

Rochester Institute of Technology

RIT Digital Institutional Repository

Theses

11-20-2015

The Morozov Discrepancy Principle for the Elliptic Inverse Problem

Peter Caya
poc1673@rit.edu

Follow this and additional works at: <https://repository.rit.edu/theses>

Recommended Citation

Caya, Peter, "The Morozov Discrepancy Principle for the Elliptic Inverse Problem" (2015). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the RIT Libraries. For more information, please contact repository@rit.edu.

The Morozov Discrepancy Principle for the Elliptic Inverse Problem

By

Peter Caya

A thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science
in Applied Mathematics
from the School of Mathematical Sciences
College of Science
Rochester Institute of Technology

November 20, 2015

Advisor: Dr. Akhtar Khan

Co-Advisor: Dr. Baasansuren Jadamba

Committee members: Dr. Patricia Clark

Dr. Bonnie Jacob

Abstract

Inverse problems of parameter identification and source identification in partial differential equations are highly ill-posed problems and for their satisfactory theoretical and numerical treatment some sort of regularization is necessary. In this thesis, we pose this inverse problem as an optimization problem and perform the regularization in Tikhonov sense. The most crucial aspect of the study of the regularized optimization problem is a proper selection of the regularization parameter. Although the theory for one of the most efficient methods for choosing an optimal regularization parameter, the so-called Morozov discrepancy principle, is well-developed for linear inverse problems, its use for nonlinear inverse problems is rather heuristic. In this thesis, we investigate the inverse problem of parameter identification using an equation error approach in which the coefficient appear linearly. Using the results known for linear inverse problems, we develop a rigorous Morozov discrepancy principle for nonlinear inverse problems. We present a detailed computational experimentation to test the feasibility of the developed approach. We also study the inverse problem of source identification in fourth-order boundary value problem.

Dedication

This thesis is dedicated to my father and the memory of my mother. Through their examples I have learned to pursue knowledge, appreciate adversity, and to cultivate patience. Their support and love has been irreplaceable.

Acknowledgement

I owe my thanks to my father, mentors, and friends for their help and guidance in this learning process.

I am especially thankful to my advisors, Dr. Akhtar Khan and Dr. Baasansuren Jadamba who have been vital to my studies at RIT. They began advising me during my second semester and the generosity they have shown with their time and expertise is remarkable. Pursuing this research under them has broadened my understanding of mathematics and life.

I would also like to thank the School of Mathematical Sciences within the College of Science at RIT for their help throughout my time here. I would like to thank, in particular, Anna Fiorucci, and Carrie Koneski who have been instrumental throughout this journey.

Most importantly, I would like to thank my father. His unflagging encouragement has helped motivate me through my entire education. I am immensely grateful for the advice and support he has given me my whole life.

Special thanks to my committee members: Prof. Patricia Clark and Dr. Bonnie Jacob, and the former, and current director of graduate programs, Dr. Nathan Cahill, and Dr. Elizabeth Cherry.

Contents

List of Figures	vii
List of Tables	xi
1 Introduction	1
2 An Equation Error Approach with H_1 Regularization	5
2.1 Introduction	5
2.2 Equation Error Approach	7
2.3 Stability of the Equation Error Method	9
2.4 Numerical Results	11
3 Source Term Identification	14
3.1 The Optimal Regularization Parameter	14
3.1.1 The Morozov Discrepancy Principle	16
3.2 Numerical Methods	17
3.2.1 The Brute Force Method	17
3.2.2 Newton's Method	17
3.2.3 Cubically Convergent Methods	18
3.3 Computational Framework	18
3.3.1 The Discrete Morozov Equation	19
3.3.2 The Discrete Source Term	19
3.4 Numerical Experiments	20
3.4.1 Second Order Examples	20
3.4.2 Fourth Order Examples	30
3.5 Conclusion	39

4	Morozov Principle for the Equation Error Approach	40
4.1	The Equation Error Functional	40
4.2	Morozov Principle	44
4.3	Discretization and Implementation Details	45
4.4	Data Smoothing	48
4.5	Numerical Experiments	48
	4.5.1 Second Order Examples	49
4.6	Conclusion	71

List of Figures

2.1	Parameter recovery using the EE method and cgtrust with $\varepsilon = 10^{-6}$.	13
3.1	Reconstructions by Using β_{opt}	23
3.2	Noise of .025	23
3.3	Noise of .05	23
3.4	Noise of .07	23
3.5	Noise of .1	23
3.6	Reconstructions by Using β_M	23
3.7	Noise of .025	23
3.8	Noise of .05	23
3.9	Noise of .07	23
3.10	Noise of .1	23
3.11	Reconstructions by Using β_{opt}	26
3.12	Noise of .025	26
3.13	Noise of .05	26
3.14	Noise of .07	26
3.15	Noise of .1	26
3.16	Reconstructions by Using β_M	26
3.17	Noise of .025	26
3.18	Noise of .05	26
3.19	Noise of .07	26
3.20	Noise of .1	26
3.21	Reconstructions by Using β_{opt}	29
3.22	Noise of .025	29
3.23	Noise of .05	29
3.24	Noise of .07	29

3.25 Noise of .1	29
3.26 Reconstructions by Using β_M	29
3.27 Noise of .025	29
3.28 Noise of .05	29
3.29 Noise of .07	29
3.30 Noise of .1	29
3.31 Reconstructions by Using β_{opt}	32
3.32 Noise of .01	32
3.33 Noise of .025	32
3.34 Noise of .05	32
3.35 Noise of .07	32
3.36 Reconstructions by Using β_M	33
3.37 Noise of .01	33
3.38 Noise of .025	33
3.39 Noise of .05	33
3.40 Noise of .07	33
3.41 Reconstructions by Using β_{opt}	35
3.42 Noise of .025	35
3.43 Noise of .05	35
3.44 Noise of .07	35
3.45 Noise of .1	35
3.46 Reconstructions by Using β_M	36
3.47 Noise of .025	36
3.48 Noise of .05	36
3.49 Noise of .07	36
3.50 Noise of .1	36
3.51 Reconstructions by Using β_{opt}	38
3.52 Noise of .025	38
3.53 Noise of .05	38
3.54 Noise of .07	38
3.55 Noise of .1	38
3.56 Reconstructions by Using β_M	39
3.57 Noise of .025	39
3.58 Noise of .05	39
3.59 Noise of .07	39
3.60 Noise of .1	39

4.1	H^1 Error for Damped Parameters	51
4.2	Reconstructions Using β_{opt}	52
4.3	Noise of .005	52
4.4	Noise of .0075	52
4.5	Noise of .01	52
4.6	Noise of .05	52
4.7	Reconstructions Using β_M	52
4.8	Noise of .001	52
4.9	Noise of .005	52
4.10	Noise of .0075	52
4.11	Noise of .01	52
4.12	Reconstructions Using β_{opt}	55
4.13	Noise of .005	55
4.14	Noise of .0075	55
4.15	Noise of .01	55
4.16	Noise of .05	55
4.17	Reconstructions Using β_M	55
4.18	Noise of .005	55
4.19	Noise of .0075	55
4.20	Noise of .01	55
4.21	Noise of .05	55
4.22	Noise of .02	58
4.23	Noise of .04	58
4.24	Noise of .06	58
4.25	Noise of .08	58
4.26	Noise of .1	58
4.27	Noise of .12	58
4.28	Noise of .02	60
4.29	Noise of .04	60
4.30	Noise of .06	60
4.31	Noise of .08	60
4.32	Noise of .1	60
4.33	Noise of .12	60
4.34	Noise of .02	62
4.35	Noise of .04	62
4.36	Noise of .06	62

4.37 Noise of .08	62
4.38 Noise of .1	62
4.39 Noise of .12	62
4.40 Reconstructions Using β_{opt}	64
4.41 Noise of .001	64
4.42 Noise of .005	64
4.43 Noise of .0075	64
4.44 Noise of .01	64
4.45 Reconstructions Using β_M	65
4.46 Noise of .001	65
4.47 Noise of .005	65
4.48 Noise of .0075	65
4.49 Noise of .01	65
4.50 Noisy Data With and Without Smoothing	68
4.51 Noise of .02	68
4.52 Noise of .04	68
4.53 Noise of .06	68
4.54 Noise of .08	68
4.55 Noise of .02	70
4.56 Noise of .04	70
4.57 Noise of .06	70
4.58 Noise of .08	70
4.59 Noise of .02	71
4.60 Noise of .04	71
4.61 Noise of .06	71
4.62 Noise of .08	71

List of Tables

3.1	Brute Force Results	21
3.2	Newton-Type Results	22
3.3	Cubic Method Results	22
3.4	Damped Regularization Parameters	22
3.5	L^2 Error for Damped Parameters	22
3.6	Brute Force Results	24
3.7	Newton-Type Results	25
3.8	Cubic Method Results	25
3.9	Damped Regularization Parameters	25
3.10	L^2 Error for Damped Parameters	25
3.11	Brute Force Results	27
3.12	Newton-Type Results	27
3.13	Cubic Method Results	28
3.14	Damped Regularization Parameters	28
3.15	L^2 Error for Damped Parameters	28
3.16	Brute Force Results	31
3.17	Newton-Type Results	31
3.18	Cubic Method Results	31
3.19	Damped Regularization Parameters	32
3.20	L^2 Error for Damped Parameters	32
3.21	Brute Force Results	34
3.22	Newton-Type Results	34
3.23	Cubic Method Results	34
3.24	Damped Regularization Parameters	34
3.25	L^2 Error for Damped Parameters	35
3.26	Brute Force Results	37
3.27	Newton-Type Results	37

3.28	Cubic Method Results	37
3.29	Damped Regularization Parameters	37
3.30	L^2 Error for Damped Parameters	38
4.1	Brute Force Results	50
4.2	Newton-Type Results	51
4.3	Cubic Method Results	51
4.4	Brute Force Results	53
4.5	Newton-Type Results	53
4.6	Cubic Method Results	54
4.7	Damped Regularization Parameters	54
4.8	Error from Noisy Data	57
4.9	Error from Smoothing Data	57
4.10	Coefficient Reconstructions with Noisy Data	59
4.11	Coefficient Reconstructions with Smoothed Data	59
4.12	Morozov Results with Noisy Data	61
4.13	Morozov Results with Smoothed Data	61
4.14	Brute Force Results	63
4.15	Newton-Type Results	63
4.16	Cubic Method Results	63
4.17	Damped Regularization Parameters	63
4.18	Error from Noisy Data	67
4.19	Error from Smoothed Data	67
4.20	Coefficient Reconstructions with Noisy Data	69
4.21	Coefficient Reconstructions with Smoothed Data	69
4.22	Morozov Results with Noisy Data	70
4.23	Morozov Results with Smoothed Data	70

Chapter 1

Introduction

The conventional way a problem in applied mathematics is presented involves the presentation of a model from which a solution to the problem is determined. In applications this is not always the case. Often a practitioner is presented with a body of information that often incorporates inconsistent information (noise). It can often prove valuable to use this information to identify the parameters for a model which may be used to better understand the phenomena producing the initial data.

A simple example of this is determining the solution to a differential equation. In this case the differential equation is referred to as a forward problem whereas the case where the model is specified from data is the inverse problem. The importance of accurately estimating the model parameters based on data is manifested in a wide range of fields, most notably statistics, geophysics, engineering, and imaging.

For further clarification, we consider the following elliptic boundary value problem (BVP)

$$-\nabla \cdot (q\nabla u) = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \quad (1.1)$$

where Ω is a suitable domain in \mathbb{R}^2 or \mathbb{R}^3 and $\partial\Omega$ is its boundary. The above BVP models interesting real-world problems and has been studied in great detail. For instance, in (1.1), $u = u(x)$ may represent the steady-state temperature at a given point x of a body; then q would be a variable thermal conductivity coefficient, and f the external heat source. The system (1.1) also models underground steady state aquifers in which the parameter q is the aquifer transmissivity coefficient, u is the hydraulic head, and f is the recharge. The inverse problem in the context of the above BVP is to estimate the coefficient

q from a measurement z of the solution u . This inverse problem has been the subject of numerous papers, see [1, 5, 21]. Numerous other inverse problems for complicated boundary value problems and diverse applications can be found in [3, 4, 8, 17, 18, 26, 30, 27, 31, 33].

In recent years, the field of inverse problems has emerged as among one of the most vibrant and expanding branches of applied and industrial mathematics. Certainly the primary reason behind this is the ever-growing number of real-world situations that are being modeled and studied in a unified framework of inverse problems. However, the theoretical aspects of inverse problems are also challenging and require a fine blending of various branches of mathematics.

A number of approaches to the aforementioned inverse problem have been proposed in the literature; most involve either regarding (1.1) as a hyperbolic PDE in q or posing an optimization problem whose solution is an estimate of q . The work by Richter [28], who used a finite difference scheme to solve the PDE for q , is an example of the first approach. Furthermore, the approach of reformulating (1.1) as an optimization problem is divided into two possibilities, namely either formulating the problem as an unconstrained optimization problem or treating it as a constrained optimization problem, in which the PDE itself is the constraint. Among the optimization-based techniques the output least-squares method is among the most widely investigated methods. The output least-squares approach minimizes the functional

$$q \rightarrow \|u(q) - z\|^2, \quad (1.2)$$

where z is the data (the measurement of u) and $u(q)$ solves the variational form of (1.1) given by

$$\int_{\Omega} q \nabla u \cdot \nabla v = \int_{\Omega} f v, \quad \text{for all } v \in H_0^1(\Omega). \quad (1.3)$$

A known deficiency of the output least squares functional is that it often fails to be convex.

There are other functionals that have been used for the numerical solvability of the above inverse problem. For example, the equation error method (cf. [1, 2, 19]), consists of minimizing the functional

$$q \rightarrow \frac{1}{2} \|\nabla \cdot (q \nabla z) + f\|_{H^{-1}(\Omega)}^2$$

where $H^{-1}(\Omega)$ is the topological dual of $H_0^1(\Omega)$ and z is the data. In the same vein, Kohn and Lowe [22] proposed a variational method that combines features of the OLS and equation error methods. Ito and Kunisch [15] and Chen and Zou [6] developed an augmented Lagrangian algorithm to solve the OLS problem by treating the PDE as an explicit constraint.

In a related work, Knowles [21] proposed using a coefficient-dependent norm in the OLS setting

$$q \rightarrow \int q \nabla(u(q) - z) \cdot \nabla(u(q) - z), \quad (1.4)$$

where z is the data (the measurement of u) and $u(a)$ solves (1.3). Knowles [21] established that the above functional is convex. We note that the above functional, although in a discrete setting, was first proposed by Tucciarelli and Ahlfeld [29], who also explored its convexity. In [13], a new modified output least-squares (MOLS) was proposed to extend (1.4) and its convexity was proved in an abstract setting. Studies related to MOLS functional and its extensions can be found in [10, 12, 14, 16, 34].

Nonlinear inverse problems of parameter identification and source identification in partial differential equations are highly ill-posed problems and for their satisfactory theoretical and numerical treatment some sort of regularization is necessary. For example, the regularized analogue of the output least-squares approach results in the following optimization problem: Find q by solving

$$\min J(q) := \|u(q) - z\|^2 + \epsilon \|q\|^2, \quad (1.5)$$

where z is the data, $\epsilon > 0$ is the regularization parameter, and $\|q\|^2$ is the regularizer.

One of the crucial aspect of solving various optimization problems, emerging from different formulations, is a proper selection of the regularization parameter. If the parameter is too small then it does not provide the much needed stability, moreover, if it is too large then the solution of the regularized problem might not offer a good approximation.

Although the theory for one of the most efficient methods for choosing an optimal regularization parameter, the so-called Morozov discrepancy principle, is well-developed for linear inverse problems, its use for nonlinear inverse problems is rather heuristic. In this thesis, we investigate the inverse problem of parameter identification using an equation error approach in which

the coefficient appears linearly. Using the results known for linear inverse problems, we develop a rigorous Morozov discrepancy principle for nonlinear inverse problems. We present a detailed computational experimentation to test the feasibility of the developed approach. We also study the inverse problem of source identification in fourth-order boundary value problem.

The contents of this thesis are organized into four sections. In Chapter 2 we consider an equation error approach for the fourth-order boundary value problems. The objective of this chapter is show that under stronger conditions on the data, weaker conditions can be imposed on the regularization space. This chapter also shows that the equation error formulation results in an unconstrained optimization problem where in the main term the coefficient appears linearly. This formulation is precisely what is needed to extend the Morozov principle from linear inverse problems to the nonlinear ones.

In Chapter 3 we discuss the Morozov principle developed by Kunisch and Zou [24]. We conduct some numerical experiments for the source identification problem in certain boundary value problems.

Chapter 4 presents a Morozov principle for nonlinear inverse problems through the equation error formulation. We present a rigorous treatment of the Morozov principle. We present numerical examples to show the feasibility of the proposed framework.

Chapter 2

An Equation Error Approach with H_1 Regularization

This chapter deals with the nonlinear inverse problem of identifying a variable parameter in fourth-order partial differential equations using an equation error approach. These equations arise in several important applications such as car windscreen modeling, deformation of plates, etc. To counter the highly ill-posed nature of the considered inverse problem, a regularization must be performed. The main contribution of this work is to show that the equation error approach permits the use of H^1 regularization whereas other optimization-based formulations commonly use H_2 regularization. We give the existence and convergence results for the equation error formulation. An illustrative numerical example is given to show the feasibility of the approach.

2.1 Introduction

Let Ω be a bounded open domain in \mathbb{R}^2 with a sufficiently smooth boundary Γ and let $f \in L^2(\Omega)$ be a given function. Consider the following fourth-order elliptic boundary value problem

$$\Delta(a\Delta u) = f \quad \text{in } \Omega, \tag{2.1}$$

augmented with the clamped boundary conditions,

$$u = 0 \quad \text{on } \Gamma, \quad (2.2a)$$

$$\frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma. \quad (2.2b)$$

In this work, our objective is to study the inverse problem of identifying the material parameter a from a measurement z of u . Applications of this study are in beam and plate models as well as car windshield modeling (see [23, 31]). This nonlinear inverse problem has been explored using the output least squares (OLS) approach in which one attempts to find a minimizer of the functional

$$J(a) := \frac{1}{2} \|u(a) - z\|^2,$$

defined by using a suitable norm (see White [32]). Here z is the data (a measurement of u) and $u(a)$ is the unique solution of (2.1) that corresponds to the material parameter a ,

One of the primary obstacles in a satisfactory treatment of the OLS-based optimization framework is due to the fact that the OLS, in general, is non-convex. Our objective then is to investigate an equation error approach for solving the nonlinear inverse problem of identifying the material parameter a . In contrast to the OLS based optimization approach, the equation error approach results in solving a convex optimization problem. See also [2, 7, 9, 10, 11, 12, 13, 14, 11, 17, 16] for recent developments in parameter identification problems.

We emphasize that the equation error approach has two advantages over the OLS approach. Firstly, it leads to a convex optimization problem and hence it only possesses global solutions. Secondly, the equation approach is computationally quite inexpensive as there is no underlying variational problem to be solved. On the other hand, a deficiency of the approach is that, due to the fact that it relies on differentiating the data, it is quite sensitive to data contamination.

The equation error approach has been studied in the context of the following simpler second-order elliptic boundary valued problem:

$$-\nabla \cdot (a\nabla u) = f \quad \text{in } \Omega, \quad (2.3a)$$

$$u = 0 \quad \text{on } \Gamma. \quad (2.3b)$$

For (2.3), the equation error approach consists of finding a minimizer of the functional

$$a \rightarrow \frac{1}{2} \|\nabla \cdot (a \nabla z) + f\|_{H^{-1}(\Omega)}^2,$$

where $H^{-1}(\Omega)$ is the topological dual of $H_0^1(\Omega)$ and z is again the measured data.

In this chapter, we extend the equation error approach to identify the coefficient a in the fourth-order boundary value problem (2.1). Our strategy is motivated by the ideas presented originally by Acar [1] and Kärkkäinen [19] for (2.3) (see also [5]). Besides giving an existence theorem and a convergence result for the discretized problem, we also give some numerical examples.

This chapter is divided into four main sections. Section 2 provides essential background material for the problem and poses the solution of the inverse problem as a solvable minimization problem. Section 3 examines the stability of the equation error method and Section 4 provides a brief numerical example to show the preliminary computational feasibility of the proposed method.

