Bayesian Inverse Problems and Seismic Inversion

Sean Wei Xinq Lim

St Hilda's College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Mathematics

Trinity Term 2016

Acknowledgements

I would like to extend my gratitude to the following people:

- My supervisors, Dr Chris Farmer and Dr Irene Moroz for their patience, guidance and support.
- My industrial collaborators, Dr Zhongmin Song and Dr York Zheng, for arranging visits to the BP Sunbury office for various trainings and visits.
- Dave Ellis, the manager for the UK Seismic Imaging of BP Exploration UK in Sunbury for approving the funding for my entire DPhil period.
- My transfer and confirmation assessors, Dr David Allwright, Dr Tarje Nissen-Meyer, and Dr Patrick Farrell, as well as my examiners, Dr Andrew Thompson and Dr Amos Lawless, for their useful and stimulating criticism.
- My colleagues for many useful discussions.

Abstract

The Bayesian formulation for inverse problems gives a way of making inferences about unknown quantities not directly observable. The application of Bayes' Theorem combines the prior information and the observation to give a posterior measure, which contains information about the quantity we are trying to estimate. In this thesis, we review a particular formulation, conventionally known as the strong constraint formulation of inverse problems. We describe methods to obtain summaries of information from the posterior measure. We also describe how prior measures are constructed using linear differential operators, to quantify as accurately as possible our knowledge of the parameters, independent of any observations.

Then, we note that the strong constraint formulation of inverse problems makes it hard to obtain summaries of information of the posterior measure, typically obtained through an optimization of a misfit functional. Therefore, we introduce the weak constraint formulation in a Bayesian context for inverse problems, which eases the task of optimization. We use this formulation to perform sampling of the posterior measure. This method is tested on some simple test problems. We also compare the results between a strong and weak constraint formulation of inverse problems by studying a one-dimensional example.

Finally, we apply the weak constraint formulation to the problem of full waveform inversion, which is a common problem in seismology. The forward problem we use here is the Laplace transform of the acoustic wave equation, and the inverse problem is solved in several frequencies. There are two approaches when observations at several frequencies are available. First is the sequential method, which processes the observation at different frequencies individually. The second method, which is the simultaneous method, processes the observations at all frequencies at once. We use the simultaneous method here, and used a non-trivial model problem, which yields promising results.

Contents

List of Symbols

Symbol	Description
$\mathcal{H}$	A real, separable Hilbert space
$\mathcal{H}_m$	Parameter space
$\mathcal{H}_u$	Variable space
$\mathcal{H}_s$	Observation space
u	Variables
m	Parameters
s	Observation
$\mathcal{O}$	Observation operator
$\mathcal{G}$	Observation operator as a function of m
ξ	Observation noise
Γ	Observation covariance matrix
Φ	Likelihood potential
μ	A probability measure
μ_0	Prior measure
μ^s	Posterior measure
$\mathcal{N}$	A Gaussian distribution
m_0	Prior mean
$\mathcal{C}_0$	Prior covariance operator
$\mathcal{L}_0$	Prior precision operator
Z	A normalization constant
μ_L	Lebesgue measure
π	A probability density
π_0	Prior probability density
π^s	Posterior probability density
d	Number of physical dimensions
D	A bounded open subset of $\mathbb{R}^d$
$\mathbb{K}$	An indexing set
c	A covariance function
E	Cameron-Martin space
K_r	The r-th modified Bessel function of the second kind
$\mathcal{J}$	Strong constraint misfit functional
$\breve{m}$	Probability maximizer
α	Acceptance probability in a Markov chain Monte Carlo sampler
$\mathbf{1}$	Identity matrix
λ	Nominal prior correlation length

λ_x	Nominal prior correlation length in the x-direction
λ_y	Nominal prior correlation length in the y-direction
σ^2	Nominal prior variance
Δ_B	Absolute correlation of a prior
ν_B	Relative correlation of a prior
J_r	The r-th Bessel function of the first kind
p	Observation precision
ζ	Dynamical noise
$\mathcal{K}$	Weak constraint misfit functional
$\mathcal{I}$	Reduced weak constraint misfit functional

Chapter 1

Introduction

1.1 Motivation

Seismology is the study of the subsurface of the Earth using seismic waves. This can be used for various purposes, such as locating mineral deposits, determining potential archaeological excavation sites, predicting earthquakes, and locating potential hydrocarbon reservoirs. Exploration seismology is a branch of seismology, which differs from the more common and older field of earthquake seismology, in that it is active seismic as opposed to the passive seismic of earthquake seismology. In active seismic, seismic waves are artificially generated using a seismic source, such as an airgun in marine seismic, or a land seismic source like a Vibroseis truck or dynamite. This is different from passive seismic, where the waves being observed are generated by the Earth itself, when tectonic plates move. Also, exploration seismology is performed on a local scale as opposed to the global scale of earthquake seismology. Mathematically, this allows for assumptions such as a flat Earth to be made, which would have been invalid on a global scale.

In active seismic, a source is set off into the Earth's subsurface. The waves then travel through the rocks, and interact according to the various properties of the rocks. Receivers are placed at various locations to detect these waves. This process is known as *seismic acquisition*. The most common type of acquisition, performed before drilling wells to locate more precisely potential locations of hydrocarbon reservoirs, is where receivers are placed on the Earth's surface in land seismic, or on the surface of the waters in marine. There are also marine seismic acquisitions where receivers are placed on the seabed.

There are two major categories of seismic acquisition, each requiring different sets of equipment. Land seismic typically requires a source such as dynamite or a Vibroseis truck, with the receivers being geophones. On the other hand, marine seismic requires an airgun source with hydrophones as receivers.

Once the process of acquisition is done, the data is then analysed to make inferences about the Earth's subsurface. This involves the solution of an inverse problem, but in order to solve the inverse problem, one must understand the forward problem first. In this case, this is the wave equation. For the purposes of this problem, we will consider the acoustic wave equation, given by

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \nabla^2 \phi = f, \tag{1.1}$$

where ϕ is the pressure perturbation, f is a source term and c is the wave speed. For many years, this was used as the model for the physical system, although there have been attempts to use instead the Navier equation as the forward problem, as it describes both the compressional, or P-waves, as well as the shear, or S-waves (see for instance, Alkhalifah, 2014, Brossier *et al.*, 2009, Raknes and Arntsen, 2014, Son *et al.*, 2014).

In Virieux and Operto (2009), three main challenges in the seismic inversion problem were outlined. These were:

1. To build a sufficiently good starting iterate for the optimization algorithm.

2. To develop a computationally tractable optimization algorithm, one which scales with the number of parameters, and is able to find the global minimum reasonably well.

3. To develop effective methods for processing seismic data.

1.2 The Bayesian framework

Traditionally, the inverse problem of parameter estimation is solved by finding a minimum of the misfit functional, which quantifies the difference between the observations and the forecasts from a forward model. We will call this approach the deterministic approach. In this thesis, we will adopt a statistical approach to inverse problems, and the framework used is the Bayesian framework. This brings with it two main benefits. Firstly, it helps us to keep track of the assumptions made. Secondly, it enables us to quantify uncertainty, as quantities are then described as random variables, governed by probability measures.

It should be noted that the Bayesian framework is by no means a superior approach in finding the truth in an inverse problem. However, it is a more general framework compared to the deterministic framework and describes how much can be inferred based on the assumptions made and information collected.

It should also be noted that there have been many criticisms of the Bayesian framework in solving inverse problems. In particular, the statistics community is known to be divided between the Bayesian approach and the frequentist approach. A frequentist criticism of the Bayesian approach is that the construction of the prior measure is difficult and can be essentially arbitrary.

However, from the Bayesian point of view, the limitations of a frequentist approach are clearly seen when there are limited observations available. When this happens, it is difficult for the frequentist approach to estimate parameters, in particular when computing a maximum likelihood estimate. This is because we would be solving an ill-posed mathematical problem. This leads to efforts to regularize the problem, which is equivalent to specifying assumptions about the parameters, as is done within the Bayesian framework. In particular, we assume that the parameters are realizations of a given probability measure. This assumption allows us to construct first the probability measure, called the prior measure, which is then updated as observations are being processed.

Therefore, there is not really a conflict between the Bayesian and frequentist approach, but rather the Bayesian framework provides a reasonable solution within its assumptions, when there are limited observations available. In small scale applications such as medical imaging, the solution of an inverse problem within a Bayesian framework may not be required. However, large scale applications such as in the fields of meteorology and seismology, where there are limited observations, this is necessary. In this thesis, we will only consider Gaussian prior measures as any non-Gaussian behaviour is modelled via a nonlinear transformation included in the forward model. Methods for the construction of priors are described in more detail in Chapter 3.

Others are skeptical about the computational feasibility of the Bayesian approach to fully characterize the posterior measure. For instance, in (Fichtner, 2011, p. 4), Fichtner writes, "Thus, despite its unquestionable conceptual beauty, probabilistic inverse theory is not well suited for full waveform inversion where the number of model parameters is on the order of several thousand to several million."

While this may be indeed true today, it is still important to pursue this field of study for two reasons. Firstly, it should be acknowledged that computing power will eventually improve, and one day, these methods will become feasible. Secondly, in the context of simpler problems, the Bayesian framework provides a good basis on which we can evaluate other methods used in various industries, which may sometimes seem ad hoc.

1.3 Literature Overview

In this section, we provide a review of the work done in full waveform inversion, and also of the theory of Bayesian inverse problems. Note that in the past, full waveform inversion was solved within a deterministic framework, i.e., it was mainly focused on the optimization of the misfit functional. Our contribution in this thesis unites these two approaches. In particular, we address the problem of optimization, and provide a Bayesian interpretation of this approach as well.

1.3.1 Full waveform inversion

The field of exploration seismology has a rich history, dating back as far as the 1920s. Most of the details and references can be found in Sheriff and Geldart (1995, Ch. 1). On the other hand, the theory for full waveform inversion started to develop in the 1970s (see Virieux and Operto, 2009, and references therein). At that time, full waveform inversion was dismissed as it was not practical as local optimization algorithms were converging to a local minimum which was not the global minimum. This was known as the "cycle-skipping problem".

In the 1990s, Pratt and Worthington (1990) reformulated full waveform inversion using the wave equation in the frequency domain instead of the time domain as the forward problem. Further, a frequency-continuation strategy was introduced to find the global minimum (Bunks *et al.*, 1995, Sirgue and Pratt, 2004). See Chapter 6 for more information.

Other forms of the continuation strategy have been explored in Burstedde and Ghattas (2009). Nevertheless, these approaches still suffered from the cycle-skipping problem. Since then, there has been a huge amount of effort in the literature to address this issue. The interested reader should consult Virieux and Operto (2009) and references within for a review of the developments.

Crase *et al.* (1990) and Amundsen (1991) have independently considered non-Gaussian noise assumptions in order to use different norms in the misfit functional. Shin and Cha (2008, 2009), Son *et al.* (2014), Pyun *et al.* (2011) have considered a lognormal error on the observation, working instead in the more general Laplace domain rather than the frequency domain.

More recently, van Leeuwen and Herrmann (2015) used a penalty function in which the PDE-constraints are relaxed. In particular, they have applied this method to the problem of full waveform inversion. This method is similar to the weak constraint methods known to meteorologists (Fisher *et al.*, 2005, Trémolet, 2006, 2007, Zupanski, 1996) and will be discussed in detail in Chapter 4 of this thesis. van

Leeuwen and Herrmann (2015) proved some theorems on convergence and introduced some useful approximations for the optimization as well.

1.3.2 Bayesian inverse problems

The preceding literature review concerned the traditional deterministic approach to inverse problems. In 1987, Tarantola published a book outlining the Bayesian framework to solve inverse problems. The second edition of the book was released in 2005 (Tarantola, 2005).

A more theoretical perspective of inverse problems, dealing with function spaces and existence of local minima are explored in Kirsch (1996). Kaipio and Somersalo (2006) developed a framework for finite-dimensional inverse problems within the Bayesian framework. An infinite-dimensional version of this framework using Gaussian priors can be found in Stuart (2010), while sampling algorithms in infinite dimensions are described in Cotter *et al.* (2013). A detailed overview of inverse problems and uncertainty quantification can be found in Sullivan (2015).

In the minimization of the misfit functional, often the computation of the gradient of the misfit functional is required. A direct computation of this gradient involves the computation of the forward model, which is computationally impractical. Using the adjoint of a differential operator, the gradient of the misfit functional can be computed without explicitly computing the gradient of the forward model. This is described in, for instance, Tarantola (1984a,b, 2005), Fichtner (2011).

The framework introduced in Stuart (2010) was used by Tan *et al.* (2013) to solve inverse problems in global seismology problems, and efforts have been made to make it more computationally tractable (Tan *et al.*, 2012, Martin *et al.*, 2012). In particular, Martin *et al.* (2012) introduces a Stochastic Newton MCMC. Here, an MCMC method is used where the proposals were generated from a Gaussian approximation of a minimum of the misfit functional. The precision matrix used here is a low-rank approximation of the Hessian.

The statistical method described in Tan *et al.* (2012, 2013), Martin *et al.* (2012) generates only one Markov chain, where the sample is taken from. In contrast, the method that is introduced in this thesis begins with a prior ensemble, and the posterior ensemble is taken from the end of each individual Markov chains, starting from a randomized maximum likelihood ensemble. Since the success of this algorithm relies heavily on the solution of multiple optimization problems, it is crucial that this issue is also addressed. This is done through the weak constraint formulation of an inverse problem, which convexifies the misfit, and makes this problem compu-

tationally feasible. The Stochastic Newton MCMC method as introduced in Martin *et al.* (2012) does not address this problem.

Note that the subject of the weak constraint formulation of an inverse problem introduced in Chapter 4 is related to the work in van Leeuwen and Herrmann (2015). However, in this thesis, the weak constraint formulation is introduced within a Bayesian framework, which is novel.

1.4 Thesis outline

In Chapter 2, we will describe a Bayesian framework of inverse problems, which is valid in both finite and infinite dimensions, using Gaussian measures as the prior measure. A description of the prior measures and how they can be constructed is found in Chapter 3. Here, we will also address the problem of modified Helmholtz measures, and why they may pose a problem in infinite dimensions when used as prior measures. In Chapter 4, we describe the weak constraint formulation for inverse problems. The idea here is to use a non-deterministic forward model. In this chapter, we also describe how sampling can be performed using techniques introduced in petroleum reservoir characterization. Chapter 5 deals with a simple model problem, which is the inverse problem for the simple harmonic oscillator. In Chapter 6, we apply the weak constraint formulation for the problem of full waveform inversion. We make some concluding remarks in Chapter 7, suggesting future directions of study.

1.5 Author's contribution

The main contribution of the author is in Chapter 4, where we introduce the weak constraint formulation for inverse problems. In particular, we have provided a Bayesian interpretation of the weak constraint formulation, which is novel. The sampling algorithm introduced in this chapter is a combination of the three ingredients:

1. The weak constraint formulation of the inverse problem.
2. The Randomized Maximum Likelihood method as introduced in Oliver *et al.* (2008).
3. The preconditioned Crank-Nicolson random walk Markov chain Monte Carlo algorithm found in Cotter *et al.* (2013).

The author's novel contribution here is a unification of these methods, in order to generate a sample of the posterior measure.

Chapter 5, which compares the results of the strong and weak constraint formulations of inverse problems using the inverse problem on the simple harmonic oscillator, is a novel work of the author.

Chapter 6, which applies the weak constraint formulation on the problem of full waveform inversion is also novel.

In Section 3.3 of Chapter 3, there is a discussion on the relation between the regularity of the samples and infinite variances. This is a novel work of the author as well. Note that the subject of regularity of samples from a probability measure has already been discussed in Stuart (2010), but this work provides a way to examine this property in practice.

Chapter 2

Bayesian Inverse Problems

In this chapter, we provide an exposition of inverse problem theory. We will adopt the statistical approach to inverse problems, working within the Bayesian framework. We will state results without proving them, and give references for more details. Most of the presentation in this chapter is based on Stuart (2010).

2.1 General Framework

Let $\mathcal{H}_m$, $\mathcal{H}_u$ and $\mathcal{H}_s$ be real separable Hilbert spaces, equipped with their respective inner products, $\langle \cdot, \cdot \rangle_m$, $\langle \cdot, \cdot \rangle_u$, $\langle \cdot, \cdot \rangle_s$, as well as their corresponding norms $\|\cdot\|_m$, $\|\cdot\|_u$ and $\|\cdot\|_s$. Suppose that the quantities u and m are related by the implicit relation

$$\mathcal{E}(m, u) = 0, \tag{2.1}$$

where $m \in \mathcal{H}_m$ and $u \in \mathcal{H}_u$, and $\mathcal{E} : \mathcal{H}_m \times \mathcal{H}_u \to \mathbb{R}$ is a given function. The quantities m are known as the *parameters* of the model while u are the *variables*.

Suppose also that each m in (2.1) gives rise to a unique set of values, u, i.e., $u = \mathcal{F}(m)$, for some $\mathcal{F} : \mathcal{H}_m \to \mathcal{H}_u$. This is an application of the implicit function theorem (Apostol, 1981, Theorem 7.6). The task of computing u, given m is known as the *forward problem*, and its solution is called the *forward model*. In practical applications, (2.1) is a model of the physical system in question, typically written as an integral or differential equation.

In some cases, m is not known and is not directly observable. In such instances, an inverse problem is solved, where the task is to recover m from partial or full observations of u. We write

$$s = \mathcal{O}(u) + \xi, \tag{2.2}$$

where $s \in \mathcal{H}_s$ is known as the *observation*, $\mathcal{O} : \mathcal{H}_u \to \mathcal{H}_s$ is the observation operator, and ξ is additive noise. Typically, we assume that ξ is normally distributed, i.e., $\xi \sim$

$\mathcal{N}(0, \Gamma)$, for some covariance operator Γ. Equation (2.2) is known as the *observation model*.

We may also write the observation model as a pure function of m. Formally, given m, we may solve (2.1) for u and apply the observation operator to u. Thus, an observation may be thought of as depending purely on the parameters, m, that is,

$$s = \mathcal{O}(\mathcal{F}(m)) + \xi = \mathcal{G}(m) + \xi, \tag{2.3}$$

where $\mathcal{G} : \mathcal{H}_m \to \mathcal{H}_s$. It is this form of the observation model that will be more commonly referred to in our further discussion instead of (2.2). Moreover, while we acknowledge that, in principle, an observation could be infinite-dimensional, we will only consider finite-dimensional observations. Hence, we have $\mathcal{H}_s = \mathbb{R}^S$ for some positive integer S, and a covariance matrix Γ.

Combining the forward and observation model, we obtain the conditional density of s given m, which statisticians call the "likelihood of m given s",

$$\rho(s \mid m) \propto \exp\left(-\Phi(m, s)\right), \tag{2.4}$$

where $\Phi : \mathcal{H}_m \times \mathcal{H}_s \to \mathbb{R}$, is called the *potential* of the likelihood, and is given by

$$\Phi(m, s) = \frac{1}{2} \left\| s - \mathcal{G}(m) \right\|_{\Gamma}^2. \tag{2.5}$$

For a particular value of s, $\rho(s \mid m)$ is not a probability density in terms of m. Note also that for the purposes of this chapter, whenever we write down a probability density as in (2.4), we mean a density with respect to the usual Lebesgue measure.

The Bayesian framework involves the specification of a prior measure, which is a quantification of our knowledge of the system prior to making observations. Throughout this thesis, we will specify a Gaussian prior. Thus, the prior measure, denoted by μ_0, is a Gaussian measure with mean m_0 and covariance operator $\mathcal{C}_0$, defined over a real separable Hilbert space, $\mathcal{H}$ such that $\mathcal{H}_m$ is compactly embedded in $\mathcal{H}$ and $\mu_0(\mathcal{H}_m) = 1$. We write $\mu_0 = \mathcal{N}(m_0, \mathcal{C}_0)$. Note that in some cases, a non-Gaussian prior may be converted into a Gaussian prior through a modification of $\mathcal{G}$. For instance, if we specify a lognormal prior on a parameter $\hat{m}$, then a Gaussian prior can be imposed by applying a change of variable $m := \ln(\hat{m})$. We also note that in cases where some limited observations of m are available using additional observation models, it is assumed that such information has already been used in building the prior, for instance using Bayesian interpolation as described in Chapter 3.

Through an informal application of Bayes' theorem (see Section 2.3 for a rigorous justification), we obtain the posterior measure, denoted by μ^s, and its *Radon-Nikodym derivative* with respect to the prior measure is given by

$$\frac{\mathrm{d}\mu^s}{\mathrm{d}\mu_0} = \frac{1}{Z(s)} \exp(-\Phi(m, s)), \tag{2.6}$$

where the normalization constant $Z(s)$ is given by the integral

$$Z(s) = \int_{\mathcal{H}_m} \exp(-\Phi(m, s)) \, \mathrm{d}\mu_0(m). \tag{2.7}$$

In the classical approach to inverse problems, a least-squares problem is solved to seek a point, $\breve{m}$, which minimizes the discrepancy as quantified by Φ between the observed data and the forward model. In contrast, the statistical formulation of inverse problems allows us to treat relevant quantities as random variables, governed by appropriate probability measures. The solution of an inverse problem is then given by the posterior measure, defined through the Radon-Nikodym derivative with respect to the prior measure as given in (2.6). Note that (2.6) is a well-defined application of Bayes' Theorem even in infinite dimensions.

In finite dimensions, we may use the Lebesgue measure as a reference measure. Then, the Radon-Nikodym derivatives of the posterior and prior measures with respect to the Lebesgue measure give us the posterior and prior probability densities, π^s and π_0, respectively, i.e.,

$$\frac{\mathrm{d}\mu^s}{\mathrm{d}\mu_L} = \pi^s(m), \quad \text{and} \quad \frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_L} = \pi_0(m), \tag{2.8}$$

where μ_L denotes the Lebesgue measure. Then, multiplying (2.6) by $\mathrm{d}\mu_0/\mathrm{d}\mu_L$ and using the fact that

$$\frac{\mathrm{d}\mu^s}{\mathrm{d}\mu_0}\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_L} = \frac{\mathrm{d}\mu^s}{\mathrm{d}\mu_L}, \tag{2.9}$$

we have

$$\frac{\mathrm{d}\mu^s}{\mathrm{d}\mu_L} \propto \exp(-\Phi(m, s))\frac{\mathrm{d}\mu_0}{\mathrm{d}\mu_L}, \tag{2.10}$$

or

$$\pi^s(m) \propto \exp(-\Phi(m, s))\pi_0(m), \tag{2.11}$$

which is Bayes' Theorem in finite dimensions.

2.2 Prior measure

In this section, we give a more in-depth discussion on our choice of prior measure. As mentioned above, we will consider only Gaussian prior measures.

Since a Gaussian measure is fully characterized by its mean and covariance, we now describe how one might specify a prior covariance operator. To this end, let μ be a Gaussian measure, i.e., $\mu = \mathcal{N}(m_0, \mathcal{C})$ on a real separable Hilbert space, $(\mathcal{H}, \langle \cdot, \cdot \rangle)$. Without loss of generality, we may assume that $m_0 = 0$. We define the inverse covariance operator to be the *precision operator*, denoted by $\mathcal{L}$.

One way to specify a prior covariance operator is by specifying $\mathcal{L}$ as a second-order positive definite elliptic operator, so that $\mathcal{C}$ is its inverse.

In particular, let $D \subseteq \mathbb{R}^d$, where $d = 1$, 2 or 3, be a bounded open set, and $\mathcal{H} \subseteq L^2(D)$, equipped with the L^2-inner product, denoted by $\langle \cdot, \cdot \rangle$. We want to specify a second-order linear differential operator $\mathcal{L}$ such that Assumption 2.9 of Stuart (2010) is satisfied. We specify this assumption below.

Assumption 2.2.1 (Assumption 2.9 of Stuart (2010))**.** Let $\mathbb{K} \subseteq \mathbb{Z}^d \backslash \{0\}$ be an indexing set, where $d > 0$. We write $k = (k_1, k_2, \cdots, k_d)^T \in \mathbb{K}$. Assume that the operator $\mathcal{L}$, densely defined on a real separable Hilbert space $\mathcal{H} \subseteq L^2(D)$, satisfies the following properties.

1. $\mathcal{L}$ is positive definite, self-adjoint and invertible.

2. The eigenfunctions, $\{\phi_k\}_{k \in \mathbb{K}}$ and corresponding eigenvalues, $\{\lambda_k\}_{k \in \mathbb{K}}$ of $\mathcal{L}$ form an orthonormal basis for $\mathcal{H}$.

3. There exist $C^+ > 0$ and $C^- > 0$ such that the eigenvalues satisfy, for all $k \in \mathbb{K}$,

$$C^- \leq \frac{\lambda_k}{\prod_{i=1}^d k_i^2} \leq C^+. \tag{2.12}$$

4. There exists $C > 0$ such that

$$\sup_{k \in \mathbb{K}} \left(\|\phi_k\|_{L^\infty} + \left\| \frac{1}{k} \cdot \nabla \phi_k \right\|_{L^\infty} \right) \leq C, \tag{2.13}$$

   where $1/k = (1/k_1, 1/k_2, \cdots, 1/k_d)^T$.

Under the conditions of Assumption 2.2.1, the covariance operator can be expressed as an integral operator, and furthermore its action can be computed without explicitly calculating the integral operator itself.

To see this, suppose that $\mathcal{L}$ is a precision operator which satisfies Assumption (2.2.1). Then, the covariance operator, $\mathcal{C}$ is well-defined by the first assumption. The second, third and fourth assumption imply that for χ, $\psi \in \mathcal{H}$, the integral

$$\langle \chi, \mathcal{C}\psi \rangle = \int_{\mathcal{H}} \langle \chi, m \rangle \langle \psi, m \rangle \, \mathrm{d}\mu(m), \tag{2.14}$$

which is the definition of a covariance operator, exists. Here, μ is the prior measure. Then, using the definition of the L^2-inner product, we have

$$\langle \chi, \mathcal{C}\psi \rangle = \int_{\mathcal{H}} \int_D \int_D \chi(x) m(x) m(y) \psi(y) \, \mathrm{d}y \, \mathrm{d}x \, \mathrm{d}\mu(m). \tag{2.15}$$

Interchanging the order of integration, we have

$$\langle \chi, \mathcal{C}\psi \rangle = \int_D \chi(x) \int_D c(x, y) \psi(y) \, \mathrm{d}y \, \mathrm{d}x, \tag{2.16}$$

where the function $c(x, y)$ is

$$c(x, y) = \int_{\mathcal{H}} m(x) m(y) \, \mathrm{d}\mu(m). \tag{2.17}$$

Therefore, the action of the covariance operator on a given function ψ is given by

$$(\mathcal{C}\psi)(x) = \int_D c(x, y) \psi(y) \, \mathrm{d}y, \tag{2.18}$$

that is, the covariance operator is an integral operator with kernel $c(x, y)$. The function $c(x, y)$ is known as the *covariance function*.

Since the covariance operator is the inverse of the precision operator, it follows that the covariance function is the Green's function of $\mathcal{L}$, so that, for all $m \in \mathcal{H}$, we have

$$\mathcal{L}(\mathcal{C}m)(x) = \mathcal{L}\left(\int_D c(x, y) m(y) \, \mathrm{d}y \right) = \int_D \delta(x - y) m(y) \, \mathrm{d}y = m(x). \tag{2.19}$$

Therefore, specifying a Gaussian prior measure can be done either by specifying the precision operator, covariance operator or the covariance function (Farmer, 2007, Stuart, 2010). In practical applications, it is increasingly common to specify the prior through its precision operator, due to its simplicity in specifying its correlation points within a neighborhood. Moreover, from a computational perspective, since a precision operator is specified using a differential operator, it can be approximated by a sparse matrix. In contrast, storing and working with its inverse, which is a dense matrix in general, requires significantly more computer memory and time, and is essentially infeasible, for high-dimensional problems.

Furthermore, if $\{\phi_k\}_{k \in \mathbb{K}}$ is a set of orthonormal eigenfunctions of $\mathcal{L}$ with $\{\lambda_k\}_{k \in \mathbb{K}}$ as its corresponding eigenvalues, then for any $m \in \mathcal{H}$, we may write m as a weighted sum of the ϕ_k's, or

$$m(x) = \sum_{k \in \mathbb{K}} \left(\int_D \phi_k(y) m(y) \, \mathrm{d}y \right) \phi_k(x). \tag{2.20}$$

Applying the covariance operator gives

$$(\mathcal{C}m)(x) = \sum_{k\in\mathbb{K}} \frac{1}{\lambda_k} \left(\int_D \phi_k(y) m(y) \, \mathrm{d}y \right) \phi_k(x), \tag{2.21}$$

since $\mathcal{C}\phi_k = (1/\lambda_k)\phi_k$ for all k. Interchanging the summation and integration, we have

$$(\mathcal{C}m)(x) = \int_D c(x,y)m(y) \, \mathrm{d}y = \int_D \sum_{k\in\mathbb{K}} \frac{\phi_k(x)\phi_k(y)}{\lambda_k} m(y) \, \mathrm{d}y. \tag{2.22}$$

Since this holds for all $m \in \mathcal{H}$, we deduce that

$$c(x,y) = \sum_{k\in\mathbb{K}} \frac{\phi_k(x)\phi_k(y)}{\lambda_k}. \tag{2.23}$$

Equation (2.21) gives us a way of computing the action of the covariance operator on a function without explicitly computing the covariance operator itself. In particular, if we can compute the eigenfunctions and eigenvalues of $\mathcal{L}$, then we may truncate the sum in (2.21) and evaluate the integrals to get a numerical approximation of the action of the covariance on a function.

Furthermore, in sampling, it is of interest to work with the square root of the covariance operator, $\mathcal{C}^{1/2}$ rather than the covariance operator itself. The action of $\mathcal{C}^{1/2}$ on a function $m \in \mathcal{H}$ is given by

$$(\mathcal{C}^{1/2}m)(x) = \sum_{k\in\mathbb{K}} \frac{1}{\lambda_k^{1/2}} \left(\int_D \phi_k(y) m(y) \, \mathrm{d}y \right) \phi_k(x), \tag{2.24}$$

if the sum converges. We define the range of $\mathcal{C}^{1/2}$ as the *Cameron-Martin space*, and denote it by E.

In our preceding discussion, we mentioned that a way to specify the covariance of a Gaussian prior measure is through the specification of its precision operator. One way to do this in practical applications is to use the *modified Helmholtz operator*, defined by

$$\mathcal{L} = ab^2 I - a\nabla^2, \tag{2.25}$$

where $a > 0$, $b > 0$, and I is the identity operator. On an infinite domain with vanishing Dirichlet boundary conditions at infinity, the Green's function for $\mathcal{L}$ in 3-D, given in Farmer (2007), is

$$g(r) = \frac{e^{-rb}}{4\pi a r}, \tag{2.26}$$

where $r = |x - y|$. In 2-D, this is

$$g(r) = \frac{1}{2\pi a} K_0(br), \tag{2.27}$$

where K_0 is the modified Bessel function of the second kind. In 1-D, the Green's function is

$$g(r) = \frac{e^{-br}}{2ab}. \tag{2.28}$$

As an example, we consider the modified Helmholtz operator in 1-D. Then, the correlation function is given by

$$c(x, y) = g(|x - y|) = \frac{e^{-b|x-y|}}{2ab}. \tag{2.29}$$

The pointwise variance is $\sigma^2 := c(x, x) = g(0) = 1/(2ab)$, and the correlation length is defined as the integral of the normalized correlation function. Therefore, the correlation length, denoted by λ, is given by

$$\lambda = \frac{1}{g(0)} \int_0^\infty g(r) \, \mathrm{d}r = \frac{1}{b}. \tag{2.30}$$

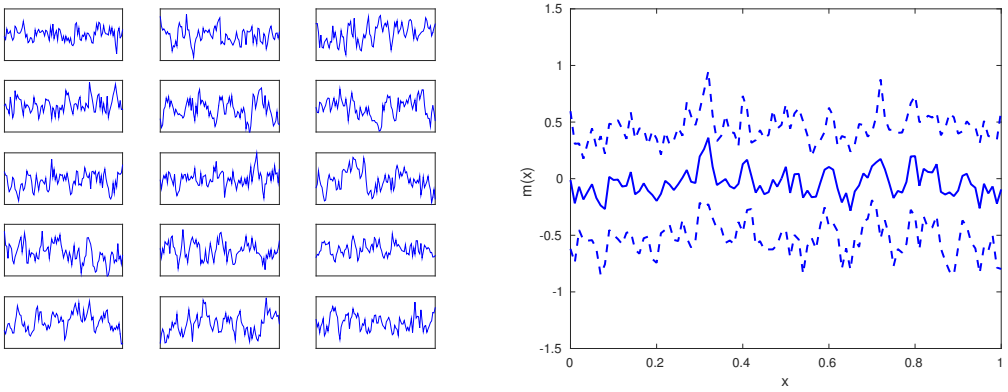
As an illustration, Figures 2.1, 2.2 and 2.3 show 15 random draws for different choices of correlation lengths. The prior measures here were sampled on the domain $[0, 1]$ on a uniform grid spacing of length h and $N + 1$ grid points, where $h = 1/N$. For this simulation, we chose $N = 100$. If L is the modified Helmholtz operator, discretized using the finite difference scheme, with homogeneous Neumann boundary conditions, and A is the Cholesky factor of L so that $L = AA^T$, a draw, ζ was generated using

$$\zeta = A^{-T}\eta, \tag{2.31}$$

where η is random vector drawn from the standard normal distribution. Note that this is an exact sample since we have

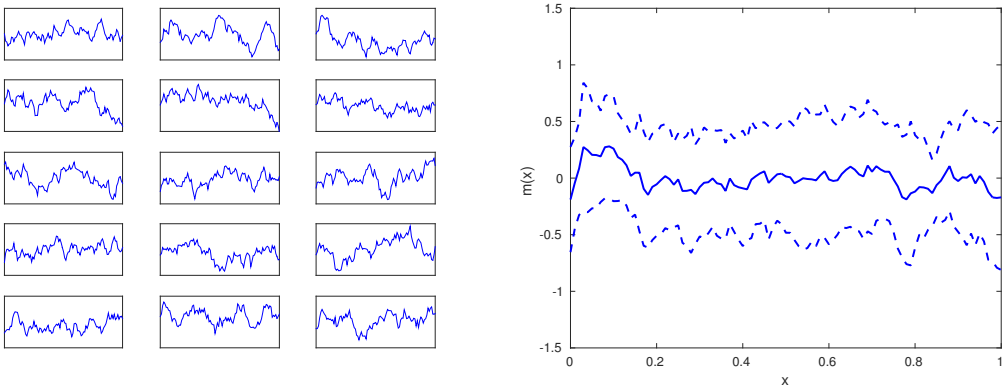
$$\mathbb{E}[\zeta] = \mathbb{E}[A^{-T}\eta] = 0, \quad \text{and} \quad \mathbb{E}[\zeta\zeta^T] = A^{-T}\mathbb{E}[\eta\eta^T]A^{-1} = L^{-1}. \tag{2.32}$$

In Figure 2.1, when the correlation length is short, a realization of the prior resembles white noise. When the correlation length is increased to 0.1, we begin to see the correlation structure between grid points, although the draws still contain some roughness (Figure 2.2). When $\lambda = 0.5$, the correlation length is half the domain, and each random draw is significantly smoother (Figure 2.3).



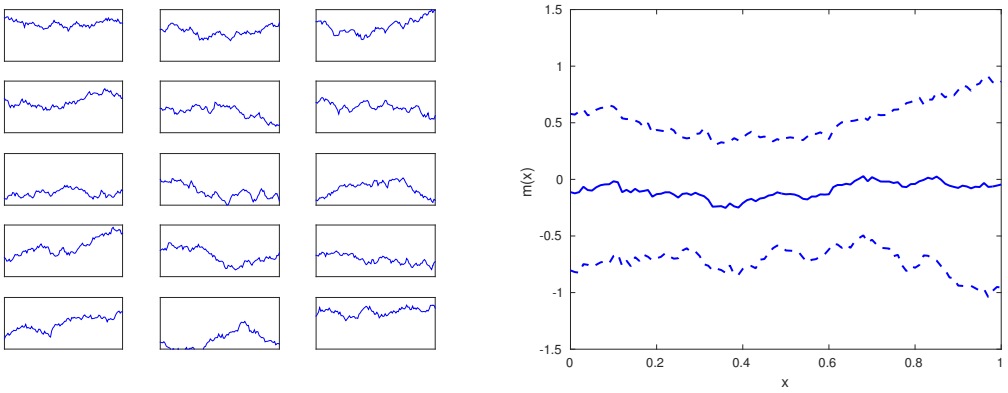
(a) (b)

Figure 2.1: (a) 15 realizations of the modified Helmholtz prior with $\lambda = 0.02$, $\sigma = 0.5$. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 2.2: (a) 15 realizations of the modified Helmholtz prior with $\lambda = 0.1$, $\sigma = 0.5$. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 2.3: (a) 15 realizations of the modified Helmholtz prior with $\lambda = 0.5$, $\sigma = 0.5$. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

2.3 Posterior measure

2.3.1 Justification

In general, the posterior measure whose Radon-Nikodym derivative is given by (2.6) may not be a well-defined probability measure. In order to demonstrate the existence of the posterior measure, the observation operator, $\mathcal{G}(m)$ has to satisfy the following assumption.

Assumption 2.3.1 (Assumption 2.7 of Stuart (2010))**.** The function $\mathcal{G}(m)$ satisfies the following.

1. For all $\varepsilon > 0$, there exists an $M = M(\varepsilon) \in \mathbb{R}$ such that for all $m \in \mathcal{H}_m$,

$$\|\mathcal{G}(m)\|_s \leq \exp\left(\varepsilon \|m\|_m^2 + M\right). \tag{2.33}$$

2. For all $r > 0$, there exists $K = K(r) > 0$ such that for all $m_1, m_2 \in \mathcal{H}_m$,

$$\|\mathcal{G}(m_1) - \mathcal{G}(m_2)\|_s \leq K \|m_1 - m_2\|_m. \tag{2.34}$$

We then have the following lemma.

Lemma 2.3.2 (Lemma 2.8 of Stuart (2010))**.** *Let $\mathcal{G} : \mathcal{H}_m \to \mathcal{H}_s$ satisfy Assumption 2.3.1. Then, the function $\Phi(m, s)$ as defined in (2.5) satisfies the following properties.*

1. *For every* $\varepsilon > 0$ *and* $r > 0$*, there exists* $M = M(\varepsilon, r) \in \mathbb{R}$ *such that, for all* $m \in \mathcal{H}_m$ *and* $s \in \mathcal{H}_s$ *with* $\|s\|_s < r$*,*

$$\Phi(m, s) \geq M - \varepsilon \|m\|_m^2 . \tag{2.35}$$

2. *For every* $r > 0$*, there exists* $K = K(r) > 0$ *such that for all* $m \in \mathcal{H}_m$ *and* $s \in \mathcal{H}_s$*, with* $\max\{\|m\|_m, \|s\|_s\} < r$*,*

$$\Phi(m, s) \leq K. \tag{2.36}$$

3. *For every* $r > 0$*, there exists* $L = L(r) > 0$ *such that for all* $m_1, m_2 \in \mathcal{H}_m$*, and* $s \in \mathcal{H}_s$ *with* $\max\{\|m_1\|_m, \|m_2\|_m, \|s\|_s\} < r$*,*

$$|\Phi(m_1, s) - \Phi(m_2, s)| \leq L \|m_1 - m_2\|_m . \tag{2.37}$$

4. *For every* $\varepsilon > 0$ *and* $r > 0$*, there exists* $C = C(\varepsilon, r) \in \mathbb{R}$ *such that for all* $s_1, s_2 \in \mathcal{H}_s$ *with* $\max\{\|s_1\|_s, \|s_2\|_s\} < r$*, and for all* $m \in \mathcal{H}_m$*,*

$$|\Phi(m, s_1) - \Phi(m, s_2)| \leq \exp(\varepsilon \|m\|_m^2 + C) \|s_1 - s_2\|_s . \tag{2.38}$$

The properties (1)–(4) in Lemma 2.3.2 are found in Assumption 2.6 of Stuart (2010), and they are the boundedness and Lipschitz properties of $\Phi(m, s)$.

Using Lemma 2.3.2, we are able to establish some results concerning the well-definedness and stability of μ^s with respect to perturbations in the observations. Before we state the result, we define the *Hellinger distance*.

Definition 2.3.3. Let μ and μ' be two probability measures which are absolutely continuous with respect to a reference measure ν. Then, the *Hellinger distance* between μ and μ' is defined by (Stuart, 2010)

$$d_{\text{Hell}}(\mu, \mu') = \sqrt{\frac{1}{2} \int_{\mathcal{H}} \left(\sqrt{\frac{\mathrm{d}\mu}{\mathrm{d}\nu}} - \sqrt{\frac{\mathrm{d}\mu'}{\mathrm{d}\nu}} \right)^2 \mathrm{d}\nu}. \tag{2.39}$$

Note that the Hellinger distance is invariant with respect to the reference measure, and that

$$0 \leq d_{\text{Hell}}(\mu, \mu') \leq 1. \tag{2.40}$$

We are now ready to state the result regarding well-definedness of the posterior measure and its stability with respect to perturbations in the observations.

Theorem 2.3.4 (Corollary 4.4 of Stuart (2010)). *Let $\mathcal{G}$ satisfy Assumption 2.3.1 and assume that μ_0 is a Gaussian measure satisfying $\mu_0(\mathcal{H}_m) = 1$. Then, the measure μ^s given by (2.6) is a well-defined probability measure on $\mathcal{H}$ and is Lipschitz in the data s, with respect to the Hellinger distance: if μ^s and $\mu^{s'}$ are two measures corresponding to the data s and s', then there exists $C = C(r) > 0$ such that, for all s, s' with $\max\{s, s'\} < r$,*

$$d_{Hell}(\mu^s, \mu^{s'}) \leq C \|s - s'\|_s . \tag{2.41}$$

Consequently, the expectation of any polynomially bounded function $f : \mathcal{H}_m \to E$ is continuous in s. In particular, the mean and the covariance operator are continuous in s.

Finally, an application of Theorem 6.31 of Stuart (2010) justifies the formula in (2.6), which, as previously pointed out, is Bayes' Theorem in infinite or finite dimensions, reproduced below.

Theorem 2.3.5 (Theorem 6.31 of Stuart (2010)). *Assume that $\mathcal{G}$ is continuous and that $\mu_0(\mathcal{H}_m) = 1$. Then, the conditional distribution of m given s, $m \mid s$ is given by the measure $\mu^s(\mathrm{d}m)$, which is absolutely continuous with respect to $\mu_0(\mathrm{d}m)$ and has the Radon-Nikodym derivative given by (2.6).*

In Section 3 of Stuart (2010), some examples are constructed to show that these assumptions are satisfied, and hence the posterior measure is a well-defined probability measure. More examples of such proofs can be found in Cotter *et al.* (2009) and Dashti and Stuart (2011). For our purposes, we will always assume without proof that the posterior measure is indeed a well-defined probability measure with its Radon-Nikodym derivative given by (2.6).

