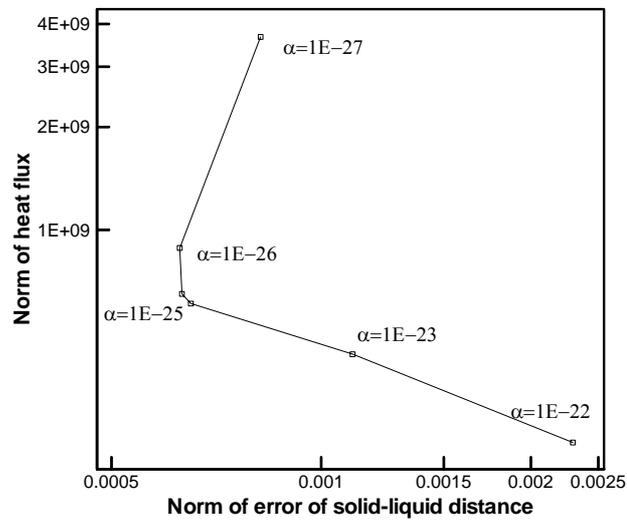
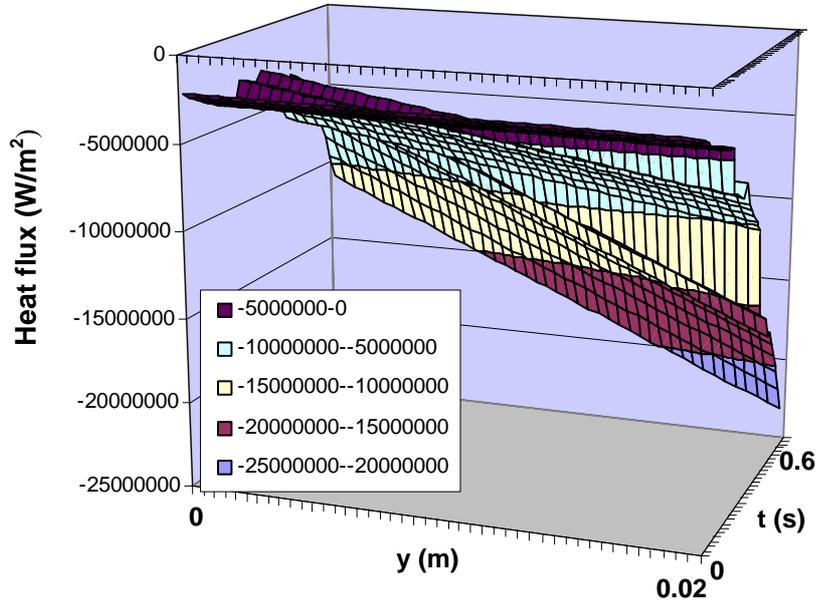
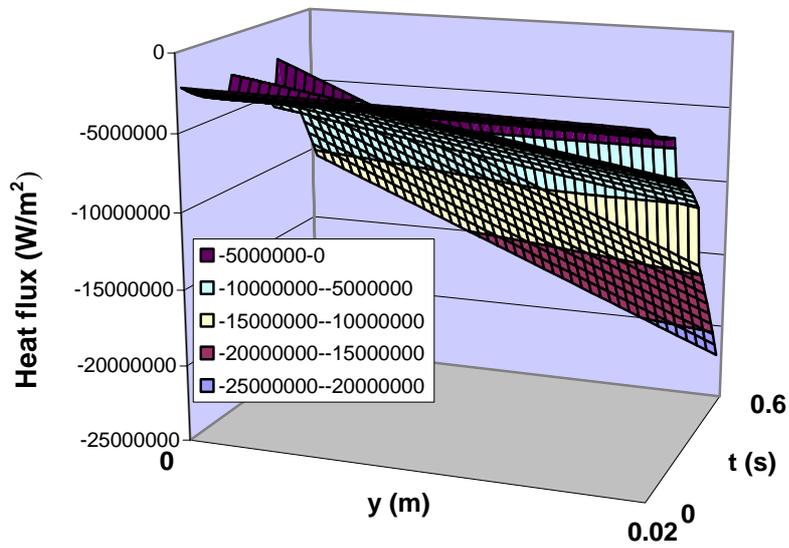


Figure 4.2.26. L-curve plots for the case 1 for the whole time-domain method. The regularization parameter $\alpha=1E-25$ is selected for the optimal solution.

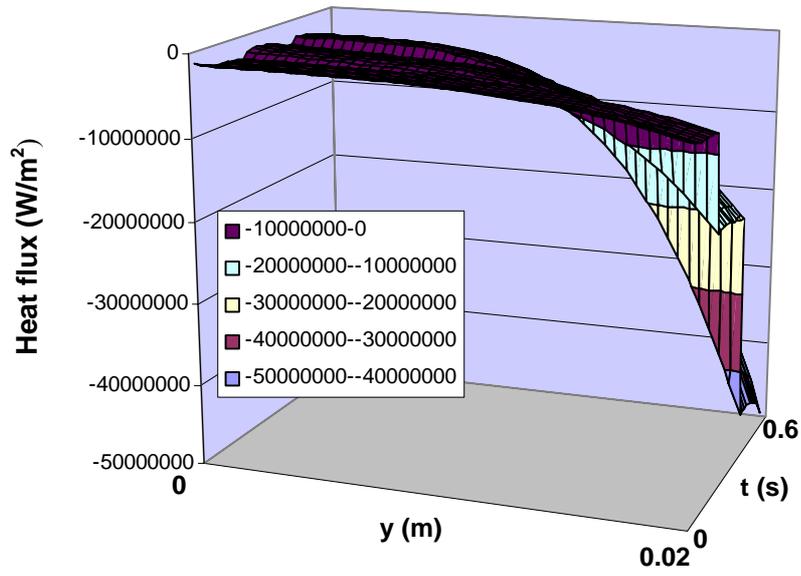


(a)

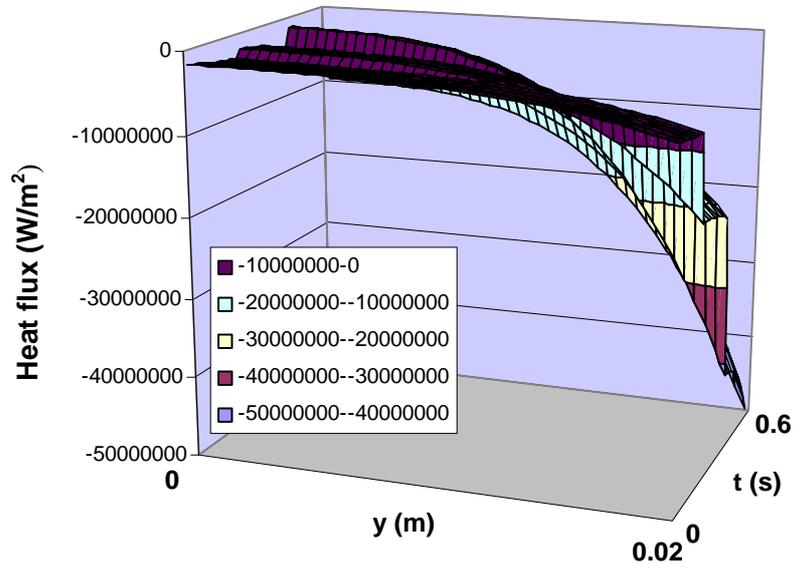


(b)

Figure 4.2.27. Optimal heat flux solutions in the case 1 for (a) the sequential method and (b) the whole time-domain method with the piecewise polynomial function.