2.2 Equation Error Approach

The variational formulation of (2.1) will be instrumental in formulating the equation error approach. The space suitable for the variational formulation is given by

$$V := \{v \in H^2(\Omega) : u = \frac{\partial u}{\partial n} = 0 \text{ on } \Gamma\}.$$

By multiplying (2.1) by a test function $v \in V$ and repeatedly using the well-known Green's formula we obtain the following variational formulation of (2.1): Find $u \in V$ such that

$$\int_{\Omega} a \Delta u \Delta v = \int_{\Omega} f v, \quad \text{for every } v \in V. \quad (2.4)$$

For a fixed pair $(a, w) \in L^\infty(\Omega) \times V$, we define the maps $E(a, w) : V \rightarrow V^*$ and $m : V \rightarrow \mathbb{R}$ by

$$\begin{aligned} E(a, w)(v) &= \int_{\Omega} a \Delta w \Delta v, \\ m(v) &= \int_{\Omega} f v. \end{aligned}$$

We note that, although the functional $E(a, w)$ was defined for fixed $a \in L^\infty(\Omega)$, $w \in V$, it remains well-defined for $a \in L^2(\Omega)$ and $w \in V \cap W^{2,\infty} := V^\infty$. In other words, we can sacrifice some regularity in a by requiring more regularity of u . This fact will play an important role below.

We first prove the following technical result for later use.

Lemma 2.2.1. *Assume that $u \in V^\infty$, $a \in L^2(\Omega)$, and $\{a_n\} \subset L^2(\Omega)$ is a sequence such that $a_n \rightarrow a$ in $L^2(\Omega)$. Then $E(a_n, u) \rightarrow E(a, u)$ in V^* .*

Proof. We begin by showing that the following inequality holds:

$$\|E(a, u)\|_{V^*} \leq \|a\|_{L^2} \|u\|_{V^\infty}. \quad (2.5)$$

In fact, using the definition of E , we have

$$|E(a, u)(v)| \leq \left| \int_{\Omega} a \Delta u \Delta v \right| \leq \|a \Delta u\|_{L^2} \|\Delta v\|_{L^2},$$

where

$$\|a \Delta u\|_{L^2}^2 = \int_{\Omega} a^2 (\Delta u)^2 \leq \|u\|_{V^\infty}^2 \|a\|_{L^2}^2,$$

and because $\|\Delta v\|_{L^2} \leq \|v\|_V$, we at once obtain (2.5).

To prove the main argument, we note that

$$(E(a_n, u) - E(a, u))(v) = \int_{\Omega} a_n \Delta u \Delta v - \int_{\Omega} a \Delta u \Delta v = \int_{\Omega} (a_n - a) \Delta u \Delta v,$$

which by using (2.5) implies that

$$|(E(a_n, u) - E(a, u))(v)| \leq \|u\|_{V^\infty} \|a_n - a\|_{L^2} \|v\|_V,$$

and consequently $\|E(a_n, u) - E(a, u)\|_{V^*} \leq \|u\|_{V^\infty} \|a_n - a\|_{L^2}$. The proof is complete. \square

Since the inverse problem at hand is ill-posed, some regularization is necessary. For this, we first define $A \subset H^1(\Omega)$ to be the closed and convex set of admissible coefficients. We consider the following regularized equation error functional to estimate a^* from a measurement z of u^* by minimizing

$$J(a; z, \varepsilon) = \|E(a, z) - m\|_{V^*}^2 + \varepsilon \|a\|_{H^1}^2. \quad (2.6)$$

Here it is assumed that $a^* \in A$ and $u^* \in V$ satisfy (2.1), $\varepsilon > 0$ is a regularization parameter, $z \in V$ is the data, and $\|\cdot\|_2^2$ is the regularization term.

Assuming that the data z is sufficiently smooth, we show that $J(\cdot; z, \beta)$ has a unique minimizer in $H^1(\Omega)$ for each $\beta > 0$.

Theorem 2.2.2. *Suppose $z \in W^\infty$. Then, for each $\beta > 0$, there exists a unique a_β satisfying*

$$J(a_\beta; z, \beta) \leq J(a; z, \beta), \text{ for all } a \in H^1(\Omega).$$

Proof. Since the functional J is bounded below, there exists a minimizing sequence $\{a_n\}$ for J . We have $\beta \|a_n\|_{H^1}^2 \leq J(a_n; z, \beta)$ for all n which implies that $\{a_n\}$ is bounded in $H^1(\Omega)$. Therefore, there exists $a_\beta \in H^1(\Omega)$ and a subsequence of $\{a_n\}$ (still denoted by $\{a_n\}$) such that $a_n \rightarrow a_\beta$ weakly in $H^1(\Omega)$ and, by Rellich's theorem, strongly in $L^2(\Omega)$. Since $z \in V^\infty$ and $a_n \rightarrow a_\beta$ in $L^2(\Omega)$, Lemma 2.2.1 confirms that $E(a_n, z) \rightarrow E(a_\beta, z)$ and since the norm is weakly lower semicontinuous, it follows that

$$\begin{aligned} \inf_{a \in H^1(\Omega)} J(a; z, \beta) &= \lim_{n \rightarrow \infty} J(a_n; z, \beta) \\ &= \lim_{n \rightarrow \infty} (\|E(a_n, z) - m\|_{V^*}^2 + \beta \|a_n\|_{H^1}^2) \\ &\geq \|E(a_\beta, z) - m\|_{V^*}^2 + \beta \|a_\beta\|_{H^1}^2 \\ &= J(a_\beta; z, \beta), \end{aligned}$$

confirming that a_β is a minimizer of $J(\cdot; z, \beta)$. The uniqueness of a_β follows from the fact that the regularized equation error functional is strictly convex. The proof is complete. \square

Since $J(a_\beta; z, \beta) \geq \inf_{a \in H^1(\Omega)} J(a; z, \beta)$, the last inequality in the above proof must actually hold as an equality and hence $\lim_{n \rightarrow \infty} \|a_n\|_{H^1} = \|a_\beta\|_{H^1}$ must remain valid. This, in view of the fact $a_n \rightarrow a_\beta$ weakly in $H^1(\Omega)$, ensures that $\{a_n\}$ actually converges to a_β strongly in $H^1(\Omega)$. Consequently any minimizing sequence of $J(\cdot; z, \beta)$ converges in $H^1(\Omega)$ to the unique minimizer a_β of $J(\cdot; z, \beta)$.

2.3 Stability of the Equation Error Method

Recall that $a^* \in A$ and $u^* \in V$ are assumed to satisfy (2.1). However, since a^* is not unique, we define the convex set $S = \{a \in H^1(\Omega) : E(a, u^*) = m\}$.

We can now prove the following stability result for the equation error approach.

Theorem 2.3.1. *Assume that $u^* \in V^\infty$ and $a^* \in H^1(\Omega)$ satisfy (2.1). Let $\{z_n\} \subset V^\infty$ be a sequence of observations of u^* that satisfy, with the sequences $\{\delta_n\}$, $\{\beta_n\}$, the conditions*

1. $\delta_n^2 \leq \beta_n \leq \delta_n$ for all $n \in \mathbb{Z}^+$;
2. $\delta_n^2/\beta_n \rightarrow 0$ as $n \rightarrow \infty$;
3. $\|z_n - u^*\|_{V^\infty} \leq \delta_n$ for all $n \in \mathbb{Z}^+$;
4. $\delta_n \rightarrow 0$ as $n \rightarrow \infty$.

For each $n \in \mathbb{Z}^+$, let a_n be the unique solution of

$$\min_{a \in H^1(\Omega)} J(a; z_n, \beta_n).$$

Then, there exists $\tilde{a} \in S$ such that $a_n \rightarrow \tilde{a}$ in $H^1(\Omega)$. Moreover, \tilde{a} satisfies $\|\tilde{a}\|_{H^1} \leq \|a\|_{H^1}$, for all $a \in S$.

Proof. Let $a \in S$ be arbitrary. Then,

$$\begin{aligned} \beta_n \|a_n\|_{H^1}^2 &\leq \|E(a, z_n) - m\|_{V^*}^2 + \beta_n \|a\|_{H^1}^2 \\ &= \|E(a, z_n - u^*)\|_{V^*}^2 + \beta_n \|a\|_{H^1}^2 \\ &\leq c \|a\|_{L^2}^2 \|z_n - u^*\|_{V^\infty}^2 + \beta_n \|a\|_{H^1}^2, \end{aligned}$$

implying that

$$\|a_n\|_{H^1}^2 \leq \|a\|_{L^2}^2 \frac{\delta_n^2}{\beta_n} + \|a\|_{H^1}^2, \quad (2.7)$$

and, in particular,

$$\|a_n\|_{H^1}^2 \leq \|a^*\|_{L^2}^2 \frac{\delta_n^2}{\beta_n} + \|a^*\|_{H^1}^2 \leq \|a^*\|_{L^2}^2 + \|a^*\|_{H^1}^2,$$

where we used the assumption $\delta_n^2 \leq \beta_n$. This proves that $\{a_n\}$ is bounded in $H^1(\Omega)$. Hence, by Rellich's lemma, there exists $\tilde{a} \in H^1(\Omega)$ and a subsequence $\{a_{n_k}\}$ such that $a_{n_k} \rightarrow \tilde{a}$ weakly in $H^1(\Omega)$ and strongly in $L^2(\Omega)$.

We claim that $\tilde{a} \in S$. Indeed, for any $\hat{a} \in S$, we have

$$\begin{aligned} \|E(a_{n_k}, u^*) - m\|_{V^*}^2 &= \|E(a_{n_k}, u^*) - E(a_{n_k}, z_{n_k}) + E(a_{n_k}, z_{n_k}) - m\|_{V^*}^2 \\ &\leq 2\|E(a_{n_k}, u^* - z_{n_k})\|_{V^*}^2 + 2\|E(a_{n_k}, z_{n_k}) - m\|_{V^*}^2 \\ &\leq 2\|a_{n_k}\|_{L^2}^2 \|z_{n_k} - u^*\|_{V^\infty}^2 + 2\|E(\hat{a}, z_{n_k}) - m\|_{V^*}^2 + 2\beta_{n_k} \|\hat{a}\|_{H^1}^2 \\ &\leq 2\|a_{n_k}\|_{L^2}^2 \delta_{n_k}^2 + 2\|\hat{a}\|_{L^2}^2 \delta_{n_k}^2 + 2\beta_{n_k} \|\hat{a}\|_{H^1}^2 \\ &\leq 2\|a_{n_k}\|_{L^2}^2 \delta_{n_k}^2 + 4\|\hat{a}\|_{H^1}^2 \delta_{n_k}, \end{aligned}$$

where we used $\delta_{n_k}^2 \leq \beta_{n_k} \leq \delta_{n_k}$ and the following inequality which remains true for any $\hat{a} \in S$:

$$\|E(\hat{a}, z_{n_k}) - m\|_{V^*}^2 + \beta_{n_k} \|\hat{a}\|_{H^1}^2 \leq \|\hat{a}\|_{L^2}^2 \delta_{n_k}^2 + \beta_{n_k} \|\hat{a}\|_{H^1}^2.$$

Because $\{\|a_{n_k}\|_{L^2}\}$ is bounded and $\delta_{n_k} \rightarrow 0$ as $k \rightarrow \infty$, this ensures that $\|E(a_{n_k}, u^*) - m\|_{V^*} \rightarrow 0$. Since we also have $E(a_{n_k}, u^*) \rightarrow E(\tilde{a}, u^*)$ by Lemma 2.2.1, this shows that $E(\tilde{a}, u^*) = m$ and hence that $\tilde{a} \in S$.

Using the fact that $a_{n_k} \rightarrow \tilde{a}$ weakly in $H^1(\Omega)$, we have $\|\tilde{a}\|_{H^1} \leq \liminf_{k \rightarrow \infty} \|a_{n_k}\|_{H^1}$. Moreover, by (2.7),

$$\beta_{n_k} \|a_{n_k}\|_{H^1}^2 \leq \|\tilde{a}\|_{L^2}^2 \delta_{n_k}^2 + \beta_{n_k} \|\tilde{a}\|_{H^1}^2,$$

which implies that

$$\|a_{n_k}\|_{H^1}^2 \leq \|\tilde{a}\|_{L^2}^2 \frac{\delta_{n_k}^2}{\beta_{n_k}} + \|\tilde{a}\|_{H^1}^2.$$

Since $\delta_{n_k}^2 / \beta_{n_k} \rightarrow 0$ as $k \rightarrow \infty$, this shows that $\limsup_{k \rightarrow \infty} \|a_{n_k}\|_{H^1} \leq \|\tilde{a}\|_{H^1}$. Therefore,

$$\|\tilde{a}\|_{H^1} \leq \liminf_{k \rightarrow \infty} \|a_{n_k}\|_{H^1} \leq \limsup_{k \rightarrow \infty} \|a_{n_k}\|_{H^1} \leq \|\tilde{a}\|_{H^1},$$

which shows that $\|a_{n_k}\|_{H^1} \rightarrow \|\tilde{a}\|_{H^1}$, and hence that $a_{n_k} \rightarrow \tilde{a}$ strongly in $H^1(\Omega)$ as $k \rightarrow \infty$.

Using (2.7),

$$\|\tilde{a}\|_{H^1}^2 \leq \lim_{k \rightarrow \infty} \|a_{n_k}\|_{H^1}^2 \leq \lim_{k \rightarrow \infty} \left(\|a\|_{L^2}^2 \frac{\delta_{n_k}^2}{\beta_{n_k}} + \|a\|_{H^1}^2 \right) = \|a\|_{H^1}^2$$

holds for every $a \in S$.

Finally, since the set S is a convex, there is a unique minimal H^1 -norm element, and we have shown that every convergent subsequence of $\{a_n\}$ converges to this unique element \tilde{a} . Thus the whole sequence $\{a_n\}$ must converge to \tilde{a} . This completes the proof. \square

2.4 Numerical Results

To test the preliminary effectiveness of the equation error approach for this inverse problem, we consider an example boundary value problem derived

from (2.1):

$$\begin{aligned} \Delta [a(x, y) \Delta u(x, y)] &= f(x, y) && \text{in } \Omega \\ u(x, y) = \frac{\partial u}{\partial n} &= 0 && \text{on } \Gamma \end{aligned} \tag{2.8}$$

where the solution u and parameter a are defined as

$$u(x, y) = 16x^2(1-x)^2y^2(1-y)^2 \quad \text{and} \quad a(x, y) = 4 + \sin(2\pi x) \sin(3\pi y).$$

For means of this numerical experiment, we take $f(x, y)$ as subsequently defined by (2.8). The domain Ω is taken as the unit square, $\Omega = (0, 1) \times (0, 1)$ with the boundary Γ as the square's outside edges.

Discretization of the solution was performed using cubic Hermite finite elements on a 20×20 mesh consisting of 882 triangles and 2,048 degrees of freedom.

The discretized optimization problem was solved using a conjugate-gradient trust-region method (`cgtrust`) with a stopping criteria on $\|\nabla J\|$ of 10^{-12} . Using a value of $\varepsilon = 10^{-6}$ for the regularization parameter with the H^1 -norm, `cgtrust` converged in 38 iterations. The computed solution at several iterations of the algorithm along with the output of the optimization are summarized in Figure 2.1. We note that this method provides a good reconstruction of the parameter in the interior of Ω with reconstruction error concentrated mostly along the boundaries.

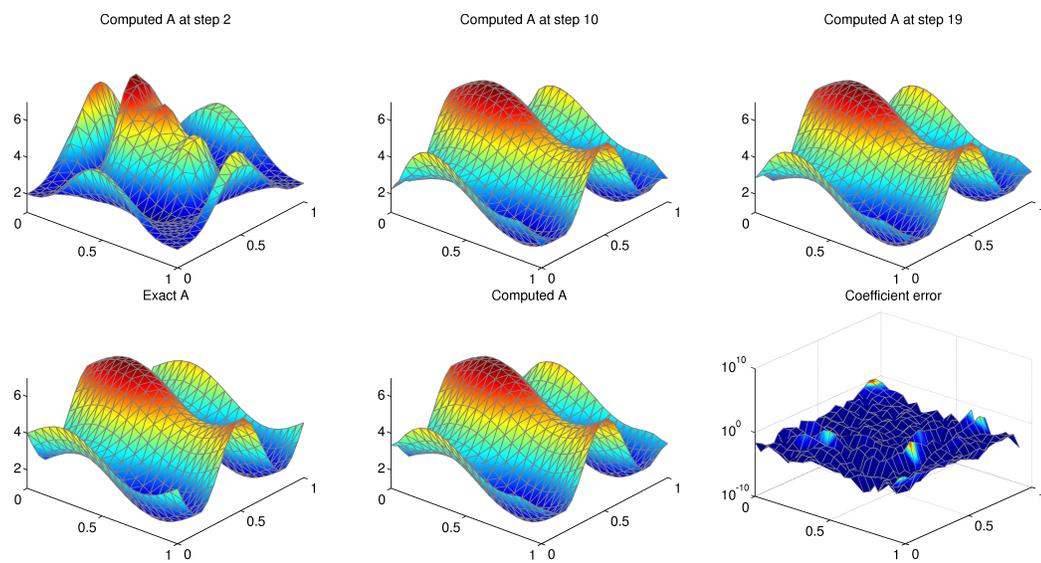


Figure 2.1 Parameter recovery using the EE method and `cgtrust` with $\varepsilon = 10^{-6}$.

Chapter 3

Source Term Identification

In this chapter, we conduct some numerical experiments for the theory of the Morozov discrepancy principle developed by Ito and Kunisch [25] where the authors focused on the linear inverse problems with an application to source term identification in partial differential equations. This chapter is a preparation for the next chapter where the approach will be extended to nonlinear inverse problems using an equation error formulation developed in the previous chapter.

This chapter will commence with a discussion of the mathematical framework employed to create a root finding problem balancing regularization error against noise error. This will be followed with numerical tests on several examples focusing on the inverse problem involving source term identification. The root finding problem will be compared to the brute force method by comparing the size of the L^2 errors as well as the speed of parameter identification. The convergence speed of the bisection method, Newton-type methods, and cubically convergent methods discussed in [35] will be compared.

3.1 The Optimal Regularization Parameter

Consider inverse problems of the form

$$Tf = z,$$

where T is a bounded operator mapping from a parameter space X to the observation space Y . Both spaces are taken to be real Hilbert spaces. We seek

to solve for the term f by posing an optimization problem defined through the regularized output least squares functional:

$$\min_{f \in X} J(f, \beta) = \frac{1}{2} \|Tf - z^\delta\|_Y^2 + \frac{\beta}{2} \|f\|_X, \quad (3.1)$$

where z^δ is noisy data and $\beta > 0$ is the regularization parameter.