2.3.2 Interpretation

The posterior measure encapsulates neatly the information we have about the parameters, given the observations and the prior measure. Nevertheless, in practical applications, merely having the posterior measure is not sufficient. We need to probe further the posterior measure in order to seek useful information.

Deterministic approach

One way to do this is to seek a mode $\breve{m}$ of the posterior measure μ^s. In finite dimensions, this is equivalent to seeking a maximum a posteriori, or MAP estimator.

It is a peak of the posterior density. The MAP estimator is a local minimizer of the functional

$$\mathcal{J}^s(m) = \Phi(m, s) + \frac{1}{2} \|m - m_0\|^2_{\mathcal{C}_0} . \tag{2.42}$$

In infinite dimensions, there is no notion of a "peak" of a probability density. Instead, we consider the probability of a ball centred at a point $m \in \mathcal{H}_m$ of radius ε. This is

$$\mu^s(B(m; \varepsilon)) = \frac{1}{Z(s)} \int_{B(m;\varepsilon)} \exp(-\Phi(m', s)) \, \mathrm{d}\mu_0(m'). \tag{2.43}$$

Then, we seek a point $\breve{m}$ where this probability is maximized, asymptotically as $\varepsilon \to 0$. This is equivalent to finding $\breve{m}$ which minimizes $\mathcal{J}^s$. The point $\breve{m}$ is then a mode of the posterior measure μ^s.

The task of minimizing the functional in (2.42) corresponds to the solution of the classical approach to inverse problems, which is the least-squares approach. The addition of the quadratic penalization term in (2.42) is known as *Tikhonov regularization*, and the functional $\mathcal{J}^s$ is called the *regularized misfit functional*, or simply the misfit functional. Within the Bayesian framework, this has a meaningful interpretation as a result of using Gaussian prior measures.

Theorem 5.4 of Stuart (2010) states that if $\mathcal{G}$ satisfies Assumptions 2.3.1, and $\mu_0(\mathcal{H}_m) = 1$, then a minimizer of $\mathcal{J}^s$ exists and lies in the Cameron-Martin space, E. Note that the same conditions assumed here are the conditions which lead to the well-definedness of the posterior measure. Therefore, in practical applications, the existence of a minimizer is not usually a concern. However, in a nonlinear inverse problem, the misfit functional is non-convex, which implies that $\mathcal{J}^s$ possesses several local minima. This fact implies that merely seeking a point estimate of the posterior measure may not sufficiently characterize the posterior measure.

Sampling approach

This leads to an effort to characterize the posterior measure more fully through sampling, so that we can also perform uncertainty quantification. Here, a realization drawn from the posterior measure represents a possible vector of parameters. One way to perform sampling is to use Markov chain Monte Carlo (MCMC) methods.

The basic idea of MCMC methods is to generate a sequence of draws that follows a target distribution μ. A commonly used method is the Metropolis-Hastings method. A key ingredient in the Metropolis-Hastings method is a Markov transition kernel, $q(m, \mathrm{d}\psi)$, which is a probability measure on $\mathcal{H}_m$. The Markov transition kernel is used to propose a move from a current state m_n that is distributed according to $q(m_n, \cdot)$. This proposal is then accepted or rejected based on the target distribution μ. The Metropolis-Hastings method is summarized as follows:

1. Let m_0 be a given initial state.

2. Generate a proposal state ψ_n from a Markov transition kernel $q(m, \mathrm{d}\psi)$.

3. Compute the acceptance probability, $\alpha(m, \psi)$, defined using the target distribution, μ.

4. Set the new state, $m_{n+1} = \psi_n$ with probability $\alpha(m_n, \psi_n)$ and $m_{n+1} = m_n$ otherwise.

To be precise, let $q(m, \mathrm{d}\psi)$ be a transition kernel. We define the measure ν on $\mathcal{H}_m \times \mathcal{H}_m$ by

$$\nu(\mathrm{d}m, \mathrm{d}\psi) = q(m, \mathrm{d}\psi)\mu(\mathrm{d}m). \tag{2.44}$$

We can also define the same measure by reversing the roles of m and ψ, i.e.,

$$\nu^T(\mathrm{d}\psi, \mathrm{d}m) = q(\psi, \mathrm{d}m)\mu(\mathrm{d}\psi). \tag{2.45}$$

Then, if ν^T is absolutely continuous with respect to ν, we can define

$$\alpha(m, \psi) = \min\left\{1, \frac{\mathrm{d}\nu^T}{\mathrm{d}\nu}(m, \psi)\right\}. \tag{2.46}$$

Define a random variable $\gamma(m, \psi)$, by

$$\gamma(m, \psi) = \begin{cases} 1 & \text{with probability } \alpha(m, \psi) \\ 0 & \text{otherwise.} \end{cases} \tag{2.47}$$

Then, given a proposal ψ_n, generated by the transition kernel q, we define

$$m_{n+1} = \gamma(m_n, \psi_n)\psi_n + (1 - \gamma(m_n, \psi_n))m_n, \quad n \geq 1. \tag{2.48}$$

If the random proposal ψ_n and the random variable $\gamma(m_n, \psi_n)$ are independent for all n, then the resulting sequence $\{m_n\}_{n=0}^{\infty}$ is a Markov chain. Furthermore, this Markov chain converges weakly to the target distribution μ, as explained in the following theorem.

Theorem 2.3.6 (Theorem 5.1 of Stuart (2010) and Proposition 3.11 of Kaipio and Somersalo (2006))**.** *Under the above assumptions, the Markov chain as defined in* (2.48) *is invariant for μ, i.e., if $m_0 \sim \mu$, then $m_n \sim \mu$ for all $n \geq 0$. Furthermore, if $\{m_n\}_{n=0}^{\infty}$ is ergodic, then for any continuous bounded function $f : \mathcal{H}_m \to \mathbb{R}$, any $M \geq 0$, and for μ_0-almost surely,*

$$\frac{1}{N}\sum_{n=0}^{N} f(m_{n+M}) \to \int_{\mathcal{H}_m} f(m)\ \mu(\mathrm{d}m) \quad \textit{as } N \to \infty. \tag{2.49}$$

The following is an example of a Metropolis-Hastings MCMC method.

Example 2.3.7 (Cotter *et al.* (2009), Example 5.2 and 5.3 of Stuart (2010))**.** Suppose we want to generate a sample from the posterior measure μ, defined by (2.6). The dependence on s is not indicated as it is not relevant to our discussion here. Since $\mu_0 = \mathcal{N}(0, \mathcal{C}_0)$, we consider proposals generated using

$$\psi_n = \sqrt{1-\beta^2} m_n + \beta \xi_n, \quad \xi_n \sim \mathcal{N}(0, \mathcal{C}_0), \tag{2.50}$$

where $\beta \in [0, 1]$. The acceptance probability is

$$\alpha(m, \psi) = \min\left\{1, \exp(\Phi(m_n) - \Phi(\psi_n))\right\}. \tag{2.51}$$

Note that here, a proposal ψ_n is always accepted if $\Phi(\psi_n) \leq \Phi(m_n)$. This is known as a random walk. Observe that it may seem more natural to define the proposals as

$$\psi_n = m_n + \beta \xi_n, \tag{2.52}$$

with acceptance probability

$$\alpha(m, \psi) = \min\left\{1, \exp(\mathcal{J}(m_n) - \mathcal{J}(\psi_n))\right\}. \tag{2.53}$$

In finite dimensions this will sample the target distribution μ^s. However, in infinite dimensions, the measure ν^T will, in general not be absolutely continuous with respect to ν, and therefore the expression in (2.46) is not well-defined. Intuitively, if $m \sim \mathcal{N}(0, \mathcal{C}_0)$ and the proposals, ψ are generated using (2.50), then $\psi \sim \mathcal{N}(0, \mathcal{C}_0)$, and therefore the prior measure is preserved. This is not the case if the proposals were generated using (2.52), since then we will have $\psi \sim \sqrt{1+\beta^2}\mathcal{N}(0, \mathcal{C}_0)$.

In practice, two main questions to consider are the size of the step, β, and the length of the Markov chain.

Consider the following example, based on Figure 3.14 of Kaipio and Somersalo (2006) to illustrate the effect of β on a Markov chain. Suppose we would like to sample from the probability density

$$\pi(x, y) \propto \exp\left(-10(x^2 - y)^2 - \left(x - \frac{1}{4}\right)^4\right). \tag{2.54}$$

Figure 2.4 shows the plot of this probability density. We use the standard normal distribution as the proposal density, and generate samples for different values of β and a Markov chain length of 5000. The samples are shown in Figure 2.5 are sampled for $\beta = 0.01$, 0.1, 0.75, and 1. For small values of β, we see that the support of

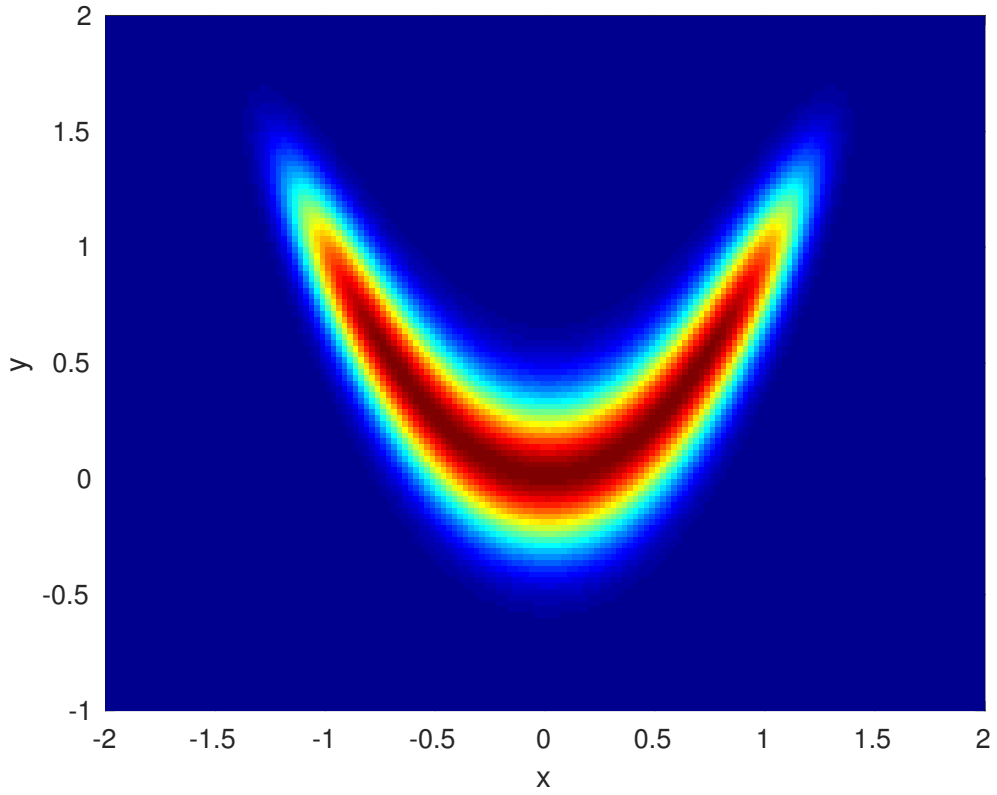


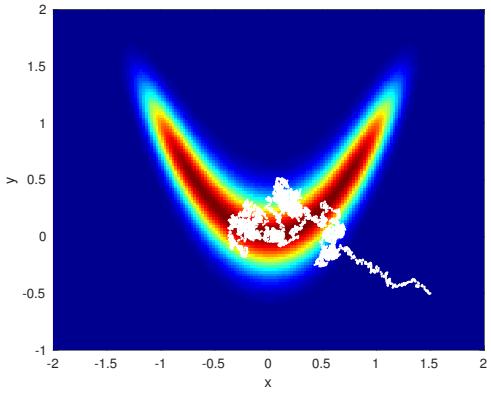
Figure 2.4: Plot of the probability density in (2.54) to be sampled from.

the probability density has not been explored fully (Figure 2.5(a)). This is because small steps are being taken to get from the initial point into the central support itself. The small value of β yields a high acceptance rate, but would require a longer chain to explore the probability density. The process of searching for the support is known as the *burn-in*.

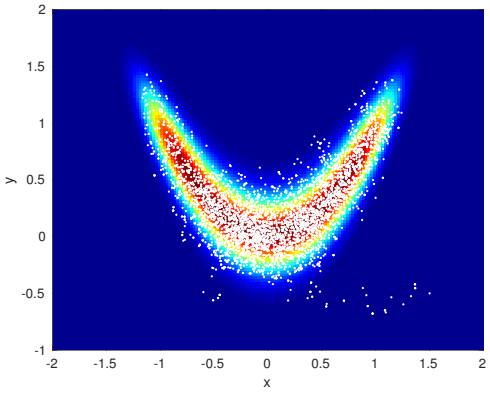
As the value of β increases, the burn-in period decreases, and the central support of the probability density can be explored more fully (Figures 2.5(b)–(c)). However, if the value of β is too large, the acceptance rate drops, and many proposals will be rejected. This can be seen in Figure 2.5(d), where there are fewer points exploring the space compared to the sample in Figure 2.5(c).

Currently, there is no known theory of how a value of β can be chosen so that the support of a probability density can be explored fully. However, Robert and Casella (2004, p. 316) gives a guide of an acceptance rate of equal to 50% for problems of dimensions 1 or 2, and an acceptance rate of close to 25% for problems of higher dimensions.

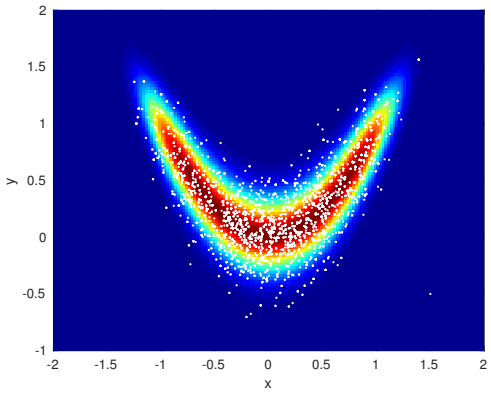
Next, we address the question of the length of the Markov chain required for convergence. Here, we turn to the central limit theorem (see for instance, Kaipio



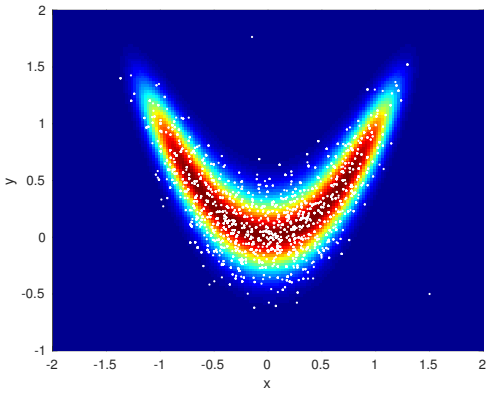
(a) $\beta = 0.01$. Acceptance rate: 97%.



(b) $\beta = 0.1$. Acceptance rate: 82%.



(c) $\beta = 0.75$. Acceptance rate: 25%.



(d) $\beta = 1$. Acceptance rate: 18%.

Figure 2.5: Samples from the probability density in (2.54) for different values of β, and Markov chain length of 5000.

and Somersalo (2006, Appendix B)), which says that the convergence is of order $O(1/\sqrt{N})$, where N is the size of the sample. However, the central limit theorem requires that the sample is independent. In practice, consecutive draws in a Markov chain will show correlation. Therefore, it is common to pick a set of draws at periodic intervals as the sample from a probability density.

Here, a higher value of β will give less correlation between consecutive draws, and hence requires a shorter periodic interval for draws to be picked. Conversely, a lower value of β gives higher correlation between consecutive draws and therefore, requires a longer periodic interval.

One way to check convergence is to evaluate the chain mean and standard deviation as the length increases. As the chain length increases, the mean and standard deviation would stabilize. Practically, this can be checked by verifying if consecutive chain means and standard deviations were within a specified tolerance for a specified length of the chain. If they are, convergence is assumed and the sampling stops. Note that in some applications, the dimensions of the problem is so high that it is not feasible to check each entry of a vector. In such cases, we would check this for several components of the vector to obtain a heuristic test for convergence.

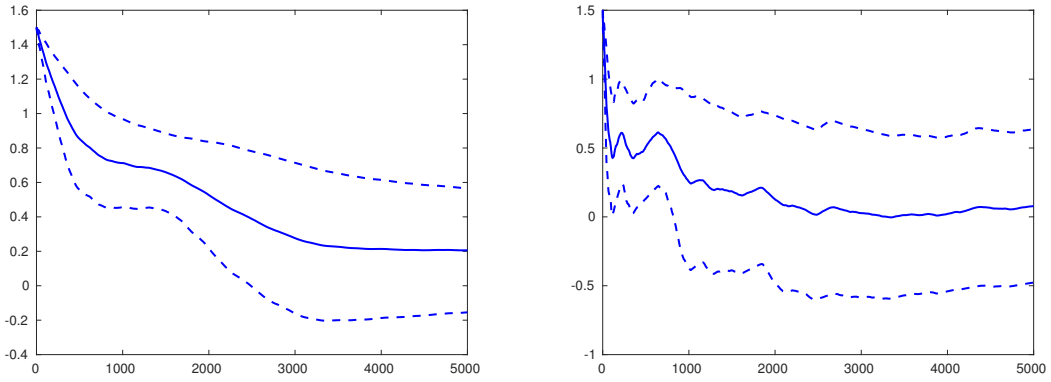
As in illustration, we consider our previous example. Figures 2.6(a)–(d) show the diagnostics for the samples generated in the previous example. Note that the mean and standard deviation has yet to stabilize for $\beta = 0.01$ and $\beta = 0.1$, while they stabilize quickly in the case of $\beta = 0.75$ and $\beta = 1$. Note also that while it appears that the mean has stabilized for $\beta = 0.01$, convergence is not yet achieved, as if we were to continue sampling, it will eventually stabilize at 0, which is the mean of the first component in (2.54), as can be seen in Figures 2.6(c)–(d).

Another diagnostic one could compute is the time-lag autocovariance, which evaluates the independence of draws. Again, in practice, this diagnostic is only computed for several components of a high-dimensional vector. For more details on MCMC sampling, we refer the reader to Kaipio and Somersalo (2006), Robert and Casella (2004) and references therein.

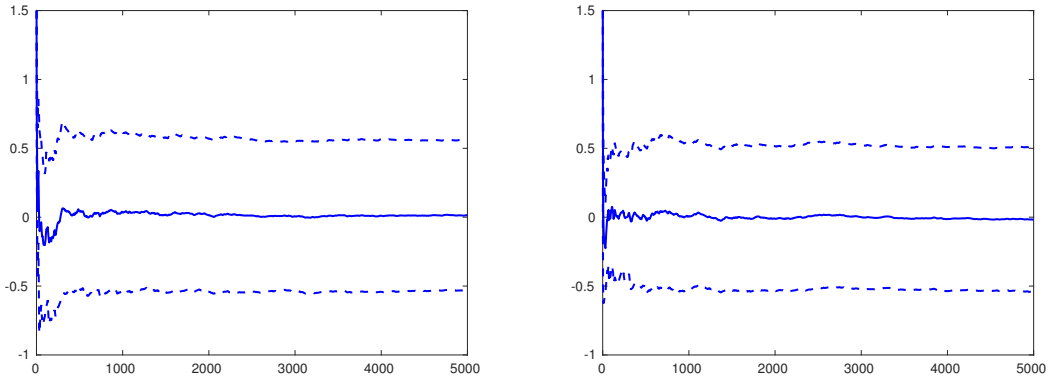
2.4 Finite-Dimensional Approximation

The implementation of algorithms requires a discretization of functions and operators. In this section, we will consider such discretizations.

We recall our general framework of inverse problems that applies in infinite as well as finite dimensions, where the Radon-Nikodym derivative of the posterior measure



(a) $\beta = 0.01$. (b) $\beta = 0.1$.



(c) $\beta = 0.75$. (d) $\beta = 1$.

Figure 2.6: Chain means and standard deviations for the first component of the samples from the probability density in (2.54) for different values of β, and Markov chain length of 5000. The abcissa is the length of the chain while the ordinate is the diagnostic of the chain. The solid line shows the chain mean, while the dotted lines show the departure of the chain mean by one standard deviation.

with respect to the prior measure is given by

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0} = \frac{1}{Z}\exp(-\Phi(m)). \tag{2.55}$$

Here, μ_0 is the prior measure, $\mu_0 = \mathcal{N}(0, \mathcal{C}_0)$, $\Phi(m)$ is the potential of the likelihood, given by (2.5), Z is the normalization constant given by (2.7), and μ is the posterior measure with μ_0 as the reference measure. Note that the dependence on s is not indicated here, since it has no relevance in our discussion.

Let $\Phi^N(m)$ denote the potential which approximates $\Phi(m)$ on some appropriate N-dimensional discretization. Note that since $\mathcal{H}_s$ is finite-dimensional, it follows that the discretization $\Phi^N(m)$, given by

$$\Phi^N(m) = \frac{1}{2}\left\|s - \mathcal{G}^N(m)\right\|_s^2, \tag{2.56}$$

depends solely on the numerical scheme used to solve the forward problem.

Then, $\Phi^N(m)$ induces an approximation of the posterior measure, denoted by μ^N, to the exact posterior measure, μ. We have

$$\frac{\mathrm{d}\mu^N}{\mathrm{d}\mu_0} = \frac{1}{Z^N}\exp(-\Phi^N(m)), \tag{2.57}$$

where Z^N is the normalization constant

$$Z^N = \int_{\mathcal{H}_m} \exp(-\Phi^N(m))\ \mathrm{d}\mu_0(m). \tag{2.58}$$

2.4.1 Accuracy and Convergence

The convergence of the approximated posterior measure μ^N to the exact posterior measure μ depends on the convergence of the approximated potential Φ^N to the exact potential Φ. In particular, we have the following theorem.

Theorem 2.4.1 (Corollary 4.9 of Stuart (2010))**.** *Assume that the measures μ and μ^N are absolutely continuous with respect to the prior measure μ_0, with $\mu_0(\mathcal{H}_m) = 1$, and the Radon-Nikodym derivatives of μ and μ^N with respect to μ_0 are given by* (2.55) *and* (2.57) *respectively. Assume further that $\mathcal{G}$ is approximated by $\mathcal{G}^N$ with the property that for $\varepsilon > 0$, there exists $K' = K'(\varepsilon) > 0$ such that*

$$\left\|\mathcal{G}(m) - \mathcal{G}^N(m)\right\|_s \leq K' \exp(\varepsilon \left\|m\right\|_m^2)\psi(N), \tag{2.59}$$

where $\psi(N) \to 0$ as $N \to \infty$. If $\mathcal{G}$ and $\mathcal{G}^N$ satisfy Assumption 2.3.1 uniformly in N, then there is a constant C, independent of N such that

$$d_{Hell}(\mu, \mu^N) \leq C\psi(N). \tag{2.60}$$

Consequently, the expectation under μ and μ^N of any polynomially bounded function $f : \mathcal{H}_m \to E$ are $O(\psi(N))$ close. In particular, the mean and covariance operator are $O(\psi(N))$ close.

The proof of Theorem 2.4.1 can be found in Stuart (2010).

2.4.2 Discretization

Recall the misfit functional as given in (2.42), which is

$$\mathcal{J}(m) = \Phi(m) + \mathcal{J}_p(m), \tag{2.61}$$

where $\mathcal{J}_p(m)$ is the functional defined by

$$\mathcal{J}_p(m) = \frac{1}{2} \|m\|^2_{\mathcal{C}_0}, \tag{2.62}$$

where we re-emphasize that we have assumed without loss of generality that $m_0 = 0$.

We approximate $\Phi(m)$ and $\mathcal{J}_p(m)$ on an N-dimensional space, denoted by $\Phi^N(m)$ and $\mathcal{J}_p^N(m)$ respectively. Then, the approximated misfit functional, denoted by $\mathcal{J}^N(m)$ is given by

$$\mathcal{J}^N(m) = \Phi^N(m) + \mathcal{J}_p^N(m). \tag{2.63}$$

As mentioned before, the discretization of the likelihood potential depends on the numerical scheme used to solve the forward problem. The regularization term, $\mathcal{J}_p(m)$ may be discretized using a different numerical scheme. For instance, if D is a bounded, open subset of $\mathbb{R}^d$, where $d = 1$, 2, or 3, and $\mathcal{H} \subseteq L^2(D)$, equipped with the L^2-inner product, then we have

$$\mathcal{J}_p(m) = \frac{1}{2} \int_D \int_D m(x)(\mathcal{L}_0 m)(y) \, \mathrm{d}y \, \mathrm{d}x. \tag{2.64}$$

Under an appropriate discretization, this gives

$$\mathcal{J}_p^N(m) = \frac{1}{2} m_N^T L_N m_N, \tag{2.65}$$

where L_N is an appropriate precision matrix and m_N is an appropriate vector.

As an example, suppose we specify the prior precision to be

$$\mathcal{L}_0 m = \alpha m - \beta \nabla^2 m, \tag{2.66}$$

where $\alpha > 0$ and $\beta > 0$. This is the modified Helmholtz operator, as introduced in (2.25)). Then, a finite difference scheme on a uniform grid with grid size h will yield a matrix L_N of the form

$$L_N = \alpha h^d \mathbf{1} + \beta h^{d-2} A, \tag{2.67}$$

where **1** is the appropriate identity matrix and A is the discretized negative Laplacian.

On the other hand, under a finite element discretization with piecewise linear basis functions, we would have

$$L_N = \alpha h^d M + \beta h^{d-2} A, \tag{2.68}$$

where M is now the appropriate mass matrix. A more detailed discussion on the finite element discretization can be found in Tan *et al.* (2013).

2.5 Concluding Remarks

In this chapter, we presented the framework of Bayesian inverse problems in infinite dimensions, which will form the basis of our approach for the inverse problems studied here. Two clear benefits of working within the Bayesian framework are that it allows us to keep track of our assumptions more easily, and also allows us to perform uncertainty quantification.

In the past, the Bayesian framework for inverse problems has been developed within a finite-dimensional context (see for instance, Kaipio and Somersalo, 2006). While we understand that computational resources are finite, working within an infinite-dimensional framework brings some advantages. It allows us to be clear about the assumptions on the function spaces we are working in, and therefore the type of solutions we might expect, for instance, in terms of continuity and boundedness of the solutions. Furthermore, the choice of norms and inner products are guided by the function space used. This, in turn, allows us to discretize the problem in a consistent manner, ensuring convergence in the limit (at least, in model problems). Another distinct advantage of an infinite-dimensional setting can be seen from Example 2.3.7, where it also guides our sampling algorithms. More detailed work regarding modifying our finite-dimensional sampling algorithms so that they also work in infinite dimensions can be found in Cotter *et al.* (2013).

While we acknowledge that errors are incurred in the process of discretization, we will use the approach in Kaipio and Somersalo (2005) in practical applications, and we will ignore errors arising from the discretization and assume that the grid we are using is sufficiently fine that it can be used as the reference model. In most cases, we will be working under the perfect model scenario, where the same model used to generate the observations is also used for the inversion. Doing this allows us to evaluate the effectiveness of our methods for solving the inverse problem without having to worry about the errors arising from discretizing the forward problem.

With regard to the notion "perfect model scenario", a term introduced in Smith (2006), and the associated term "inverse crime", as coined in (Kaipio and Somersalo, 2006, Chapter 1), two tasks arise in evaluating the effectiveness of algorithms in solving inverse problems. The first is the evaluation of the accuracy and efficiency of numerical methods for computing characteristic properties of the posterior measure. The second is the comparison of such characteristic properties with data obtained from a real system.

When these two tasks are not distinguished from each other, then indeed, one is committing an inverse crime. Thus, it is to be emphasized that in this thesis, the main aim is to perform the former task rather than the latter, which is to evaluate the accuracy and efficiency of numerical methods for summarizing a posterior measure. In this case, it is quite appropriate to work within the perfect model scenario so that one can use knowledge of the details of the observation process to assess the accuracy and efficiency of the method of producing the summary.

Chapter 3

Construction of Priors

3.1 Introduction

A prior measure forms the starting point in solving an inverse problem, in the sense that it is a quantification of our knowledge of the parameters prior to making observations. Then, the observations and forward model is used to update the knowledge of the parameters in the prior. This gives the posterior measure, which we can then use to extract useful information. Hence, specifying a prior measure is an important task in solving an inverse problem.

A prior measure is also central in the framework of inverse problems as it plays many different roles in different contexts. In a deterministic framework for inverse problems, a prior acts as a regularizer of the minimization problem. Consider, for instance, the misfit function

$$\mathcal{J}(m_1, m_2) = \frac{p}{2}(m_1 + m_2 - s)^2, \tag{3.1}$$

for some $p > 0$ and $s \in \mathbb{R}$, as well as parameters m_1 and m_2. The solution of an inverse problem in the deterministic framework is the global minimum of this misfit function. The gradient of $\mathcal{J}$ is given by

$$\nabla\mathcal{J}(m_1, m_2) = \begin{pmatrix} p(m_1 + m_2 - s) \\ p(m_1 + m_2 - s) \end{pmatrix}. \tag{3.2}$$

Setting $\nabla\mathcal{J}(\breve{m}_1, \breve{m}_2) = 0$, we have the linear system

$$\breve{m}_1 + \breve{m}_2 - s = 0, \tag{3.3}$$

$$\breve{m}_1 + \breve{m}_2 - s = 0, \tag{3.4}$$

which has infinitely many solutions. However, if we regularize the problem by minimizing instead the function

$$\mathcal{J}_r(m_1, m_2) = \frac{p}{2}(m_1 + m_2 - s)^2 + \frac{\varepsilon}{2}(m_1 - m_{10})^2 + \frac{\varepsilon}{2}(m_2 - m_{20})^2, \tag{3.5}$$

for some $m_{10} \in \mathbb{R}$, $m_{20} \in \mathbb{R}$ and $\varepsilon > 0$, then the gradient is now given by

$$\nabla \mathcal{J}_r(m_1, m_2) = \begin{pmatrix} p(m_1 + m_2 - s) + \varepsilon(m_1 - m_{10}) \\ p(m_1 + m_2 - s) + \varepsilon(m_2 - m_{20}) \end{pmatrix}. \tag{3.6}$$

Setting $\nabla \mathcal{J}_r(\breve{m}_1, \breve{m}_2) = 0$ and solving the resulting linear system, we have the unique solution

$$\breve{m}_1 = \frac{p(m_{10} - m_{20} + s) + \varepsilon m_{10}}{2p + \varepsilon}, \quad \breve{m}_2 = \frac{p(-m_{10} + m_{20} + s) + \varepsilon m_{20}}{2p + \varepsilon}. \tag{3.7}$$

It is interesting to note that if we take the limit as $\varepsilon \to 0$, we have

$$\lim_{\varepsilon \to 0} \breve{m}_1 = \frac{m_{10} - m_{20} + s}{2}, \quad \lim_{\varepsilon \to 0} \breve{m}_2 = \frac{-m_{10} + m_{20} + s}{2}, \tag{3.8}$$

where the terms m_{10} and m_{20} have not vanished. Here, we see the importance of the prior in improving the conditioning of the minimization problem.

In the theory of infinite-dimensional Bayesian inverse problems, it is well-known that there is no natural generalization of the Lebesgue measure. In this framework, the prior measure plays the crucial role of being the reference measure.

Therefore, we can conclude that the prior measure is an important ingredient in the framework of inverse problems, and the choice of prior measure is an important decision to make in the solution of an inverse problem. The discussion in this chapter is on the choice of priors. In particular, we will focus on Gaussian prior measures, where the choice of prior is then determined by the choice of the precision operator. Without loss of generality, we may assume that the prior measure has zero mean, since any prior with nonzero mean can be translated into a prior measure with zero mean.

We also remark that in general, it is possible to specify non-Gaussian prior measures as well. In Chapter 3 of Kaipio and Somersalo (2006), other examples of priors are discussed. In the infinite-dimensional setting, some theory on Besov priors has been developed. We refer the interested reader to Dashti *et al.* (2012) for further details.

3.2 Modified Helmholtz Measures

In this section, we will focus on Gaussian measures with zero mean and a modified Helmholtz operator as its precision. This will be known as a *modified Helmholtz measure*. As mentioned in Chapter 2, a modified Helmholtz measure can be specified as a prior in solving an inverse problem as its precision operator satisfies Assumption 2.2.1. The discussion that follows concerns the specification of the hyperparameters in a modified Helmholtz measure.

3.2.1 1-D modified Helmholtz measures

Consider a one-dimensional modified Helmholtz operator, defined by

$$\mathcal{L}m = a(b^2 m - \nabla^2 m), \tag{3.9}$$

where $a > 0$, $b > 0$. Recall that on an infinite domain with vanishing far-field conditions, the Green's function, which is also the correlation function of a 1-D modified Helmholtz operator, is given by

$$c(x, y) = g(r) = \frac{1}{2ab} e^{-br}, \tag{3.10}$$

where $r = |x - y|$. We define the pointwise variance as $\sigma^2 := c(x, x) = g(0) = 1/(2ab)$, while the correlation length, λ is defined as the integral

$$\lambda = \frac{1}{g(0)} \int_0^\infty g(r) \, \mathrm{d}r = \frac{1}{b}. \tag{3.11}$$

We may take a and b as our hyperparameters, and we can express them in terms of σ and λ as

$$a = \frac{\lambda}{2\sigma^2}, \quad b = \frac{1}{\lambda}. \tag{3.12}$$

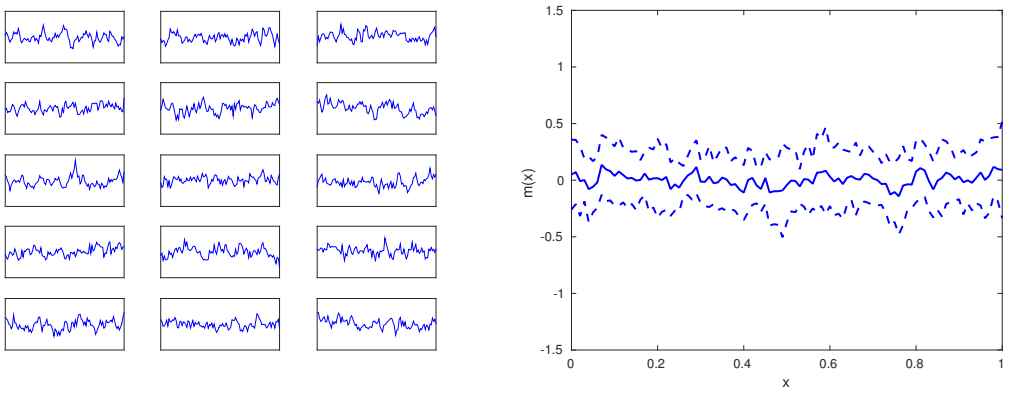
In order to understand better the nature of these measures, we sample from them. Figures 3.1, 3.2 and 3.3 show 15 realizations drawn from the modified Helmholtz prior with $\lambda = 0.02$ and $\sigma = 0.25$, 0.5 and 0.75 respectively.

Sampling was performed on the domain $[0, 1]$ on a uniform grid with spacing of length h and with $N + 1$ grid points, where $h = 1/N$. For this simulation, we chose $N = 100$, and the modified Helmholtz operator was discretized using finite differences with homogeneous Neumann boundary conditions. If L is the discretized modified Helmholtz operator, and A is the Cholesky factor of L so that $L = AA^T$, a draw, ζ was generated using

$$\zeta = A^{-T}\eta, \tag{3.13}$$

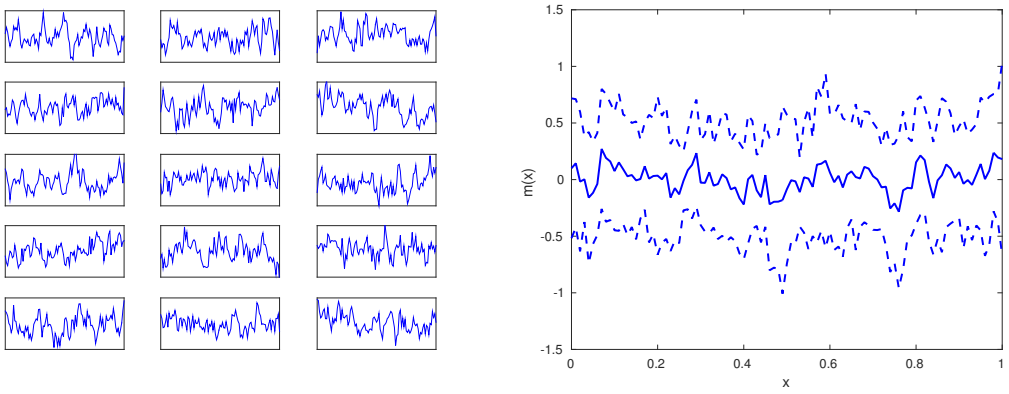
where η is a random vector drawn from the standard normal distribution. Note that this is an exact sample since we have

$$\mathbb{E}[\zeta] = A^{-T}\mathbb{E}[\eta] = 0, \quad \text{and} \quad \mathbb{E}[\zeta\zeta^T] = A^{-T}\mathbb{E}[\eta\eta^T]A^{-1} = L^{-1}. \tag{3.14}$$



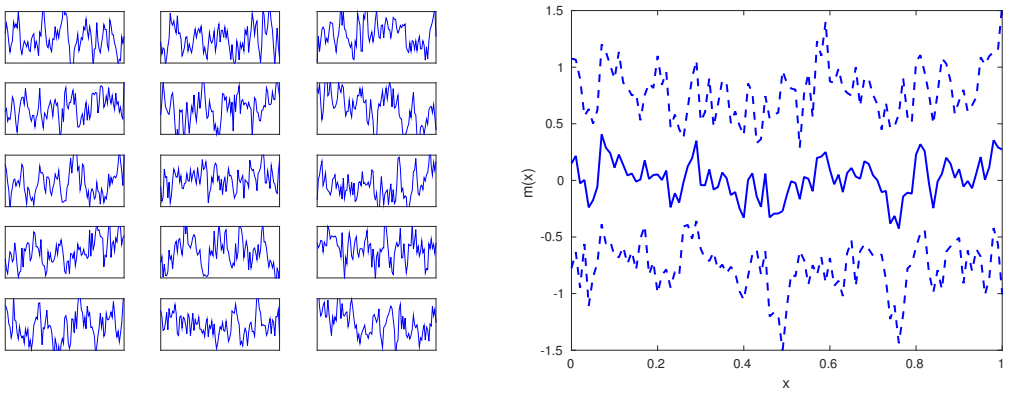
(a) (b)

Figure 3.1: (a) 15 realizations of 1-D modified Helmholtz measures with $\lambda = 0.02$, $\sigma = 0.25$. The realizations were all plotted on the same scale, with the same ordinate and abscissa as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 3.2: (a) 15 realizations of 1-D modified Helmholtz measures with $\lambda = 0.02$, $\sigma = 0.5$. The realizations were all plotted on the same scale, with the same ordinate and abscissa as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 3.3: (a) 15 realizations of 1-D modified Helmholtz measures with $\lambda = 0.02$, $\sigma = 0.75$. The realizations were all plotted on the same scale, with the same ordinate and abscissa as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

Figures 2.1, 2.2 and 2.3 show 15 realizations drawn from the modified Helmholtz measure with $\sigma = 0.02$ and $\lambda = 0.02$, 0.1 and 0.5 respectively.

In Figure 2.1, when the correlation length is short, a realization of the prior resembles white noise. When the correlation length is increased to 0.1, we begin to see the correlation structure between grid points, although the draws still contain some roughness (Figure 2.2). When $\lambda = 0.5$, each random draw becomes smoother (Figure 2.3).

3.2.2 2-D modified Helmholtz measures

In two spatial dimensions, the correlation function of the modified Helmholtz measure with precision operator defined in (3.9) is given by

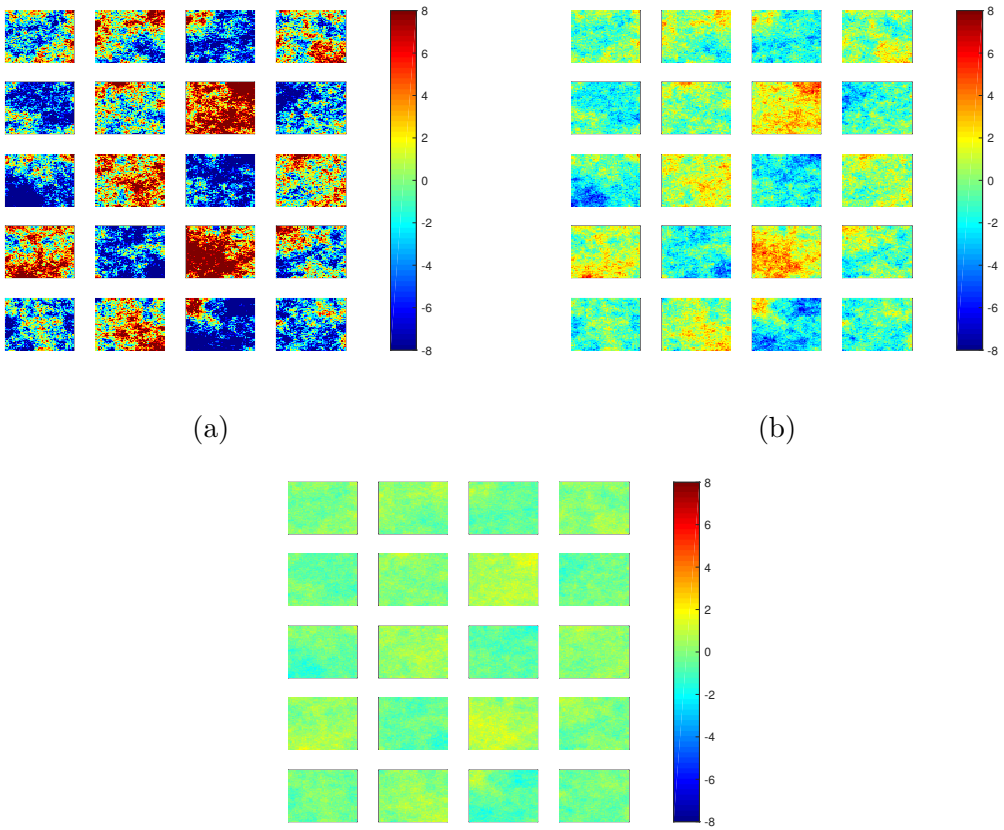
$$c(x, y) = g(r) = \frac{1}{2\pi a} K_0(br), \tag{3.15}$$

where K_n is the nth modified Bessel function of the second kind and $r = |x - y|$. The function $K_0(br)$ decays as $r \to \infty$. Here, the hyperparameter b controls the correlation length while a controls the variance.

Once again, we sample from this measure. This was performed on the unit square $[0, 1] \times [0, 1]$ on a uniform grid spacing of length h and $N + 1$ grid points in each direction, where $h = 1/N$. For this simulation, we chose $N = 100$.