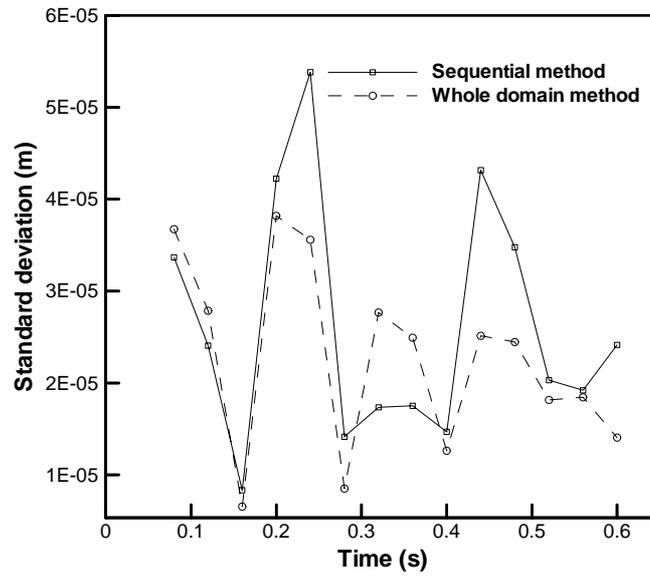


(a)

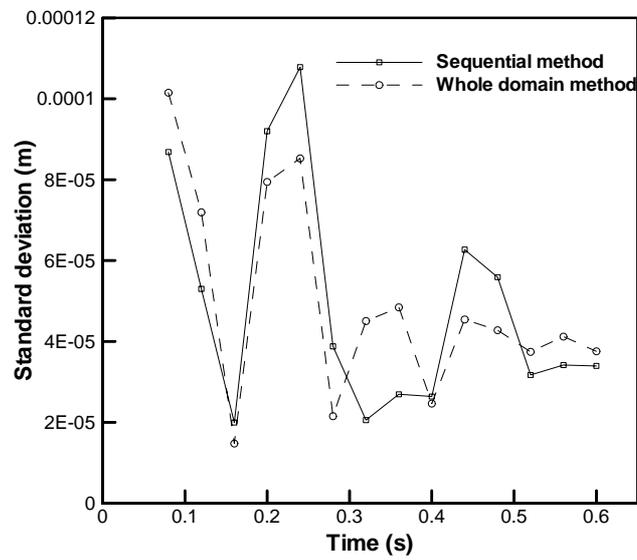


(b)

Figure 4.2.28. Optimal heat flux solution in the case 2 for (a) the sequential method and (b) the whole time-domain method with the piecewise polynomial function



(a)



(b)

Figure 4.2.29. Standard deviations σ (Eq.4.2.2) for the error of the solidification distance for (a) the case 1 and (b) the case 2

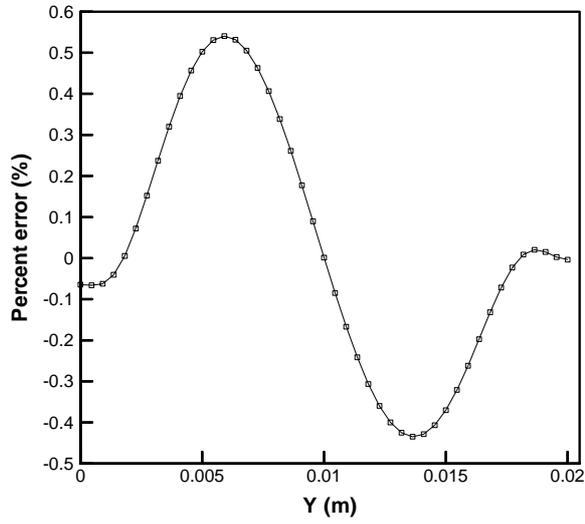


Figure 4.2.30. Percent errors of the solid-liquid distance at $t=0.6$ for the sequential method for the case1

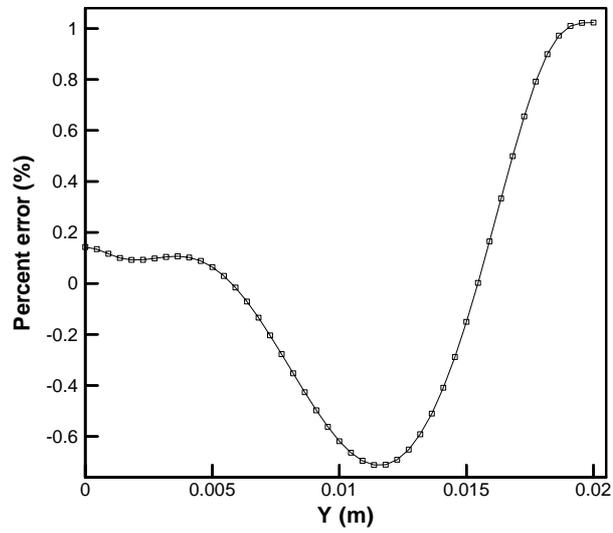


Figure 4.2.31. Percent errors of the solid-liquid distance at $t=0.6$ for the whole time-domain method for the case2