The derivative of the above functional at f in any direction δf is given by:

$$DJ(f, \beta)(\delta f) = \langle Tf - z^\delta, T\delta f \rangle_Y + \beta \langle f, \delta f \rangle_X.$$

Therefore, the minimizer of the above optimization problem can be computed by using the following variational equation:

$$\langle Tf, Tg \rangle_Y + \beta \langle f, g \rangle_X = \langle z^\delta, Tg \rangle_Y \quad \forall g \in X. \quad (3.2)$$

In the following, we view f as a function of the regularization parameter β and proceed to obtain its derivative characterization. To derive an expression for the derivative a term t is added to β for (3.2):

$$\langle Tf(\beta + t), Tg \rangle_Y + (\beta + t) \langle f(\beta + t), g \rangle_X = \langle z^\delta, Tg \rangle_Y. \quad (3.3)$$

By subtracting (3.2) from (3.3), we obtain

$$\langle T(f(\beta + t) - f(\beta)), Tg \rangle_Y + \beta \langle (f(\beta + t) - f(\beta)), g \rangle_X + t \langle f(\beta + t), g \rangle_X = 0.$$

By dividing both sides by t , we get

$$\left\langle T \left(\frac{f(\beta + t) - f(\beta)}{t} \right), Tg \right\rangle_Y + \beta \left\langle \frac{f(\beta + t) - f(\beta)}{t}, g \right\rangle_X + \langle f(\beta + t), g \rangle_X = 0,$$

which when passed to the limit $t \rightarrow 0$, yields

$$\langle T(f'(\beta)), Tg \rangle_Y + \beta \langle (f'(\beta)), g \rangle_X + \langle f(\beta), g \rangle_X = 0.$$

Rearranging the system above produces the equivalent:

$$\langle Tf'(\beta), Tg \rangle_Y + \beta \langle f'(\beta), g \rangle_X = - \langle f(\beta), g \rangle_X.$$

It is further shown in [25] that $f(\beta)$ is infinitely differentiable, and that by proceeding as we have above $f^{(n)}(\beta)$ can be determined by solving the equation:

$$\langle Tf^{(n)}(\beta), Tg \rangle_Y + \beta \langle f^{(n)}(\beta), g \rangle_X = -n \langle f^{(n-1)}(\beta), g \rangle_X$$

3.1.1 The Morozov Discrepancy Principle

A popular criterion to estimate the best regularization parameter is the Morozov discrepancy principle which states that β should be chosen such that the error from regularization is equal to the error from the noise in the data. In context to the OLS functional this is expressed as the equality between the residual term and the noise level:

$$\|Tf(\beta) - z^\delta\|_Y^2 = \|z - z^\delta\|_Y^2.$$

We now define the value function F and recall derivative formulae:

$$\begin{aligned} F(\beta) &= \frac{1}{2}\|Tf - z^\delta\|_Y^2 + \frac{\beta}{2}\|f\|_X^2, \\ F'(\beta) &= \frac{1}{2}\|f(\beta)\|_X^2, \\ F''(\beta) &= \langle f_h(\beta), f'_h(\beta) \rangle_X, \\ F'''(\beta) &= \langle f'(\beta), f'(\beta) \rangle_X + \langle f(\beta), f''(\beta) \rangle_X. \end{aligned}$$

The Morozov equation may be rewritten in these terms to give a function and its derivatives in terms of the regularization parameter β :

$$\begin{aligned} G(\beta) &= F(\beta) - \beta F'(\beta) - \frac{1}{2}\|z - z^\delta\|_Y^2 \\ G'(\beta) &= -\beta F''(\beta) = -\beta \langle f'(\beta), f(\beta) \rangle \\ G''(\beta) &= -\beta F'''(\beta) - F''(\beta) \\ &= -\langle f_h(\beta), f'_h(\beta) \rangle_X - \beta(\langle f'(\beta), f'(\beta) \rangle_X + \langle f(\beta), f''(\beta) \rangle_X). \end{aligned}$$

The details can be found in [25].

Furthermore, we also recall that the damped version of the Morozov principle and its derivatives are given by:

$$\begin{aligned} G_\gamma(\beta) &= F(\beta) + (\beta^\gamma - \beta)F'(\beta) - \frac{1}{2}\|z - z^\delta\|_Y^2 \\ G'_\gamma(\beta) &= \gamma\beta^{\gamma-1}F'(\beta) + (\beta^\gamma - \beta)F''(\beta) \\ G''_\gamma(\beta) &= (\beta^\gamma - \beta)F'''(\beta) + (\gamma\beta^{\gamma-1} - 1)F''(\beta) + (\gamma - 1)\gamma\beta^{\gamma-2}F'(\beta) + \gamma\beta^{\gamma-1}F''(\beta). \end{aligned}$$

Here γ indicates the damping parameter. The damping parameter γ decreases the value of the solution β_M and will be denoted separately as β_M^γ .

Also note:

$$\lim_{\gamma \rightarrow \infty} G_\gamma(\beta) = G(\beta).$$

This convergence was observed to occur quickly as γ increases. For the tests in the work, γ parameters ranging from one to two were tested. In all numerical tests, $\gamma = 2$ produced results very close to the undamped equation.

3.2 Numerical Methods

3.2.1 The Brute Force Method

The most simple method to approximate the best regularization parameter is to calculate $f(\beta)$ for a large number of uniformly spaced regularization parameters, and then to record the L^2 error. One-thousand equally spaced regularization parameters were tested. The parameters near the regularization parameter producing the lowest L^2 error were then tested to refine the estimate of β_{opt} . The best estimate is used as the benchmark to compare the accuracy of the other methods and is referred to as the solution by brute force.

What solution by brute force gains in simplicity of implementation and accuracy given enough time it loses in speed. The brute force method was able identify a good regularization parameter for the examples in this paper, but as the scope of the problem becomes more demanding brute force becomes less capable of finding a good estimate in a timely manner.

3.2.2 Newton's Method

Since the Morozov principle is now a root finding problem, there are several different methods which can be considered in order to identify the value of β_M . The most basic method that may be applied is the bisection method. This paper follows the framework used by Kunisch and Zou [25] which concentrates on the quadratic convergence offered by Newton's method.

$$\beta_{k+1} = \beta_k - \frac{G(\beta_k)}{G'(\beta_k)}.$$

Two ways of implementing Newton's method are considered. The first involves calculating $f'(\beta)$ explicitly. The calculation of the derivative of the co-

efficient can be approximated taking the finite difference quotient:

$$\tilde{f}'_k(\beta_k, \beta_{k-1}) = \frac{f(\beta_k) - f(\beta_{k-1})}{\beta_k - \beta_{k-1}}$$

When $\tilde{f}'_k(\beta_k, \beta_{k-1})$ is used the method is referred to as the quasi-Newton method.

3.2.3 Cubically Convergent Methods

Three cubically convergent methods for the Morozov equation are considered in [35]. For the sake of brevity, the expression below is defined:

$$L(\beta_k) = G(\beta_k)[G'(\beta_k)]^{-2}G''(\beta_k).$$

The methods are as follows:

1. Chebyshev's Method:

$$\beta_{k+1} = \beta_k - \left(1 + \frac{1}{2}L(\beta_k)\right)[G'(\beta_k)]^{-1}G(\beta_k).$$

2. Halley Method:

$$\beta_{k+1} = \beta_k - \left(1 + \frac{1}{2}L(\beta_k)\left[1 - \frac{1}{2}L(\beta_k)\right]^{-1}\right)[G'(\beta_k)]^{-1}G(\beta_k).$$

3. Super Halley Method:

$$\beta_{k+1} = \beta_k - \left(1 + \frac{1}{2}L(\beta_k)\left[1 - L(\beta_k)\right]^{-1}\right)[G'(\beta_k)]^{-1}G(\beta_k).$$

Zou, Wang, and Zhang [35] noted that the cubic convergence of these methods requires the evaluation of $\frac{1}{2}n^2 + \frac{21}{6}n$ more multiplications and divisions than Newton's method.

3.3 Computational Framework

A finite element framework was used to approximate the equations in the finite dimensional subspace using the GETFEM library. The continuous piecewise finite element space will be represented as V^h . For brevity the finite dimensional analogue of $f(\beta)$ is represented as f_h while the stiffness matrix will be represented as A .

3.3.1 The Discrete Morozov Equation

$$\begin{aligned} J(f_h, \beta) &= \frac{1}{2} \langle U(f_h) - Z^\delta, U(f_h) - Z^\delta \rangle + \frac{\beta}{2} \langle f_h, f_h \rangle \\ &= \frac{1}{2} (U(f_h) - Z^\delta)^T M (U(f_h) - Z^\delta) + \frac{\beta}{2} f_h^T M f_h, \end{aligned}$$

where $U(g)$ solves the equation:

$$AU(g) = Mg \quad \forall g \in V^h$$

The discrete form of the derivatives of the OLS functional above are:

$$\begin{aligned} F'(\beta) &= \frac{\beta}{2} \langle f_h, f_h \rangle_X = \frac{\beta}{2} f_h^T M f_h. \\ F''(\beta) &= \langle f'_h, f_h \rangle_X = f_h^T M f'_h. \\ F'''(\beta) &= \langle f'(\beta), f'(\beta) \rangle_X + \langle f(\beta), f''(\beta) \rangle_X = f_h'^T M f'_h + f_h'^T M f_h. \end{aligned}$$

3.3.2 The Discrete Source Term

To find the value of $f(\beta)$, we substitute (3.2) into the discrete version of 3.2:

$$u(g)^T M u(f_h) + \beta g^T M f_h = g^T M Z^\delta$$

Substituting $u(g) = A^{-1} M g$ yields a system which can be used to solve for f_h .

$$g^T M A^{-1} M A^{-1} M f_h + \beta g_h^T M f_h = A^{-1} M A^{-1} M f_h + \beta f_h = A^{-1} M Z^\delta = g^T M Z^\delta.$$

Multiply both sides by $AM^{-1}A$:

$$\begin{aligned} M f_h + \beta A M^{-1} A f_h &= A Z^\delta \\ (M + \beta A M^{-1} A) f_h &= A Z^\delta. \end{aligned}$$

The discrete analogue for $f_h^{(n)}$ can be determined by similar operations:

$$\begin{aligned} \langle T f_h^{(n)}, T g \rangle_Y + \beta \langle f_h^{(n)}, g \rangle_X &= -n \langle f_h^{(n-1)}, g \rangle_X \\ \langle u(f_h^{(n)}), u(g) \rangle_Y + \beta \langle f_h^{(n)}, g \rangle_X &= -n \langle f_h^{(n-1)}, g \rangle_X \\ \langle A^{-1} M f_h^{(n)}, A^{-1} M g \rangle_Y + \beta \langle f_h^{(n)}, g \rangle_X &= -n \langle f_h^{(n-1)}, g \rangle_X \\ g^T M A^{-1} M A^{-1} M f_h^{(n)} + \beta g^T M f_h^{(n)} &= -n g^T M f_h^{(n)}. \end{aligned}$$

Eliminating g^T , and multiplying both sides by $AM^{-1}AM^{-1}$ yields:

$$(M + \beta A M^{-1} A) f_h^{(n)} = -n A M A f_h^{(n-1)}.$$

3.4 Numerical Experiments

The effectiveness of the brute force method and the Morozov principle were compared through tests on second and fourth order partial differential equations. The brute force method should produce the most accurate estimate of the best regularization parameter and will be relied upon as the benchmark for numerical results. The speed of the root finding methods used to solve the Morozov principle are also compared.

Noisy data was generated using the following method:

$$z^\delta(x) = z(x) + \hat{\delta}R * \max|z(x)|,$$

where R is a randomly generated number between -1 and 1 .

The Morozov principle was tested using the bisection, Newton-type, and cubically convergence methods with a stopping criteria of $\frac{|\beta_n - \beta_{n-1}|}{\beta_{n-1}} < 10^{-3}$. The reconstructions using β_M approached the accuracy of the brute force method generated parameters which were larger than those produced by brute force leading to less accurate reconstructions.

3.4.1 Second Order Examples

We now focus on the following second-order boundary value problem:

$$\begin{aligned} -\frac{d}{dx}\left(a(x)\frac{du}{dx}(x)\right) &= f(x) \\ u(0) &= u(1) = 0, \end{aligned}$$

where we are interested in identifying f from a measurement of u .

Example 3.4.1. *In this example, we take*

$$\begin{aligned} a(x) &= \exp(1 + x^2) \\ u(x) &= \exp(-x)\sin(\pi x). \end{aligned}$$

The aforementioned methods employing the Morozov principle were used in order to estimate the optimal regularization parameter. Regularization parameters ranging from .01 to .1 were used to test the way that the regularization parameters changed.

Displayed below is a progression showing the deformation of the source term as the noise is added. It can be observed that when no noise is present the formula for $f(\beta)$ produces very accurate estimates of the source term. As the level of noise is increased towards 10%, a gradual deformation of the source term occurs. Despite this, the source term recovery is recognizable.

The Morozov principle was able to identify good substitutes for β_{opt} . Although the L^2 error is larger, β_M was identified much more quickly and was the same general size as β_{opt} . The time to identify β_{opt} took well over ten seconds while the bisection method took a third of a second or less to converge. This reduction in time is more marked for the Newton-type methods which took fractions of a second. The cubic methods required fewer iterations to meet the termination criteria, and were generally faster than the Newton type methods though not dramatically so.

Using the damped Morozov principle for this example does not produce uniformly more accurate estimations of the regularization parameters. Despite this, the damped Morozov principle carries the useful option in that it has the potential of letting the practitioner make some limited adjustments if it is determined that β_{opt} is too large.

The smaller β_{opt} is, the faster the damped Morozov equation converges to the undamped version. The approximations created through damping were smaller than the undamped approximations and would be recommended if the undamped solutions are uniformly too large over several different noise levels. The convergence speed for all the numerical methods was observed to be similar to the undamped version.

Table 3.1 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0.01	4.6009e-07	2.1433
0.025	1.1402e-06	2.5392
0.05	2.1804e-06	2.9105
0.075	3.1806e-06	3.1854
0.1	4.2408e-06	3.419

Table 3.2 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	6.3325e-07	5.3844	47	0.32769	15	0.041561	10	0.016275
0.025	1.7193e-06	7.5812	46	0.25906	14	0.029702	9	0.017075
0.05	3.7747e-06	10.0475	49	0.25915	12	0.028729	9	0.017581
0.075	6.1463e-06	12.2177	48	0.33018	12	0.055068	9	0.015245
0.1	8.7055e-06	14.1111	45	0.2493	12	0.033118	8	0.01636

Table 3.3 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.01	8	0.011131	8	0.016402	7	0.014371
0.025	7	0.009576	7	0.01452	7	0.014419
0.05	7	0.012938	7	0.014206	6	0.011959
0.075	7	0.011225	6	0.012141	6	0.011855
0.1	6	0.011607	6	0.01182	6	0.015955

Table 3.4 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	1.0076e-08	1.4732e-07	5.1815e-07	6.2486e-07	6.3275e-07	6.3322e-07
.025	4.4229e-08	4.9648e-07	1.4262e-06	1.6935e-06	1.7175e-06	1.7192e-06
.05	1.3078e-07	1.2019e-06	3.116e-06	3.7064e-06	3.7689e-06	3.7742e-06
.075	2.4408e-07	2.0159e-06	5.0508e-06	6.0221e-06	6.1348e-06	6.1453e-06
.1	3.7913e-07	2.9284e-06	7.1718e-06	8.5221e-06	8.6872e-06	8.7037e-06

Table 3.5 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	73.5511	6.9715	5.3299	5.3787	5.3841	5.3844
.025	90.4508	10.0929	7.5214	7.5718	7.5805	7.5811
.05	101.6111	12.2118	9.8147	10.0182	10.045	10.0473
.075	105.0999	13.5819	11.8375	12.1707	12.2133	12.2173
.1	105.9693	14.8735	13.6869	14.0583	14.1058	14.1106

Figure 3.1 Reconstructions by Using β_{opt}

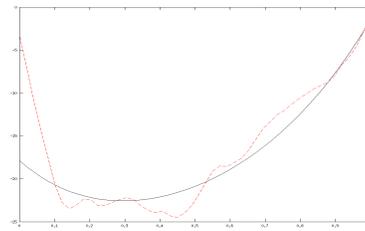
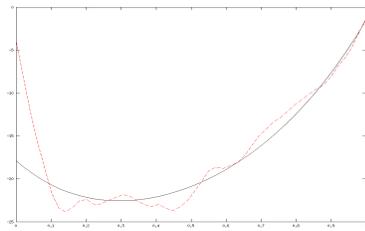


Figure 3.2 Noise of .025

Figure 3.3 Noise of .05

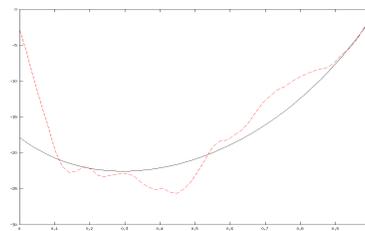
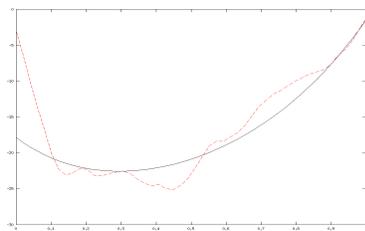


Figure 3.4 Noise of .07

Figure 3.5 Noise of .1

Figure 3.6 Reconstructions by Using β_M

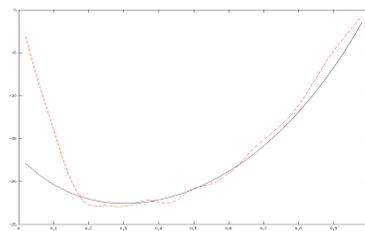
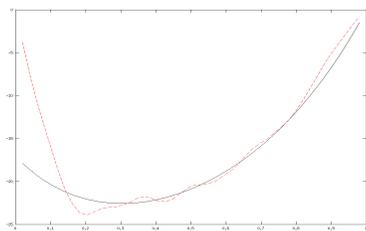


Figure 3.7 Noise of .025

Figure 3.8 Noise of .05

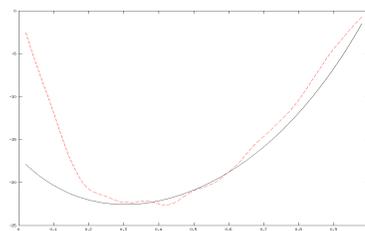
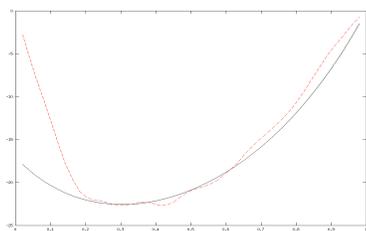


Figure 3.9 Noise of .07

Figure 3.10 Noise of .1

Example 3.4.2. *In the second example was tested using the equations:*

$$\begin{aligned} a(x) &= .5\cos(4\pi x) + 1 \\ u(x) &= \sin(4\pi x) \end{aligned}$$

Using the brute force method in the same manner as the first example produced accurate reconstructions with evident roughness for larger noise levels. The Morozov results showed similar performance to the brute force estimate. In this case the Morozov principle identified a slightly under-regularized parameter. Despite the increase in the L^2 error term in the case of Morozov, the illustrations show that the estimated parameter from this method still compensate for the added noise very well.

Since the regularization parameters were generally too small the damped Morozov principle failed to yield improvements on the undamped results. The scenario where $\gamma = 1$ produced particularly inaccurate results. With the exception of the scenario where the noise level is .1, the best results were given using little or no damping at all.

The performance of the root finding algorithms is similar to those from example one. The Newton method greatly outperforms the quasi-Newton method. The cubic methods require fewer iterations than the Newton-type methods but yielded no appreciable difference in performance time. Of the cubic methods, Halley's is the only one to repeatedly outperform the quadratic methods.

