UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Trans-dimensional Bayesian Inversion of Controlled Source
Electromagnetic Data**

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Earth Sciences

by

Anandaroop Ray

Committee in charge:

Kerry Key, Chair
Duncan Agnew
Michael Buckingham
Catherine Constable
Peter Gerstoft
Daniel Tartakovsky

2014

UMI Number: 3623036

UMI

Dissertation Publishing

UMI 3623036

ProQuest®

The dissertation of Anandaroop Ray is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2014

DEDICATION

To Chandra Nath Chatterjee – for contributing the most important equation in

this dissertation:

$$down \neq out$$

EPIGRAPH

When tackling any new skill it is common sense to expect and accept that lots of mistakes will be made. It's part of the learning process.

—Mark Linley,

'How to Draw Anything'

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

In my former professional life in the oil and gas exploration industry, a very astute observation was drilled into me by my unit manager. 'All geophysics is decision support,' they said. Over the years, as I timorously dared squeak out the results of my subsurface geophysical investigations to DRBs (Decision Review Boards, another TLA or Three Letter Acronym), I realized that an enormous amount of geological, geophysical, economic, political and safety considerations went into making a hydrocarbon exploration investment decision.

I had to justify my results, as did a bevy of other geoscientists, in front of a panel of experts whose job it was to find out what errors we may have made both in our individual analyses, and in how we had linked them together. Each scenario the exploration teams came up with, was scrutinized at an exhaustively detailed level. At first, I quailed at the prospect of having to present a panel of experts with inferences about parts of the near subsurface (down to $\sim$10 km depth and $\sim$100 km laterally), derived from observations of data collected on the surface. I then came to the realization that the more experienced hands showed more aplomb not because their work was better, but because they had embraced the idea that their results were uncertain. Of course, as the reader will no doubt imagine, this made me feel a lot better, but not perhaps because of what one may at first be inclined to think. As we will rigorously expound upon in the next few chapters, uncertainty is an inescapable aspect of our lives. The geoscientists in my organization did not do shoddy and uncertain work. It was just that they were well aware that their notions of the buried sediments, were merely interpretations. They knew qualitatively, that there were other possible hypotheses for sediment accumulation, charge formation (the natural chemical process of oil and gas formation), hydrocarbon source migration (movement of hydrocarbons as restricted by geological layering), and structural trap and seal formation (which ensures that the hydrocarbons are ensconced in a space we can drill) – than the ones they had drawn up.

Typically, geophysicists are involved in two stages of this process. First, at a very large spatial scale ($>$ 100 km laterally), using topographic, gravity and magnetic data in helping outline the area and depths of interest. This directly influences the hypotheses of sediment accumulation and charge formation. Second, they are also involved at the

more detailed kilometer scales (both in depth and area), for the imaging of structures (using seismic data) and analyses of signatures which indicate hydrocarbon presence (using controlled source electromagnetic or CSEM data, for instance).

However, until a well is safely drilled after negotiating the high pressures found in deep rock formations, one cannot be sure that we have indeed struck oil. This is because only a successful 'down-hole' signature from the well using sophisticated instruments and the return of hydrocarbons in the drilling fluid indicate hydrocarbon discovery. Then of course, there is the issue of the volumes discovered and their economic profitability. A rule of thumb, averaging worldwide, is that only one-in-three wells are successful, and it is getting harder and harder to find oil and gas in easy to access locations.

This one-in-three statistic fully reflects the uncertain nature of the oil and gas exploration industry. Our observations of the data reflecting the presence of oil and gas are indirect, incomplete, and noisy. From these independent data, we must multiplicatively infer what is known as a POS or a probability of success for the prospect. Then there are other probabilistic measures of the oil and gas reserves in the given field and what that ultimately translates to, also probabilistically, in terms of estimated demand, production costs, and ultimately, cost recovery over timelines that are a significant portion of a human lifetime. The monetary figures involved easily go into billions of dollars.

At present, the idea of the uncertainties involved in exploration and reservoir management are mostly qualitative, and inferred relative to experience with previous exploration and production campaigns. While this experience of 'for the previous oilfield, statement X held true, therefore for this one ....' generally works, the process of uncertainty appraisal is in dire need of quantification (pers. comm. with G. Michael Hoversten). Integration of all the involved data, geological, geophysical, political, economic and of all other sorts, in a waterproof quantitative framework which accounts for all uncertainty – is probably impossible. However, we can characterize the risks and costs involved in oil drilling. For example, we can try and reliably provide the chance of encountering hazardous shallow gas or the presence of economically viable hydrocarbons, by analyzing the available geophysical data in a rigorous probabilistic framework. This is possible as it involves rigorous physics, statistics, and some well established laws of chance. Once we have characterized the uncertainty, we are then in a position to try and

address it in our operations.

ORGANIZATION OF THE DISSERTATION

In this dissertation, we will attempt to demonstrate uncertainty quantification in subsurface geophysical analysis. We will do so by focusing on using geophysical data derived from CSEM surveys within a Bayesian probabilistic context. A CSEM survey and the subsequent 'inversion' of CSEM data allows us to infer subsurface resistivity in the survey area. Since oil and gas reservoirs are highly resistive relative to their surrounding sediments, CSEM is an exploration tool that is a good indicator of hydrocarbon presence. In this vein, the dissertation is organized as follows:

Chapter 1 provides a brief introduction to Bayesian inversion and the physics of marine CSEM. Chapter 2 develops the mathematical machinery for a trans-dimensional algorithm, which spans multiple hypothesis spaces. The algorithm is applied to the CSEM inverse problem for anisotropic subsurface resistivity. A 1D implementation with real data examples from the Pluto gas field is presented. Chapter 3 implements an algorithm using 'replica exchange' or 'parallel tempering' to speed up convergence to the desired resistivity uncertainty in CSEM inversion. In Chapter 4, the trans-dimensional method is implemented using a 2D parameterization with Voronoi cells. It is then applied to real CSEM data from the Scarborough gas field. Chapter 5 is a brief essay with concluding remarks.

ACKNOWLEDGEMENTS

I suppose that thanks first and foremost should go to Chandra Nath Chatterjee, for convincing me more than a decade ago that a future as a historian was probably not one I would be able to reconcile myself to. I would like to thank all my teachers and mentors who over the years did nothing to cure me of my insufferable enthusiasm for whatever it was they had the misfortune to be teaching me. Chief among them of course, is my advisor Kerry Key, whose passion for EM geophysics and inversion not only parallels but far surpasses mine. Not only does he live with the 'curse of the dipole,' he thrives in it. He has been a great mentor who gave me just enough rope, not to hang myself with. For getting a doctoral degree in something one really enjoys, I can't think of a better place to get it in than coastal California, and I must with deep gratitude thank Steve Constable for welcoming me to the Marine EM lab at UCSD.

Life isn't always a bed of roses, as we well know. I for one have never actually seen a bed of roses. Peter, iHeatYou, Anita (and gang), Preeti & Co., Sir Rat, Tam and Tildie (and the rest of the gang), and Bikram – thank you for having stood with me through thick and thin. If I've ever needed to head shorewards, you've more than provided safe haven.

Of course, none of this work would ever have seen the light of day had my unflappable room-mate Madush not put up with my incessant playback of Sigur Rós and Explosions in the Sky well past the witching hour and proclivity to discuss my misconceptions about Buddhism, Life, The Universe and Everything Else with him. Not that he knows much about these things either. It's just that he put up with it, is the point I'm trying to make.

I must thank my committee members for being incredibly enthusiastic about me as a person and not just a researcher through various stages of my research, whether it be connected to inverse theory, underwater acoustics, statistical signal processing, agreeing to videoconference when not physically available, high-fiving me to express solidarity in our lack of understanding of quantum mechanics, or taking me out flying in order for me to impress a wonderful young lady.

Mike Hoversten at Chevron, is of course accountable for first piquing my interest in MCMC sampling, and then letting it run unbounded from there. Two early career

scientists who have selflessly helped me come to grips with the deceptive simplicity that is trans-dimensional Bayesian inversion are Thomas Bodin (UC Berkeley) and Jan Dettmer (University of Victoria). Thanks go out to them for their prompt responses to my emails, meeting me when I've had the good fortune to attend the same conferences as them, as well as sharing office space at UC Berkeley.

————————————

Chapter 2, in full, is a reprint of material as it appears in the Geophysical Journal International, Ray, A., and Key, K., Blackwell Publishing Ltd., 2012. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in full, is a reprint of material as it appears in Geophysics, Ray A., Alumbaugh D. L., Hoversten, G. Michael, and Key K., Society of Exploration Geophysicists, 2013. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in full, has been submitted for publication of the material as it may appear in the Geophysical Journal International, Ray, A., Key, K., Bodin, T., Myer D., and Constable S., John Wiley and Sons Ltd., 2014. The dissertation author was the primary investigator and author of this paper.

2004	B. Sc. in Physics with Honours, University of Delhi
2007	M. Tech. in Applied Geophysics, Indian Institute of Technology, Roorkee
2007–2010	Potential Fields Geophysicist with Shell Exploration and Production
2012	Researcher in the Modeling & Inversion group, Chevron ETC
2013	Researcher in the Reservoir Properties from Seismic group, Chevron ETC
2010–2014	Research Assistant, University of California, San Diego
2014	Ph. D. in Earth Sciences, University of California, San Diego

PUBLICATIONS

Ray, A., Alumbaugh, D. L., Hoversten, G. M., and Key, K., 2013: Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering. Geophysics, 78(6), E271–E280.

Ray, A., and Key, K., 2012: Bayesian inversion of marine CSEM data with a transdimensional self parametrizing algorithm. Geophysical Journal International, 191, 1135–1151.

Cullen, A., Reemst, P., Henstra, G., Gozzard, S. and Ray, A., 2010: Rifting of the South China Sea: new perspectives. Petroleum Geoscience, 16(3), 273–282.

Ray, A., and Srivastava, D. C., 2008: Non-linear least squares ellipse fitting using the genetic algorithm with applications to strain analysis. Journal of Structural Geology, 30(12), 1593–1602.

ABSTRACT OF THE DISSERTATION

**Trans-dimensional Bayesian Inversion of Controlled Source
Electromagnetic Data**

by

Anandaroop Ray

Doctor of Philosophy in Earth Sciences

University of California, San Diego, 2014

Kerry Key, Chair

This dissertation is an attempt to apply the powerful tool of Bayesian inference to the elegant physics of controlled source electromagnetic (CSEM) propagation through marine sediments with high resistivity contrasts. CSEM is highly effective in the detection of resistive hydrocarbon accumulations in conductive sediments, but its use requires careful analysis owing to the problems of sparsely sampled data, noisy observations and non-uniqueness. Bayesian inference allows us to tackle these limitations in a quantitative, probabilistic framework by using prior information about the geology in question and the statistics of the data noise. We extend conventional Bayesian analyses to a second level, where we infer the complexity of subsurface resistivity models required to explain the observed noisy CSEM data, from the observations themselves. Although we

have focused on the CSEM method, the techniques discussed are generally applicable to many problems in the geosciences.

Chapter 1

Introduction

1.1 Why use Bayesian inference in geophysics?

If we examine Figure 1.1, we can see a progression from left to right. The first image shows an artist's caricature of a person, sketched within ten minutes. Impressive as it may be, the next image shows the caricature, now barely recognizable and pixelated. Only every eighth row and column of the original image has been preserved and noise has been added to it. The next three images show fairly similar attempts at reconstructing the original caricature. The last image to the right, now shows the mean of many reconstructions, and restores a smidgeon of dignity to the artist's work and her subject's visage. Omitting the details for now, the mean reconstructed image was formed by sampling many images like the three shown to its left, and then summing them up. The sampling process used a Bayesian philosophy, which only required *a priori* knowledge of

Figure 1.1: From left to right: A caricature drawn by an artist, the same image decimated and now with noise added to it, three reconstructions that fit the noisy image and finally, the mean of many reconstructed images.
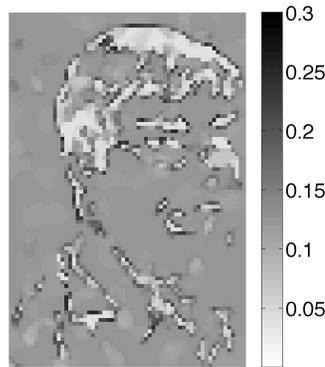
Figure 1.2: The standard deviation of reconstructed images which fit the decimated noisy image. The darker parts of the original image have less variance.

the kind of noise corrupting the image, and some information regarding the geometric extents of the true image and the range of its brightness values.

In the field of geophysics, an integral part of our analysis of the earth's structure first begins with the collection of data using instruments of some kind such as magnetometers, seismometers or gravimeters, to name just a few. The data are always noisy and sampled in a spatially discontinuous manner, just like in the caricature. The reason behind the imperfect sampling is simple – we humans are tiny, the earth is large and we can't possibly land our instruments in every nook and cranny of the earth's surface. There is also the issue of the significant monetary cost involved in acquiring data in a denser fashion.

The reasons behind the presence of noise in the observations are more subtle, and indeed more philosophical. What constitutes 'noise' and 'signal' essentially lie in the eye of the geophysical beholder. For many purposes, including ours in this dissertation – it is adequate to treat as noise, fluctuations in the observed geophysical signal that are not of our interest. Further, since we cannot possibly model these fluctuations at all locations and at all possible times, we treat them in a statistical fashion as being random. Once we are satisfied that we have an adequate statistical representation of the noise, and possess some deterministic physical mechanism that allows us to emulate the observed signal, we are in business and able to begin geophysical analysis. The difference between the emulated signal and the observed signal allows us to quantify a residual quantity –

which we interpret as noise. Given our knowledge of the noise statistics, we know how likely a given residual should be.

Bayes' theorem (Bayes and Price, 1763) is essentially a mathematical formulation that lets us translate the observed noise into an ensemble of models compatible with the data, just like the three similar images in columns 3, 4 and 5 of Figure 1.1. If we have a sufficiently large ensemble of models, most of which fit the data within the observed noise – we can then find statistical properties of the ensemble such as the mean (Figure 1.1, extreme right) or standard deviation as shown in Figure 1.2.

For geophysical data, the statistical properties of the reconstructed model ensemble - known in Bayesian parlance as the 'posterior' model ensemble, contain valuable information about the earth (Tarantola and Valette, 1982; Aster et al., 2005). From such a model ensemble, one can obtain an idea of the uncertainty with which we are able to reconstruct physical models of the earth, given noisy geophysical observations at discrete locations such as satellites, marine ship tracks, ground or airborne surveys, etc.

1.2 The methodology of geophysical inversion

When applied to geophysical data, the process referred to in the previous section for illuminating earth structure is known as geophysical inversion. Although the methodology outlined so far has been probabilistic and Bayesian, it does not need to be. At the risk of invoking recursive logic at this point, we can state that solving an inverse problem requires solving a forward problem and an inverse problem. The forward problem, is of the following nature, for example – 'given the shape and size of a lion's vocal chords, what will it sound like when it roars?' A corresponding inverse problem could be, 'given the sound of a lion's roar, what can we say about the shape and size of its vocal chords?' Of course, to answer the latter question, we must have some idea of what kind of vocal chord structure produces a certain kind of roar. We could make a very simple approximation of a vocal chord as a tube of a given length, and use the physics of wave propagation to map a vocal chord of given length to an ensemble of sounds at various frequencies. This would be the solving of the forward problem. Presumably, we could then listen to different lions roaring at a safe distance away in the African savannas, and

then be able to say something about the lengths of their vocal chords, given that longer vocal cavities in larger lions would produce lower frequency roars.

In geophysics, we ask this sort sort of question all the time – given observations of the earth's gravity field on the surface, what can we say about the structure of the basin buried in Cretaceous sediments? Given the seismograms providing earthquake ground motions, what can be inferred about the source geometry?

To formally go about this process, we first need a model of the earth, $\mathbf{m}$, which we can can map to the observed data $\mathbf{d}$. For any general subsurface geophysics problem, we could discretize the earth's subsurface into m intervals and observe the data at n points on the earth that are easily accessible to us. We could then use a functional f which maps these m parameters in $\mathbf{m}$ to the n observed points in $\mathbf{d}$ in the following manner:

$$\mathbf{d} = f(\mathbf{m}). \tag{1.1}$$

So it might appear that geophysical inversion is simply a case of finding $\mathbf{m}$ which minimizes the 'size' of the quantity

$$\mathbf{d} - f(\mathbf{m}), \tag{1.2}$$

but this is where it begins to get tricky. There are all kinds of questions hidden in the minimization of the 'size' of expression (1.2). Does a model even exist? If so, what is the most appropriate functional f? What kind of *norm* (measure of length) should we use to minimize the 'distance' between the observed data $\mathbf{d}$ and the prediction $f(\mathbf{m})$? How should we discretize the model vector $\mathbf{m}$? Is the ensuing model from this minimization sensible? These and many other questions have been answered at length in the classic texts on geophysical inversion such as those by Parker (1994), Menke (2012), Tarantola (2005) and Aster et al. (2005) to name just a few.

If the problem is linear, such that $\mathbf{d} = \mathbf{Gm}$ where $\mathbf{G}$ is an $n \times m$ matrix that linearly maps the $m \times 1$ vector of model parameters to the $n \times 1$ vector of noiseless, perfect, observed data, there are 'one shot' methods of finding $\mathbf{m}$ which minimizes

the Euclidean norm of expression (1.2). The classic Gaussian least squares solution is $\mathbf{m} = (\mathbf{G^tG})^{-1}\mathbf{G^td}$, but as is quickly found out by first year grad school geophysicists, we can run into unrealistic looking solutions that fit the data well, but make no physical sense. The problem may be ill conditioned, with small errors in the observations leading to models that are very dissimilar to the true model. The observations may not be independent. Worse, if there are more unknowns than data observations then additional information may be required to obtain a solution. One sensible idea to adopt in this case is to find a model of 'least length' or the 'minimum norm' solution. We may also wish to trust certain data points more than others and weight them accordingly. All these cases have been dealt with insightfully and in detail in Menke (2012).

Many geophysical inverse problems are not well posed in the classical sense. Solutions, if they exist may not be unique, and can vary wildly for slight perturbations in observed data. For many geophysical problems, especially those in electromagnetics, $f(\mathbf{m})$ is not a linear transformation – consequently, acceptable solutions to the inverse problem may not vary continuously with the observed data and parts of the model space may be 'forbidden', as shown for example by Sneider and Trampert (2000) and in section 3.2.1 of this dissertation (Ray et al., 2013).

One strategy for solving a non-linear inverse problem is to use a Gauss-Newton approach – from an initial starting model, $f(\mathbf{m})$ is approximated using a Taylor series expansion (to first order) about the starting model. An update to the model is found by solving a linear problem, and this process is continued until the algorithm converges upon a minimum of the surface defined by the Euclidean norm of the expression (1.2). However, as pointed out by Parker (1983), the model with minimum misfit may not be a very geologically sensible model. Indeed, if the observation errors are Gaussian, the sum squared errors will be χ^2 distributed and the χ^2 distribution for more than two independent observations is absolutely improbable at zero sum squared error. There are various possible solutions which restrict overfitting the data, by trading off the data misfit given by a norm of the expression (1.2) with a norm of the model vector $\mathbf{m}$. This is done by simultaneously minimizing the data misfit norm and the model norm, and choosing a weight λ which adequately balances the model complexity (or size) and the data fit. This is known as model *regularization* and λ is known as the *trade-off* or

regularization parameter. For example, with the Euclidean norm, instead of minimizing only expression (1.2), we minimize the following objective function

$$u = ||\mathbf{W}\mathbf{d} - \mathbf{W}f(\mathbf{m})||^2 + \lambda||\mathbf{R}\mathbf{m}||^2, \qquad (1.3)$$

where $\mathbf{R}$ is a matrix that operates on $\mathbf{m}$ to provide a measure of smoothness such as a first or second derivative, or it could simply be an identity matrix. $\mathbf{W}$ is a matrix of data weights, typically the inverse covariance matrix of data errors. If λ is large, the model size is penalized heavily relative to the data misfit, and conversely a smaller λ will result in better fit data, but with perhaps unreasonably complicated models. Model regularization involves choosing λ, the operator matrix $\mathbf{R}$ and even a suitable model norm, which need not be L2 or Euclidean. There are various guiding principles in making sensible regularization choices, such as choosing the smoothest model which fits the data within a given tolerance (e.g., Constable et al., 1987).

In a Gauss-Newton type approach, to calculate the Taylor expansion of the forward function, a first derivative of $f(\mathbf{m})$ with respect to $\mathbf{m}$ needs to be calculated. This derivative is known as the $n \times m$ 'Jacobian' or 'model sensitivity' matrix. For large problems with many data and model parameters this presents both a computational and storage challenge. Conjugate gradient methods (Mackie and Madden, 1993; Rodi and Mackie, 2001) calculate only the action of the Jacobian matrix on an arbitrary vector and thus mitigate these computational issues. However, the method of choosing the regularization parameter λ in these methods is rather arbitrary.

Both the Gauss-Newton and conjugate gradient approaches involve following a gradient down a surface defined by some objective function as in equation (1.3) to a point of minimum. They rely on being close to the global minimum (the 'true' solution) of the objective function instead of an isolated minimum pocket. If there are various minima with acceptable levels of data misfit, these methods provide only one model solution. Further, the process of performing model regularization in both of these approaches is rather an art, and can produce sensible, but very different types of models with the same level of data misfit. Bayesian methods as we will see in the later sections of this chapter, though far more computationally expensive, do not require any linearization nor the difficult calculation of numerical derivatives. They produce results in the form of model
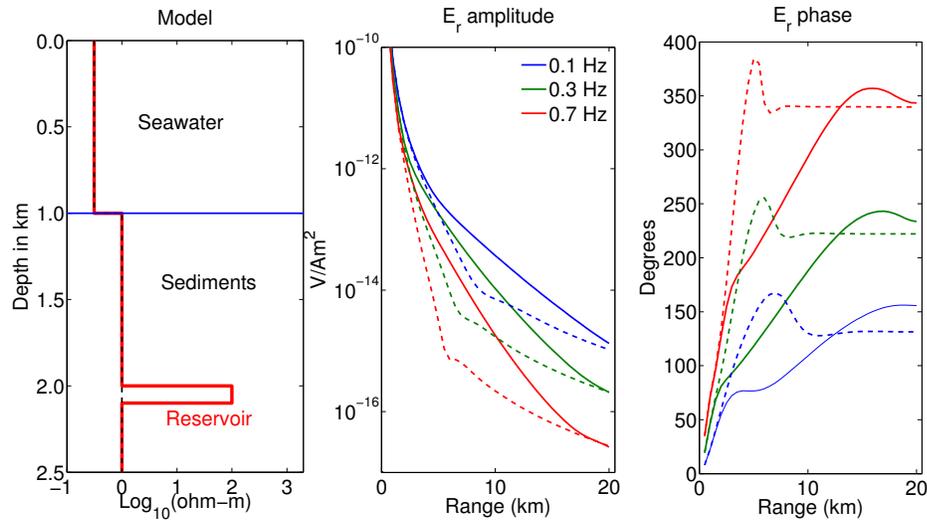
Figure 1.3: CSEM radial electric field E_r responses at the seafloor in 1 km deep sea water at three transmission frequencies. The left pane shows a 100 m thick resistive reservoir embedded in conducting sediments (red model). The background model without a reservoir is shown in dashed black. The middle pane shows the amplitude of the electric field and the right pane its phase. Solid lines correspond to the reservoir response and dashed lines indicate the background response in the absence of reservoir.

ensembles that are not shy of highlighting non-uniqueness in the model space. Further, given sensible prior boundaries derived from what is known about the subsurface, they do not require any subjective model regularization.

1.3 Why use controlled source electromagnetic sounding (CSEM)?

As a hydrocarbon exploration tool for subsurface imaging of geological structures – the reflection seismic method is unparalleled. The physics of seismic wave propagation obey the wave equation, and the accompanying high resolution is what would be expected of wave-like phenomena. However, the seismic method is extremely sensitive to contrasts in the product of velocity and density (acoustic impedance) between geological layers in the subsurface. For hydrocarbons, especially gas – as a seismic wave propagates from a saline water bearing medium to a hydrocarbon bearing medium, there is a very large

reflection response observed at the surface that saturates with the presence of even minute amounts of gas. On the other hand, electromagnetic energy, when diffusing through conductive media at low frequencies, is sensitive to high resistivity contrasts without the response being saturated. An oil-water or gas-water interface presents precisely such a contrast, given that saline sediments and hydrocarbon bearing sediments are highly conductive and resistive respectively. Of course, in any realistic exploration scenario, one must combine information from many different geophysical and geological sources before making a decision to drill a well.

Marine CSEM is a low frequency active source sounding method, the use of which is motivated by its sensitivity to resistivity, which is more indicative of hydrocarbon presence (as opposed to geological structure). The marine CSEM response at the seafloor in 1 km deep sea water due to a 100 m thick reservoir at 1 km depth below the seafloor is shown in Figure 1.3. Solid lines represent responses in the presence of a reservoir, and dashed lines indicate CSEM response in the absence of a reservoir.

Much of the pioneering work to do with the marine CSEM method was carried out at the Scripps Institution of Oceanography, the National Oceanography Centre Southampton, and Cambridge University. A comprehensive review can be found in Constable (2010). The physics of marine CSEM propagation has been analyzed in great detail in previous work (e.g., Chave and Cox, 1982; Evans, 1991; MacGregor, 1997; Loseth, 2007; Key, 2009) and a recent review can be found in Key (2012). A key point of note is that the problem is highly non-linear, and model responses cannot linearly be related to model parameters. The basic principles of electromagnetic wave propagation as applicable to the marine CSEM problem are discussed in the next subsection.