Figure 3.4 shows 20 draws from modified Helmholtz measures with $b = 1$ and $a = 0.1$, 1 and 10 respectively. Note that as a increases, the fluctuations from the

prior mean in the ensemble reduce. Note also that the pointwise variance decreases by the same factor by which a increases (Figure 3.5), indicating that the relationship between $1/a$ and the variance is linear.



(a) (b)

(c)

Figure 3.4: 20 realizations of 2-D modified Helmholtz measures on the unit square with $b = 1$ and (a) $a = 0.1$, (b) $a = 1$, (c) $a = 10$.

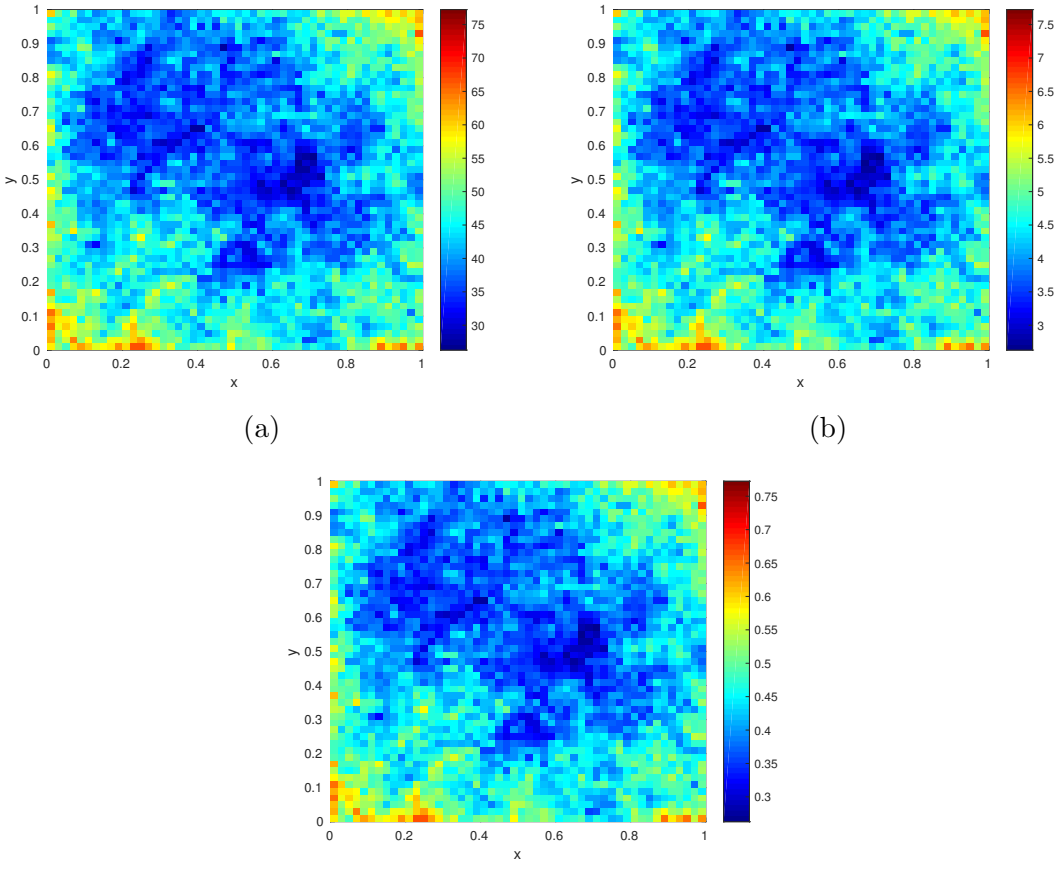
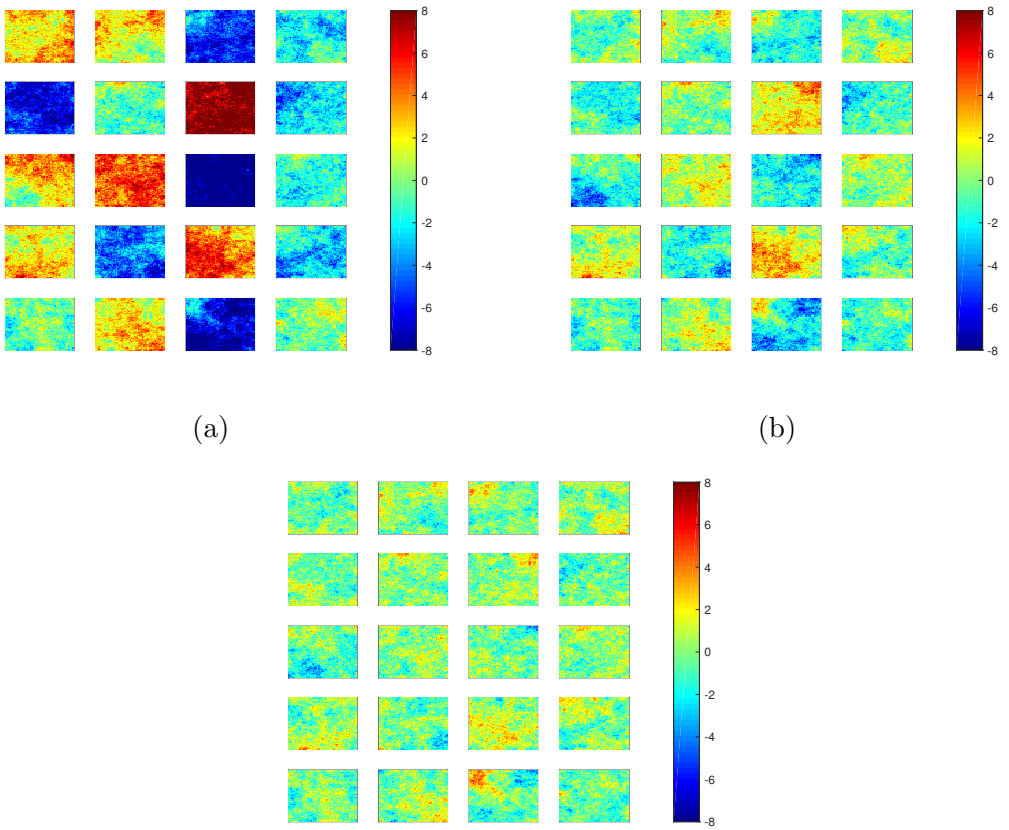


Figure 3.5: Pointwise variance of the realizations of 2-D modified Helmholtz measures on the unit square with $b = 1$ and (a) $a = 0.1$, (b) $a = 1$, (c) $a = 10$.



(a) (b)

(c)

Figure 3.6: 20 realizations of 2-D modified Helmholtz measures on the unit square with $a = 1$ and (a) $b = 0.1$, (b) $b = 1$, (c) $b = 10$.

Figure 3.6 shows 20 draws from modified Helmholtz measures with $a = 1$ and $b = 0.1$, 1 and 10 respectively. Note that as b increases, the correlation structure decreases. This is an indication that b is inversely proportional to the correlation length. As b increases, the first term in (3.9) dominates, so that the precision operator looks like a multiple of the identity map, which is the white noise operator. Therefore, each draw will resemble white noise, as can be seen here.

In some cases, one may wish to use anisotropic priors, where the correlation in one direction may be stronger than in another. In two-dimensional systems, we may generalize the modified Helmholtz measure to incorporate such anisotropy. The modified Helmholtz operator then generalizes to

$$\mathcal{L}m = a\Delta_B \left(m - \nabla \cdot (B^{-2}\nabla m)\right), \tag{3.16}$$

where $a > 0$ and Δ_B is the determinant of a symmetric, positive definite matrix B, with entries

$$B = \begin{pmatrix} b_x & c \\ c & b_y \end{pmatrix}. \tag{3.17}$$

If we set $c = 0$, then the correlation in the horizontal and vertical directions are independent of each other. In this case, we have

$$\mathcal{L}m = ab_x b_y \left(m + \frac{1}{b_x^2}\frac{\partial^2 m}{\partial x^2} + \frac{1}{b_y^2}\frac{\partial^2 m}{\partial y^2}\right). \tag{3.18}$$

We define the quantity

$$\nu_B := b_x/b_y. \tag{3.19}$$

Then, we have

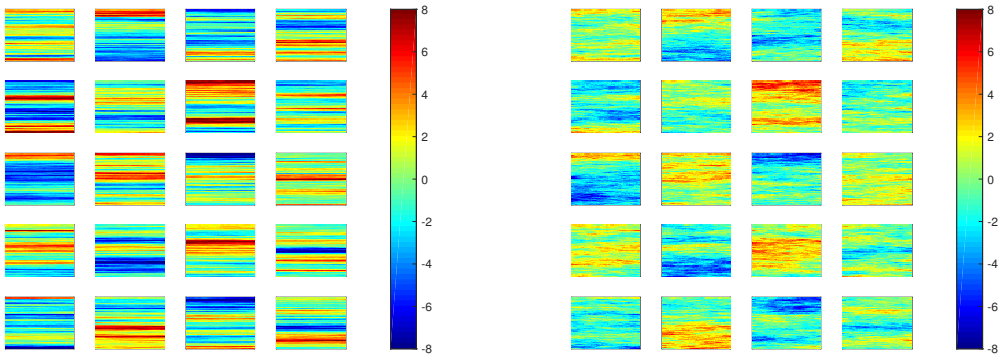
$$\mathcal{L}m = a\left(\Delta_B m + \frac{1}{\nu_B}\frac{\partial^2 m}{\partial x^2} + \nu_B\frac{\partial^2 m}{\partial y^2}\right). \tag{3.20}$$

Note that if $b_x = b_y = b$, this reduces to the isotropic prior case as in (3.9).

In practice, it is sufficient to consider anisotropic modified Helmholtz measures with $c = 0$ and varying b_x and b_y. Note that b_x and b_y are uniquely determined by the values of Δ_B and ν_B.

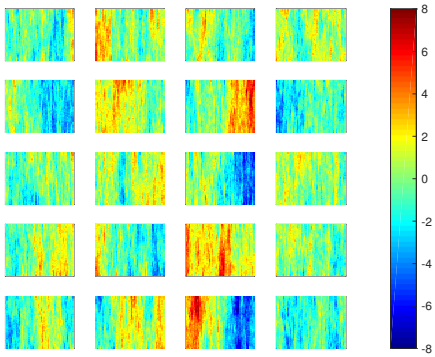
In order to illustrate how an anisotropic modified Helmholtz measure varies with the hyperparameters, we sample from these measures, varying Δ_B and ν_B.

Figure 3.7 shows 20 draws from the 2-D anisotropic modified Helmholtz measure on the unit square. We set $a = 1$, $\Delta_B = 0.1$ and vary ν_B. If $\nu_B > 1$, then the correlation in the x-direction is stronger than in the y-direction (Figure 3.7(a), (b)). Conversely, if $\nu_B < 1$, then the correlation in the y-direction is stronger than in the x-direction (Figure 3.7(c), (d)). Therefore, we see that the hyperparameter ν_B determines the relative strength of the correlation between the x- and y-directions.

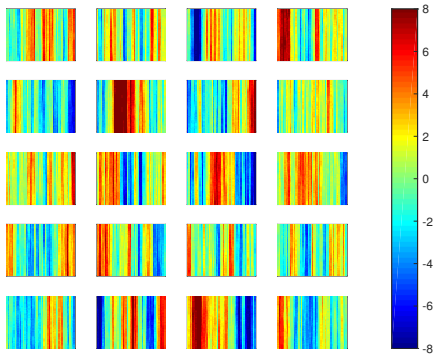


(a)

(b)

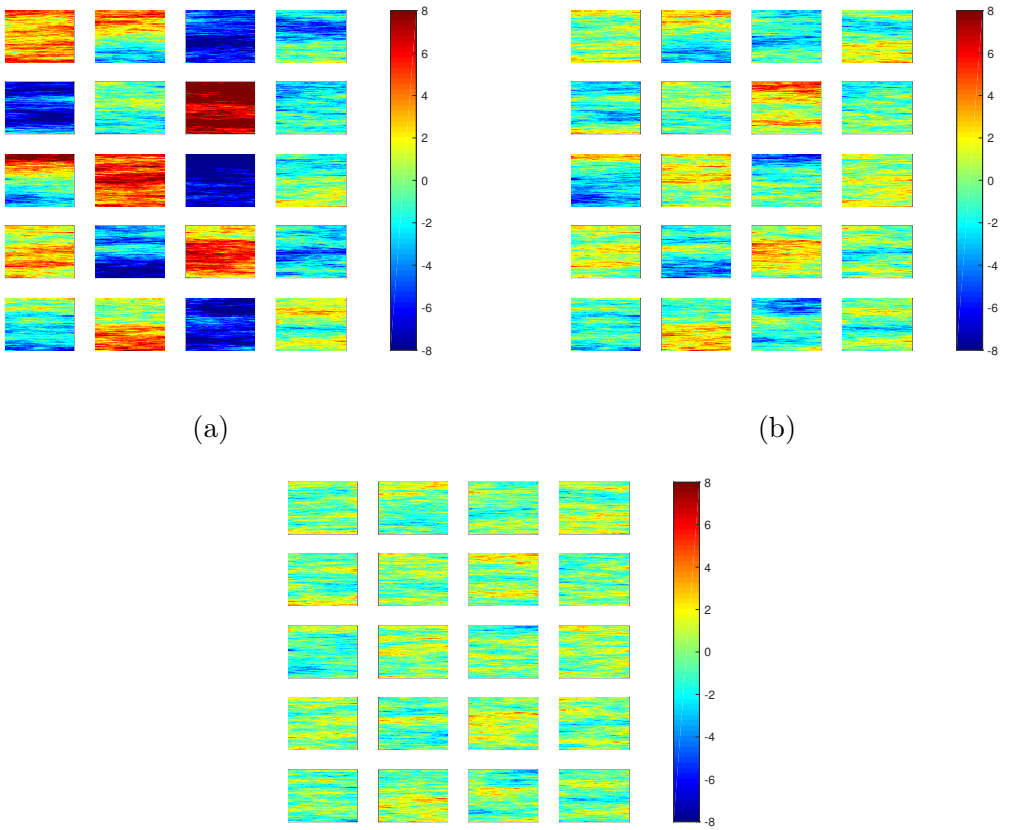


(c)



(d)

Figure 3.7: 20 realizations of 2-D anisotropic modified Helmholtz measures on the unit square with $a = 1$, $\Delta_B = 1$ and (a) $\nu_B = 0.01$, (b) $\nu_B = 0.1$, (c) $\nu_B = 10$, and (d) $\nu_B = 100$.



(a) (b)

(c)

Figure 3.8: 20 realizations of 2-D anisotropic modified Helmholtz measures on the unit square with $a = 1$, $\nu_B = 0.1$ and (a) $\Delta_B = 0.1$, (b) $\Delta_B = 1$, (c) $\Delta_B = 10$.

Figure 3.8 shows 20 draws from the 2-D anisotropic modified Helmholtz measure on the unit square. Here, we set $a = 1$, $\nu_B = 0.1$ and vary Δ_B. The hyperparameter Δ_B determines the overall strength of the correlation. In particular, a higher value of Δ_B gives a lower correlation, and vice versa. Hence, we call the quantity Δ_B the *absolute correlation*, while ν_B is called the *relative correlation*.

3.3 Precision Operators

In Chapter 2, it was mentioned that one class of precision operators that can be specified are second order positive definite elliptic operators. However, one may also define the precision operator through powers of an operator which satisfies Assumption 2.2.1.

Let D be a bounded open set, where $D \subseteq \mathbb{R}^d$, for $d = 1$, 2 or 3, and let $\mathcal{H} \subseteq L^2(D)$ be a Hilbert space. Let $\mathcal{L}$ be an operator which satisfies Assumption 2.2.1, defined on D. Then, we may define a precision operator $\mathcal{L}_0 := \mathcal{L}^\alpha$, for some $\alpha \in \mathbb{Z}$, and $\alpha \geq 1$. An important task which follows is to find conditions on α which ensure the regularity of samples drawn from this probability measure. It was proved in Stuart (2010) that a sufficient condition for this to hold is that we require $\alpha > d/2$.

An immediate consequence of this result is that it further guides our choice of prior precision operator in solving inverse problems. For instance, in one-dimensional problems, specifying a modified Helmholtz measure as the prior measure, i.e. taking $\alpha = 1$, is sufficient to ensure regularity. However, in inverse problems involving two or more spatial dimensions, this condition is not satisfied.

This becomes apparent when one tries to evaluate the pointwise variance for the 2-D isotropic modified Helmholtz measure, which is infinite, since

$$\lim_{r \to 0} g(r) = \lim_{r \to 0} \frac{1}{2\pi a} K_0(br) = \infty. \tag{3.21}$$

In the discretized case, the infinite variance manifests itself as we refine the grid, as the pointwise variance grows as a finer grid is used. Therefore, for an arbitrarily fine grid, the pointwise variance is finite, but in infinite dimensions, using such a prior for inverse problems on a 2-D system is not possible.

This issue can be easily remedied by taking $\alpha = 2$ in a two-dimensional system. In the isotropic case, the precision operator then becomes

$$\mathcal{L}_0 m = a^2(b^4 m - 2b^2 \nabla^2 m + \nabla^4 m), \tag{3.22}$$

which is the same form as the biharmonic-Helmholtz operator as in Farmer (2007), given by

$$\mathcal{L}_0 m = a^2(b^4 m - 2b^2 \cos(2t) \nabla^2 m + \nabla^4 m), \tag{3.23}$$

with $t = 0$. For this reason, we shall call the Gaussian measure with the precision operator in (3.22) the *biharmonic measure*.

In Farmer (2007), it was claimed that the Green's function for this operator can be reduced to the integral

$$g(r) = \frac{1}{2\pi a^2} \int_0^\infty \frac{kJ_0(kr)}{(k^2 + b^2)^2} \, \mathrm{d}k, \tag{3.24}$$

where J_0 is the Bessel function of the first kind. In fact, this integral is the Hankel transform of $1/(k^2 + b^2)^2$, and can be found in a table in Piessens (2000), which gives

$$g(r) = \frac{1}{2\pi a^2} \left(\frac{r}{2b} K_1(br) \right) = \frac{r}{4\pi a^2 b} K_1(br), \tag{3.25}$$

Now, the variance is well-defined, as we have

$$\sigma^2 = \lim_{r \to 0} g(r) = \frac{1}{4\pi a^2 b} \lim_{r \to 0} (rK_1(br)) = \frac{1}{4\pi a^2 b^2}. \tag{3.26}$$

The correlation length is then defined as

$$\lambda = \int_0^\infty \frac{g(r)}{g(0)} \, \mathrm{d}r = b \int_0^\infty rK_1(br) \, \mathrm{d}r = \frac{\pi}{2b}, \tag{3.27}$$

where the integral is evaluated using integration by parts and the fact that

$$\int_0^\infty K_0(r) \, \mathrm{d}r = \frac{\pi}{2}. \tag{3.28}$$

The hyperparameters a and b are therefore uniquely determined by λ and σ. We have

$$a = \frac{\lambda}{\pi^{3/2}\sigma}, \quad \text{and} \quad b = \frac{\pi}{2\lambda}. \tag{3.29}$$

In the anisotropic case with $c = 0$, we may define the correlation lengths and variance similarly. If λ_x and λ_y denote the correlation lengths in the x- and y-direction respectively, then we have

$$a = \frac{\sqrt{\lambda_x \lambda_y}}{\pi^{3/2}\sigma}, \quad b_x = \frac{\pi}{2\lambda_x}, \quad b_y = \frac{\pi}{2\lambda_y}. \tag{3.30}$$

The variance and correlation lengths are given by

$$\sigma^2 = \frac{1}{4\pi a^2 \Delta_B}, \quad \lambda_x = \frac{\pi}{2b_x}, \quad \lambda_y = \frac{\pi}{2b_y}, \tag{3.31}$$

and we also have

$$\Delta_B = b_x b_y = \frac{\pi^2}{4\lambda_x \lambda_y}, \quad \text{and} \quad \nu_B = \frac{b_x}{b_y} = \frac{\lambda_y}{\lambda_x}. \tag{3.32}$$

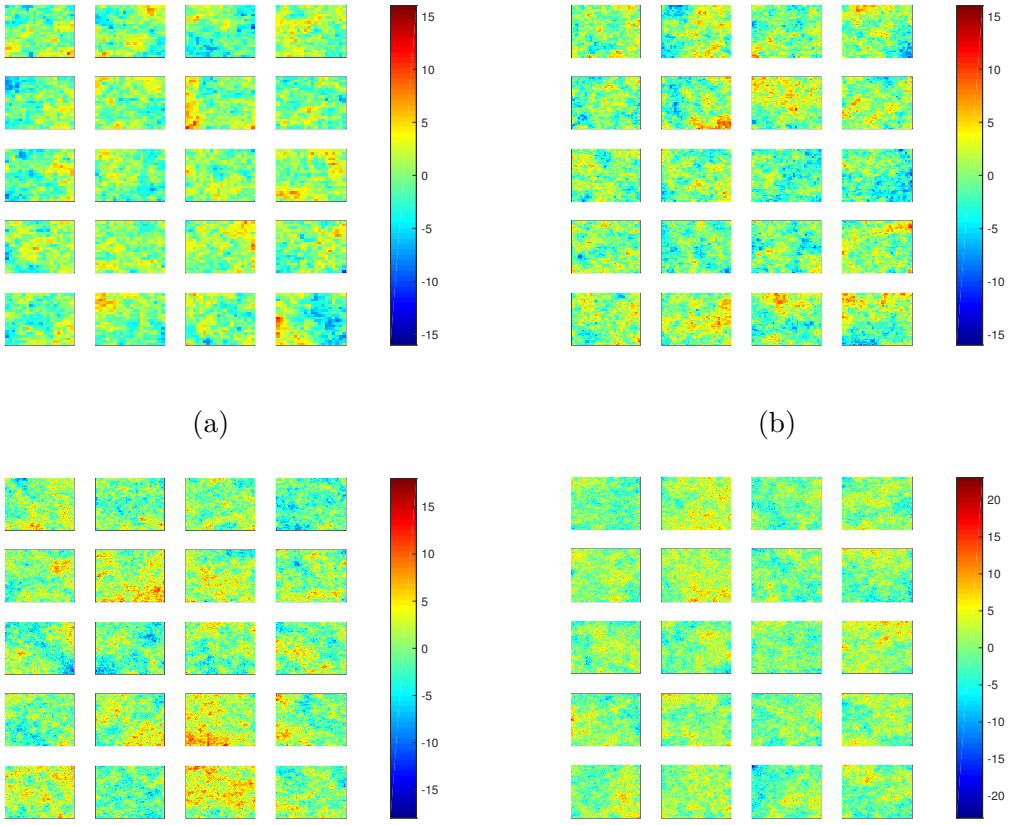


Figure 3.9: 20 realizations of 2-D isotropic modified Helmholtz measures on the unit square with $\sigma = 1$ and $\lambda = 0.5$ for (a) $N = 20$, (b) $N = 40$, (c) $N = 80$, and (d) $N = 160$.

We now demonstrate this effect by comparing draws from the modified Helmholtz measure and the biharmonic measure. The measures are specified through the specification of σ, λ_x and λ_y.

Figures 3.9 and 3.11 show 20 draws from a 2-D isotropic modified Helmholtz measure, as described in (3.16) and biharmonic measure respectively, as described in (3.22), with $\sigma = 1$, $\lambda = 0.5$, while Figures 3.10 and 3.12 show the corresponding standard deviations. Note that the draws from the biharmonic measure are significantly smoother than the draws from the modified Helmholtz measure. Note also that as the number of gridpoints increases, the standard deviations for the draws from the modified Helmholtz measure increase, while the standard deviations for the biharmonic measure remain the same.

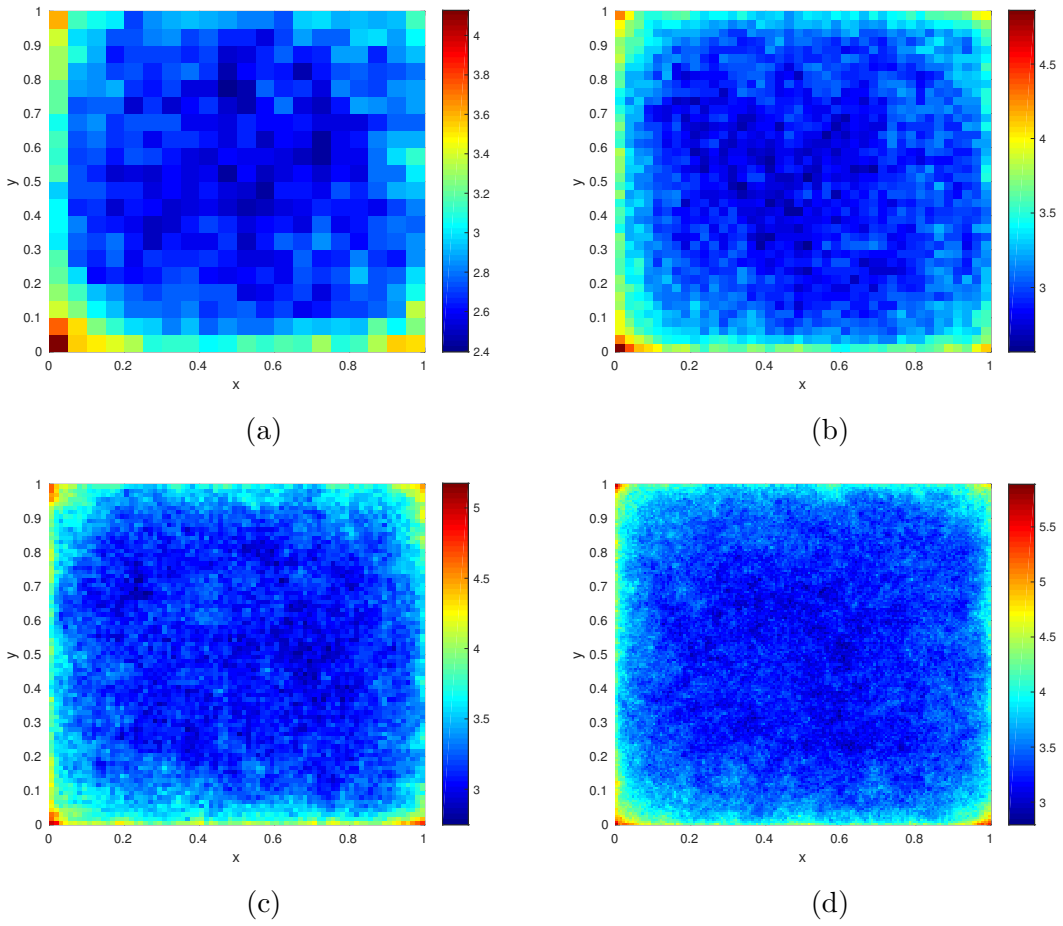
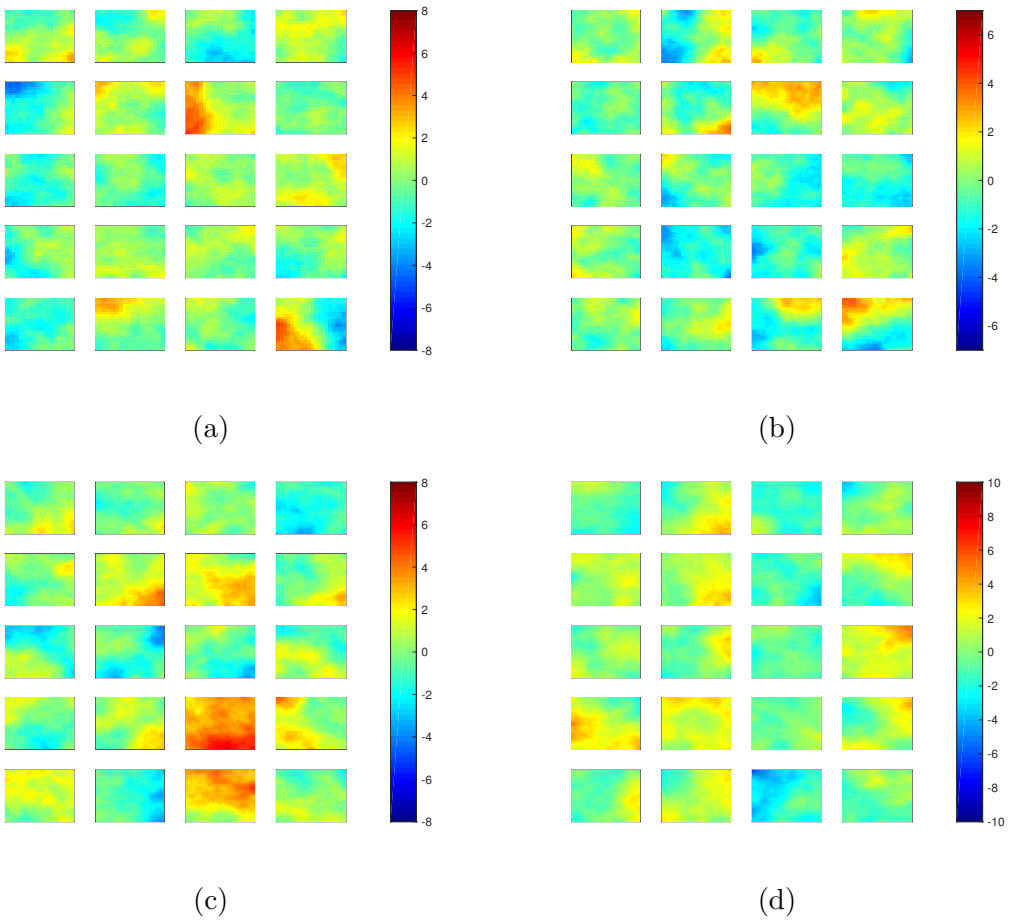


Figure 3.10: Standard deviations of 2-D isotropic modified Helmholtz measures on the unit square with $\sigma = 1$ and $\lambda = 0.5$ for (a) $N = 20$, (b) $N = 40$, (c) $N = 80$, and (d) $N = 160$.



(a) (b)

(c) (d)

Figure 3.11: 20 realizations of 2-D isotropic biharmonic measures on the unit square with $\sigma = 1$ and $\lambda = 0.5$ for (a) $N = 20$, (b) $N = 40$, (c) $N = 80$, and (d) $N = 160$.

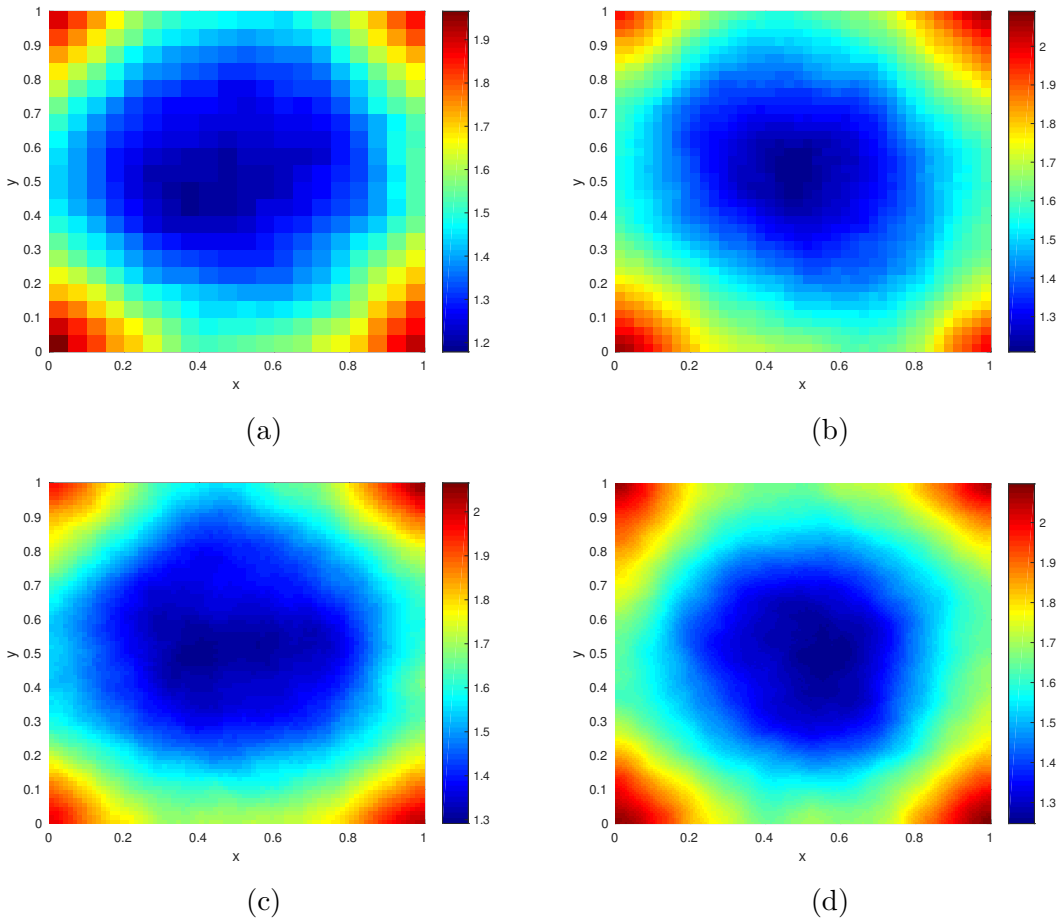


Figure 3.12: Standard deviations of 2-D isotropic biharmonic measures on the unit square with $\sigma = 1$ and $\lambda = 0.5$ for (a) $N = 20$, (b) $N = 40$, (c) $N = 80$, and (d) $N = 160$.

3.4 Bayesian Interpolation

3.4.1 Motivation

In solving inverse problems, on occasion, we might have limited observations on the parameter, m which we wish to incorporate in the prior measure. Bayesian interpolation, which is the focus of this section, gives us a way of doing this.

3.4.2 Formulation and Solution

Let D be a bounded open set, where $D \subseteq \mathbb{R}^d$, with $d = 1$, 2 or 3, and suppose we wish to solve an inverse problem for $m(x)$, for $x \in D$, where we observe the parameter instead of the variable. Let $\mu_0 = \mathcal{N}(0, \mathcal{C}_0)$ be the prior measure for this problem, for some prior covariance $\mathcal{C}_0$, defined through an elliptic operator $\mathcal{L}_0$ such that $\mathcal{C}_0 = \mathcal{L}_0^{-1}$.

Suppose we observe the parameter, m at several points, i.e., we have

$$s_k = m(x_k) + \xi_k, \quad k \in \mathbb{K},\ \xi_k \sim \mathcal{N}\left(0, \frac{1}{p}\mathbf{1}\right), \tag{3.33}$$

where s_k is the observation of the function at point x_k, $\mathbb{K}$ is an indexing set, and $\mathbf{1}$ is the identity operator. Then, the likelihood potential is given by

$$\Phi(m) = \frac{p}{2}\sum_{k\in\mathbb{K}}(s_k - m(x_k))^2. \tag{3.34}$$

Then, an application of Bayes' theorem gives a posterior measure, μ_1, whose Radon-Nikodym derivative with respect to μ_0, is given by

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}\mu_0} = \frac{1}{Z}\exp(-\Phi(m)), \tag{3.35}$$

where Z is a normalization constant. We may also write down the misfit functional, which is given by

$$\mathcal{J}(m) = \frac{p}{2}\sum_{k\in\mathbb{K}}(s_k - m(x_k))^2 + \frac{1}{2}\int m(\mathcal{L}_0 m)\,\mathrm{d}x, \tag{3.36}$$

We may rewrite the first term in the misfit functional as an integral involving a Dirac delta function. This yields

$$\mathcal{J}(m) = \frac{p}{2}\sum_{k\in\mathbb{K}}\int \delta(x - x_k)(s_k - m(x))^2\,\mathrm{d}x + \frac{1}{2}\int m(\mathcal{L}_0 m)\,\mathrm{d}x. \tag{3.37}$$

Combining the integrals gives

$$\mathcal{J}(m) = \int \left[\frac{p}{2}\sum_{k\in\mathbb{K}}\delta(x - x_k)(s_k - m(x))^2 + \frac{1}{2}m(\mathcal{L}_0 m)\right]\,\mathrm{d}x. \tag{3.38}$$

Note that this misfit functional is quadratic. Therefore, the posterior measure is Gaussian, and in order to fully characterize the posterior measure, it suffices to compute the mean and precision. This is done by minimizing the functional $\mathcal{J}(m)$.

We begin by computing the first variation of $\mathcal{J}(m)$. For a test function, ψ, the first variation of $\mathcal{J}(m)$, $\delta\mathcal{J}(m;\psi)$ is given by

$$\delta\mathcal{J}(m;\psi) = \int \left[p \sum_{k\in\mathbb{K}} \delta(x - x_k)(m(x) - s_k) + \mathcal{L}_0 m \right] \psi(x)\, \mathrm{d}x \tag{3.39}$$

Setting $\delta\mathcal{J} = 0$ and note that the equation holds for all ψ, the fundamental theorem of the calculus of variations gives the elliptic partial differential equation

$$\left(p \sum_{k\in\mathbb{K}} \delta(x - x_k) + \mathcal{L}_0 \right) \breve{m}(x) = p \sum_{k\in\mathbb{K}} s_k. \tag{3.40}$$

The solution to (3.40), $\breve{m}(x)$, is the mean of the posterior measure, while the operator

$$\mathcal{L} := p \sum_{k\in\mathbb{K}} \delta(x - x_k) + \mathcal{L}_0 \tag{3.41}$$

is the precision operator of the posterior measure.

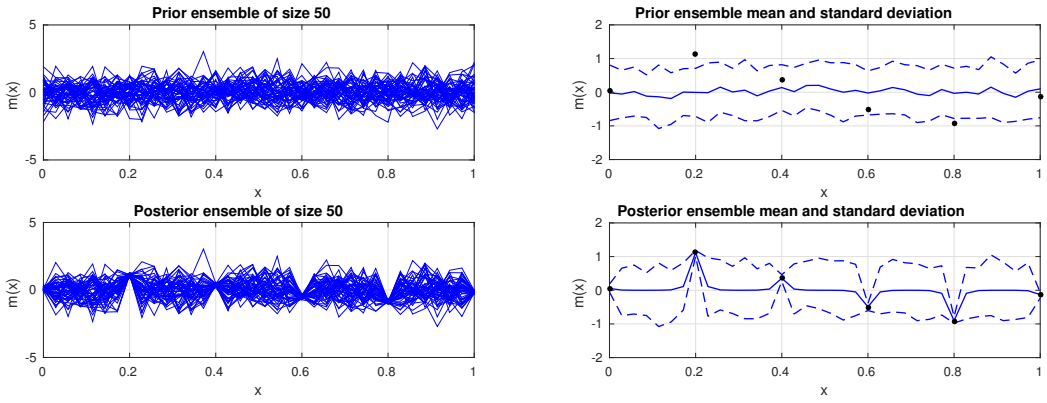
Thus, Bayesian interpolation gives us a way of updating a prior measure by incorporating information regarding limited observations of the parameters.

3.4.3 Examples

Next, we present some numerical experiments in 1-D to demonstrate the effects of Bayesian interpolation. For these simulations, we make $K = 6$ observations on the interval $[0, 1]$, discretized unifomly into K^2 gridpoints, so that the grid size, $h = 1/(K^2 - 1)$. Let the truth be $m(x) = \sin(2\pi x)$, and the observation precision is $p = 100$.

We perform Bayesian interpolation on the modified Helmholtz measure with the same variance and different correlation lengths. Figure 3.13 shows the results for Bayesian interpolation on a modified Helmholtz measure with $\sigma^2 = 1$ and $\lambda = 0.01$. Note here that a short correlation length implies that the prior measure behaves like white noise. The posterior mean is similar to the prior mean except at points where the observation is made, as there is little correlation structure between adjacent grid points.

When the correlation length is increased, we begin to see a greater influence exerted by the observations. This can be seen in Figure 3.14 where $\lambda = 0.2$. Here,



(a) (b)

Figure 3.13: Bayesian interpolation with a modified Helmholtz measure with nominal variance, $\sigma^2 = 1$ and nominal correlation length, $\lambda = 0.01$. The figures show (a) 50 realizations of the prior and posterior measures, and (b) the ensemble mean (blue continuous line) and its departure by one standard deviation (blue dashed line).

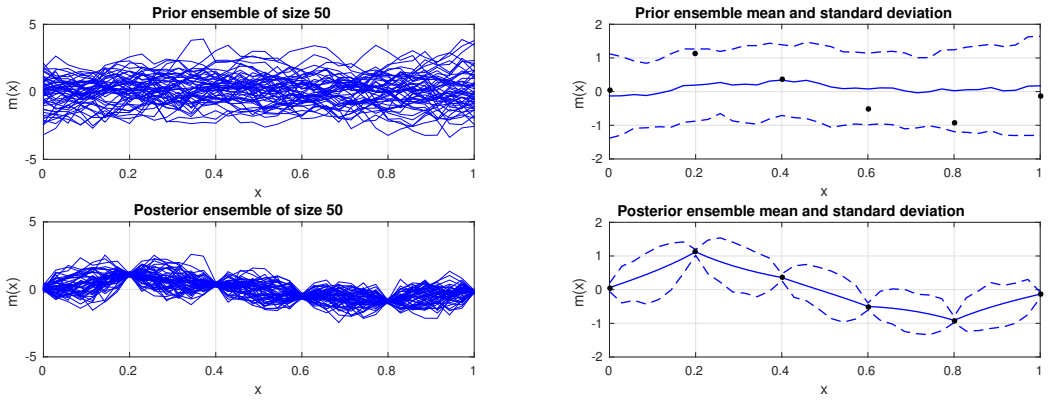
there is a higher degree of correlation between the observed points, and the posterior mean resembles a piecewise linear function connected by the points where observations were made.

Figure 3.15 shows the results for Bayesian interpolation for a modified Helmholtz measure with an even higher correlation length. Here, λ was chosen to be 0.8. Note that while the overall structure of each realization and the posterior mean is similar to that in Figure 3.14, the standard deviation is reduced. The reduction in fluctuations from the posterior mean is a result of the higher correlation structure.

These results demonstrate that for a prior with low correlation length, having limited direct observations on the parameters will not change the prior measure, except at the points where observations were made. However, for priors with higher correlation structure, even incorporating very few observations can alter the prior significantly. In general, one should be careful not to specify a high correlation length, as this leads to a dogmatic prior, unless of course, one has a firm belief that the correlation length is indeed large.