Table 3.6 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0	2.0004e-08	0.85309
0.01	1.0002e-07	4.8669
0.025	2.0004e-07	8.9274
0.05	3.2006e-07	13.5338
0.075	4.6009e-07	16.9697
0.1	5.6011e-07	19.7684

Table 3.7 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	6.3129e-08	28.4478	54	0.36534	15	0.036433	10	0.017254
0.025	1.5164e-07	83.5626	55	0.36888	15	0.038394	9	0.020121
0.05	2.9813e-07	184.4505	55	0.32009	15	0.041738	8	0.014713
0.075	4.472e-07	287.8678	57	0.35604	15	0.030054	8	0.013196
0.1	6.0026e-07	391.659	57	0.39746	15	0.041048	7	0.016241

Table 3.8 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.01	8	0.018758	7	0.016512	7	0.015369
0.025	7	0.01342	7	0.014417	6	0.012006
0.05	6	0.011069	6	0.012281	6	0.014298
0.075	6	0.016636	6	0.011909	6	0.011853
0.1	6	0.022234	6	0.012937	5	0.013483

Table 3.9 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	9.0955e-10	2.2351e-08	5.9482e-08	6.2988e-08	6.3124e-08	6.3129e-08
.025	4.8813e-09	7.7432e-08	1.4587e-07	1.5138e-07	1.5163e-07	1.5164e-07
.05	1.6668e-08	1.8377e-07	2.8945e-07	2.9769e-07	2.9811e-07	2.9813e-07
.075	3.3416e-08	2.9875e-07	4.358e-07	4.4657e-07	4.4717e-07	4.472e-07
.1	5.4072e-08	4.1995e-07	5.8615e-07	5.9944e-07	6.0021e-07	6.0026e-07

Table 3.10 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	1059.5511	70.9781	29.7451	28.4938	28.4495	28.4479
.025	1712.8158	140.8563	85.0657	83.6241	83.5652	83.5627
.05	2310.354	240.1195	185.5074	184.4964	184.4527	184.4506
.075	2833.989	333.4067	288.1503	287.874	287.8681	287.8678
.1	3237.6296	424.1377	390.9766	391.6086	391.656	391.6588

Figure 3.11 Reconstructions by Using β_{opt}

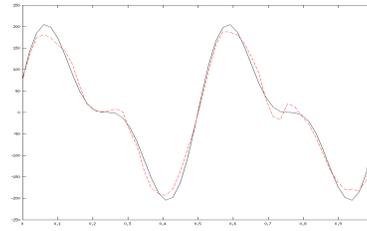
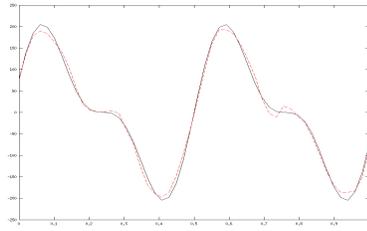


Figure 3.12 Noise of .025

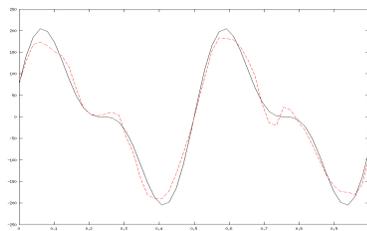


Figure 3.13 Noise of .05

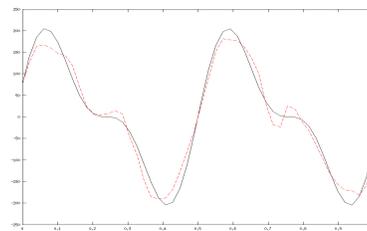


Figure 3.14 Noise of .07

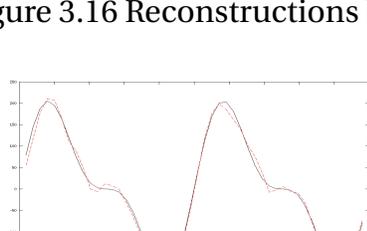


Figure 3.15 Noise of .1

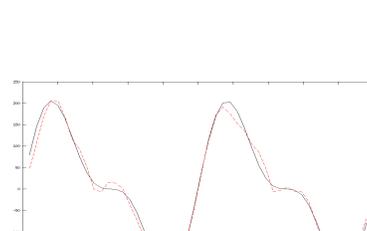


Figure 3.16 Reconstructions by Using β_M

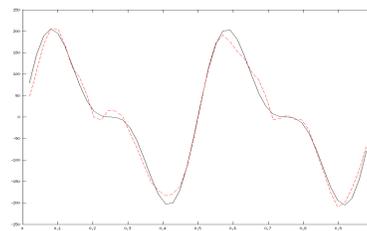
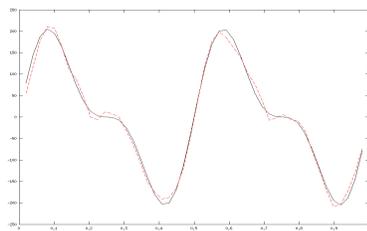


Figure 3.17 Noise of .025

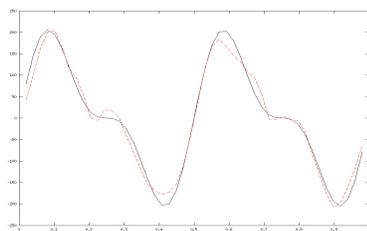


Figure 3.18 Noise of .05

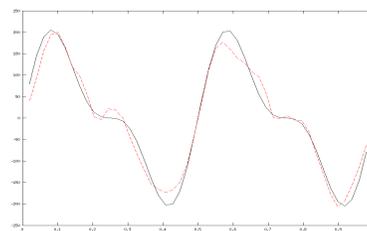


Figure 3.19 Noise of .07

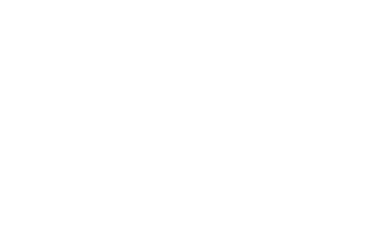


Figure 3.20 Noise of .1



Example 3.4.3. *In this example we take the data:*

$$\begin{aligned} a(x) &= \log(x + 2) \\ u(x) &= -(x^4 - x^3) \end{aligned}$$

The brute force method produces satisfactory results with the shortfall of noticeable oscillations in the reconstructions. The Morozov principle produces results of similar accuracy to the brute force method in terms of error. In this case β_M yielded smoother but less accurate results. Using the damped Morozov principle with parameter $\gamma = 1.2$ compensates for this overregularization and was able to produce superior results to β_M in some cases.

The performance of the root finding methods for the Morozov principle was similar to the previous two examples with drastic improvements found using the Newton-type methods, and more modest ones with the cubic methods. The cubic methods generally reduced the number of steps to convergence by one or two which was enough to ensure that these methods outperformed the Newton-type methods

Table 3.11 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0	1.0002e-08	0.088554
0.01	1.8004e-07	0.55166
0.025	8.4017e-07	0.7959
0.05	2.7005e-06	0.99174
0.075	5.3411e-06	1.1136
0.1	8.6817e-06	1.2037

Table 3.12 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	8.9217e-07	0.42687	43	0.3383	20	0.056131	7	0.014417
0.025	3.6837e-06	0.78618	44	0.34239	39	0.10613	6	0.011469
0.05	1.0713e-05	1.1647	44	0.2467	13	0.032145	5	0.008464
0.075	1.9921e-05	1.4324	43	0.27972	14	0.042103	6	0.015632
0.1	3.0548e-05	1.6348	45	0.32304	13	0.02761	6	0.011481

Table 3.13 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal, Time
0.01	6	0.012873	6	0.012598	5	0.010036
0.025	5	0.008981	5	0.010575	5	0.02391
0.05	4	0.007096	4	0.007342	4	0.007466
0.075	5	0.009337	4	0.007941	4	0.007882
0.1	5	0.008891	5	0.009701	5	0.013512

Table 3.14 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	4.188e-08	4.6136e-07	8.4333e-07	8.8896e-07	8.9197e-07	8.9216e-07
.025	2.4703e-07	2.0349e-06	3.4572e-06	3.6639e-06	3.6821e-06	3.6836e-06
.05	9.506e-07	6.2076e-06	9.9955e-06	1.0635e-05	1.0705e-05	1.0713e-05
.075	2.0907e-06	1.1924e-05	1.855e-05	1.9752e-05	1.9901e-05	1.9918e-05
.1	3.6489e-06	1.888e-05	2.8466e-05	3.0271e-05	3.0513e-05	3.0544e-05

Table 3.15 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.01	0.53389	0.35247	0.41935	0.42638	0.42684	0.42686
.025	0.81846	0.69234	0.7743	0.78515	0.7861	0.78617
.05	1.1487	1.0541	1.1482	1.1629	1.1645	1.1647
.075	1.3895	1.318	1.4141	1.4302	1.4322	1.4324
.1	1.589	1.5266	1.6166	1.6324	1.6345	1.6348

Figure 3.21 Reconstructions by Using β_{opt}

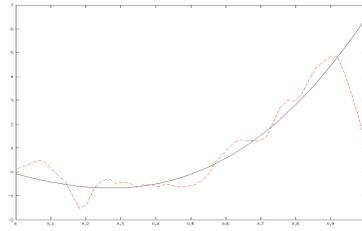
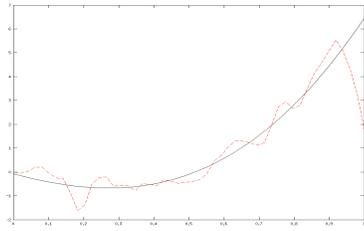


Figure 3.22 Noise of .025

Figure 3.23 Noise of .05

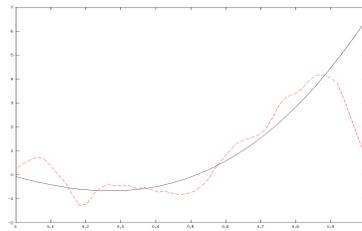
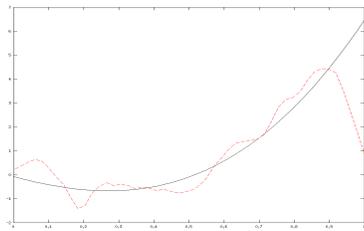


Figure 3.24 Noise of .07

Figure 3.25 Noise of .1

Figure 3.26 Reconstructions by Using β_M

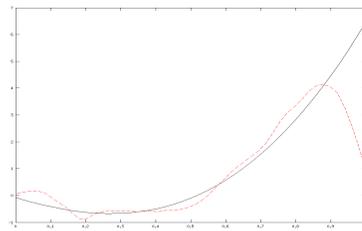
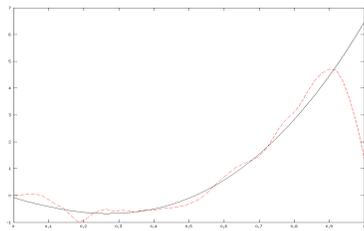


Figure 3.27 Noise of .025

Figure 3.28 Noise of .05

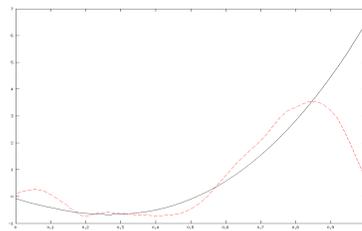
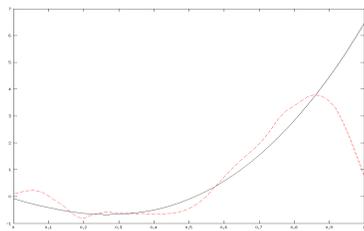


Figure 3.29 Noise of .07

Figure 3.30 Noise of .1

3.4.2 Fourth Order Examples

In this section, we focus on the inverse problem of identifying the source term f in the beam equation. To be specific, we focus on the following static fourth-order boundary value problem in $\Omega := (0, 1)$:

$$(a(x)u'')'' = f(x), \quad (3.4)$$

$$u(0) = u'(0) = 0, \quad (3.5)$$

$$u(1) = u'(1) = 0, \quad (3.6)$$

where $a(x)$ is a variable coefficient and f is the source term.

The boundary conditions (3.5)-(3.6) are the so-called clamped boundary conditions. However, our approach can easily be carried over to other types of boundary conditions as well.

In this case, the problems were interpolated on a grid of fifty points. Implementing the fourth order problem is made more computationally complex due to the necessity for cubic interpolating functions instead of the simple hat function considered in the second order examples. In this case the ill-posed derivative operation is performed four times instead of two making the inaccuracy introduced by noise especially egregious. For this set of problems the damped or undamped Morozov principles were found to be able to identify satisfactory regularization parameters when used with the same tolerance criteria as the second order problems.

Example 3.4.4. *The first fourth order example is:*

$$a(x) = \left(2x - \frac{1}{2}\right)^3 + 2$$

$$u(x) = \sin\left(4\pi x - \frac{\pi}{2}\right) + 1.$$

Selecting the regularization parameter by brute force produces results which begin to depart from the actual source term as the noise level is increased. Despite this, the results stay reasonably accurate as the level of noise is increased to ten-percent after which significant flattening was noted for the reconstruction.

The undamped Morozov principle repeats its excellent performance in the estimation of the best regularization parameter discussed for the second order examples. In all cases β_M is comparable to β_{opt} . This follows for the

associated L^2 errors. While producing results which are nearly the same as using the brute force method, identifying β_M requires fractions of a second while finding β_{opt} for this more difficult problem required minutes of computation time. Using the damped Morozov principle gives reconstructions which are much poorer quality than the undamped version due to underestimating the regularization parameter.

For this example the cubic methods performed particularly well. Convergence took six iterations at most with the cubic methods while Newton's method and the quasi-Newton methods took twelve and fourteen iterations respectively. Due to the larger decrease in necessary iterations for the cubic methods only a fraction of the time was needed when compared to the Newton-type methods. Newton's method converges much more quickly than the quasi-Newton method.

Table 3.16 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0	1e-13	24212.8767
0.01	1.1262e-12	27805.9067
0.025	1.7286e-11	28923.8263
0.05	4.7291e-11	29270.0184
0.075	7.3159e-11	29535.5923
0.1	9.7472e-11	29811.8629

Table 3.17 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	5.6995e-12	27825.0414	55	1.0014	14	0.14168	12	0.08534
0.025	1.925e-11	28755.8655	54	0.74379	14	0.097445	12	0.062245
0.05	4.0442e-11	29188.2471	53	0.72623	14	0.1035	12	0.056391
0.075	6.0821e-11	29397.7266	52	0.81902	14	0.10052	12	0.069541
0.1	8.0774e-11	29538.6465	52	0.93815	14	0.13344	12	0.08215

Table 3.18 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.01	6	0.037872	6	0.042068	5	0.03531
0.025	5	0.021909	5	0.030133	5	0.030193
0.05	4	0.02176	4	0.024303	4	0.025639
0.075	4	0.021718	4	0.024334	4	0.024623
0.1	4	0.015978	4	0.020681	4	0.019554

Table 3.19 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	3.3824e-13	5.3824e-12	5.8616e-12	5.8646e-12	5.8646e-12	5.8646e-12
0.025	1.7679e-12	1.9502e-11	2.0551e-11	2.0559e-11	2.0559e-11	2.0559e-11
0.05	5.9217e-12	4.2436e-11	4.3854e-11	4.3867e-11	4.3867e-11	4.3867e-11
0.075	1.1588e-11	6.4188e-11	6.5862e-11	6.5878e-11	6.5878e-11	6.5878e-11
0.1	1.8191e-11	8.5311e-11	8.7205e-11	8.7224e-11	8.7224e-11	8.7224e-11

Table 3.20 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	27221.9282	27782.7191	27857.8645	27858.3131	27858.3156	27858.3156
0.025	29190.2758	28915.8508	28944.8006	28945.0156	28945.0171	28945.0171
0.05	30957.3184	29632.2977	29635.5921	29635.6257	29635.6261	29635.6261
0.075	32799.478	30027.359	30022.8509	30022.8167	30022.8164	30022.8164
0.1	34292.116	30311.9284	30304.4664	30304.4023	30304.4015	30304.4015

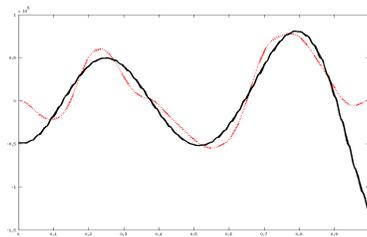
Figure 3.31 Reconstructions by Using β_{opt} 

Figure 3.32 Noise of .01

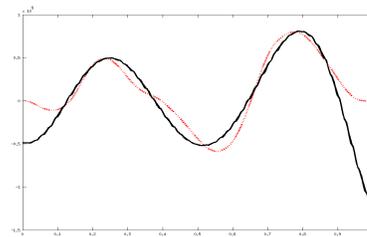


Figure 3.33 Noise of .025

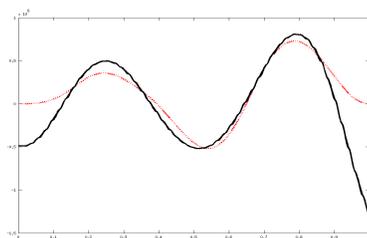


Figure 3.34 Noise of .05

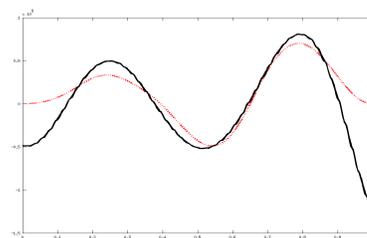


Figure 3.35 Noise of .07

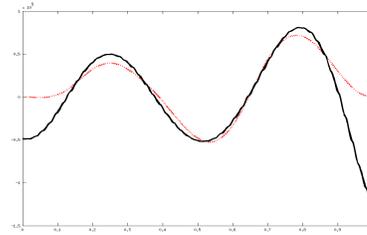
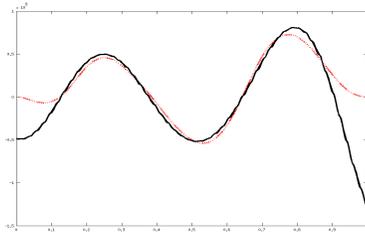
Figure 3.36 Reconstructions by Using β_M 

Figure 3.37 Noise of .01

Figure 3.38 Noise of .025

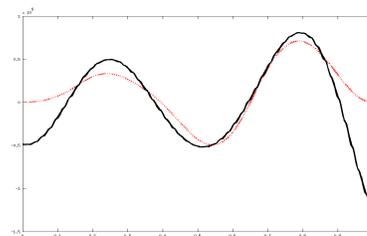
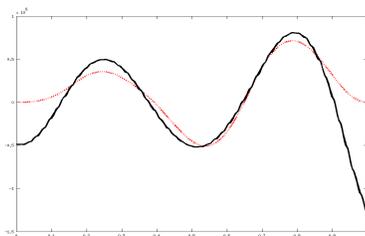


Figure 3.39 Noise of .05

Figure 3.40 Noise of .07

Example 3.4.5. A second fourth order example reinforces the observations made with the first fourth order problem:

$$a(x) = (1 + x)$$

$$u(x) = -\cos(2\pi x) + 1.$$

The reconstructions using the brute force method are recognizable though significant flattening is observed for the larger noise levels. The Morozov principle begins by over-regularizing the problem but for larger noise terms under-estimates the best parameter. The additional L^2 error from the Morozov equation is not significant however, and the illustration of the reconstructions using β_M show that they are comparable to those generated using brute force. Using the damped Morozov principle with a parameter $\gamma = 1.2$ gives minor improvements to the reconstructions in terms of the error.

Unlike example one, the cubic methods fail to converge quickly enough to warrant their preference over Newton's method. Chebyshev's method was able match or beat Newton's method by a small margin, but not to an appreciable degree. Newton's method outperformed the quasi-Newton method to a greater degree in this case than in the previous example.