1.3.1   Maxwell's equations in a conducting medium

Maxwell's equations in a source free, isotropic, homogenous conducting medium for the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$ are given as:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.4}$$

$$\nabla \cdot \mathbf{E} = 0, \tag{1.5}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.6}$$

$$\nabla \times \mathbf{B} = \mu\sigma\mathbf{E} + \mu\epsilon\frac{\partial \mathbf{E}}{\partial t}, \tag{1.7}$$

where μ is the permittivity of the medium, ϵ its permeability and σ its conductivity. For CSEM applications, it is the conductivity σ or its inverse, resistivity ρ that we are most interested in. Taking the curl of equation (1.4), using the identity, $\nabla \times \nabla \times \mathbf{F} = -\nabla^2\mathbf{F} + \nabla(\nabla \cdot \mathbf{F})$ for a vector field $\mathbf{F}$, from equations (1.5) and (1.7) we obtain a general version of the wave equation with an attendant diffusion term proportional to the first derivative of the field with time:

$$\nabla^2\mathbf{E} = \mu\sigma\frac{\partial \mathbf{E}}{\partial t} + \mu\epsilon\frac{\partial^2\mathbf{E}}{\partial t^2}, \tag{1.8}$$

Defining a Fourier transform pair according to the following convention:

$$\mathbf{F}(\mathbf{k}, \omega) = \int\int \mathbf{F}(\mathbf{r}, t)e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}d\mathbf{r}\,dt, \tag{1.9}$$

$$\mathbf{F}(\mathbf{r}, t) = \frac{1}{8\pi^3}\int\frac{1}{2\pi}\int \mathbf{F}(\mathbf{k}, \omega)e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}d\mathbf{k}\,d\omega, \tag{1.10}$$

$$\tag{1.11}$$

where $\mathbf{k}$ is the spatial wavenumber and ω the angular frequency, we apply the forward transform as defined in equation (1.9) to equation (1.8). Given the properties of the Fourier transform whereby derivatives are transformed into multiplicative quantities (e.g., Arfken et al., 2011), we obtain the following relation:

$$k^2 = i\omega\mu\sigma + \epsilon\mu\omega^2. \tag{1.12}$$

The first term in equation (1.12) is called the conduction term and the second term is known as the displacement term. Note that the displacement term does not depend on conductivity σ. The wavenumber k is complex, with the real and imaginary parts given as follows (Ward and Hohmann, 1987):

$$Re\{k\} = \omega\sqrt{\frac{\epsilon\mu}{2}}\left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} + 1\right]^{1/2}, \tag{1.13}$$

$$Im\{k\} = \omega\sqrt{\frac{\epsilon\mu}{2}}\left[\sqrt{1 + \left(\frac{\sigma}{\epsilon\omega}\right)^2} - 1\right]^{1/2}, \tag{1.14}$$

We can see from the inverse transform in equation (1.10) that solutions to the general wave equation (1.8) are superpositions of the form

$$\mathbf{E} = \mathbf{E_0}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}. \tag{1.15}$$

Without loss of generality, if we are to consider propagation now only in the $\hat{\mathbf{z}}$ direction, we can admit solutions of the form

$$\mathbf{E} = \mathbf{E_0}e^{i(kz-\omega t)}, \tag{1.16}$$

$$\mathbf{E} = \mathbf{E_0}e^{-Im\{k\}z}e^{i(Re\{k\}z-\omega t)}. \tag{1.17}$$

Thus we can see that there is an exponential spatial decay associated with the imaginary part of the wavenumber, and the propagation phase is controlled by the real part of the wavenumber. For a typical 1 ohm-m marine sediment, a plot of k^2 vs frequency is provided, with the separate contributions from the conduction and displacement terms shown in Figure 1.4. Thus we can see that until we are in the GHz range, the conduction term is orders of magnitude larger than the displacement term. To keep the spatial decay small yet allow propagation through marine sediments such that we are able to sense the presence of roughly kilometer deep discontinuities in resistivity, we must limit the frequency range of operation for marine CSEM. Frequencies from 0.1 to 10 Hz are used, as shown by the pink bar on the frequency axis. This is still the range where $\omega\mu\sigma >> \epsilon\mu\omega^2$
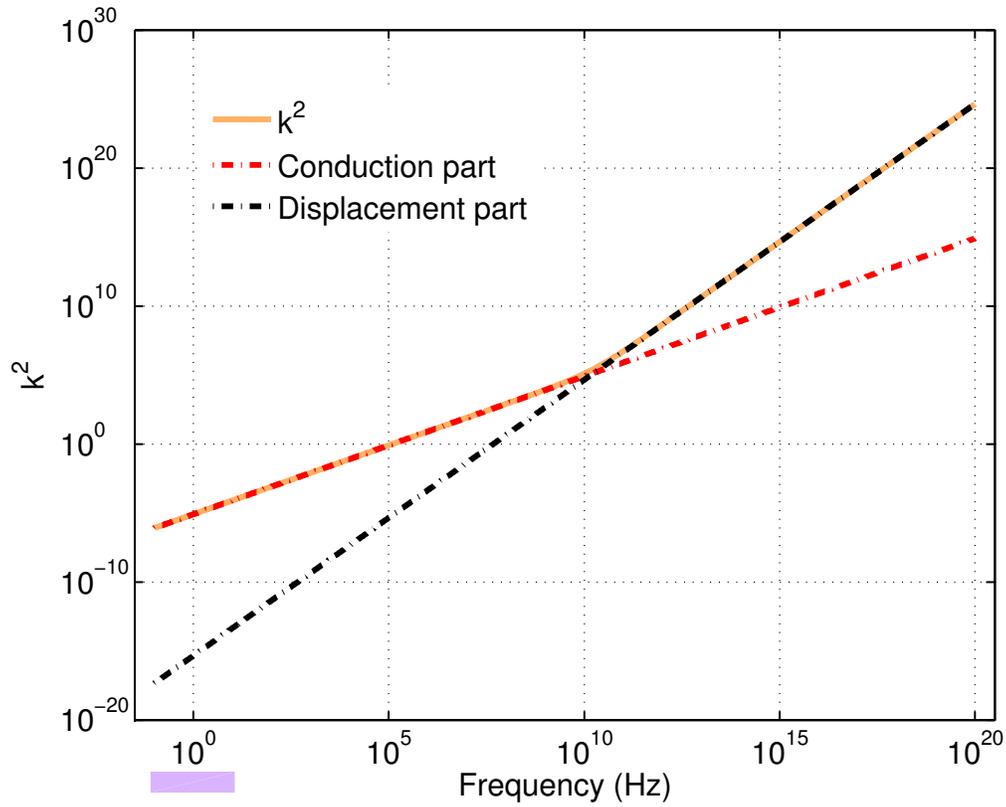
Figure 1.4: Squared magnitude of the wavenumber k shown with its constituent parts, as defined in equation (1.12) through a 1 ohm-m sediment. CSEM frequencies of operation are shown by the pink bar on the frequency axis. In this frequency range, the displacement term is negligible. Note that the displacement term is independent of conductivity and is the only term present in propagation through air.

implying that we are in the 'quasi-static' regime with

$$\frac{\sigma}{\epsilon\omega} >> 1, \tag{1.18}$$

and from equations (1.13) and (1.14) we obtain the following expression for the wavenumber:

$$Re\{k\} = Im\{k\} = \sqrt{\frac{\mu\omega\sigma}{2}} = \frac{1}{\delta}, \tag{1.19}$$

where $\delta = \sqrt{\frac{2}{\mu\omega\sigma}}$ is known as the 'skin depth.' It governs the propagation distance at which the wave has an amplitude that is $1/e$ of its value at the origin. We can also define a wavelength λ from the definition of the wavenumber:

$$Re\{k\} = \frac{2\pi}{\lambda}, \tag{1.20}$$

$$\lambda = 2\pi\delta, \tag{1.21}$$

$$\lambda = 2\pi\sqrt{\frac{2}{\mu\omega\sigma}}. \tag{1.22}$$

Thus we see a peculiar characteristic of diffusion problems and why they provide such low resolution. Before a plane wavefront has travelled a full wavelength λ, its amplitude attenuates to a value which is $1/e$ of its original value in a distance $\delta = \lambda/2\pi$ (Griffiths, 1999). This is why we cannot think of the CSEM method as being similar to the seismic method – although the math is the same, the governing physics are different. In Figure 1.5, a plot of skin depths with geologically relevant resistivities is provided.

1.4 Uncertainty in resistivity from CSEM inversion

As a consequence of its diffusive nature, robust inferences made from a CSEM survey are necessarily from inversion of the data, and not merely from examination of the data itself (Weiss, 2007). Typically, regularized and linearized gradient based inversion methods have been used to arrive at models that minimize data misfit and are also 'optimal' in some user-defined sense. For instance, models can be pre-determinedly smooth or prejudiced to be close to a reference model. By means of regularization, highly oscillatory features in the model that are thought to be outside the resolution of CSEM
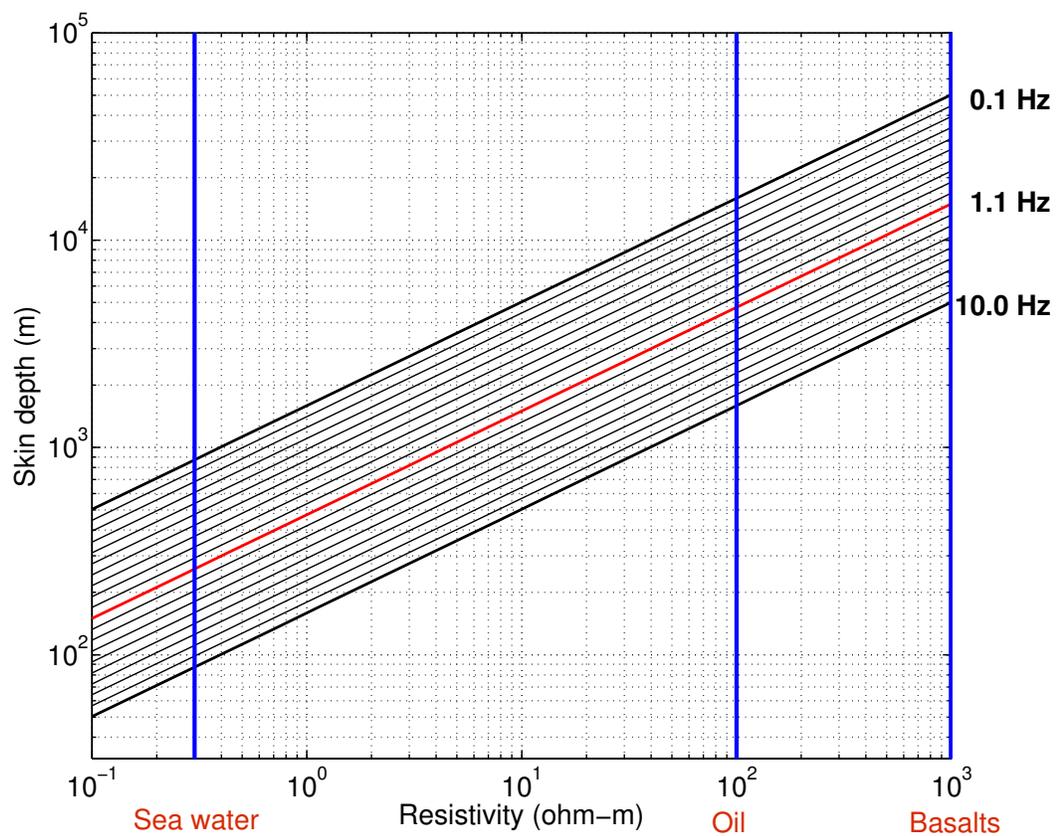
Figure 1.5: A plot of skin depth with resistivity at several different frequencies. Geologically relevant values of resistivity have been marked on the resistivity axis with blue lines.

are suppressed (e.g., Constable et al., 1987; Newman and Alumbaugh, 2000; MacGregor and Sinha, 2000; Abubakar et al., 2008; Key, 2009; Sasaki, 2013; Mittet and Gabrielsen, 2013). Although gradient based inversion methods methods are highly efficient and well understood, they provide either a single smooth model as a result, or a suite of smooth models. These models provide a limited insight into the various classes of models that are compatible with the observed data given the noise. Furthermore, a clear understanding of the resolvability of subsurface resistivity and non-uniqueness of the final solution does not emerge from a linearized treatment of the non-linear CSEM problem.

To quantify the uncertainty inherent in the inversion of CSEM data, one can utilize a Bayesian framework where information is expressed as probability density functions or PDFs (Tarantola and Valette, 1982; Scales and Sneider, 1997). As alluded to in Section 1.1, in such a framework, model parameters are treated as random variables, and their fit to the observed data given the observed statistical noise allows one to formulate a model likelihood. To make the connection with deterministic inversion methods, to first order, models with low misfit possess a higher likelihood. After incorporating prior knowledge of the models that is independent of the data, the product of the prior model probability and the likelihood is known as the posterior model probability. This posterior PDF describes the full solution to the inverse problem — it represents the probability of the model, given the observed data. Those parts of the model space that are more frequently required by the data than other parts manifest with greater posterior probability, and hence are more certain to be properties of the earth (Backus, 1988).

Before moving on to the specifics of determining uncertainty in CSEM inversion, it must be understood here that uncertainty is an inescapable aspect of geophysical inversion. Whether we are interested in drilling a multi-million dollar well based on the hypothesis that there is an oil reservoir at a certain depth at the proposed drill location, or putting forward a hypothesis for tremor source locations in a resistive region of the lithosphere, we should be aware that alternative hypotheses can exist (Wheelock, 2012), making it important to evaluate the uncertainty associated with the hypothesis model space (MacKay, 2003).

1.4.1 Putting CSEM inversion and Bayes' theorem together with Markov Chain Monte Carlo (MCMC) sampling

In the inversion of CSEM data, as for any realistic geophysical problem there are numerous parameters to determine. A Bayesian analysis requires the sampling of a large number of models, and consequently the forward evaluation of these models can be computationally very demanding. An exhaustive grid search over model parameters requires an exponentially increasing number of computations as the number of model parameters to solve for is increased. Hence, for Bayesian inversion, owing to this 'curse of dimensionality' we use Markov Chain Monte Carlo (MCMC) sampling which is essentially a guided random walk through the more probable parts of the posterior model space (e.g., Liang et al., 2011). An MCMC sampling scheme is orders of magnitude less computationally expensive than a corresponding grid search when the number of parameters is high.

We will demonstrate the use of MCMC in this section with a toy two parameter 1D CSEM inverse problem where the depth of the reservoir bottom and its resistivity are considered to be unknown. As shown in Figure 1.3 the true model values are 2100 m and 100 ohm-m. Synthetic CSEM data were calculated and 4% Gaussian noise was added to the data. Before we jump into the inversion, a few terms need to be defined.

Bayesian information is contained in probability density functions (PDFs) represented by $p(\cdot)$. Using Bayes' theorem, (Bayes and Price, 1763) we write

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m}) \cdot p(\mathbf{m})}{p(\mathbf{d})}, \tag{1.23}$$

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior assumptions}}{\text{evidence}}. \tag{1.24}$$

For Bayesian geophysical inversion, the data vector $\mathbf{d}$ is a constant. All PDFs with a model dependence are functions of the random variable $\mathbf{m}$. The term $p(\mathbf{d}|\mathbf{m})$ can then be interpreted as the model likelihood, the functional form of which depends on the statistics of the noise distribution, and the value of which depends on the model $\mathbf{m}$ being sampled and its misfit. For Gaussian noise, the model likelihood is given as:

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(- \frac{[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1}[\mathbf{d} - f(\mathbf{m})]}{2} \right). \tag{1.25}$$

Here $f(\mathbf{m})$ corresponds to the modeled data and $\mathbf{C}_d$ is the data covariance matrix and $[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1}[\mathbf{d} - f(\mathbf{m})]$ is the χ^2 misfit for the evaluated model $\mathbf{m}$. Thus knowledge of the data noise statistics allows us to transform calculated misfit into likelihood, which is a probability. The prior model distribution $p(\mathbf{m})$ represents our state of knowledge *independent* of the survey data. The evidence term $p(\mathbf{d})$ corresponds to a constant PDF normalizing factor equal to the integral over all possible models of the numerator in equation (1.23).

For the two parameter problem we will discuss in this section, all models are assigned equal prior probability and treating the evidence as a constant of proportionality, from equation (1.23) we can see that the posterior model PDF will be equal to the model likelihood PDF.

Just so that we know the 'answer' i.e., the posterior PDF, before carrying out MCMC sampling of the posterior, since there are only 2 parameters to search for, we carried out a grid search of the χ^2 misfits using $200{\times}200$ samples. The grid search extents were in the prior model space defined between 63 to 159 ohm-m and 2070 to 2150 m, as shown in Figure 1.6a. The truth is shown at 2100 m and 100 ohm-m with an encircled cross. Using equation (1.25) and the fact that the likelihood in this case equals the posterior distribution, we arrive at the posterior model PDF in Figure 1.6b. Using MCMC sampling, taking steps in the model space as shown in Figure 1.6c we are able to reproduce this posterior PDF as we have done in Figure 1.6d with 20,000 samples, half the number required for an adequate grid search.

Here we have used an MCMC sampler called the Metropolis-Hastings (MH) algorithm (Hastings, 1970; Metropolis et al., 1953) to sample the posterior model PDF. This is done using an acceptance probability α (e.g., Liang et al., 2011). At every step of the Markov Chain, a candidate model is sampled by perturbing the current model from a known distribution (the proposal distribution q) and the acceptance α is calculated. A random number r is then sampled uniformly from the interval [0,1]. If $r < \alpha$ the proposed perturbation is accepted, else the old model is retained. The rationale behind this algorithm can be explained by examining in more detail the expression for α where
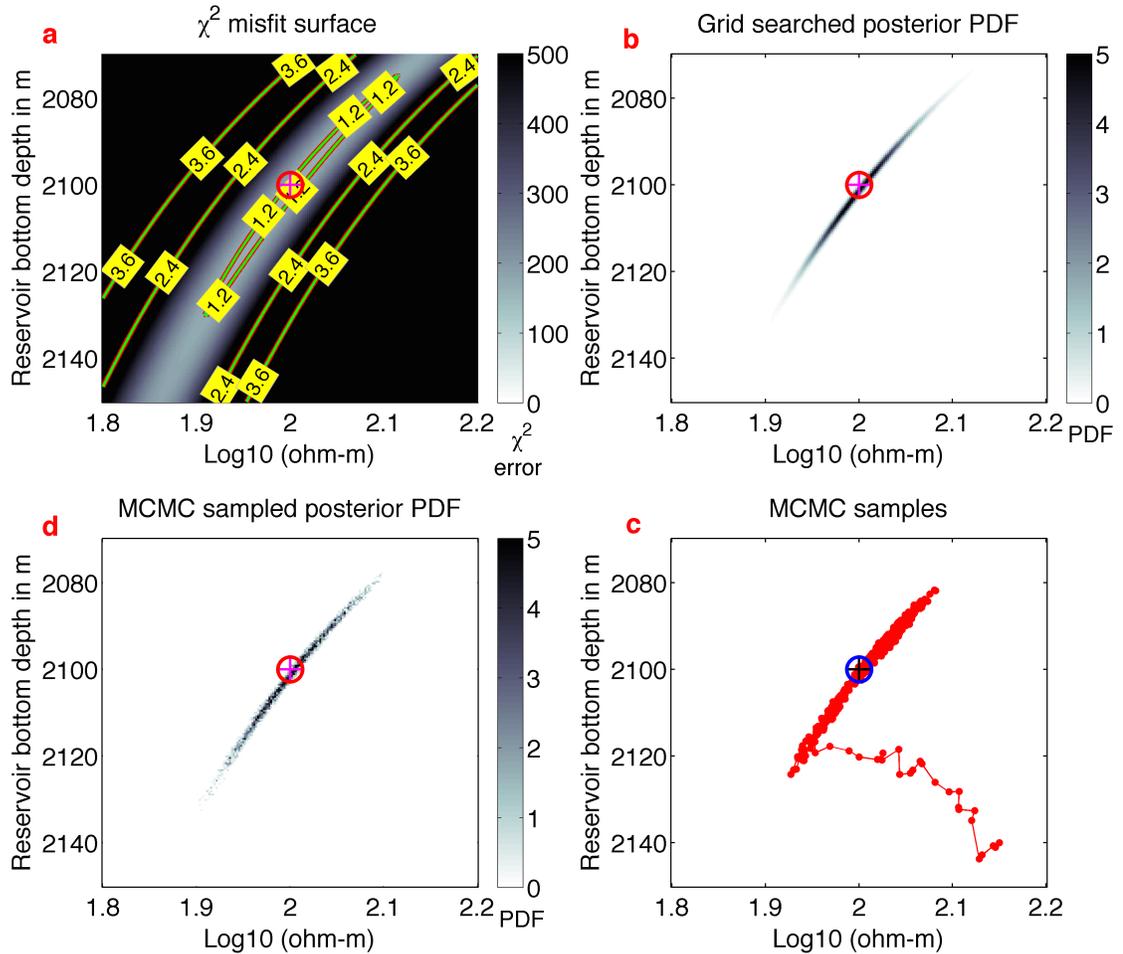
Figure 1.6: A toy 2 parameter CSEM problem with true model parameters at 2100 m and 100 ohm-m (encircled cross in all the plots), with 4% Gaussian noise added to the synthetic data. a) A grid searched χ^2 misfit surface. Contours of RMS misfit have been overlain on the misfit grid. b) The χ^2 misfit translated into likelihood which in this case is the posterior PDF, using equation (1.25). c) First 2000 steps taken by the Markov chain in sampling the model space. d) PDF of the scatterplot in c, but from all 20,000 samples barring the first 100 steps. Note how b) and d) are very similar, but have been arrived at by using different techniques.
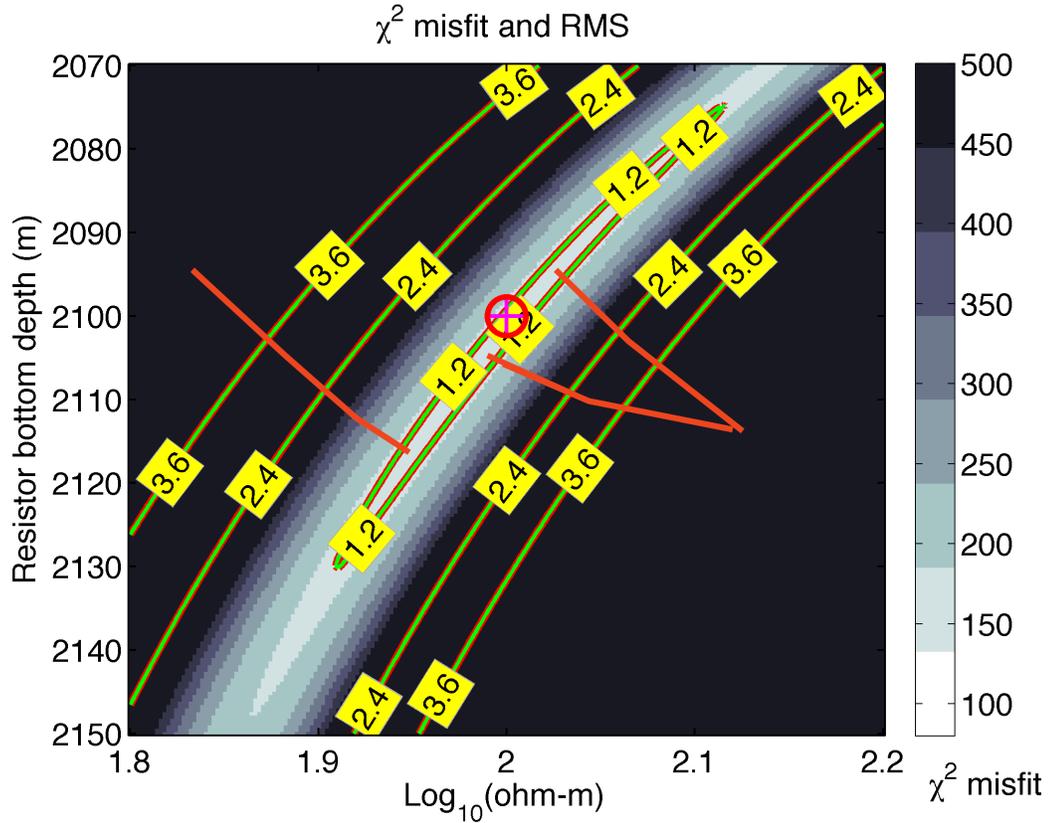
Figure 1.7: Three hypothetical gradient based inversion paths shown in red. Depending on the start model, or even with the same start model and different regularization choices, at the end they produce models that may be at very different locations in the model space. The true parameters as before, are indicated with an encircled cross.

$$\alpha(\mathbf{m}'|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]. \tag{1.26}$$

Here $\mathbf{m}'$ is the new proposed model and $\mathbf{m}$ is the old model (throughout this chapter, primes will denote new model values). Specifically, $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$ is the likelihood ratio and $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}$ is the proposal ratio. In a classic MH algorithm with a fixed number of dimensions, using uniform priors and symmetric proposals the acceptance term α in equation (1.26) is reduced to a ratio of likelihoods (Yardim et al., 2006; Bodin and Sambridge, 2009; Dettmer et al., 2010). Hence the algorithm always moves towards areas of higher posterior probability if the data misfit improves (likelihood ratio > 1). However, it can also move to areas of lower posterior probability with a probability α if the misfit

does not improve (likelihood ratio < 1). This is another key difference between MCMC sampling and optimization methods, both of which navigate a misfit / probability space – the optimization methods always follow the gradient of u in equation (1.3), producing a single model at the end (Figure 1.7) which may be highly dependent on start location or the choices made in model regularization.

1.4.2   Marginal PDFs

With larger numbers of model parameters, it is not possible to visualize their joint probability densities. In such cases, we are able to look at their marginal probabilities by integrating out the dependence on other variables. For instance, in this two parameter case, with the two variables z for bottom depth and ρ for reservoir resistivity, we can integrate out the distributions for either variable in the following manner:

$$p(z) = \int p(z, \rho)d\rho, \tag{1.27}$$

$$p(\rho) = \int p(z, \rho)dz. \tag{1.28}$$

The above marginalization also generalizes to n-dimensions. In MCMC sampling, the histogram of any sampled variable is the marginal distribution of that variable. The joint and marginal densities for the two parameter case are shown in Figure 1.8. Also shown in the plot for the joint PDF, are 70 randomly chosen1D models (in blue) from the posterior PDF. Since we are going to ultimately look at multiple parameter problems, where it will be impossible to visualize a histogram per parameter, we are going to switch to the use of color intensity to represent histograms of resistivity at different depths, as shown in Figure 1.9. In the left panel, we have binned all the vertical lines corresponding to model resistivities (the vertical blue lines show 70 such model resistivities in Figure 1.8) and colored the histogram heights according to 'color hotness' at every depth. To show variability in the placement of interfaces with resistivity contrasts, in the right panel of Figure 1.9, we have binned all the horizontal lines corresponding to interfaces in 1D models (again, the horizontal blue lines show 70 such interfaces in Figure 1.8). A delta function like spike is seen at 2000 m – the top depth of the reservoir was not allowed

Posterior PDF and its marginals
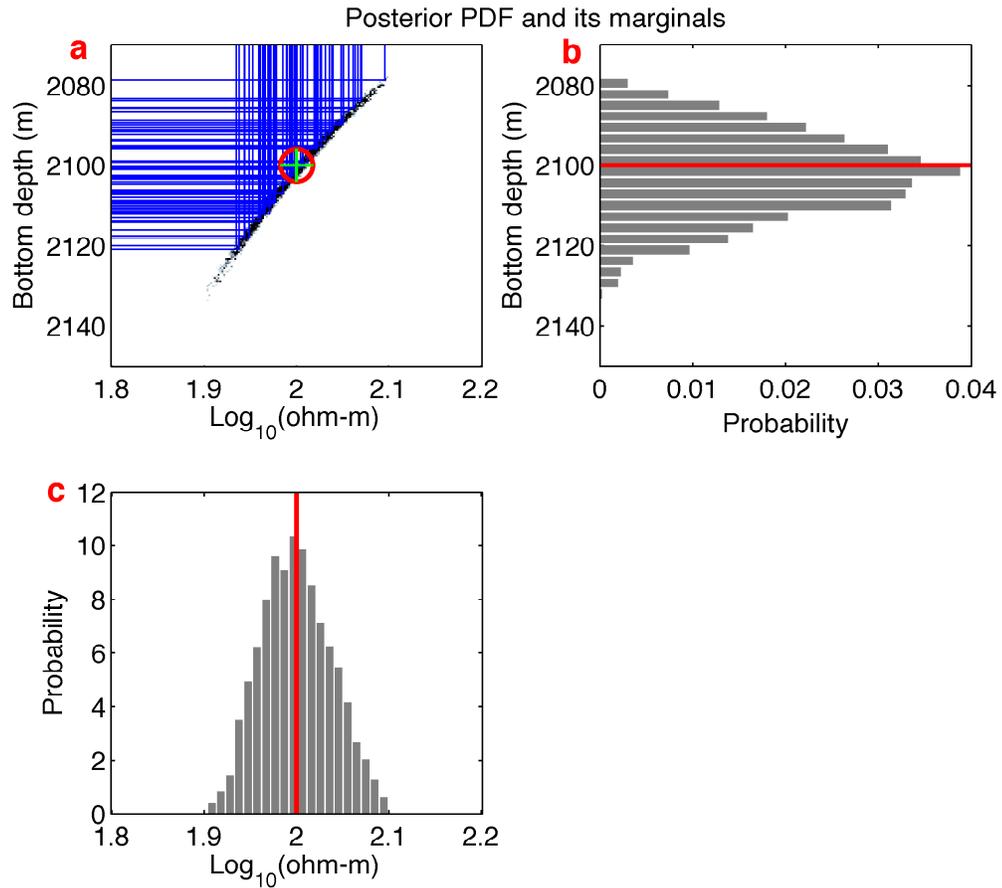
Figure 1.8: a) The blue lines show 70 random models associated with the posterior PDF. True value shown by an encircled cross. b) Binning the horizontal blue lines, we obtain a marginal PDF on the probability of placing the reservoir bottom depth. The true value is shown with a red line. c) Binning the vertical blue lines, we obtain a marginal PDF on the probability of the reservoir resistivity. Again, the true value is shown with a red line.
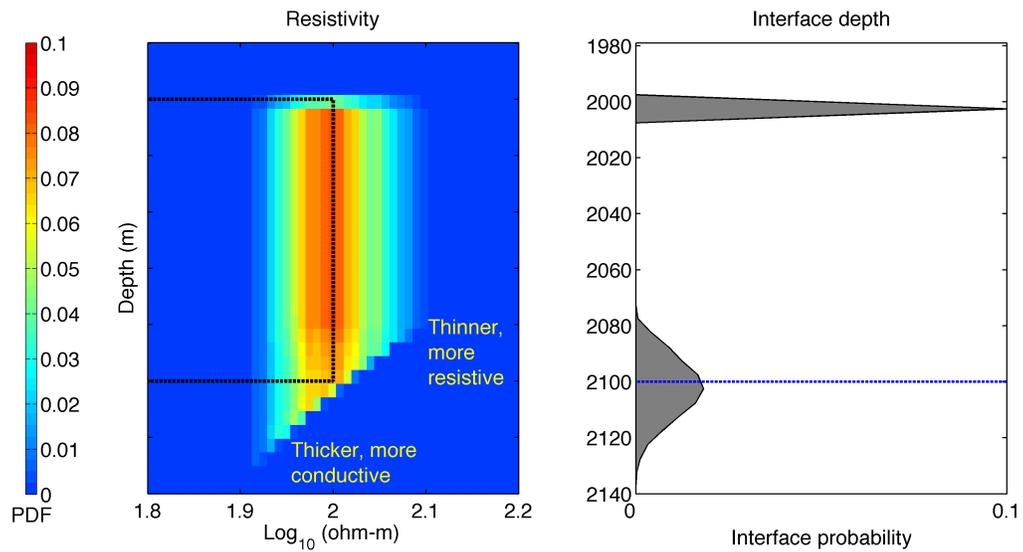
Figure 1.9: In the left pane, the model resistivity values sampled have been binned every 5 m and colored according to the height of the PDF within a particular depth range. Hotter colors are more probable. The true model is shown with a dashed black line. Note how the distributions of resistivity with depth are *not* very Gaussian below 2100 m. The right pane bins the locations of interfaces from the sampled models in 5 m intervals.

to vary. A Gaussian-like distribution is seen centered around 2100 m depth, indicating the uncertainty in placement of the reservoir bottom. The well known CSEM trade-off between thicker more conductive layers and thinner more resistive layers is clearly visible in the left panel. Note that we are now able to see how uncertain our solution to the problem is, without making any assumptions about the 'Gaussianity' of our solution, as would be required by a model covariance matrix approach (e.g., Menke, 2012). In fact, we see that the PDFs of resistivity with depth are not very Gaussian at depths below 2100 m in the left panel of Figure 1.9.