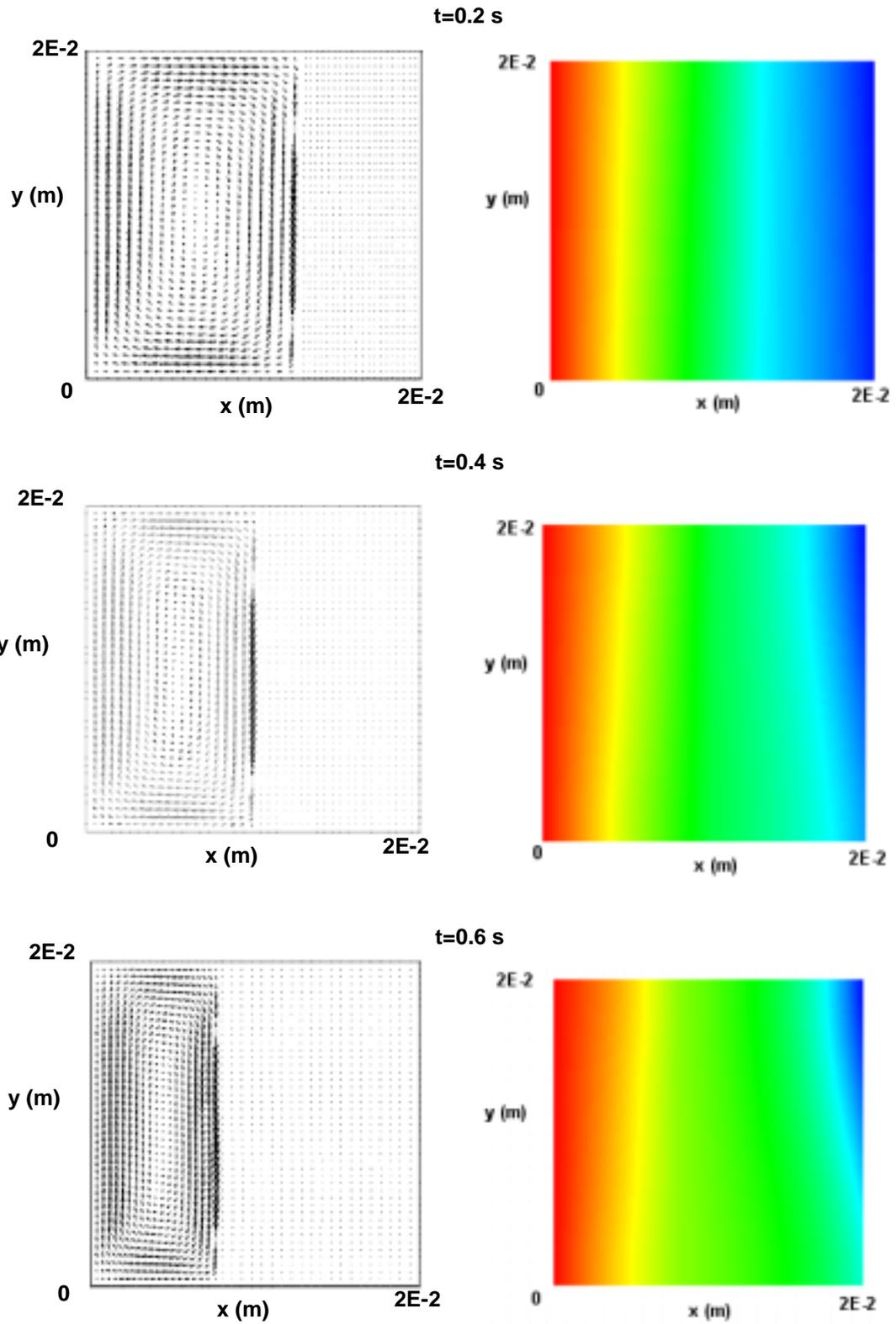


Figure 4.2.32. Velocity (left) and temperature (right) distributions at times $t=0.2, 0.4, 0.6$ for the case1

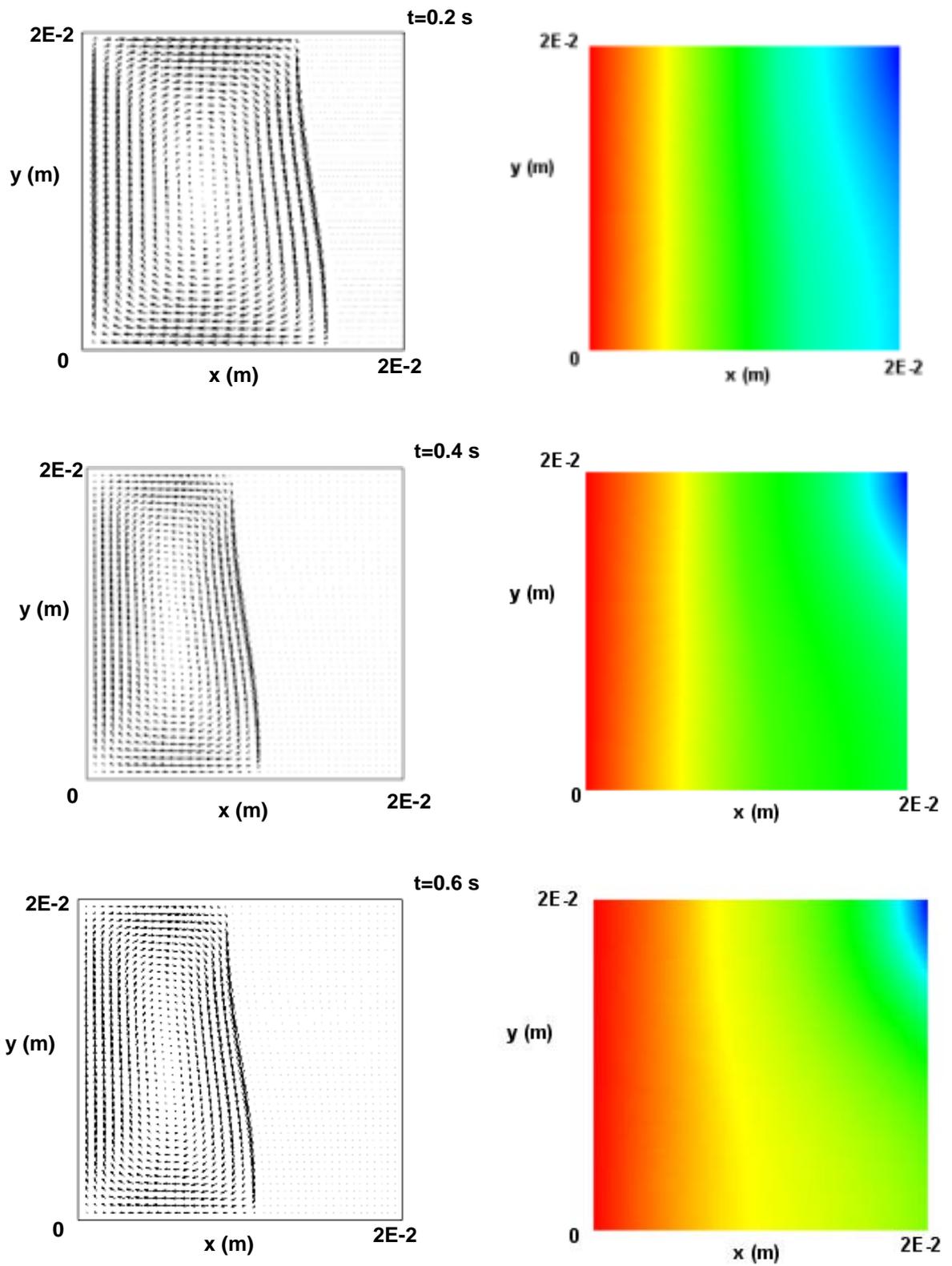


Figure 4.2.33. Velocity (left) and temperature (right) distributions at times $t=0.2, 0.4, 0.6$ for the case2