Table 3.21 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0	1e-13	608.7772
0.01	7.024e-11	959.2497
0.025	7.3019e-10	1116.3738
0.05	1.768e-09	1214.4901
0.075	3.0341e-09	1279.5203
0.1	4.9815e-09	1323.6757

Table 3.22 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	3.706e-10	1024.4985	66	1.0818	17	0.12142	5	0.020392
0.025	1.2823e-09	1129.4963	64	1.1428	17	0.12151	4	0.015404
0.05	3.9448e-09	1244.9522	63	1.2225	17	0.16595	5	0.028324
0.075	8.7322e-09	1313.1052	61	0.87199	17	0.12383	4	0.0157
0.1	1.4388e-08	1345.2415	61	0.86456	17	0.1216	4	0.015664

Table 3.23 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.01	4	0.015341	4	0.019397	4	0.021323
0.025	4	0.015551	4	0.021344	4	0.019529
0.05	4	0.021246	5	0.032064	5	0.032426
0.075	4	0.015701	4	0.020688	4	0.021398
0.1	4	0.015649	4	0.019893	4	0.020298

Table 3.24 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	3.6063e-11	3.3642e-10	3.7011e-10	3.7059e-10	3.706e-10	3.706e-10
0.025	1.8434e-10	1.1851e-09	1.2806e-09	1.2823e-09	1.2823e-09	1.2823e-09
0.05	5.7337e-10	3.4947e-09	3.9338e-09	3.9446e-09	3.9448e-09	3.9448e-09
0.075	1.0895e-09	7.4124e-09	8.6931e-09	8.7311e-09	8.7322e-09	8.7322e-09
0.1	1.7176e-09	1.2136e-08	1.4315e-08	1.4386e-08	1.4388e-08	1.4388e-08

Table 3.25 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	1000.7046	1020.1553	1024.4393	1024.4978	1024.4985	1024.4985
0.025	1199.066	1126.1546	1129.4337	1129.4952	1129.4962	1129.4963
0.05	1336.4395	1237.7582	1244.7808	1244.9487	1244.952	1244.9522
0.075	1383.3637	1305.9523	1312.9082	1313.1	1313.105	1313.1052
0.1	1413.0761	1340.3796	1345.0933	1345.238	1345.2416	1345.2415

Figure 3.41 Reconstructions by Using β_{opt}

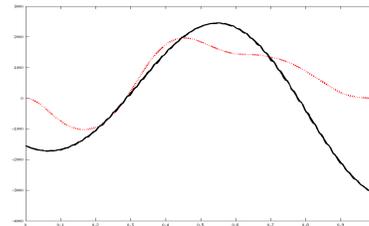
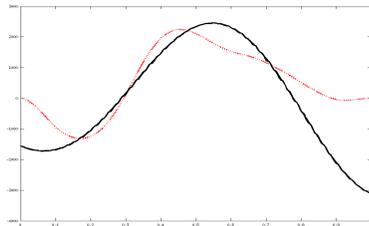


Figure 3.42 Noise of .025

Figure 3.43 Noise of .05

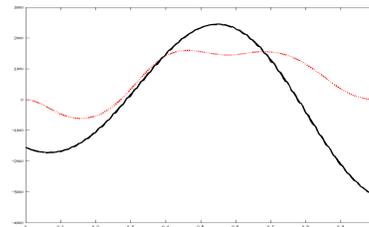
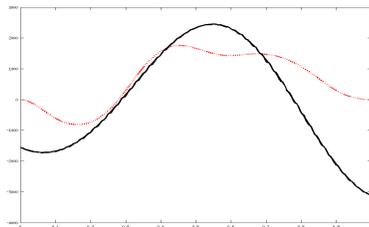


Figure 3.44 Noise of .07

Figure 3.45 Noise of .1

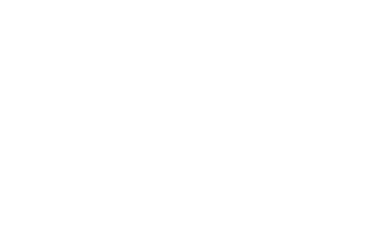


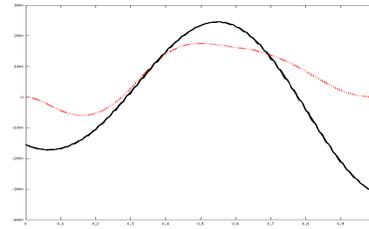
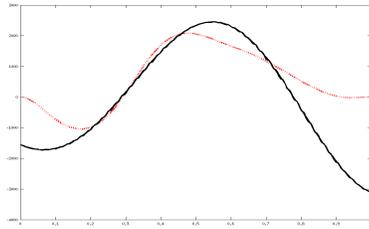
Figure 3.46 Reconstructions by Using β_M 

Figure 3.47 Noise of .025

Figure 3.48 Noise of .05

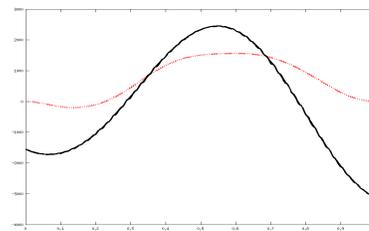
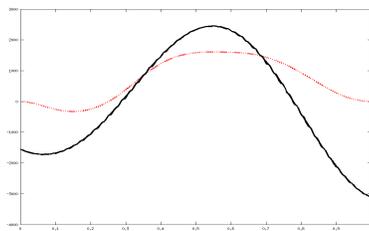
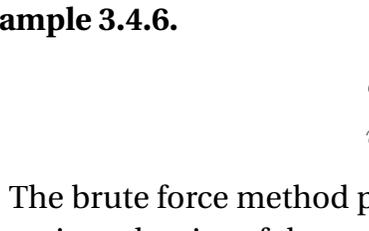


Figure 3.49 Noise of .07

Figure 3.50 Noise of .1

**Example 3.4.6.**

$$a(x) = (x^2 + 1)$$

$$u(x) = \cos(2\pi x) - 1.$$

The brute force method provides very good reconstructions of the source term given the size of the error term introduced. This accuracy was followed by the Morozov principle which generated reconstructions with errors only slightly larger than those produced by the brute force method. In this case the Morozov principle over-regularized the problem but this did not lead to significant reductions in the accuracy of the reconstructions. Damping with $\gamma = 1.2$ produced better results than the undamped method, but this benefit is miniscule.

The Chebyshev method is able to converge to β_M with three or four iterations and has a marked improvement over Newton's method. This superiority in performance is not replicated by the other cubically convergent methods which generally took slightly more time than Newton's method. Newton's method is able to converge to the solution in about one-fifth of the time of the quasi-Newton method .

Table 3.26 Brute Force Results

$\hat{\delta}$	β_{opt}	L^2 Error
0	1e-13	576.6874
0.01	9.1689e-10	939.2465
0.025	2.3267e-09	954.3663
0.05	4.3193e-09	963.0226
0.075	6.2154e-09	967.9941
0.1	8.067e-09	972.0326

Table 3.27 Newton-Type Results

$\hat{\delta}$	β_M	L^2 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.01	1.0696e-09	939.737	38	0.65415	6	0.052969	3	0.015346
0.025	2.8667e-09	956.4537	37	0.49931	6	0.033722	4	0.017222
0.05	6.8103e-09	976.0295	35	0.45531	9	0.054012	4	0.014652
0.075	1.2456e-08	998.9455	34	0.43026	7	0.040854	4	0.014924
0.1	2.0838e-08	1025.8118	34	0.64214	6	0.050635	5	0.028855

Table 3.28 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.01	3	0.016049	3	0.014506	3	0.014588
0.025	3	0.009632	4	0.01713	4	0.018167
0.05	3	0.009839	3	0.010883	4	0.017129
0.075	3	0.009726	3	0.011104	3	0.012062
0.1	4	0.021029	4	0.022653	5	0.030553

Table 3.29 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	1.1561e-10	9.7313e-10	1.0628e-09	1.0644e-09	1.0644e-09	1.0644e-09
0.025	4.9255e-10	2.6649e-09	2.8614e-09	2.8656e-09	2.8657e-09	2.8657e-09
0.05	1.3669e-09	6.2424e-09	6.7439e-09	6.7569e-09	6.7572e-09	6.7572e-09
0.075	2.5011e-09	1.1258e-08	1.2389e-08	1.2423e-08	1.2424e-08	1.2424e-08
0.1	3.9041e-09	1.8404e-08	2.0842e-08	2.0928e-08	2.093e-08	2.093e-08

Table 3.30 L^2 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
0.01	999.4081	939.3201	939.6974	939.7065	939.7067	939.7067
0.025	1098.3753	955.265	956.4174	956.4458	956.4464	956.4464
0.05	1101.4939	971.8765	975.5324	975.6296	975.6318	975.6318
0.075	1077.4348	991.9241	998.5574	998.7535	998.7588	998.7593
0.1	1050.2028	1015.8537	1025.8248	1026.1568	1026.1665	1026.1664

Figure 3.51 Reconstructions by Using β_{opt}

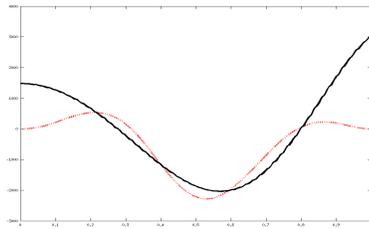


Figure 3.52 Noise of .025

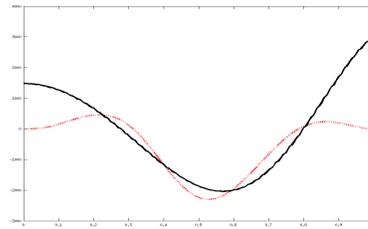


Figure 3.53 Noise of .05

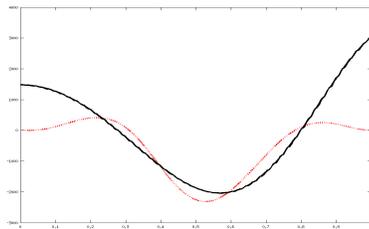


Figure 3.54 Noise of .07

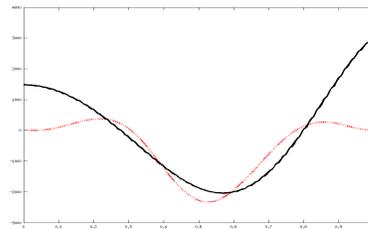


Figure 3.55 Noise of .1

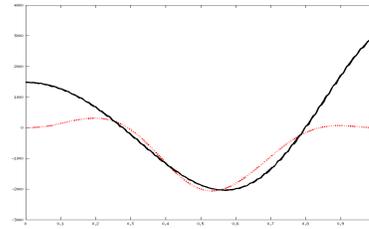
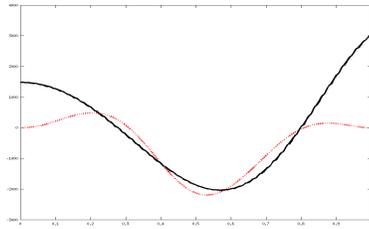
Figure 3.56 Reconstructions by Using β_M 

Figure 3.57 Noise of .025

Figure 3.58 Noise of .05

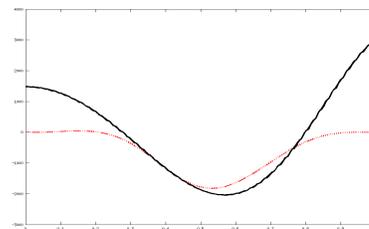
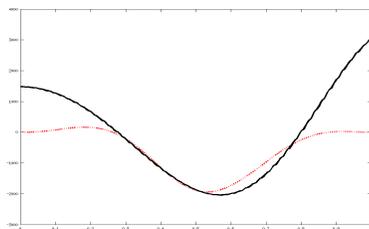
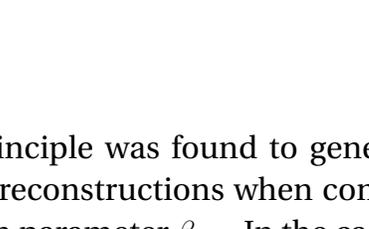
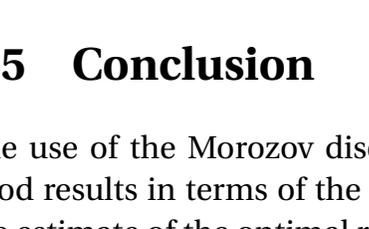


Figure 3.59 Noise of .07

Figure 3.60 Noise of .1



3.5 Conclusion

The use of the Morozov discrepancy principle was found to generate very good results in terms of the error in the reconstructions when compared to the estimate of the optimal regularization parameter β_{opt} . In the cases where the Morozov principle overestimated the size of the regularization parameter the damped Morozov principle was able to compensate by adjusting the estimated regularization parameter downwards. Newton's method was able to converge to β_M much more quickly than the quasi-Newton method. While the cubically convergent methods often converged more quickly than Newton's method, they did not consistently outperform Newton's method in terms of time.

Chapter 4

Morozov Principle for the Equation Error Approach

The general methodology introduced by Kunsich and Zou [25], and tested numerically in Chapter 2 for source identification problem will be extended in this chapter for the equation error formulation for the inverse problem of identifying variable parameters in general partial differential equations. The Morozov principle allows us to set up a root finding problem to approximate the regularization parameter minimizing. As before, we use finite element framework to discretize the equations. The bisection, Newton-type, and cubic methods are tested on a set of second-order elliptic problem and comparisons are made with the identification of the optimal parameter through brute force.

4.1 The Equation Error Functional

Let B be a Banach space and let A be a nonempty, closed, and convex subset of B . Let V be a Hilbert space which will be identified with its topological dual V^* in the usual manner. Let $T : B \times V \times V \rightarrow \mathbb{R}$ be a trilinear form with $T(a, u, v)$ symmetric in u, v . Let $m : V \rightarrow \mathbb{R}$ be a bounded linear functional. Assume that there are constants $\alpha > 0$ and $\beta > 0$ with

$$T(a, u, v) \leq \beta \|a\|_B \|u\|_V \|v\|_V, \quad \text{for all } u, v \in V, a \in B, \quad (4.1)$$

$$T(a, u, u) \geq \alpha \|u\|_V^2, \quad \text{for all } u \in V, a \in A. \quad (4.2)$$

We consider the following variational problem: Given $a \in A$, find $u = u(a) \in V$ such that

$$T(a, u, v) = m(v), \quad \text{for every } v \in V. \quad (4.3)$$

Due to the symmetry, continuity, and ellipticity of T , it follows from the Riesz representation theorem that for every $a \in A$, the variational problem (4.3) admits a unique solution $u(a)$. In this work, our interest is in the inverse problem of identifying the parameter a when a measurement z of the solution $u(a)$ is available. This inverse problem is often posed as an optimization problem and there are many optimization formulations that have been used. This includes, the output least-squares, the modified output least-squares, and the equation error approach, among others. In this work, our focus is on the equation error approach.

Given any pair $(a, w) \in A \times V$, we define $e(a, w) \in V$ through the variational equation:

$$\langle e(a, w), v \rangle_V = T(a, w, v) - m(v), \quad \text{for every } v \in V. \quad (4.4)$$

Using the identifiability of Hilbert spaces, we have

$$e(a, w) = T(a, w) - m,$$

where the elements are the corresponding Riesz elements. Clearly, for any direction $b \in H$, we get

$$D_a e(a, u)(b) = T(b, u).$$

Given $\varepsilon > 0$, we consider the optimization problem via the regularized equation error formulation:

$$\text{minimize } J(a, \varepsilon) := \frac{1}{2} \|e(a, z_\delta)\|_V^2 + \frac{\varepsilon}{2} \|a\|_H^2, \quad (4.5)$$

where H is a suitable Hilbert space, $\varepsilon > 0$ is the regularization parameter, and $z_\delta \in V$ is the noisy data with noise level $\delta > 0$, that is,

$$\|z_0 - z_\delta\| \leq \delta.$$

By assuming that H is compactly embedded into B , an existence result can be given. This assumption holds, for example for the choices $H := H_2(\Omega)$

and $B = L^\infty(\Omega)$. However, we have seen that existence results can be obtained without the compact embedding assumption. This has been shown for the choices $H = H^1(\Omega)$ and $a \in L^2(\Omega)$ but at the expense of higher smoothness requirement on the data. Therefore, without distinguishing between these two case for the existence, we simply assume that (4.5) is solvable. Clearly, due to the strongly convex regularizer, such a solution is unique.

The following result gives an optimality condition for the above problem:

Theorem 4.1.1. *For every $\varepsilon > 0$, the solution $a(\varepsilon)$ of the optimization problem (4.5) is unique and can be characterized as the unique solution of the following variational equation: Find $a \in H$ such that*

$$\langle e(a, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle a, b \rangle_H = 0, \quad \text{for every } b \in H, \quad (4.6)$$

or, equivalently,

$$\langle T(a, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle a, b \rangle_H = \langle m, T(b, z_\delta) \rangle_V, \quad \text{for every } b \in H. \quad (4.7)$$

Proof. For any direction $b \in H$, the solution $a(\varepsilon)$ satisfies the following optimality condition:

$$D_a J(a, \varepsilon)(b) = 0.$$

Since $e(\cdot, z)$ is affine, we have

$$D_a J(a, \varepsilon)(b) = \langle e(a, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle a, b \rangle_H,$$

and the proof follows at once. \square

The following result, where the regularized solution $a(\varepsilon)$ is seen as the function of the regularization parameter ε , embarks on the differentiability of the regularized solutions:

Theorem 4.1.2. *For every $\varepsilon > 0$, the map $\varepsilon \rightarrow a(\varepsilon)$, where $a(\varepsilon)$ solves the variational equation (4.6) (or equivalently (4.5)) is differentiable. Moreover, the derivative $a'(\varepsilon)$ is characterized as the unique solution of the following variational equation: Find $w \in H$ such that*

$$\langle T(w, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle w, b \rangle_H = - \langle a(\varepsilon), b \rangle_H \quad \text{for every } b \in H. \quad (4.8)$$

Proof. Let $\varepsilon > 0$ be fixed. Let w be the solution of variational equation (4.8). For any $b \in H$, we have

$$\begin{aligned}\langle e(a(\varepsilon + t), z_\delta), T(b, z_\delta) \rangle_V + (\varepsilon + t) \langle a(\varepsilon + t), b \rangle_H &= 0, \\ \langle e(a(\varepsilon), z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle a(\varepsilon), b \rangle_H &= 0.\end{aligned}$$

A simple rearrangement of the above two equations implies that

$$\left\langle \frac{e(a(\varepsilon + t), z_\delta) - e(a(\varepsilon), z_\delta)}{t}, T(b, z_\delta) \right\rangle_V + \varepsilon \left\langle \frac{a(\varepsilon + t) - a(\varepsilon)}{t}, b \right\rangle_H + \langle a(\varepsilon + t), b \rangle_H = 0,$$

or equivalently

$$\langle T(\delta_t, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle \delta_t, b \rangle_H + \langle a(\varepsilon + t) - a(\varepsilon), b \rangle_H = 0,$$

where $\delta_t := t^{-1}(a(\varepsilon + t) - a(\varepsilon))$. By subtracting (4.8) from the above identity, we deduce

$$\langle T(\delta_t - w, z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle \delta_t - w, b \rangle_H + \langle a(\varepsilon + t) - a(\varepsilon), b \rangle_H = 0.$$

By setting $b = \delta_t - w$, in the above, we get

$$\begin{aligned}\|\delta_t - w\|_H^2 &= -\langle a(\varepsilon + t) - a(\varepsilon), \delta_t - w \rangle_H - \langle T(\delta_t - w, z_\delta), T(\delta_t - w, z_\delta) \rangle_V \\ &= -\langle a(\varepsilon + t) - a(\varepsilon), \delta_t - w \rangle_H - \|T(\delta_t - w, z_\delta)\|_V^2 \\ &\leq -\langle a(\varepsilon + t) - a(\varepsilon), \delta_t - w \rangle_H.\end{aligned}$$

By Cauchy-Schwarz inequality, we get $\|\delta_t - w\|_H^2 \leq \|a(\varepsilon + t) - a(\varepsilon)\|_H \|\delta_t - w\|_H$, and hence

$$\|\delta_t - w\|_H \leq \|a(\varepsilon + t) - a(\varepsilon)\|_H,$$

and by taking limits $t \rightarrow 0^+$, we get $\|\delta_t - w\|_H \rightarrow 0$. This proves that a is differentiable at ε and $a'(\varepsilon) = w$. The proof is complete. \square

Let $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ be the value function defined by

$$\Phi(\varepsilon) := J(a(\varepsilon), \varepsilon) := \frac{1}{2} \|e(a(\varepsilon), z_\delta)\|_V^2 + \frac{\varepsilon}{2} \|a(\varepsilon)\|_H^2. \quad (4.9)$$

We have the following result concerning the smoothness of the map Φ :

Theorem 4.1.3. *The map Φ is twice differentiable for every $\varepsilon > 0$, with derivatives given by*

$$\Phi'(\varepsilon) = \frac{1}{2} \|a(\varepsilon)\|_H^2, \quad (4.10)$$

$$\Phi''(\varepsilon) = \langle a(\varepsilon), a'(\varepsilon) \rangle_H. \quad (4.11)$$

Moreover, if $T(\cdot, z_\delta, m) \neq 0_{H^*}$, then $\Phi(\varepsilon)$ is strictly increasing and strictly concave.