1.4.3   Monitoring the progress of sampling

We show the sampling of model parameters by the Markov chain in the first two rows of Figure 1.10. The third row shows the acceptance rate for sampling, which is a measure of the efficiency of the sampler. Too high or too low an acceptance rate implies inefficient sampling and will be discussed in detail later in Sections 2.2.8 and 4.A.6. A good rule of thumb for higher dimensional problems is to aim for about 23% for more than 6 parameters, based on considerations of convergence to a Langevin diffusion process (Chib and Greenberg, 1995; Roberts et al., 1997). The fourth row shows the range of sampled root mean square (RMS) misfit values. These are obtained by taking the χ^2 misfit for every model, dividing by the number of observed data points and taking square root.

We must be careful to avoid the 'burn-in' samples (Figure 1.11) that do not form part of a stationary Markov chain before forming our posterior model ensemble. These samples are usually located in the low probability (or high misfit) areas of the model space as shown in Figure 1.11a. Only the first 2000 samples have been shown in this figure for clarity, as the Markov chain keeps traveling up and down the 'probability crest' over the next 18,000 samples. Zooming in to the first 100 samples (Figure 1.11b), the true parameter values are shown with a dashed black line in the first two rows.
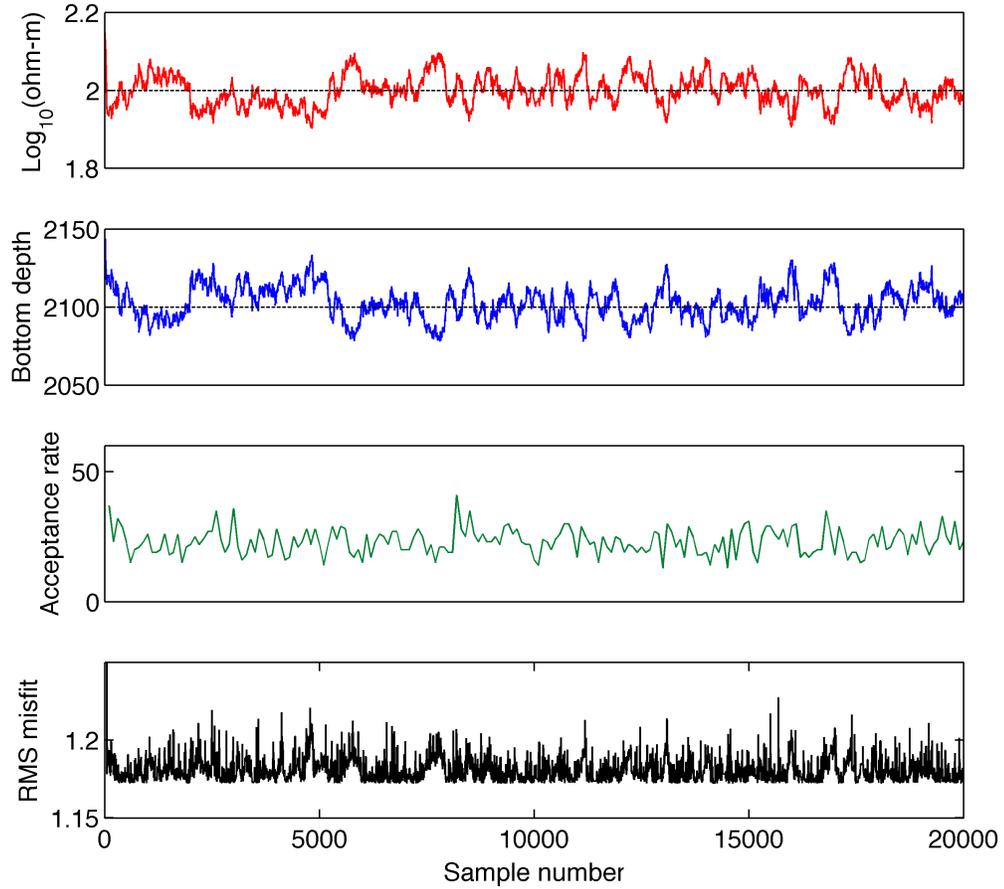
Figure 1.10: The top two rows show the sampled values for reservoir resistivity and its bottom depth, respectively. True values are shown with a dashed black line. The third row shows the percentage of accepted model space steps made by the Markov chain in every 100 steps. Too low or too high a value both indicate inefficient sampling. The last row shows the range of RMS misfits sampled by the Markov chain.
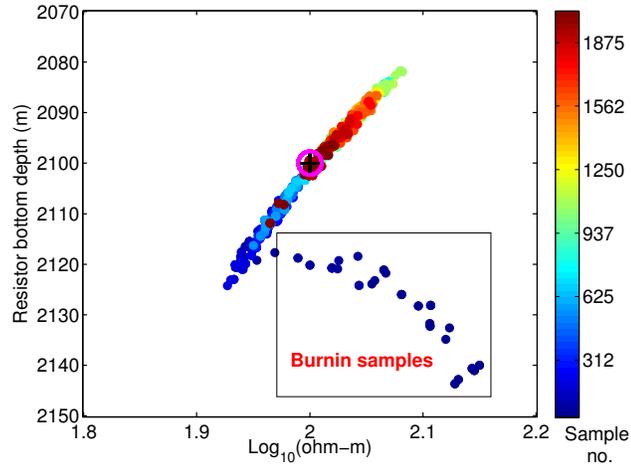
(a)
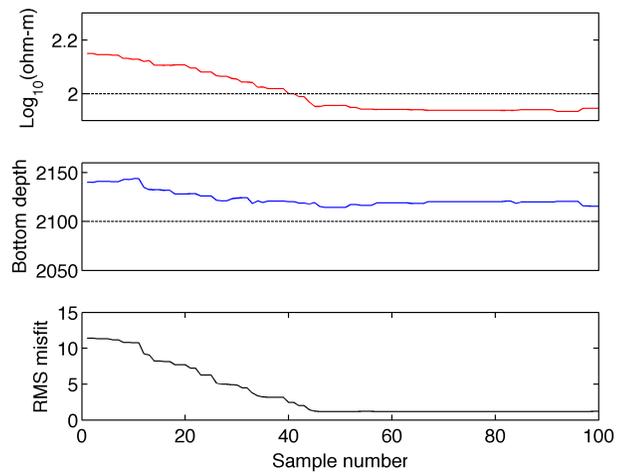
(b)

Figure 1.11: The 'burnin-in' period of an MCMC chain illustrated. a) The first 2000 MCMC samples colored in sequence. True value shown by an encircled cross. b) The same as Figure 1.10, but zoomed into the first 100 steps. True values are shown with a dashed black line in the first two rows. We can see that the first 50 or so samples are in high misfit, low probability regions of the model space. These are the 'burn-in' samlpes.
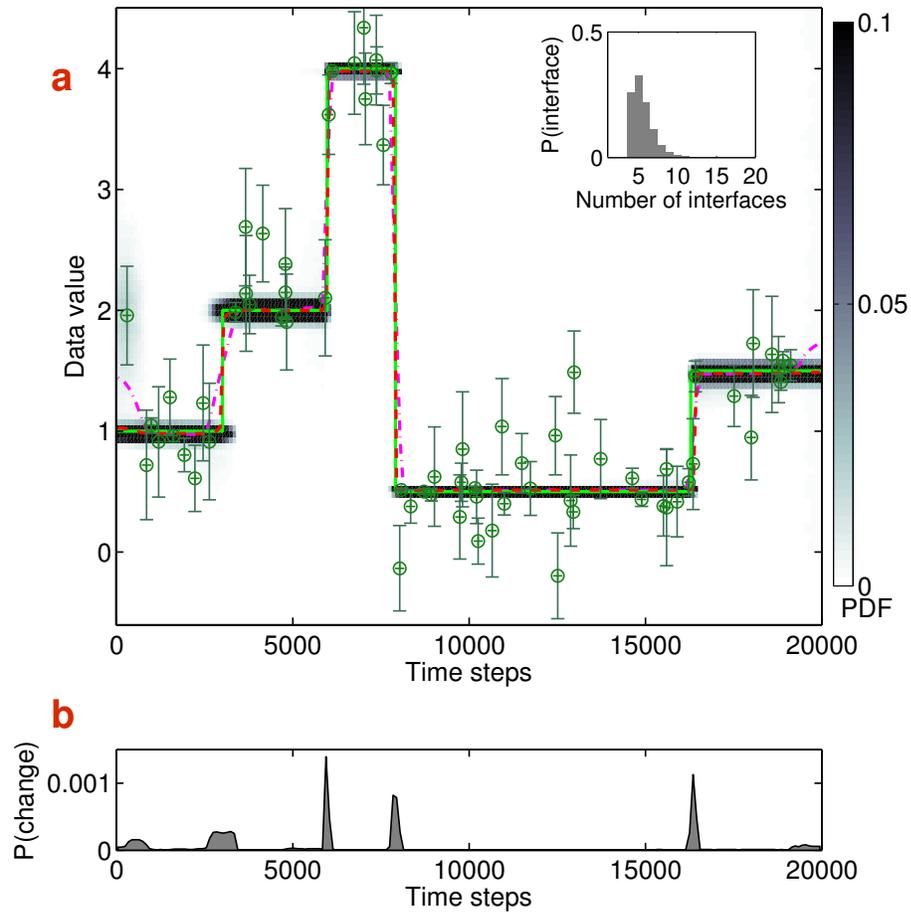
Figure 1.12: The change point identification problem for a noisy synthetic time series. a) The bright green line represents a 'true model' with abrupt 'change-points' in time. It is perturbed with Gaussian noise to create a synthetic 'time series' (dark green points). Using a partition modeling approach and Bayesian RJ-MCMC, we obtain a PDF on the time series that can fit the noisy data. Darker colors are more probable. The mean (dashed magenta) and median (dashed red) models are two statistics displayed from the ensemble solution. The inset box shows a PDF on the number of interfaces required to solve the problem. b) Posterior PDF of change points with time.

1.5 Dealing with uncertain parameterizations

All too often in geophysical analysis, we are faced with a situation in which we are not sure of how to parameterize the problem we are interested in. To be more specific, we know the parameters we are interested in, but we are not sure how many there are, or how these parameters are distributed in time or space. A prime example arises in studying events as seen in a geochemical record that is not evenly sampled. This time series analysis problem is analogous to what is known in statistical parlance as the 'change-point analysis problem' (CHIP) (e.g. Liang et al., 2011) and can be dealt with by using a partition modeling approach (e.g., Gallagher et al., 2011). In this approach, any model which explains the observed noisy data is divided into a number of partitions, the locations and numbers of which are determined by the data themselves. In a Bayesian framework, we will end up with an ensemble of partition models which explain the data. From this posterior ensemble, once again we can derive important information regarding the uncertainty about our model estimates. The CHIP problem can be effectively solved using the Reversible Jump Markov Chain Monte Carlo (RJ-MCMC) technique Green (1995) or the trans-dimensional MCMC method Sisson (2005). All the codes we have developed in this dissertation use a trans-dimensional Bayesian mechanism. Our solution to a CHIP problem for an irregularly sampled, noisy synthetic time series using Bayesian RJ-MCMC is shown in Figure 1.12a. The 'observed' data with 1σ error bars are shown in dark green. A posterior PDF on the time series that fit the data has been shown proportional to color. Darker colors correspond to higher probabilities. The number of partitions are allowed to vary between 2 and 61 (inset figure), when in truth there are 5. The dashed magenta line corresponds to the mean model and the red line corresponds to the median of the ensemble models at every time. In this case, the mean and median models are good estimates of the true model (bright green), and in Figure 1.12b we can see that the locations of the change points have been inferred well despite the noisy data.

If we turn Figure 1.12a clockwise by 90° then it begins to look quite similar to the left panel of Figure 1.9, albeit with a different color scale. Both plots are marginal PDFs of models with time or depth. Thus, we can look at the CSEM inverse problem, as another kind of problem suitable for treatment with the Bayesian RJ-MCMC or trans-dimensional method. This is all the more appropriate as we really do not know where in

the earth our layers or cells of resistivity are, or even how many there are. This problem is not unique to CSEM, and RJ-MCMC has been applied to various geophysical problems in the recent past. A detailed review of such work can be found in Sambridge et al. (2013).

In later chapters, we deal in detail with the issues of subjective model regularization and parameterization and how we can get around it using trans-dimensional Bayes'. For now, we demonstrate very simply with Figure 1.13 how the chosen parameterization affects our final posterior ensemble solution. The true model is an anisotropic CSEM resistivity model shown with a thick red line. In Figure 1.13a the subsurface has been exactly parameterized to have the true number of layers, three. In Figure 1.13b, the number of layers has been extended to 16. In Figure 1.13c, we have used the partition modeling approach and RJ-MCMC where the number of layers is variable. In all cases, darker colors in the left two panes correspond to higher probabilities on horizontal and vertical resistivity. The blue lines correspond to the 5% and 95% quantiles of resistivity with depth. The first two rows, with the exactly parameterized and over-parametrized solutions were derived using a fixed dimensional MCMC sampler. When the number of layers is known and fixed at the right positions (Figure 1.13a) – the problem does not seem to be that uncertain - all distributions of resistivity with depth are quite narrow. However, in geophysics, this level of prior knowledge about the subsurface is not available and we usually over-parameterize the problem (Figure 1.13b). However, now the distributions of resistivity with depth are far wider and the model uncertainties seem far larger. This is one of the reasons why over-parameterized gradient based inversions use regularization to smooth a deterministic solution. If that were not the case, any individual solution that fits the data from within the posterior ensemble shown in Figure 1.13b may be unrealistically oscillatory (Constable et al., 1987). Finally, when we allow the number of resistivity interfaces and their locations to vary between 1 and 30, we obtain the uncertainties shown in Figure 1.13c. In this case, we have let trans-dimensional Bayes' determine the layer parameterization from the data itself. We have sampled over a number of parameterizations, as can be seen in Figure 1.14.

Finally, we mention here that trans-dimensional Bayes' follows a 'parsimony' rule whereby models are preferred neither to be too simple nor too complicated. This
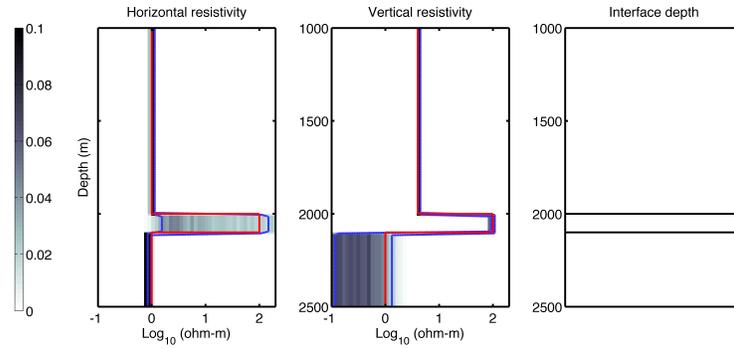
(a) Exactly parameterized
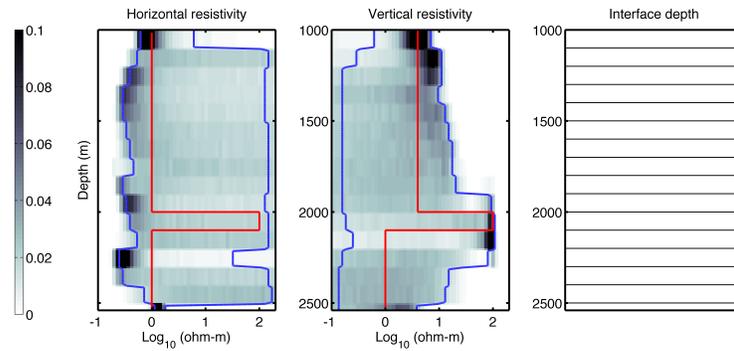
(b) Over-parameterized
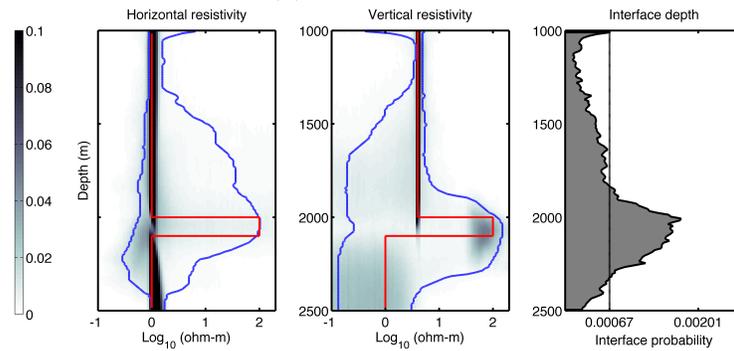
(c) Trans-dimensionally parameterized

Figure 1.13: MCMC inversion results for a synthetic CSEM resistivity model (red). Darker colors correspond to higher probabilities. The 5% and 95% quantiles on resistivity with depth are shown in blue. A more realistic view of uncertainty is obtained when we allow the number of resistivity contrasts to be variable as in c).
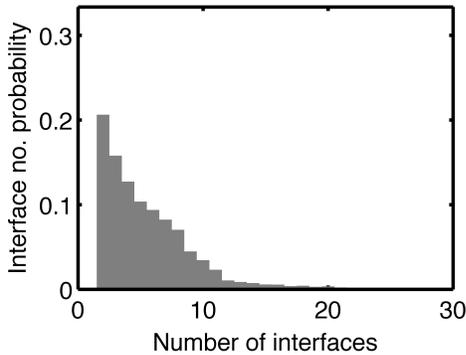
Figure 1.14: The number of interfaces required to likely fit the CSEM data for the trans-dimensional inversion result shown in Figure 1.13c

is ensured by a synergy between adding more model parameters and deleting them. For instance, if two spatially adjacent resistivities are similar, the algorithm will merge them into one resistivity if the resulting data likelihood is reasonable. Conversely, if splitting a resistive layer in two will likely fit the observed CSEM data, the two layers are encouraged to have very different resistivities for the split to be acceptable. The mathematical formulation of this idea is fully developed in Sections 4.A.4 and 4.A.5.

1.6 Full circle: connection between data and models

Finally, at this juncture, we would like to show the spread in data fit from the sampled models in the RJ-MCMC posterior ensemble corresponding to Figure 1.13c. The data fits from 100 randomly chosen posterior models are shown in Figure 1.15. Most of these models fit the data well within the error bars, with the exception of a few, which are required to form the tails of the posterior distribution.

Acknowledgments

This chapter has been influenced in no small fashion by discussions with Peter Gerstoft, Thomas Bodin and Kerry Key.
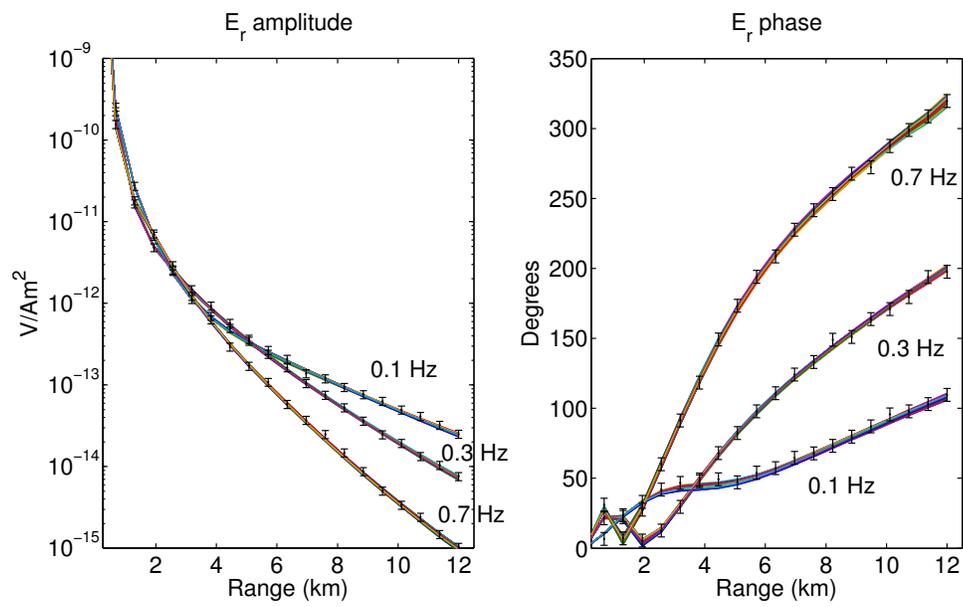
Figure 1.15: CSEM radial electric field E_r data fit from 100 randomly selected models in the posterior ensemble corresponding to the inversion results shown in Figure 1.13c.

References

Abubakar, A., Habashy, T. M., Druskin, V. L., Knizhnerman, L., and Alumbaugh, D., 2008: 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements. *Geophysics*, **73**(4), F165–F177.

Arfken, G. B., Weber, H. J., and Harris, F. E., 2011: *Mathematical methods for physicists: a comprehensive guide.* Academic press.

Aster, R. C., H Thurber, C., and Borchers, B., 2005: *Parameter estimation and inverse problems.* Elsevier, Burlington.

Backus, G. E., 1988: Bayesian inference in geomagnetism. *Geophysical Journal International*, **92**(1), 125–142.

Bayes, T., and Price, R., 1763: An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions*, **53**, 370–418. doi:10.1098/rstl.1763.0053.

Bodin, T., and Sambridge, M., 2009: Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, **178**(3), 1411–1436.

Chave, A. D., and Cox, C. S., 1982: Controlled electromagnetic sources for measuring electrical conductivity beneath the oceans, 1. Forward problem and model study. *Journal of Geophysical Research*, **87**(B7), 5327–5338.

Chib, S., and Greenberg, E., 1995: Understanding the metropolis-hastings algorithm. *American Statistician*, 327–335.

Constable, S., 2010: Ten years of marine CSEM for hydrocarbon exploration. *Geophysics*, **75**(5), 75A67–75A81.

Constable, S. C., Parker, R. L., and Constable, C. G., 1987: Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(3), 289–300.

Dettmer, J., Dosso, S. E., and Holland, C. W., 2010: Trans-dimensional geoacoustic inversion. *The Journal of the Acoustical Society of America*, **128**(6), 3393–3405.

Evans, R. L., 1991: *Electrical Resistivity Structure of the East Pacific Rise Near 13 N.* Ph.d. thesis, University of Cambridge.

Gallagher, K., Bodin, T., Sambridge, M., Weiss, D., Kylander, M., and Large, D., 2011: Inference of abrupt changes in noisy geochemical records using transdimensional changepoint models. *Earth and Planetary Science Letters*, **311**(1-2), 182–194. ISSN 0012821X. doi:10.1016/j.epsl.2011.09.015.

Green, P. J., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.

Griffiths, D. J., 1999: Introduction to electrodynamics.

Hastings, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.

Key, K., 2009: 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, **74**(2), F9.

Key, K., 2012: Marine electromagnetic studies of seafloor resources and tectonics. *Surveys In Geophysics*, **33**(1), 135–167.

Liang, F., Liu, C., and Carroll, R., 2011: *Advanced Markov chain Monte Carlo methods: learning from past samples*, volume 714. Wiley. com.

Loseth, L. O., 2007: *Modelling of Controlled Source Electromagnetic Data*. Ph.d. thesis, Norwegian University of Science and Technology.

MacGregor, L., and Sinha, M., 2000: Use of marine controlled-source electromagnetic sounding for sub-basalt exploration. *Geophysical Prospecting*, **48**(6), 1091–1106.

MacGregor, L. M., 1997: *Electromagnetic investigation of the Reykjanes Ridge near 58 North*. Ph.d. thesis, University of Cambridge.

MacKay, D. J. C., 2003: *Information theory, inference and learning algorithms*. Cambridge University Press.

Mackie, R. L., and Madden, T. R., 1993: Three-dimensional magnetotelluric inversion using conjugate gradients. *Geophysical Journal International*, 215–229.

Menke, W., 2012: *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., 1953: Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, **21**(6), 1087–1092.

Mittet, R., and Gabrielsen, P. T., 2013: Decomposition in upgoing and downgoing fields and inversion of marine CSEM data. *Geophysics*, **78**(1), E1–E17. ISSN 0016-8033. doi:10.1190/geo2011-0520.1.

Newman, G. A., and Alumbaugh, D. L., 2000: Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, **140**(2), 410–424.

Parker, R. L., 1983: The magnetotelluric inverse problem. *Geophysical Surveys*, **6**(1-2), 5–25.

Parker, R. L., 1994: *Geophysical Inverse Theory*. Princeton University Press.

Ray, A., Alumbaugh, D. L., Hoversten, G. M., and Key, K., 2013: Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering. *Geophysics*, **78**(6), E271–E280.

Roberts, G. O., A., G., and Gilks, W. R., 1997: Weak convergence and optimal scaling of random walk Metropolis algorithms. *Annals of Applied Probability*, **7**(1), 110–120.

Rodi, W., and Mackie, R. L., 2001: Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion. *Geophysics*, **66**(1), 174–187.