Table 4.1.1 Input heat generation and heat flux at x=1m

	Heat generation (W/m ³)	Heat flux (W/m ²) at x=1m
Case1	0	100
Case2	1	95
Case3	100	50

Table 4.1.2 The regularization parameters chosen in each method and the corresponding least square error for the 2-D problem

	α	$\ q_{analytic} - q_{computed}\ ^2$
Analytic	3.52×10^{-3}	7.17×10^1
DP	1.44×10^{-3}	8.59×10^1
OCV	2.42×10^{-4}	1.25×10^2
GCV	1.55×10^{-4}	1.43×10^2
ML	8.51×10^{-6}	6.33×10^3
ML-variance	1.06×10^{-5}	4.34×10^3
L-curve	5.07×10^{-5}	3.48×10^2

Table 4.1.3 Behavior of the Levenberg-Marquardt method using the More's implementation for axisymmetric over-specified problems: $m=36$, $\sigma=0.05$

Iteration	μ	Δ	$\ \{P\}\ $	$\ \{\bar{T}\} - \{T\}\ $	$\ \{q_{exact}\} - \{q_{computed}\}\ $
1	6.187E-01	3.300E+00	3.300E+00	2.683E+00	3.554E+00
2	6.518E-01	1.650E+00	1.650E+00	1.434E+00	1.907E+00
3	7.183E-01	8.250E-01	8.250E-01	8.244E-01	1.086E+00
4	8.517E-01	4.125E-01	4.125E-01	5.388E-01	6.810E-01
5	1.120E+00	2.063E-01	2.063E-01	4.126E-01	4.844E-01
6	1.656E+00	1.031E-01	1.031E-01	3.589E-01	3.906E-01
7	2.732E+00	5.156E-02	5.156E-02	3.357E-01	3.461E-01
8	4.884E+00	2.578E-02	2.578E-02	3.253E-01	3.247E-01
9	9.188E+00	1.289E-02	1.289E-02	3.204E-01	3.143E-01
10	1.780E+01	6.445E-03	6.445E-03	3.181E-01	3.092E-01
11	3.502E+01	3.223E-03	3.223E-03	3.169E-01	3.067E-01
12	6.946E+01	1.611E-03	1.611E-03	3.164E-01	3.054E-01
13	1.383E+02	8.057E-04	8.057E-04	3.161E-01	3.048E-01
14	2.761E+02	4.028E-04	4.028E-04	3.159E-01	3.044E-01
15	5.516E+02	2.014E-04	2.014E-04	3.159E-01	3.043E-01

Table 4.1.4 The regularization parameters chosen in each method and the corresponding least square error for the axisymmetric problem

	α	$\ q_{analytic} - q_{computed}\ ^2$
Analytic	1.59×10^{-2}	1.66×10^{-1}
DP	1.16×10^{-2}	1.73×10^{-1}
OCV	4.50×10^{-3}	3.06×10^{-1}
GCV	4.05×10^{-3}	3.33×10^{-1}
ML	3.23×10^{-4}	4.19
ML-variance	1.91×10^{-4}	8.23
L-curve	3.64×10^{-3}	3.63×10^{-1}

Table 4.1.5 Sum of square errors in calculated heat flux values in $\sigma=0$ for example problem 1

	Regularization	SVD
Approach 1	0.24985	0.36176
Approach 2	0.13214	0.13214
Approach 3	0.06138	0.06138
Approach 4	0.01602	N/A

Table 4.1.6 Sum of square errors in calculated heat flux values in $\sigma=0.01$ for example problem 1

	Regularization	SVD
Approach 1	2.38464	119.978
Approach 2	0.35157	0.35157
Approach 3	0.06349	0.06349
Approach 4	0.01695	N/A

Table 4.1.7 Sum of square errors in calculated heat flux values in $\sigma=0$ for example problem2

	Regularization	SVD
Approach 1	0.22678	0.22717
Approach 2	1.11153	1.11153
Approach 3	24.4941	24.5335

Table 4.1.8 Sum of square errors in calculated heat flux values in $\sigma=0.01$ for example problem2

	Regularization	SVD
Approach 1	3.08138	125.073
Approach 2	1.37455	1.37812
Approach 3	24.4729	24.4953

Table 4.2.1 Thermal properties used in the example problem

	Symbol	Value	Units
Heat conductivity in solid	K_s	10	W/mK
Heat conductivity in liquid	K_L	5	W/mK
Specific heat in solid	c_s	390	J/kgK
Specific heat in liquid	c_L	1000	J/kgK
Latent heat	H	1.97×10^9	J/kg
Density	ρ	4300	kg/m ³
Viscosity	μ	0.04	Ns/m ²
Thermal expansion coefficient	β	2.70×10^{-5}	1/K
Melting temperature	T_m	2243.15	K

Table 4.2.2 Thermal properties of aluminum

	Symbol	Value	Units
Heat conductivity in solid	K_s	218	W/mK
Heat conductivity in liquid	K_L	100	W/mK
Specific heat in solid	c_s	897	J/kgK
Specific heat in liquid	c_L	897	J/kgK
Latent heat	H	3.97×10^5	J/kg
Density	ρ	2700	kg/m ³
Viscosity	μ	3.39×10^{-3}	Ns/m ²
Thermal expansion coefficient	β	2.25×10^{-5}	1/K
Melting temperature	T_m	933	K

CHAPTER FIVE

CONCLUSIONS

In this section, the conclusions of the inverse heat conduction problems and the inverse design solidification problems are addressed.

5.1 INVERSE HEAT CONDUCTION PROBLEMS

This paper presented a numerical study on inverse heat conduction problems. Three different methods, the Tikhonov regularization method, the SVD method, and the Levenberg-Marquardt method, were discussed and their performance was assessed comparatively. Five different schemes for choosing an optimal regularization parameter α for inverse heat transfer calculations were also discussed and evaluated using 2-D steady-state heat conduction cases. In addition, parameter estimation and function estimation were discussed using 1-D transient heat conduction problems. It is found that the discrepancy principle (DP) gives the best estimate of α based on the testing of the 2-D problems with various conditions. The ML method is very stable but always estimates smaller regularization parameters than in the analytic solution. The L-curve method is similar to the ML method. In many cases, the OCV and GCV perform well, but not as good as the DP; in a few cases, however, they specify more than one parameter. Based on the regularization case studies, the DP was used to estimate the singularity threshold value for the SVD method, and good results were obtained for all the cases studied. For all the cases tested except the simple cases, the Levenberg-Marquardt method with the Ozisik's implementation did not perform well; however, the method with the More's implementation has the same accuracy of the regularization method if the trust-region

radius is carefully chosen. For the cases studied, the truncated SVD method with the threshold value determined by the DP appears to be the best, although in many cases it gave results very comparative with the regularization method. As for the cases studied in the 1-D transient problems, the parameter estimation is superior to the function estimation when the optimal heat flux solution has a smooth curve. On the other hand, when the heat flux solution contains discontinuities and sharp corners, the function estimation is superior to the parameter estimation. The inverse algorithms were also applied to estimate the heat flux experienced due to droplet spray cooling of a microelectronic processor. The truncated SVD and regularization methods give almost identical results.