Proof. The differentiability of Φ follows from the differentiability of $a(\cdot)$. Moreover, we have

$$\Phi'(\varepsilon) = \langle e(a'(\varepsilon), z_\delta), e(a(\varepsilon), z_\delta) \rangle_V + \varepsilon \langle a'(\varepsilon), a(\varepsilon) \rangle_H + \frac{1}{2} \|a(\varepsilon)\|_H^2 = \frac{1}{2} \|a(\varepsilon)\|_H^2,$$

where we used (4.6) with $b = a(\varepsilon)$. The proof of (4.11) is then immediate. Finally, taking $b = a'(\varepsilon)$ in (4.8), for every $\varepsilon > 0$, we have

$$\begin{aligned} \Phi''(\varepsilon) &= \langle a(\varepsilon), a'(\varepsilon) \rangle_H \\ &= - \langle e(a'(\varepsilon), z_\delta), e(a'(\varepsilon), z_\delta) \rangle_V - \varepsilon \langle a'(\varepsilon), a'(\varepsilon) \rangle_H \\ &= - \|e(a'(\varepsilon), z_\delta)\|_V^2 - \varepsilon \|a'(\varepsilon)\|_H^2 \\ &\leq 0. \end{aligned}$$

Furthermore, $\Phi''(\varepsilon) < 0$, for every $\varepsilon > 0$, on the contrary this would imply $a'(\bar{\varepsilon}) = 0$ for some $\bar{\varepsilon} > 0$. By (4.8) $a(\bar{\varepsilon}) = 0$, and by (4.12)

$$\langle m, T(b, z_\delta) \rangle_V = 0, \quad \text{for every } b \in H. \quad (4.12)$$

equivalently $T(b, z_\delta, m) = 0$, for every $b \in H$, and a contradiction completes the proof. \square

4.2 Morozov Principle

The Morozov principle suggests solving the following nonlinear equation: Given δ find ε such that

$$\|e(a(\varepsilon), z_\delta)\|_V^2 = \delta^2. \quad (4.13)$$

Evidently, the above equation can equivalently be posed as the problem of finding ε such that

$$\Phi(\varepsilon) - \varepsilon \Phi'(\varepsilon) = \frac{1}{2} \delta^2. \quad (4.14)$$

The Damped Morozov's principle suggest solving the following equation: Given δ find ε such that

$$\|e(a(\varepsilon), z_\delta)\|_V^2 + \varepsilon^\gamma \|a(\varepsilon)\|_H^2 = \delta^2 \quad (4.15)$$

where $\gamma \in [1, \infty]$. Equivalently, find ε such that

$$\Phi(\varepsilon) + (\varepsilon^\gamma - \varepsilon) \Phi'(\varepsilon) = \frac{1}{2} \delta^2. \quad (4.16)$$

Clearly, (4.13) corresponds to the case $\gamma = \infty$.

We define

$$\Phi(0) := \inf_{a \in H^2} \frac{1}{2} \|e(a, z^\delta)\|_V^2 = \inf_{a \in H^2} \frac{1}{2} \|T(a, z_\delta) - m\|_V^2.$$

We have the following result:

Theorem 4.2.1. *If $\Phi(0) < \frac{1}{2} \delta^2 < \Phi(1)$, then (4.16) has a unique solution $\beta^* \in (0, 1]$.*

4.3 Discretization and Implementation Details

We need to discretize variational equations (4.6) and (4.8) which give us $a(\varepsilon)$ and $a'(\varepsilon)$. In this section, we describe the finite element framework that will be used for discretization.

Let \mathcal{T}_h be a triangulation of the domain Ω . We define \mathcal{A}_h to be the space of all continuous piecewise polynomials of degree d_a relative to \mathcal{T}_h . Similarly, \mathcal{V}_h will be the space of all continuous piecewise polynomials of degree d_u relative to \mathcal{T}_h , subject to the constraint that the Dirichlet boundary conditions are satisfied.

Bases for \mathcal{A}_h and \mathcal{V}_h will be represented by $\{\psi_1, \psi_2, \dots, \psi_m\}$ and $\{\varphi_1, \varphi_2, \dots, \varphi_n\}$, respectively. The space \mathcal{A}_h is then isomorphic to \mathbb{R}^m , and for any $a \in \mathcal{A}_h$, we define $A \in \mathbb{R}^m$ by $A_i = a(x_i)$, $i = 1, 2, \dots, m$, where $\{\psi_1, \psi_2, \dots, \psi_m\}$ is a nodal basis corresponding to the nodes $\{x_1, x_2, \dots, x_m\}$. Conversely, each $A \in \mathbb{R}^m$ corresponds to $a \in \mathcal{A}_h$ defined by $a = \sum_{i=1}^m A_i \psi_i$. Similarly, $u \in \mathcal{V}_h$ will correspond to $U \in \mathbb{R}^n$, where $U_i = u(y_i)$, $i = 1, 2, \dots, n$ and $u = \sum_{i=1}^n U_i \varphi_i$. Here y_1, y_2, \dots, y_n are the nodes of the mesh defining \mathcal{V}_h . Note that although both \mathcal{A}_h and \mathcal{V}_h are defined relative to the same triangles, the nodes are different.

By $M_V, K_V \in \mathbb{R}^{n \times n}$ we denote the mass and the stiffness matrix with respect to V_h , that is,

$$\begin{aligned}(M_V)_{i,j} &= \int_{\Omega} \varphi_i \varphi_j dx, \\ (K_V)_{i,j} &= \int_{\Omega} \nabla \varphi_i \nabla \varphi_j dx.\end{aligned}$$

Similarly, $M_H, K_H \in \mathbb{R}^{m \times m}$ denote the mass and the stiffness matrix with respect to A_h , that is,

$$\begin{aligned}(M_H)_{i,j} &= \int_{\Omega} \psi_i \psi_j dx, \\ (K_H)_{i,j} &= \int_{\Omega} \nabla \psi_i \nabla \psi_j dx.\end{aligned}$$

Furthermore, for every $A \in \mathbb{R}^m$, $K(A) \in \mathbb{R}^{n \times n}$ is the matrix defined by

$$K(A)_{i,j} = T(a, \varphi_i, \varphi_j),$$

and $F \in \mathbb{R}^m$ is the vector defined by

$$F_i = m(\varphi_i).$$

We will also use the the adjoint matrix $L(\cdot) \in \mathbb{R}^{n \times m}$ defined by

$$L(V)A = K(A)V, \text{ for every } A \in \mathbb{R}^m, V \in \mathbb{R}^n.$$

We now proceed to discretize $a(\varepsilon)$ and $\Phi(\varepsilon)$. For any $a \in A$, $z \in Z$, the discretization of the Riesz element $T(a, z)$ is given by the vector $T(A, Z) \in \mathbb{R}^n$ verifying

$$K_V T(A, Z) = K(A)Z,$$

and consequently

$$T(A, Z) = K_V^{-1} L(Z)A.$$

In the same way the discretization of $e(a, z)$ is given by

$$E(A, Z) = K_V^{-1} (L(Z)A - F).$$

Applying the discretization scheme to the variational equation

$$\langle T(a(\varepsilon), z_\delta), T(b, z_\delta) \rangle_V + \varepsilon \langle a, b \rangle_H = \langle m, T(v, z_\delta) \rangle_V \text{ for every } b \in H,$$

we get

$$\begin{aligned} & \widehat{B}^T L(Z\delta)^T K_V^{-1} (K_V + M_V) K_V^{-1} L(Z\delta) a_h[\varepsilon] + \varepsilon \widehat{B} (K_H + M_H) a_h[\varepsilon] \\ & = \widehat{B}^T L(Z\delta)^T K^{-1} (M + K) F. \end{aligned}$$

where \widehat{B} corresponds to the arbitrary b and by $a_h[\varepsilon] \in \mathbb{R}^m$ we denote the discretization of $a(\varepsilon)$.

This last expression is equivalent to solving the linear system: Find $a_h[\varepsilon] \in \mathbb{R}^m$ such that

$$\begin{aligned} & [L(Z\delta)^T K_V^{-1} (K_V + M_V) K_V^{-1} L(Z\delta) + \varepsilon (K_H + M_H)] a(\varepsilon) \\ & = L(Z\delta)^T K_V^{-1} (K_V + M_V) F. \end{aligned} \quad (4.17)$$

In the same way, the derivative $a'(\varepsilon)$ corresponds to solving the following variational equation

$$\langle T(a'(\varepsilon), z_\delta), T(v, z_\delta) \rangle_V + \varepsilon \langle a'(\varepsilon), v \rangle_H = - \langle a(\varepsilon), v \rangle_H \text{ for every } \delta a \in V.$$

Following the same ideas as before, its discrete version is given by: Find $Da_h[\varepsilon] \in \mathbb{R}^m$ such that

$$\begin{aligned} & [L(Z\delta)^T K_V^{-1} (K_V + M_V) K_V^{-1} L(Z\delta) + \varepsilon (K_H + M_H)] Da_h[\varepsilon] \\ & = - (K_H + M_H) a_h[\varepsilon]. \end{aligned} \quad (4.18)$$

Recall that the value function $\Phi : \mathbb{R}_+ \rightarrow \mathbb{R}$ is defined by

$$\begin{aligned} \Phi_h(\varepsilon) & := J(a_h(\varepsilon), \varepsilon) = \frac{1}{2} \|e(a_h(\varepsilon), Z_\delta)\|_V^2 \\ & + \frac{\varepsilon}{2} \|a_h(\varepsilon)\|_H^2 \end{aligned}$$

and consequently, its discrete analogue reads:

$$\Phi_h(\varepsilon) = \frac{1}{2} (L(Z\delta) a_h(\varepsilon) - F)^T K_V^{-1} (K_V + M_V) K_V^{-1} (L(Z\delta) a_h(\varepsilon) - F) + \frac{\varepsilon}{2} a_h(\varepsilon)^T (K_H + M_H) a_h(\varepsilon). \quad (4.19)$$

Using (4.10), its discrete derivative is given by

$$D\Phi_h(\varepsilon) = \frac{1}{2} \|a_h(\varepsilon)\|_H^2 = \frac{1}{2} a_h(\varepsilon)^T (K_H + M_H) a_h(\varepsilon), \quad (4.20)$$

and the second derivative by

$$D^2\Phi_h(\varepsilon) = \langle a_h(\varepsilon), Da_h(\varepsilon) \rangle_H = a_h(\varepsilon)^T (K_H + M_H) Da_h(\varepsilon). \quad (4.21)$$

Recall that the Morozov principle corresponds to solving the following nonlinear scalar equation

$$\text{Find } \varepsilon \in (0, 1) \text{ such that } G^\gamma(\varepsilon) := \Phi(\varepsilon) - (\varepsilon^\gamma - \varepsilon)\Phi'(\varepsilon) - \frac{1}{2}\delta^2.$$

The natural discretization of this problem is given by

$$\text{Find } \varepsilon \in (0, 1) \text{ such that } G_h^\gamma(\varepsilon) := \Phi_h(\varepsilon) - (\varepsilon^\gamma - \varepsilon)\Phi_h'(\varepsilon) - \frac{1}{2}\delta^2.$$

To apply the known method (Bisection, Newton, QuasiNewton) we only need to evaluate $G_h^\gamma(\varepsilon)$ and the derivative $DG_h^\gamma(\varepsilon)$ for the methods where this derivative is used:

$$DG^\gamma(\varepsilon) = D\Phi_h(\varepsilon) - (\varepsilon^\gamma - \varepsilon)D^2\Phi_h(\varepsilon). \quad (4.22)$$

4.4 Data Smoothing

To improve the efficiency of the parameter identification process, it is natural to perform some data smoothing before using the data as an input for the optimization formulation of the inverse problem. In this chapter, we obtain a set of smooth data z_s by solving the following optimization problem

$$\min_{z_s} \|z_s - z_\delta\|_{H^1}^2 + \alpha \|\nabla z_s\|_{L^2}^2, \quad (4.23)$$

where α is the smoothing parameter. This strategy for error smoothing was explored in [20] where smoothing was shown to yield increased accuracy in the coefficient reproductions.

By standard arguments, we obtain the following finite dimensional version optimality condition for 4.23:

$$(K + M + \alpha K)Z_s = (K + M)Z_\delta. \quad (4.24)$$

4.5 Numerical Experiments

A series of numerical tests concentrating on second order elliptic partial differential equations are considered. These examples were interpolated using

one-hundred basis functions. The brute force method is used in a manner similar to the chapter where the source term is considered. The Morozov principle is also compared using the same tolerances previously discussed.

The problem for coefficient identification is a nonlinear problem. The result of this is that adding noise will have more of an effect on the results for this problem than in the source term problem. V levels of noise were used for these problems and limit on the size of the added noise was determined by the degree to which the reconstruction deviated from the actual solution.

Error smoothing was used on second order examples two and three to demonstrate how the coefficient reconstruction could be improved by compensating for the added noise. The best smoothing parameter α_{opt} was identified by successively identifying Z_s for a range of smoothing parameters while calculating $\|z - z_s\|_{H^1}$ for each. After the best version of the smoothed data was identified, the best regularization parameter was identified for both the smoothed and unsmoothed data. These are known as β_{opt}^s and β_{opt}^δ respectively. Finally, the Morozov principle was used to estimate the best regularization parameter for both cases which are labeled as β_M^s and β_M^δ . The L^2, H^1 semi-norm, and H^1 errors were calculated and compared to assess the performance of the parameter identification. The results for the coefficient reconstruction are also compared.

In general, the Morozov principle was able to find comparable results when compared to the brute force method. Where the undamped version of the Morozov equation fell short, the damped version was generally able to provide a more satisfactory result. Furthermore, the Morozov principle became very accurate when smoothing was applied to the data. Several Newton-type and cubically convergent algorithms were implemented. In these tests it was found that Newton's method converged to β_M in fractions of a second. The cubic methods often shortened the time required to do this, but not in a significant manner.

4.5.1 Second Order Examples

Example 4.5.1. Consider the second order problem in one-dimension where:

$$\begin{aligned} a(x) &= \exp(1 + x^2) \\ u(x) &= \exp(-x)\sin(\pi x). \end{aligned}$$

The brute force method was employed in the same manner as the source term section. Since this is a nonlinear problem employing the equation error functional, the regularization parameters are much larger, and the effect adding noise has on the reproductions is much more noticeable. It can be seen by the brute force reproductions below, the equation error functional works well for small noise, but when $\hat{\delta} = .05$ the noise introduces noticeable inaccuracy into the reconstruction.

For this example, the Morozov equation produces over-regularized values. This same problem was observed in [25] where in some cases β_M was of a much larger magnitude than β_{opt} . While the reproductions start off as fairly accurate, largest noise parameter gives a significant deviation from the brute force method. The damped version of the Morozov principle provides an alternative which can reduce the estimation to a similar magnitude to the brute force parameter. It can be seen that a damping parameter of $\gamma = 1$ generates a set of parameters with much lower error values for all noise levels but the smallest where damping actually under-regularizes the problem. The larger noise levels contain noticeable levels of flattening when the Morozov principle is used.

The cubically convergent methods took the fewest iterations and the least amount of time to converge to a satisfactory estimation of β_M . Despite this, the decrease in time for the cubic methods over Newton's method was not significant. In some cases the cubic methods either took about as long, or slightly longer due to the larger number of necessary calculations.

Table 4.1 Brute Force Results

$\hat{\delta}$	β_{opt}	H^1 Error
0.0001	0.00020965	0.30564
0.001	0.0028665	0.51348
0.005	0.010663	0.70489
0.0075	0.014764	0.80309
0.01	0.019451	0.91302

Table 4.2 Newton-Type Results

$\hat{\delta}$	β_M	H^1 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.0001	0.0006029	0.33704	32	0.11487	39	0.054092	8	0.010742
0.001	0.01319	0.68537	35	0.15354	10	0.013266	6	0.007805
0.005	0.20813	1.2514	36	0.11213	10	0.008747	5	0.006001
0.0075	0.58111	1.4835	29	0.085677	11	0.01401	7	0.006122
0.01	1.8623	1.6794	34	0.081709	12	0.017727	8	0.007819

Table 4.3 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.0001	7	0.014318	7	0.014385	7	0.013638
0.001	5	0.009088	5	0.009986	5	0.009116
0.005	5	0.009274	5	0.010092	5	0.009162
0.0075	5	0.006162	5	0.01036	5	0.009832
0.01	6	0.008257	6	0.012774	6	0.008989
$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.0001	1.9175e-05	0.00011024	0.00030784	0.00048997	0.00057152	0.00059532
.001	0.0014809	0.0040079	0.0070892	0.0096992	0.011412	0.012356
.005	0.05155	0.081123	0.10677	0.12787	0.14494	0.15864
.0075	0.16589	0.21911	0.26175	0.29671	0.32601	0.351
.01	0.46161	0.52277	0.56926	0.6063	0.63674	0.66235

Figure 4.1 H^1 Error for Damped Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.0001	0.41949	0.31655	0.30983	0.32605	0.334	0.33631
.001	0.57325	0.52565	0.58724	0.63417	0.66074	0.67415
.005	0.94859	1.0443	1.1035	1.1429	1.1705	1.1906
.0075	1.213	1.2732	1.3121	1.3396	1.3602	1.3763
.01	1.4438	1.469	1.4859	1.4983	1.5078	1.5154

Figure 4.2 Reconstructions Using β_{opt}

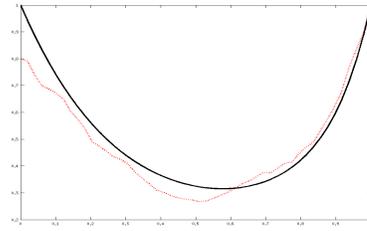
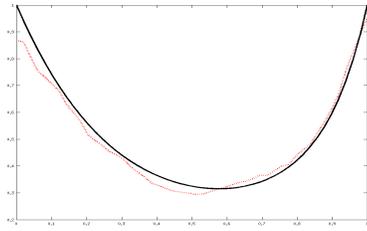


Figure 4.3 Noise of .005

Figure 4.4 Noise of .0075

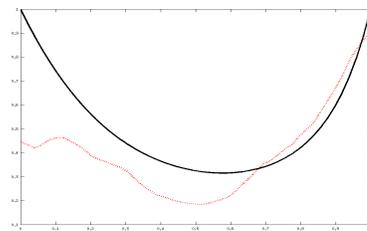
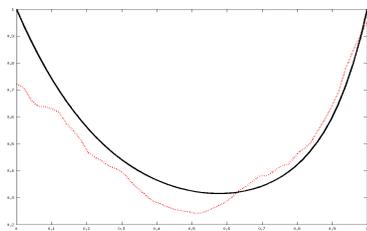


Figure 4.5 Noise of .01

Figure 4.6 Noise of .05

Figure 4.7 Reconstructions Using β_M

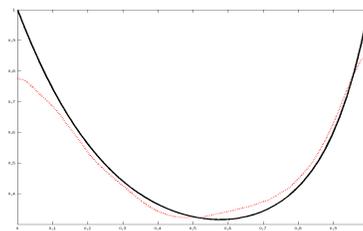
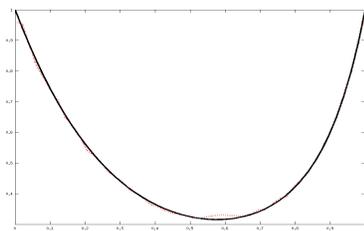


Figure 4.8 Noise of .001

Figure 4.9 Noise of .005

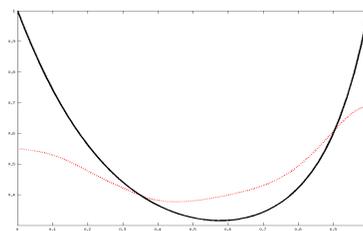
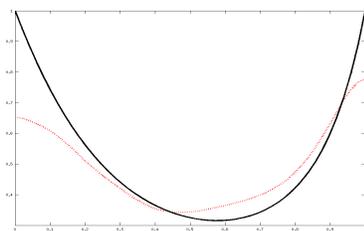


Figure 4.10 Noise of .0075

Figure 4.11 Noise of .01

Example 4.5.2.

$$a(x) = .5\cos(4\pi x) + 1,$$

$$u(x) = \sin(4\pi x).$$

This second example yields much more favorable results in terms of the error levels. Five percent noise causes a noticeable malformation of the coefficient reconstruction, but a recognizable reconstruction is found for all of the noise levels. The Morozov principle is able to identify a parameter preserving the which produces a reconstruction which is similar to. In example one, β_M was on separate orders of magnitude from β_{opt} . In this case, β_M was close enough to β_{opt} that the reconstructions were nearly identical to those produced using the brute force method. Even in the case of the largest noise level β_M is only twice as large as β_{opt} which is a very small deviation when compared to other tests in this paper and [25]. The damped results were also comparable, but no consistent improvement over the undamped method was produced using damping.

Newton's method converged to β_M much more quickly than the quasi-Newton method. The cubically convergent methods were able to outperform the Newton type methods in several examples, but this was a marginal improvement in terms of convergence speed.