Sambridge, M., Bodin, T., Gallagher, K., and Tkalčić, H., 2013: Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A*, (December 2012).

Sasaki, Y., 2013: 3D inversion of marine CSEM and MT data : An approach to shallow-water problem. *Geophysics*, **78**(1), E59–E65.

Scales, J. A., and Sneider, R., 1997: To Bayes or not to Bayes? *Geophysics*, **62**(4), 1045–1046.

Sisson, S. a., 2005: Transdimensional Markov Chains. *Journal of the American Statistical Association*, **100**(471), 1077–1089. ISSN 0162-1459. doi:10.1198/016214505000000664.

Sneider, R., and Trampert, J., 2000: Linear and nonlinear inverse problems. In *Geomatic methods for the analysis of data in the earth sciences*, editors A. Dermanis, A. Grun, and F. Sanso, 93–164. Springer, Berlin.

Tarantola, A., 2005: *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadephia.

Tarantola, A., and Valette, B., 1982: Inverse problems= quest for information. *Journal of Geophysics*, **50**, 159–170.

Ward, S. H., and Hohmann, G. W., 1987: Electromagnetic theory for geophysical applications. In *Electromagnetic Methods in Applied Geophysics*, editor M. N. Nabighian, 131–312. Soc. Exploration Geophysicists, Tulsa.

Weiss, C., 2007: The fallacy of the "shallow-water problem" in marine CSEM exploration. *Geophysics*, **72**(6), A93–A97.

Wheelock, B. D., 2012: *Electromagnetic Imaging of the Crust and Upper Mantle across the Continental Margin in Central California*. Ph.D. thesis, University of California, San Diego.

Yardim, C., Gerstoft, P., and Hodgkiss, W. S., 2006: Estimation of Radio Refractivity From Radar Clutter Using Bayesian Monte Carlo Analysis. *IEEE Transactions on Antennas and Propagation*, **54**(4), 1318–1327.

Chapter 2

Bayesian inversion of marine CSEM data with a trans-dimensional self parameterizing algorithm

Anandaroop Ray and Kerry Key
Scripps Institution of Oceanography, La Jolla, CA, USA

Abstract. The posterior distribution of earth models that fit observed geophysical data convey information on the uncertainty with which they are resolved. From another perspective, the non-uniqueness inherent in most geophysical inverse problems of interest can be quantified by examining the posterior model distribution converged upon by a Bayesian inversion. In this work we apply a reversible jump Markov chain Monte Carlo method to sample the posterior model distribution for the anisotropic 1-D seafloor conductivity constrained by marine controlled source electromagnetic data. Unlike conventional gradient based inversion approaches, our algorithm does not require

any subjective choice of regularization parameter, and it is self parameterizing and trans-dimensional in that the number of interfaces with a resistivity contrast at depth is vari-able, as are their positions. A synthetic example demonstrates how the algorithm can be used to appraise the resolution capabilities of various electromagnetic field components for mapping a thin resistive reservoir buried beneath anisotropic conductive sediments. A second example applies the method to survey data collected over the Pluto gas field on the Northwest Australian shelf. A benefit of our Bayesian approach is that subsets of the posterior model probabilities can be selected to test various hypotheses about the model structure, without requiring further inversions. As examples, the subset of model probabilities can be viewed for models only containing a certain number of layers, or for models where resistive layers are present between a certain interval as suggested by other geological constraints such as seismic stratigraphy or nearby well logs.

2.1 Introduction

The ability of marine controlled-source electromagnetic (CSEM) data to detect thin resistive layers such as hydrocarbon formations trapped in conductive sediments is now well-known due to more than a decade of research conducted in conjunction with the commercialization of CSEM technologies (e.g., Ellingsrud et al., 2002; Um and Alumbaugh, 2007; Constable, 2010). Given a CSEM data set, the inverse problem is to find a resistivity model compatible with the data. This non-linear ill-posed inverse problem can be tackled in a number of ways. The traditional approaches used during the past few decades for various EM geophysics applications have relied on regularization methods that serve to stabilize the problem, usually through the use of norms that penalize the model roughness or deviations from a preferred model (e.g., Constable et al., 1987; Newman and Alumbaugh, 2000; Abubakar et al., 2008). Though highly efficient and well understood, these gradient based optimization methods only provide a single smooth model as a result, or provide a suite of smooth models as a function of data misfit. While such results can be considered robust in the sense that the smoothness constraint tends to eliminate unnecessary or spurious features and retains only those smooth features required to fit the data, these methods have a very narrow view of the full range of models compatible with the data. Furthermore, questions about the resolution of the resulting inverse models remain unanswered.

A solution to this problem is to utilize Bayesian methods to characterize model resolution in terms of marginal probability density functions for each model parameter (e.g., Sen and Stoffa, 1995). Since Bayesian probability is a statement of information, it is natural to postulate geophysical inversion in a Bayesian framework. Inversion thus implies sampling a posterior model distribution, which depends on available prior information on likely models and the likelihood of the data fit for sampled models. The posterior distribution of model parameters conveys information on the uncertainty with which they are resolved. From another perspective, the non-uniqueness inherent in most geophysical problems of interest can be quantified by examining the posterior model distribution.

Our study is not the first to apply Bayesian methods for 1-D inversion of marine CSEM problems. One of the earliest applications focused on joint inversion of

CSEM and seismic data in order to improve estimates of reservoir properties (Hou et al., 2006; Chen et al., 2007). Trainor-Guitton and Hoversten (2011) use a sampling scheme which involves both the Metropolis-Hastings algorithm and slice sampling (Neal, 2003) in order to improve convergence upon the distribution of solution models. Buland and Kolbjornsen (2012) apply the Metropolis-Hastings algorithm (Hastings, 1970) to invert marine CSEM data together with magnetotelluric (MT) data in order to constrain the range of likely resistivities as a function of depth.

While these inversions have indeed been probabilistic, they have not addressed the issue of model selection, or scalability of the model space. That is to say, either the number of layers and their depths have been fixed a priori using only a few parameters, or the number of depth intervals is very large in order to allow for more structural detail (requiring considerably more computational effort). Having too few or too many parameters will bias the inversion results (Dettmer et al., 2010). Tompkins et al. (2011) address the issue of model space scalability for the marine CSEM problem by using either a principle component analysis approach or singular value decomposition to transform their model domain. This may result in a more efficient search since the model space is greatly reduced (Collins and Fishman, 1995), and has similarly been applied to geo-acoustic inversion (Dosso and Dettmer, 2011). However, the approach of Tompkins et al. (2011) is not Bayesian and requires a starting regularization as well as subjective layer parameterization. Gunning et al. (2010) developed a 'Bayesianized' parametric bootstrap method that splits layers to create resolution at depths statistically justified by the CSEM data.

Our approach is to essentially use nothing more complicated than Bayes' theorem (Bayes, 1763) itself, and the idea of 'Bayesian parsimony' which states that the evidence term in Bayes' theorem prefers models with simpler model parameterizations (Dettmer et al., 2010). To this end we apply the reversible jump Markov Chain Monte Carlo (RJ-MCMC) method to characterize model uncertainty for marine CSEM data. This approach was first introduced to the statistics community for change point analysis by Green (1995) and deals with a variable number of unknowns, which in our case are the number and positions of the model layers and their resistivities. The RJ-MCMC method has found widespread use in a diverse number of fields from genetics (Huelsenbeck

et al., 2004) to automated surveillance (Smith et al., 2005). In the geophysical realm, Malinverno (2002) has applied it to DC resistivity data inversion and more recently Agostinetti and Malinverno (2010) and Bodin et al. (2012) have used it for seismic receiver function analysis. A recent application to airborne EM can be found in Minsley (2011). Reversible jump MCMC methods are often said to be 'trans-dimensional', where trans-dimensionality implies that the number of model parameters is variable. The algorithm we follow has been used for the geo-acoustic problem (Dettmer et al., 2010) and seismic surface wave tomography (Bodin and Sambridge, 2009), using the 'partition modelling' approach outlined by Bodin and Sambridge (2009).

The manner in which this algorithm avoids regularization yet provides reasonable solutions is elegantly simple. The results of the inversion, as in any Bayesian framework, are not limited to a single 'best-fit' solution. Instead, the algorithm provides a distribution of model solutions, the large majority of which fit the data to within its uncertainty. The non-uniqueness inherent in most geophysical inverse problems due to both the physics of the problem and a finite set of noisy observations implies that there are an infinite number of models that fit the data. However, some parts of the model space are more frequently required by models that fit the data than other parts of the model space. It is these parts of the model space which are deemed to be more probable (Backus, 1988; Tarantola, 2005). Since the model space is often high-dimensional, the ensemble of models can be conveniently viewed via marginal distributions of the parameters of interest (e.g., Sen and Stoffa, 1995).

Here we present the theory for a trans-dimensional RJ-MCMC with an adaptation that mixes local and global steps in the model space to ensure thorough sampling of the model space. We then proceed with two example applications for marine CSEM inversion. The first is a resolution test for resistivity anisotropy in a model of a thin resistive layer representing an offshore hydrocarbon reservoir and the second is an application to CSEM survey data collected over the Pluto gas field in the Northwest Shelf of Australia.

2.2 Theory and Algorithm

2.2.1 Anisotropic forward modelling

The analytic expressions for the electromagnetic (EM) fields in media that are transversely isotropic about the vertical (TIV) due to electric and magnetic dipole transmitters are given in Appendix F of Løseth and Ursin (2007). As an illustrative example, we provide here the integral transform for the radial electric field due to a horizontal electric dipole (HED) at an azimuth β and range r from the dipole, normalized by the source dipole moment:

$$E_r = -\frac{1}{4\pi}\cos\beta\left[\mathcal{I}_0^{TM} + \frac{1}{r}(\mathcal{I}_1^{TE} - \mathcal{I}_1^{TM})\right], \tag{2.1}$$

where

$$\mathcal{I}_0^{TM} = \int_0^\infty dk_r k_r J_0(k_r r) g^{TM}(p_{zII}), \tag{2.2}$$

$$\mathcal{I}_1^{TE} = \int_0^\infty dk_r J_1(k_r r) g^{TE}(p_{zI}), \tag{2.3}$$

$$\mathcal{I}_1^{TM} = \int_0^\infty dk_r J_1(k_r r) g^{TM}(p_{zII}). \tag{2.4}$$

In the above equations, k_r is the radial wavenumber, J_0 and J_1 are Bessel functions of order 0 and 1, respectively, and the integrations are over all possible wavenumbers k_r. The functions $g^{TE}(p_{zI})$ and $g^{TM}(p_{zII})$, which are functions of k_r, contain information about the seabed resistivity structure in the form of the generalized transverse electric (TE) and transverse magnetic (TM) reflection coefficients of a stack of conducting layers, including the air, sea and seafloor sediments. Full expressions for these functions can be found in Løseth and Ursin (2007) or derived in a slightly different form in Chew (1995).

For the case of TIV anisotropy, from equations 119 in Løseth and Ursin (2007), it can be shown that within a given layer

$$p_{zI}^2 = \mu\tilde{\varepsilon}_h - \frac{k_r^2}{\omega^2}, \tag{2.5}$$

$$p_{zII}^2 = \mu\tilde{\varepsilon}_h - \frac{\tilde{\varepsilon}_h}{\tilde{\varepsilon}_v}\frac{k_r^2}{\omega^2}, \tag{2.6}$$

where $\tilde{\varepsilon} = \varepsilon + \frac{i\sigma}{\omega}$ is the complex permittivity with the corresponding conductivity σ_h or σ_v for the horizontal or vertical directions within the layer. The terms ε and μ are the permittivity and permeability of free space, respectively. The term p_{zI} is the vertical 'slowness' (vertical wavenumber divided by the angular frequency ω) associated purely with the TE mode while the term p_{zII} is associated only with the TM mode. Thus, from equations (2.5) and (2.6) it is apparent that the TE mode depends only on the horizontal conductivity while the TM mode depends on both the horizontal and vertical conductivity (e.g., Ramananjaona et al., 2011); this is expected given that the TE mode in 1-D layered structures is associated only with horizontal currents while the TM mode is associated with both horizontal and vertical currents.

The complete suite of fundamental dipole sources also includes the vertical electric dipole (VED), the horizontal magnetic dipole (HMD) and the vertical magnetic dipole (VMD). Each source produces particular modes in the EM field components, as listed in Table 2.1. The specific integral expressions each field component and source type can be found in Løseth and Ursin (2007). For both the HED and HMD, the expressions for the radial electric field E_r, azimuthal electric field E_β, the radial magnetic field H_r and the azimuthal magnetic field H_β are mixtures of TE and TM modes, thus they contain sensitivity to both the horizontal and vertical conductivity of the seabed. Likewise, the three field components produced by a VED are entirely TM mode fields, which equation 2.6 shows to contain sensitivity to the conductivity anisotropy. Conversely, the E_β, H_r and H_z components produced by a VMD are only TE mode fields and therefore are only sensitive to horizontal conductivity. Similarly, the H_z produced from either a HED or HMD is TE mode only and hence it is completely blind to the vertical conductivity.

While our understanding of the CSEM sensitivity to anisotropy is aided by knowledge of the modes present in each component, this knowledge does not address how well each component can *resolve* anisotropic conductivity. In an earlier study for isotropic conductivity, smooth inversions of synthetic data from each field component were used to identify which components can best recover a thin resistive layer representing a hydrocarbon reservoir, finding the HED transmitter to be overall superior for recovering the resistive layer as compared to the VED (Key, 2009). While the smooth inversions

Table 2.1: Modes present in the various EM field components produced by each of the four fundamental electric and magnetic dipoles sources when embedded in 1-D layered media.

Component	HED	VED	HMD	VMD
E_r	TE,TM	TM	TE,TM	
E_β	TE,TM		TE,TM	TE
E_z	TM	TM	TM	
H_r	TE,TM		TE,TM	TE
H_β	TE,TM	TM	TE,TM	
H_z	TE		TE	TE

were adequate for this purpose, they only produce a single conductivity model and offer no insights on the full range of models that might fit a given data set. Later in this work, we expand on this study by carrying out a Bayesian uncertainty analysis to characterize the resolution possible with each field component and also extend the model to include a modest amount of anisotropy in the overburden.

While either of the HED and HMD sources might be best for resolving anisotropy since they have the richest combination of modes, for practical reasons most marine CSEM surveys use a HED source. Very long HED antennas (100-500 m) can be deep-towed just above the seabed, creating a large source dipole moment that offers good signal-to-noise ratio data to long source-receiver offsets, whereas practically achievable HMD dipole moments result in much lower amplitude electric and magnetic fields and hence lower signal-to-noise ratios and a more limited range of usable source-receiver offsets. Hence the resolution study we present will focus only on the HED source.

2.2.2 Bayesian inversion

In the most general sense, the goal of geophysical studies is to provide information about the subsurface geology. Since Bayesian probability is a statement of information (e.g., Scales and Snieder, 1997), a Bayesian framework for geophysical inversion seems natural to postulate. We begin with Bayes' Theorem,

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \tag{2.7}$$

where **d** is the vector of observed data and **m** is a vector of model parameters. However, in a Bayesian approach, all information is contained in probability density functions (PDFs) represented by the notation $p(.)$. The idea is to estimate the posterior model distribution $p(\mathbf{m}|\mathbf{d})$ (i.e., the probability of the model given the data) starting from the knowledge of a prior model distribution $p(\mathbf{m})$ given the observed data **d**. The information contained in the data is inserted into Bayes' theorem using the likelihood function $p(\mathbf{d}|\mathbf{m})$ which is the probability of the data given the model, (essentially the PDF of the data misfit as a function of **m**). Assuming the noise in the data is Gaussian, we have the following likelihood function:

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(- \frac{[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1}[\mathbf{d} - f(\mathbf{m})]}{2} \right), \tag{2.8}$$

where $f(\mathbf{m})$ is the predicted data due to model **m** and $\mathbf{C_d}$ is the data covariance matrix, assumed here to be a diagonal matrix of data noise variances. The likelihood function in equation (2.8) is formulated such that models with lower misfit demand a higher likelihood while poorly fitting models have less likelihood.

The evidence term $p(\mathbf{d})$ is equal to an integration of the numerator in the RHS of (2.7) over all possible models, and is therefore only a normalizing constant to ensure that $p(\mathbf{m}|\mathbf{d})$ integrates to 1; consequently it can be ignored, providing the following relationship:

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m})p(\mathbf{m}). \tag{2.9}$$

Thus it is now a matter of sampling the right hand side of equation (2.9) to obtain the solution to the inverse problem and obtain the posterior model distribution. However, it is not possible to exhaustively sample the model space for more than a few parameters, hence in practice one resorts to efficient sampling methods like the Markov chain Monte Carlo (MCMC) method (e.g., Liang et al., 2010).

2.2.3   Reversible Jump MCMC

The Reversible Jump MCMC method (RJ-MCMC) is a particular type of Metropolis-Hastings (MH) algorithm (e.g., Liang et al., 2010) that allows one to search the model space such that models are not restricted in the number of parameters they

may posses. Thus, they are said to be 'trans-dimensional' in the number of model parameters. In the context of the CSEM problem, the number of resistivity contrasts in the seabed is usually unknown a priori and should also be considered a part of the inverse problem. Malinverno (2002) considered this aspect of 'model selection' for the DC resistivity problem and Minsley (2011) applied this idea to the airborne frequency-domain EM problem. Formulating the inverse problem in a trans-dimensional manner allows for a broader characterization of the posterior model distribution, as demonstrated by Dettmer et al. (2010) for the geo-acoustic problem and Bodin and Sambridge (2009) for seismic tomography. There is a natural 'parsimony' (Malinverno, 2002) built into the Bayesian RJ-MCMC framework. Simple models that fit the data well by using a few parameters are favoured over models with a larger number of parameters; this aspect of the RJ-MCMC Bayesian approach is similar to Occam's inversion, in which a penalty against model roughness is used find a smooth model that only contains features required to fit the data (Constable et al., 1987). However, the trans-dimensional algorithm does not require any regularization or smoothing parameters, which are usually subjective choices for gradient based inversion methods. In short, a Bayesian RJ-MCMC framework is a truly data driven inversion scheme that parameterizes itself as demanded by the data.

The remainder of this section describes the details of the RJ-MCMC approach for the interested reader, but one could without loss of continuity skip to the example applications given in section 3.

2.2.4   Prior information and model parameterization

In a pure Bayesian framework, prior information encapsulates knowledge about the model which is independent of the data. In order to be as general as is physically feasible, we use bounded uniform priors for all the model parameters of interest. In this section we closely follow the prior parameterization approach of Bodin and Sambridge (2009) but use interfaces (Dettmer et al., 2010) instead of Voronoi cells. We parameterize each model $\mathbf{m}$ as having n interfaces at depths z, and therefore $n+1$ resistivities each for the horizontal resistivity ρ_h and for the vertical resistivity ρ_v, where resistivity $\rho = \sigma^{-1}$.

The number of layers is fixed for a given model $\mathbf{m}$, hence

$$p(\mathbf{m}) = p(\mathbf{m}|n)p(n). \tag{2.10}$$

We use a uniform prior on n, given by

$$p(n) = \begin{cases} \frac{1}{n_{max}-n_{min}} & \text{if } n_{min} \le n \le n_{max} \\ 0 & \text{else} \end{cases}. \tag{2.11}$$

We assume no *a priori* knowledge between the depth of interfaces in the model and the layer resistivities (also considered independent), therefore their PDFs can be separated in the following product form,

$$p(\mathbf{m}|n) = p(\mathbf{z}|n)p(\boldsymbol{\rho}_h|n)p(\boldsymbol{\rho}_v|n). \tag{2.12}$$

Assuming interfaces can be located anywhere in the subsurface between z_{max} and z_{min} such that an interface can be at any of N (temporarily discrete) depths, for n interfaces, we can arrange them without paying attention to their ordering in $\frac{N!}{n!(N-n)!}$ ways. Note that this unspecified variable N will cancel out of the expressions we need to use in the algorithm and is only introduced for ease of mathematical derivation. Thus,

$$p(\mathbf{z}|n) = \begin{cases} \left[\frac{N!}{n!(N-n)!}\right]^{-1} & \text{if } z_{min} \le z \le z_{max} \\ 0 & \text{else} \end{cases}. \tag{2.13}$$

Assuming that all $n+1$ resistivities, ρ_h and ρ_v lie uniformly between ρ_{min} and ρ_{max}, independent of each other, we write

$$p(\boldsymbol{\rho}_h|n)p(\boldsymbol{\rho}_v|n) = \begin{cases} \left[\frac{1}{\rho_{max}-\rho_{min}}\right]^{2(n+1)} & \text{if } \rho_{min} \le \rho \le \rho_{max} \\ 0 & \text{else} \end{cases}. \tag{2.14}$$

To obtain the explicit expression for the prior model probability, we write $\Delta\rho = \rho_{max} - \rho_{min}$ and $\Delta n = n_{max} - n_{min}$ and substitute equations (2.11) – (2.14) into (2.10) to get

$$p(\mathbf{m}) = \begin{cases} \frac{n!(N-n)!}{N!\Delta n(\Delta\rho)^{2n+2}} & \text{if } z \in [z_{min}, z_{max}] \text{ and } \rho \in [\rho_{min}, \rho_{max}], \, \forall n \in [n_{min}, n_{max}] \\ 0 & \text{else} \end{cases}.$$

(2.15)

2.2.5 MH algorithms and the acceptance probability

What guides an MCMC sampler like the MH algorithm to convergence upon the posterior distribution is the acceptance probability α (e.g., Liang et al., 2010). At every step of the Markov Chain, a candidate model is sampled by perturbing the current model from a known distribution (the proposal distribution q) and the acceptance α is calculated. A random number r is then sampled uniformly from the interval [0,1]. If $r < \alpha$ the proposed perturbation is accepted, else the old model is retained. The rationale behind this algorithm can be explained by examining in more detail the expression for α (Bodin and Sambridge, 2009), where

$$\alpha(\mathbf{m'}|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{m'})}{p(\mathbf{m})} \times \frac{p(\mathbf{d}|\mathbf{m'})}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m'})}{q(\mathbf{m'}|\mathbf{m})} \times |\mathbf{J}|\right].$$

(2.16)

Here $\mathbf{m'}$ is the new proposed model and $\mathbf{m}$ is the old model (throughout this paper, primes will denote new model values). Specifically, $\frac{p(\mathbf{m'})}{p(\mathbf{m})}$ is the prior ratio, $\frac{p(\mathbf{d}|\mathbf{m'})}{p(\mathbf{d}|\mathbf{m})}$ is the likelihood ratio and $\frac{q(\mathbf{m}|\mathbf{m'})}{q(\mathbf{m'}|\mathbf{m})}$ is the proposal ratio. The Jacobian term $|\mathbf{J}|$ is not to be confused with the model Jacobian needed for gradient based inversions (e.g., Constable et al., 1987), but is a matrix that incorporates changes in model dimension when moving from $\mathbf{m}$ to $\mathbf{m'}$. In a classic MH algorithm with a fixed number of dimensions, the prior ratio, proposal ratio (for symmetric proposals), and Jacobian term are all 1 (Dettmer et al., 2010). Hence the algorithm always moves towards areas of higher posterior probability if the data misfit improves (likelihood ratio > 1). However, it can also move to areas of lower posterior probability with a probability α if the misfit does not improve (likelihood ratio < 1).

To be able to compare likelihoods between models with different numbers of parameters (i.e., with different dimensions), the Jacobian in the acceptance term in

equation (2.16) needs to be evaluated. There are various implementations of RJ-MCMC, and in all the examples cited so far, a 'birth-death' scheme has been used. As shown in Bodin and Sambridge (2009) and Dettmer et al. (2010) for the 'birth-death' RJ-MCMC scheme, this Jacobian term is unity. We have adopted the 'birth-death' algorithm in this paper and shall not concern ourselves with this Jacobian term any further.

As to why the algorithm should not always look to improve the data fit by simply increasing the number of parameters (interfaces in the seabed), if we examine equation (2.16) we find that even if the likelihood ratio times the proposal ratio is greater than one for a proposed move that inserts a new interface into the model, the prior ratio will be less than one owing to the fact that the new prior PDF $p(\mathbf{m}')$ needs to integrate over a larger number of parameters to equal 1. Hence, there is an opposition to the 'birth' of a new layer (which may lead to improvement of data fit) by the prior ratio.

2.2.6 Outline of our algorithm

We start the algorithm with a 'minimum structure model,' following the approach of Agostinetti and Malinverno (2010). Placing one layer above a halfspace (i.e., $n = 1$) we start with a single interface and then allow the algorithm to iteratively add interfaces ('birth') or remove them ('death'), perturbing the layer resistivities, as the data may demand via the acceptance probability α in (2.16).

We then closely follow the 'partition modelling' approach of Bodin and Sambridge (2009), but use layer interfaces instead of their Voronoi cells (Dettmer et al., 2010). We perturb candidate model resistivities just as Agostinetti and Malinverno (2010) perturb elastic properties for their seismic receiver function problem. We depart from the above methods in our implementation by randomly using a local (small) or global (large) proposal for perturbing resistivities when the number of dimensions is fixed (i.e., it is not a birth or death move) as suggested in Andriéu et al. (2003). In brief, this is how we proceed:

Initialization

Start the algorithm with $n = 1$ and layer resistivities sampled randomly from the uniform prior distribution represented by equation (2.14),

Update layer properties

At every even numbered step we perturb *all* the $2(n+1)$ layer resistivities about their current values using a Gaussian proposal $q(\mathbf{m}'|\mathbf{m})$ with a standard deviation Σ_ρ, where

$$q(\mathbf{m}'|\mathbf{m}) = \left(\frac{1}{\sqrt{2\pi}\Sigma_\rho}\right)^{2n+2} \exp\left[-\frac{1}{2\Sigma_\rho^2}(\boldsymbol{\rho}' - \boldsymbol{\rho})^T(\boldsymbol{\rho}' - \boldsymbol{\rho})\right]. \qquad (2.17)$$

Note that this update move does not involve a change in the number of interfaces.

It is in this move that we depart from previously carried out work, by randomly switching between an update proposal step size of Σ_ρ to a smaller step size $f\Sigma_\rho$. f is a fraction to be determined by the problem at hand and is constant throughout the algorithm. The randomly selected step size allows the algorithm to take global or local steps while sampling the model space during fixed dimension moves (Andriéu et al., 2003).

Perturb model geometry

At every odd numbered step we allow one of the following 4 moves:

1) Birth of an interface: $n' = n + 1$. Uniformly between z_{min} and z_{max}, we randomly select an unoccupied depth and insert an interface. The layer above or below the newly created interface is randomly selected and *only* its horizontal and vertical resistivity are perturbed according to a 2D Gaussian proposal with standard deviation Σ'_ρ.

2) Death of an interface: $n' = n - 1$. An existing interface is selected at random and deleted. The new horizontal and vertical resistivity in the layer are randomly copied from one of the two layers that used to exist on either side of the deleted interface.

3) Move an interface: $n' = n$. An existing interface is selected at random and its depth perturbed by a 1-D Gaussian proposal with standard deviation Σ_m. If the new interface depth is shallower or deeper than the interfaces currently above or below it, this move can be thought of as a simultaneous death/birth move. i.e., properties from one of the layers on either side of the old interface are copied (at random) and assigned to one of the layers on either side of the new interface position (also at random).

4) No perturbation: $n' = n$ and $\mathbf{z}' = \mathbf{z}$.

Each of these model geometry perturbation moves is applied with a certain prob-

ability, such that their sum equals 1. In addition, the birth and death probabilities are set equal. We set the probabilities as follows: [birth, death, move, no perturbation] $\equiv$ $\left[\frac{3}{8}, \frac{3}{8}, \frac{1}{8}, \frac{1}{8}\right]$

At each step of the Markov chain, the proposed model is evaluated for acceptance. If it is accepted, it becomes the current model. If it is rejected, the current model is preserved and the algorithm moves on to the next step. In order to compute the acceptance, one needs to evaluate equation (2.16), for which we explicitly describe the proposal distributions and their ratios in the next section.