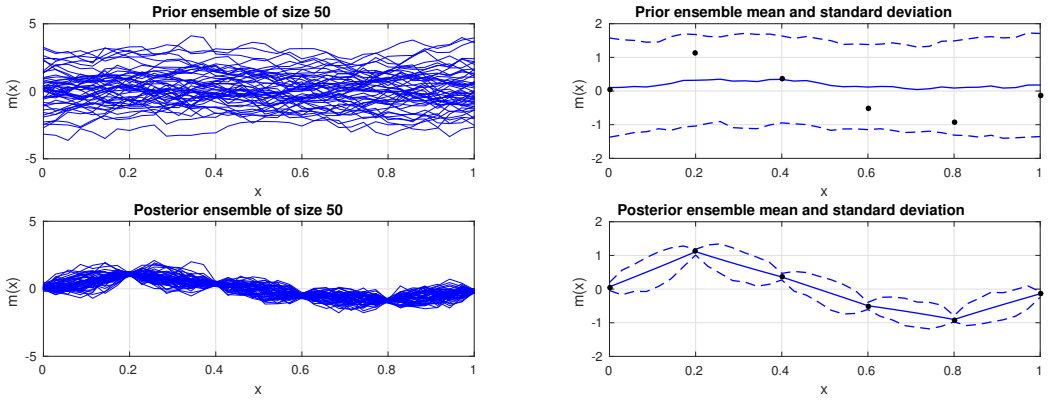
3.5 Concluding Remarks

In this chapter, we have presented various choices of Gaussian measures that can be used as priors in solving inverse problems. In particular, we focused on modified Helmholtz measures. We have also seen yet another advantage of working within a Bayesian framework which is valid in both finite and infinite dimensions, in that we have explained the phenomenon of an infinite variance of a covariance operator,



(a) (b)

Figure 3.14: Bayesian interpolation with a modified Helmholtz measure with nominal variance, $\sigma^2 = 1$ and nominal correlation length, $\lambda = 0.2$. The figures show (a) 50 realizations of the prior and posterior measures, and (b) the ensemble mean (blue continuous line) and its departure by one standard deviation (blue dashed line).



(a) (b)

Figure 3.15: Bayesian interpolation with a modified Helmholtz measure with nominal variance, $\sigma^2 = 1$ and nominal correlation length, $\lambda = 0.5$. The figures show (a) 50 realizations of the prior and posterior measures, and (b) the ensemble mean (blue continuous line) and its departure by one standard deviation (blue dashed line).

defined as the inverse of a modified Helmholtz operator, in systems with two or higher spatial dimensions.

It was noted that using modified Helmholtz measures for problems with higher spatial dimensions is not justified in infinite dimensions, as the draws from this measure are not almost surely regular. This led us to consider Gaussian measures with precision operators chosen as integer or fractional powers of modified Helmholtz operators. One such class of measures discussed here is that of biharmonic measures, which are measures that can be used as priors for systems in two and three spatial dimensions.

Finally, we presented a way to update our prior measure using Bayesian interpolation, in cases where limited direct observations of the parameters are available. Here, we have shown that careful consideration should be put into choosing the hyperparameters, as it will affect the results of Bayesian interpolation, as well as any further update on the prior measure by the observation and the forward model.

In the following chapters, we will use modified Helmholtz measures and biharmonic measures as priors to solve our inverse problems, applying Bayesian interpolation where limited direct observations are available.

Chapter 4

Weak Constraint Formulation

4.1 Introduction

In this chapter, we will introduce the weak constraint formulation of an inverse problem. This term originates from data assimilation applied to meteorology, where the method is called "Weak Constraint 4D-Var" (Fisher *et al.*, 2005, Trémolet, 2006, 2007, Zupanski, 1996), where 4D refers to three space and one time dimension, and "Var" is an abbreviation for "Variational". The basic idea of this method, as usually presented, arises from the imperfectness of the forward model. In some cases, the forward model does not describe a physical system sufficiently well. In such instances, it might be better to use a non-deterministic forward model. In Weak Constraint 4D-Var, model noise is added to the solution of the forward problem. In particular, if the variables, u at time step n, denoted by u^n, are related to the variables at the previous time step via the equation

$$u^n = \mathcal{F}(m)u^{n-1}, \tag{4.1}$$

for some given function $\mathcal{F}$ and parameters m, then the weak constraint approach involves the addition of a random term, so that the forward model becomes

$$u^n = \mathcal{F}(m)u^{n-1} + \zeta^n, \tag{4.2}$$

where ζ^n is a random term, governed by a suitable probability measure.

Here, we present a weak constraint formulation of an inverse problem, which differs somewhat from the Weak Constraint 4D-Var method. Instead of adding the model noise to the solution of the forward problem, we will directly include the model noise in the forward problem. As we will see, in terms of deterministic inverse problems, this results in an optimization problem where the constraint of the forward model is relaxed. Note that the deterministic version of the weak constraint formulation has been presented in van Leeuwen and Herrmann (2015).

The novelty of this approach is a rigorous formulation within the Bayesian framework, and a way to characterize the posterior using sampling. In the weak constraint formulation, the forward problem as well as the prior information on the parameters are regarded as the prior of an inverse problem. This differs from the usual Bayesian framework for inverse problems as described in Chapter 2 in two ways. Firstly, since the forward problem is no longer deterministic, we solve not only for m, but for u as well. Secondly, the prior measure does not only describe prior information on m, but on u as well. This is done through the forward problem. Then, it follows that the prior measure is no longer Gaussian in general. Nevertheless, the Girsanov formula (Theorem 4.2.1) ensures that the prior measure is always well-defined, if the measure induced by Brownian motion is used as the initial reference measure.

We begin by introducing the weak constraint formulation of an inverse problem in the fully discrete case, and then extending it to a semidiscrete case. This is done by taking the limit as $\tau \to 0$. Then, we describe methods to interpret the posterior measure, and apply them in two model problems – the 1-D well-testing problem as well as the inverse problem on the 2-D wave equation.

Finally, we remark that the theory developed here is restricted to forward problems which are linear in u. A similar theory might also be developed for any general nonlinear forward problem, although this might not be as straightforward as in the case of linear forward problems.

4.2 Weak Constraint Framework

4.2.1 Fully Discrete Case

Consider a forward model in a discrete space and discrete time setting of the form

$$u^n = u^{n-1} + \tau \mathcal{F}(m) u^{n-1} + \sqrt{\tau}\zeta^n, \quad n = 1, \cdots, N, \tag{4.3}$$

where $u^n \in \mathbb{R}^K$, $m \in \mathbb{R}^K$, $\mathcal{F}(m) \in \mathbb{R}^{K \times K}$, $\zeta^n \sim \mathcal{N}(0, \Gamma_f)$ for some covariance matrix Γ_f, with $\Gamma_f^{-1} = L_f$, and τ is the time step. Note that the function $\mathcal{F}(m)$ here is different from the function $\mathcal{F}(m)$ in (4.1) and (4.2). The vectors u^n and ζ^n denote their values at time step n.

For a symmetric, positive definite matrix A, let the inner product $\langle \cdot, \cdot \rangle_A$ and norm $\|\cdot\|_A$ be defined as

$$\langle u, v \rangle_A := u^T A^{-1} v, \quad \text{and} \quad \|u\|_A := \sqrt{\langle u, u \rangle_A}. \tag{4.4}$$

The joint density of m, u^n and ζ^n is given by

$$\pi_1(u^{0:N}, m, \zeta^{1:N}) = \frac{1}{Z_f} \prod_{n=1}^{N} \left[\delta(u^n - u^{n-1} - \tau \mathcal{F}(m) u^{n-1} - \sqrt{\tau}\zeta^n) e^{-\frac{1}{2}\|\zeta^n\|^2_{\Gamma_f}} \right] \pi_0(u^0, m), \tag{4.5}$$

where the superscript on $u^{0:N}$ denotes the variables from time step 0 to N, and similarly for $\zeta^{1:N}$. The density $\pi_0(u^0, m)$ is the prior density on m and u^0. The normalization constant Z_f is given by

$$Z_f = (2\pi)^{NK/2} \det(\Gamma_f)^{N/2}. \tag{4.6}$$

Then, the joint density of m and $u^{0:N}$ is given by

$$\pi_1(u^{0:N}, m) = \int \pi_1(u^{0:N}, m, \zeta^{1:N}) \, \mathrm{d}\zeta^{1:N}. \tag{4.7}$$

The integral is evaluated by performing the change of variable

$$\eta^n = \sqrt{\tau}\zeta^n, \quad n = 1, \cdots, N. \tag{4.8}$$

Then, we have

$$\pi_1(u^{0:N}, m) = \frac{\pi_0(u^0, m)}{\tau^{N/2} Z_f} \int \prod_{n=1}^{N} \left[\delta(u^n - u^{n-1} - \tau \mathcal{F}(m) u^{n-1} - \eta^n) e^{-\frac{1}{2\tau}\|\eta^n\|^2_{\Gamma_f}} \right] \, \mathrm{d}\eta^{1:N}, \tag{4.9}$$

which gives

$$\pi_1(u^{0:N}, m) = \frac{1}{\tau^{N/2} Z_f} \exp\left(-\mathcal{K}_f(u^{0:N}, m)\right) \pi_0(u^0, m), \tag{4.10}$$

where $\mathcal{K}_f(u^{0:N}, m)$ is

$$\mathcal{K}_f(u^{0:N}, m) = \frac{1}{2\tau} \sum_{n=1}^{N} \left\| u^n - u^{n-1} - \tau \mathcal{F}(m) u^{n-1} \right\|^2_{\Gamma_f}. \tag{4.11}$$

Expanding the terms in the summation gives

$$\mathcal{K}_f(u^{0:N}, m) = \frac{1}{2\tau} \sum_{n=1}^{N} \Bigg(\left\| u^n - u^{n-1} \right\|^2_{\Gamma_f} - 2\tau \left\langle u^n - u^{n-1}, \mathcal{F}(m) u^{n-1} \right\rangle_{\Gamma_f} + \tau^2 \left\| \mathcal{F}(m) u^{n-1} \right\|^2_{\Gamma_f} \Bigg). \tag{4.12}$$

Suppose we specify a Gaussian density on m and u^0, with mean and covariance pairs (m_0, C_0^m) and (u_0, C_0^u) respectively. Then, we have

$$\pi_1(u^{0:N}, m) \propto \exp\left(-\mathcal{K}_f(u^{0:N}, m) - \mathcal{K}_p(u^0, m)\right), \tag{4.13}$$

where $\mathcal{K}_p(u^0, m)$ is given by

$$\mathcal{K}_p(u^0, m) = \frac{1}{2}\|m - m_0\|^2_{C_0^m} + \frac{1}{2}\left\|u^0 - u_0\right\|^2_{C_0^u}. \tag{4.14}$$

Then, the probability density $\pi_1(u^{0:N}, m)$ may be used as the prior to solve an inverse problem. Given an observation s, corrupted by additive noise with zero mean and covariance Γ, the Bayes update gives the posterior density

$$\pi(u^{0:N}, m \mid s) = \frac{1}{Z}\exp(-\mathcal{K}(u^{0:N}, m)), \tag{4.15}$$

where Z is a normalization constant and the function $\mathcal{K}(u^{0:N}, m)$ is

$$\mathcal{K}(u^{0:N}, m) = \mathcal{K}_\ell(u^{0:N}) + \mathcal{K}_f(u^{0:N}, m) + \mathcal{K}_p(u^0, m), \tag{4.16}$$

where $\mathcal{K}_\ell(u^{0:N})$ is the functional

$$\mathcal{K}_\ell(u^{0:N}) = \frac{1}{2}\left\|\mathcal{O}(u^{0:N}) - s\right\|^2_\Gamma, \tag{4.17}$$

for an observation operator $\mathcal{O}$.

4.2.2 Semidiscrete Case

We now consider the weak constraint formulation in the semidiscrete case, which is the continuous time setting. This is the case where $\tau \to 0$. In the fully discrete case, since the inverse problem is finite-dimensional, the probability densities defined are with respect to the Lebesgue measure. However, in the infinite-dimensional case, there is no natural generalization of the Lebesgue measure. Therefore, a new reference measure is required. We first define this reference measure in the discrete case, and then consider the limit as $\tau \to 0$. This will then give us a similar form of (4.15) in the continuous time setting.

We begin by defining a reference measure according to the dynamical system

$$u^n = u^{n-1} + \sqrt{\tau}\zeta^n, \quad \zeta^n \sim \mathcal{N}(0, \Gamma_f), \tag{4.18}$$

with the same initial condition as (4.3). Then, by a similar argument as above, the joint density of m and $u^{0:N}$, based on (4.18) is given by

$$\tilde{\pi}_1(u^{0:N}, m) = \frac{1}{\tau^{N/2} Z_f}\exp\left(-\frac{1}{2\tau}\sum_{n=1}^{N}\left\|u^n - u^{n-1}\right\|^2_{\Gamma_f}\right)\pi_0(u^0, m), \tag{4.19}$$

and we have

$$\pi_1(u^{0:N}, m) = \tilde{\pi}_1(u^{0:N}, m)\exp\left(\sum_{n=1}^{N}\left(\left\langle u^n - u^{n-1}, \mathcal{F}(m)u^{n-1}\right\rangle_{\Gamma_f} - \frac{\tau}{2}\left\|\mathcal{F}(m)u^{n-1}\right\|^2_{\Gamma_f}\right)\right), \tag{4.20}$$

as motivated by (4.12).

Note that, in the finite-dimensional case, the probability densities as defined in Subsection 4.2.1 are expressed in terms of the usual Lebesgue measure, that is, if ν_1 and $\tilde{\nu}_1$ are the measures induced by the equations (4.3) and (4.18) respectively, then the densities $\pi_1(u^{0:N}, m)$ and $\tilde{\pi}_1(u^{0:N}, m)$ are given by the Radon-Nikodym derivatives of ν_1 and $\tilde{\nu}_1$ with respect to the Lebesgue measure, i.e.

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\mu_L} = \pi_1(u^{0:N}, m), \quad \text{and} \quad \frac{\mathrm{d}\tilde{\nu}_1}{\mathrm{d}\mu_L} = \tilde{\pi}_1(u^{0:N}, m), \tag{4.21}$$

where μ_L denotes the Lebesgue measure. However, we may perform a change of reference measure using the chain rule

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\mu_L} = \frac{\mathrm{d}\nu_1}{\mathrm{d}\tilde{\nu}_1}\frac{\mathrm{d}\tilde{\nu}_1}{\mathrm{d}\mu_L}. \tag{4.22}$$

Then, substituting (4.21) and (4.22) into (4.20) and eliminating $\mathrm{d}\tilde{\nu}_1/\mathrm{d}\mu_L$ on both sides, we have

$$\frac{\mathrm{d}\nu_1}{\mathrm{d}\tilde{\nu}_1} = \exp\left(\sum_{n=1}^{N}\left(\left\langle u^n - u^{n-1}, \mathcal{F}(m)u^{n-1}\right\rangle_{\Gamma_f} - \frac{\tau}{2}\left\|\mathcal{F}(m)u^{n-1}\right\|_{\Gamma_f}^2\right)\right), \tag{4.23}$$

which is *the Girsanov formula* (Law *et al.*, 2015) in finite dimensions. It is this form that generalizes into the continuous time setting in the limit as $\tau \to 0$.

In particular, consider the dynamical system, written informally in Langevin form as

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \mathcal{F}(m)u(t) + \Gamma_f^{1/2}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{4.24}$$

where $u : [0, T] \to \mathbb{R}^K$, $\mathcal{F}(m) \in \mathbb{R}^{K\times K}$ and $\mathrm{d}W/\mathrm{d}t$ is distributed according to the standard normal distribution. This is the continuous form of (4.3). The equation (4.24) induces the measure μ_1 on the Hilbert space $\mathcal{H} := L^2([0, T]; \mathbb{R}^K)$, $T > 0$, with reference measure $\tilde{\mu}_1$, induced by the dynamical system

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t} = \Gamma_f^{1/2}\frac{\mathrm{d}W}{\mathrm{d}t}, \tag{4.25}$$

where both dynamical systems have the same initial condition and $\mathcal{F}(m)u(t)$ is Lipschitz on bounded sets in both u and m. Then, the Girsanov formula can be derived from (4.23) by taking $\tau \to 0$ and is stated in the following theorem.

Theorem 4.2.1 (Girsanov Formula, Theorem 6.9 of Law *et al.* (2015))**.** *Suppose $\mathcal{F}(m)$ and $\mathrm{d}W/\mathrm{d}t$ are such that both equations* (4.24) *and* (4.25) *have solutions on $t \in [0, T]$ that do not blow up almost surely. Then, the measures μ_1 and $\tilde{\mu}_1$ on $\mathcal{H}$ generated by* (4.24) *and* (4.25) *respectively are equivalent with Radon-Nikodym derivative*

$$\frac{\mathrm{d}\mu_1}{\mathrm{d}\tilde{\mu}_1} = \exp\left(-\frac{1}{2}\int_0^T \|\mathcal{F}(m)u(\tau)\|_{\Gamma_f}^2\,\mathrm{d}\tau + \langle\mathcal{F}(m)u, \mathrm{d}u\rangle_{\Gamma_f}\right). \tag{4.26}$$

It is the measure μ_1 that is used as the prior measure in solving an inverse problem when the forward model involves stochastic noise. Note here that the prior measure is not Gaussian. The Girsanov formula stated in (4.26) ensures that the prior measure, μ_1 is well-defined, as it is equivalent to a scaled Brownian motion.

Then, an informal application of Bayes' Theorem gives the posterior measure, μ, whose Radon-Nikodym derivative is given by

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_1} = \frac{1}{Z} \exp\left(-\frac{1}{2} \left\|\mathcal{O}(u(t)) - s\right\|_{\Gamma}^2\right), \tag{4.27}$$

for some observation operator $\mathcal{O}$ and observation covariance Γ, noting that the values of $\mathcal{O}(u(t))$ are a finite-dimensional vector.

Note that while we have justified the use of the prior measure μ_1, we have yet to justify the posterior measure μ with μ_1 as the reference measure. This will involve proving the following:

1. μ is a well-defined measure.
2. μ is a probability measure, i.e., Z is bounded above and below.
3. μ is stable with respect to small perturbations in the observation, s.
4. μ is indeed the probability measure for m and u, given s, with Radon-Nikodym derivative with respect to μ_1 given by (4.27).

However, we conjecture that these results can be generalized from the results in Stuart (2010), as the weak constraint requires a two-step Bayes update, beginning from a Gaussian prior measure, μ_0. Therefore, conditions required for μ to be a well-defined probability measure should also follow from the conditions outlined in Chapter 2.

4.2.3 Fully Continuous Setting

Girsanov's formula in Theorem 4.2.1 readily generalizes to a continuous space and time setting as well. In this case, the Hilbert space in consideration is $\mathcal{H} = L^2([0,T], H)$, where H is an infinite-dimensional real, separable Hilbert space. One will then be able to prove a more general form of Theorem 4.2.1. For a technical discussion on this matter, we refer the reader to Da Prato and Zabczyk (2014).

4.3 Analyzing The Posterior

The weak constraint formulation differs from the framework outlined in Chapter 2, in that since the forward model is no longer deterministic, one would need to find

a probability measure that describes both the parameters m and variables u. In our preceding discussion, we used the superscript notation to describe u at a particular time step. For the rest of this chapter, we write u to denote either a vector of the variables at every time step, ordered using the standard ordering, $((u^1)^T, (u^2)^T, \cdots, (u^N)^T)^T$, or a time-dependent function in the infinite-dimensional setting. In the most general form, given the model noise ζ, normally distributed with zero mean and covariance $\mathcal{C}_f$, we may write the forward model as

$$\mathcal{A}(m)u - q(m) = \zeta, \tag{4.28}$$

where $q(m)$ is an appropriate source term and $\mathcal{A}(m)$ is a matrix defined blockwise, augmented according to the time-stepping scheme employed in solving the forward problem.

Then, the weak constraint misfit functional to be considered can be written as

$$\mathcal{K}(u, m) = \frac{1}{2} \|\mathcal{O}(u) - s\|_{\Gamma}^2 + \frac{1}{2} \|\mathcal{A}(m)u - q(m)\|_{\mathcal{C}_f}^2 + \frac{1}{2} \|m - m_0\|_{\mathcal{C}_0}^2, \tag{4.29}$$

where $\mathcal{O}$ is an observation operator and Γ is the observation covariance. Here, note that the initial condition, u^0 in (4.16) is treated as a parameter.

4.3.1 Probability Maximizers

In the classical Bayesian framework for solving inverse problems, it was explained in Subsection 2.3.2 that one could find a probability maximizer by solving the optimization problem

$$\breve{m} = \underset{m \in \mathcal{H}_m}{\arg\min}\, \mathcal{J}^s(m), \tag{4.30}$$

where $\mathcal{J}^s(m)$ is given in (2.42). This is known as the strong constraint misfit functional. In the weak constraint formulation, one could similarly find a probability maximizer by solving

$$(\breve{u}, \breve{m}) = \underset{u \in \mathcal{H}_u, m \in \mathcal{H}_m}{\arg\min}\, \mathcal{K}(u, m), \tag{4.31}$$

with $\mathcal{K}(u, m)$ as given in (4.29).

Note that minimizing (4.29) is easier than minimizing the strong constraint misfit as described in (2.42), as the optimizer is not required to fit the forward model exactly.

Furthermore, an evaluation of the weak constraint misfit functional would not require the solution of a forward problem. Consequently, any gradient computations of $\mathcal{K}(u, m)$ can be done without using the adjoint method, where additional computations are required to solve a linear system for the adjoint states.

Moreover, examining the terms in (4.29), we note that if $\mathcal{O}$ is linear, then the first and third terms are quadratic, and only the second term contributes to the non-convexity of $\mathcal{K}(u, m)$. Therefore, it is plausible that model noise with a higher variance will dampen the non-convexity and allow the functional $\mathcal{K}(u, m)$ to be minimized more easily.

Note however, that a straightforward implementation of the minimization of $\mathcal{K}(u, m)$ requires substantial computer memory. For instance, if $m \in \mathbb{R}^K$ and each forward run has N time steps, then $u \in \mathbb{R}^{NK}$, so that the memory requirement has grown by a factor of $N + 1$ for the storage of each state vector. If the forward model is time independent, then the amount of storage memory required for a state vector is doubled. It is thus necessary to consider optimization schemes that allow us to hold only the variables at two time steps in fast memory, whilst the remaining variables can be held on auxiliary storage.

In van Leeuwen and Herrmann (2015), it was recommended that one should first solve for the optimal values for the variables and subsequently work only within the parameter space. It was asserted in van Leeuwen and Herrmann (2015) that for this approach, called the reduced penalty method, only half of the number of PDE solutions would be required in the minimization compared to the minimization problem of the conventional misfit functional.

4.3.2 Sampling

Sampling is typically performed using Markov chain Monte Carlo (MCMC) methods. For instance, one could start from a random point and employ a random walk to explore the state space. For this case, there would be a burn-in period before the Markov chain is within the support of the probability measure. One could also start from a probability maximizer and begin the random walk from such a local optimum. In some cases, the chain could be stuck within the neighborhood of the probability maximizer. In (Oliver *et al.*, 2008, p. 331–333), a 1-D example was constructed where the posterior density was bimodal. The method used in Oliver *et al.* (2008) was to approximate the posterior density by a Gaussian around the probability maximizer, and use that approximation to generate proposals. It was found that the proposals generated could not reasonably explore the neighborhood of the other mode because the acceptance probability of a proposal to transition to that neighborhood was too low.

Randomized Maximum Likelihood (RML) is a method introduced to sample from a probability density. The method of RML is described in Oliver *et al.* (2008, Chapter 10). If the inverse problem is linear and the prior is Gaussian, it can be proven

that RML generates exact samples Oliver *et al.* (2008). However, for a nonlinear inverse problem or a non-Gaussian prior, the samples are only approximate. In fact, Oliver *et al.* (2008, p. 328–329) show that the uncertainty is sometimes overestimated. Nevertheless, one can use the ensemble obtained from RML as a starting point for a random walk to obtain exact samples. We describe this method in Algorithm 1.

Algorithm 1 is divided into two parts. The first part is the Randomized Maximum Likelihood. The RML procedure begins by generating a prior ensemble of independent realizations. The RML ensemble is then generated using optimization from the prior ensemble. Note here that since we are using the RML ensemble as the starting point in the Markov chain, it is not necessary for the optimization problem to be solved to a high accuracy.

The second part of the algorithm is the Markov chain Monte Carlo method. In particular, the proposal used is the preconditioned Crank-Nicolson proposal in a random walk, as describe in Cotter *et al.* (2013). Using the RML ensemble as the starting point in a Markov chain gives the benefit that we are able to use shorter chains to explore the state space. Furthermore, the ensemble obtained will honour the observations.

The end result is an independent posterior ensemble, generated using shorter Markov chains, which honours the observations, and allows us to explore the state space more thoroughly, but at the cost of solving more optimization problems.

4.3.3 Relation to the Penalty Method

Note that the weak constraint approach has been considered in PDE-constrained optimization, for instance, in van Leeuwen and Herrmann (2015). There, the covariance of the forward model was chosen to be $(1/\lambda)\mathbf{1}$, with $\mathbf{1}$ as the identity matrix corresponding to zeroth order Tikhonov regularization, and the observation error was assumed to be standard normal, with a linear observation operator. Then, the weak constraint misfit becomes

$$\mathcal{K}(u,m) = \frac{1}{2}\left\|\mathcal{O}(u) - s\right\|_2^2 + \frac{\lambda}{2}\left\|\mathcal{A}(m)u - q\right\|_2^2 + \frac{1}{2}\left\|m - m_0\right\|_{\mathcal{C}_0}^2, \tag{4.40}$$

where λ was called the penalty parameter.

In Lemma 4.1, Theorem 4.3, and Section 5.3 of van Leeuwen and Herrmann (2015), it was suggested that the penalty parameter λ should be scaled according to $\|\mathcal{O}(\mathcal{A}^{-1})\|_2^2$. In particular the scaling suggested was of the form

$$\lambda = \tilde{\lambda}\left\|\mathcal{O}(\mathcal{A}^{-1})\right\|_2^2, \tag{4.41}$$

Algorithm 1 RML with MCMC

Input R: number of ensemble members, N_c: length of Markov chain.
for $r = 1$ **to** R **do**
Draw $m_r \sim \pi_0$, and solve $\mathcal{A}(m_r)u_r - q = 0$.
Solve

$$(\breve{u}_r, \breve{m}_r) = \arg\min_{u,m} \mathcal{K}_r(u, m), \tag{4.32}$$

where

$$\mathcal{K}_r(u, m) = \frac{1}{2} \|\mathcal{O}(u) - s\|_\Gamma^2 + \frac{1}{2} \|\mathcal{A}(m)u - q\|_{\mathcal{C}_f}^2 + \frac{1}{2} \|m - m_r\|_{\mathcal{C}_0}^2 . \tag{4.33}$$

end for
for $r = 1$ **to** R **do**
Set the realization $(u_r^*, m_r^*) = (\breve{u}_r, \breve{m}_r)$.
Compute

$$\zeta_r^* = \mathcal{A}(\breve{m}_r)\breve{u}_r - q. \tag{4.34}$$

for $i = 1$ **to** N_c **do**
Draw $\hat{m} \sim \pi_0(m)$ and $\hat{\zeta} \sim \mathcal{N}(0, \mathcal{C}_f)$.
Set $\beta \in [0, 1]$ and generate random walk proposals

$$\tilde{m} = \sqrt{1 - \beta^2} m_r^* + \beta \hat{m} \tag{4.35}$$

$$\tilde{\zeta} = \sqrt{1 - \beta^2} \zeta_r^* + \beta \hat{\zeta}. \tag{4.36}$$

Compute $\tilde{u}$ by solving $\mathcal{A}(\tilde{m})\tilde{u} = q + \tilde{\zeta}$.
Compute acceptance probability, α given by

$$\alpha = \min(1, \exp(\Psi_2 - \Psi_1)), \tag{4.37}$$

with

$$\Psi_1 = \frac{1}{2} \|\mathcal{O}(\tilde{u}) - s\|_\Gamma^2 , \tag{4.38}$$

$$\Psi_2 = \frac{1}{2} \|\mathcal{O}(u_r^*) - s\|_\Gamma^2 . \tag{4.39}$$

Accept the proposal $(\tilde{u}, \tilde{m})$ with probability α.
if accept **then**
Set $(u_r^*, m_r^*, \zeta_r^*) = (\tilde{u}, \tilde{m}, \tilde{\zeta})$.
end if
end for
end for

with $\tilde{\lambda} > 1$ considered large and $\tilde{\lambda} < 1$ considered small. From a Bayesian perspective, this is a discussion on the relative magnitude of the precision of the forward model and the observation precision. In particular, if we set $\tilde{\lambda} > 1$, then we are specifying a higher precision on the forward problem than on the observations. Conversely, setting $\tilde{\lambda} < 1$ implies that the observations have a higher precision than the forward problem.

Similarly, one could also consider the relative size of the precision of the forward problem and the prior precision, $\mathcal{L}_0$ by considering the scaling

$$\lambda = \overline{\lambda} \left\| \mathcal{L}_0(\mathcal{A}^{-1}) \right\|_2^2. \tag{4.42}$$

Here, specifying $\overline{\lambda} > 1$ would put more emphasis on satisfying the forward model instead of staying close to the prior mean, while specifying $\overline{\lambda} < 1$ would lead to an optimizer which prefers to be closer to the prior mean rather than satisfying the forward model.

4.4 Applications

In this section, we apply the weak constraint formulations to two model problems. The first is the 1-D well testing problem. Well-testing is an important problem in petroleum engineering, where a well test is used in order to learn about the characteristics in the region near the well in a reservoir where hydrocarbons are trapped. The second problem is an inverse problem on the 2-D wave equation. This is related to the seismic inversion problem.

In both problems, we will perform the RML-MCMC method for sampling described in Subsection 4.3 above. In order to evaluate more accurately the effectiveness of the weak constraint method and the RML-MCMC algorithm, we will work within the perfect model scenario, where the same grid as used to generate the observations is also used for inversion, and the noise in the forward problem is not added into the observations. Note that for both model problems, the step size of the RML-MCMC sampling method was chosen such that the acceptance rate for the random walk proposals were roughly 0.25. All optimization problems were solved with Matlab's trust-region method with a linear model function, and tolerance of 10^{-6}.

4.4.1 1-D Well-testing

Consider the partial differential equation

$$\gamma \frac{\partial u}{\partial t} - \frac{\partial}{\partial x}\left(\frac{\kappa}{\mu}\frac{\partial u}{\partial x}\right) = q, \tag{4.43}$$

on the unit interval $[0,1]$ with homogeneous Neumann boundary conditions. Here, γ is the porosity, κ is the permeability, μ is the viscosity, u represents the pressure and q is a source term. We assume that $\kappa/\mu \equiv 1$, $q \equiv 0$ and we write the porosity in the form

$$\gamma = \frac{e^m}{1+e^m}, \tag{4.44}$$

where $m : \mathbb{R} \to \mathbb{R}$, so that $\gamma \in (0,1)$. Then, (4.43) becomes

$$\frac{e^m}{1+e^m}\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = 0. \tag{4.45}$$

We prescribe the initial condition

$$u(x,0) = 1 - \delta(x - x_w), \tag{4.46}$$

where x_w is the position of the well.

The equation (4.45) is discretized using centred finite differences in space and implicit Euler in time. We use a uniform grid of size h, so that the grid points, x_i are given by

$$x_i = ih, \quad i = 0, \cdots K, \tag{4.47}$$

and $h = 1/K$. We also use a uniform time step, Δt and N time steps.

This gives N linear systems of the form

$$F(m)U^n = U^{n-1}, \quad n = 1, \cdots, N, \tag{4.48}$$

where the vector $U^n = (u_0^n, u_1^n, \cdots, u_{K-1}^n, u_K^n)^T$, and the matrix F is the $(K+1) \times (K+1)$ matrix

$$F(m) = \mathbf{1} - \frac{\Delta t}{h^2} A(m) B_h, \tag{4.49}$$

and $\mathbf{1}$ is the identity matrix. The matrices $A(m)$ and A_h are

$$A(m) = \begin{pmatrix} 1+e^{-m_0} & & & & & \\ & 1+e^{-m_1} & & & & \\ & & \ddots & & & \\ & & & 1+e^{-m_{K-1}} & & \\ & & & & 1+e^{-m_K} \end{pmatrix}, \tag{4.50}$$

$$B_h = \begin{pmatrix} -1 & 1 & & & & & & \\ 1 & -2 & 1 & & & & & \\ & 1 & -2 & 1 & & & & \\ & & \ddots & \ddots & \ddots & & & \\ & & & & 1 & -2 & 1 & \\ & & & & & 1 & -2 & 1 \\ & & & & & & 1 & -1 \end{pmatrix}, \tag{4.51}$$

and entries not displayed are 0.

Let i_w be the index of the position of the well. Then, the initial condition is

$$U_i^0 = 1 - \delta_{i,i_w}. \tag{4.52}$$

The flux around the well is observed at every time step, i.e.,

$$s^n = \frac{u_{i_w+1}^n - 2u_{i_w}^n + u_{i_w-1}^n}{h} + \xi^n, \quad \xi^n \sim \mathcal{N}(0, \sigma_{obs}^2). \tag{4.53}$$

We perform RML-MCMC sampling as described in Subsection 4.3 above to generate samples for the porosity profile on the grid. For the forward model, we set $K = 50$, $\Delta t = 1/400$ and $N = 10$. The well is placed at the centre of the domain, $x_w = 1/2$. The observation noise is white noise with precision 10^4. The precision of the forward model was specified according to (4.41) with $\tilde{\lambda} = 10$. For this problem, no noise was added when the observations were generated, although noise is assumed to be present when solving the inverse problem.

The prior specified has zero mean and a modified Helmholtz prior precision, constructed with $\sigma = 1.5$ and $\lambda = 0.2$, as described in Chapter 2. The RML ensemble has 100 members, and each ensemble member undergoes a random walk of step size $\beta = 0.02$ and a Markov chain of length 1000.

Figures 4.2(a)–(b) show 10 realizations of the prior as well as the ensemble mean and standard deviation. The prior realizations have a rough structure due to the short correlation length. This feature is also present in the realizations of the RML ensemble and the posterior (Figures 4.3(a) and 4.4(a)).

Figures 4.3(a)–(b) show 10 realizations of the RML ensemble as well as the ensemble mean and standard deviation, and Figures 4.4(a)–(b) show 10 realizations of the RML-MCMC ensemble as well as the ensemble mean and standard deviation.

Note that for both the RML ensemble and the RML-MCMC ensemble, there is a smaller degree of uncertainty around $x = 0.5$ than the other parts of the domain (Figures 4.3(b) and 4.4(b)). This corresponds to the position of the well, as observations were made at that position. Both the RML ensemble mean and the RML-MCMC ensemble mean have a similar structure to the probability maximizer (Figure 4.1). Here, since the observations were only made at one point, the most probable porosity profile is one which produces an observation which is closer to the observation at that point ($x = 0.5$), while at the rest of the gridpoints, there is a tendency to move toward the structure of the prior. In this case, since the prior has a zero mean, the probability maximizer is closer to zero, with the correlation structure prescribed in the prior precision operator preserved.

Observe that the RML-MCMC ensemble mean and standard deviation did not change a lot when the MCMC algorithm was applied on the RML ensemble. This is a result of only small corrections made to the RML ensemble, as can be seen between Figure 4.3(a) and Figure 4.4(a), and is an indication of the effectiveness of the method of RML for this problem. Two factors contribute to this. Firstly, this has to do with the convexity of the misfit function, $\mathcal{K}(m, u)$. As mentioned before, a linear inverse problem with Gaussian prior will produce exact samples under the method of RML. The fact that the RML ensemble has performed well, in that it has not been moved too much under the MCMC corrections could indicate that $\mathcal{K}(m, u)$ has a convex structure. The second factor that contributes to the effectiveness of the performance of RML has to do with the dimension of the problem. Since the problem is in one spatial dimension, the dimension of the inverse problem is relatively small. In smaller dimensions, it is easier to explore the state space, as there are fewer perturbation directions. Therefore, a smaller sample size would be required for one to generate a sample which covers the support of the posterior density.

Next, we evaluate the quality of the realizations of the RML and the RML-MCMC ensemble. This is done by computing the error between the observations generated using the realizations and the observations used for the inverse problem. The results can be seen in Figure 4.5. Here, we observe that both the errors of the RML ensemble and the RML-MCMC ensemble are centred around 10^{-2}, which is precisely the standard deviation of the observations, with the errors coming from the posterior realizations having less fluctuations. This is an indication that the random walk Metropolis-Hastings method did a good job of moving the RML ensemble into one that contains exact realizations of the posterior. Note that although the RML ensemble seemed to be a sufficiently good sample to start with, there are a few realizations that did not produce observations close to the data. For instance, an ensemble member between 20 and 30 produced an observation that was of the order 0.1 away from the data generated, indicating that the prior realization generated to begin with could have been two or three standard deviations away from the prior mean. Nevertheless, the MCMC procedure moved this ensemble member into one which better honoured the observations.

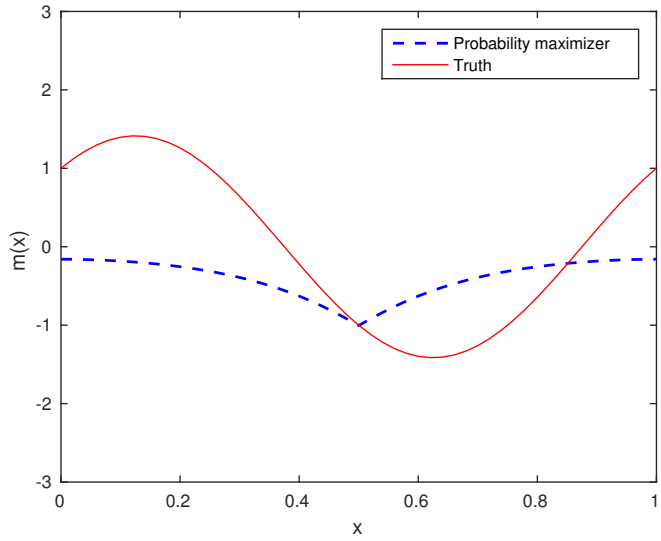
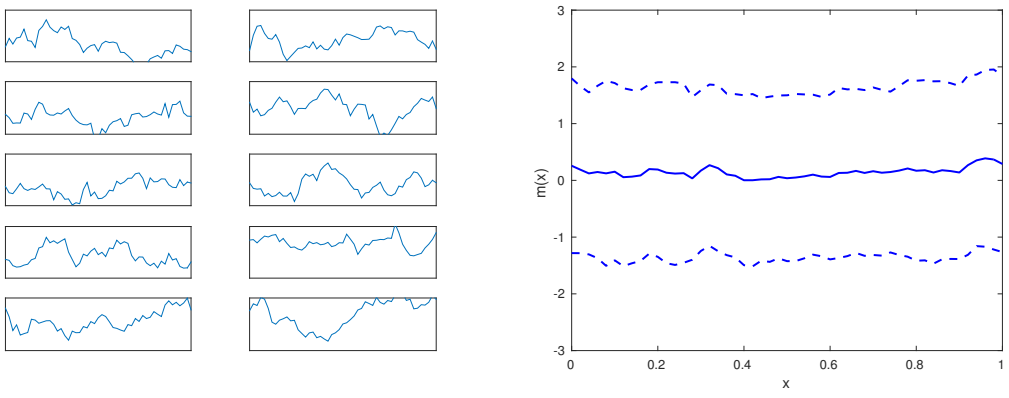
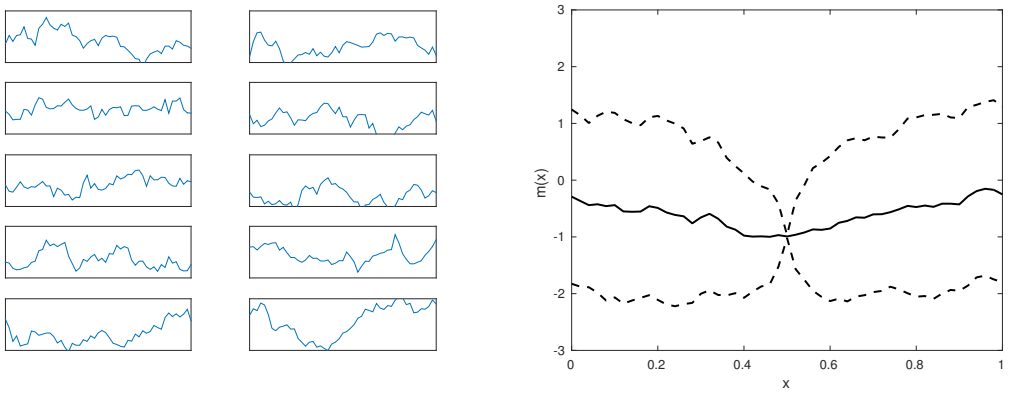


Figure 4.1: Probability maximizer and truth for the 1-D well test problem. The probability maximizer was generated by solving an optimization on the misfit functional of the form (4.29) for this inverse problem.



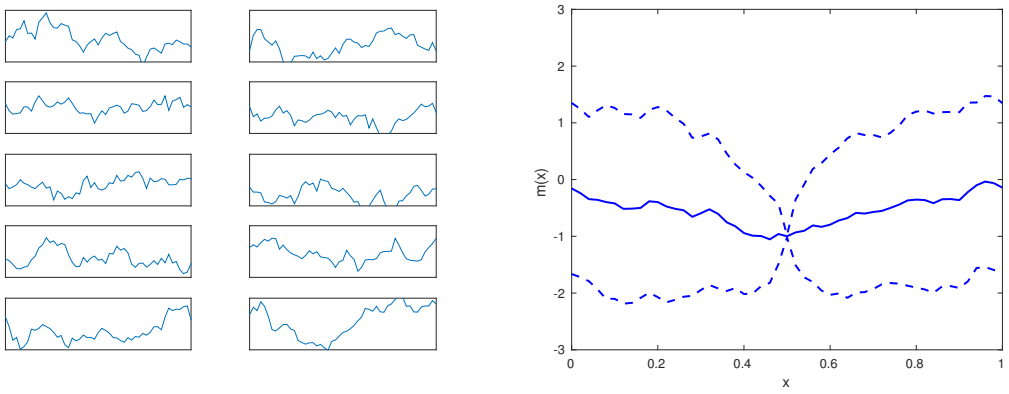
(a) (b)

Figure 4.2: (a) 10 realizations of the prior for the 1-D well test problem. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 4.3: (a) 10 realizations of the RML ensemble for the 1-D well test problem. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.



(a) (b)

Figure 4.4: (a) 10 realizations of the RML-MCMC ensemble for the 1-D well test problem. The realizations were all plotted on the same scale, with the same ordinate and abscissa axes as in (b). (b) The ensemble mean is indicated by the solid line, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

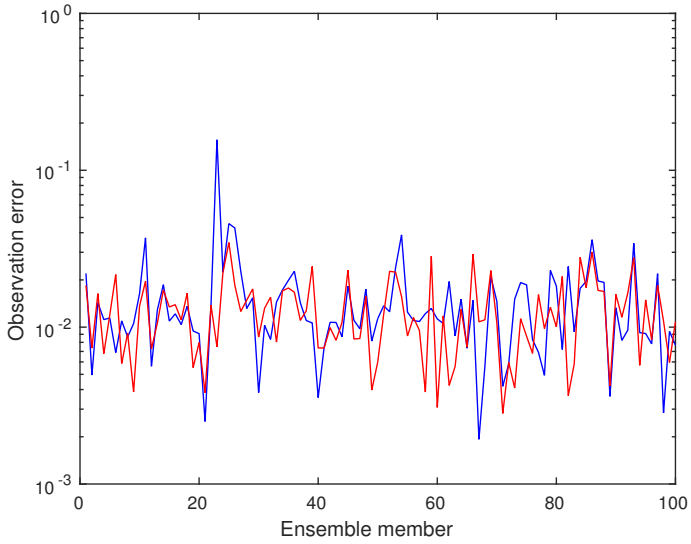


Figure 4.5: Observation errors for the 1-D well test problem. The red line shows $\|\mathcal{O}(u(m^r_{MCMC})) - s\|_\infty$ and the blue line shows $\|\mathcal{O}(u(m^r_{RML})) - s\|_\infty$, where m^r_{MCMC} denotes the r-th realization of the RML-MCMC ensemble, and m^r_{RML} denotes the r-th realization of the RML ensemble, where $r = 1, 2, \cdots, 100$.