5.2 INVERSE DESIGN SOLIDIFICATION PROBLEMS

This paper also presented a computational algorithm for the inverse steady-state solidification problem and the inverse design of solidification processing systems. The algorithm entails the use of the Tikhonov regularization method, along with an appropriately selected regularization parameter based on the L-curve method. The direct solution of the moving boundary problem was solved using the deforming finite element method. The direct and inverse formulations were presented. The new finite difference scheme for determining the sensitivity coefficients was also proposed in the inverse steady-state solidification problems. The determination of the optimal regularization parameter α using the L-curve method was also given. The design algorithm was applied to determine the appropriate boundary heat flux distribution to obtain prescribed solid-liquid interfaces in a 2-D cavity. The whole time-domain method and the sequential

method were used to approximate the optimal heat flux solutions in the inverse design of solidification processes. Based on the cases studied, the proposed finite difference scheme is superior to the conventional finite difference scheme in determination of the sensitivity coefficients for the inverse steady-state solidification problems. The results of the inverse design of solidification processes show that the sequential method is comparative to the whole time-domain method if the diffusion time between the boundary and the interface is carefully considered. We also find that the L-curve based regularization method is reasonably accurate for both the inverse steady-state solidification problem and the inverse design of solidification processing systems.

BIBLIOGRAPHY

- Allen, D. M., "The relationship between variable selection and data augmentation and a method for prediction", *Technometrics*, Vol.16 (1974) 125-127.
- Barnett, S., "Matrices: Methods and Applications", Oxford University Press, 1990.
- Bartoszynski, R., "Probability and statistical inference", Wiley, 1996.
- Beck, J. V., Blackwell, B., and St. Clair, C. R., "Inverse heat conduction, Ill-posed problems", Wiley-Interscience, New York, 1985.
- Beck, J. V. and Blackwell, B., "in Handbook of Numerical Heat Transfer", Minkowycz, W. J., Sparrow, E. M., Schneider, G. E., and Pletcher, R. H., Wiley-Interscience, New York, (1988) 787-834
- Cui, X. and Li, B.Q., "A Parallel Galerkin Boundary Element Method For Surface Radiation And Mixed Heat Transfer Calculations In Complex 3-D Geometries", *International Journal for Numerical Methods in Engineering*, 61 (2004) 1-25.
- Demoment, G., "Image reconstruction and restoration: Overview of common estimation structures and problems", *IEE transactions on acoustics, speech, and signal processing*, Vol. 37 (1989) 2024-2036.
- Dulikrabich, G. S., Colaco, M. J., Martin, T. J., and Lee, S., "An inverse method allowing user-specified layouts of magnetized microfibers in solidifying composites", *J. Composite Materials*, Vol.37 (2003) 1351-1365.
- Fletcher, C. A. J., "Computational Techniques for Fluid Dynamics", Springer-Verlag, 1991.
- Fortier, N., Demoment, G., and Goussard, Y., "GCV and ML methods of determining parameters in image restoration by regularization: Fast computation in the spatial domain and experimental comparison", *Journal of visual communication and image representation*, Vol. 4 (1993) 157-170.
- Galatsanos, N. P. and Katsaggelos, A. K., "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation", *IEEE Transactions on image processing*, Vol. 1 (1992) 322-336.
- Golub, G. H., Heath, M., and Wahba, G., "Generalized cross-validation as a method for choosing a good ridge parameter", *Technometrics*, Vol. 21 (1979) 215-223.

- Hale, S. W., Keyhani, M., and Frankel, J. I., "Design and control of interfacial temperature gradients in solidification", *I. J. Heat and Mass Transfer* 43 (2000) 3795-3810.
- Hansen, P. C. and O'Leary, D. P., "The use of the L-curve in the regularization of discrete ill-posed problems", *SIAM J. Sci Comput.* Vol. 14 (1993) 1487-1503.
- Hollingsworth, K. G. and Johns, M. L., "Measurement of emulsion droplet sizes using PFG NMR and regularization methods", *J. colloid and interface science*, 258 (2003) 383-389.
- Horacek, B. M. and Clements, J. C., "The inverse problem of electrocardiography: A solution in terms of single- and double-layer sources on the epicardial surface", *Mathematical biosciences*, 144 (1997) 119-154.
- Johnson, C. R. and MacLeod R. S., "Adaptive local regularization methods for the inverse ECG problem", *Progress in Biophysics & Molecular Biology* 69 (1998) 405-423.
- Kang, S. and Zabaras, N., "Control of the freezing interface motion in two-dimensional solidification processes using the adjoint method", *I. J. Numer. Meth. Engng.* 38 (1995) 63-80.
- Kim, S. K., Jung, B. S., Kim, H. J., and Lee, W., "Inverse estimation of thermophysical properties for anisotropic composite", *Experiment Thermal and Fluid Science*, 27 (2003) 697-704.
- Kitagawa, G. and Gersch, W., "A smoothness priors long AR model method for spectral estimation", *IEEE transactions on automatic control*, Vol.AC-30 (1985) 57-65.
- Fitzpatrick, B. G., "Bayesian analysis in inverse problems", *Inverse problems* 7 (1991) 675-702.
- Krishnan, M. and Sharma, D. G. R., "Determination of the interfacial heat transfer coefficient in unidirectional heat flow by Beck's non linear estimation procedure", *I. Comm. Heat Mass Transfer*, Vol. 23, No.2 (1996) 203-214.
- Li, K., Li, B. Q., and de Groh, H. C., "Numerical analysis of double-diffusive convection/solidification under g-jitter/magnetic fields", *Journal of Thermophysics and Heat Transfer*, Vol. 17, No.2 (2003) 199-209.