2.2.7 Proposal distributions and acceptance probabilities

Our proposal distributions are similar to those described in Bodin and Sambridge (2009) and Dettmer et al. (2010), but since the trans-dimensional approach to the best of our knowledge has not been used for solving the CSEM problem we review the model proposals in modest detail here.

Fixed dimension moves

For all moves that are neither birth nor death, the number of interfaces remain fixed. In these moves, we have elected to use Gaussian proposals to suggest the new model parameters by centering the proposals on the old parameters and drawing a random number from a normal distribution with a given standard deviation (step size). As can be seen from equation (2.17) these kinds of moves are symmetric, implying that the probability to go from the old state to the new state is the same as it would be in going from the new state to the old state:

$$\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{fixed}} = 1. \tag{2.18}$$

Since the number of dimensions remains constant, the prior ratio in equation (2.16) is 1. Hence for fixed dimension moves, we find that the acceptance probability is simply the

ratio of the likelihoods:

$$
\alpha_{\mathrm{f}} = \begin{cases} \min\left[1, \dfrac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \text{if } z \in [z_{min}, z_{max}] \text{ and } \rho \in [\rho_{min}, \rho_{max}], \ \forall n \in [n_{min}, n_{max}] \\ \\ 0 & \text{else} \end{cases}.
$$

$$(2.19)$$

Birth move

For a birth move, one can select from out of $N - n$ unoccupied spaces. The perturbations for the birthed layer's horizontal and vertical resistivity are drawn from a 2D Gaussian with standard deviation Σ_{bd}, centered about the old values in the layer. Since the selection of an interface depth and the perturbations are independent, we can write

$$
\begin{aligned}
q(\mathbf{m}'|\mathbf{m}) &= q(\mathbf{z}'|\mathbf{m})q(\boldsymbol{\rho}'|\mathbf{m}) & (2.20) \\
&= \frac{1}{(N-n)}\frac{1}{2\pi\Sigma_{bd}^2}\exp\left[-\frac{(\rho_h' - \rho_h)^2 + (\rho_v' - \rho_v)^2}{2\Sigma_{bd}^2}\right]. & (2.21)
\end{aligned}
$$

For the reverse move in a birth, keeping in mind that the current state has n interfaces, a birth would have $n + 1$ interfaces to delete from, and the probability of removing resistivities in a layer in the reverse move is 1, thus we can say that

$$
\begin{aligned}
q(\mathbf{m}|\mathbf{m}') &= q(\mathbf{z}|\mathbf{m}')q(\boldsymbol{\rho}|\mathbf{m}') & (2.22) \\
&= \frac{1}{(n+1)} \times 1. & (2.23)
\end{aligned}
$$

Thus in a birth move, from equations (2.21) and (2.23), the proposal ratio can be written as

$$
\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{birth}} = \frac{(N-n)2\pi\Sigma_{bd}^2}{n+1}\exp\left[\frac{(\rho_h' - \rho_h)^2 + (\rho_v' - \rho_v)^2}{2\Sigma_{bd}^2}\right].
$$

$$(2.24)$$

Finally from equations (2.15), (2.16) and (2.24) we get for the birth moves, the following acceptance probability

$$
\alpha_{\mathrm{b}} = \begin{cases} \min\left[1, \frac{2\pi\Sigma_{bd}^2}{\Delta\rho^2}\exp\left[\frac{(\rho_h'-\rho_h)^2+(\rho_v'-\rho_v)^2}{2\Sigma_{bd}^2}\right]\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \text{if } \rho \in [\rho_{min}, \rho_{max}], \\ & \forall n \in [n_{min}, n_{max}] \\ 0 & \text{else} \end{cases} \qquad (2.25)
$$

Death move

In a death move, one can select one of n places for deletion. Further, the probability of removing resistivities in a layer is certain. Thus,

$$
\begin{aligned}
q(\mathbf{m}'|\mathbf{m}) &= q(\mathbf{z}'|\mathbf{m})q(\boldsymbol{\rho}'|\mathbf{m}) && (2.26) \\
&= \frac{1}{n} \times 1. && (2.27)
\end{aligned}
$$

In the reverse move for death, since the reference state has n interfaces, there are $N - (n - 1)$ sites at which to add an interface. Further, the resistivity perturbations are proposed using a 2D Gaussian centered around the current values. Hence

$$
\begin{aligned}
q(\mathbf{m}|\mathbf{m}') &= q(\mathbf{z}|\mathbf{m}')q(\boldsymbol{\rho}|\mathbf{m}') && (2.28) \\
&= \frac{1}{N-n+1} \times \frac{1}{2\pi\Sigma_{bd}^2}\exp\left[-\frac{(\rho_h-\rho_h')^2+(\rho_v-\rho_v')^2}{2\Sigma_{bd}^2}\right]. && (2.29)
\end{aligned}
$$

Thus we can see from equations (2.27) and (2.29) that the proposal ratio for death can be written as:

$$
\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{death}} = \frac{n}{(N-n+1)2\pi\Sigma_{bd}^2}\exp\left[-\frac{(\rho_h-\rho_h')^2+(\rho_v-\rho_v')^2}{2\Sigma_{bd}^2}\right]. \qquad (2.30)
$$

Again from equations (2.15), (2.16) and (2.30) we get for the death moves, the following acceptance probability

$$
\alpha_{\mathrm{d}} = \begin{cases} \min\left[1, \frac{\Delta\rho^2}{2\pi\Sigma_{bd}^2}\exp\left[-\frac{(\rho_h'-\rho_h)^2+(\rho_v'-\rho_v)^2}{2\Sigma_{bd}^2}\right]\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \forall n \in [n_{min}, n_{max}] \\ 0 & \text{else} \end{cases} \qquad (2.31)
$$

It should be noted that the derived expressions for α in equations (2.19), (2.25) and (2.31) do not involve the variable N (as promised) and are very similar in form to the expressions derived in Bodin and Sambridge (2009). This demonstrates how flexible the algorithm is when solving completely different kinds of geophysical problems.

2.2.8   Convergence to the posterior distribution

The algorithm is run for a given number of steps until it is deemed to have collected enough samples to provide a reasonable estimate of the posterior. There are a couple of caveats in this regard, as there are with any MCMC sampler (Liang et al., 2010). If the algorithm is seeded with an initial model that is in a low posterior probability region, it may take quite a few steps till it reaches a region of high posterior probability, such that it begins to sample models which fit the data within the given data error. The number of such required steps (which are subsequently discarded in the final chain) is known in MCMC parlance as the 'burn-in' period, which depends on how well the proposal distributions have been scaled (Chib and Greenberg, 1995). This brings us to the step sizes (scaling) in the proposal distributions in the form of the standard deviations $\Sigma_\rho, \Sigma_m, \Sigma_{bd}$ and the fraction f required in the various proposals to generate a new candidate model. The form of the proposal distributions should 'emulate' the posterior for efficient sampling, but since the posterior distribution may be complicated (and unknown *a priori*), any kind of simple distribution, symmetric where possible, can be used. The exact form of the proposal does not affect convergence. The suitability of the step size for the problem at hand can be examined by looking at the number of samples accepted in a large interval of steps, referred to as the acceptance rate. If the acceptance rate is too low, it means that the step sizes are too large as lots of steps are falling outside the prior bounds or are being rejected as they land in low probability (high misfit) areas. If the acceptance rate is too high, then it implies that the algorithm isn't exploring the model space enough and will again be slow to converge upon the posterior distribution. While the sampled posterior should not depend on the size of the steps taken, one has to factor in the optimality of the step size as otherwise convergence will be very slow (Bodin and Sambridge, 2009). For an illuminating discussion on this matter, one can refer to Chib and Greenberg (1995) or Trainor-Guitton and Hoversten

(2011) for a more recent discussion relevant to marine CSEM. For further discussions on convergence diagnostics and the practicality of their application, one can refer to Liang et al. (2010) The algorithm should be run long enough at the lowest acceptable RMS (achieved after the burn-in period) such that there is at least stationarity achieved in the square misfit with iteration number. Further, to ensure that the inferred posterior is not biased due to being trapped in local maxima (of the posterior probability), we recommend that the algorithm be run from many different starting points, ideally in parallel for computational efficiency. The final ensemble for posterior inference can be constructed by concatenating the various parallel chains (e.g., Dettmer et al., 2010; Agostinetti and Malinverno, 2010; Bodin et al., 2012).

2.2.9 Implementation Details

We implemented the trans-dimensional Bayesian inversion algorithm in MAT-LAB since it allowed for rapid prototyping, testing and modification. Specifically, we coded the comprehensive suite of equations for magnetic and electric dipole sources embedded in TIV anisotropic media that are listed in Appendix F of Løseth and Ursin (2007). The Hankel transform integrals are evaluated using the digital filters and lagged convolution MATLAB codes that are freely-available in Key (2012a). In order to speed up the program's performance, the portion handling the 1-D CSEM forward calls was converted to C and precompiled as a MEX (MATLAB Executable) using the MATLAB Coder utility. The numerical CSEM responses from the forward code were validated by comparisons with suites of anisotropic responses given in Løseth and Ursin (2007) and Li and Dai (2011). We validated the reliable sampling of the trans-dimensional Bayesian algorithm by ensuring that we could sample and recover the specified uniform prior model distribution by removing the data and running the RJ-MCMC algorithm. Since resistivity can span many orders of magnitude, we follow the usual approach for EM geophysics and parameterized the code to invert for $\log_{10}$(resistivity) rather than its linear counterpart.

2.3   Example Applications

2.3.1   Synthetic studies for the resolution of a thin resistive layer

In this example application, we use the trans-dimensional inversion algorithm to study how well each CSEM field component can resolve the anisotropic resistivity of an offshore hydrocarbon reservoir model. In a previous study conducted for a simple isotropic reservoir model, the Occam inversion method was used to find the smoothest inversion model for each data component, generally showing that the field components from a HED are best able to recover the sediment and reservoir resistivity as compared to a VED (Key, 2009). Additionally, because the HED data are the most commonly collected on commercial surveys, we restrict our synthetic studies here to data generated from a HED.

The model under consideration is based on the well-studied canonical 1-D reservoir model (e.g., Constable and Weiss, 2006) but here has been updated to include a moderate amount of anisotropy in the overburden sediments, as shown in Fig. 2.1. The main feature of the model is a 100 m thick, 100 ohm-m isotropic reservoir located 1 km beneath the seafloor. The overburden resistivity is 1 ohm-m in the horizontal direction but 4 ohm-m in the vertical direction, representative of bulk anisotropy due to sediment layering.

We modeled 20 source-receiver offsets spaced evenly from 0 to 12 km, with the transmitter positioned 10 m above the seafloor and broadcasting frequencies 0.1, 0.3 and 0.7 Hz. As an example of the contributions of the various model features to the CSEM responses, Fig. 2.2 shows the inline radial electric field E_r magnitude versus offset (MVO) and phase versus offset (PVO) responses for the anisotropic model (red), a similar model but where the sediments are isotropic with 1 ohm-m resistivity (green), and a model with a uniform 1 ohm-m isotropic sediments (dashed blue). The resistive reservoir can be seen to be the largest contributor to the responses, but the overburden anisotropy also contributes substantially to the increased MVO responses accompanied by markedly different PVO responses.

Synthetic data were generated for both inline and broadside transmitter geometries by adding 5% random Gaussian noise to the model responses and then removing
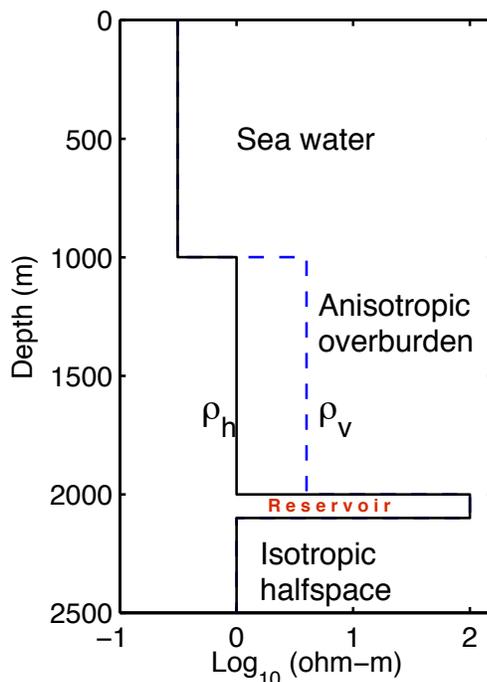
Figure 2.1: Synthetic model of a thin resistive hydrocarbon reservoir with overlying anisotropic conductive sediments. The horizontal resistivity is shown in solid black, and the vertical resistivity with blue dashes.
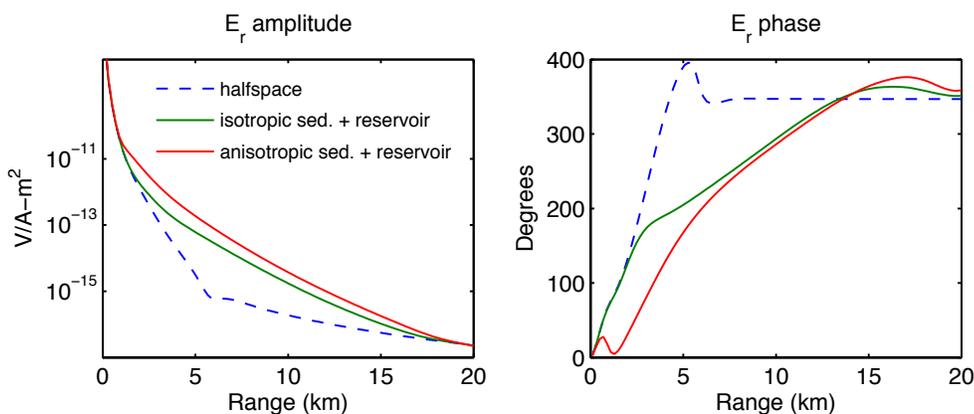
Figure 2.2: Source normalized CSEM responses shown as magnitude versus offset (MVO) and phase versus offset (PVO) curves for the radial electric field at 0.7 Hz produced by a horizontal electric dipole (HED). Plotted responses correspond to an isotropic 1 ohm-m seafloor in 1 km of water (dashed blue), the isotropic model corresponding only to ρ_h in Fig. 2.1 (green) and the full anisotropic model in Fig. 2.1 (red).

data below amplitudes of 10^{-15} V/Am^2 for electric fields and 10^{-18} T/Am for magnetic fields, representative of the electric and magnetic field noise floors for currently available CSEM transmitter-receiver systems with 60 second data stacks.

The trans-dimensional algorithm was independently applied to the synthetic noisy data for each of the 6 EM field components in order to obtain the marginal PDFs on the model parameters, which then can be used to compare the resolution properties of each component. We allowed the number of interfaces to range from 1-30, and the range of depths at which to place interfaces from 1005 m to 2500 m. We used uniform prior distributions on the log-resistivities in the seafloor that allowed them to vary between -1 and 2.3 (corresponding to linear resistivities of 0.1 to 200 ohm-m). For each field component, a total of 200 MCMC chains were run in parallel (each starting point was randomly sampled from the prior distribution), with 2 million steps taken in each each chain. Approximately 1/3 of the chain length was discarded as part of the initial 'burn-in' process (Dettmer et al., 2010) and the remaining steps were collected. The length of the 'burn-in' period was determined by monitoring for when the algorithm reached an RMS data misfit value of ≈ 1 (corresponding to a χ^2 value equal to the number of data points); after the burn-in period, the algorithm continued to sample models that fit the data within an acceptable data RMS misfit. The final concatenated MCMC chain is constructed from all 200 parallel runs, but with care to only concatenate chains that have converged upon the region of acceptable RMS. Those chains, which even after 'burn-in' failed to achieve an RMS close to 1 were discarded (Trainor-Guitton and Hoversten, 2011). For each field component's inversion, the final concatenated chains consists of roughly 200,000 models after thinning the chain to retain 1 out of every 1000 samples (as a means of reducing the storage requirements and the burden for later plotting algorithms). We found that a further decimation of 1 in every 4 models produced acceptable PDFs, with a final sample size of around 50,000 models. Since this trans-dimensional method has not been applied to marine CSEM data before (to the best of our knowledge), we preferred to be cautious rather than undersample the parameter PDFs. The implication is that we potentially could have used a smaller number of parallel MCMC chains or a shorter chain length, or some combination of both to significantly reduce the computational load of this task, given that we achieved acceptance rates between 12%-22%. These acceptance

rates are in agreement with those reported by Dettmer et al. (2010) and Agostinetti and Malinverno (2010).

The final ensembles for each data component consist of about 50,000 1-D models, each of which contains a variable number of layers, the layer interface depths and the horizontal and vertical resistivities of each layer. To visualize the range of acceptable resistivities, we binned these parameters into a grid of fine depth intervals to produce a histogram of resistivities at each depth, normalized by the total number of resistivities encountered in a depth bin. Higher histogram values at a particular depth bin indicate a higher probability of the resistivity being at the corresponding value at that depth. Likewise, the probability that an interface exists at a certain depth is shown by binning the interface depths required by the various models over a range of fine depth intervals.

The results are shown in Fig. 2.3, where each row corresponds to an inversion of a certain EM field component. In every row, the first two columns correspond to PDFs of the horizontal and vertical resistivity as a function of depth, with the 5% and 95% quantiles at each depth shown in blue. The true model is shown in red, and higher values of probability at a depth are shown by darker shading. The third column in each row corresponds to the probability of a layer interface at a given depth, with the true values at which interfaces are located indicated by red horizontal lines. The fine black vertical line in this column corresponds to a uniform probability on the location of interfaces. Again, higher values indicate a greater probability for the presence of a resistivity contrast at the corresponding depth. Fig. 2.4 shows the probability of the number of interfaces required for each particular field component, where the vertical red line denotes the true value of 2 interfaces for our 3 layered anisotropic model.

From the synthetic inversions we can make a number of observations. The vertical resistivities for all components (with the exception of H_z which is a pure TE mode) have the most probable value aligned with the true value at shallower depths. At the reservoir and greater depths, the well-known trade off between reservoir thickness and resistivity (Constable and Weiss, 2006; Key, 2012b) results in the PDFs being more diffuse, yet still concentrated near the true resistivity. Further, there is a marked increase in the 95% quantile line in the vicinity of the reservoir whereas at much shallower depths the 95% quantiles suggest there is little probability of a resistive layer.
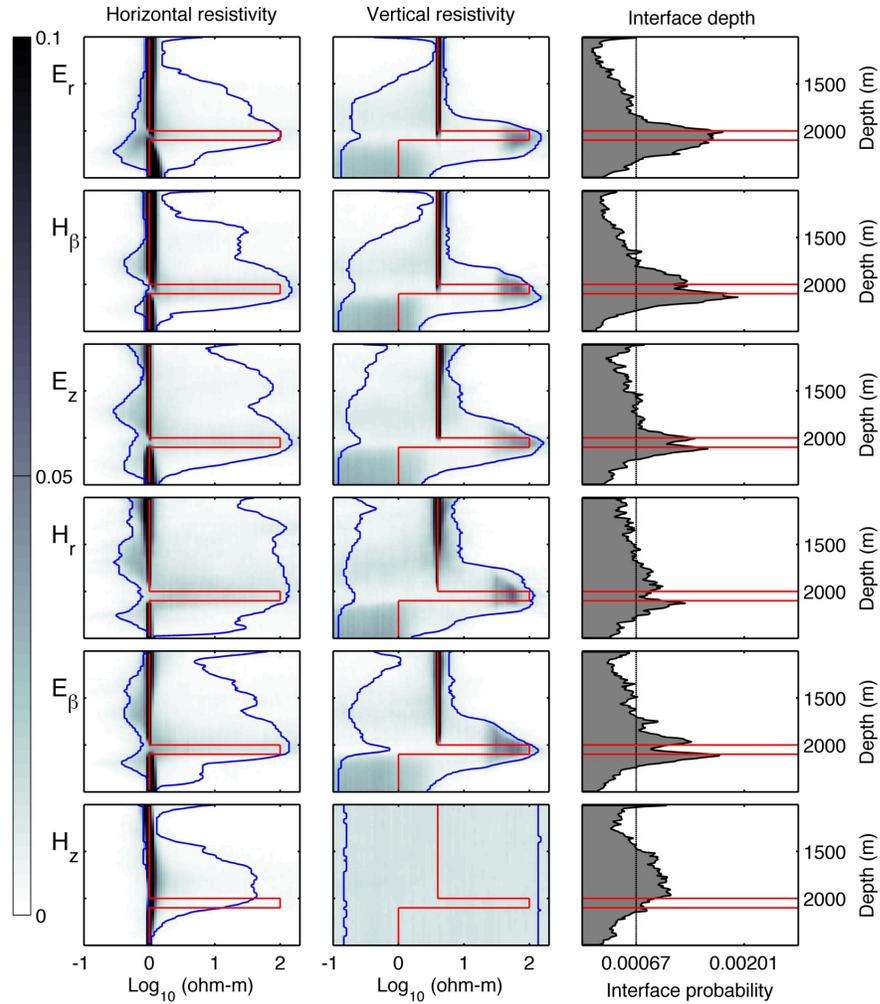
Figure 2.3: Trans-dimensional Bayesian inversion of synthetic HED data with 5% added noise. The results from independent inversion of each field component are shown rowwise. Marginal distributions on anisotropic resistivity values at depth are shown as shaded colors in the first two columns while the third column indicates the probability of an interface. 5% and 95% quantile lines are shown in blue and the true values are indicated by red lines. The top three rows correspond to the components measured by receivers inline with the HED while the bottom three rows are for components measured by receivers broadside to the HED.
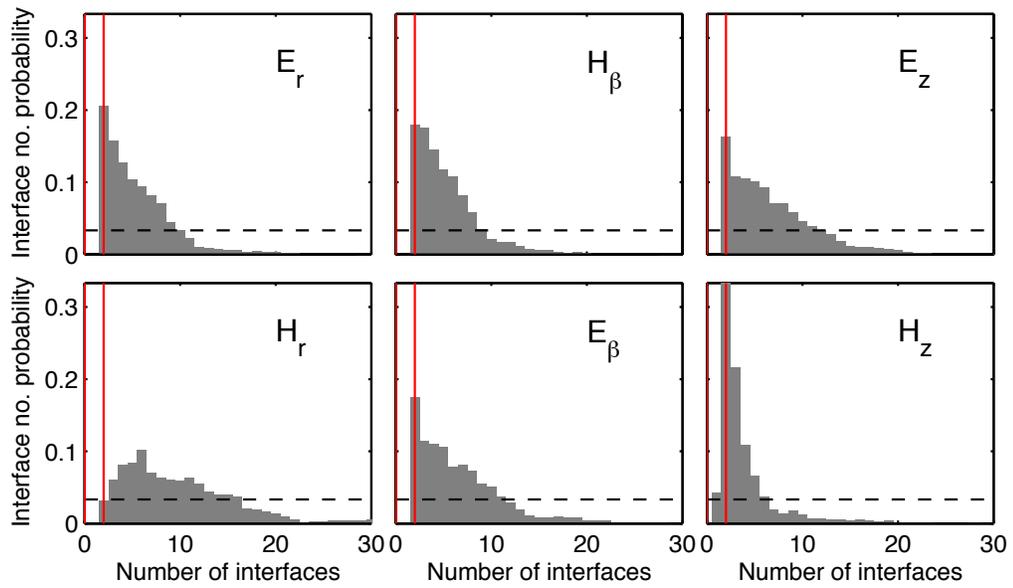
Figure 2.4: Probability of the number of interfaces determined from independent inversions of synthetic data for each field component. The true number of interfaces (2) is shown by the red line. The fine dashed black horizontal line corresponds to a uniform probability on the number of interfaces.

For the horizontal resistivities, the most probable values at shallow depths are also closely aligned with the true value until the reservoir level, where there is an abrupt decrease as the PDF becomes greatly spread out with no significant peak and the confidence of the estimate (indicated by the width between the 5% and 95% quantiles) at these depths is quite low. This inability to resolve the horizontal resistivity within the reservoir is not surprising, as the current density in the resistive reservoir layer becomes predominantly vertically polarized with little horizontal flow and hence contains greatly diminished sensitivity to the horizontal resistivity at this depth (Brown et al., 2012). There is however, a slight increase in the 95% limit around the reservoir depth at least suggesting the probability of an increased resistivity at this depth. Perhaps surprisingly, the lower halfspace resistivities are well imaged by the horizontal resistivity PDFs, probably owing to the fact that to complete the circuit, current must flow horizontally at this depth.

The interface depth probabilities in Fig. 2.3 show that all components (with the exception of H_z) have a high probability for layer interfaces near the correct depths of the reservoir. Some estimate well either the reservoir top (by E_r) or the reservoir bottom (by H_β or H_r) or both the top and bottom (by E_β and E_z). The vertical magnetic field H_z seems unable to clearly estimate either the reservoir top or bottom, but this is unsurprising given that it only contains the TE mode which is well-known to be insensitive to thin resistive layers. The ability of the E_z to recover both the top and bottom depths of the reservoir is intriguing, particularly since previous smooth inversion studies suggested it has slightly lower resolution for the reservoir thickness (see Fig. 7 in Key, 2009). While all other components (besides H_z) contain a mixture of TE and TM modes, the E_z component only contains the TM mode; perhaps its ability to resolve the reservoir boundaries is due to galvanic charge accumulation effects. Further exploration of this result is clearly warranted but is beyond the scope of this work.

Anisotropy in the shallower section is well resolved by all components (again with the exception of H_z), with the tightest bounds on anisotropy being given by the inline radial electric field. For all these components, a good indicator for the presence of a resistive reservoir seems to be the combination of the peak in the interface depth probabilities aligning with a high probability of large vertical resistivity. As can be seen

in Fig. 2.3, this alignment occurs around the true reservoir vertical resistivity and depth.

The probability for the number of interfaces shown in Fig. 2.4 illustrates that the algorithm did not spend much time testing a large number of layers (which require longer times per forward call), suggesting a computational efficiency advantage to using the RJ-MCMC sampler over alternative approaches using a large number of fixed layer interfaces. In fact, one suggested approach when using this algorithm is to first look at the ensemble of all inverted models to identify the most probable number of layer interfaces, and then to examine the PDFs for the subset of models with that number of layers.

Fig. 2.5 shows that when we subset the posterior ensemble to display only models which require 2 interfaces (i.e. 3 layers) to fit the data, which of course is the true number of interfaces in the synthetic model, we get very tight constraints on the ensuing distributions for the radial inline electric field data. This agrees with a previous studying showing a tight region of low misfit in the two parameter model space of reservoir resistivity and thickness (Key, 2012b). In an exploration scenario, the 'data mining' or post-processing of the trans-dimensional inversion results opens up the possibility of incorporating more information as and when it becomes available during the interpretation phase, in order to obtain more certainty in the inverted results without having to perform any further inversion – as the full posterior model solution only needs to be subsetted for this kind of exploration of model space. This ability is further explored in our application to real data, as described in the next section.