4.4.2 2-D Acoustic Wave Equation

We now consider the inverse problem on the 2-D acoustic wave equation,

$$\frac{\partial^2 \hat{u}}{\partial t^2} - e^m \nabla^2 \hat{u} = 0, \quad x \in \Omega := [0,1] \times [0,1] \tag{4.54}$$

with initial conditions

$$\hat{u}(x, y, 0) = u_0(x, y), \quad \hat{u}_t(x, y, 0) = v_0(x, y) \tag{4.55}$$

and homogeneous Neumann boundary conditions. For this problem, we employ the real temporal Laplace transform

$$u(x, y, \gamma) = \int_0^\infty e^{-\gamma t} \hat{u}(x, y, t)\, \mathrm{d}t, \tag{4.56}$$

where $\gamma > 0$ is the Laplace parameter. Note here that the real Laplace transform corresponds to the zero frequency component of the waveform, and that it is difficult to invert this transform, as this is an ill-posed problem (see for instance, Chen (1993)). Nevertheless, we may still use this transform as the focus is on solving the inverse problem of estimating the parameter, and therefore an inverse transform is not necessary.

Applying the Laplace transform to (4.54), we have

$$-e^m \nabla^2 u + \gamma^2 u = \gamma u_0 + v_0. \tag{4.57}$$

We may solve (4.57) using a centred finite difference scheme on a uniform grid. Suppose h is the grid size in both the x- and y-directions, so that we have

$$x_i = ih, \quad i = 0, 1, \cdots, K, \tag{4.58}$$

$$y_j = jh, \quad j = 0, 1, \cdots, K, \tag{4.59}$$

and $h = 1/K$. This yields the linear system

$$F(m)U = f, \tag{4.60}$$

where $F(m) = -e^m A_h + \gamma^2 \mathbf{1}$, A_h is the 2-D discretized Laplacian, and $\mathbf{1}$ is the $(K+1)^2 \times (K+1)^2$ identity matrix. We employ the standard ordering for the vectors U and f, i.e.

$$U_{(K+1)i+j} = u(x_i, y_j), \quad \text{and} \quad f_{(K+1)j+i} = \gamma u_0(x_i, y_j) + v_0(x_i, y_j), \tag{4.61}$$

for $i = 0, \cdots, K$ and $j = 0, \cdots, K$.

We observe the Laplace transform of the displacement on the boundary of the unit square, $\partial\Omega$,

$$s = \mathcal{O}u + \xi, \tag{4.62}$$

where the vector ξ is white noise with component variance σ^2, and $\mathcal{O}$ is the observation operator, given by the block diagonal matrix

$$\mathcal{O} = \begin{pmatrix} \mathbf{1}_{K+1} & & & & \\ & \mathcal{O}' & & & \\ & & \ddots & & \\ & & & \mathcal{O}' & \\ & & & & \mathbf{1}_{K+1} \end{pmatrix}, \tag{4.63}$$

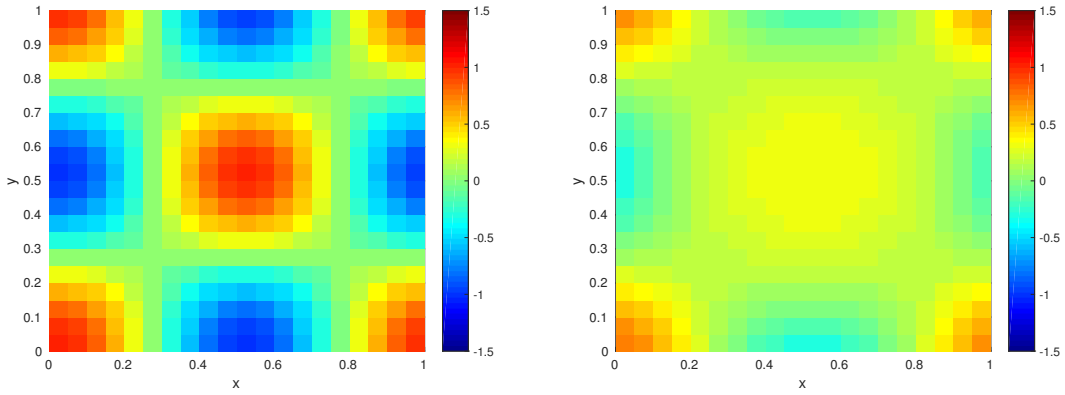
where $\mathcal{O}'$ is a $2 \times (K+1)$ matrix defined by

$$\mathcal{O}' = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}. \tag{4.64}$$

The matrix $\mathcal{O}'$ extracts the points along the line $y = 0$ and $y = 1$ while the identity matrices extract the points along the line $x = 0$ and $x = 1$.

For this problem, we set $K = 20$, with $\gamma = 1$. The truth and probability maximizer is shown in Figures 4.6(a)–(b). The initial conditions are $u_0(x, y) = e^{-2((x-1/2)^2+(y-1/2)^2)}$ and $v_0(x, y) \equiv 0$. We use the modified Helmholtz operator to build the prior precision. Hence, the precision, $\mathcal{L}_0$ takes the form

$$\mathcal{L}_0 = ab^2 - a\nabla^2, \tag{4.65}$$



(a) (b)

Figure 4.6: Plot of $m(x)$ for the inverse problem on the 2-D wave equation: (a) Truth and (b) probability maximizer, obtained by solving an optimization on the misfit functional of the form (4.29) for this problem.

where $a = \lambda/(2\sigma^2)$ and $b = 1/\lambda$. Here, we specify $\sigma = 1/2$ and $\lambda = 1/5$. Recall that the value of a controls the variance, while b controls the correlation length. Higher values of a give lower variances while higher values of b give shorter correlation lengths. We set the prior mean to be identically zero. The observation noise is white noise with precision 10^4, and no noise was added when the observations were generated, but the presence of observation noise is assumed when solving the inverse problem.

For the random walk MCMC, we use the step length, $\beta = 0.075$ and produce chains of length 100.

The precision of the forward problem is scaled according to (4.41). Here, we solve the inverse problem for $\tilde{\lambda} = 1$.

Figures 4.7(a)–(c) show 20 realizations of the prior along with the ensemble mean and standard deviation. Note the rough structure of each realization due to a short correlation length.

Figures 4.8(a)–(c) shows 20 realizations of the RML ensemble, with its ensemble mean and standard deviation, and Figures 4.9 shows 20 realizations of the RML-MCMC ensemble, along with its ensemble mean and standard deviation.

Notice first that the mean of the RML ensemble has begun to take on the structure of the probability maximizer (see Figure 4.8(b) and Figure 4.6(b)), and the standard deviation has not changed much in the middle of the domain (Figure 4.8(c)). This is due to the fact that no observations were made in the interior of the domain. It is therefore difficult to reduce the uncertainty in that region.

The RML-MCMC ensemble produces an ensemble mean which contains more

detail in the interior of the domain (see Figure 4.8(b) and Figure 4.9(b)). Note also that the overall standard deviation of the RML-MCMC ensemble on the boundary is lower than that of the RML ensemble (see Figure 4.8(c) and Figure 4.9(c)).

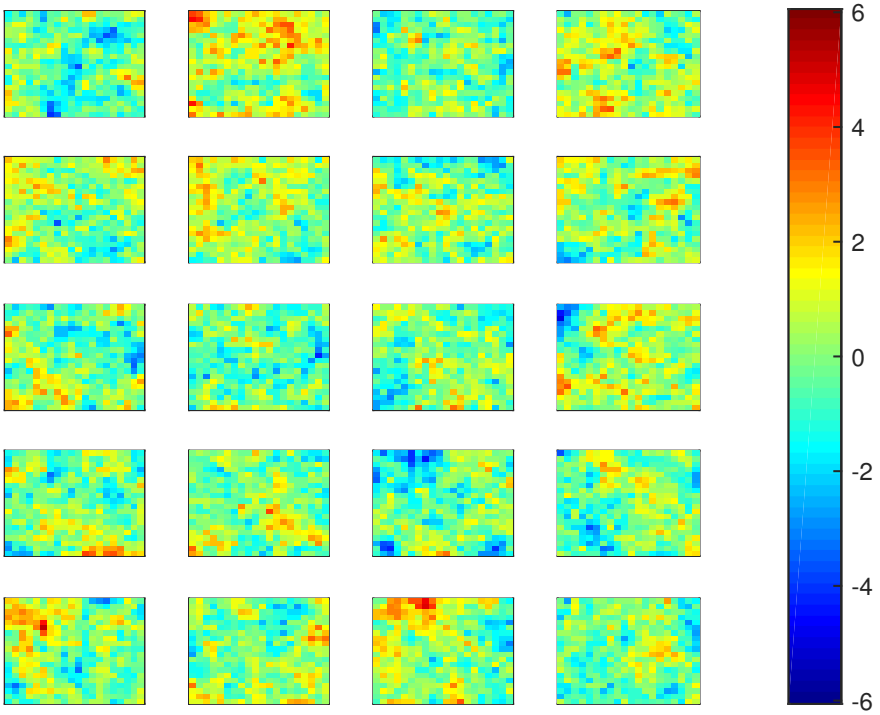
Finally, each realization in Figure 4.9(a) is an equiprobable profile for the parameter $m(x, y)$, which honours the observations more faithfully than the RML ensemble (Figure 4.10). Here, note that the RML ensemble did not achieve an observation error of order 10^{-2} as the well-testing problem did. The increased number of dimensions here has made the optimization problem harder to solve, and therefore one would need to use MCMC to correct the approximate sample generated by the method of RML to get a more accurate sample of the posterior. In this problem, it appears that longer Markov chains are necessary to perturb the RML ensemble so that the observation errors fall to within the standard deviation of the observation noise.

4.5 Concluding Remarks

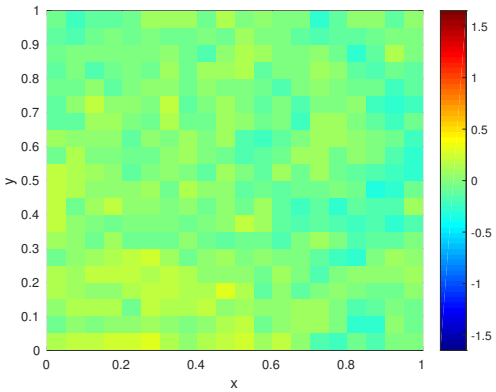
In this chapter, we presented the weak constraint formulation of an inverse problem. The weak constraint approach is an attempt to convexify the misfit functional to ease the task of optimization. Here, we explored the ability to sample from the posterior measure obtained from the weak constraint formulation.

The method chosen for sampling in this chapter is the Randomized Maximum Likelihood (RML) method with a random walk Metropolis-Hastings method, using preconditioned Crank-Nicolson proposals. According to Oliver *et al.* (2008), the method of Randomized Maximum Likelihood allows us to explore the state space more easily at the cost of solving more optimization problems, and allows us to generate samples which honour the observations more readily. Then, in order to make the sample more accurate, a random walk method was used to perturb the ensemble.

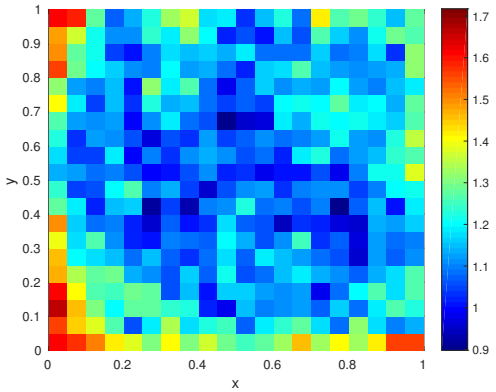
Finally, we applied this method to two problems: the 1-D well testing problem and an inverse problem on the 2-D wave equation. While the preliminary results obtained here are encouraging, in Chapter 6, we will study further the inverse problem on the 2-D wave equation.



(a)

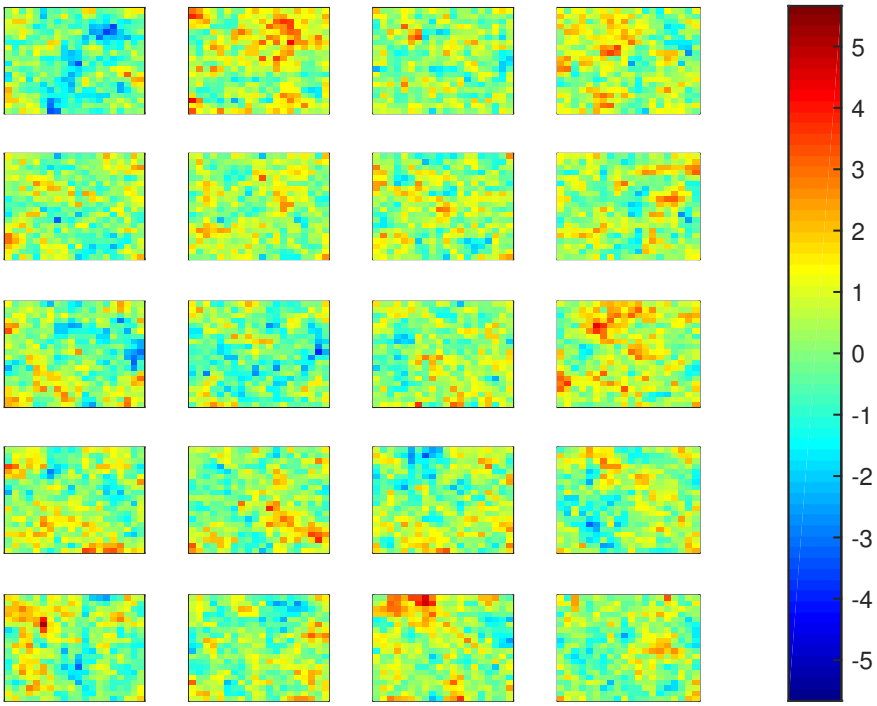


(b)

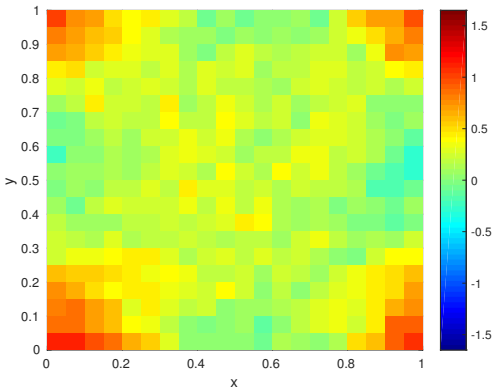


(c)

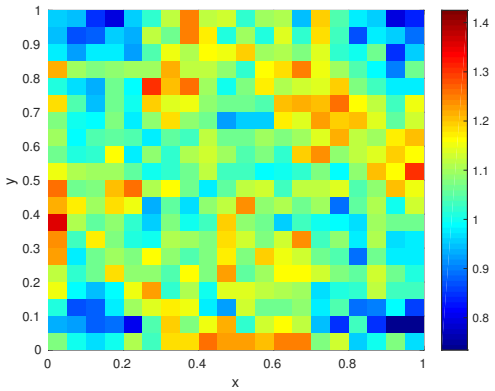
Figure 4.7: (a) 20 realizations of the prior sample; (b) prior ensemble mean, and (c) prior ensemble standard deviation of the inverse problem on the 2-D wave equation.



(a)

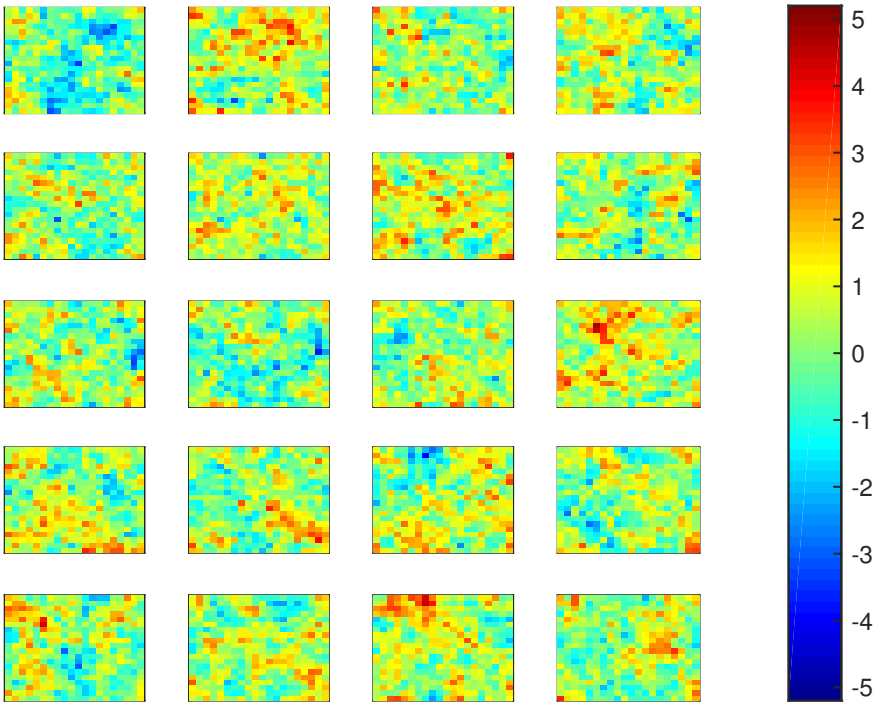


(b)

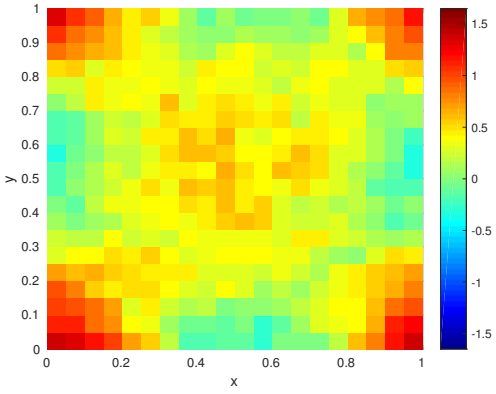


(c)

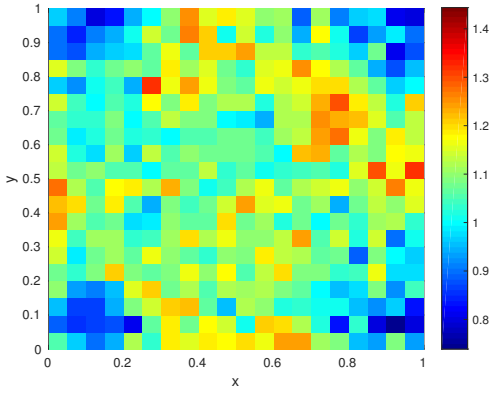
Figure 4.8: (a) 20 realizations of the RML ensemble; (b) RML ensemble mean, and (c) RML ensemble standard deviation of the inverse problem on the 2-D wave equation.



(a)



(b)



(c)

Figure 4.9: (a) 20 realizations of the MCMC ensemble; (b) MCMC ensemble mean, and (c) MCMC ensemble standard deviation of the inverse problem on the 2-D wave equation.

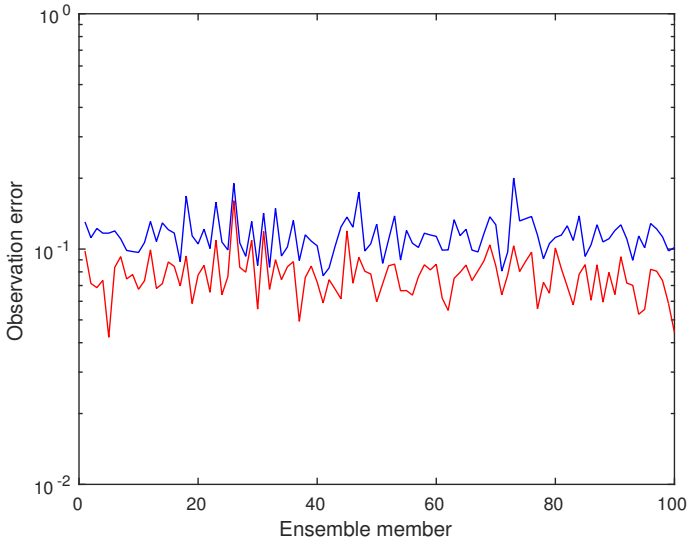


Figure 4.10: Observation errors for the inverse problem on the 2-D wave equation. The red line shows $\|\mathcal{O}(u(m^r_{MCMC})) - s\|_\infty$ and the blue line shows $\|\mathcal{O}(u(m^r_{RML})) - s\|_\infty$, where m^r_{MCMC} denotes the r-th realization of the RML-MCMC ensemble, and m^r_{RML} denotes the r-th realization of the RML ensemble, where $r = 1, 2, \cdots, 100$.

Chapter 5

Comparison between Strong and Weak Constraint Formulations

In this chapter, we consider the inverse problem for the simple harmonic oscillator, and attempt to solve it within the Bayesian framework. The inverse problem for the simple harmonic oscillator is chosen as a model problem as it is both low-dimensional and the exact solution to the forward problem is known. We will compare the results from both the strong and weak constraint formulation of the inverse problem.

5.1 Formulation

Consider a vibrating spring system, where an object is attached to a spring as shown in Figure 5.1. For small amplitude oscillations, the governing equation, given in non-dimensional form, is

$$\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} + 4\pi^2\omega^2 u = 0, \quad t \geq 0, \tag{5.1}$$

with initial conditions $u(0) = 1$, $u'(0) = 0$. Here, the quantity ω is the frequency of the vibration and $u(t)$ is the amplitude. Note that the quantity $4\pi^2$ appears as a results of using frequency rather than angular frequency, which is $2\pi\omega$.

Since we have $\omega > 0$, we may write $\omega = e^m$ for some $m \in \mathbb{R}$. Then, equation (5.1) becomes

$$\frac{\mathrm{d}^2 u}{\mathrm{d}t^2} + 4\pi^2 e^{2m} u = 0. \tag{5.2}$$

Furthermore, we can write (5.2) as a system of first order ordinary differential equations. If we define the quantity

$$v = \frac{\mathrm{d}u}{\mathrm{d}t}, \tag{5.3}$$

we have

$$\begin{cases} \dfrac{\mathrm{d}u}{\mathrm{d}t} - v = 0, \\ \dfrac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m} u = 0, \end{cases} \tag{5.4}$$

with $(u(0), v(0))^T = (1, 0)^T$.

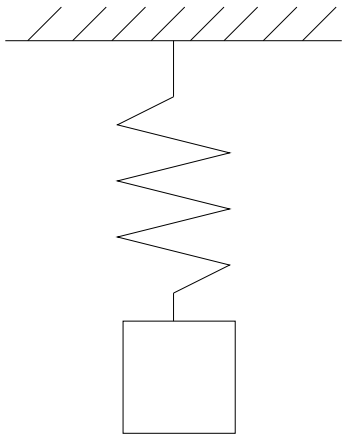


Figure 5.1: Simple harmonic oscillator.

We observe the displacement of the object at discrete times t_n, where $n = 1, \cdots, N$, with observations made at uniform time intervals. We write

$$t_n = n\tau, \quad n = 1, \cdots, N, \tag{5.5}$$

with $\tau = 1/N$.

Then, our observation, s is a vector of size N, with entries given by

$$s_n = u(t_n) + \xi_n, \quad \xi_n \sim \mathcal{N}(0, 1/p), \tag{5.6}$$

for $n = 1, \cdots, N$, and some $p > 0$.

We will apply both the strong and weak constraint methods, as described in Chapter 4, and compare their results below. For each simulation below, we will assume that $m_t = 0$, where m_t is the truth, and is equivalent to $\omega = 1$. Therefore, we have

$$s_n = \cos(2\pi t_n) + \xi_n = \cos(2\pi n\tau) + \xi_n. \tag{5.7}$$

5.2 Solution by strong constraint

The system (5.4) has an exact solution, given by

$$u(t) = \cos(2\pi e^m t). \tag{5.8}$$

Then, the likelihood is given by

$$\rho(s \mid m) \propto \exp(-\Phi(m)), \tag{5.9}$$

where the likelihood potential, $\Phi(m)$, is given by

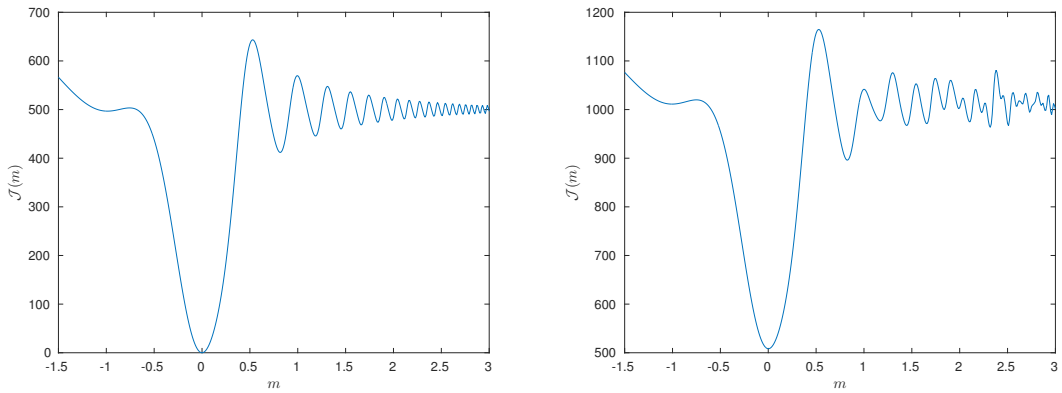
$$\Phi(m) = \frac{p}{2}\sum_{n=1}^{N}\left(\cos(2\pi n\tau e^m) - s_n\right)^2 = \frac{p}{2}\sum_{n=1}^{N}\left(\cos(2\pi n\tau e^m) - \cos(2\pi n\tau)\right)^2. \tag{5.10}$$

Let $p_0 > 0$ be the prior precision for this problem, so that the misfit functional is

$$\mathcal{J}(m) = \frac{p}{2}\sum_{n=1}^{N}\left(\cos(2\pi n\tau e^m) - s_n\right)^2 + \frac{p_0}{2}(m - m_0)^2, \tag{5.11}$$

where m_0 is the prior mean.

We show plots for the misfit for $p = 1$, $p_0 = 10^{-4}$, $m_0 = 1/2$, and $N = 1000$. Figure 5.2(a) shows the plot of the misfit when noise has not been added into the observations, while in Figure 5.2(b), noise has been added into the observations.



(a) (b)

Figure 5.2: Plots of the strong constraint misfit functional, $\mathcal{J}(m)$ of (5.10). In (a), no noise was added to the observation, while noise was added in (b).

In both cases, the posterior is multimodal, since it contains multiple local minima. Note that the oscillations in the misfit are due to the cosine terms coming from the likelihood potential. From Figure 5.2(a), we can see that a global minimum is attained near $m = 0$. In fact, there are infinitely many global minima of the likelihood potential, due to aliasing. This occurs whenever each term in the summation of (5.10) vanishes, i.e.,

$$\cos(2\pi n\tau e^m) - \cos(2\pi n\tau) = 0, \quad n = 1, \cdots, N. \tag{5.12}$$

Applying the sum-to-product formula for trigonometric identities, we have, for all n,

$$\sin(\pi n\tau(e^m + 1)) \sin(\pi n\tau(e^m - 1)) = 0. \tag{5.13}$$

If we have

$$\sin(\pi n\tau(e^m + 1)) = 0, \tag{5.14}$$

then this implies that

$$n\tau(e^m + 1) \in \mathbb{Z}. \tag{5.15}$$

Since multiplication is closed under $\mathbb{Z}$, we deduce that

$$\tau(e^m + 1) = k, \quad k \in \mathbb{Z}, \tag{5.16}$$

and therefore

$$e^m = Nk - 1, \quad k \in \mathbb{Z}. \tag{5.17}$$

Similarly, if we have

$$\sin(\pi n\tau(e^m - 1)) = 0, \tag{5.18}$$

then this gives

$$e^m = Nk + 1, \quad k \in \mathbb{Z}. \tag{5.19}$$

Therefore, the frequencies which are aliased are at $Nk \pm 1$, for $k \in \mathbb{Z}$, $k > 0$. Hence, the global minima for the likelihood potential are given by $\ln(Nk \pm 1)$ for $k \in \mathbb{Z}$, $k > 0$. Note that this is not visible in Figure 5.2(a) as the first aliased point is at $\ln(99) \approx 4.5951$, which is outside the domain of the plot of the misfit functional. However, in the misfit, due to the addition of the prior term, the aliased points will have a higher function value, leaving the regularized misfit with only one global minimum.

Turning our attention to the plot of the misfit with noisy observations (Figure 5.2(b)), once again we see the multimodal nature of the posterior. In this case, since noise has been added into the observations, the oscillations are more irregular. Note also that due to the discrepancy between the noisy observation and the observation predicted by the forward model, the misfit has shifted away from zero, as can be seen on the vertical axis, in particular at $m = 0$.

5.3 Solution by weak constraint

Solving the inverse problem using the weak constraint formulation gives rise to the misfit functional

$$\mathcal{K}(u,v,m) = \frac{p}{2}\sum_{n=1}^{N}(u(t_n)-s_n)^2 + \frac{p_f}{2}\left(\left\|\frac{\mathrm{d}u}{\mathrm{d}t} - v\right\|_{L^2}^2 + \left\|\frac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m}u\right\|_{L^2}^2\right) + \frac{p_0}{2}(m-m_0)^2, \tag{5.20}$$

for some $p_f > 0$.

The misfit functional can also be written as

$$\mathcal{K}(u,v,m) = \frac{p}{2}\sum_{n=1}^{N}(u(t_n)-s_n)^2 + \frac{p_f}{2}\int\left(\left(\frac{\mathrm{d}u}{\mathrm{d}t} - v\right)^2 + \left(\frac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m}u\right)^2\right)\,\mathrm{d}t + \frac{p_0}{2}(m-m_0)^2. \tag{5.21}$$

Rewriting the first term as an integral involving Dirac delta functions, we have

$$\mathcal{K}(u,v,m) = \frac{p}{2}\sum_{n=1}^{N}\int(u(t)-s_n)^2\delta(t-t_n)\,\mathrm{d}t + \frac{p_f}{2}\int\left(\left(\frac{\mathrm{d}u}{\mathrm{d}t} - v\right)^2 + \left(\frac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m}u\right)^2\right)\,\mathrm{d}t + \frac{p_0}{2}(m-m_0)^2. \tag{5.22}$$

Combining the integrals gives

$$\mathcal{K}(u,v,m) = \frac{1}{2}\int\left[p\sum_{n=1}^{N}(u(t)-s_n)^2\delta(t-t_n) + p_f\left(\left(\frac{\mathrm{d}u}{\mathrm{d}t} - v\right)^2 + \left(\frac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m}u\right)^2\right)\right]\,\mathrm{d}t + \frac{p_0}{2}(m-m_0)^2. \tag{5.23}$$

Given an m, we can compute the optimal values of the parameters, $\breve{u}(m)$ and $\breve{v}(m)$ respectively. For test functions η_1 and η_2, the first variations of $\mathcal{K}$ with respect to u and v, respectively, are given by

$$\delta_u\mathcal{K}(u,v,m) = \int\left[p\sum_{n=1}^{N}(u-s_n)\delta(t-t_n) + p_f\left((-u'' + 16\pi^4 e^{4m}u) + (1+4\pi^2 e^{2m})v'\right)\right]\eta_1\,\mathrm{d}t, \tag{5.24}$$

and

$$\delta_v\mathcal{K}(u,v,m) = \int p_f\left(-v'' + v - (1+4\pi^2 e^{2m})u'\right)\eta_2\,\mathrm{d}t \tag{5.25}$$

In order to solve for $\breve{u}(m)$ and $\breve{v}(m)$ for a given m, we set both $\delta_u\mathcal{K}$ and $\delta_v\mathcal{K}$ to zero. This yields a system of differential equations, given by

$$p\sum_{n=1}^{N}(\breve{u}-s_n)\delta(t-t_n)+p_f\left((-\breve{u}''+16\pi^4e^{4m}\breve{u})+(1+4\pi^2e^{2m})\breve{v}'\right)=0 \tag{5.26}$$

$$p_f\left(-\breve{v}''+\breve{v}-(1+4\pi^2e^{2m})\breve{u}'\right)=0, \tag{5.27}$$

where $'$ denotes the time derivatives. Rearranging gives

$$\left(p\sum_{n=1}^{N}\delta(t-t_n)+16\pi^4p_fe^{4m}\right)\breve{u}-p_f\breve{u}''+p_f(1+4\pi^2e^{2m})\breve{v}'=p\sum_{n=1}^{N}s_n\delta(t-t_n) \tag{5.28}$$

$$p_f(-\breve{v}''+\breve{v})-p_f(1+4\pi^2e^{2m})\breve{u}'=0. \tag{5.29}$$

In matrix-vector form, this is

$$\begin{pmatrix}\delta^2_{u,u}\mathcal{K} & \delta^2_{u,v}\mathcal{K}\\ \delta^2_{v,u}\mathcal{K} & \delta^2_{v,v}\mathcal{K}\end{pmatrix}\begin{pmatrix}\breve{u}\\ \breve{v}\end{pmatrix}=\begin{pmatrix}p\sum_n s_n\delta(t-t_n)\\ 0\end{pmatrix}, \tag{5.30}$$

where the terms in the matrix are given by

$$\delta^2_{u,u}\mathcal{K}=p\sum_n\delta(t-t_n)+16\pi^4p_fe^{4m}-p_f\frac{\mathrm{d}^2}{\mathrm{d}t^2}, \tag{5.31}$$

$$\delta^2_{u,v}\mathcal{K}=p_f(1+4\pi^2e^{2m})\frac{\mathrm{d}}{\mathrm{d}t}, \tag{5.32}$$

$$\delta^2_{v,u}\mathcal{K}=-p_f(1+4\pi^2e^{2m})\frac{\mathrm{d}}{\mathrm{d}t}, \tag{5.33}$$

$$\delta^2_{v,v}\mathcal{K}=p_f-p_f\frac{\mathrm{d}^2}{\mathrm{d}t^2}. \tag{5.34}$$

We can then solve this linear system, which gives the optimal values of u and v in terms of m. Then, an optimal value of m can be obtained by solving the optimization problem

$$\breve{m}=\arg\min_m\mathcal{I}(m), \tag{5.35}$$

where $\mathcal{I}(m)$ is defined as

$$\mathcal{I}(m):=\mathcal{K}(m,\breve{u}(m),\breve{v}(m)). \tag{5.36}$$

The functional $\mathcal{I}(m)$ is called the *reduced misfit functional*, or simply the reduced functional. Note that in practice, the reduced misfit functional is defined in a finite-dimensional setting, where linear system in (5.30) and the integrals in (5.21) are discretized.

Therefore, if we discretize the domain $[0,1]$ into uniform intervals of size τ, the integrals discretize into

$$\int_0^1 \left(\frac{\mathrm{d}u}{\mathrm{d}t} - v\right)^2 \,\mathrm{d}t \approx \sum_{n=1}^{N} \left(\frac{u_n - u_{n-1}}{\tau} - v_n\right)^2 \tau, \tag{5.37}$$

and

$$\int_0^1 \left(\frac{\mathrm{d}v}{\mathrm{d}t} + 4\pi^2 e^{2m} u\right)^2 \,\mathrm{d}t \approx \sum_{n=1}^{N} \left(\frac{v_n - v_{n-1}}{\tau} + 4\pi^2 e^{2m} u_n\right)^2 \tau, \tag{5.38}$$

where we now use u and v to denote either column vectors of size N with entries u_n and v_n respectively for $n = 1, \cdots, N$, or their continuous forms, whichever is appropriate in the context.

The discretized integrals in (5.37) and (5.38) can be written compactly in matrix and vector forms. We have

$$\sum_{n=1}^{N} (u_n - u_{n-1} - \tau v_n)^2 = \|Du - u_0 - \tau v\|_2^2, \tag{5.39}$$

, and

$$\sum_{n=1}^{N} \left(v_n - v_{n-1} + 4\pi^2 \tau e^{2m} u_n\right)^2 = \left\|Dv + 4\pi^2 \tau e^{2m} u\right\|_2^2, \tag{5.40}$$

where the matrix $D \in \mathbb{R}^{N\times N}$ and vector $u_0 \in \mathbb{R}^N$ have entries

$$D = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{pmatrix}, \quad u_0 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{5.41}$$

and entries not displayed are zero.

Hence, the discretized weak constraint misfit functional, denoted by $\mathcal{K}_\tau$, is given by

$$\begin{aligned} \mathcal{K}_\tau(u, v, m) = \frac{p}{2} \|u - s\|_2^2 + \frac{p_f}{2\tau} \left(\|Du - u_0 - \tau v\|_2^2 + \left\|Dv + 4\pi^2 \tau e^{2m} u\right\|_2^2\right) \\ + \frac{p_0}{2}(m - m_0)^2. \end{aligned} \tag{5.42}$$

The linear system in (5.30) discretizes to be

$$\begin{pmatrix} \nabla^2_{u,u}\mathcal{K}_\tau & \nabla^2_{u,v}\mathcal{K}_\tau \\ (\nabla^2_{u,v}\mathcal{K}_\tau)^T & \nabla^2_{v,v}\mathcal{K}_\tau \end{pmatrix} \begin{pmatrix} \breve{u} \\ \breve{v} \end{pmatrix} = \begin{pmatrix} ps + (p_f/\tau) D^T u_0 \\ -p_f u_0 \end{pmatrix}, \tag{5.43}$$

where the second derivatives are given by

$$\nabla^2_{u,u}\mathcal{K}_\tau = p\mathbf{1} + (p_f/\tau)(D^T D + 16\pi^4 \tau^2 e^{4m} \mathbf{1}), \tag{5.44}$$

$$\nabla^2_{u,v}\mathcal{K}_\tau = p_f(4\pi^2 e^{2m} D - D^T), \tag{5.45}$$

$$\nabla^2_{v,v}\mathcal{K}_\tau = (p_f/\tau)(\tau^2 \mathbf{1} + D^T D), \tag{5.46}$$

and **1** is the appropriate identity matrix. We can then define the reduced functional

$$\mathcal{I}_\tau(m) = \mathcal{K}_\tau(m, \breve{u}(m), \breve{v}(m)). \tag{5.47}$$

We plot the reduced functional $\mathcal{I}_\tau(m)$ using the same values of p, p_0, m_0 and N as used in the strong constraint formulation of Section 5.2. Figures 5.3(a)–(d) show the plots of $\mathcal{I}_\tau(m)$ for different values of p_f for noise-free observations. Note that for smaller values of p_f, the reduced misfit functional is smoother and more convex than in the strong constraint case. Note also that as p_f increases, $\mathcal{I}_\tau(m)$ begins to resemble $\mathcal{J}(m)$ (cf. Figure 5.2(a)). For this problem, note that for $p_f = 1000$, the reduced misfit tends to a value near 250 as m increases. This is due to the increased numerical errors from the discretization of the integrals.

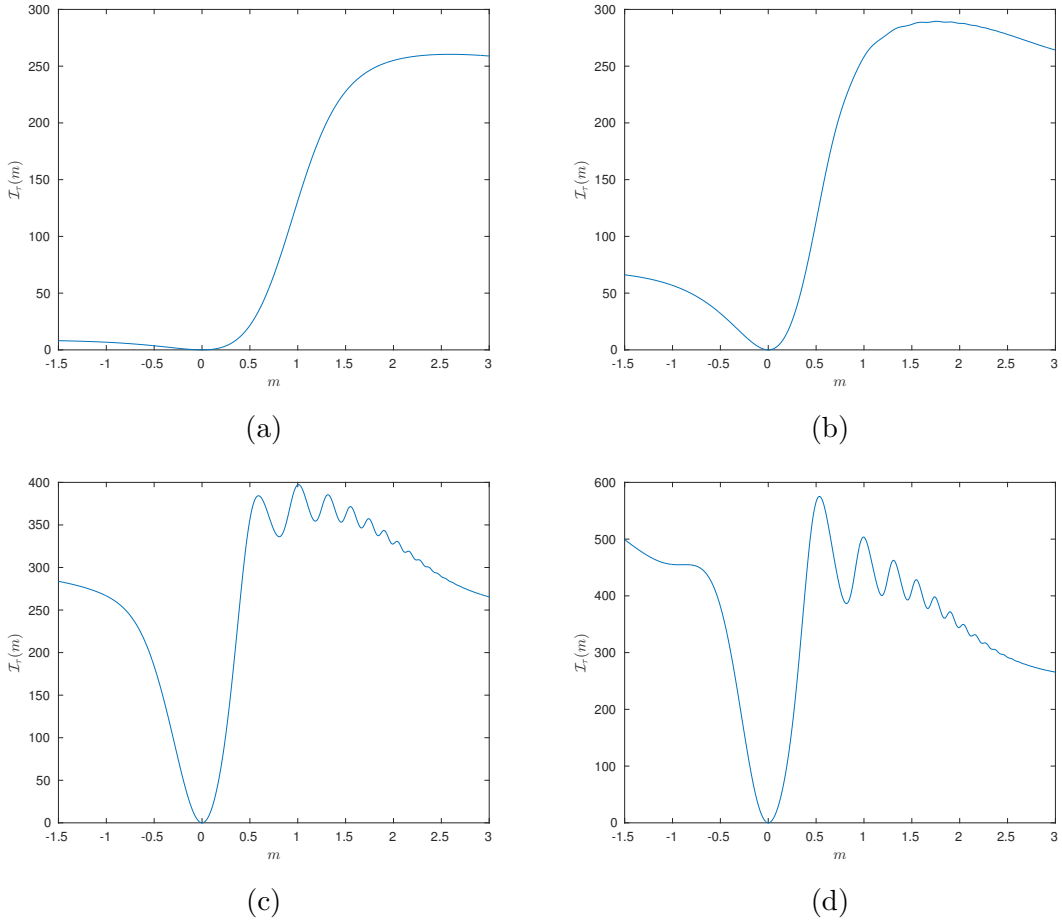


Figure 5.3: Plot of the weak constraint reduced misfit functional, $\mathcal{I}_\tau(m)$ with noise-free observation for (a) $p_f = 1$, (b) $p_f = 10$, (c) $p_f = 100$, and (d) $p_f = 1000$.

Figure 5.4(a)–(d) show the plots of the reduced misfits where noise has been added to the observation. We see that the noise in the observation has not affected the convexity near the truth, and we are still able to capture the overall features of the misfit.