- Marin, L. and Lesnic, D., "BEM first-order regularization method in linear elasticity for boundary identification", *Comput. Methods Appl. Mech. Engrg.* 192 (2003) 2059-2071.
- Martin, T. J. and Dulikravich, G. S., "Inverse determination of boundary conditions and sources in steady heat conduction with heat generation", *Transaction of the ASME Vol. 118* (1996) 546-554.
- More, J. J., "The Levenberg-Marquardt algorithm implementation and theory", in *Lecture Notes in Mathematics*, No. 630, Numerical Analysis, G. Watson, ed., Springer-Verlag, (1978) 105-116
- Morozov, V. A., "Methods for Solving Incorrectly Posed Problems", Springer-Verlag, New York, 1984.
- Nagle R. K. and Saff, E. B., "Fundamentals of Differential Equations & Boundary Value Problems", Addison-Wesley, Reading, MA, 1993.
- Nocedal, J. and Wright, S. J., "Numerical optimization", Springer-Verlag, New York, 1999.
- Okamoto, K. and Li, B. Q., "Optimal regularization methods for inverse heat transfer problem" ASME Summer Heat Transfer Conference, Charlotte, NC, CD ROM, July, 2004.
- O'Mahoney, D. and Browne, D. J., "Use of experiment and an inverse method to study interface heat transfer during solidification in the investment casting process", *Experimental Thermal and Fluid Science*, 22 (2000) 111-122.
- Ozisik, M. N., "Heat conduction", John Wiley & Sons, New York, 1993.
- Ozisik, M. N. and Orlande, H. R. B., "Inverse heat transfer", Taylor & Francis, New York, 2000.
- Pitman, J., "Probability", Springer-Verlag, New York, 1993.
- Reginska, T., "A regularization parameter in discrete ill-posed problems", *SIAM J. Sci. Comp* Vol. 17 (1996) 740-749.
- Ruan, Y., "Analysis of temperature fields and thermal stresses in solidifying bodies using the finite element method", Ph.D. thesis, University of Minnesota, 1990.

- Sawaf, B., Ozisik, M. N., and Jarny, Y., "An inverse analysis to estimate linearly temperature dependent thermal conductivity components and heat capacity of an orthotropic medium", *Int. J. Heat Mass Transfer*, Vol. 38 (1995) 3005-3010.
- Schwarzkopf, J., Cader, T., Okamoto, K., Li, B. Q., and Ramaprian, B., "Effect of spray angle in spray cooling thermal management of electronics", *ASME Summer Heat Transfer Conference*, Charlotte, NC, CD ROM, July, 2004.
- Shu, Y., Li, B. Q., and de Groh, H. C., "Magnetic damping of g-jitter induced double diffusive convection in microgravity", *Numerical Heat Transfer*, Vol. 42, No.4 (2002) 345-364.
- Song, S. P. and Li, B. Q., "A hybrid boundary /finite element method for simulating viscous flows and shapes of droplets in electric fields", *I. J. of Computational Fluid Dynamics*, Vol. 15, No. 1 (2002) 125-144.
- Tautenhahn, U. and Jin, Q., "Tikhonov regularization and *a posteriori* rules for solving nonlinear ill posed problems", *Inverse problems* 19 (2003) 1-21.
- Trujillo, D. M. and Busby, H. R., "Optimal regularization of the inverse heat-conduction problem", *J. Thermophysics*, Vol.3 (1989) 423-427.
- Wahba, G., "A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem", *Annals of Statistics*, Vol.13 (1985) 1378-1402.
- Xu, R. and Naterer, G. F., "Inverse method with heat and entropy transport in solidification processing of materials", *J. Materials Processing Tech.* 112 (2001) 98-108.
- Yagola, A. G., Leonov, A. S., and Titarenko, V. N., "Data errors and an error estimation for ill-posed problems", *Inverse problems in engng*, vol. 10, No. 2, (2002), 117-129
- Yang, G. Z. and Zabaras, N., "An adjoint method for the inverse design of solidification processes with natural convection", *I. J. Numer. Meth. Engng.* 42 (1998) 1121-1144.
- Yoon, S. H. and Nelson, P. A., "Estimation of acoustic source strength by inverse methods: part II, Experimental investigation of methods for choosing regularization parameters", *J. sound and vibration* 233(4) (2000) 669-705.

- Zabaras, N., "Inverse finite element techniques for the analysis of solidification process",
I. J. Numer. Meth. Engng. 29 (1990) 1569-1587.
- Zabaras, N. and Kang, S., "On the solution of an ill-posed design solidification problem
using minimization techniques in finite and infinite-dimensional function spaces",
I. J. Numer. Meth. Engng. 36 (1993) 3973-3990.
- Zabaras N. and Nguyen, T. H., "Control of the freezing interface morphology in
solidification process in the presence of natural convection", I. J. Numer. Meth.
Engng. 38 (1995) 1555-1578.

APPENDIX A

DERIVATION OF MAXIMUM LIKELIHOOD METHOD

In this appendix, we show that maximizing $L(\{\bar{T}\}/\{q\}, \sigma^2)$ in Eq. (3.2.20) yields Eq. (3.2.21), and is equivalent to minimizing Eq. (3.2.22). The integrand in Eq. (3.2.20) is equal to

$$\begin{aligned}
 p(\{\bar{T}\}/\{q\}, \sigma^2) \cdot p(\{q\}/\alpha, \sigma^2) &= \left(\frac{1}{2\pi\sigma^2}\right)^M \cdot \det[\alpha[I]]^{1/2} \\
 &\cdot \exp\left[\left(\frac{-1}{2\sigma^2}\right)(\{q\} - \{\hat{q}\})^T ([J]^T [J] + \alpha[I])(\{q\} - \{\hat{q}\})\right] \\
 &\cdot \exp\left[\left(\frac{-1}{2\sigma^2}\right)R(\alpha)\right]
 \end{aligned} \tag{A.1}$$

where

$$\{\hat{q}\} = ([J]^T [J] + \alpha[I])^{-1} [J]^T [\{\bar{T}\} - \{T(\{q_0\})\}] \tag{A.2}$$

$$[A(\alpha)] = ([J]^T [J] + \alpha[I])^{-1} [J]^T \tag{A.3}$$

$$R(\alpha) = [\{\bar{T}\} - \{T(\{q_0\})\}]^T [[I] - [J][A(\alpha)]] [\{\bar{T}\} - \{T(\{q_0\})\}] \tag{A.4}$$

Integrating Eq. (A.1), we have

$$\begin{aligned}
 L(\{\bar{T}\}/\alpha, \sigma^2) &= \\
 &\left(\frac{1}{2\pi\sigma^2}\right)^{M/2} \det[\alpha[I]]^{1/2} \cdot \det[[J]^T [J] + \alpha[I]]^{-1/2} \cdot \exp\left[\frac{-R(\alpha)}{2\sigma^2}\right]
 \end{aligned} \tag{A.5}$$

where the following relation has been used,

$$\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(\{x\} - \{\eta\})^T [D]^{-1} (\{x\} - \{\eta\})\right] dx = (\sqrt{2\pi})^n \det[D]^{1/2} \tag{A.6}$$

Taking the logarithm of both sides of Eq. (A.5) yields

$$\begin{aligned} \log(L(\{\bar{T}\} / \alpha, \sigma^2)) &= \left(\frac{-M}{2}\right) \log(2\pi\sigma^2) + \\ &\log(\det[\alpha[I]])^{1/2} (\det[[J]^T [J] + \alpha[I]])^{-1/2} - \left(\frac{1}{2\sigma^2}\right) R(\alpha) \end{aligned} \quad (\text{A.7})$$