2.3.2 Application to data from the Pluto gas field, Northwest Shelf of Australia

In this section, we apply the trans-dimensional inversion to real CSEM survey data collected over the Pluto gas field on the Northwest Shelf of Australia as part of a commercial survey in early 2007. A previous analysis of this single site of CSEM data was presented in the context of solving for unknown receiver orientations in Key and Lockwood (2010). Pluto together with the close by Xena field comprise a 5 TCF gas accumulation (Tilbury et al., 2009). The discovery well Pluto-1 found a gas column of 209 m gross thickness approximately 3100 m below sea level in a sandy tilted Triassic fault
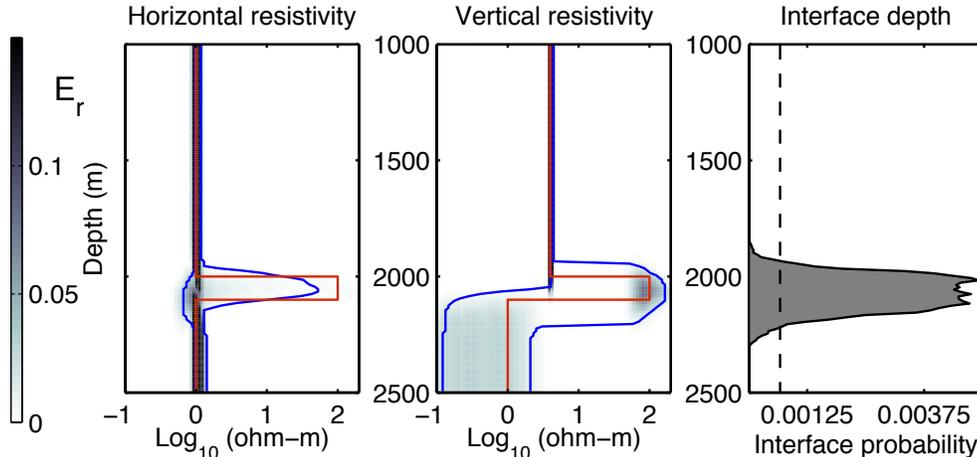
Figure 2.5: Subset of the posterior ensemble for the E_r synthetic data inversion where the number of interfaces is equal to the true value of 2. The 5% and 95% quantiles are far tighter than for the full ensemble shown in Fig. 2.3.

block, with an average net porosity of 28% with gas saturations in the main reservoir of around 93%. Radial inline electric field data collected at 0.24 and 0.4 Hz for a single site were inverted using the trans-dimensional RJ-MCMC method described in this paper.

The trans-dimensional inversion results for the resistivity PDFs and the number of interfaces required are shown in Figs. 2.6 and 2.7, respectively. The horizontal and vertical resistivity as a function of depth shows little evidence of anisotropy from the seafloor down to 2000 m, with the most probable resistivity value being 0.5 ohm-m. However, the 5% and 95% quantile lines on the horizontal resistivity are larger than the very narrow PDFs of the vertical resistivity in this depth interval. The interface probability as a function of depth (third column in Figs. 2.6) suggests a large probability of an interface at 2150 m, followed by another peak at 2950 m. The vertical resistivity jumps to a most probable value around 10 ohm-m accompanied by an increase in the 95% quantile on resistivities at depth. This is very similar to what we observed in our synthetic studies on E_r, as can be seen by comparing Fig. 2.3 with Fig. 2.6. Further, at depths beyond 2000 m the horizontal resistivity PDFs are very diffuse, again similar to what was seen in the synthetic studies in resistive layers. However, as there seem to multiple resistive targets at depths beyond 2000 m, the PDFs do not show a narrowing again as they they did for the single resistive layer synthetic data, but instead are more
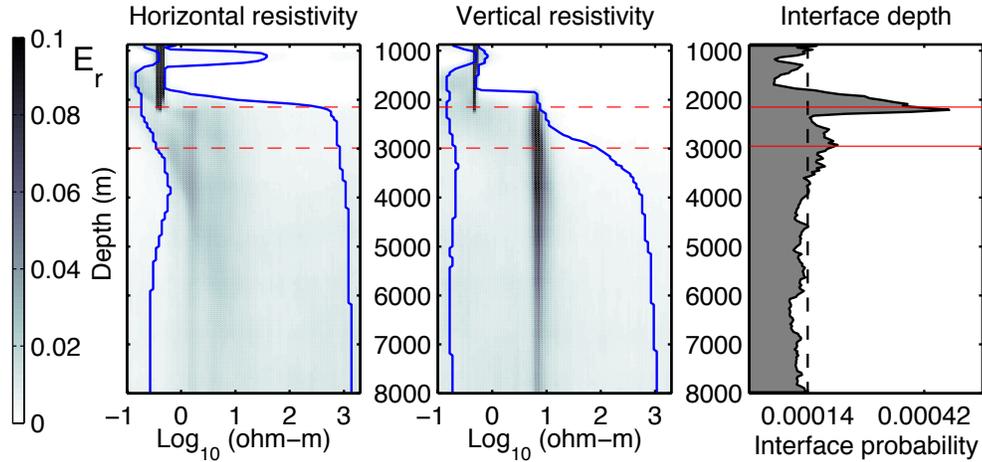
Figure 2.6: Results of the trans-dimensional inversion of the Pluto inline radial electric field data for a single receiver. Interfaces with high resistivity contrast are indicated and marked with red horizontal lines at 2150 m and 2950 m depth (right pane), with the values of most probable vertical resistivity at these depths being 10 ohm-m (middle pane). Note the similarities with the synthetic studies of E_r on encountering a resistor at depth – the horizontal resistivity PDFs become diffuse (left pane) but there is a pronounced peak in the vertical resistivity PDFs.

diffuse indicating that resolution decreases with depth, as expected for CSEM on the conductive sediments of the continental shelves. The most probable number of interfaces, from Fig. 2.7 is 3.

The data fit for 50 randomly sampled models (out of a total of 94,450) in the final posterior model distribution is shown in Fig. 2.8, along with the corresponding models. The scatter in the resistivity models mirrors the probabilities shown in the PDF plots of Figs. 2.6, where the vertical resistivity at 1-2 km depth has the least scatter, in association with its tightest probabilities. Conversely, the horizontal resistivity at these depths can be seen to be less well resolved.

In Fig. 2.9, a number of different scenarios have been subsetted (rows b-d) from the full posterior model distribution (row a). The inline radial electric field has been previously inverted using the Occam method to generate a smooth isotropic 1-D model (Key and Lockwood, 2010); this result is shown as the red line in Fig. 2.9. The Occam inversion model exhibits a smooth resistivity peak in the vicinity of the highest resistivities observed in the Pluto-1 well log. It appears that the highest probability of an
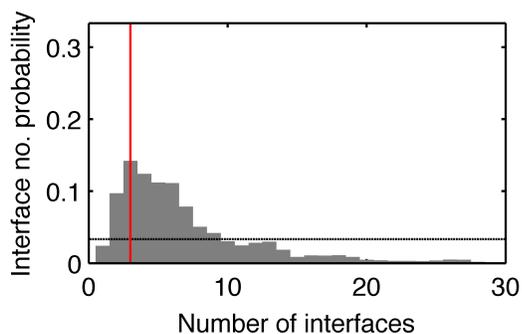
Figure 2.7: PDF of the number of interfaces required to fit the Pluto data. The most probable number of interfaces is 3.
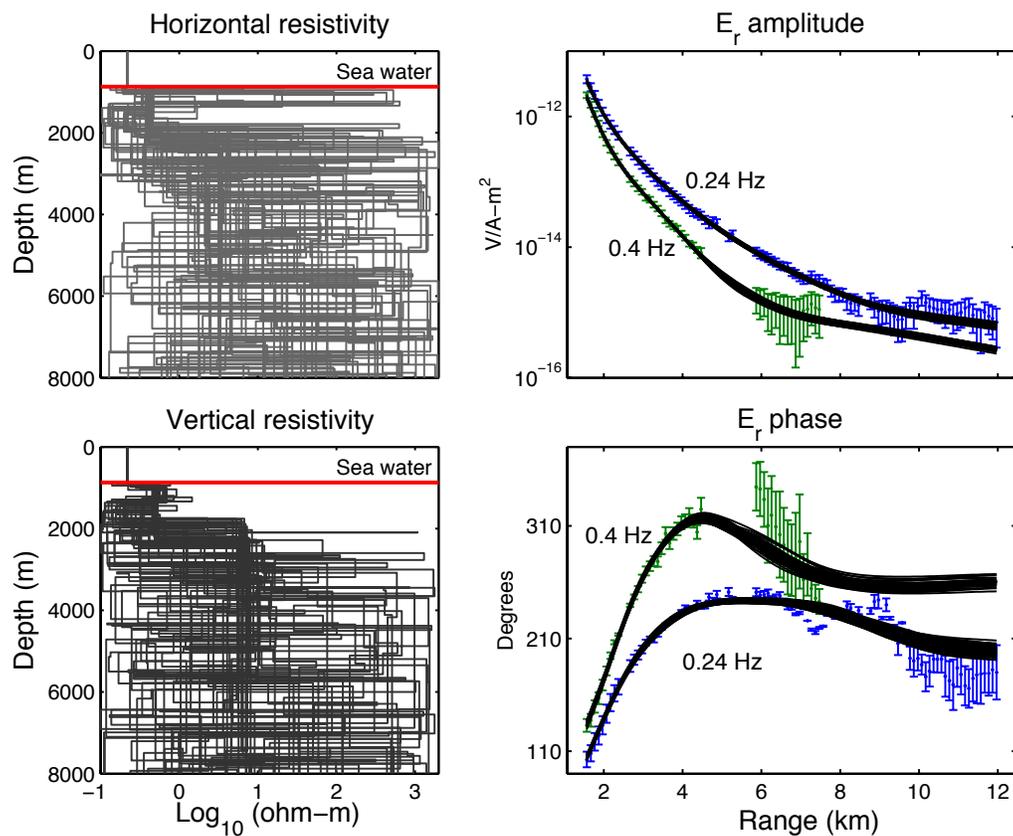
Figure 2.8: A randomly chosen set of 50 models (left) fitting to RMS 1 from the full posterior model distribution and their data fits (right). Note how similar the sampled models when plotted together begin to look like the PDFs of resistivity shown in Fig. 2.6.

interface in the full posterior solution shown in Fig. 2.9a is shallower at close to 2150 m depth. There is another interface probability peak close to 2950 m, although it is of a lower probability. This does not imply that there needs to be interfaces exclusively at one depth or the other, as we shall show with the following analysis.

Following the same methodology as introduced for the synthetic studies, we look at the posterior distribution of the number of interfaces, and pick out the most probable number, 3. If we only look at this subset of models with exactly 3 interfaces (wherever they may be) from the final posterior solution ensemble, we obtain the scenario in (b). With this subset the 5% and 95% quantiles become closer in the resistivity PDFs and the probability of an interface near 2150 m depth is now accentuated. 14% of the 94,450 models in the full posterior correspond to this scenario. In (c), we only look at the subset of models from the full posterior ensemble that require interfaces between the depth range 2000 m to 2450 m. 77% of the models in the full posterior correspond to this scenario. Notice how all these models, which require there to be interfaces between 2000 m to 2450 m, also require a high probability of there being interfaces between 2450 m to 3300 m. Finally, we could choose to incorporate our knowledge of the nearby well-log which says there is an interface at 3000 m. Or we could also look at it as if we were examining the probability of there being models which require interfaces between 2450 m to 3300 m and what features that would require in the rest of the model. Subsetting such models, which require interfaces between 2450 m to 3300 m, gives us scenario (d), corresponding to 67% of the models in the full posterior ensemble. Notice how there is still a high probability of interfaces at shallower depths between 2000 m to 2450 m for all these models. Given that 96% of models require interfaces between 2000 m to 3300 m, one can calculate the probability on the number of models that require interfaces to be present in both intervals = (77% + 67%) - 96% = 48%, also borne out by subsetting the full model posterior. Thus 48% of models in the final posterior solution ensemble require there to be interfaces in both intervals, as we set out to demonstrate.

To further illustrate the high-probability scenario (d) shown in Fig. 2.9, we zoom in to the depths of interest as shown in Fig. 2.10, where hotter colours correspond to higher probabilities. The Pluto-1 resistivity well log has also been included in the middle pane, and both resistivity panes show the smooth isotropic inversion result from Key
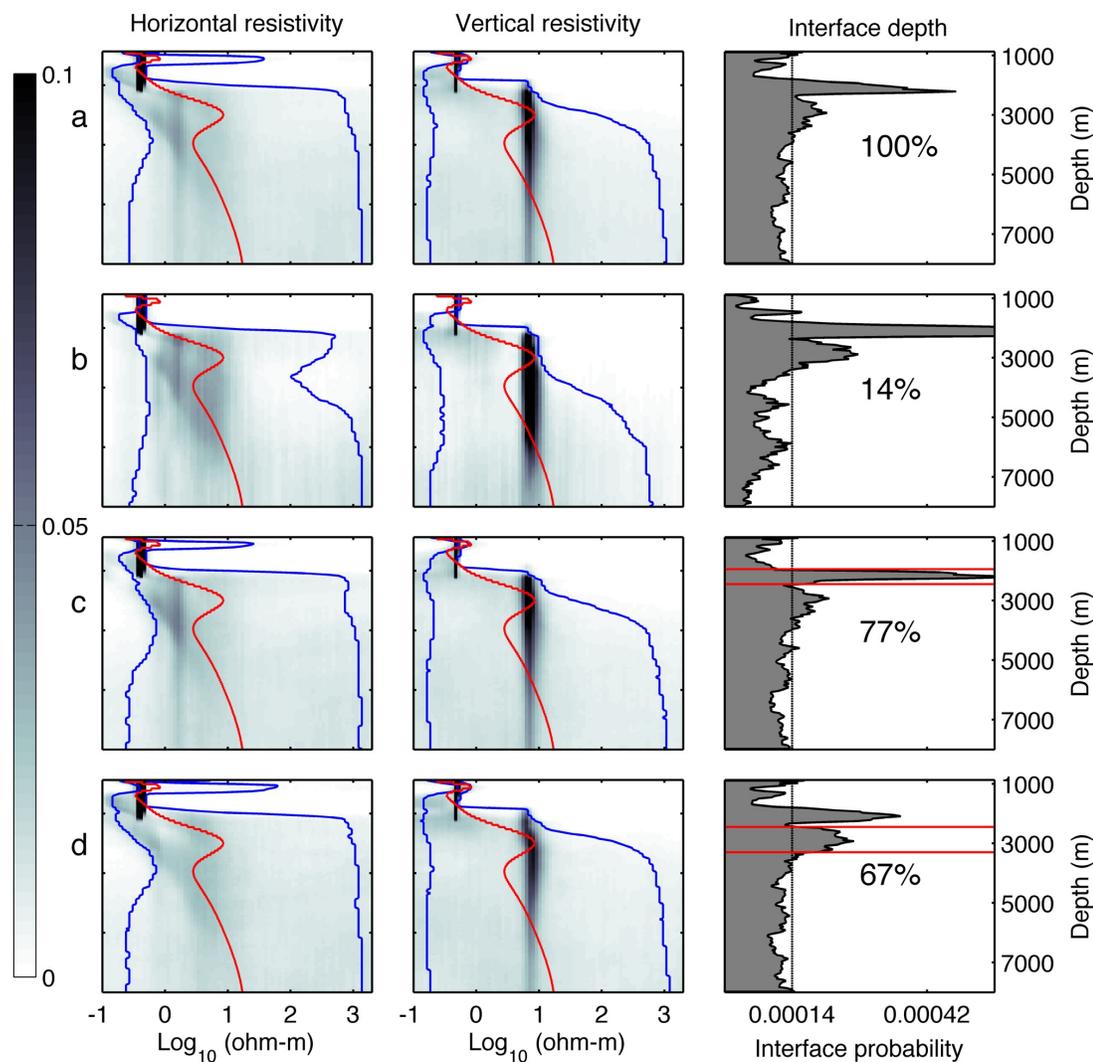
Figure 2.9: Different scenarios mined from the full posterior ensemble. a) The full posterior solution and accompanying PDFs. b) Subset of solutions where the number of interfaces is 3 (the most probable number). c) Subset of solutions which require interfaces between 2000 to 2450 metres, shown by red horizontal lines. d) Subset of solutions which require interfaces between 2450 to 3300 metres, shown by red horizontal lines. The fraction of models associated with each scenario out of the full posterior are given as percentages in the last column.

and Lockwood (2010). It is encouraging to see that the peak resistivity in the smooth inversion coincides with the high probability value of about 10 ohm-m at around 2950 m depth, as is also suggested by the well log. However, it is rather intriguing that there is a greater probability for another interface near the shallower depth of 2100 m. Given that the isotropic smooth inversion (Key and Lockwood, 2010) was regularized with a penalty on model roughness, it is not surprising that this interface was not found given that it would be difficult for such an inversion to resolve a sharp jump in resistivity at depth and the possibility of two or more closely spaced resistive layers.

An explanation may be that the high vertical resistivity at these shallower depths is caused by the presence of regionally extensive shales above the Pluto gas field (Tilbury et al., 2009). Shales can exhibit a large vertical resistivity and small horizontal resistivity due to microscopic grain layering and macro scale layer interbedding; since such shales could potentially mask or be mistaken for a resistive hydrocarbon layer, the identification of anisotropic shales is of considerable concern for CSEM exploration (e.g., Brown et al., 2012). One cannot also discount the possibilities of there being a multi-level reservoir or the reservoir top being shallower in the survey area than in Pluto-1. Lastly, this feature could also be the effect of 2-D or 3-D geology, but these possibilities need to be fully modelled by 2-D or 3-D forward solvers to see if the feature is still retained in the resulting posterior distributions. Here our interpretation is limited by having only a single site of relatively noisy data (compared to data from more recent industry surveys). Bayesian analysis of data from additional CSEM receivers located both on and off the Pluto reservoir, along with integration of the stratigraphy determined from seismic imaging and petrophysical analyses from the many nearby wells, would almost certainly shed more light on the nature of the shallower resistive interface detected here.

2.4 Conclusions

In this study we have developed a Bayesian approach for the inversion of CSEM data. We used this approach to quantify the resolvability of anisotropic conductivity for a 1D model representative of an offshore hydrocarbon reservoir, and to quantify the resolution of marine CSEM data collected at the Pluto gas field offshore Northwest
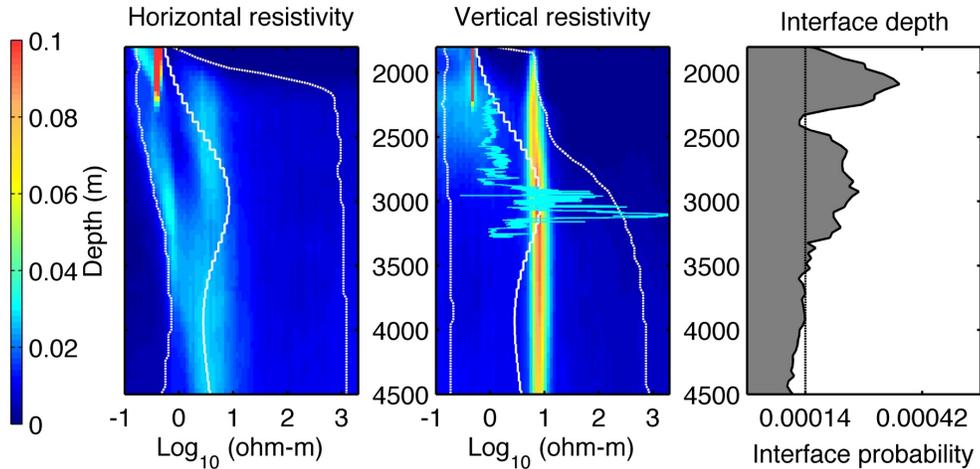
Figure 2.10: Close up view of scenario d) from Fig. 2.9 which requires there to be interfaces between 2450 m to 3300 m. In the middle pane, the Pluto-1 well log is shown as the thin blue line. The smooth inversion result from Key and Lockwood (2010) is shown as the thin white line. Hotter colours in the resistivity PDFs represent higher probabilities. Note how the probability of interfaces at shallower depths around 2150 m still remains. The probability of this scenario is 67%.

Australia.

Our work with synthetic models show that anisotropy can indeed be well resolved in the shallower sediments, at least when the ratio of vertical to horizontal resistivities is near 4:1. Highly resistive isotropic reservoir layers seem to be characterized by very wide distributions with only the peak in the distribution of vertical resistivities at reservoir depth being trustworthy. All field components from the horizontal electric dipole except H_z show a similar resolution of the reservoir vertical resistivity, whereas H_z is completely insensitive to the reservoir. The radial inline electric field E_r seems best at constraining the shallow anisotropy while the vertical electric field E_z seems best at resolving both the reservoir top and bottom depths (Fig. 2.3). Given that both these field components can be acquired with standard inline geometry, it augurs that with relatively straightforward survey designs over a 1-D target, anisotropy and the reservoir can be resolved if the target geology is not significantly different from what has been studied here.

Our approach to data inversion uses a Bayesian framework and a 'birth-death' RJ-MCMC sampler which has not been used for inverting marine CSEM data yet. The algorithm is trans-dimensional in that the number of model parameters is not fixed *a*

priori. Further, the algorithm is self-parametrizing in that the interfaces in a particular model are not fixed to be at a specific location and can move. We differ from other trans-dimensional RJ-MCMC algorithms in mixing a global and local 'update' move for fixed dimension steps instead of using more computationally intensive approaches such as delayed rejection. The algorithm is highly flexible and this method of Bayesian inversion is the most truly 'data driven' approach that the authors have yet converged upon. A further improvement in this direction would be adding the data noise parameters themselves as unknowns (Bodin et al., 2012), which would tackle the crucial issue of estimating data noise, which directly affect the misfit function and thus the posterior solution converged upon.

The full posterior ensemble can be subsetted to display model distributions corresponding to a most probable parameter such as the number of interfaces (Bodin et al., 2012), to incorporate new information as and when it may become available, or, for hypothesis testing of different plausible geological scenarios. Subsetting is rather easily carried out and does not require additional inversion. Further, different scenarios can be quantitatively compared by the fraction of models required in each scenario out of the total number of models in the full posterior distribution of models.

Application of the trans-dimensional inversion to marine CSEM data from the Pluto field are in general agreement with previous results obtained with smooth inversion. The differences, where they exist, highlight the fact that smooth models which are subjectively regularized in order to obtain a stable and meaningful result out of the many highly oscillatory ones that all fit the data (Fig. 2.8) – may miss features that are not smooth variations in resistivity, but nonetheless are likely geological features from a data misfit point of view.

Despite its numerous advantages, an obvious drawback to using this method is the fact that a large number of forward calls need to be made to sample the posterior model distribution adequately. For 1-D earth models, the problem is tractable in that reasonable estimates of the posterior solution can be found within a couple of hours when running on parallel computer systems. The rate limiting step is essentially how long it takes to achieve burn-in for each chain. For Pluto, we ran 128 chains in parallel over 12 hours and oversampled the distribution by at least a factor of two. Post burn-in, if

enough chains are run in parallel, the number of further steps required can be reduced if the number of parallel chains is increased. For 2-D or 3-D geology, evaluating a forward model response can take considerably longer and the millions of forward calls may take inordinate amounts of time when done serially. However, given that parallel 2-D and 3-D forward solvers for EM geophysics problems are becoming increasingly more efficient and that independent Markov chains are highly suited to being run in parallel, this bodes well for the future of this flexible Bayesian approach to inverse problem solving.

Acknowledgments

Chapter 2, in full, is a reprint of material as it appears in the Geophysical Journal International, Ray, A., and Key, K., Blackwell Publishing Ltd., 2012. The dissertation author was the primary investigator and author of this paper.

References

Abubakar, A., Habashy, T. M., Druskin, V. L., Knizhnerman, L., and Alumbaugh, D.,
2008: 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements. *Geophysics*, **73**(4), F165–F177.

Agostinetti, N. P., and Malinverno, A., 2010: Receiver function inversion by transdimensional monte carlo sampling. *Geophysical Journal International*, **181**(2), 858–872. ISSN 1365-246X. doi:10.1111/j.1365-246X.2010.04530.x.

Andriéu, C., de Freitas, N., Doucet, A., and Jordan, M. I., 2003: An introduction to mcmc for machine learning. *Machine Learning*, **50**, 5–43. ISSN 0885-6125. 10.1023/A:1020281327116.

Backus, G. E., 1988: Bayesian inference in geomagnetism. *Geophysical Journal*, **92**(1), 125–142. ISSN 1365-246X. doi:10.1111/j.1365-246X.1988.tb01127.x.

Bayes, T., 1763: An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, **53**, 370–418.

Bodin, T., and Sambridge, M., 2009: Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, **178**(3), 1411–1436. ISSN 1365-246X. doi:10.1111/j.1365-246X.2009.04226.x.

Bodin, T., Sambridge, M., Tkalcic, H., Arroucau, P., Gallagher, K., and Rawlinson, N., 2012: Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research*, **117**(B02031), 1–24.

Brown, V., Hoversten, M., Key, K., and Chen, J., 2012: Resolution of reservoir scale electrical anisotropy from marine csem data. *Geophysics*, **77**(2), E147–E158. doi:10.1190/geo2011-0159.1.

Buland, A., and Kolbjornsen, O., 2012: Bayesian inversion of CSEM and magnetotelluric data. *Geophysics*, **77**(1), E33–E42.

Chen, J., Hoversten, G. M., Vasco, D., Rubin, Y., and Hou, Z., 2007: A Bayesian model for gas saturation estimation using marine seismic AVA and CSEM data. *Geophysics*, **72**(2), WA85–WA95.

Chew, W., 1995: *Waves and field in inhomogeneous media*. Wiley-IEEE Press.

Chib, S., and Greenberg, E., 1995: Understanding the metropolis-hastings algorithm. *The American Statistician*, **49**(4), pp. 327–335. ISSN 00031305.

Collins, M. D., and Fishman, L., 1995: Efficient navigation of parameter landscapes. *Acoustical Society of America Journal*, **98**, 1637–1644. doi:10.1121/1.413430.

Constable, S., 2010: Ten years of marine CSEM for hydrocarbon exploration. *Geophysics*, **75**(5), 75A67–75A81.

Constable, S. C., Parker, R. L., and Constable, C. G., 1987: Occam's inversion - A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(03), 289–300.

Constable, S. C., and Weiss, C. J., 2006: Mapping thin resistors and hydrocarbons with marine EM methods: Insights from 1D modeling. *Geophysics*, **71**(2), G43–G51.

Dettmer, J., Dosso, S., and Holland, C., 2010: Trans-dimensional geoacoustic inversion. *Journal of the Acoustical Society of America*, **128**, 3393–3405.

Dosso, S. E., and Dettmer, J., 2011: Bayesian matched-field geoacoustic inversion. *Inverse Problems*, **27**(5), 055009.

Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M. C., MacGregor, L. M., and Constable, S., 2002: Remote sensing of hydrocarbon layers by seabed logging (SBL): Results from a cruise offshore Angola. *The Leading Edge*, **21**, 972–982.

Green, P., 1995: Reversible jump MCMC and Bayesian model selection. *Biometrika*, **82**, 711–732.

Gunning, J., Glinsky, M. E., and Hedditch, J., 2010: Resolution and uncertainty in 1D CSEM inversion: A Bayesian approach and open-source implementation. *Geophysics*, **75**(6), F151–F171.

Hastings, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109. doi:10.1093/biomet/57.1.97.

Hou, Z., Rubin, Y., Hoversten, G. M., Vasco, D., and Chen, J., 2006: Reservoir-parameter identification using minimum relative entropy-based Bayesian inversion of seismic AVA and marine CSEM data. *Geophysics*, **71**(6), O77–O88.

Huelsenbeck, J. P., Larget, B., and Alfaro, M. E., 2004: Bayesian phylogenetic model selection using reversible jump Markov chain Monte Carlo. *Molecular Biology and Evolution*, **21**(6), 1123–1133.

Key, K., 2009: 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, **74**(2), F9–F20.

Key, K., 2012a: Is the fast Hankel transform faster than quadrature? *Geophysics*, **77**(3), F21–F30. doi:10.1190/geo2011-0237.1.

Key, K., 2012b: Marine electromagnetic studies of seafloor resources and tectonics. *Surveys In Geophysics*, **33**(1), 135–167.

Key, K., and Lockwood, A., 2010: Determining the orientation of marine csem receivers using orthogonal procrustes rotation analysis. *Geophysics*, **75**(3), F63–F70. doi:10.1190/1.3378765.

Li, Y., and Dai, S., 2011: Finite element modelling of marine controlled-source electromagnetic responses in two-dimensional dipping anisotropic conductivity structures. *Geophysical Journal International*, **185**, 622–636.