Table 4.4 Brute Force Results

$\hat{\delta}$	β_{opt}	H^1 Error
0	2.0004e-08	0.097791
0.001	0.0018536	0.44649
0.005	0.012174	0.8215
0.0075	0.017902	0.94
0.01	0.023372	1.0373
0.05	0.1151	2.068

Table 4.5 Newton-Type Results

$\hat{\delta}$	β_M	H^1 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
1e-14	3.0518e-05	0.11066	15	0.074762	12	0.026487	11	0.026087
0.001	0.0029202	0.54435	19	0.078098	10	0.019319	8	0.013634
0.005	0.014763	0.96871	17	0.078988	9	0.01515	6	0.009948
0.0075	0.02214	1.1187	20	0.092362	9	0.013203	6	0.008873
0.01	0.029608	1.2416	19	0.093537	8	0.012256	5	0.01173
0.05	0.21437	2.4953	22	0.099909	12	0.022766	6	0.009903

Table 4.6 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal, Time
1e-14	9	0.030169	8	0.031503	8	0.023043
0.001	6	0.01433	6	0.015527	6	0.015517
0.005	5	0.011845	5	0.012469	5	0.012899
0.0075	5	0.010763	5	0.010421	5	0.019068
0.01	5	0.012122	5	0.012696	5	0.01302
0.05	4	0.009317	5	0.02096	5	0.012128

Table 4.7 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.0001	1.9175e-05	0.00011024	0.00030784	0.00048997	0.00057152	0.00059532
.001	0.0014809	0.0040079	0.0070892	0.0096992	0.011412	0.012356
.005	0.05155	0.081123	0.10677	0.12787	0.14494	0.15864
.0075	0.16589	0.21911	0.26175	0.29671	0.32601	0.351
.01	0.46161	0.52277	0.56926	0.6063	0.63674	0.66235

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.0001	0.41949	0.31655	0.30983	0.32605	0.334	0.33631
.001	0.57325	0.52565	0.58724	0.63417	0.66074	0.67415
.005	0.94859	1.0443	1.1035	1.1429	1.1705	1.1906
.0075	1.213	1.2732	1.3121	1.3396	1.3602	1.3763
.01	1.4438	1.469	1.4859	1.4983	1.5078	1.5154

Figure 4.12 Reconstructions Using β_{opt}

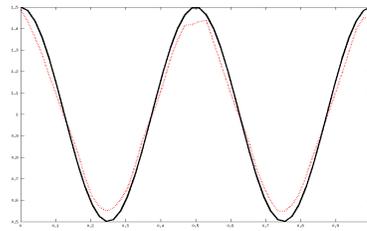
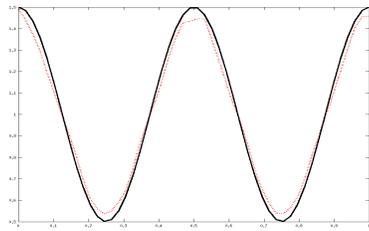


Figure 4.13 Noise of .005

Figure 4.14 Noise of .0075

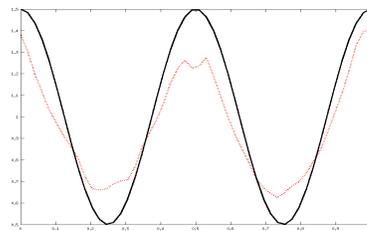
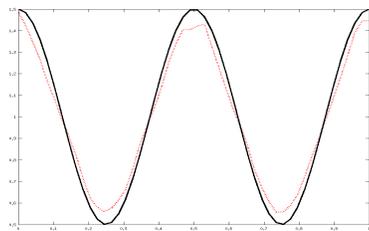


Figure 4.15 Noise of .01

Figure 4.16 Noise of .05

Figure 4.17 Reconstructions Using β_M

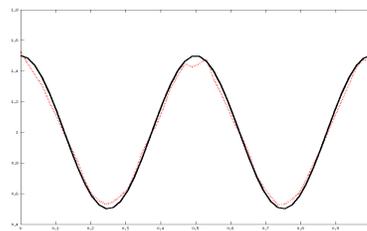
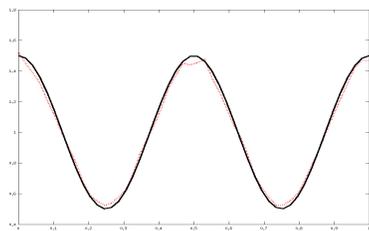


Figure 4.18 Noise of .005

Figure 4.19 Noise of .0075

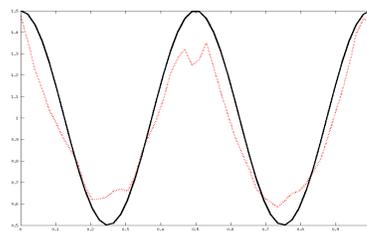
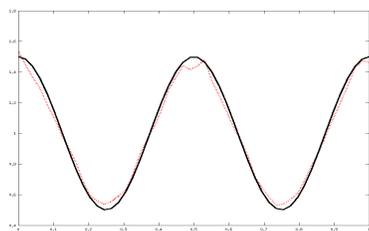


Figure 4.20 Noise of .01

Figure 4.21 Noise of .05

In this instance smoothing had a very beneficial impact on the calculation of both the optimal regularization parameter, and the Morozov parameter. A cursory glance on the results from the smooth and unsmoothed data shows that smoothing produces a noticeable reduction in the H^1 error produced by comparing the noisy data to Z . This comes at the expense of a large increase in the L^2 error. This latter impact is evident in the graphs on the next page. The red line which represents the noisy data without smoothing exhibits sharp gyrations. The black line denoting the smoothed data is flattened by with a reduction in the sudden changes introduced by noise.

Using the brute force method shows that smoothing produces a noticeable effect on the quality of the reconstructions of the source term. First, the H^1 error is reduced. Additionally, the L^2 error is also reduced meaning that not only will the reconstructions using smoothed data exhibit a smoother general shape, but that the flattening noted for the smoothed data will not be repeated. In the graphs of these reconstructions it can be seen that the blue line representing $a(\beta_{opt}^s)$ resembles a (black) much more than $a(\beta_M^s)$ (red).

In terms of the quality of the reconstructions, it can be seen that those produced with smoothing and the Morozov principle are much higher quality. The estimates of the regularization parameter are much closer to the optimal parameter and the overcompensation that occurs with large amounts of noise is also curbed. With the largest level of added noise β_M^s is 0.25758 which is comparable to $\beta_{opt}^s = 0.21088$. On the other hand, $\beta_M^\delta = 0.66454$ while the best parameter found with noise (β_{opt}^δ) is much smaller at 0.23952. As a result, for this noise level the level of H^1 error for the Morozov principle on smoothed data is about three-quarters of the result if no processing of the data is utilized.

Table 4.8 Error from Noisy Data

Noise	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0.0045342	0.30283	0.30287
0.02	0.0090684	0.60566	0.60573
0.03	0.013603	0.90849	0.9086
0.04	0.018137	1.2113	1.2115
0.05	0.022671	1.5142	1.5143
0.06	0.027205	1.817	1.8172
0.07	0.031739	2.1198	2.1201
0.08	0.036274	2.4226	2.4229
0.09	0.040808	2.7255	2.7258
0.1	0.045342	3.0283	3.0287
0.11	0.049876	3.3311	3.3315
0.12	0.05441	3.634	3.6344

Table 4.9 Error from Smoothing Data

Noise	α	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0	0.0045342	0.30283	0.30287
0.02	0	0.0090684	0.60566	0.60573
0.03	0.005005	0.01471	0.90688	0.907
0.04	0.015015	0.022558	1.206	1.2062
0.05	0.02002	0.028613	1.5017	1.502
0.06	0.035035	0.039104	1.793	1.7934
0.07	0.05005	0.049748	2.0793	2.0799
0.08	0.065065	0.060294	2.3595	2.3603
0.09	0.085085	0.073101	2.6331	2.6341
0.1	0.10511	0.085609	2.8993	2.9005
0.11	0.13013	0.10012	3.1577	3.1593
0.12	0.15516	0.11414	3.4078	3.4097

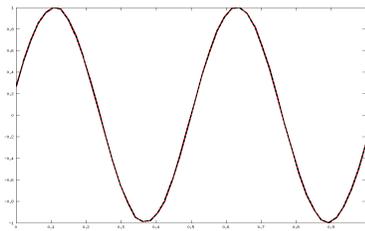


Figure 4.22 Noise of .02

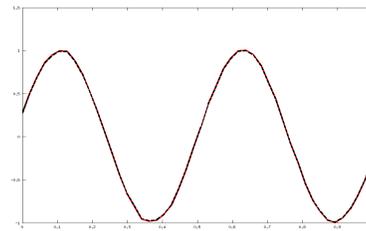


Figure 4.23 Noise of .04

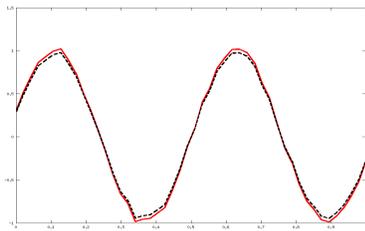


Figure 4.24 Noise of .06

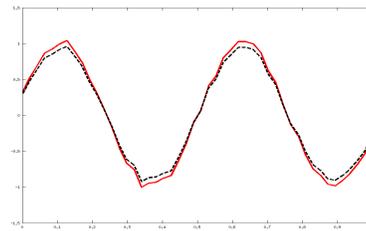


Figure 4.25 Noise of .08

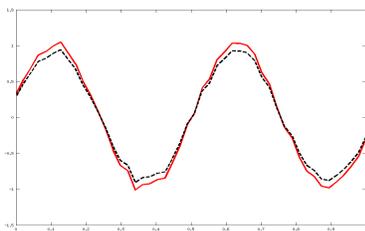


Figure 4.26 Noise of .1

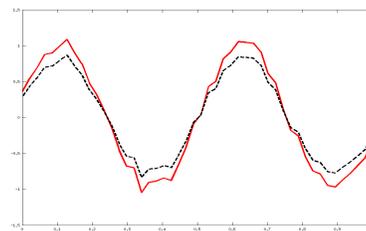


Figure 4.27 Noise of .12

Table 4.10 Coefficient Reconstructions with Noisy Data

Noise	β_{opt}^δ	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0.017413	0.045041	1.1067	1.1076
0.02	0.041529	0.071009	1.4458	1.4475
0.03	0.063285	0.088959	1.6591	1.6615
0.04	0.083438	0.10357	1.8216	1.8245
0.05	0.10279	0.11688	1.9566	1.9601
0.06	0.12182	0.12989	2.0744	2.0784
0.07	0.14083	0.1432	2.1803	2.185
0.08	0.15999	0.15712	2.2773	2.2828
0.09	0.1794	0.17182	2.3674	2.3736
0.1	0.19912	0.18739	2.4516	2.4587
0.11	0.21916	0.20381	2.5307	2.5389
0.12	0.23952	0.22106	2.6053	2.6146

Table 4.11 Coefficient Reconstructions with Smoothed Data

Noise	β_{opt}^s	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0.017413	0.045041	1.1067	1.1076
0.02	0.041529	0.071009	1.4458	1.4475
0.03	0.063279	0.087928	1.6536	1.656
0.04	0.083072	0.10056	1.8055	1.8083
0.05	0.10183	0.11215	1.9358	1.939
0.06	0.11938	0.12143	2.0397	2.0433
0.07	0.13628	0.13025	2.1332	2.1372
0.08	0.15273	0.1387	2.2193	2.2236
0.09	0.16826	0.14646	2.2956	2.3003
0.1	0.18336	0.15402	2.3676	2.3726
0.11	0.19736	0.16107	2.4324	2.4378
0.12	0.21088	0.16795	2.4944	2.5

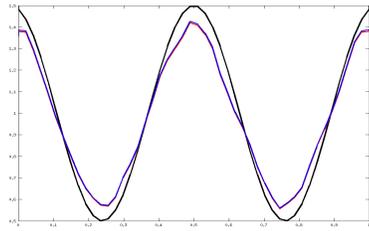


Figure 4.28 Noise of .02

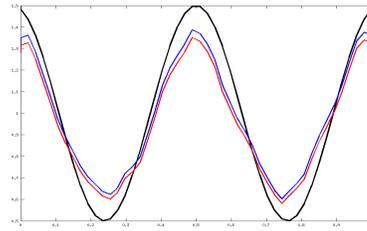


Figure 4.29 Noise of .04

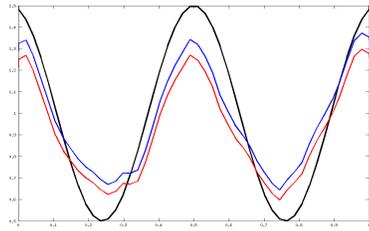


Figure 4.30 Noise of .06

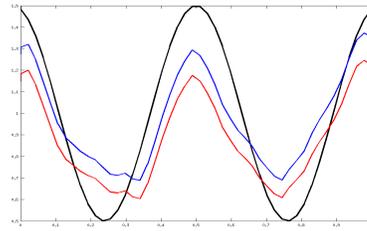


Figure 4.31 Noise of .08

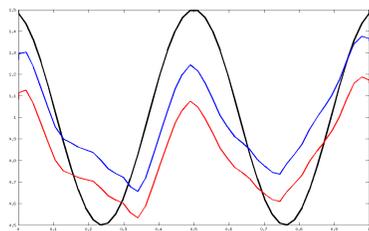


Figure 4.32 Noise of .1

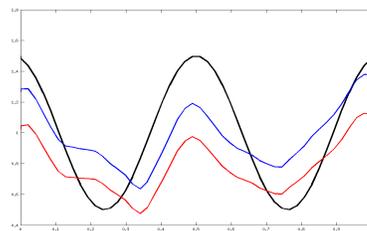


Figure 4.33 Noise of .12

Table 4.12 Morozov Results with Noisy Data

Noise	β_M^δ	L^2 Error	H^1S Error	H^1 Error
0.01	0.022645	0.051818	1.119	1.1202
0.02	0.045969	0.075161	1.4491	1.4511
0.03	0.069723	0.09387	1.6633	1.666
0.04	0.095096	0.11107	1.831	1.8343
0.05	0.12339	0.12814	1.9772	1.9814
0.06	0.15617	0.14586	2.1152	2.1202
0.07	0.19547	0.16472	2.2533	2.2593
0.08	0.24413	0.18503	2.398	2.4051
0.09	0.30631	0.20699	2.554	2.5624
0.1	0.38851	0.23075	2.7254	2.7352
0.11	0.50147	0.25639	2.9149	2.9262
0.12	0.66454	0.28398	3.1241	3.1369

Table 4.13 Morozov Results with Smoothed Data

Noise	β_M^s	L^2 Error	H^1S Error	H^1 Error
0.01	0.022645	0.051818	1.119	1.1202
0.02	0.045969	0.075161	1.4491	1.4511
0.03	0.068583	0.092021	1.6565	1.6591
0.04	0.090581	0.10556	1.8096	1.8127
0.05	0.11388	0.11914	1.9437	1.9473
0.06	0.13545	0.12973	2.0508	2.0549
0.07	0.15734	0.13998	2.1487	2.1532
0.08	0.17951	0.14982	2.2397	2.2447
0.09	0.20027	0.15853	2.3201	2.3255
0.1	0.22096	0.16693	2.3962	2.402
0.11	0.23951	0.17437	2.4637	2.4699
0.12	0.25758	0.18154	2.5281	2.5346

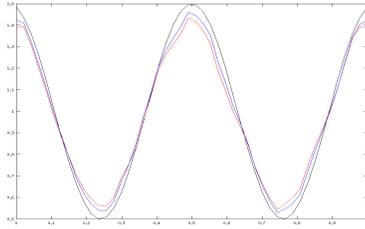


Figure 4.34 Noise of .02

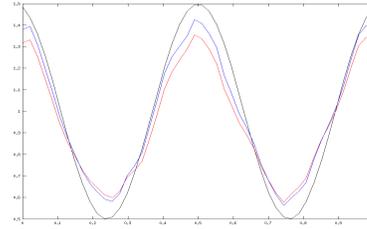


Figure 4.35 Noise of .04

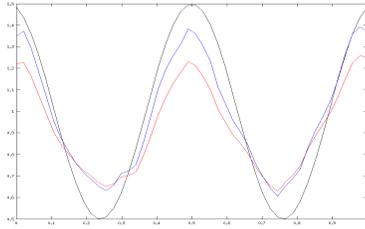


Figure 4.36 Noise of .06

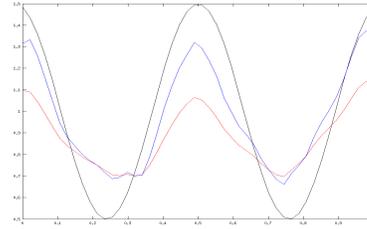


Figure 4.37 Noise of .08

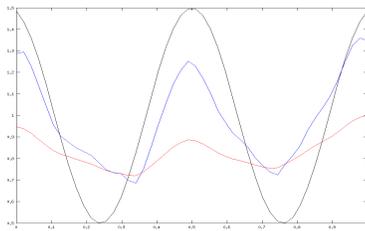


Figure 4.38 Noise of .1

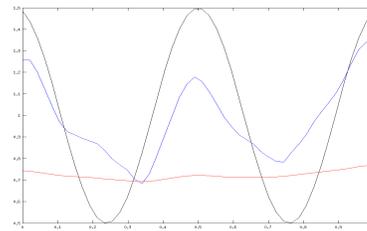


Figure 4.39 Noise of .12

Example 4.5.3. A third second order example reinforces the first two:

$$\begin{aligned} a(x) &= \log(x + 2), \\ u(x) &= -(x^4 - x^3), \end{aligned}$$

In the case of the brute force method, notable flattening of the coefficient reconstructions was evident for noise levels above 1% and 5%. The Morozov principle was able to identify β_M which was about one order of magnitude larger than β_{opt} . As a result a noticeably larger error is evident for larger levels of noise. The damped version of the Morozov equation performs better for the larger noise levels, but did not offer any consistent improvements. Of the numerical methods, Newton's method performs the best of the in this example. The cubic methods require fewer iterations but the extra computa-

tion time makes them fall short in most cases. The cubic methods were able to outperform Newton's method for 1% noise. The quasi-Newton method performs better in this problem in context to other methods, but still takes two or three more times to converge than Newton's method.