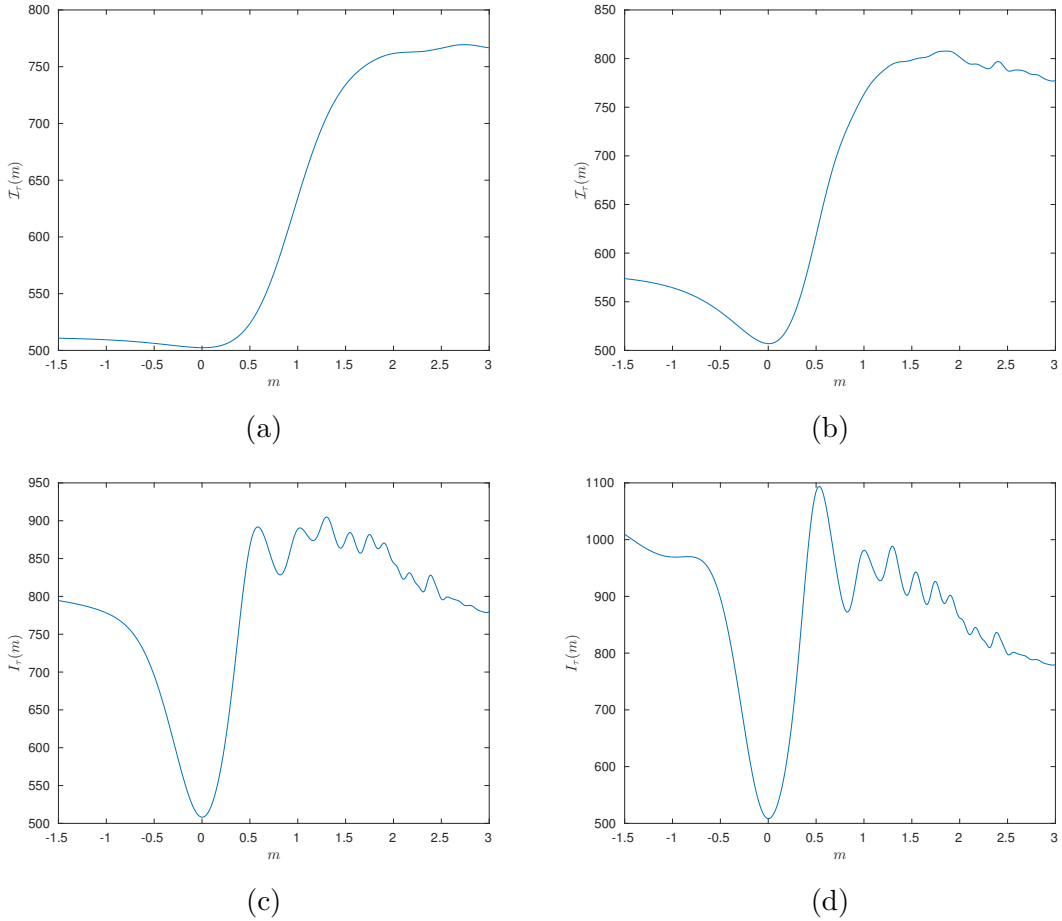


Figure 5.4: Plot of the weak constraint reduced misfit $\mathcal{I}_\tau(m)$ with noisy observation for (a) $p_f = 1$, (b) $p_f = 10$, (c) $p_f = 100$, (d) $p_f = 1000$.

5.4 Concluding Remarks

In this chapter we presented both the strong and weak constraint approach for the inverse problem on the simple harmonic oscillator. While the inverse problem is simple and low-dimensional, it serves as a model problem to study and provides a good basis for which new methods can be tested. Due to its simplicity, we were able to perform some analysis in continuous form before finally discretizing the misfits in order to implement them. The analysis here shows that the weak constraint formulation of an inverse problem indeed does lead to a more convex misfit functional than the misfit functional arising from the strong constraint formulation. Note that if we had discretized the integrals in (5.21) and performed the analysis in a finite-dimensional setting, we would have obtained the same results as above, provided the same schemes were used in the discretization.

Finally, note that one could also form the weak constraint misfit functional as

$$\mathcal{K}(u,m) = \frac{p}{2}\sum_{n=1}^{N}(u(t_n)-s_n)^2 + \frac{p_f}{2}\int (u(t)-\cos(2\pi e^m t))^2 \,\mathrm{d}t + \frac{p_0}{2}(m-m_0)^2, \quad (5.48)$$

where the solution of the forward problem has been used in the second term of the misfit functional instead of the forward problem itself. Its discretized form is given by

$$\mathcal{K}_\tau(u,m) = \frac{p}{2}\left\|u-s\right\|_2^2 + \frac{p_f\tau}{2}\left\|u-\psi\right\|_2^2 + \frac{p_0}{2}(m-m_0)^2 \quad (5.49)$$

where $\psi(m)$ is a vector of size N with entries

$$\psi_n(m) = \cos(2\pi n\tau e^m). \quad (5.50)$$

Then, for a given m, we have

$$\breve{u}(m) = \frac{1}{p+p_f\tau}(ps + p_f\tau\psi(m)). \quad (5.51)$$

Therefore the discrete reduced functional, $\mathcal{I}_\tau(m)$, can be written as

$$\mathcal{I}_\tau(m) = \frac{p}{2}\left\|\breve{u}(m)-s\right\|_2^2 + \frac{p_f\tau}{2}\left\|\breve{u}(m)-\psi(m)\right\|_2^2 + \frac{p_0}{2}(m-m_0)^2, \quad (5.52)$$

Substituting (5.51) into (5.52) and simplifying, we have

$$\mathcal{I}_\tau(m) = \frac{pp_f\tau}{2}\sum_{n=1}^{N}(\cos(2\pi n\tau e^m)-s_n)^2 + \frac{p_0}{2}(m-m_0)^2, \quad (5.53)$$

which is exactly the same as the strong constraint misfit functional in (5.11), except that the first term, coming from the likelihood, has been scaled. Therefore, this form of the weak constraint formulation brings no advantage in terms of dealing with the non-convexity that was in the strong constraint formulation, and crucially, the computation of the gradient still requires the computation of the adjoints. This is not the case in the weak constraint formulation as presented in Chapter 4.

Chapter 6

Weak Constraint Formulation of Full Waveform Inversion

6.1 Introduction

In this chapter, we will present the weak constraint formulation of the problem of full waveform inversion. Most of the research in full waveform inversion today solves the problem in the frequency domain, employing a Fourier transform. Here, we will use a Laplace transform, where the frequency component will have non-zero real and imaginary parts, as opposed to the purely imaginary Fourier transform. This method was explored by Shin and Cha (2009). Performing the inversion in the Laplace domain gives us more flexibility in terms of choosing the frequency components, and in any case, if one prefers to use the Fourier transform instead, the real component of the frequency can be set to zero.

6.2 Formulation

6.2.1 Forward Problem

The forward problem is the wave equation

$$\frac{1}{c(x,y)^2}\frac{\partial^2\hat{\phi}}{\partial t^2} - \nabla^2\hat{\phi} = \hat{q}, \quad (x,y) \in \Omega := [0,1]\times[0,1], \tag{6.1}$$

where $c(x,y)$ is the wave speed, $\hat{\phi}(x,y,t)$ is the pressure, and $\hat{q}(x,y,t)$ is the source, with boundary conditions

$$(\nabla\hat{\phi})\cdot n = 0 \quad \text{if } y = 0, \tag{6.2}$$

$$\frac{1}{c}\frac{\partial\hat{\phi}}{\partial t} + (\nabla\hat{\phi})\cdot n = 0 \quad \text{on } \partial\Omega\backslash\{y=0\}, \tag{6.3}$$

and zero initial conditions.

Typically, the full waveform inversion is solved using the Laplace transform in t of the wave equation as the forward problem. This is given by

$$\frac{\omega^2}{c^2}\widetilde{\phi}_\omega - \nabla^2\widetilde{\phi}_\omega = \widetilde{q}_\omega, \tag{6.4}$$

where

$$\widetilde{\phi}_\omega(x,y) = \int_0^\infty \hat{\phi}(x,y,t)e^{-\omega t}\,\mathrm{d}t, \quad \text{and} \quad \widetilde{q}_\omega(x,y) = \int_0^\infty \hat{q}(x,y,t)e^{-\omega t}\,\mathrm{d}t, \tag{6.5}$$

and $\omega = \alpha + 2\pi i\beta$, for $\alpha > 0$ and $\beta > 0$. Note here that β is different from the β that is used to denote the MCMC step length.

We set $e^{m(x,y)} = 1/c(x,y)^2$, which gives

$$\omega^2 e^m\widetilde{\phi}_\omega - \nabla^2\widetilde{\phi}_\omega = \widetilde{q}_\omega. \tag{6.6}$$

We write $\widetilde{\phi}_\omega = u_\omega + iv_\omega$ and $\widetilde{q}_\omega = f_\omega + ig_\omega$. Then, expanding and comparing real and imaginary parts of (6.6), we have

$$(\alpha^2 - 4\pi^2\beta^2)e^m u_\omega - 4\pi\alpha\beta e^m v_\omega - \nabla^2 u_\omega = f_\omega, \tag{6.7}$$

$$4\pi\alpha\beta e^m + (\alpha^2 - 4\pi^2\beta^2)e^m v_\omega - \nabla^2 v_\omega = g_\omega. \tag{6.8}$$

In matrix form, this is

$$e^m \begin{pmatrix} \alpha^2 - 4\pi^2\beta^2 & -4\pi\alpha\beta \\ 4\pi\alpha\beta & \alpha^2 - 4\pi^2\beta^2 \end{pmatrix} \begin{pmatrix} u_\omega \\ v_\omega \end{pmatrix} - \nabla^2 \begin{pmatrix} u_\omega \\ v_\omega \end{pmatrix} = \begin{pmatrix} f_\omega \\ g_\omega \end{pmatrix}. \tag{6.9}$$

The Neumann boundary conditions become

$$(\nabla u_\omega)\cdot n = 0, \quad \text{and} \quad (\nabla v_\omega)\cdot n = 0, \tag{6.10}$$

on $y = 0$, and the absorbing boundary conditions become

$$e^{m/2} \begin{pmatrix} \alpha & -2\pi\beta \\ 2\pi\beta & \alpha \end{pmatrix} \begin{pmatrix} u_\omega \\ v_\omega \end{pmatrix} + \nabla \begin{pmatrix} u_\omega \\ v_\omega \end{pmatrix} \cdot n = 0, \tag{6.11}$$

on $y = 1$.

The forward model (6.9), (6.10) and (6.11) is solved using a second-order centred finite difference scheme on a uniform grid. The grid size is $h = 1/K$, where K is the number of cells in the x- and y-directions. Then, a grid point (x_i, y_j) is given by

$$x_i = ih, \quad i = 0, \cdots, K, \tag{6.12}$$

$$y_j = jh, \quad j = 0, \cdots, K. \tag{6.13}$$

This leads to the linear system

$$\left[\begin{pmatrix}(\alpha^2-4\pi^2\beta^2)E(m) & -4\pi\alpha\beta E(m)\\ 4\pi\alpha\beta E(m) & (\alpha^2-4\pi^2\beta^2)E(m)\end{pmatrix}+\begin{pmatrix}A_h & 0\\ 0 & A_h\end{pmatrix}\right]\begin{pmatrix}U_\omega\\ V_\omega\end{pmatrix}=\begin{pmatrix}F_\omega\\ G_\omega\end{pmatrix}, \tag{6.14}$$

where A_h is the discretized negative Laplacian, and U_ω and V_ω are the discretizations of u_ω and v_ω respectively. The matrix $E(m)$ is defined as follows: The vector m is defined by the standard ordering, which is

$$m_{i(K+1)+j} = m(x_i, y_j), \tag{6.15}$$

where the $m(x, y)$ on the right-hand side of (6.15) is the continuous form while the m on the left-hand side of (6.15) is the discretized form. For the rest of this chapter, unless otherwise specified, we shall use m to denote the discretized form of the parameter.

Then, we define the matrix $E(m)$ to be a diagonal matrix, which has entries

$$E(m)_{ii} = e^{m_i}. \tag{6.16}$$

The vectors F_ω and G_ω are the discretizations of f_ω and g_ω respectively. The vectors U_ω, V_ω, F_ω, G_ω are ordered by the standard ordering, as explained in Section 4.4.2. Finally, we can write (6.14) in the compact form

$$\mathcal{A}_\omega(m)\phi_\omega - q_\omega = 0. \tag{6.17}$$

For the purposes of the weak constraint formulation, we add noise and use the forward problem

$$\mathcal{A}_\omega(m)\phi_\omega - q_\omega = \zeta_\omega, \tag{6.18}$$

where $\zeta_\omega \sim \mathcal{N}(0, \Gamma_f)$, for some covariance matrix Γ_f.

6.2.2 Observation model

We observe ϕ_ω at every grid point on the surface, where $y = 0$. Let r_ω and s_ω denote the real and imaginary parts of the observed Laplace transform of the wavefield respectively. Then, we have

$$(r_\omega)_k = u_\omega(x_k, 0) + \eta_k, \tag{6.19}$$

$$(s_\omega)_k = v_\omega(x_k, 0) + \theta_k, \tag{6.20}$$

where $\eta_k \sim \mathcal{N}(0, 1/p)$ and $\theta_k \sim \mathcal{N}(0, 1/p)$.

We write the observation model compactly as

$$\psi_\omega = \mathcal{O}\phi_\omega + \xi, \tag{6.21}$$

where $\psi_\omega = (r_\omega^T, s_\omega^T)^T$, $\xi = (\eta^T, \theta^T)^T$, $\xi \sim \mathcal{N}(0, (1/p)\mathbf{1})$, where $\mathbf{1}$ is the appropriate identity matrix, the superscript T denotes a transpose, and $\mathcal{O}$ is the matrix defined by

$$\mathcal{O} = \begin{pmatrix} \mathcal{O}' & \\ & \mathcal{O}' \end{pmatrix}, \tag{6.22}$$

where the matrix $\mathcal{O}'$ is a $(K+1) \times (K+1)^2$ observation matrix defined by

$$\mathcal{O}'_{ij} = \begin{cases} 1 & \text{if } j = 1 + (i-1)(K+1), \\ 0 & \text{otherwise.} \end{cases} \tag{6.23}$$

As an example, when $K = 2$, the matrix $\mathcal{O}'$ is given by

$$\mathcal{O}' = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}. \tag{6.24}$$

6.3 Single frequency inversion

Before we can understand how inversion is done at several frequencies, we first study the inversion in a single frequency. Suppose we have a prior measure for m, $\mu_0 = \mathcal{N}(0, \mathcal{C}_0)$, and an observation ψ_ω, such that the likelihood potential is given by

$$\Phi_\omega(\phi_\omega) = \frac{p}{2} \left\| \psi_\omega - \mathcal{O}\phi_\omega \right\|_2^2. \tag{6.25}$$

This gives rise to the posterior measure, μ, characterized by the weak constraint misfit functional

$$\mathcal{K}_\omega(\phi_\omega, m) = \frac{p}{2} \left\| \psi_\omega - \mathcal{O}\phi_\omega \right\|_2^2 + \frac{1}{2} \left\| \mathcal{A}_\omega(m)\phi_\omega - q_\omega \right\|_{\Gamma_f}^2 + \frac{1}{2} \left\| m - m_0 \right\|_{\mathcal{C}_0}^2. \tag{6.26}$$

The reduced misfit functional, $\mathcal{I}_\omega(m)$, associated with $\mathcal{K}_\omega(\phi_\omega, m)$, is then given by

$$\mathcal{I}_\omega(m) = \mathcal{K}_\omega(\breve{\phi}_\omega(m), m), \tag{6.27}$$

where $\breve{\phi}_\omega$ is given by

$$\breve{\phi}_\omega(m) = (p\mathcal{O}^T\mathcal{O} + \mathcal{A}_\omega(m)^T L_f \mathcal{A}_\omega(m))^{-1}(p\mathcal{O}^T\psi_\omega + \mathcal{A}_\omega(m)^T L_f q_\omega). \tag{6.28}$$

In order to understand this misfit functional, we plot it along a line parametrized by two points in the parameter space. In particular, if we have $m_1 \in \mathcal{H}_m$, $m_2 \in \mathcal{H}_m$, where $\mathcal{H}_m$ is the parameter space as defined in Chapter 2, then we plot the one-dimensional function $\mathcal{M}(\lambda; m_1, m_2)$, defined as

$$\mathcal{M}(\lambda; m_1, m_2) = \mathcal{I}_\omega(\lambda m_1 + (1-\lambda)m_2), \quad \lambda \in [0, 1]. \tag{6.29}$$

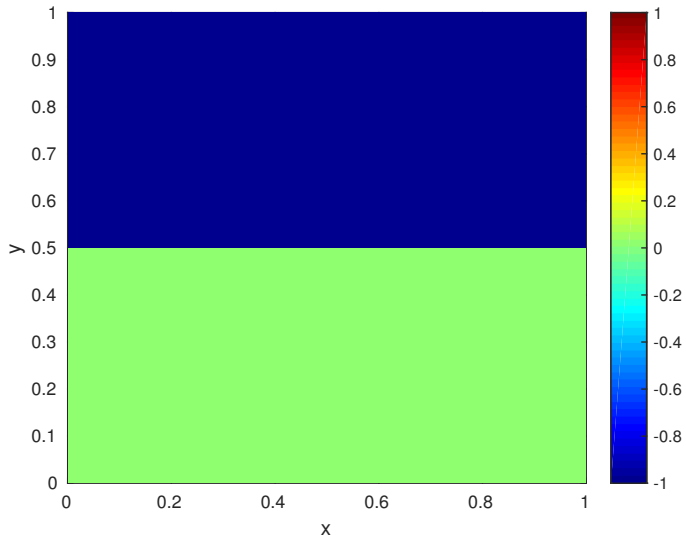


Figure 6.1: The truth used in plotting the misfit functional (6.29), where $m_2 = 0.2$ and m_1 is the truth.

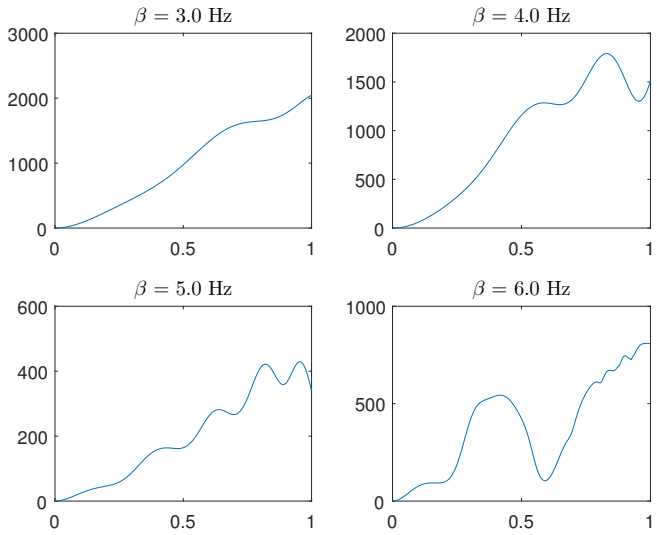


Figure 6.2: The misfit functional, $\mathcal{I}(m)$, plotted on a line segment between two points m_1 and m_2 in the parameter space for $\alpha = 0.01$ and varying values of β. The abscissa and ordinate for each plots are λ and $\mathcal{M}(\lambda)$ respectively. We use m_1 as the truth as shown in Figure 6.1 and $m_2 = 0.2$.

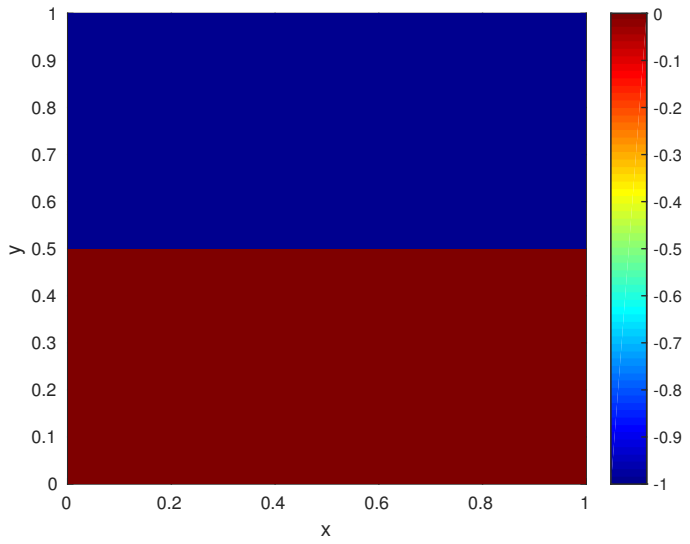


Figure 6.3: Truth used in the examples below.

In this example, we used m_1 as the truth (see Figure 6.1) and $m_2 = 0.2$. In Figure 6.2, the misfit functional (6.29) is plotted for a fixed value of α and varying values of β. We can see that as the frequency, β increases, the convexity of the misfit decreases.

Next, we would like to explore the difficulty of finding a probability maximizer. For this purpose, we use the same truth as in Figure 6.1, reproduced in Figure 6.3 using a different colormap. We solve the minimization problem to find a minimum of $\mathcal{I}(m)$, using different starting iterates, as shown in Figure 6.4(a)–(d). The prior specified is a modified Helmholtz prior measure with hyperparameters $a = 1$, $\Delta_B = 1$ and $\nu_B = 0.25$.

Figure 6.5 shows the solution for the optimization problem on $\mathcal{I}(m)$ for $\alpha = 0.01$ and $\beta = 0.1$, using the four starting iterates as shown in Figure 6.4. For these starting iterates, due to the inversion being performed in a low frequency, the misfit functional is more convex and therefore each starting iterate produces similar results.

Figures 6.6(a)–(d) show the results obtained from solving the minimization problem for $\alpha = 0.01$ and $\beta = 6$. Note that different starting iterates produced different solutions, although the third and fourth starting iterates produced the same minimizer. This is an indication of the multimodality of the posterior measure at higher frequencies.

Within the context of finding the global minimum of the misfit functional, an important subject of research in full waveform inversion is therefore to find a starting iterate that is sufficiently close to the neighbourhood of the global minimum (Virieux and Operto, 2009).

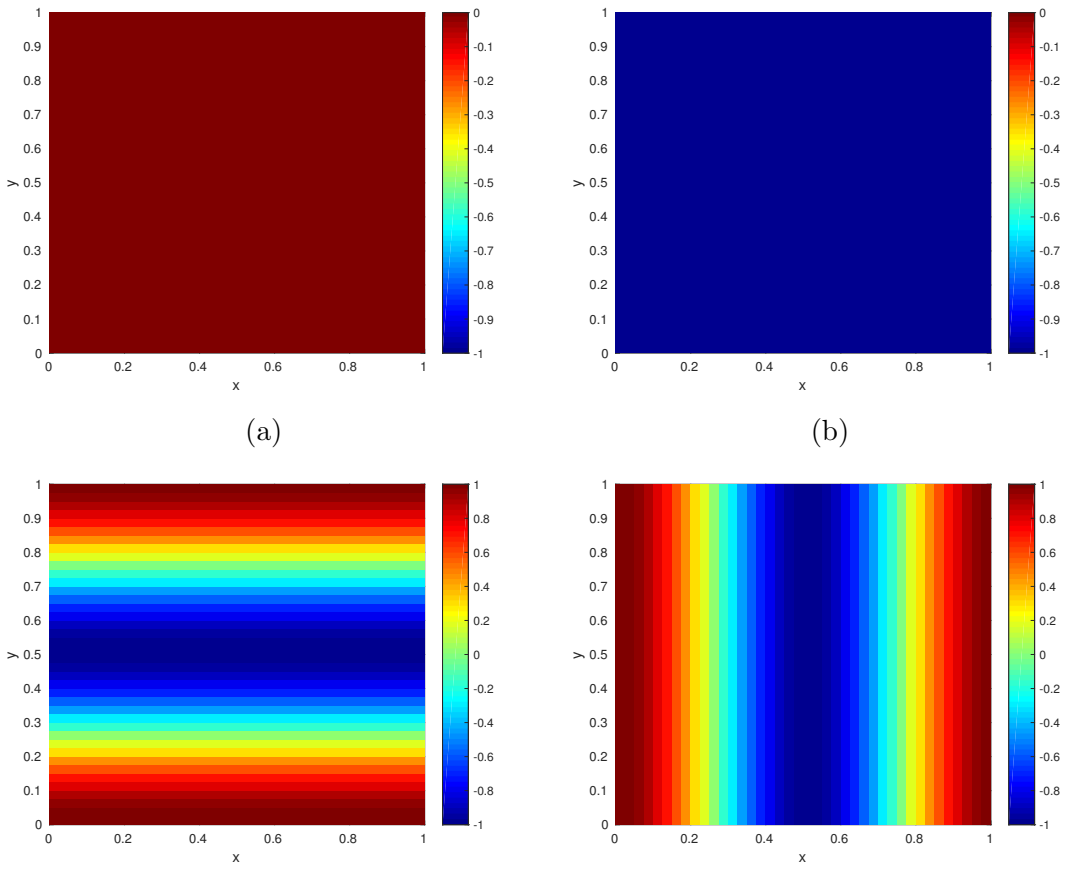


Figure 6.4: Different initial iterates for the problem of seeking a probability maximizer.

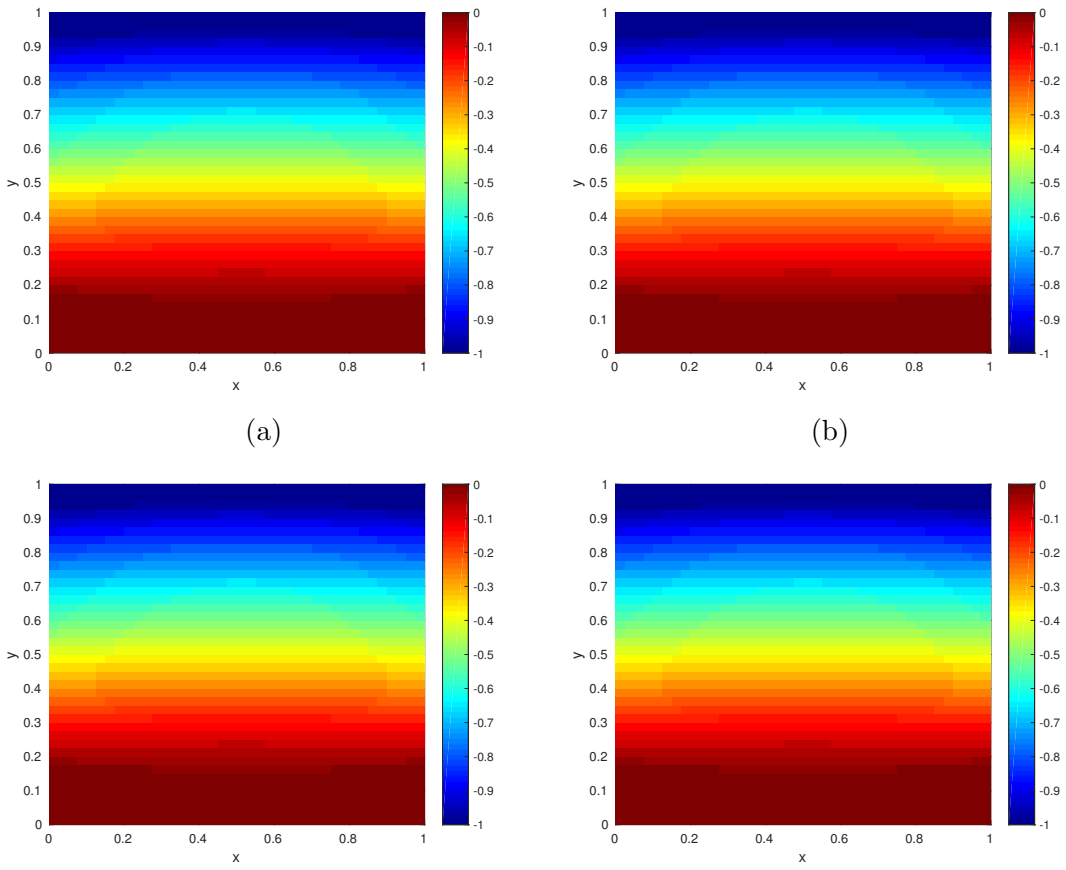


Figure 6.5: Solutions to the minimization problem for the full waveform inversion with $\alpha = 0.01$ and $\beta = 0.1$ with the four different starting iterates of Figure 6.4.

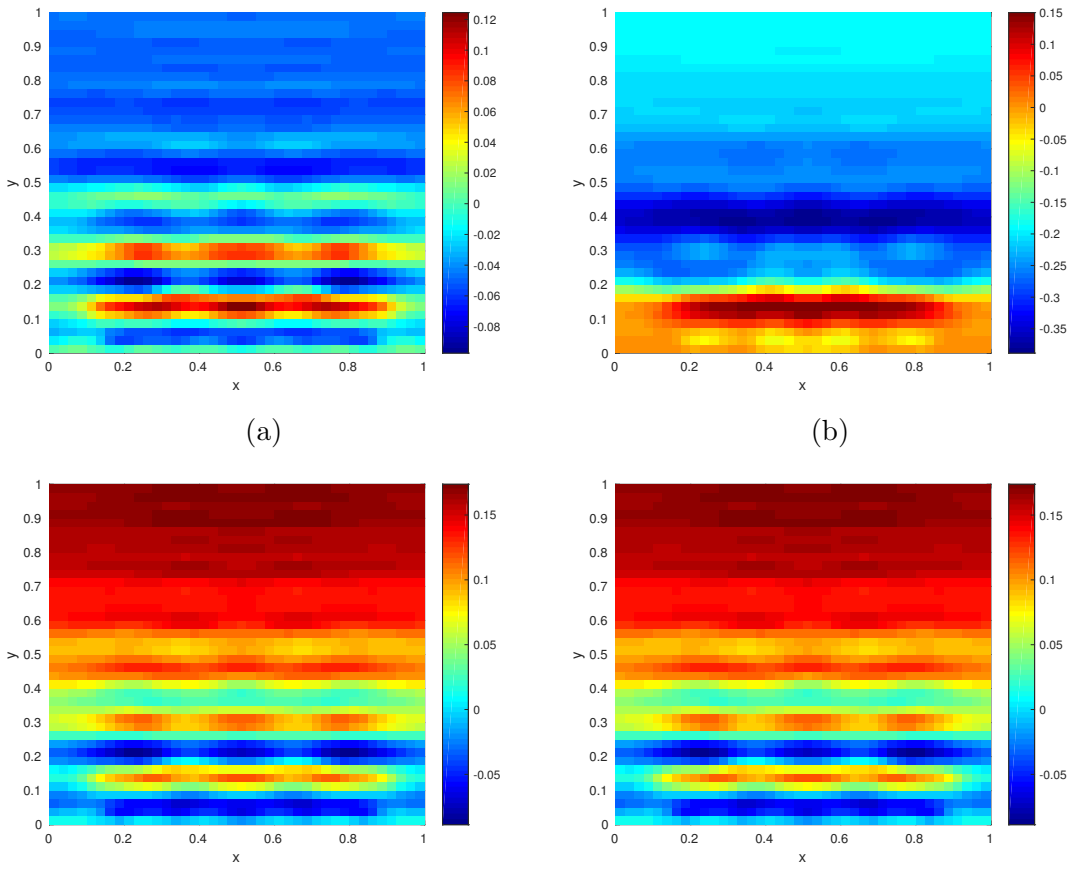


Figure 6.6: Solutions to the minimization problem for the full waveform inversion with $\alpha = 0.01$ and $\beta = 6$ with the four different starting iterates of Figure 6.4.

6.4 Multiple frequency inversion

Typically, the observations we have are available in several frequencies. There are two ways to perform inversion on this data. The first is a sequential method, where data at each frequency is processed incrementally, and the inversion is done sequentially, where the previous solution of the optimization at a lower frequency is used as the starting point in the next optimization. This is the common practice in the seismic industry today.

Alternatively, one could use a simultaneous method, which takes in the observations at all frequencies at once. We will look at both approaches.

6.4.1 Sequential method

In the sequential approach, the inversion is done frequency-by-frequency. Within the seismic industry, there is no known optimal way of choosing the frequencies. In general, the workflow for a sequential method for finding a probability maximizer is as follows:

1. Choose frequencies $\omega_1, \omega_2, \cdots, \omega_W$.

2. Solve for
$$\breve{m}_\omega^1 = \arg\min_{\phi_\omega, m} \mathcal{I}_{\omega_1}(m). \tag{6.30}$$

3. For $j = 2, \cdots, W$, solve
$$\breve{m}_\omega^j = \arg\min_{\phi_\omega, m} \mathcal{I}_{\omega_j}(m), \tag{6.31}$$
using $\breve{m}_\omega^{j-1}$ as the initial iterate.

The statistical approach to quantify uncertainty requires a small modification to the workflow, and is similar to the Kalman filter (see for instance, Chapter 4 of Kaipio and Somersalo, 2006):

1. Choose frequencies $\omega_1, \omega_2, \cdots, \omega_W$, and set $j = 1$.

2. Draw R realizations from μ_0, denote them by $m^1, m^2, \cdots, m^R$.

3. For each m^r, solve the optimization
$$\breve{m}^r = \arg\min \mathcal{I}_\omega^r(m), \tag{6.32}$$
where $\mathcal{I}_\omega^r$ is
$$\mathcal{I}_\omega^r(m) := \frac{p}{2} \left\| \psi_\omega - \mathcal{O}\breve{\phi}_\omega \right\|_2^2 + \frac{1}{2} \left\| \mathcal{A}(m)_\omega \breve{\phi}_\omega - q_\omega \right\|_{\Gamma_f}^2 + \frac{1}{2} \left\| m - m^r \right\|_{\mathcal{C}_0}^2. \tag{6.33}$$

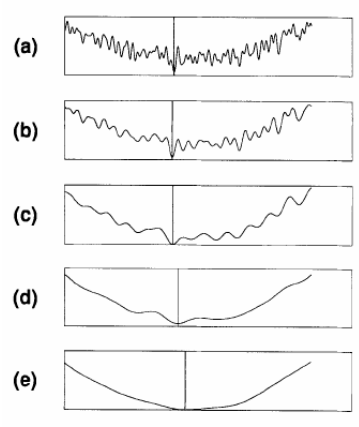


Figure 6.7: Strategy for finding a global minimum sequentially. We begin by minimizing the misfit functional for the lower frequency (e), and then gradually move up the misfit functional for a higher frequency (a). Reproduced from Bunks *et al.* (1995)

4. Run MCMC on each $\breve{m}^r_{\omega_j}$ to get $m^r_{\omega_j}$.

5. Use the new ensemble $m^r_{\omega_j}$ as the prior ensemble for next frequency ω_{j+1}.

It is argued that by starting from a low frequency and stepping up incrementally, we will gradually find the correct neighborhood of the global minimum. An illustration of this idea is given in Bunks *et al.* (1995), reproduced in Figure 6.7.

6.4.2 Simultaneous method

The simultaneous method includes observations at all frequencies at once. The misfit functional then becomes

$$\mathcal{I}(m) = \sum_{\omega\in\Omega}\left(\frac{p}{2}\left\|\psi_\omega - \mathcal{O}\breve{\phi}_\omega\right\|_2^2 + \frac{1}{2}\left\|\mathcal{A}_\omega(m)\breve{\phi}_\omega - q_\omega\right\|_{\Gamma_f}^2\right) + \frac{1}{2}\left\|m - m_0\right\|_{\mathcal{C}_0}^2. \quad (6.34)$$

As can be seen in the summation of the terms in 6.34, the misfit functional includes all the observations at once. Therefore, within the context of a probability maximizer, the solution is one which matches the high as well as the low frequencies.

Note that here, the forward precision is assumed to be the same across all frequencies. However, as we have seen, the high frequency component is highly oscillatory, and therefore, it is reasonable to prescribe a lower precision for the higher frequencies than for the lower frequencies.

6.4.3 Comparison between the two methods

In the sequential method, each step requires us to store the variables and observations at only one frequency. In contrast, the simultaneous method requires us to do

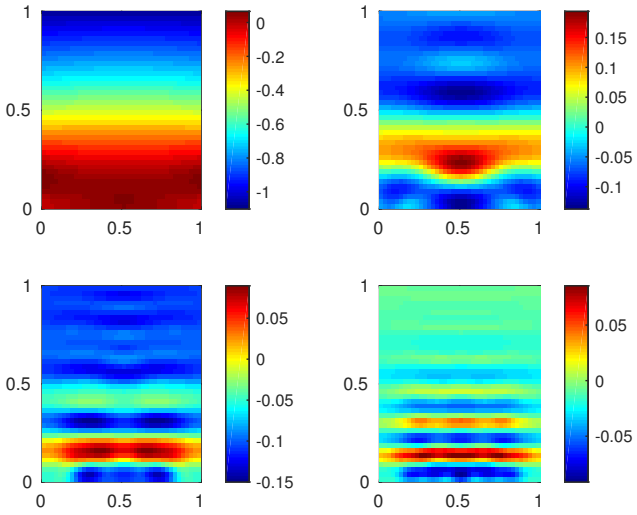


Figure 6.8: A probability maximizer for the full waveform inversion using the sequential method for $\alpha = 0.01$ and $\beta = 0.1$, 2, 4, and 6.

this for every frequency. Therefore, the sequential method would seem to require less memory to implement.

However, in evaluating a misfit functional for the simultaneous method, one could solve for the variables at each frequency, and then assemble the likelihood, gradient and Hessian of the misfit, and then step forward to the next frequency, discarding the variables at the previous frequency, since it is no longer needed. The same procedure can be applied for the observations as well. This makes the memory requirements of the simultaneous method to be the same as the sequential method, so that the only distinction is that each function evaluation, gradient and Hessian computation are slower in the simultaneous method than in the sequential method. Nevertheless, this is offset by the fact that only one optimization problem has to be solved.

Furthermore, since the simultaneous method attempts to fit the observations at all frequencies at once, it should produce minimizers which are closer to the truth. On the other hand, the sequential method merely tries to fit the observation at the highest frequency, using solutions from the previous frequency as starting iterates. The following example illustrates this fact.

We solve the optimization problem on the misfit functionals, where the same truth as in Figure 6.3 is used to generate the observations. We fix $\alpha = 1/100$ and use $\beta = 0.1$, 2, 4, and 6 as the frequencies. The sequential and simultaneous methods were applied to this problem and the results are shown in Figures 6.8 and 6.9.

Note that the simultaneous method produces a solution which is closer to the

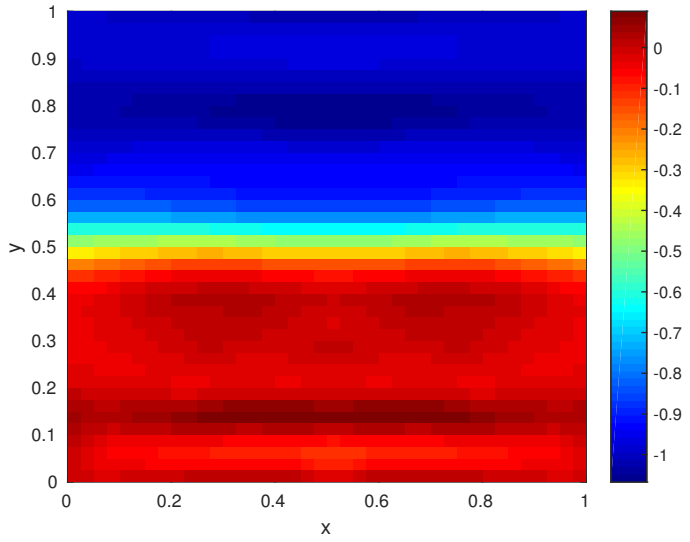


Figure 6.9: A probability maximizer for the full waveform inversion using the simultaneous method for $\alpha = 0.01$ and $\beta = 0.1$, 2, 4, and 6.

truth than the sequential method. This is because the simultaneous method has attempted to fit the observations at all frequencies, whereas in the sequential method, the solution is one which fitted the observation at the highest frequency, but the fit at a lower frequency has now been lost.

For the rest of the chapter, we will employ the simultaneous method to solve the problem of full waveform inversion. The next section will explore various forms of this inverse problem in order to evaluate our methods for solving these problems.

6.5 Applications

We will work within the perfect model scenario to evaluate the effectiveness of this method. In other words, the same grid and numerical scheme is used to generate the observations as well as to perform the inversion.

Figure 6.10 shows the truth used for this experiment. The prior specified is the biharmonic measure with $\sigma = 1/2$, $\lambda_x = 3/4$ and $\lambda_y = 1/4$. The number of cells in each direction is $K = 40$, and the observation standard deviation is 1% of $\|\phi_\omega\|_\infty$. Here, this gives a standard deviation of the order 10^{-3}. For the forward precision, we use the precision as described in (4.41) with $\tilde{\lambda} = 1/10$. The inversion is done for $\alpha = 0.01$ and $\beta = 0.1$, 2, 4, and 6. We choose a stepsize of 0.005 for the Markov chain Monte Carlo sampling, and an ensemble size of $R = 100$. The stepsize is chosen such that it yields an acceptance rate of around 0.25. All optimization problems were solved using Matlab's trust-region method with a quadratic model function,

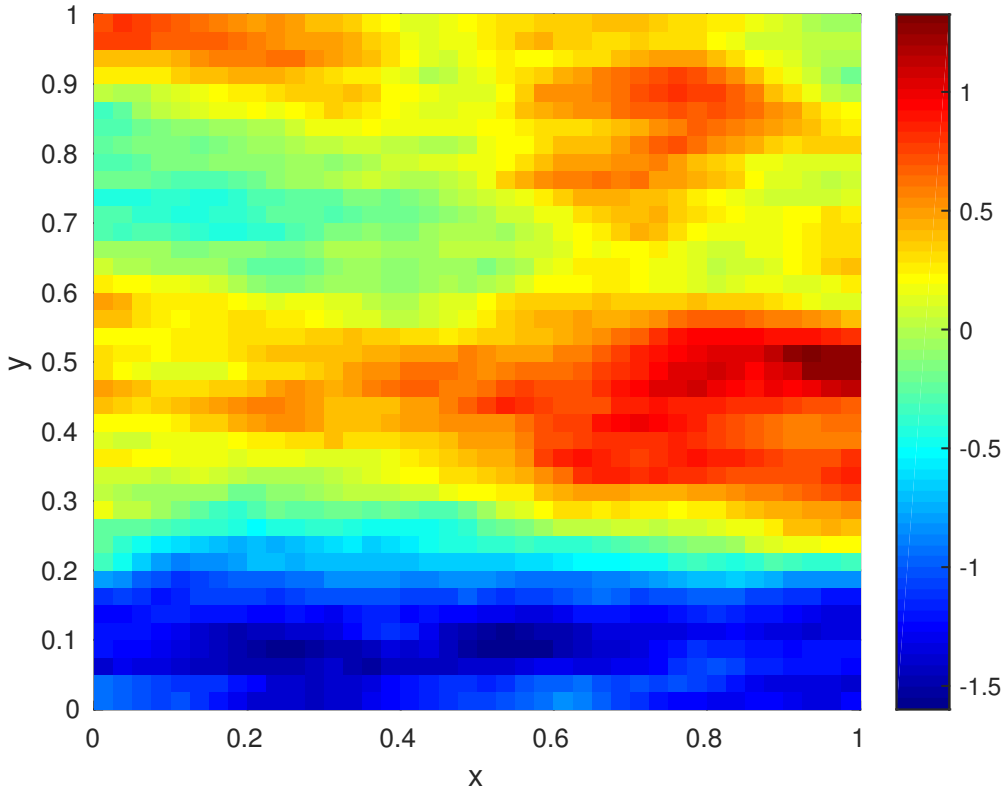


Figure 6.10: Truth used in the numerical simulations. The depth $y = 0$ is the shallow layer while $y = 1$ indicates a deeper layer.

and tolerance of 10^{-6}.