Liang, F., Liu, C., and Carrol, R. J., 2010: *Advanced Markov chain Monte Carlo methods: learning from past samples*. Wiley Series in Computational Statistics. Wiley, New York, NY.

Løseth, L. O., and Ursin, B., 2007: Electromagnetic fields in planarly layered anisotropic media. *Geophysical Journal International*, **170**(1), 44–80. ISSN 1365-246X. doi: 10.1111/j.1365-246X.2007.03390.x.

Malinverno, A., 2002: Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, **151**, 675–688.

Minsley, B. J., 2011: A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. *Geophysical Journal International*, **187**(1), 252–272. ISSN 1365-246X. doi:10.1111/j.1365-246X.2011.05165.x.

Neal, R. M., 2003: Slice sampling. *The Annals of Statistics*, **31**(3), 705–767.

Newman, G. A., and Alumbaugh, D. L., 2000: Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, **140**(2), 410–424.

Ramananjaona, C., MacGregor, L., and Andréis, D., 2011: Sensitivity and inversion of marine electromagnetic data in a vertically anisotropic stratified earth. *Geophysical Prospecting*, **59**(2), 341–360. ISSN 1365-2478. doi:10.1111/j.1365-2478.2010.00919.x.

Scales, J. A., and Snieder, R., 1997: To Bayes or not to Bayes? *Geophysics*, **62**(4), 1045–1046.

Sen, M., and Stoffa, P. L., 1995: *Global optimization methods in geophysics.* Elsevier.

Smith, K., Gatica-Perez, D., and Odobez, J.-M., 2005: Using particles to track varying numbers of interacting people. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, 962 – 969 vol. 1. ISSN 1063-6919. doi:10.1109/CVPR.2005.361.

Tarantola, A., 2005: *Inverse problem theory and methods for model parameter estimation.* Society of Industrial and Applied Mathematics.

Tilbury, L. A., Clayton, C. J., C., J., T., Philip, G., Boyd, G. A., Johnson, G. G., Rayfield, M. A., Hartanto, L., and Lance, D. P., 2009: Pluto - a major gas field hidden beneath the continental slope (dvd). *The APPEA Journal*, **49**, 243–265.

Tompkins, M. J., Fernández Martínez, J. L., Alumbaugh, D. L., and Mukerji, T., 2011: Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: Marine controlled-source electromagnetic examples. *Geophysics*, **76**(4), F263–F281.

Trainor-Guitton, W., and Hoversten, G. M., 2011: Stochastic inversion for electromagnetic geophysics: Practical challenges and improving convergence efficiency. *Geophysics*, **76**(6), F373–F386. doi:10.1190/geo2010-0223.1.

Um, E. S., and Alumbaugh, D. L., 2007: On the physics of the marine controlled-source electromagnetic method. *Geophysics*, **72**(2), WA13–WA26.

Chapter 3

Robust and accelerated Bayesian inversion of marine CSEM data using parallel tempering

Anandaroop Ray[1], David L. Alumbaugh[2,3], G. Michael Hoversten[2] and Kerry Key[1]

[1]Scripps Institution of Oceanography, La Jolla, CA, USA
[2]Chevron Energy Technology Company, San Ramon, CA, USA
[3]Now at NEOS GeoSolutions, Pleasanton, CA, USA

Abstract. Bayesian methods can quantify the model uncertainty that is inherent in inversion of highly nonlinear geophysical problems. In this approach, a model likelihood function based on knowledge of the data noise statistics is used to sample the posterior model distribution, which conveys information on the resolvability of the model parameters. Since these distributions are multi-dimensional and non-linear, we use Markov chain Monte Carlo methods for highly efficient sampling. Because a single Markov chain can become stuck in a local probability mode, we run various randomized Markov chains independently. To some extent this problem can be mitigated by running

independent Markov chains, but unless a very large number of chains are run, biased results may be obtained. We get around these limitations by running parallel, interacting Markov chains with 'annealed' or 'tempered' likelihoods which enable the whole system of chains to effectively escape local probability maxima. We test this approach using a trans-dimensional algorithm, where the number of model parameters as well as the parameters themselves are treated as unknowns during the inversion. This gives us a measure of uncertainty that is independent of any particular parameterization. We then subset the ensemble of inversion models to either reduce uncertainty based on a priori constraints or to examine the probability of various geological scenarios. We demonstrate our algorithms's fast convergence to the posterior model distribution with a synthetic 1-D marine controlled source electromagnetic data example. The speed up gained from this new approach will facilitate the practical implementation of future 2-D and 3-D Bayesian inversions, where the cost of each forward evaluation is significantly more expensive than for the 1-D case.

3.1   Introduction

Marine controlled source electromagnetic (CSEM) methods have been used to image geology with highly resistive contrasts for over three decades (Young and Cox, 1981). Extensive research and commercialization of this technology over the last ten years (Ellingsrud et al., 2002; Constable, 2010) has led to its being added to the staple suite of seismic methods as an exploration tool. Owing to the fact that electromagnetic skin depths are smaller in conductive media, marine geophysical EM methods almost always operate in the lower frequency quasi-static regime. This allows for deeper penetration of the CSEM fields into the earth, but as a consequence it is more a diffusive process than wave like (Loseth et al., 2006). Thus, the resolution of CSEM is never quite as good at depth as that of the seismic method, but the value of CSEM lies in its sensitivity to resistivity (which may be indicative of hydrocarbon saturation), and not acoustic impedance (which may be more indicative of geological structure). Owing to this diffusive characteristic of marine CSEM, robust inferences made from a survey are necessarily from inversion of the data, and not merely from examination of the data itself (Weiss, 2007). Typically, regularized and linearized gradient based inversion methods have been used to arrive at models that in addition to minimizing data misfit are 'optimal' in some user-defined sense. By means of regularization, highly oscillatory features in the model that are thought to be outside the resolution of CSEM are eliminated (e.g., Constable et al., 1987; Newman and Alumbaugh, 2000; Abubakar et al., 2008; Key, 2009; Sasaki, 2013). Though these methods are highly efficient and well understood, they provide only a single smooth model as a result, or a suite of smooth models. These models provide a limited insight into the various classes of models that are compatible with the observed data given the noise. A clear understanding of the resolvability of various parts of the modelspace does not emerge from a linearized treatment of the non-linear CSEM problem.

To quantify the uncertainty inherent in the inversion of CSEM data, one can utilize a Bayesian framework where all information is stored in probability distribution functions or PDFs. Since Bayesian probability (Bayes and Price, 1763) is a measure of information (Scales and Sneider, 1997) and since it is the aim of geophysical inversion to provide information about the earth's subsurface, it is natural to postulate geophysi-

cal inverse problems in a Bayesian framework (Tarantola and Valette, 1982). In such a framework, model parameters are treated as random variables, and their fit to the observed data given the observed statistical noise allow one to formulate a model likelihood. To make the connection with deterministic inversion methods, to first order, models with low misfit possess a higher likelihood. After incorporating prior knowledge of the models independent of the data, the product of the prior model probability and the likelihood is known as the posterior model probability. Those parts of the model space that are more frequently required by the data than other parts manifest with greater posterior probability, and hence are better resolved (Backus, 1988). However, fixing a particular model parameterization (e.g., fixing the number of layers, their thickness and locations) for the inversion is known to produce posterior distributions, *only* for the given parametrization (Dettmer et al., 2010). This is where the 'trans-dimensional' (Bodin and Sambridge, 2009) or 'reversible jump' (Green, 1995) Markov Chain Monte Carlo (RJ-MCMC) differs from traditional MCMC methods, in sampling from a posterior distribution where the number of unknowns and their positions, are also treated as part of the inverse problem (Agostinetti and Malinverno, 2010). Such algorithms have a 'parsimony' property (Malinverno, 2002), which refers to the fact that Bayes' Theorem deems models that explain the data with simpler parameterizations more probable. MacKay (2003) discusses this aspect of Bayes' Theorem in some detail. A good introduction to geophysical trans-dimensional Bayesian inversion can be found in Sambridge et al. (2013).

Parallel tempering is an accelerated MCMC technique sometimes known as 'replica exchange' (Earl and Deem, 2005; Geyer, 1991; Swendsen and Wang, 1987). Using a sequence of parallel, interacting MCMC chains with 'annealed' or 'tempered' likelihoods allows the entire system of chains to effectively sample the model space without getting trapped in local modes of posterior probability. Recent examples of parallel tempering as applied to geophysical inversion can be found in Dosso et al. (2012), who use the method to discover multiple modes in the posterior model distribution while inverting underwater acoustic reverberation data and in Dettmer and Dosso (2012), who use underwater acoustic data to invert for seafloor sediment properties.

Bayesian uncertainty estimation for the CSEM problem has indeed been carried out in the past (e.g., Chen et al., 2007; Trainor-Guitton and Hoversten, 2011; Buland and

Kolbjornsen, 2012). However, these methods considered the model parameterization to be fixed and did not address the model space at different scales. Tompkins et al. (2011) consider a scheme which does indeed address the issue of model parameterization at different scales, but their method is not Bayesian and requires the use of a starting model that fits the data well and the resulting uncertainty retains some of the characteristics of this model. Gunning et al. (2010) use a 'Bayesianized' hierarchical bootstrapping method to address the issue of model resolvability and escape from local probability maxima.

The advantage in using a trans-dimensional approach (Bodin and Sambridge, 2009) is that both the theoretical framework and the practical implementation of it are straightforward - as explicitly shown in Ray and Key (2012), nothing more than a literal interpretation of Bayes' Theorem (Bayes and Price, 1763) and the generation of random numbers is required to explore a realistic posterior model distribution. Recent applications of the trans-dimensional method to solve geophysical EM methods can be found in Minsley (2011), Brodie and Sambridge (2012), who applied it the airborne EM problem, and in Ray and Key (2012), who tackle the marine CSEM problem. In this paper, we extend the application of RJ-MCMC to CSEM by introducing parallel tempering, to hasten convergence of the RJ-MCMC chains and quickly escape local probability maxima. By using parallel tempering, we show that previously undetected modes in the posterior model distributions for CSEM data have come to light, and that the total number of forward solves is reduced to less than half the number required without using it. This is significant as a realistic uncertainty appraisal for 2D CSEM data will require fast, efficient and accurate sampling, as 2D or 3D EM problems are computationally far more expensive with complicated interactions between model parameters. We demonstrate our concept using synthetic 1-D data, and add reversible jump MCMC and parallel tempering to an arsenal of tools for tackling 2D or 3D problems in the future. Lastly, we use the reversible jump or trans-dimensional method to provide a workflow for testing geological hypotheses by subsetting the posterior distribution of inverted models by interrogating it with intelligent queries, without requiring further inversion. We see this as being particularly useful in exploration scenarios where new information keeps coming in over time, which can be used to hone the posterior model distribution without

re-inverting the acquired data.

Theory

3.1.1 Bayesian inversion, Markov chains and the reversible jump

Bayesian information is contained in probability density functions (PDFs) represented by $p(\cdot)$. Using Bayes' theorem, (Bayes and Price, 1763) we write

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m})}{p(\mathbf{d})} \tag{3.1}$$

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior assumptions}}{\text{evidence}} \tag{3.2}$$

For Bayesian geophysical inversion, the data vector $\mathbf{d}$ is a constant. All PDFs with a model dependence are functions of the random variable $\mathbf{m}$. The term $p(\mathbf{d}|\mathbf{m})$ can then be interpreted as the model likelihood, the functional form of which depends on the statistics of the noise distribution, and the value of which depends on the model $\mathbf{m}$ being sampled and its misfit. For Gaussian noise, the model likelihood is given as:

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(-\frac{[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1}[\mathbf{d} - f(\mathbf{m})]}{2}\right) \tag{3.3}$$

Here $f(\mathbf{m})$ corresponds to the modeled data and $\mathbf{C}_d^{-1}$ is the data covariance matrix and $[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1}[\mathbf{d} - f(\mathbf{m})]$ is the χ^2 misfit for the evaluated model $\mathbf{m}$. The prior model distribution $p(\mathbf{m})$ represents our state of knowledge *independent* of the survey data. The evidence term $p(\mathbf{d})$ corresponds to a constant PDF normalizing factor equal to the integral over all models of the numerator in equation 3.1. Though the evidence can help in 'model selection,' i.e., decide which model parameterization is more probable than the other, it is very challenging to compute as it requires evaluation of a multi-dimensional integral over different models, evaluated for different model parameterizations. Another means of performing model selection is to calculate the full posterior model probability distribution $p(\mathbf{m}|\mathbf{d})$ by allowing the problem to be trans-dimensional (Dettmer et al., 2010), i.e., have a varying number of model parameters. This is the RJ-MCMC approach that we have used in this paper, which is different from the usual MCMC approach in

the following manner. Treating the evidence as a proportionality constant, it follows from equation 3.1 that

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m}) \tag{3.4}$$

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m_k}|k) \times p(k) \tag{3.5}$$

What we have effectively done, is not claimed to have known the optimal model paramterization *a priori* in equation 3.4. We have not fixed the prior model probability to be a constant over all models, which fixed dimensional MCMC samplers do. We do not regularize our inversion, as we do not know what an optimal regularization which preserves resolution yet removes spurious oscillatory behavior is. Similarly we simply do not know how many layers to represent the earth with, or what their locations are (Bodin and Sambridge, 2009). For a given model $\mathbf{m}$, we split $p(\mathbf{m})$ into 2 parts. One part contains information about the number of interfaces k in the model, $p(k)$. The other part $p(\mathbf{m_k}|\mathbf{k})$ contains information about where these interfaces are in depth, and what the resistivities in between these layers, the top and the bottom halfspace are, given the number of interfaces k. Our task, is to evaluate uncertainty in the models inverted from the observed data. To this end we must arrive at the posterior distribution of models, most of which fit the data well, by evaluating their misfit and sampling models according to equation 3.5. However, it is nearly impossible to exhaustively sample the model space for more than a few parameters owing to the 'curse of dimensionality,' hence we resort to probing this highly non-linear distribution using various Markov chain Monte Carlo methods (e.g., Liang et al., 2010) and focus on the RJ-MCMC or trans-dimensional method (Sambridge et al., 2013) in this work.

3.1.2   Metropolis-Hastings MCMC and the acceptance probability

The RJ-MCMC sampler is a particular type of Metropolis-Hastings (MH) sampler (e.g., Hastings, 1970; Liang et al., 2010). A MCMC sampler like the MH algorithm converges upon the posterior distribution using an acceptance probability α. At every step of the Markov Chain, a candidate model is sampled by perturbing the current model using a known probability distribution (the proposal distribution q) and the acceptance

α is calculated. The proposal distribution q is usually a simple distribution which should be easy to draw samples from (such as a Gaussian), and also be scaled somewhat like the posterior distribution we are sampling (Ray and Key, 2012). A random number r is then drawn uniformly from the interval [0,1]. If $r < \alpha$ the proposed perturbation is accepted, else the old model is retained. The rationale behind this algorithm can be explained by examining in more detail the expression for α (Bodin and Sambridge, 2009), where

$$\alpha(\mathbf{m'}|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{m'})}{p(\mathbf{m})} \times \frac{p(\mathbf{d}|\mathbf{m'})}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m'})}{q(\mathbf{m'}|\mathbf{m})} \times |\mathbf{J}|\right]. \tag{3.6}$$

Here $\mathbf{m'}$ is the new proposed model and $\mathbf{m}$ is the old model. Specifically, $\frac{p(\mathbf{m'})}{p(\mathbf{m})}$ is the prior ratio, $\frac{p(\mathbf{d}|\mathbf{m'})}{p(\mathbf{d}|\mathbf{m})}$ is the likelihood ratio and $\frac{q(\mathbf{m}|\mathbf{m'})}{q(\mathbf{m'}|\mathbf{m})}$ is the proposal ratio. The Jacobian term $|\mathbf{J}|$ is not to be confused with the model Jacobian needed for gradient based inversions (e.g., Constable et al., 1987), but is a matrix that incorporates changes in model dimension when moving from $\mathbf{m}$ to $\mathbf{m'}$. For a fixed number of dimensions in a classic MH algorithm, the prior ratio (for uniform priors), proposal ratio (for symmetric proposals), and Jacobian term are all 1 (Dettmer et al., 2010). Hence the algorithm always moves towards areas of higher posterior probability if the data misfit improves (likelihood ratio > 1). However, it can also move to areas of lower posterior probability with a probability α if the misfit does not improve (likelihood ratio < 1). This is essentially how MCMC samplers give us 'regional' as opposed to 'point' information about models that fit the data well, unlike gradient based optimizers. Extensive details on the trans-dimensional RJ-MCMC prior and proposal ratios can be found in Ray and Key (2012), Bodin and Sambridge (2009) or Dettmer et al. (2010).

To be able to compare likelihoods between models with different numbers of parameters (i.e., with different dimensions), the Jacobian in the acceptance term in equation 3.6 needs to be evaluated. There are various implementations of RJ-MCMC, and in all the examples cited so far, a 'birth-death' scheme has been used. As shown in Bodin and Sambridge (2009) for the 'birth-death' RJ-MCMC scheme, this Jacobian term is unity. We have adopted the 'birth-death' algorithm in this paper and shall not concern ourselves with this Jacobian term any further.

As to why the algorithm should not always look to improve the data fit by simply increasing the number of parameters (interfaces in the seabed), if we examine equation

3.6 we find that even if the likelihood ratio times the proposal ratio is greater than one for a proposed move that inserts a new interface into the model, the prior ratio will be less than one owing to the fact that the new prior PDF $p(\mathbf{m}')$ needs to integrate over a larger number of parameters to equal 1. Hence, there is an opposition to the 'birth' of a new layer (which may lead to improvement of data fit) by the prior ratio.

3.1.3 Parallel tempering

If one were to examine the likelihood function for Gaussian noise in equation 3.3, it is apparent that the 'peakiness' of the likelihood function can be manipulated if one were to introduce a term that plays the statistical mechanics analogue of temperature in a partition function (Earl and Deem, 2005). Detailed balance is an aspect of MCMC which enables unbiased sampling, allowing samples to be distributed in proportion to the target posterior PDF. If we can 'anneal' or 'temper' the likelihood, without violating the detailed balance (Earl and Deem, 2005), then local probability maxima in the posterior can be overcome. That is, if we were to say that our annealed likelihood should be

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(- \frac{[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1} [\mathbf{d} - f(\mathbf{m})]}{2T_j} \right) \tag{3.7}$$

instead of the untempered version with $T_j = 1$, we would be sampling from a smoother likelihood with less extreme peaks and valleys than the target untempered likelihood. The first examples of accomplishing this without violating detailed balance can be found in Swendsen and Wang (1987) and in its more familiar form as a MCMC sampling algorithm in Geyer (1991). The basic idea is to run N parallel Markov chains in concert, with the jth chain at temperature T_j, with $T_1 = 1$ and $T_j > 1; \forall j > 1$. Temperatures are usually arranged in ascending order, with Markov chains at adjacent temperatures being allowed to exchange their states (models) with a fixed probability or over a fixed number of steps. At the end of this joint simulation of N Markov chains, the target chain at $T_1 = 1$ is used for posterior inference. An effective implementation is to have a temperature ladder of increasing temperatures, as demonstrated in Figure 3.1. This figure describes how the likelihood function can be annealed as a consequence of using equation 3.7. Note how the likelihoods corresponding to higher temperature chains
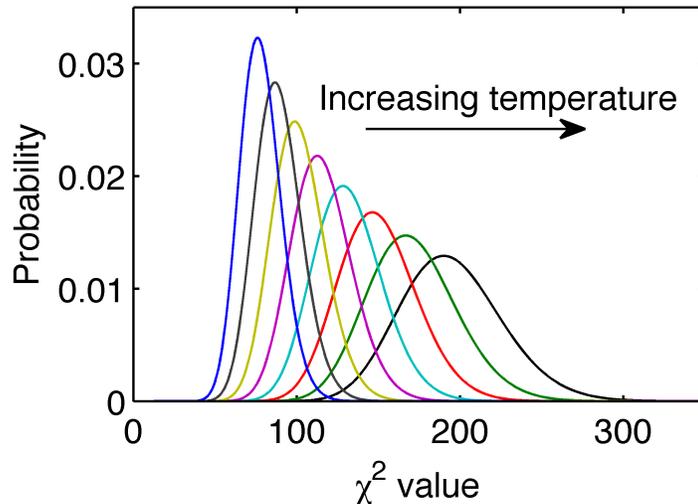
Figure 3.1: Annealed or tempered likelihood functions as a function of their χ^2 misfit. Note how the leftmost likelihood function at $T = 1$ is narrow and peaked, which can manifest as a model space which is harder to sample.

sample higher values of χ^2 misfit over a broader range of probable likelihoods. Thus, these chains are never stuck in local probability maxima. Owing to their overlap with the narrower lower temperature chains, exchange of states (models) for adjacent chains is possible using the following MH acceptance criterion:

$$\alpha_{swap} = \frac{p_{hot}(\mathbf{d}|\mathbf{m}_{cold})}{p_{hot}(\mathbf{d}|\mathbf{m}_{hot})} \times \frac{p_{cold}(\mathbf{d}|\mathbf{m}_{hot})}{p_{cold}(\mathbf{d}|\mathbf{m}_{cold})} \tag{3.8}$$

It is important to ensure while using equation 3.8, that the probability of selecting all pairs of temperatures is equal, and that the pairs are chosen at random in order to maintain detailed balance (Dettmer and Dosso, 2012). Further, to ensure that one has configured a reasonable temperature ladder, accepted exchange rates between adjacent temperatures should be close to 25% (Dosso et al., 2012). The highest temperature should be high enough to allow the chain to escape local probability maxima, yet not so high for there to be no significant overlap between adjacent chains, which will make exchanges of information improbable. For a given model space, adding more chains to the configuration once an optimal exchange acceptance rate is found does not improve chain mixing. A detailed discussion on setting the temperature ladder can be found in

Earl and Deem (2005).

The main motivation for using parallel tempering is that it is very efficient at escaping local probability modes (misfit minima) and the communication between adjacent chains significantly speeds up convergence to the posterior solution. This is an important factor in ultimately carrying out 2D inversion of CSEM data, where the computational complexity in forward model evaluation is far greater. One can escape local posterior probability modes by running lots of independent chains at the target temperature (Ray and Key, 2012; Bodin et al., 2012; Dettmer et al., 2010), but for a 2D problem where the evaluation of any one forward model itself can at present take up the resources of a cluster of computers, it is important to keep the number of parallel chains small. Parallel tempering achieves just this, bringing down the total number of evaluations as the number of communicating parallel chains required is small, a fact we demonstrate in this paper.

3.2   Synthetic inversion tests

As a test model for the algorithm, we use a 1-D model already studied by Trainor and Hoversten (2009) and Tompkins et al. (2011). The model, as well as its phase and amplitude response are shown in Figure 3.2. This model was chosen for study by Trainor and Hoversten (2009) owing to a frustrating lack of well defined convergence to a posterior distribution. 2% Gaussian noise was added to the modeled data, and a standard source normalized amplitude of 10^{-15} V/Am^2 was used as the noise floor.

3.2.1   Interface depths exactly fixed

In the first suite of tests, using the guidelines mentioned for temperature selection, we applied parallel tempering using 4 chains at temperatures T = [1.00, 1.35, 1.84, 2.50]. Note that the temperatures are equally spaced in the log domain (Dettmer and Dosso, 2012). We used an ordinary MH algorithm without the reversible jump for our initial studies. Using 30,000 samples per chain, with swaps attempted at every step, dismissing the first 3,000 as the 'burn-in' samples (low probability, high misfit region of model space), the results from the target chain at $T = 1$ are shown in Figure 3.3a, with the
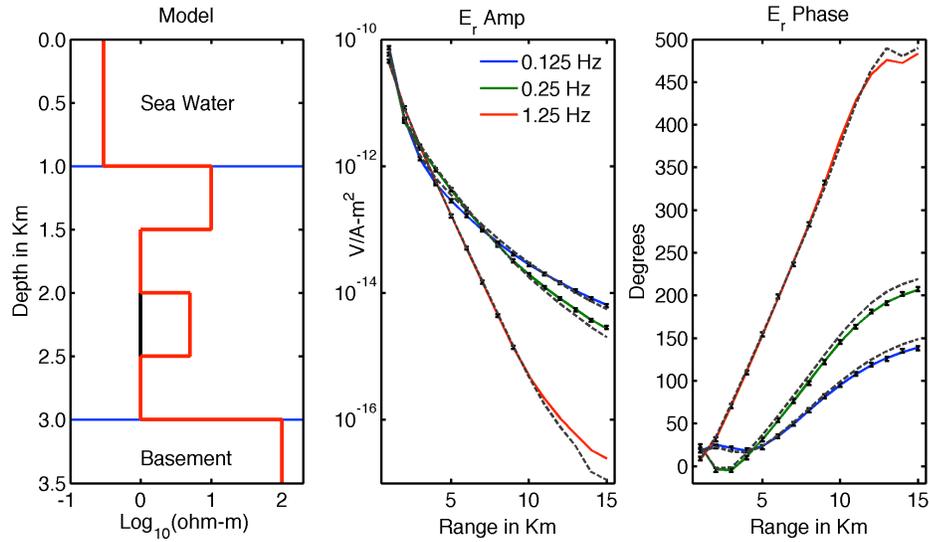
Figure 3.2: Synthetic 1-D model (red) used for this study. Alternating, moderately resistive then conductive sediments are terminated by a highly resistive basement at depth. The CSEM amplitude and phase responses at the seafloor receivers are plotted with range at 3 frequencies. The background responses in the absence of the middle layer (black model) are shown with dashed grey lines.

truth being shown in each layer by a red vertical line at the correct resistivity value. The vertical axis for all plots corresponds to probability of resistivity in the given layer. The displayed marginal probabilities are simply obtained by binning the sampled posterior model resistivities at each depth. The same histograms for each layer, translated into depth and the $\log_{10}$-resistivity domain are displayed as an image in Figure 3.3b. Hotter colors represent higher probability, and cooler colors correspond to lower probability. The plots look far more natural in the log domain, which is more representative of subsurface resistivity than its linear counterpart. The true model is shown with a dotted yellow-black 1-D model. Though we have exactly fixed the model interfaces to be at the true layer depths, we get a somewhat surprising result. In addition to the probable layer resistivities being clustered near the truth (as we expected), in layers 3 and 4, we get $\log_{10}$ resistivities clustered around 0.05 and 0.32 that are less probable than the maximum probable in the layer, but are not negligible (which we did not expect). They have a slightly higher misfit and as expected from the likelihood formulation 3.3, a lower probability. However, they represent models that are within the data error. Root mean

square (RMS) misfit values are obtained by dividing the χ^2 misfit by the number of data points and taking the square root. Models which belong to the 'true' family of models have an RMS misfit close to 0.9, whereas those models belonging to the 'shadow' family of models have an RMS closer to 1.1. Further, in carrying out untempered fixed dimensional inversions, we always chanced upon one or the other family of models. In 10 independent fixed dimensional trials, only 2 converged upon the true family. However, from Figures 3.3a and 3.3b it is apparent that the 'true' family of models is more probable. Thus we would probably have needed many more Markov chains running independently to obtain the true posterior probabilities of the 2 classes of models compatible with the data. In fact, though the results of Trainor and Hoversten (2009) and Tompkins et al. (2011) are generally compatible, albeit with certain distinct differences, neither of them show any hint of the 'shadow' family of models in their posterior distributions. The fact that separate modes in a difficult probability landscape exist is not new, and is the reason why parallel tempering is used in various sampling applications in the first place. We verified the existence of these distinct modes by computing a 5D grid searched posterior over the probable search ranges indicated by Figure 3.3b. Though the resulting grid was coarse (not shown here), it required $15 \times 15 \times 30 \times 30 \times 20 = 4,050,000$ evaluations to prove the existence of separate modes. In contrast, with parallel tempering and MCMC, we needed only $4 \times 30,000 = 120,000$ samples in total. This is a major reason why high dimensional integrals (such as those required for marginal probability distributions) are computed using stochastic methods instead of a brute-force approach. Geophysical evidence of multiple modes in the model space has been amply demonstrated by Dosso et al. (2012) for geoacoustic inversion, and by Gunning et al. (2010) for the case of 'Bayesianized' CSEM inversion. What is perhaps a little surprising is that as simple a model as the one we have studied, demonstrates such non-uniqueness with exact layer parameterization.