Table 4.14 Brute Force Results

$\hat{\delta}$	β_{opt}	H^1 Error
0	1.0802e-06	0.1623
0.001	0.00013611	0.19698
0.005	0.0010952	0.20865
0.0075	0.0020953	0.21455
0.01	0.003388	0.2239
0.05	0.025674	0.48028

Table 4.15 Newton-Type Results

$\hat{\delta}$	β_M	H^1 Error	Bis. Iter	Bis. Time	QN. Iter	QN. Time	N. Iter	N. Time
0.001	0.0029202	0.54435	19	0.078098	10	0.019319	8	0.013634
0.005	0.014763	0.96871	17	0.078988	9	0.01515	6	0.009948
0.0075	0.02214	1.1187	20	0.092362	9	0.013203	6	0.008873
0.01	0.029608	1.2416	19	0.093537	8	0.012256	5	0.01173
0.05	0.21437	2.4953	22	0.099909	12	0.022766	6	0.009903

Table 4.16 Cubic Method Results

$\hat{\delta}$	Cheb. Iter	Cheb. Time	Hal. Iter	Hal. Time	SHal. Iter	SHal. Time
0.001	6	0.015026	7	0.01846	5	0.012832
0.005	5	0.011954	6	0.015487	4	0.009434
0.0075	4	0.009135	6	0.015982	4	0.009549
0.01	4	0.010284	4	0.009561	4	0.009804
0.05	5	0.011862	4	0.020215	4	0.010956

Table 4.17 Damped Regularization Parameters

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.001	6.6757e-06	4.5776e-05	0.00015259	0.00030518	0.00048828	0.00048828
.005	0.0001297	0.00052643	0.001297	0.0021667	0.0027161	0.0029602
.0075	0.00030327	0.0010681	0.0023499	0.0036316	0.0044098	0.004776
.01	0.00057411	0.0018196	0.0036697	0.0053406	0.0063477	0.0068359
.05	0.059181	0.088432	0.1097	0.12498	0.13603	0.14406

$\hat{\delta}/\gamma$	1	1.2	1.4	1.6	1.8	2
.001	0.44009	0.21206	0.19707	0.20016	0.20355	0.20355
.005	0.3558	0.21758	0.20895	0.21254	0.21496	0.21599
.0075	0.35067	0.22461	0.21475	0.2185	0.22141	0.22281
.01	0.35096	0.23445	0.22404	0.22796	0.23144	0.23326
.05	0.53775	0.59392	0.62792	0.64923	0.66322	0.67272

Figure 4.40 Reconstructions Using β_{opt}

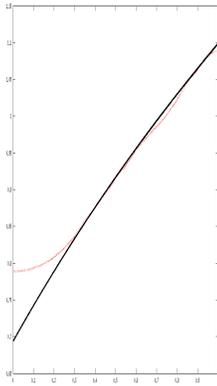


Figure 4.41 Noise of .001

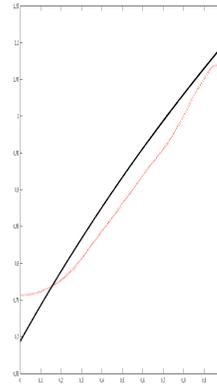


Figure 4.42 Noise of .005

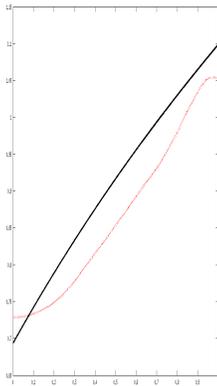


Figure 4.43 Noise of .0075

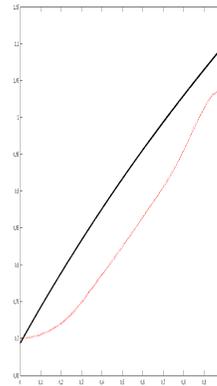


Figure 4.44 Noise of .01

Figure 4.45 Reconstructions Using β_M

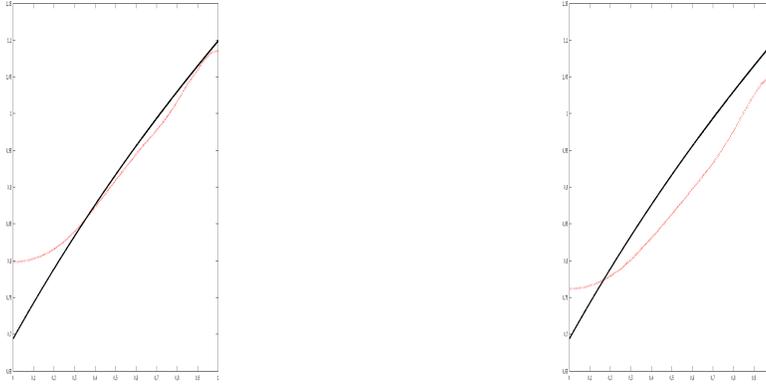


Figure 4.46 Noise of .001

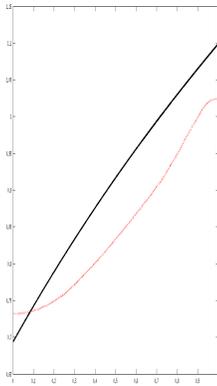


Figure 4.47 Noise of .005

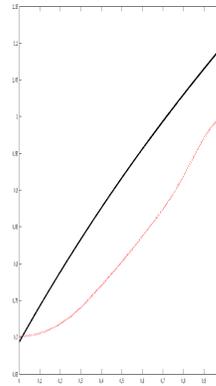


Figure 4.48 Noise of .0075

Figure 4.49 Noise of .01

Noticeable improvements in the reconstruction of the coefficient term are observable with the use of data smoothing. In the instance of 8% noise being added to the data, smoothing reduces $\|Z - Z_s\|_{H^1}^2$ to roughly 72% of $\|Z - Z_\delta\|_{H^1}^2$. Once again, part of the disadvantage of this reduction is that $\|Z - Z_s\|_{L^2}^2$ is made significantly larger than $\|Z - Z_\delta\|_{L^2}^2$. This is evident in the graphs for this example as flattening similar to that observed in example two is evident.

The use of the smoothed data in the identification of the regularization parameter by brute force shows a drastic improvement in the errors for the reconstructions. In the case where 8% noise is added smoothing gives a repro-

duction of the coefficient term which a H^1 error that is 73.8% of that obtained without smoothing. The L^2 error is also experiences a significant decrease in the coefficient reconstructions. When the reconstructions below are viewed as a survey of the reproductions shows that the coefficient reproductions derived from smoothed data maintains a fidelity to the noiseless data which is far less impacted by noise than the unprocessed noisy data.

The Morozov principle when utilized on the unsmoothed data gives parameters which show the familiar overdamping observed in earlier examples. In this case, the root finding algorithm identifies the regularization parameter as thirty for the three largest noise levels. This number was picked as an upper limit for the parameter. At this level of regularization the coefficient which is recovered (red) is not recognizeable when shown in context to the actual term (black). However, the Morozov principle along with smoothing allows for the identification of a regularization parameter which is not only much smaller, but which also carries with it a much smaller error level. The result is that the blue line representing the reconstruction using the Morozov principle with smoothed data retains much more of the actual shape for the coefficient term.

Table 4.18 Error from Noisy Data

Noise	L^2 Error	H^1S Error	H^1 Error
0.01	0.00049329	0.034377	0.03438
0.02	0.00098658	0.068753	0.068761
0.03	0.0014799	0.10313	0.10314
0.04	0.0019732	0.13751	0.13752
0.05	0.0024665	0.17188	0.1719
0.06	0.0029598	0.20626	0.20628
0.07	0.003453	0.24064	0.24066
0.08	0.0039463	0.27501	0.27504

Table 4.19 Error from Smoothed Data

Noise	α	L^2 Error	H^1S Error	H^1 Error
0.01	0.015015	0.0010059	0.03404	0.034055
0.02	0.06006	0.0034846	0.066476	0.066567
0.03	0.13514	0.007188	0.096218	0.096486
0.04	0.23023	0.011268	0.12269	0.12321
0.05	0.35536	0.015801	0.1457	0.14655
0.06	0.50551	0.020291	0.16541	0.16665
0.07	0.67568	0.024448	0.18215	0.18378
0.08	0.87087	0.028319	0.1963	0.19833

Figure 4.50 Noisy Data With and Without Smoothing

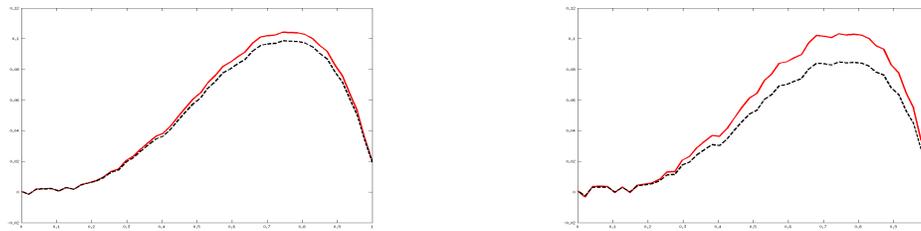


Figure 4.51 Noise of .02

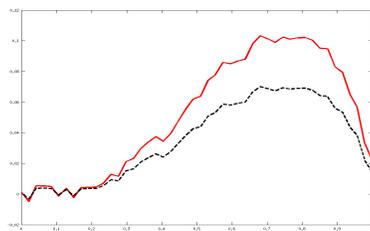


Figure 4.52 Noise of .04

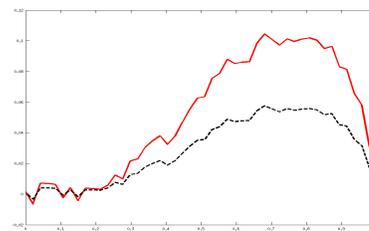


Figure 4.53 Noise of .06



Figure 4.54 Noise of .08



Table 4.20 Coefficient Reconstructions with Noisy Data

Noise	β_{opt}^δ	L^2 Error	H^1S Error	H^1 Error
0.01	0.00091724	0.027478	0.19345	0.19539
0.02	0.0028052	0.069267	0.20318	0.21466
0.03	0.0064047	0.12404	0.21303	0.24651
0.04	0.0098348	0.17761	0.21895	0.28193
0.05	0.013038	0.22819	0.22358	0.31946
0.06	0.016095	0.27558	0.22738	0.35728
0.07	0.019038	0.3197	0.23051	0.39414
0.08	0.021883	0.36056	0.23311	0.42936

Table 4.21 Coefficient Reconstructions with Smoothed Data

Noise	β_{opt}^s	L^2 Error	H^1S Error	H^1 Error
0.01	0.00091798	0.023649	0.1933	0.19474
0.02	0.0029805	0.048083	0.20488	0.21044
0.03	0.0069582	0.085604	0.21473	0.23117
0.04	0.010322	0.11833	0.22089	0.25059
0.05	0.013068	0.14813	0.22618	0.27038
0.06	0.015319	0.17065	0.23107	0.28725
0.07	0.017072	0.1893	0.23556	0.3022
0.08	0.01839	0.20659	0.23967	0.31641

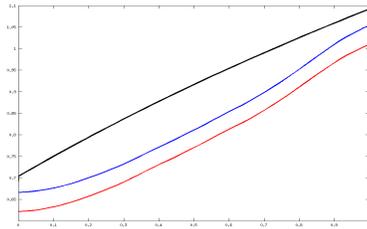


Figure 4.55 Noise of .02

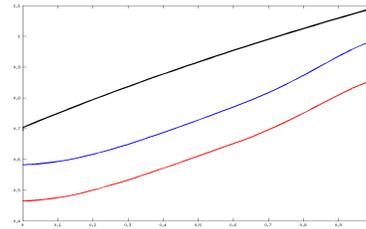


Figure 4.56 Noise of .04

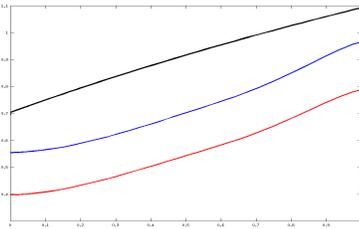


Figure 4.57 Noise of .06

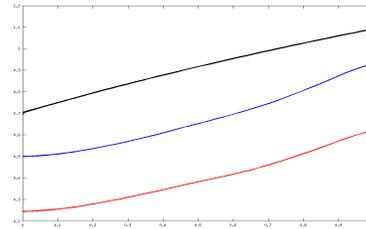


Figure 4.58 Noise of .08

Table 4.22 Morozov Results with Noisy Data

Noise	β_δ	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0.0047141	0.064324	0.19955	0.20966
0.02	0.01217	0.1565	0.20128	0.25497
0.03	0.023983	0.26309	0.2085	0.33569
0.04	0.043624	0.37761	0.22792	0.44107
0.05	0.078721	0.49645	0.26153	0.56113
0.06	0.14871	0.61627	0.30378	0.68708
0.07	0.32026	0.73391	0.34716	0.81187
0.08	1.0649	0.84653	0.38623	0.93047

Table 4.23 Morozov Results with Smoothed Data

Noise	β_s	L^2 Error	$H^1 S$ Error	H^1 Error
0.01	0.0043401	0.05153	0.19792	0.20452
0.02	0.0093036	0.10325	0.19804	0.22335
0.03	0.01388	0.14781	0.1999	0.24861
0.04	0.017708	0.1883	0.20118	0.27556
0.05	0.020334	0.21895	0.20217	0.29801
0.06	0.021905	0.24386	0.2028	0.31717
0.07	0.022714	0.26595	0.20303	0.33459
0.08	0.022822	0.28384	0.20321	0.34909

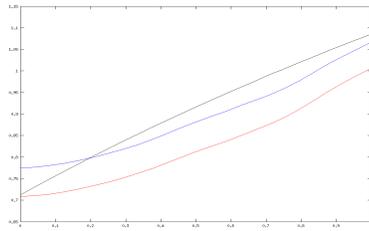


Figure 4.59 Noise of .02

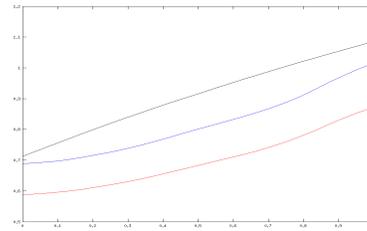


Figure 4.60 Noise of .04

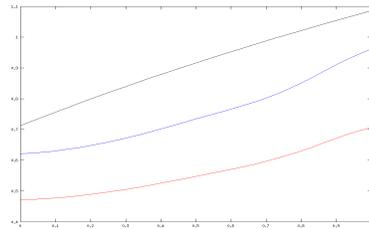


Figure 4.61 Noise of .06

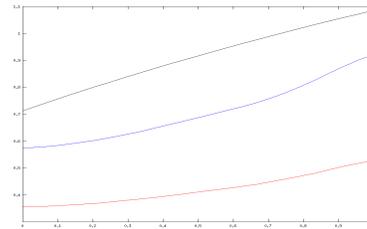


Figure 4.62 Noise of .08

4.6 Conclusion

This chapter showed how a mathematical framework analogous to that used for the source term identification problem in chapter two could be used applied to the classical inverse problem by using the equation error functional. A computational framework was given and shown to be effective on a wide range of examples involving the elliptic partial differential equation in one and two dimensions.

The Morozov principle was shown to be able to identify satisfactory substitutes for β_{opt} . The time it took to identify β_M was a fraction of the time that it

took to employ the brute force method for these simple examples and would presumably hold for larger scale or more complex problems. Where the undamped Morozov principle fell short the damped version was sometimes able to produce more accurate results. This strategy for estimating the regularization parameter was shown to be further improved using error smoothing which also allowed for much larger levels of noise to be considered.

The Morozov principle was a reliable alternative to the trial and error method of estimating the regularization parameter for most cases. The equation error approach's reliance on the data introduces susceptibility to noise, which limited the size of the noise being introduced to the data. To counteract this, error smoothing was implemented on several examples to demonstrate the improvement of the results. Error smoothing not only greatly increased the amount of noise it was possible to generate a reconstruction from, but also improved the quality of the result of the Morozov principle when compared to the optimal regularization parameter. In general, Newton's method was found to be the most effective algorithm for solving the root finding problem. While slight improvements in convergence time were gained with the cubic methods, the improvement was not enough to justify a clear preference for the cubic methods over the quadratic methods.

Bibliography

- [1] R. Acar. Identification of the coefficient in elliptic equations. *SIAM J. Control Optim.*, 31(5):1221–1244, 1993.
- [2] M. F. Al-Jamal and M. S. Gockenbach. Stability and error estimates for an equation error method for elliptic equations. *Inverse Problems*, 28(9):095006, 15, 2012.
- [3] G. Alessandrini. An identification problem for an elliptic equation in two variables. *Ann. Mat. Pura Appl. (4)*, 145:265–295, 1986.
- [4] H. Ammari, P. Garapon, and F. Jouve. Separation of scales in elasticity imaging: a numerical study. *J. Comput. Math.*, 28(3):354–370, 2010.
- [5] H. T. Banks and K. Kunisch. *Estimation techniques for distributed parameter systems*, volume 1 of *Systems & Control: Foundations & Applications*. Birkhäuser Boston, Inc., Boston, MA, 1989.
- [6] Z. Chen and J. Zou. An augmented Lagrangian method for identifying discontinuous parameters in elliptic systems. *SIAM J. Control Optim.*, 37(3):892–910, 1999.
- [7] E. Crossen, M. S. Gockenbach, B. Jadamba, A. A. Khan, and B. Winkler. An equation error approach for the elasticity imaging inverse problem for predicting tumor location. *Comput. Math. Appl.*, 67(1):122–135, 2014.
- [8] H. W. Engl and P. Kügler. The influence of the equation type on iterative parameter identification problems which are elliptic or hyperbolic in the parameter. *European J. Appl. Math.*, 14(2):129–163, 2003.

- [9] M. S. Gockenbach. The output least-squares approach to estimating Lamé moduli. *Inverse Problems*, 23(6):2437–2455, 2007.
- [10] M. S. Gockenbach, B. Jadamba, and A. A. Khan. Numerical estimation of discontinuous coefficients by the method of equation error. *Int. J. Math. Comput. Sci.*, 1(3):343–359, 2006.
- [11] M. S. Gockenbach, B. Jadamba, and A. A. Khan. Equation error approach for elliptic inverse problems with an application to the identification of Lamé parameters. *Inverse Probl. Sci. Eng.*, 16(3):349–367, 2008.
- [12] M. S. Gockenbach and A. A. Khan. Identification of Lamé parameters in linear elasticity: a fixed point approach. *J. Ind. Manag. Optim.*, 1(4):487–497, 2005.
- [13] M. S. Gockenbach and A. A. Khan. An abstract framework for elliptic inverse problems. I. An output least-squares approach. *Math. Mech. Solids*, 12(3):259–276, 2007.
- [14] M. S. Gockenbach and A. A. Khan. An abstract framework for elliptic inverse problems. II. An augmented Lagrangian approach. *Math. Mech. Solids*, 14(6):517–539, 2009.
- [15] K. Ito and K. Kunisch. The augmented Lagrangian method for parameter estimation in elliptic systems. *SIAM J. Control Optim.*, 28(1):113–136, 1990.
- [16] B. Jadamba, A. A. Khan, G. Rus, M. Sama, and B. Winkler. A new convex inversion framework for parameter identification in saddle point problems with an application to the elasticity imaging inverse problem of predicting tumor location. *SIAM J. Appl. Math.*, 74(5):1486–1510, 2014.
- [17] B. Jadamba, A. A. Khan, and M. Sama. Inverse problems of parameter identification in partial differential equations. In *Mathematics in science and technology*, pages 228–258. World Sci. Publ., Hackensack, NJ, 2011.
- [18] B. Jin and P. Maass. Sparsity regularization for parameter identification problems. *Inverse Problems*, 28(12):123001, 70, 2012.
- [19] T. Kärkkäinen. An equation error method to recover diffusion from the distributed observation. *Inverse Problems*, 13(4):1033–1051, 1997.

- [20] J. B. Khan, Akhtar. Numerical estimation of discontinuous coefficients by the method of equation error. *International Journal of Mathematics and Computer Science*, pages 343–359, 2006.
- [21] I. Knowles. Parameter identification for elliptic problems. *J. Comput. Appl. Math.*, 131(1-2):175–194, 2001.
- [22] R. V. Kohn and B. D. Lowe. A variational method for parameter identification. *RAIRO Modél. Math. Anal. Numér.*, 22(1):119–158, 1988.
- [23] P. Kügler. A parameter identification problem of mixed type related to the manufacture of car windshields. *SIAM J. Appl. Math.*, 64(3):858–877 (electronic), 2004.
- [24] K. Kunisch and J. Zou. Iterative choices of regularization parameters in linear inverse problems. *Inverse Problems*, 14(5):1247–1264, 1998.
- [25] K. Kunisch and J. Zou. Iterative choices of regularization parameters in linear inverse problems. *Inverse Problems*, 14(5):1247, 1998.
- [26] J. Li and J. Zou. A multilevel model correction method for parameter identification. *Inverse Problems*, 23(5):1759–1786, 2007.
- [27] S. Manservigi and M. Gunzburger. A variational inequality formulation of an inverse elasticity problem. *Appl. Numer. Math.*, 34(1):99–126, 2000.
- [28] G. R. Richter. An inverse problem for the steady state diffusion equation. *SIAM J. Appl. Math.*, 41(2):210–221, 1981.
- [29] T. Tucciarelli and D. P. Ahlfeld. A new formulation for transmissivity estimation with improved global convergence properties. *Water Resources Research*, 27(2):243–251, 1991.
- [30] K. van den Doel and U. M. Ascher. Dynamic level set regularization for large distributed parameter estimation problems. *Inverse Problems*, 23(3):1271–1288, 2007.
- [31] L. W. White. Estimation of elastic parameters in beams and certain plates: H^1 regularization. *J. Optim. Theory Appl.*, 60(2):305–326, 1989.

- [32] L. W. White. Estimation of elastic parameters in a nonlinear elliptic model of a plate. *Appl. Math. Comput.*, 42(2, part II):139–187, 1991.
- [33] G. Yuan and M. Yamamoto. Lipschitz stability in inverse problems for a Kirchhoff plate equation. *Asymptot. Anal.*, 53(1-2):29–60, 2007.
- [34] J. Zou. Numerical methods for elliptic inverse problems. *Int. J. Comput. Math.*, 70(2):211–232, 1998.
- [35] Y. Zou, L. Wang, and R. Zhang. Cubically convergent methods for selecting the regularization parameters in linear inverse problems. *Journal of Mathematical Analysis and Applications*, 356(1):355–362, 2009.