In order to construct a prior mean for this problem, we assume that an exploratory well has been drilled at $x = 0.5$. We then use the values obtained along this line to extrapolate to the entire domain.

6.5.1 Noise-free observations

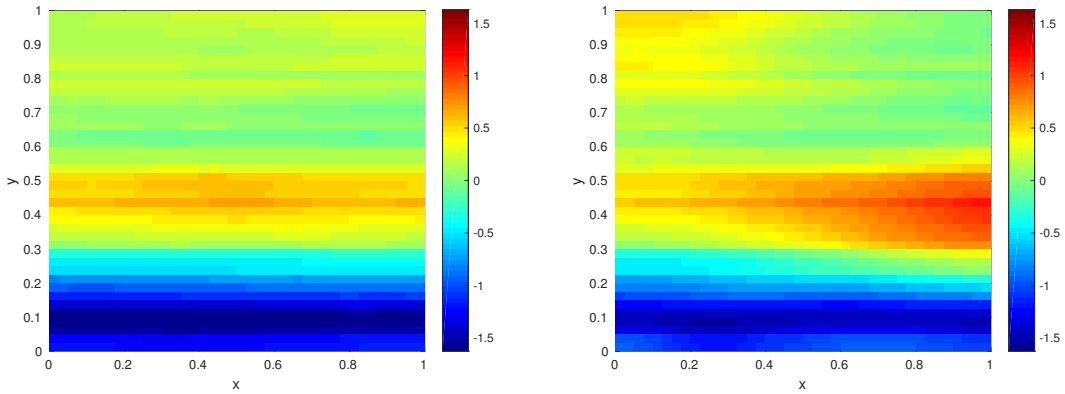
Figures 6.11(a)–(c) show the means for the prior ensemble, RML ensemble and the MCMC ensemble respectively. The RML and MCMC ensemble means are almost identical, which suggests that the RML ensemble was already a good sample from the posterior measure. Note that the MCMC ensemble mean differs slightly from the probability maximizer (Figure 6.11(d)). The probability maximizer contains more details about the subsurface. Nevertheless, a clearer picture of the posterior measure can be seen from the ensemble members.

Figures 6.12(a)–(c) show the plots for 20 ensemble members of the prior, RML, and MCMC ensemble respectively. Note that here, we see that the RML and MCMC ensemble members agree amongst themselves on the first layer, but they contain different information about the deeper parts of the subsurface. For instance, MCMC ensemble member on row 2, column 4, suggests that the wave speed in the deeper parts of the subsurface is higher than the wave speed in that same part of the MCMC ensemble member on row 2, column 3.

In general, we see that there is greater agreement in the shallower parts of the subsurface than in the deeper parts. This can be more readily seen in the plots of the ensemble standard deviation in Figures 6.13(a)–(c). Here, we see that the standard deviation is higher in the deeper parts of the subsurface than in the shallower parts. Note also that the standard deviation is lower in the middle parts of the subsurface. This is due to the fact that the prior mean was constructed using information from the middle of the domain.

Figure 6.14 shows the observation errors of the RML and the MCMC ensemble. Note that while the error in the MCMC ensemble is higher than the RML ensemble in the lower frequencies, at 0.1 Hz, the difference is around 10^{-3}, while the difference is around 2×10^{-3} at 2 Hz. At the higher frequencies, the errors are significantly smaller. This shows that the RML ensemble produced honours the observations at the higher frequencies, and is a further indication that the RML ensemble was a sufficiently good ensemble to work with practically to characterize the posterior measure.

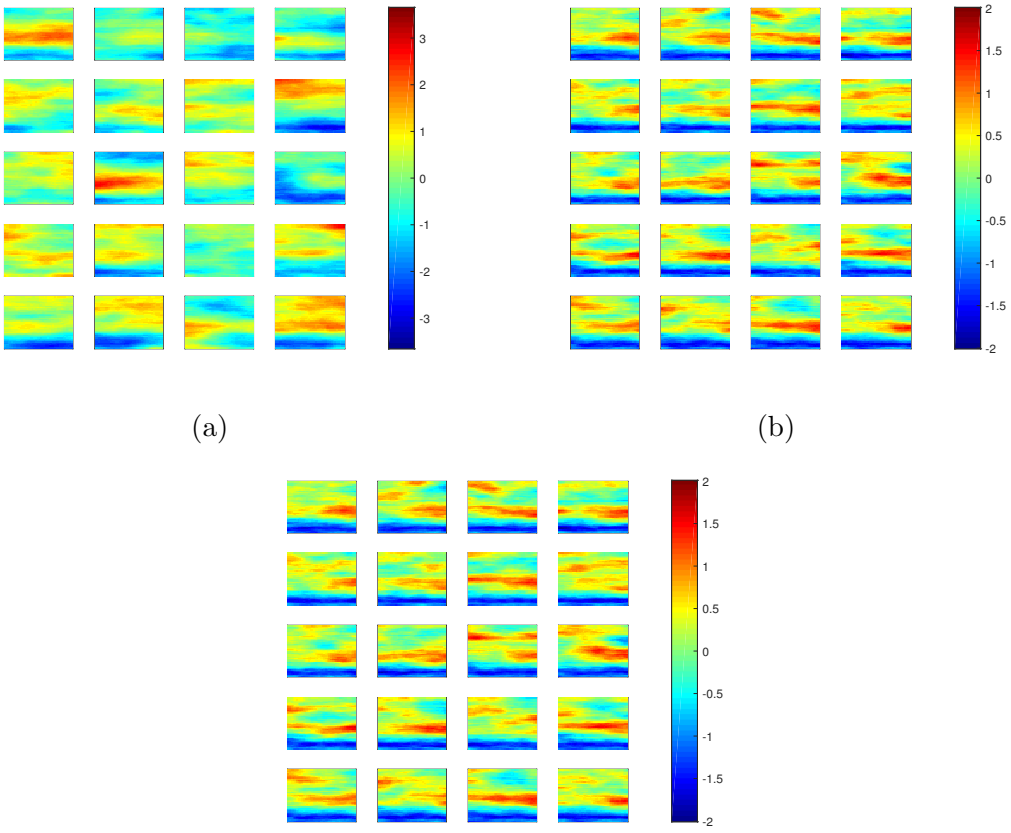
Note that one might consider using a smaller or larger ensemble size in these experiments. In order to address this question, we plot the mean and standard



(a) (b)

(c) (d)

Figure 6.11: Ensemble means for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble; as well as (d) the probability maximizer for the problem of full waveform inversion using noiseless observation for the problem discussed in Section 6.5.



(a) (b)

(c)

Figure 6.12: The (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion using noiseless observation for the problem discussed in Section 6.5.

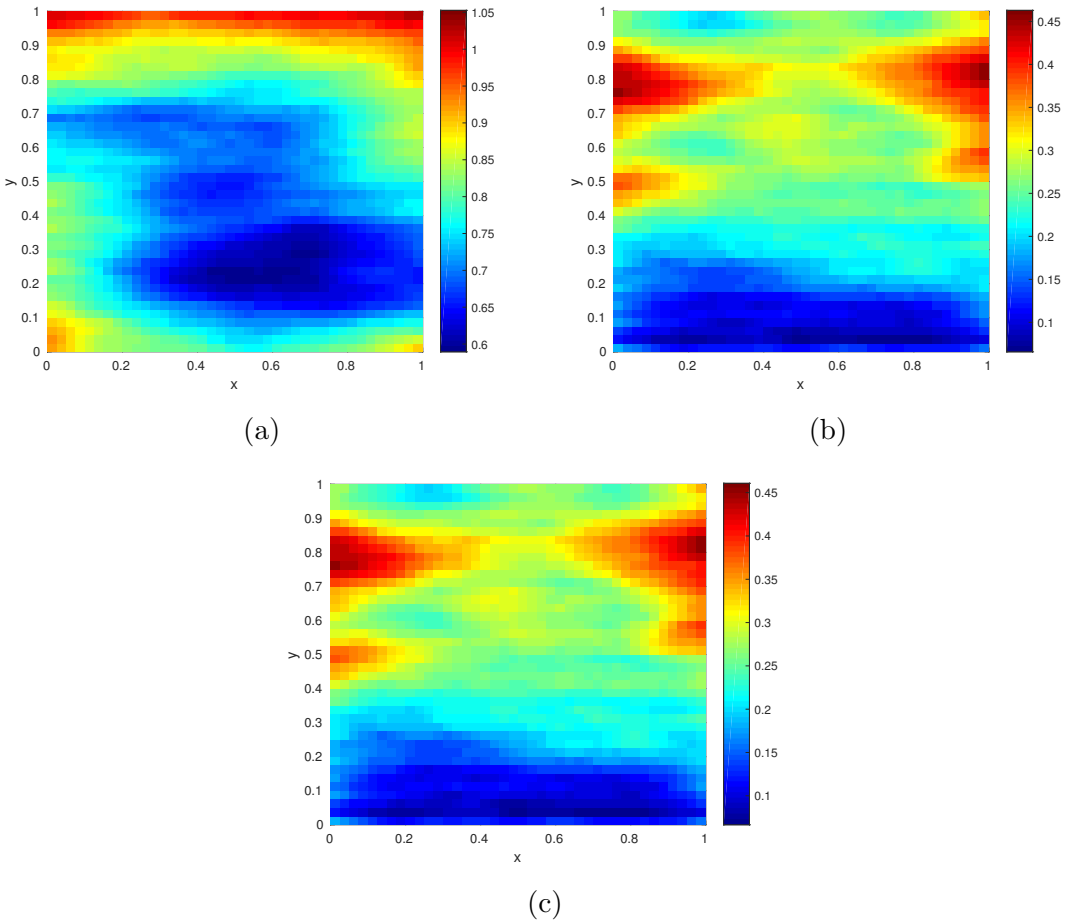


Figure 6.13: Ensemble standard deviations for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion using noiseless observation for the problem discussed in Section 6.5.

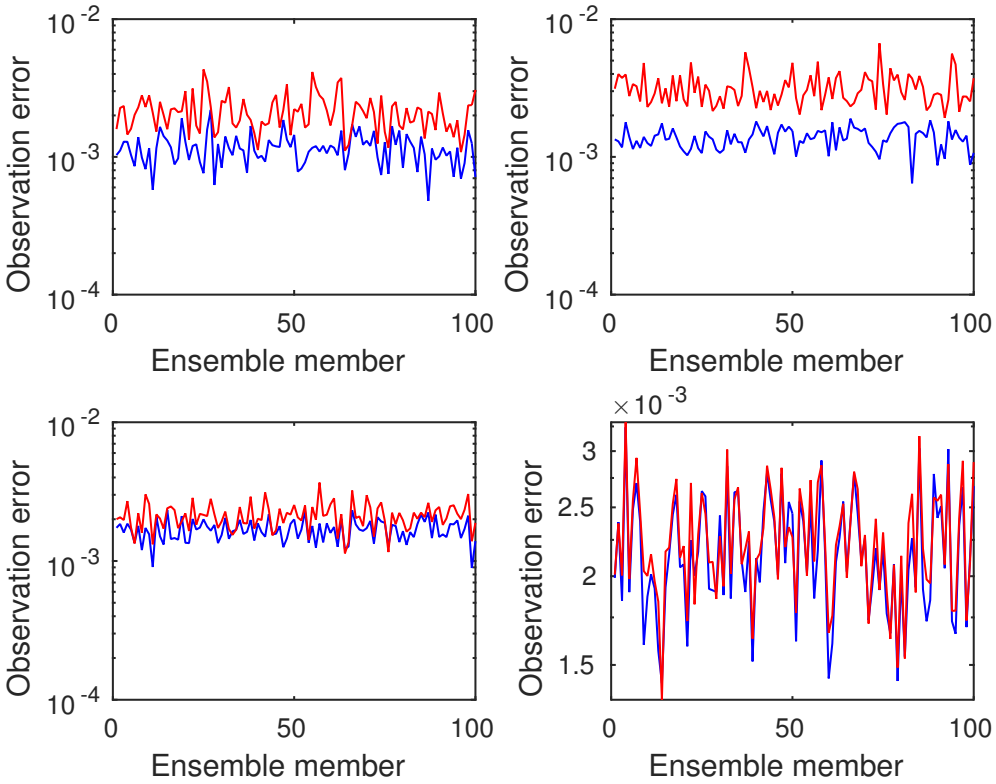


Figure 6.14: Observation errors for the problem of full waveform inversion with noiseless observation for frequencies (a) 0.1 Hz, (b) 2 Hz, (c) 4 Hz, and (d) 6 Hz. The red line shows $\|\mathcal{O}(u(m^r_{MCMC})) - s\|_\infty$ and the blue line shows $\|\mathcal{O}(u(m^r_{RML})) - s\|_\infty$, where m^r_{MCMC} denotes the r-th realization of the RML-MCMC ensemble, and m^r_{RML} denotes the r-th realization of the RML ensemble, where $r = 1, 2, \cdots, 100$.

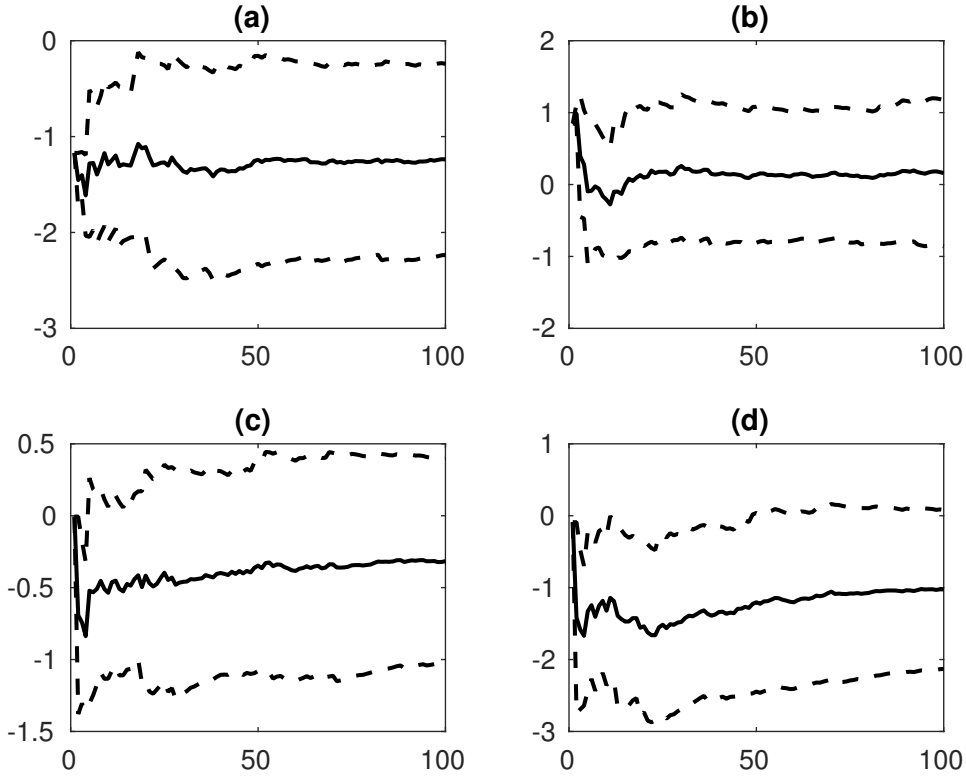


Figure 6.15: Plots of prior ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noiseless observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

deviation of the ensemble at various points. Figures 6.15, and 6.16 and 6.17 show the ensemble means and their departure by one ensemble standard deviation for the prior, RML and MCMC ensemble respectively for different entries of m, namely m_{43} (near surface point), m_{82} (surface point), m_{1200} (middle of the subsurface), m_{1600} (near the bottom of the subsurface). We see that as the ensemble size increases, the means and standard deviations eventually stabilizes. In fact, for this problem, the means and standard deviations begin to stabilize rather early, indicating that a smaller ensemble, say of size 50, could have been used.

6.5.2 Noisy observations

Figures 6.18(a)–(c) shows the means of the prior ensemble, RML ensemble, and the MCMC ensemble when 1% noise has been added into the observations. Note that the MCMC ensemble mean is smoother compared to the probability maximizer

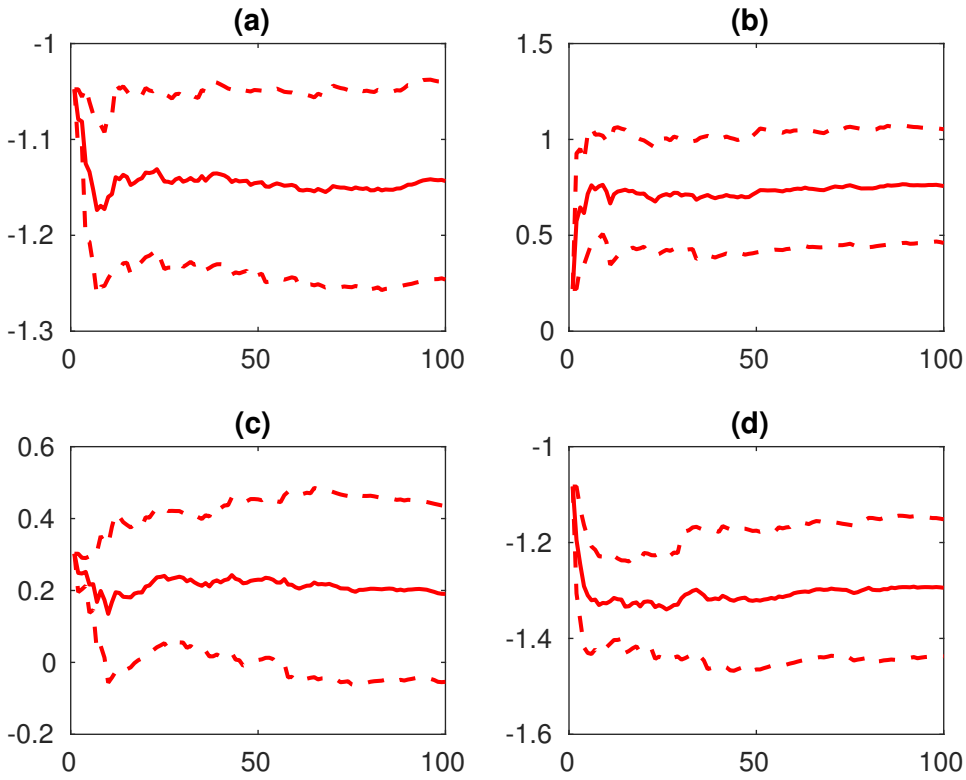


Figure 6.16: Plots of RML ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noiseless observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

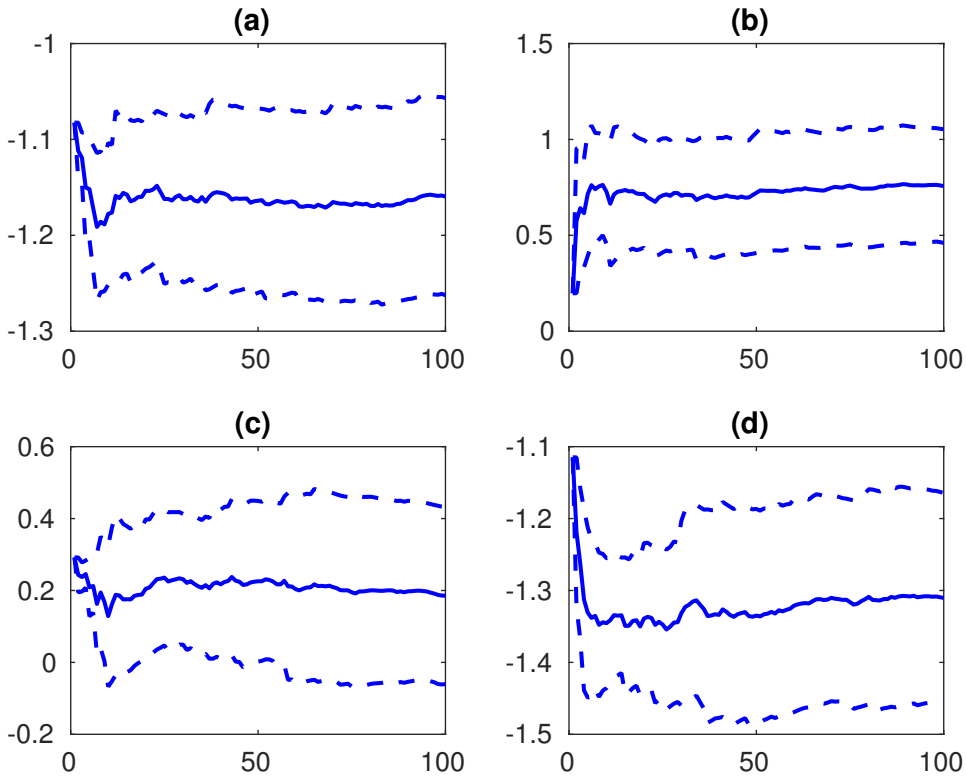


Figure 6.17: Plots of MCMC ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noiseless observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

(Figure 6.18(d)). Unlike the ensemble means, the probability maximizer shows some difficulty in detecting the first layer from $y = 0$ to around $y = 0.2$, due to the noise in the observation. This can be seen from the variations in the wavespeed from around $x = 0.1$ to around $x = 0.4$.

Figures 6.19(a)–(c) show the plots for 20 ensemble members of the prior, RML, and MCMC ensemble respectively. Here, we see that most of the RML and MCMC ensemble members detect the first layer, although the wave speeds in the next layer differ. For instance, the RML and MCMC ensemble member on row 2, column 2 shows a lower wave speed of around 0.5 in the next layer than most other ensemble members suggests that wave speeds could be higher in certain parts. We also see that once again, the RML and MCMC ensembles contain differing information about the deeper parts of the subsurface. For instance, the RML and MCMC ensemble member on row 1, column 2, suggests that wave speeds are higher in the left-hand side of the deeper parts of the subsurface, which is not visible in most other RML and MCMC ensemble members.

As is the case in the previous subsection, in general, we see that there is greater agreement between ensemble members in the shallower parts of the subsurface than in the deeper parts. This can be more readily seen in the plots of the ensemble standard deviation in Figures 6.20(a)–(c). Here, we see that the standard deviation is higher in the deeper parts of the subsurface than in the shallower parts. Note also that the standard deviation is lower in the middle parts of the subsurface. This is due to the fact that the prior mean was constructed using information from the middle of domain. Note also, that the standard deviation here is slightly higher than the standard deviations in the previous case where noise was not added to the observations.

Figure 6.21 show the observation errors of the RML and the MCMC ensemble. Note that while the error in the MCMC ensemble is higher than the RML ensemble in the lower frequencies, the difference is rather small, while the errors do not differ by much in the higher frequencies. This is a further indication that the RML ensemble was a sufficiently good ensemble to work with practically to characterize the posterior measure, even in the case where noise was added to the observations.

Comparing the RML and MCMC means in both cases, we see that they have roughly the same means (see Figures 6.11(b)–(c) compared to Figures 6.18(b)–(c)). However, through the ensembles plotted, we see that it is easier to infer the first layer in the noiseless case than it is in the noisy case, as expected.

Note also that more iterations of the optimization were required in the noisy case to find the probability maximizer than in the noiseless case.

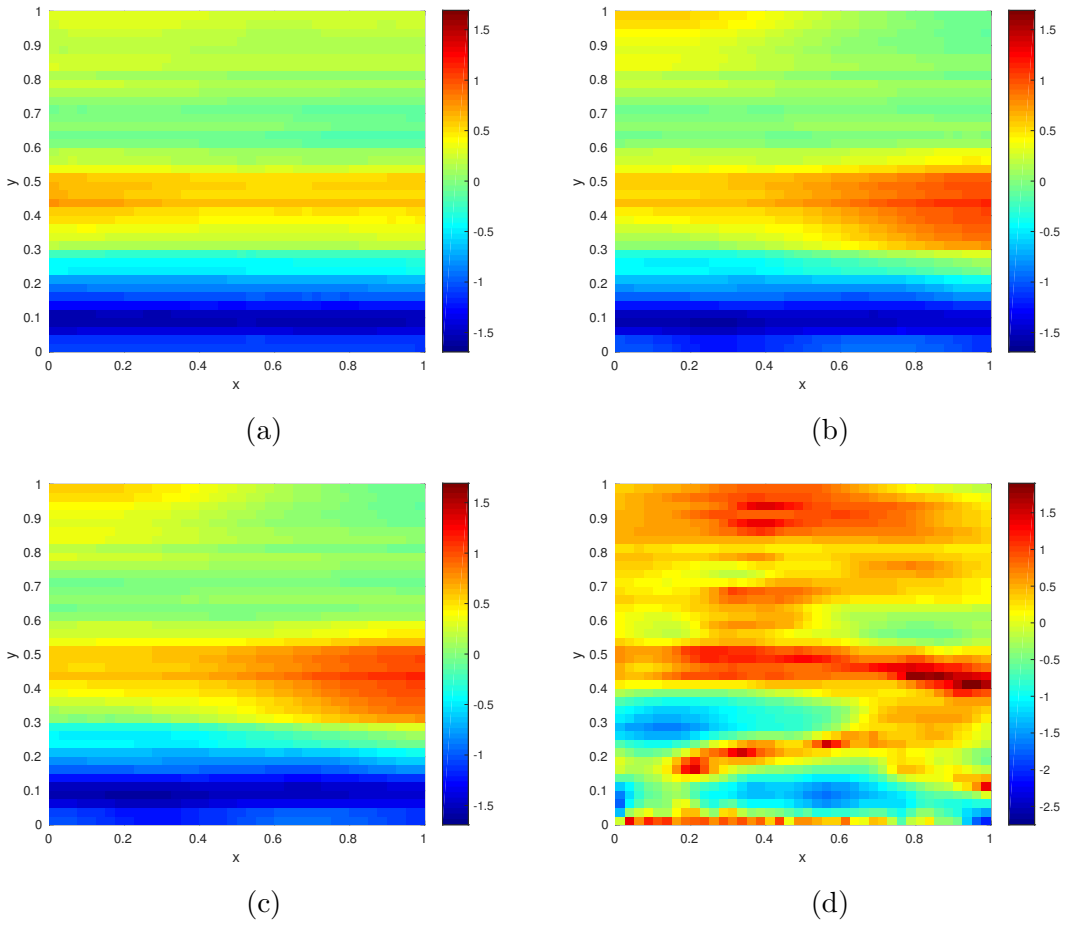


Figure 6.18: Ensemble means for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble; as well as (d) the probability maximizer for the problem of full waveform inversion using noisy observation for the problem discussed in Section 6.5.

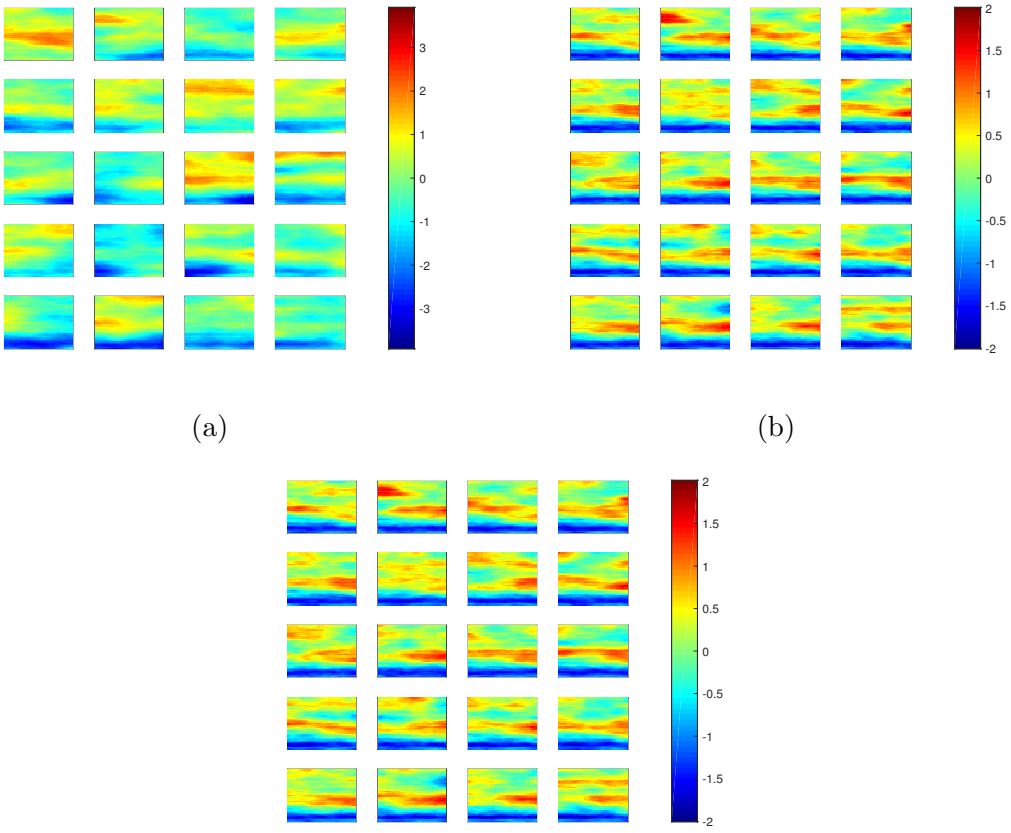


Figure 6.19: The (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion using noisy observation for the problem discussed in Section 6.5.

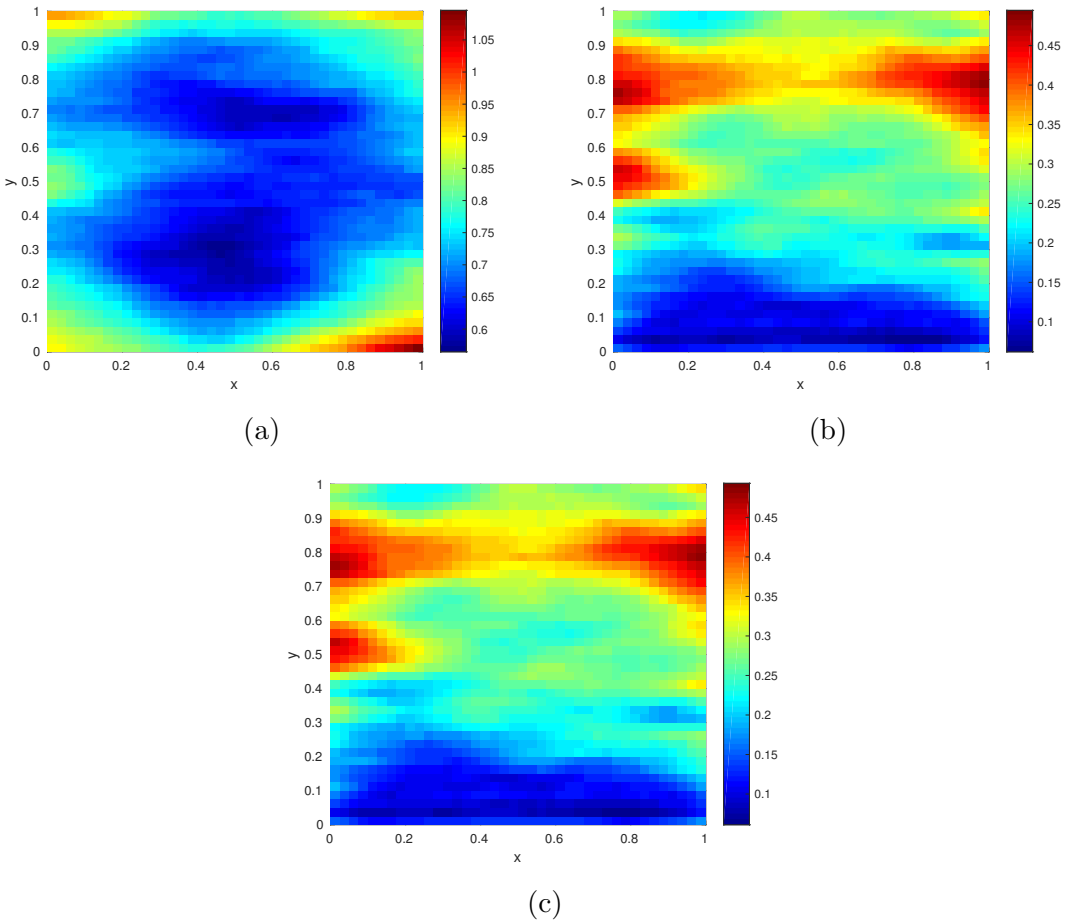


Figure 6.20: Ensemble standard deviations for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion using noisy observation for the problem discussed in Section 6.5.

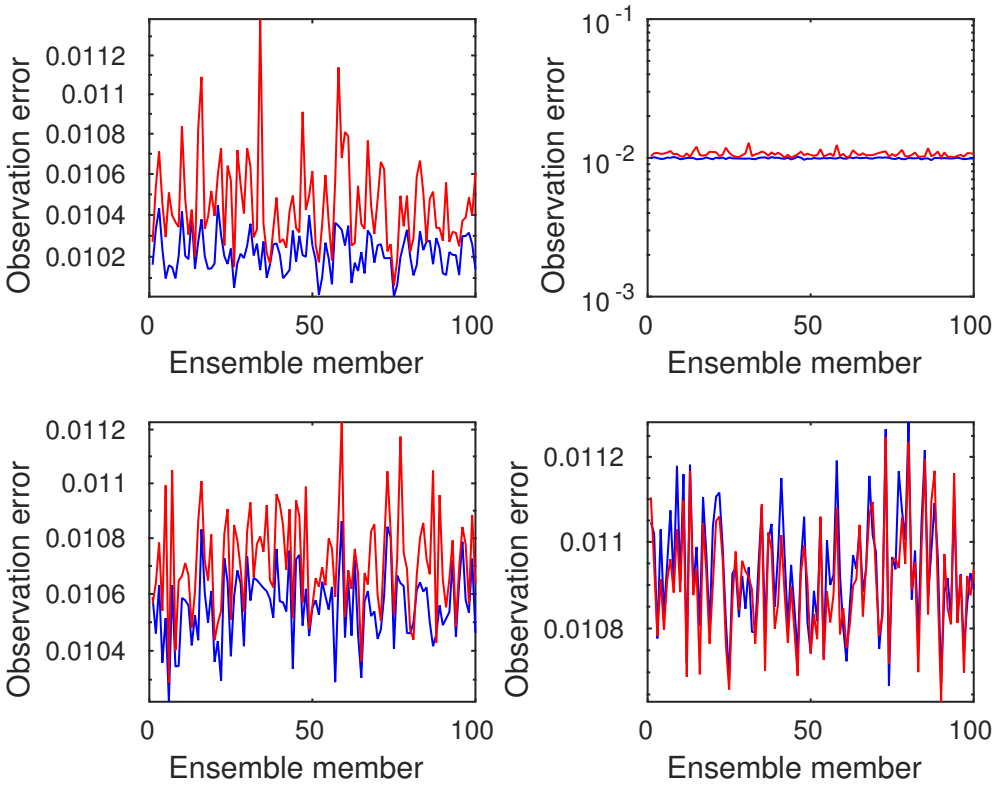


Figure 6.21: Observation errors for the problem of full waveform inversion with noisy observation for frequencies (a) 0.1 Hz, (b) 2 Hz, (c) 4 Hz, and (d) 6 Hz. The red line shows $\|\mathcal{O}(u(m^r_{MCMC})) - s\|_\infty$ and the blue line shows $\|\mathcal{O}(u(m^r_{RML})) - s\|_\infty$, where m^r_{MCMC} denotes the r-th realization of the RML-MCMC ensemble, and m^r_{RML} denotes the r-th realization of the RML ensemble, where $r = 1, 2, \cdots, 100$.

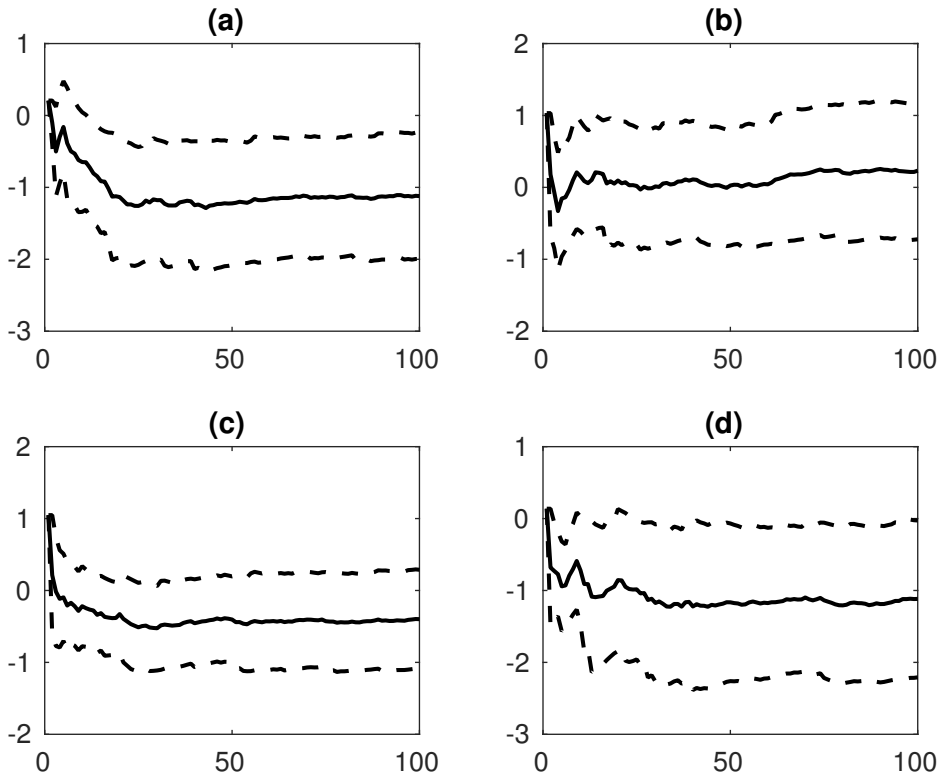


Figure 6.22: Plots of prior ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noisy observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

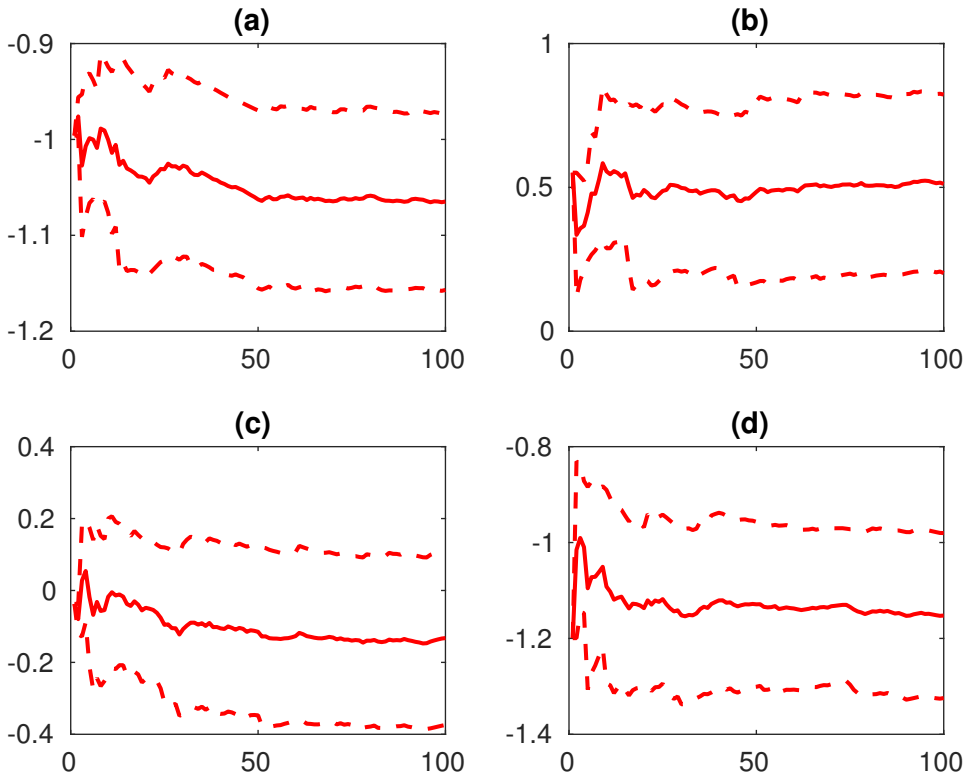


Figure 6.23: Plots of RML ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noisy observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

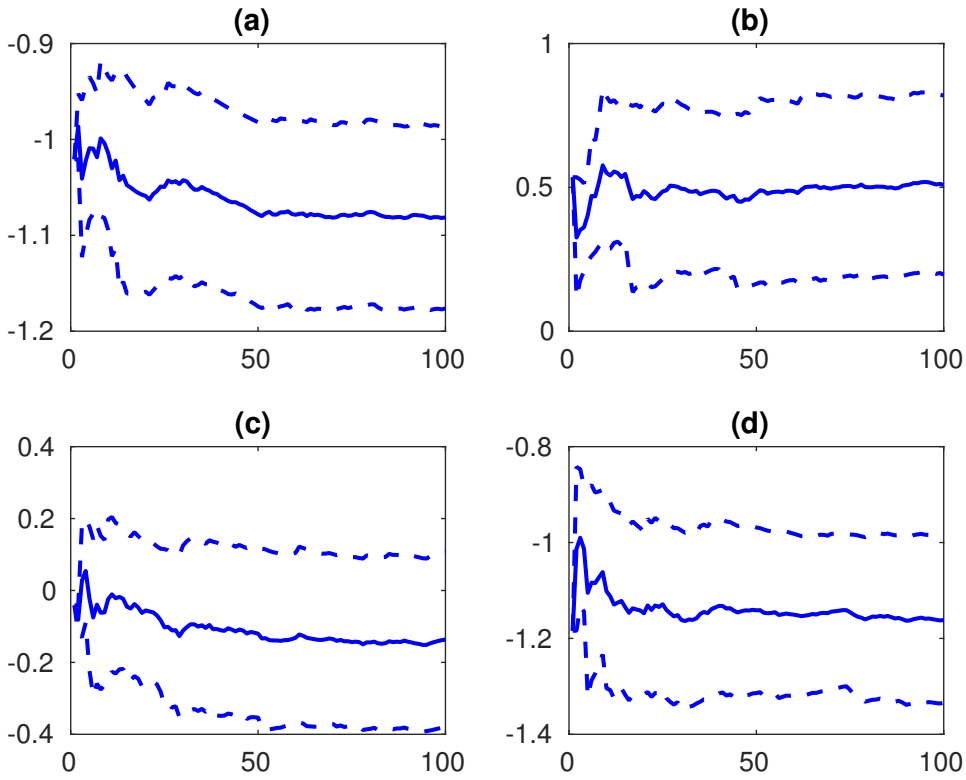


Figure 6.24: Plots of MCMC ensemble mean and standard deviation for (a) m_{43}, (b) m_{82}, (c) m_{1200}, and (d) m_{1600} for the problem of full waveform inversion with noisy observation. The abscissa is the number of ensemble members and the ordinate is the value of m at a particular point. The solid lines in the middle of the plots indicate the ensemble mean, while the dotted lines show the departure of the ensemble mean by one ensemble standard deviation.