Taking the derivative of Eq. (A.7) with respect to σ^2 and setting it equal to zero, one has

$$\sigma^2 = \left(\frac{1}{M}\right) \cdot R(\alpha) \quad (\text{A.8})$$

The above σ^2 is proposed as the estimate of the noise variance. Substituting Eq. (A.8) into Eq. (A.5)

$$\begin{aligned} L(\{\bar{T}\} / \alpha, \sigma^2) &= \\ &\left(\frac{M}{2\pi R(\alpha)}\right)^{M/2} \det[\alpha[I]]^{1/2} \det[[J]^T [J] + \alpha[I]]^{-1/2} \cdot \exp\left(-\frac{M}{2}\right) \end{aligned} \quad (\text{A.9})$$

To maximize Eq. (A.9) with respect to α , we consider the following function:

$$ML_1(\alpha) = \left[\frac{1}{R(\alpha)}\right]^{M/2} \det[\alpha[I]]^{1/2} \det[[J]^T [J] + \alpha[I]]^{-1/2} \quad (\text{A.10})$$

With the relations,

$$\det[[J]^T [J] + \alpha[I]]^{-1} = \det\left[\frac{1}{\alpha} ([I] - [J][A])\right] \quad (\text{A.11})$$

and

$$\det \alpha[I] = \alpha^M \quad (\text{A.12})$$

Eq. (A.10) is expressed by

$$ML_1(\alpha) = \left\{ \frac{[\det([I] - [J][A])^{1/M}]^M}{R(\alpha)} \right\}^{M/2} \quad (\text{A.13})$$

Maximizing the likelihood function (A.13) can be obtained by minimizing the following equation,

$$\begin{aligned}
ML_2(\alpha) &= \left\{ \frac{R(\alpha)}{[\det([I] - [J][A])]^{1/M}} \right\} = \\
&\frac{[\{\bar{T}\} - \{T(\{q_0\})\}]^T ([I] - [J][A(\alpha)]) [\{\bar{T}\} - \{T(\{q_0\})\}]}{(\det[I] - [J][A(\alpha)])^{1/M}}
\end{aligned} \tag{A.14}$$

APPENDIX B

DERIVATION OF OCV AND GCV

In this appendix, the derivation of OCV and GCV function is shown. To do that, we write the cost function given by Eq. (3.2.23) as

$$S_0(\alpha, k) = [\{\bar{T}(k)\} - \{T\}]^T [\{\bar{T}(k)\} - \{T\}] + \alpha \{q\}^T \{q\} \quad (\text{B.1})$$

where the vector of the measured temperature $\{\bar{T}(k)\}$ is given by

$$\{\bar{T}^T(k)\} = [\bar{T}_1, \bar{T}_2, \dots, \bar{T}_{k-1}, T_k(\alpha, k), \bar{T}_{k+1}, \dots, \bar{T}_M] \quad (\text{B.2})$$

The solution that minimizes $S_0(\alpha, k)$ is given by

$$\{q(\alpha, k)\} = ([J]^T [J] + \alpha [I])^{-1} [J]^T (\{\bar{T}(k)\} - \{T(\{q_0\})\}) \quad (\text{B.3})$$

We write the vector of estimated temperature as

$$\begin{aligned} \{T(\alpha, k)\} &= [J] \{q(\alpha, k)\} + \{T(\{q_0\})\} = \\ &= [J]([J]^T [J] + \alpha [I])^{-1} [J]^T (\{\bar{T}(k)\} - \{T(\{q_0\})\}) + \{T(\{q_0\})\} \\ &= [B(\alpha)](\{\bar{T}(k)\} - \{T(\{q_0\})\}) + \{T(\{q_0\})\} \end{aligned} \quad (\text{B.4})$$

where the matrix $[B(\alpha)]$ is given by $[J]([J]^T [J] + \alpha [I])^{-1} [J]^T$. This relationship, when written in full, takes the form

$$\begin{bmatrix} T_1(\alpha, k) \\ T_2(\alpha, k) \\ \vdots \\ T_k(\alpha, k) \\ \vdots \\ T_M(\alpha, k) \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1M} \\ b_{21} & b_{22} & \dots & b_{2M} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \dots & b_{kM} \\ \vdots & \vdots & & \vdots \\ b_{M1} & b_{M2} & \dots & b_{MM} \end{bmatrix} \begin{bmatrix} \bar{T}_1 - T_{01} \\ \bar{T}_2 - T_{02} \\ \vdots \\ T_k(\alpha, k) - T_{0k} \\ \vdots \\ \bar{T}_M - T_{0M} \end{bmatrix} + \begin{bmatrix} T_{01} \\ T_{02} \\ \vdots \\ T_{0k} \\ \vdots \\ T_{0M} \end{bmatrix} \quad (\text{B.5})$$

where b_{ij} are the components of matrix $[B(\alpha)]$. From Eq. (B.3), the corresponding set of estimated temperature can be written as,

$$\{T(\alpha)\} = [J]\{q(\alpha)\} + \{T(\{q_0\})\} = [B(\alpha)](\{\bar{T}\} - \{T(\{q_0\})\}) + \{T(\{q_0\})\}. \quad (B.6)$$

When written in full, this relationship can be expressed as

$$\begin{bmatrix} T_1(\alpha) \\ T_2(\alpha) \\ \vdots \\ T_k(\alpha) \\ \vdots \\ T_M(\alpha) \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1M} \\ b_{21} & b_{22} & \cdots & b_{2M} \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kM} \\ \vdots & \vdots & & \vdots \\ b_{M1} & b_{M2} & \cdots & b_{MM} \end{bmatrix} \begin{bmatrix} \bar{T}_1 - T_{01} \\ \bar{T}_2 - T_{02} \\ \vdots \\ \bar{T}_k - T_{0k} \\ \vdots \\ \bar{T}_M - T_{0M} \end{bmatrix} + \begin{bmatrix} T_{01} \\ T_{02} \\ \vdots \\ T_{0k} \\ \vdots \\ T_{0M} \end{bmatrix} \quad (B.7)$$

From the above two matrix expressions a relationship between $T_k(\alpha, k)$ and $T_k(\alpha)$ obtained. It follows that

$$\begin{aligned} T_k(\alpha, k) &= b_{k1}(\bar{T}_1 - T_{01}) + b_{k2}(\bar{T}_2 - T_{02}) + \cdots \\ &+ b_{kk}(T_k(\alpha, k) - T_{0k}) + \cdots + b_{kM}(\bar{T}_M - T_{0M}) + T_{0k} \end{aligned} \quad (B.8)$$

and also that

$$\begin{aligned} T_k(\alpha) &= b_{k1}(\bar{T}_1 - T_{01}) + b_{k2}(\bar{T}_2 - T_{02}) + \cdots \\ &+ b_{kk}(\bar{T}_k - T_{0k}) + \cdots + b_{kM}(\bar{T}_M - T_{0M}) + T_{0k} \end{aligned} \quad (B.9)$$

Taking the difference between these two equations then shows that

$$(1 - b_{kk})T_k(\alpha, k) = T_k(\alpha) - b_{kk}T_{0k} \quad (B.10)$$

It therefore follows, after some algebra, that

$$\bar{T}_k - T_k(\alpha, k) = \frac{\bar{T}_k - T_k(\alpha)}{1 - b_{kk}} \quad (B.11)$$

This enables the expressions for the ordinary cross-validation function $V_0(\alpha)$ given by equation (3.2.24) to be written as

$$V_o(\alpha) = \frac{1}{M} \sum_{k=1}^M \left[\frac{(\bar{T}_k - T_k(\alpha))}{(1 - b_{kk})} \right]^2 \quad (\text{B.12})$$

Substituting $T_k(\alpha)$ from Eq. (B.6), we may write Eq. (B.12) as

$$V_o(\alpha) = \frac{1}{M} \left\| [\text{C}]([\text{I}] - [\text{B}(\alpha)])(\{\bar{\text{T}}\} - \{\text{T}(\{q_0\})\}) \right\|^2 \quad (\text{B.13})$$

where $[\text{C}]$ is the diagonal matrix whose entries are $1/(1-b_{kk})$. Note that $V_o(\alpha)$ is not a function of either $\{q\}$ to be restored or the noise but a function of $[\text{J}]$ the sensitivity matrix, $\{\bar{\text{T}}\}$ the measured temperature and α the regularization parameter.

The singular value decomposition (SVD) of the matrix $[\text{J}]$ gives the relationship $\{\bar{\text{T}}\} = [\text{J}]\{q\} + \{\text{T}(\{q_0\})\} + \{\varepsilon\}$ in the form

$$\{\bar{\text{T}}\} = [\text{U}][\text{S}][\text{V}]^T \{q\} + \{\text{T}(\{q_0\})\} + \{\varepsilon\} \quad (\text{B.14})$$

Since the unitary matrix $[\text{U}]$ has the property $[\text{U}]^T[\text{U}] = [\text{I}]$, pre-multiplication of Eq. (B.14) by $[\text{U}]^T$ results in

$$[\text{U}]^T \{\bar{\text{T}}\} = [\text{S}][\text{V}]^T \{q\} + [\text{U}]^T \{\text{T}(\{q_0\})\} + [\text{U}]^T \{\varepsilon\} \quad (\text{B.15})$$

This equation describes the relationship between the “transformed” temperature $\{\tilde{\text{T}}\} = [\text{U}]^T \{\bar{\text{T}}\}$ and the “transformed” heat flux $\{\tilde{q}\} = [\text{V}]^T \{q\}$ as described in [Barnett, 1990]. This equation is further transformed by pre-multiplication by the matrix $[\text{W}]$, which has the ik th entry given by

$$W_{ik} = \frac{1}{\sqrt{M}} e^{2\pi j i k / M}, \quad i, k = 1, 2, \dots, M \quad (\text{B.16})$$

where $j = \sqrt{-1}$. The matrix $[\text{W}]$ is a unitary matrix ($[\text{W}]^T[\text{W}] = [\text{I}]$). Premultiplying $[\text{W}]$, Eq. (B.15) becomes

$$[\text{W}][\text{U}]^T \{\bar{\text{T}}\} = [\text{W}][\text{S}][\text{V}]^T \{q\} + [\text{W}][\text{U}]^T \{\text{T}(\{q_0\})\} + [\text{W}][\text{U}]^T \{\varepsilon\} \quad (\text{B.17})$$

The transformed model is thus written as

$$\{\bar{\mathbf{T}}_{trans}\} = [\mathbf{J}_{trans}] \{\mathbf{q}_{trans}\} + \{\mathbf{T}(\{\mathbf{q}_0\})_{trans}\} + \{\boldsymbol{\varepsilon}_{trans}\} \quad (\text{B.18})$$

where $\{\bar{\mathbf{T}}_{trans}\} = [\mathbf{W}][\mathbf{U}]^T \{\hat{\mathbf{T}}\}$ is the vector of transformed measured temperature, $\{\mathbf{q}_{trans}\} = [\mathbf{W}][\mathbf{V}]^T \{\mathbf{q}\}$ is the vector of transformed source strengths, $\{\boldsymbol{\varepsilon}_{trans}\} = [\mathbf{W}][\mathbf{U}]^T \{\boldsymbol{\varepsilon}\}$ is the vector of transformed noise components and $[\mathbf{J}_{trans}] = [\mathbf{W}][\mathbf{S}][\mathbf{W}]^T$.

The procedure adopted by Golub *et al.* (1979) is then to apply ordinary cross-validation to this transformed model. Firstly, the transformed influence matrix $[\mathbf{B}_{trans}]$ is defined by $[\mathbf{B}_{trans}] = [\mathbf{J}_{trans}]([\mathbf{J}_{trans}]^T [\mathbf{J}_{trans}] + \alpha[\mathbf{I}])^{-1} [\mathbf{J}_{trans}]^T$. Replacing the terms by terms with those subscripted by trans in the ordinary cross-validation function (i.e. Eq. (B.13)) and noting that $[\mathbf{B}_{trans}]$ is a circulant matrix and thus its diagonal terms are the same and constant, we can write the generalized cross-validation function as

$$V(\alpha) = \frac{(1/M) \left\| ([\mathbf{I}] - [\mathbf{B}_{trans}]) (\{\bar{\mathbf{T}}_{trans}\} - \{\mathbf{T}_{trans}(\{\mathbf{q}_0\})\}) \right\|^2}{[(1/M) Tr([\mathbf{I}] - [\mathbf{B}_{trans}])]^2} \quad (\text{B.19})$$

where Tr denotes the trace (sum of diagonal entries) of a matrix. By expressing $V(\alpha)$ in terms of the eigenvalues of $[\mathbf{B}_{trans}]$ and noting that these are the same as the eigenvalues of $[\mathbf{B}(\alpha)]$ also, the generalized cross-validation function can be written as

$$V(\alpha) = \frac{(1/M) \left\| ([\mathbf{I}] - [\mathbf{B}]) (\{\bar{\mathbf{T}}\} - \{\mathbf{T}(\{\mathbf{q}_0\})\}) \right\|^2}{[(1/M) Tr([\mathbf{I}] - [\mathbf{B}])]^2} \quad (\text{B.20})$$

In comparison with Eq. (B.12), we see that the general cross-validation function $V(\alpha)$ is a weighted version of the ordinary cross-validation function $V_0(\alpha)$, that is,

$$V(\alpha) = \frac{1}{M} \sum_{k=1}^M \left[\frac{(\bar{T}_k - T_k(\alpha))}{(1 - b_{kk})} \right]^2 w_{kk} \quad (\text{B.21})$$

where w_{kk} is interpreted as the weighting function,

$$w_{kk} = \left(\frac{(1 - b_{kk})}{1 - (1/M) \text{Tr}[\mathbf{B}(\alpha)]} \right)^2 \quad (\text{B.22})$$