Finally, in Figures 6.22–6.24, we see that the ensemble means and standard deviations for the RML and MCMC ensemble requires a larger ensemble in order to stabilize compared to the means and standard deviations in the noiseless case. This suggests that more ensemble members are required in practical applications to ensure convergence if the noise levels are higher.

6.6 Discussion

There are many more experiments that can be performed to ascertain the robustness of the algorithm as described in Algorithm 1. We will demonstrate the robustness of this algorithm by performing further numerical experiments on the problem of full waveform inversion, as described in Section 6.5, with different settings on:

1. Noise levels.
2. Location of observation.
3. Choice of frequencies.

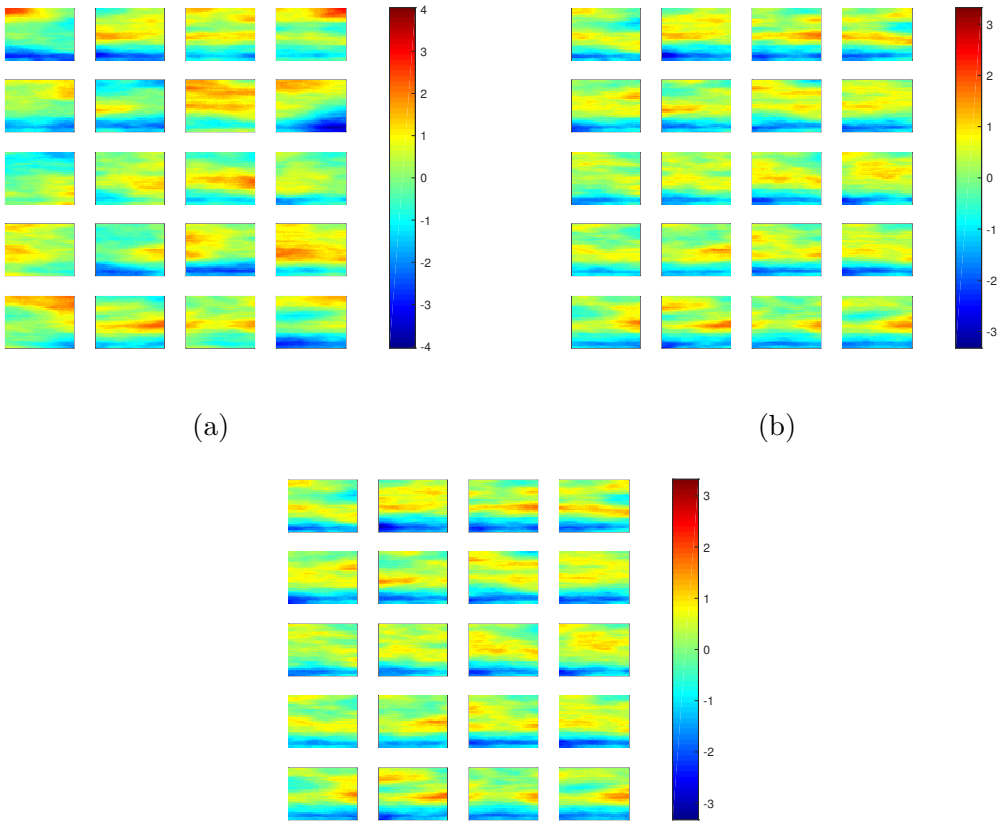
6.6.1 Noise levels

One such experiment is to consider different noise levels in the observations. In general, we would expect that higher noise levels in the observations will induce greater uncertainty. This will manifest itself through greater fluctuations in values in the RML and MCMC ensemble itself, which in turn gives a higher ensemble standard deviation for the RML and MCMC ensemble.

As an illustration, Figures 6.25(a)–(c) and Figures 6.26(a)–(c) show 20 plots of the prior, RML and MCMC ensemble as well as their respective ensemble standard deviations for the same problem as in Section 6.5 but with 10% noise instead. Note here that there are greater fluctuations in the RML and MCMC ensemble, even in detecting the first layer. This, in turn translates to a higher ensemble standard deviation. Here, the plots of the ensemble means are not shown as it does not play a relevant role in this discussion.

6.6.2 Location of observation

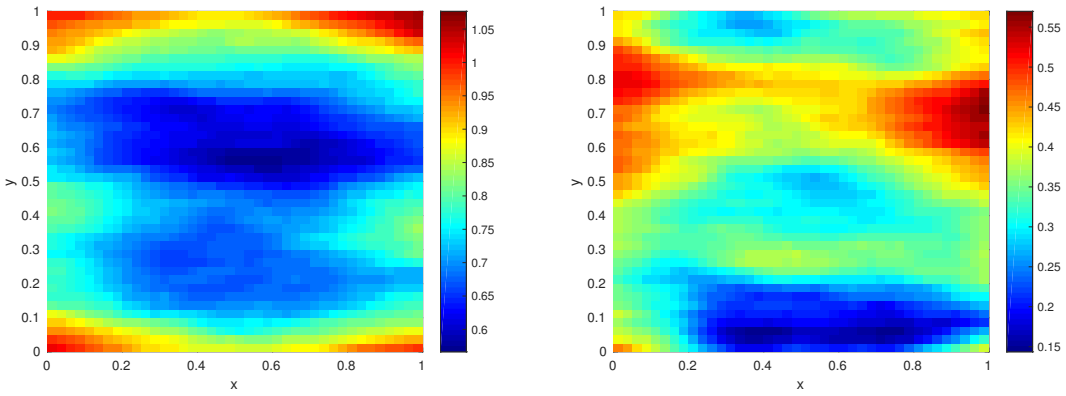
In this chapter, we set up the problem of full waveform inversion where the observations were made at every grid point on the surface. We can also consider a variation of the problem where there are fewer observations.



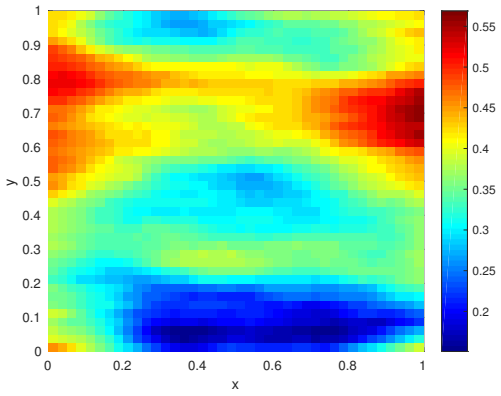
(a) (b)

(c)

Figure 6.25: The (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion with 10% noise in the observation.



(a) (b)



(c)

Figure 6.26: Ensemble standard deviations for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion with 10% noise in the observation.

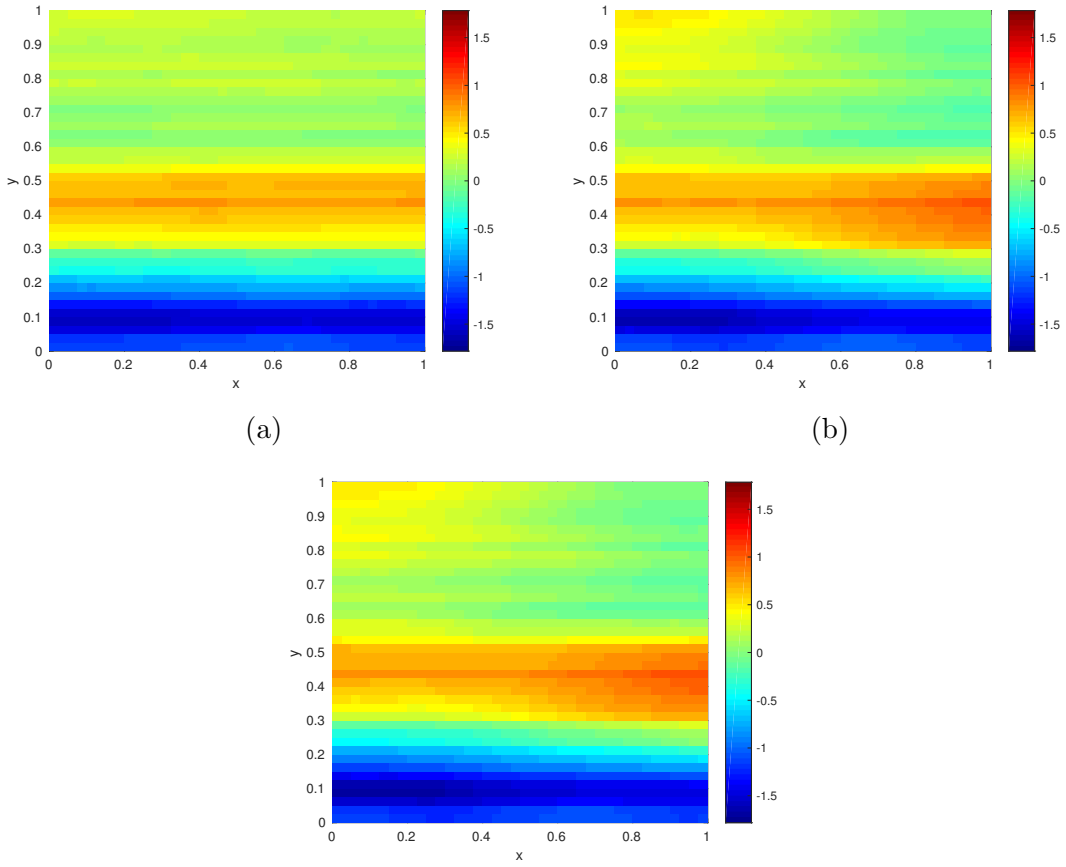


Figure 6.27: Ensemble means for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble for the problem of full waveform inversion, where observations were made at the points $(0, 0)$, $(0.5, 0)$, and $(1, 0)$, with 1% observation noise.

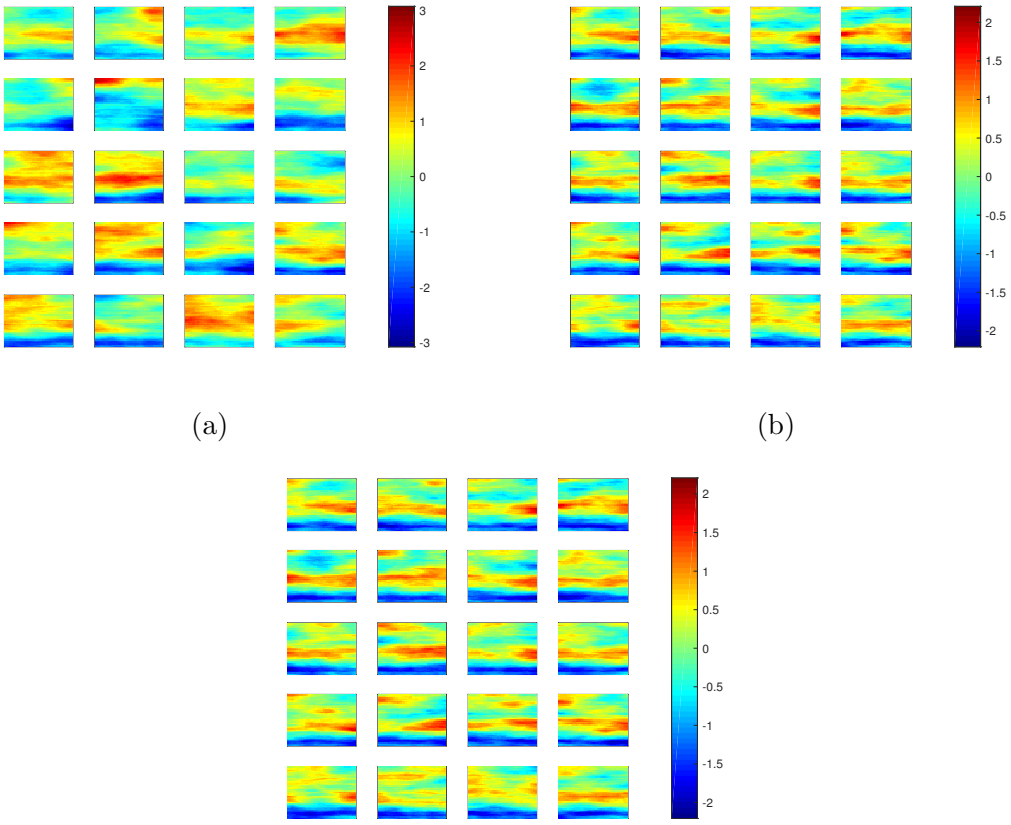
An extreme case would be to consider what happens if there are no observations at all. In this case, the first term in $\mathcal{I}(m)$ in (6.34) that comes from the likelihood vanishes. Then, from (6.28) we have

$$\breve{\phi}_\omega(m) = (\mathcal{A}_\omega(m)^T L_f \mathcal{A}_\omega(m))^{-1}(\mathcal{A}_\omega(m)^T L_f q_\omega) = \mathcal{A}_\omega(m)^{-1} q_\omega, \tag{6.35}$$

so that the second term in (6.34) also vanishes. Therefore, the misfit merely involves the term coming from the prior on m.

Therefore, we can conclude that as fewer observations are made, the results obtained will be closer to the prior.

This can be seen when we run the experiment for the same problem as defined in Section 6.5, with 1% noise in the observations. However, instead of observing at every grid point on the surface, we only observe the variables at the points $(0, 0)$, $(0.5, 0)$, and $(1, 0)$ on the surface.



(a) (b)

(c)

Figure 6.28: The (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion, where observations were made at the points $(0, 0)$, $(0.5, 0)$, and $(1, 0)$, with 1% observation noise.

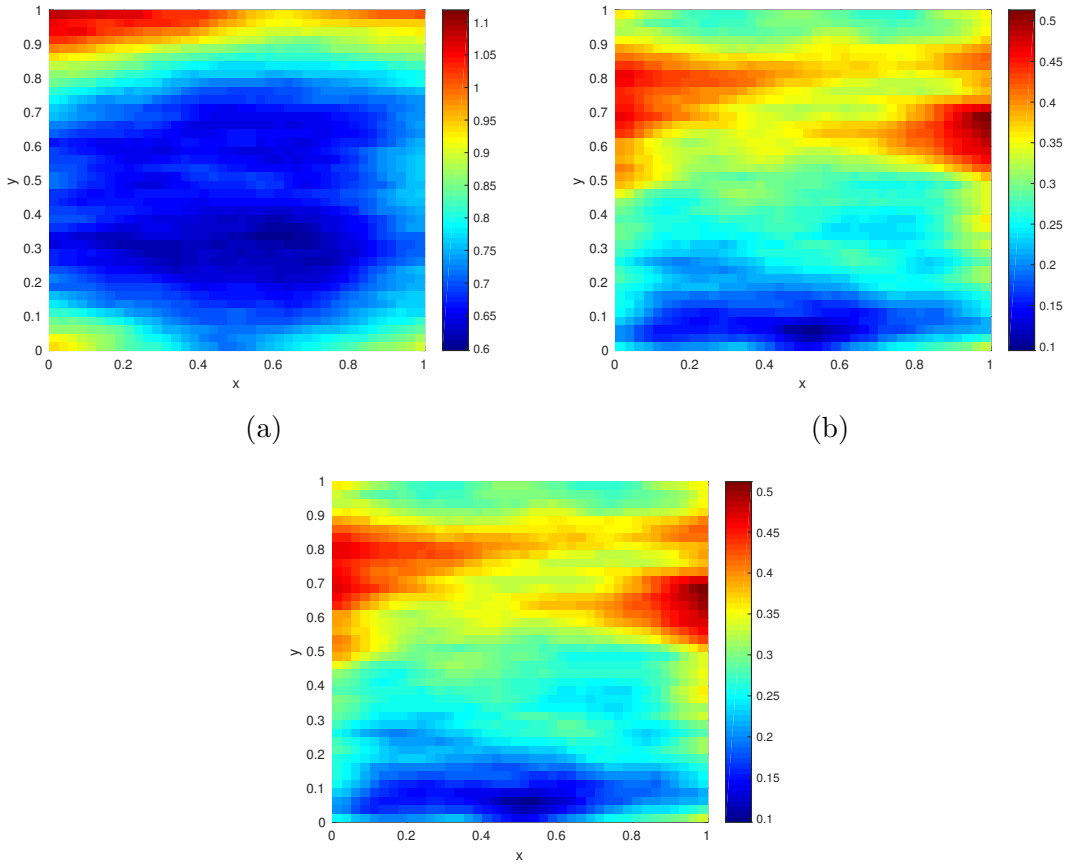


Figure 6.29: Ensemble standard deviations for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion, where observations were made at the points $(0, 0)$, $(0.5, 0)$, and $(1, 0)$, with 1% observation noise.

Figures 6.27(a)–(c) show the prior, RML, and MCMC ensemble means respectively. Here, the RML and MCMC ensemble is slightly detects the structure of the subsurface in the neighbourhood of the point $(0, 1)$. Otherwise, there is not much obvious change in the ensemble means.

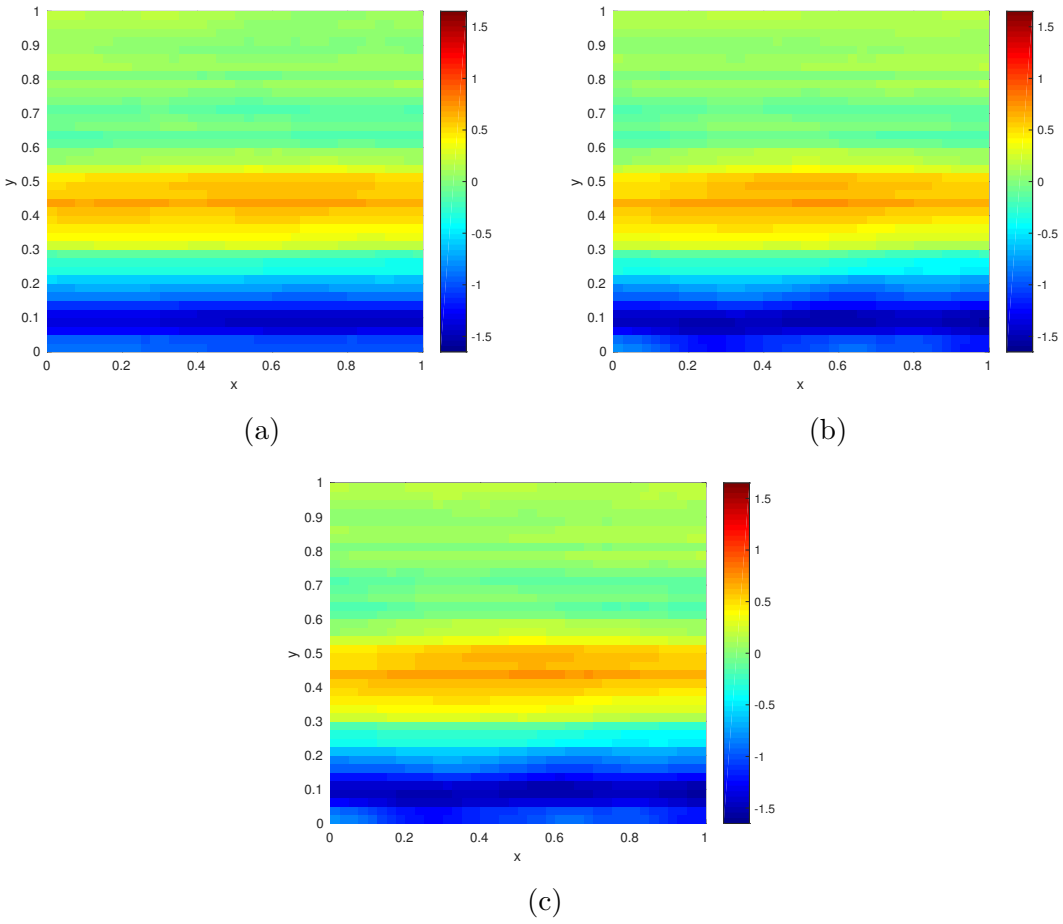
Figures 6.28(a)–(c) show the 20 plots of the prior, RML, and MCMC ensemble members. We see that the each ensemble members picks up the first layer near the subsurface, although there are greater fluctuations for this problem compared to the results obtained in Section 6.5 (see Figures 6.19(a)–(c)). This is made clearer by the plots of the standard deviation in Figures 6.29(a)–(c). Note here that the standard deviation at the point $(0.5, 0)$ is lower than other parts near the surface, as observations were made at that point. We also expect that the standard deviations at the points $(0, 0)$ and $(1, 0)$ will decrease significantly in the RML and MCMC ensemble as these were also observation points. This is indeed what is seen here as well.

6.6.3 Choice of frequencies

In the seismic industry, it is desirable to use observations from an optimal set of frequencies which gives the best results in terms of efficiency and information. However, there is no known theory on how this is to be done. Nevertheless, it is well-known that the low frequencies are crucial in imaging the subsurface of the Earth. However, the current technology available in collecting seismic data does not give low frequency data. This hinders the ability to image the subsurface such that the result is close to the truth, due to the nonconvexity of the misfit functional. Here, we demonstrate the robustness of the weak constraint formulation, as well as the RML-MCMC sampling algorithm in quantifying uncertainty. We solve the same problem with the same settings as defined in Section 6.5, with 1% noise in the observations, but only with $\beta = 4$ and 6.

The mean of the prior, RML, and MCMC ensembles are shown in Figures 6.30. Here, we see that not much progress has been made in the ensemble mean, as there is hardly any visible difference in the ensemble means. This is what we would expect. However, if we look at the ensemble members in Figures 6.31, we see that the optimization has done a good job in detecting the first layer in the ensemble members while producing different images showing different structures deeper in the subsurface.

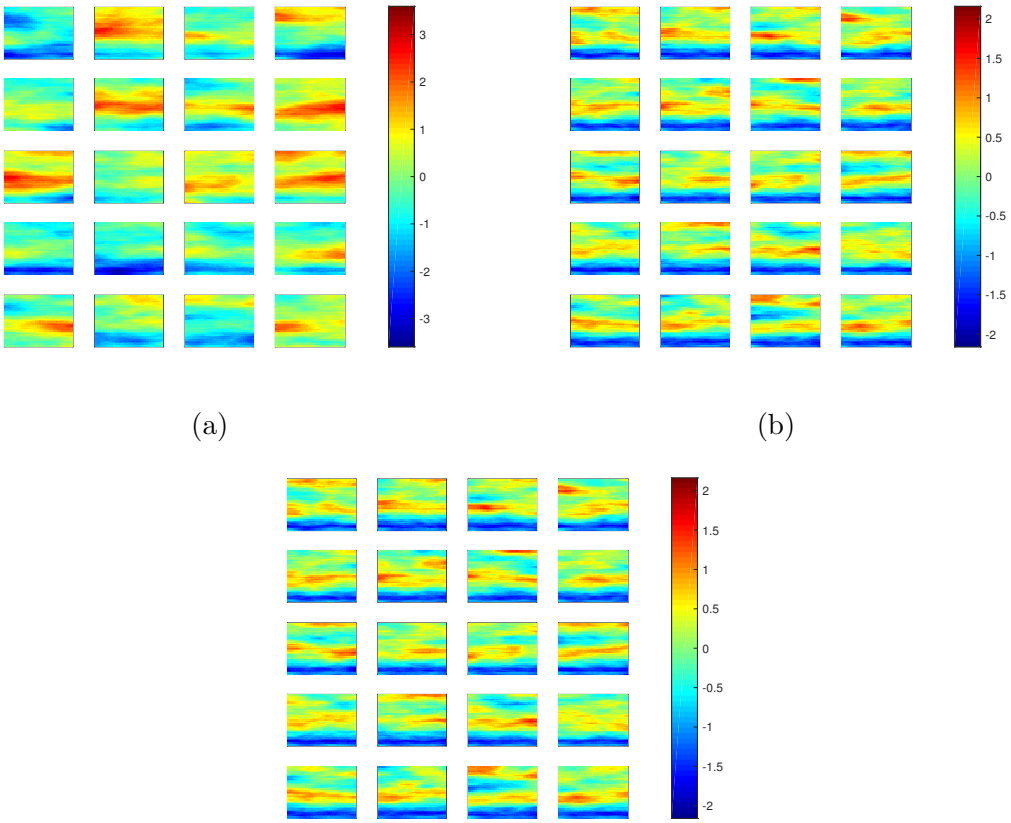
Figures 6.32(a)–(c) show the plot of the standard deviations for the prior, RML, and MCMC ensembles respectively. Note that the ensemble standard deviation here reveals greater fluctuations in the deeper parts of the subsurface. Note also



(a) (b)

(c)

Figure 6.30: Ensemble means for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble for the problem of full waveform inversion, as described in Section 6.5, but with $\beta = 4$ and 6, with 1% observation noise.



(a) (b)

(c)

Figure 6.31: The (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion, as described in Section 6.5, but with $\beta = 4$ and 6, with 1% observation noise.

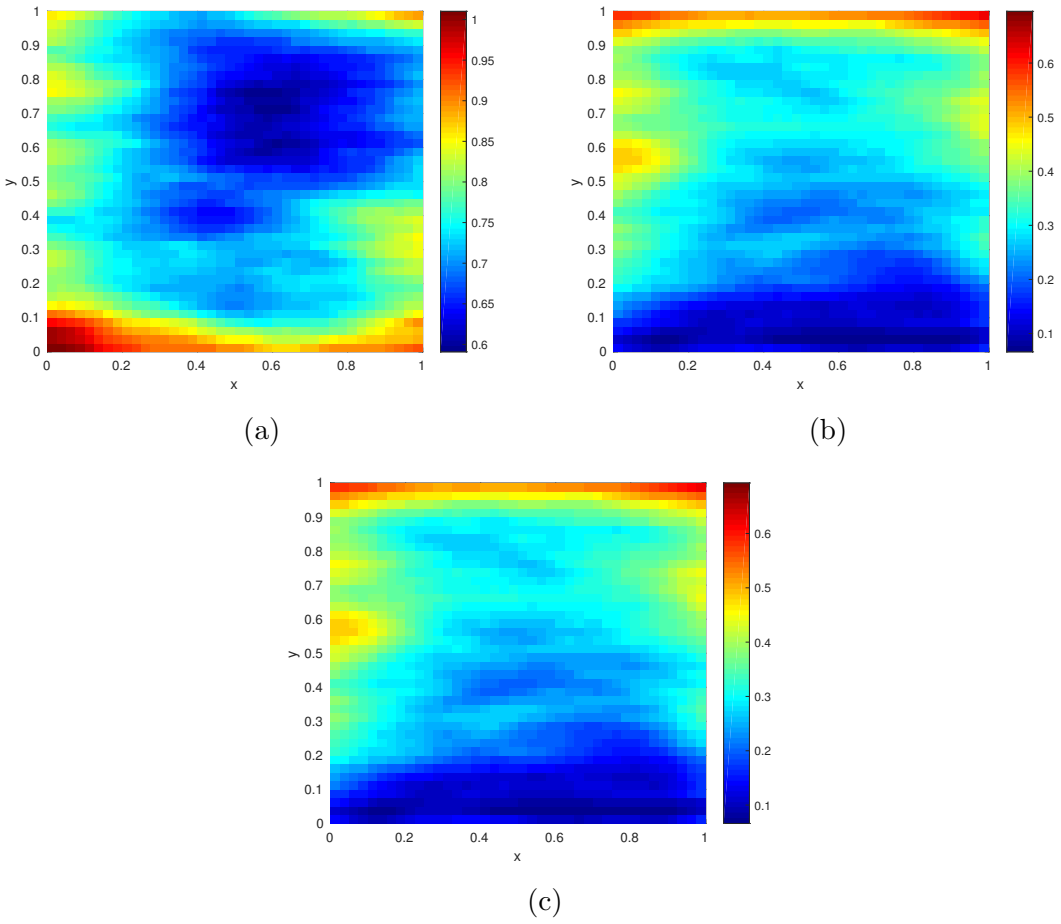


Figure 6.32: Ensemble standard deviations for the (a) prior ensemble, (b) RML ensemble, and (c) MCMC ensemble, for the problem of full waveform inversion, as described in Section 6.5, but with $\beta = 4$ and 6, with 1% observation noise.

that the RML and MCMC ensemble standard deviations are significantly different in structure from the RML and MCMC ensemble standard deviations of the problem as described in Section 6.5 (Figures 6.20(b)–(c)). This highlights the difficulty in detecting uncertainty in the deeper parts of the subsurface without information from the lower frequencies.

The results obtained in this section confirm that, regardless of different settings used in the problem of full waveform inversion, the ability of the algorithm as presented in Algorithm 1 to quantify uncertainty is not hindered.

6.7 Concluding Remarks

In this chapter, we presented the weak constraint formulation for the problem of full waveform inversion. The problem was formulated in the frequency domain through the application of a Laplace transform in time. We have also adopted a simultaneous approach to solve this problem, which processes the observation from all frequencies at once. We applied the RML-MCMC sampling algorithm as described in Algorithm 1 in Chapter 4 to sample from the posterior measure. This was applied to a nontrivial example. Two cases where considered, one where noise was not added to the observations and one where noise was added to the observations. We see that in both cases, the probability maximizers were different due to the absence and presence of observation noise, while the RML and MCMC ensembles shared similar features.

Finally, we demonstrated the robustness of the weak constraint formulation and the RML-MCMC sampling algorithm by changing the settings of the problem as described in Section 6.5. We find that the results obtained were according to what we would have expected, while showing that increasing the difficulty of the problem of full waveform inversion, whether by higher noise levels, fewer observation points, or by using fewer frequencies, does not hinder the ability of the algorithm proposed to generate multiple images and quantify uncertainty.

The application of the RML-MCMC algorithm on the weak constraint formulation for the problem of full waveform inversion gives us a way to generate multiple images, which are realizations of the posterior measure of the speed of the wave in the subsurface, rather than just one image, which is common practice in the industry today. The generation of multiple images allows us to evaluate uncertainty more concretely in order to make more informed decisions.

Chapter 7

Conclusion

7.1 Final remarks

In this thesis, we have presented a framework for solving inverse problems. Namely, we have presented the Bayesian framework in both finite and infinite dimensions. In the case of infinite dimensions, we used Gaussian prior measures as the reference measures. This framework gives us a way to solve any inverse problem, so long as the assumptions are satisfied.

We then discussed the construction of prior measures. In particular, we showed that the hyperparameters in the modified Helmholtz measure and the biharmonic measure have an intuitive physical interpretation. We also showed that, while the modified Helmholtz measure is perfectly viable in finite dimensions, it should not be used in the infinite-dimensional case, as the variance diverges as the grid size is refined.

The framework, as presented in Chapter 2, is known as the strong constraint formulation. This is the usual framework used to solve inverse problems today, and is the one used to solve the problem of full waveform inversion. However, the highly nonconvex property of the misfit functional makes it difficult to obtain useful summaries of the information available in the posterior measure. In particular, it is difficult to minimize the misfit functional, or functionals that take the form of the misfit functional.

Therefore, we instead implemented the weak constraint formulation of inverse problems in Chapter 4. This is an approach that has been used before by meteorologists, and has just been recently introduced into the seismic industry (see van Leeuwen and Herrmann, 2015), but with a full Bayesian interpretation in this thesis. However, the difference between this version of the weak constraint formulation and the version used in meteorology, is that the model noise is added into the differential equation rather than to the solution of the differential equation. This brings the

benefit that the optimization problem that is solved does not require the computation of adjoints, which is a huge computational overhead when such problems are being solved.

We then introduced the RML-MCMC algorithm, which is an algorithm that samples from the posterior measure. This algorithm is a modification of the Randomized Maximum Likelihood method introduced in reservoir simulation by Oliver *et al.* (2008, Chapter 10). Note that the Randomized Maximum Likelihood method as introduced there does not give an exact sample of the posterior measure. Nevertheless, this can be made exact by using the RML ensemble as the starting point in an MCMC simulation. The advantage here is that there is no concern about the independence of the samples and hence the convergence of the MCMC algorithm, since the starting point was independent to begin with, and the optimization has provided a good starting point in the Markov chain.

Then, in Chapter 5, we compared the results between the weak and strong constraint formulations. This was done by examining a model problem, which is the inverse problem on the simple harmonic oscillator. This is a simple, one-dimensional inverse problem, which gives us great insight into the difference between the posterior densities obtained in both the strong and weak constraint formulations.

Finally, we proceeded to apply the weak constraint formulation to the problem of full waveform inversion (FWI). In this formulation, we used the Laplace transform of the acoustic wave equation as the forward model. We employed the simultaneous method, which takes in the observations at several frequencies at once, and showed the effectiveness of the RML-MCMC algorithm in sampling from the posterior measure.

From our simulations, we find that a simultaneous approach could work better in solving the problem of full waveform inversion, rather than using a sequential approach. Further, we also find that a standard FWI could be inferior to a full uncertainty quantification using the Bayesian framework, even if only estimates of the mean are used. This can be seen by comparing the results of the posterior ensemble mean and the probability maximizer in Chapter 6.

Based on our research, recommendations to the seismic industry can be classified into the short, medium, and long term goals. In terms of the short term goals, the recommendation is first to implement a fully Bayesian framework, in particular, the implementation of a prior, or regularization term in the misfit functional.

The medium term goal would be to shift the approach of solving the inverse problem from a strong constraint formulation to a weak constraint formulation.

Long term goals would include:

1. The implementation of a simultaneous approach instead of a sequential one.

2. Drawing samples from the posterior using the RML-MCMC method, and using the ensemble mean instead of the probability maximizer in light of the results in Chapter 6.

It is believed that these recommendations would further improve the results obtained from solving the problem of full waveform inversion, while the sampling algorithm presented will provide an avenue to perform uncertainty quantification, from which an economic decision concerning drilling exploratory wells can be made. Note also that the recommendation of the simultaneous approach has been suggested before in Brossier *et al.* (2009).

7.2 Future work

There are many natural extensions arising from the study of this subject, both in terms of forward modelling and the inverse problem.

In terms of the forward modelling, one can question the accuracy of the numerical solution method of the wave equation, or one could question the accuracy of the wave equation itself, and instead use the full elastic wave equation.

From Chapter 5, an interesting extension would be to consider a forward problem involving two, three or even higher dimension system of harmonic oscillators. This is also relevant as the wave equation itself can be thought of as a coupled system of harmonic oscillators.

From the perspective of the inverse problem, one could start by constructing a more accurate observation model. For example, the geophones used in seismic acquisition do not record pressure perturbations directly. Geophones work by converting vibrations in the earth into electric signals, and this is modelled by a damped harmonic oscillator (Sheriff and Geldart, 1995, Sect. 7.5). Thus, it would be more accurate to take this into account while constructing an observation model in the context of land seismic.

Moreover, the simulations performed here were in a perfect model scenario, where the same equations used to generate the observation were also used to solve the inverse problem. As discussed at the end of Chapter 2, doing this helps us to evaluate the accuracy of the numerical algorithms for obtaining summaries of the posterior measure.

One could construct extensions of this problem to test the effectiveness of the weak constraint formulation further. Two such extensions would be:

1. To use a finer grid to generate the observations, and use a coarser grid to solve the inverse problem.

2. To generate the observations in the time domain, and then solve the inverse problem in the frequency domain.

Both of these extensions are harder inverse problems, as they simulate the solving of an inverse problem without knowing the exact forward model. In these cases, one could consider increasing the variance on the dynamical noise to allow for higher discrepancies between the predicted observation from solving the forward model and the actual observation.

In terms of the weak constraint formulation of inverse problems, one could examine further the probability measure on the model noise. In this thesis, the noise was modelled using a Gaussian measure, but in principle, many other measures could be used.

Specific to the full waveform inversion, it will be interesting to study the frequencies used for the inversion, since there is no known theory for which frequencies would produce the best results. Currently, in practice, the frequencies are chosen arbitrarily, as was done in Chapter 6. However, it would be good if we are able to somehow determine the optimal frequencies to solve the problem that will give the best results in terms of efficiency.

Finally, a theoretical extension of this thesis would be to consider analysing the weak constraint formulation in infinite dimensions. This would involve proving the statements mentioned at the end of Subsection 4.2.2.

Bibliography

Alkhalifah, T. (2014). *Full Waveform Inversion: Where are the Anisotropic Parameters hiding?* EAGE Publications bv.

Amundsen, L. (1991). Comparison of the least-squares criterion and the Cauchy criterion in frequency-wavenumber inversion. *Geophysics*, **56**(12), 2027–2035.

Apostol, T. (1981). *Mathematical Analysis, 2nd Edition*. Addison-Wesley Publishing Company.

Brossier, R., Operto, S., and Virieux, J. (2009). Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, **74**(6), WCC105–WCC118.

Bunks, C., Saleck, F. M., Zaleski, S., and Chavent, G. (1995). Multiscale seismic waveform inversion. *Geophysics*, **80**(5), 1457–1483.

Burstedde, C. and Ghattas, O. (2009). Algorithmic strategies for full waveform inversion: 1d experiments. *Geophysics*, **74**(6), WCC37–WCC46.

Chen, W. D. (1993). A regularization method for the numerical inversion of the Laplace transform. *SIAM J. Numer. Anal.*, **30**(3), 759–773.

Cotter, S. L., Dashti, M., Robinson, J. C., and Stuart, A. M. (2009). Bayesian inverse problems for functions and applications to fluid mechanics. *Inverse Problems*, **25**, 115008.

Cotter, S. L., Roberts, G. O., Stuart, A. M., and White, D. (2013). MCMC Methods for Functions: Modifying Old Algorithms to Make Them Faster. *Statistical Science*, **28**, 424–446.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A. (1990). Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, **55**(5), 527–538.

Da Prato, G. and Zabczyk, J. (2014). *Stochastic Equations in Infinite Dimensions*. Cambridge University Press.

Dashti, M. and Stuart, A. M. (2011). Uncertainty Quantification and Weak Approximation of an Elliptic Inverse Problem. *SIAM Journal on Numerical Analysis*, **49**(6), 2524–2542.

Dashti, M., Harris, S., and Stuart, A. (2012). Besov priors for Bayesian inverse problems. *Inverse Problems and Imaging*, **6**, 183–200.

Farmer, C. L. (2007). Bayesian field theory applied to scattered data interpolation and inverse problems. In A. Iske and J. Levesley, editors, *Algorithms for Approximation*, pages 147–166. Springer.

Fichtner, A. (2011). *Full Seismic Waveform Modelling and Inversion*. Springer-Verlag.

Fisher, M., Leutbecher, M., and Kelly, G. A. (2005). On the Equivalence Between Kalman Smoothing and Weak-Constraint Four-dimensional Variational Data Assimilation. *Q. J. R. Meterol. Soc.*, **131**, 3235–3246.

Kaipio, J. and Somersalo, E. (2005). Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, pages 493–504.

Kaipio, J. and Somersalo, E. (2006). *Statistical and Computational Inverse Problems*. Springer-Verlag.

Kirsch, A. (1996). *An Introduction to the Mathematical Theory of Inverse Problems*. Springer-Verlag.

Law, K., Stuart, A., and Zygalakis, K. (2015). *Data Assimilation: A Mathematical Introduction*. Springer.

Martin, J., Wilcox, L. C., Burstedde, C., and Ghattas, O. (2012). A Stochastic Newton MCMC Method for large-scale statistical inverse problems with application to seismic inversion. *SIAM Journal on Scientific Computing*, **34**(3), A1460–A1487.

Oliver, D. S., Reynolds, A. C., and Liu, N. (2008). *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press.

Piessens, R. (2000). The Hankel Transform. In A. D. Poularikas, editor, *Transforms and Applications Handbook: Second edition*, chapter 9. CRC Press LLC.

Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave equation method. *Geophysical Prospecting*, **38**, 287–310.

Pyun, S., Son, W., and Shin, C. (2011). 3d acoustic waveform inversion in the Laplace domain using an iterative –solver. *Geophysical Prospecting*, **59**, 386–359.

Raknes, E. B. and Arntsen, B. (2014). Strategies for elastic full waveform inversion. In *SEG Technical Program Expanded Abstracts 2014*, pages 1222–1226.

Robert, C. P. and Casella, G. (2004). *Monte Carlo Statistical Methods (2nd. edition)*. Springer Science+Business Media, LLC.

Sheriff, R. E. and Geldart, L. P. (1995). *Exploration Seismology*. Cambridge University Press.

Shin, C. and Cha, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysics J. Int.*, **173**, 922–931.

Shin, C. and Cha, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. *Geophysics J. Int.*, **177**, 1067–1079.

Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, **69**(1), 231–248.

Smith, L. A. (2006). Predictibility past predictibility present. In T. Palmer and R. Hagedorn, editors, *Predictibility of Weather and Climate*, pages 217–250. Cambridge University Press.

Son, W., Pyun, S., Shin, C., and Kim, H.-J. (2014). Laplace-domain wave equation modeling and full waveform inversion in 3D isotropic elastic media. *Journal of Applied Geophysics*, **105**, 120–132.

Stuart, A. M. (2010). Inverse problems: A Bayesian perspective. *Acta Numerica*, **19**, 451–559.

Sullivan, T. J. (2015). *Introduction to Uncertainty Quantification*. Springer-Verlag.

Tan, B.-T., Burstedde, C., Ghattas, O., Martin, J., Stadler, G., and Wilcox, L. C. (2012). Extreme-scale UQ for Bayesian inverse problems governed by PDEs. In *High Performance Computing, Networking, Storage and Analysis (SC), 2012 International Conference for*, pages 1–11.

Tan, B.-T., Ghattas, O., Martin, J., and Stadler, G. (2013). A Computational Framework for Infinite-Dimensional Bayesian Inverse Problems. Part I: The Linearized Case, with Applications to Global Seismic Inversion. *SIAM J. Sci. Comput.*, **35**(6), A2494–A2523.

Tarantola, A. (1984a). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266.

Tarantola, A. (1984b). Linearized inversion of seismic reflection data. *Geophysical Prospecting*, **32**, 998–1015.

Tarantola, A. (2005). *Inverse Problem Theory*. SIAM.

Trémolet, Y. (2006). Accounting for an imperfect model in 4D-Var. *Q. J. R. Meteorol. Soc.*, **132**, 2483–2504.

Trémolet, Y. (2007). Model-error estimation in 4D-Var. *Q. J. R. Meteorol. Soc.*, **133**, 1267–1280.

van Leeuwen, T. and Herrmann, F. J. (2015). A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems*, **32**, 015007.

Virieux, J. and Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC127–WCC152.

Zupanski, D. (1996). A General Weak Constraint Applicable to Operational 4DVAR Data Assimilation Systems. *Monthly Weather Review*, **125**, 2274–2292.