3.2.2   Trans-dimensional inversion with parallel tempering

Given that we don't always know the exact layer parameterization required to perform an inversion, we elected to perform a trans-dimensional inversion with the noisy synthetic data. The number of interfaces is allowed to vary from 1 to 15, and the

(a)   (b)
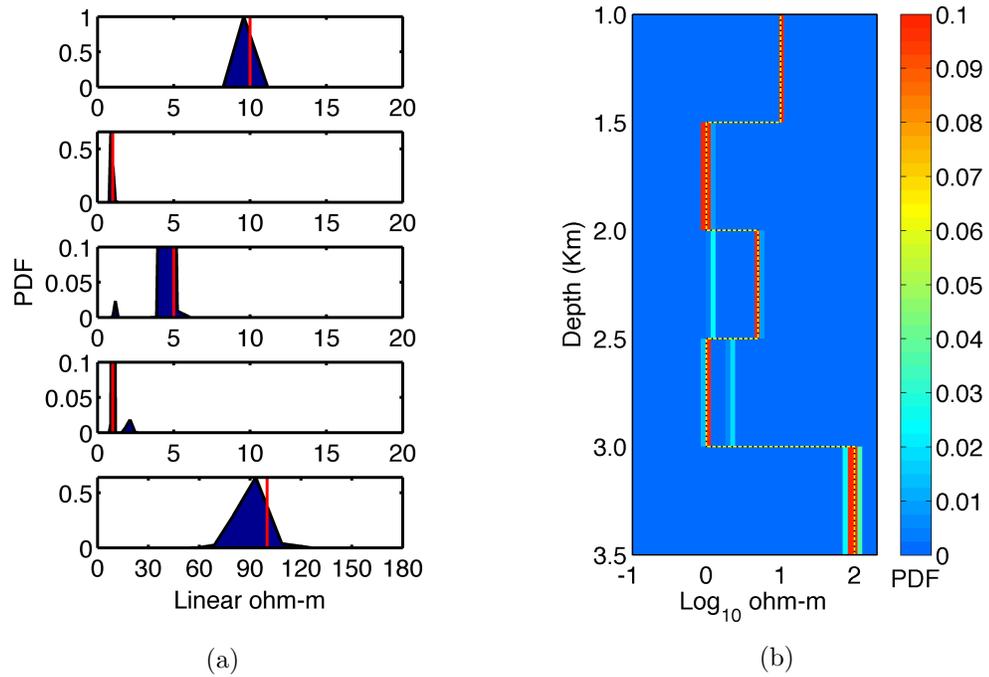
Figure 3.3: a) Marginal resistivity distributions in each layer for a fixed dimensional inversion. The truth is shown with a red vertical line. b) Same as a) except now higher probabilities are shown in hotter colors, and the resistivity scale is in $\log_{10}$ ohm-m. The true model is shown with a dotted yellow-black line.

interfaces can be placed anywhere between 1002 to 3500 meters. Models are allowed to have any $\log_{10}$ resistivities between -1 to 2.3 (0.1 to 200 linear ohm-m). In this section, we also applied parallel tempering to the RJ-MCMC framework. A remarkable aspect of most MCMC algorithms is their flexibility, which allowed us to run a trans-dimensional RJ-MCMC algorithm within a parallel tempering framework with very little modification to either set of codes.

Effect of different temperatures

For the RJ-MCMC application, we used 8 different temperatures at $T = [1.00, 1.14, 1.30, 1.48, 1.69, 1.92, 2.19, 2.50]$ and swapped randomly between 2 chains at every single step. For flexible model parameterization, 8 chains were used as opposed to the 4 used earlier for fixed dimensional MCMC. This is because we now have a different and larger posterior model probability space to sample. The results are shown in Figure 3.4a, where each row corresponds to the indicated temperature. The left panels show probability of resistivity at depth, and the right panels show the probability of the presence of interfaces at depth. At each depth in the left panels, the 5% and 95% quantile lines of resistivity have been indicated with a dashed black line, and the truth with a solid black line. The dotted vertical line in the right panel corresponds to a uniform probability of interfaces between 1002 to 3500 m. As shown in Figure 3.1, the hotter temperature chains sample higher χ^2 misfit values. Returning to Figure 3.4a, clearly, the hotter chains are not able to resolve well the middle $\log_{10}$ resistivity of 0.69 (5 ohm-m linear), preferring to be closer to the 'shadow' value of $\log_{10}$ resistivity of 0.05 (1.12 ohm-m linear) with smoother posterior model ensembles. As we progress downward towards the cooler chains, we see sensitivity to the true value of the middle layer, though it is not uniquely resolved. This is in line with the observations of both Bodin et al. (2012) and Gunning et al. (2010) where they explicitly state that larger data errors than actually observed in the data will lead to less structure (smoother model ensembles) in the inferred posterior. Since increasing the data errors for inversion is in some sense, increasing the temperature in the annealed likelihood, the two statements are equivalent. Thus, the higher temperature chains should be less sensitive to the middle layer, which is indeed what we observe in this study.
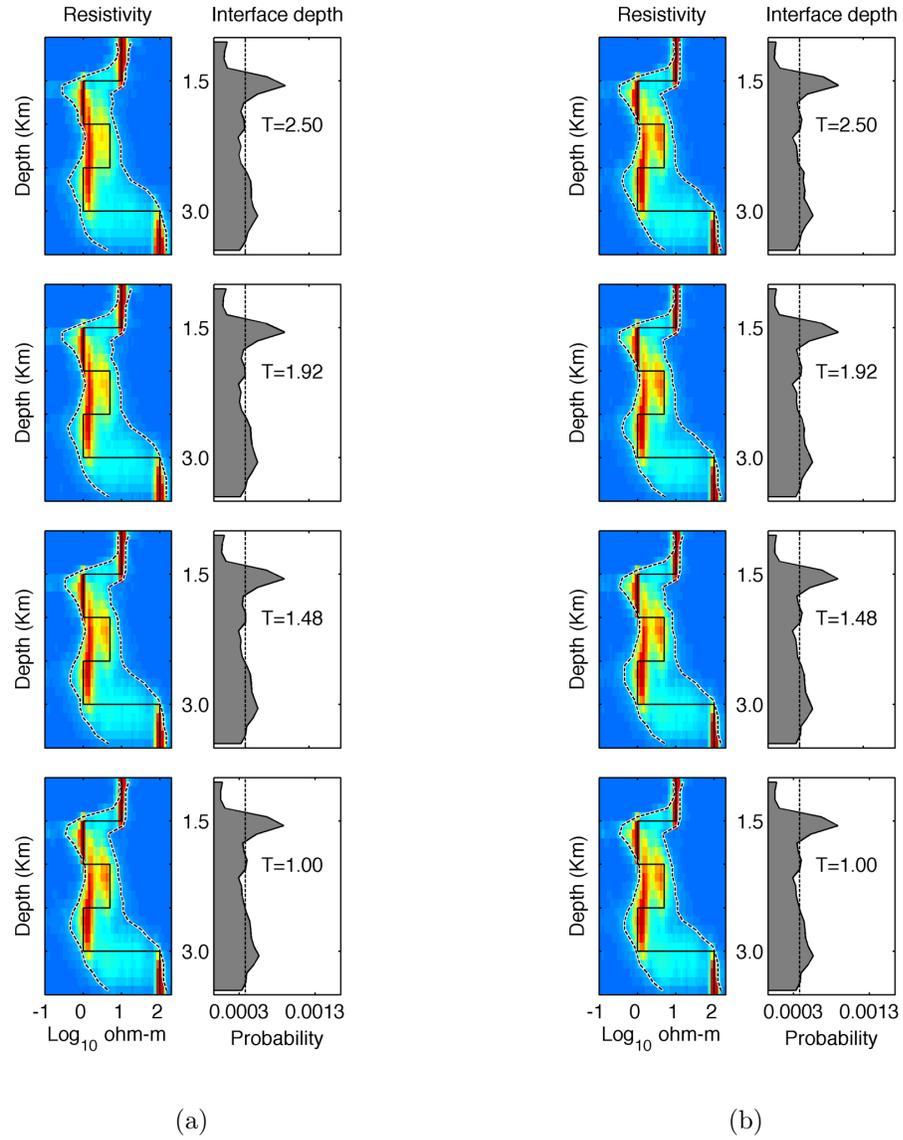
Figure 3.4: a) Unweighted trans-dimensional posterior probability distributions for resistivity at depth (left panel) and interface probability with depth (right panel). Each row corresponds to the indicated temperature. The 5% and 95% quantile lines of resistivity at depth are shown with dashed black lines and the truth with a solid black line. b) Samples at each temperature re-weighted to the target temperature $T = 1$ to remove bias.

Looking at the target chain at $T = 1$, at some depths, the truth is in fact less probable than the 'shadow' value. Given the results of our fixed dimensional modeling, this should not come as a surprise. In this light, the result is also fairly intuitive - when we allow the layer parameterization to be variable, we are looking at a more realistic uncertainty estimate which should clearly be less certain than if we had fixed the interfaces to be at their true positions. Without more information, this is all the subsurface information content that we can glean from the observed data, within the relatively large prior bounds. In the following section, we deal with trying to reduce this uncertainty, *a posteriori*.

Sample re-weighting

We can also re-weight the chains from different temperatures to obtain samples at the target temperature of $T = 1$ using a weighting factor (Brooks and Neil Frazer, 2005; Dosso et al., 2012). For $\mathbf{m}_i^j$, which is the ith model in a chain at temperature T_j with a sampled misfit $\chi^2(\mathbf{m}_i^j)$, the weight is given as:

$$w(\mathbf{m}_i^j) = \frac{\exp\left[-\frac{\chi^2(\mathbf{m}_i^j)}{2}\left(1 - \frac{1}{T_j}\right)\right]}{\sum_i \exp\left[-\frac{\chi^2(\mathbf{m}_i^j)}{2}\left(1 - \frac{1}{T_j}\right)\right]} \tag{3.9}$$

Notice that at a higher temperature, a higher misfit implies a smaller weight. Further, the weight for all models in a chain are the same if $T_j = 1$, implying no re-weighting. For all models in a chain at a particular temperature, the samples can be re-weighted to the target temperature $T = 1$ using equation 3.9. For instance, marginal probabilities of resistivity at depth can be found by binning the resistivity values as before, but each histogram count needs to be multiplied by the weight corresponding to the model being binned. The results of this operation are shown in Figure 3.4b, where each row corresponds to the re-weighted samples at a given temperature. The color scales are the same for all images in Figure 3.4a and Figure 3.4b. All re-weighted distributions look fairly similar, but re-weighted samples from $T = 2.5$ have a slightly rougher posterior distribution – borne out by a minute observation of the interface probability curve at this temperature in Figure 3.4b and comparison with the interface probability curve for the $T = 1$ case. The $T = 2.5$ chain has undersampled the model space to a very

small extent, a similar observation being borne out by Dosso et al. (2012). The extent of undersampling is problem dependent and it should not be a cause for alarm if it is greater than shown here. This undersampling merely illustrates that there is a tradeoff between traversing greater distances in the model space and fine sampling of the target areas. Since parallel tempering allows for the trickling down of information at higher temperatures to lower temperatures, the lower temperature chains will have sampled the model space adequately.

Posterior on the number of interfaces

The marginal posterior distribution on the number of interfaces for the target temperature $T = 1$ is shown in Figure 3.5. The true number of interfaces is shown with a red vertical line. The dashed black horizontal line corresponds to a uniform probability on the number of interfaces. The number of interfaces is a variable that changes from model to model and models with different numbers of interfaces can be swapped between different chains. Since the true number of interfaces is not known by the inversion *a priori*, the most probable number of interfaces is not equal to the true value of 4 but turns out to be 8. This is not unusual and a similar phenomenon has also been observed by Minsley (2011).

Proposal variances at different temperatures

If the posterior probability space has various different length scales associated with multiple modes, as long as equation 3.8 has been obeyed, different proposal distribution step sizes can be assigned to each chain. For instance, we can assign larger proposal variances to the higher temperature chains to ensure that they explore the model space with large steps. The step size for a chain at a particular temperature should be made smaller if the acceptance rates for that chain become very low (e.g., Ray and Key, 2012; Bodin et al., 2012) or if the accepted swap rate between adjacent chains falls far below 25%.
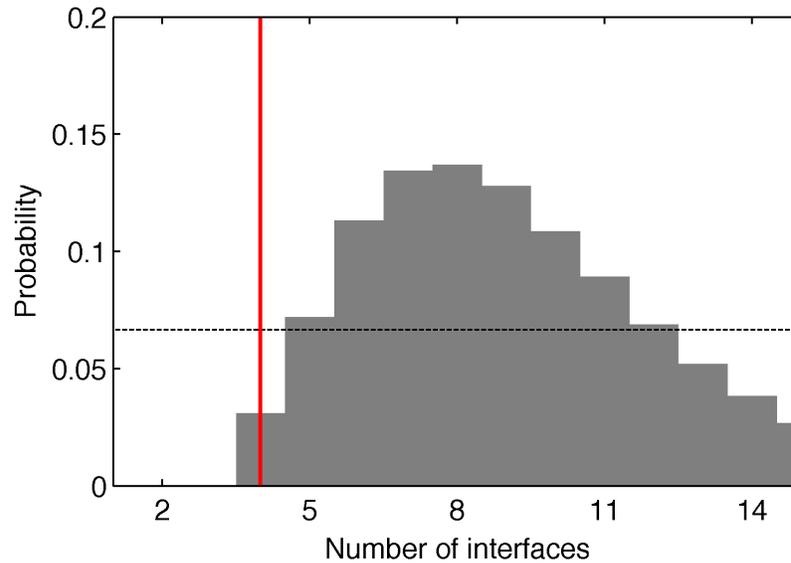
Figure 3.5: Marginal probabilities on number of interfaces required by the observed data at the target sampling temperature. The true number, 4 is marked with a vertical red line. The dashed black horizontal line corresponds to a uniform probability on the number of interfaces.

3.2.3 Convergence to the posterior

A definite statement about convergence to the true posterior is difficult to make, and is still an area of active research, especially with RJ-MCMC where the number of parameters may change in the next step (Bodin and Sambridge, 2009). However, a method that works well is to keep sampling until it is apparent that the posterior distribution does not change appreciably by adding more samples (Dettmer and Dosso, 2012). For the trans-dimensional case, with a burn-in length of 5000 samples, we oversampled by 500,000 samples in each chain in order to test for convergence and stopped sampling at 1.75 million samples per chain.

3.2.4 Comparison between ordinary and parallel tempered MCMC

As is evident from Figures 3.3a and 3.3b, parallel tempering reveals hidden modes in the fixed-dimensional case by transitioning between them with relative ease. Further, since this problem can be mitigated by running numerous independent, non-communicating chains, it is only fair to compare parallel tempering with the case when

we run various independent MCMC chains. For the trans-dimensional case, to obtain the level of detail and convergence found in the last panel of Figure 3.4a, with parallel tempering we ran 8 chains with 1.75 million samples in each chain. In total we utilized 14 million samples with a slightly lower number of total forward evaluations (as proposed samples falling outside the prior bounds are not evaluated). For an equivalent level of detail without using parallel tempering, we needed at least 56 independent RJ-MCMC chains with 500,000 samples each, a total of 28 million samples, with a similar number of forward evaluations. Thus we see that parallel tempering requires both less chains, and a lower total number of forward evaluations by a factor of 2. To illustrate how much better chain mixing is using parallel tempering, a comparison has been made between sampled posteriors from parallel tempering and independent chains, both for 8 chains and 500,000 samples in each chain. Figure 3.6a shows the result from all 8 independent chains and Figure 3.6b shows the results using parallel tempering, inference made from only the target chain at $T = 1$. Though neither ensemble has achieved stationarity yet, an examination of the last panel of Figure 3.4a (which uses the same color scale) shows that the result with parallel tempering is closer to the final posterior distribution, as is evident in the shape of the interface probability curves and the smoothness of the 5% and 95% quantile lines of resistivity at depth. In fact, it may even appear that the interface probability bumps in the independent chain results are closer to the truth. This is not because they find the truth better than parallel tempering does, but merely because they have not sampled enough of the model space in an equivalent number of samples. Very similar behavior has been reported by Dosso et al. (2012).

3.3 Reducing uncertainty post-inversion and scenario evaluation

If one has used large enough prior bounds, a lot of information is contained in PDFs pertaining to subsurface resistivity. In a typical exploration scenario, as a prospect is being evaluated, more information such as seismic data, well information, or even geological models become available over the course of time. One can then incorporate this information when analyzing the posterior model distribution, without performing
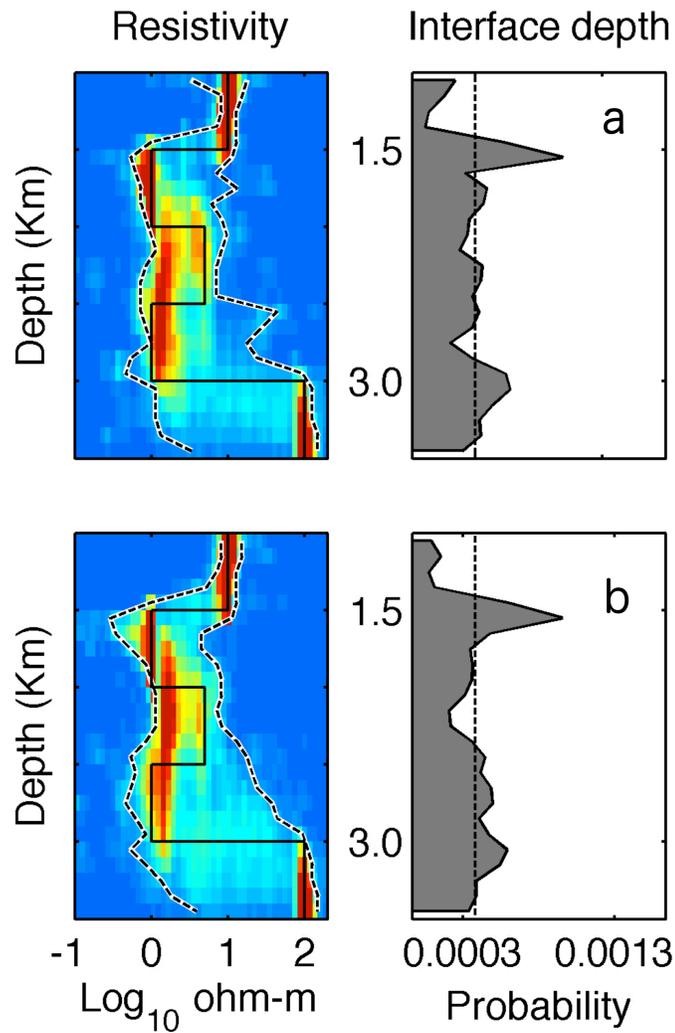
Figure 3.6: a) Posterior from 8 independent, non interacting chains, 500,000 samples in each chain. b) Posterior from parallel tempering with the same number of chains and samples, inference using only the target chain. As evidenced by the smoother 5% and 95% quantile lines of resistivity at depth, parallel tempering estimates are closer to the final sampled posterior.

further inversion. When utilizing information from geological models, one is effectively examining different geological scenarios embedded in the posterior. An example application is shown in Figures 3.7 and 3.8. In step 1, we set up reasonable uniform bounded prior distributions (Ray and Key, 2012) for resistivity at depth (Figure 3.7, left panel), interfaces at depth (Figure 3.7, right panel) and the number of layers required by a model (Figure 3.8). In step 2, we introduce the data, and through the data misfit translated into a likelihood, we sample the product of likelihood and prior, to obtain a posterior model distribution. Posterior inference is made only from the chain at the target temperature $T = 1$. Step 2 produces the marginal posterior distribution for the given prior bounds and data errors as shown in Figures 3.7 and 3.8. Step 2 in these figures represent the same case as Figures 3.4a and 3.5 for the target temperature $T = 1$. We could stop here, if this was all we knew about the subsurface, and indeed we should if that is the case. However, if new information from well logs or seismic imaging confirms the presence of certain horizons in the subsurface, we could mine the posterior model distribution using this information and go to step 3. In this step we have selected all models in the posterior ensemble which have interfaces within ± 75 m of the true location of interfaces, and the resulting posteriors from this subset have been shown in Figures 3.7 and 3.8. Note how the presence of a 5 layered structure is becoming quite apparent in Figure 3.7 step 3, and the middle layer is more likely to be resistive. All rows of Figure 3.7 are normalized to the same color scale. The last row shows the step 3 posterior overlain with examples of sampled models similar to both the true (black) as well as 'shadow' (purple) family of models, given that sampled models can have different numbers of interfaces at different depths. The 5% and 95% quantile lines have been omitted for clarity. Note the similarity of this last row to Figure 3.3b for the fixed dimensional case. The fixed dimensional case shown in Figure 3.3b is a highly specific case of Figure 3.7 step 3, and MCMC samplers are not grid searches. However, that the fixed dimensional case, which is a subset of the full posterior should be so similar to the case shown in step 3 of Figure 3.7 is indeed re-assuring.

This methodology can be encapsulated in the following manner: 1) Start by setting reasonably wide bounded prior distributions. 2) Introduce the data to obtain the full posterior. 3) Subset the data based on available information. Geological models

can also be tested in this fashion. By subsetting parts of the model, we introduce a conditioning on the posterior. This conditioning affects other parts of the posterior model distribution. Depending on the available knowledge about the subsurface, this method can be used to examine different geological hypotheses. At this juncture, all sampled likely models obey the model physics. Further, this step requires no further inversion, but simply mining of the posterior distribution (Ray and Key, 2012). We must caution though, that MCMC samplers are not exhaustive grid searches, so sampling for too specific a query may turn up no results, which does not imply that these specific cases do not exist. A corollary to post-inversion conditioning on the posterior, is that we could use a more restrictive prior distribution to reduce the model space being searched. However, this approach has a serious shortcoming, in that one can 'tune' the inversion drastically by overspecifying the prior distribution, when in reality our prior knowledge is limited. This can lead to serious bias in the resulting posterior distribution, as has been demonstrated by Minsley (2011).

3.4 Conclusions

Uncertainty is an inescapable aspect of geophysical inversion. Noisy observations, incomplete data sampling, insufficient knowledge of a suitable subsurface parameterization as well as model physics contribute to uncertainty and non uniqueness in the inverted models. In this work, we have attempted to fully address these issues using a trans-dimensional Bayesian technique. To overcome strong non-linearities in the posterior model probability distributions with multiple probability maxima, we have used parallel tempering to reliably transition between different modes. We have also demonstrated that parallel tempering speeds up convergence to the posterior distribution by reducing the total number of required samples to less than half of what would be required otherwise. We have attempted to use as few parallel chains as possible with as few forward calls per chain, in order to keep the total number of forward evaluations low. With the advent of cluster computing, computation times for accurate 2D and 3D model responses are rapidly getting smaller. However, the process of evaluating these model responses in themselves use up significant cluster resources, and it is not yet feasible, for
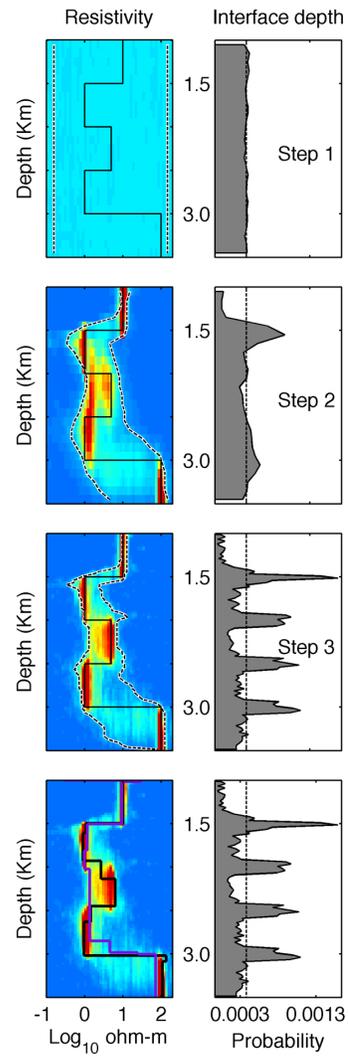
Figure 3.7: Workflow for inversion: 1) Start with broad uniform priors. 2) Introduce the data and obtain full posterior model distribution. 3) Subset posterior using available knowledge of interfaces. The last row shows the step 3 posterior overlain with sampled models that are similar to the true model (in black) and the 'shadow' model (in purple).
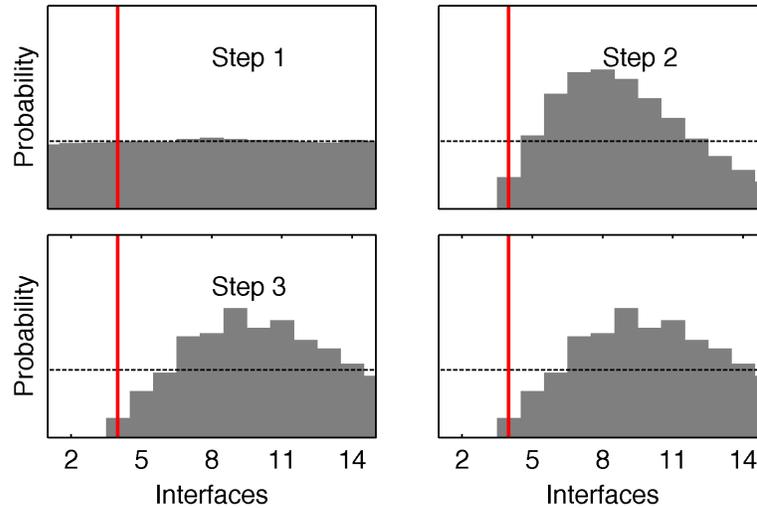
Figure 3.8: Marginal posterior probability distributions on the number of interfaces for the cases described in Figure 3.7.

tens of independent (non interacting) Markov chains, to assign each Markov chain its own cluster for forward computations. This is where parallel tempering drastically reduces the number of required chains (8 for the RJ-MCMC), yet keeps the number of forward evaluations per chain down to about 1.25 million for convergence. Given a hypothetical 2D forward evaluation time of 1 second, 1.25 million computations will take 14.5 days. While this may seem like a significant amount of time, investments for drill decisions are significantly more expensive than this amount of computer time. With advances in GPU computing and the availability of fast parallel 2.5D forward solvers (Key and Ovall, 2011), we have demonstrated that when the jump to quick, higher dimensional forward modeling is made, parallel tempering and trans-dimensional RJ-MCMC will be valuable tools with which to evaluate the full uncertainty associated with observed data.

Acknowledgements

Chapter 3, in full, is a reprint of material as it appears in Geophysics, Ray A., Alumbaugh D. L., Hoversten, G. Michael, and Key K., Society of Exploration Geophysicists, 2013. The dissertation author was the primary investigator and author of this paper.

References

Abubakar, A., Habashy, T. M., Druskin, V. L., Knizhnerman, L., and Alumbaugh, D., 2008: 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements. *Geophysics*, **73**(4), F165–F177.

Agostinetti, N. P., and Malinverno, A., 2010: Receiver function inversion by transdimensional Monte Carlo sampling. *Geophysical Journal International*, **181**, 858–872.

Backus, G. E., 1988: Bayesian inference in geomagnetism. *Geophysical Journal International*, **92**(1), 125–142.

Bayes, T., and Price, R., 1763: An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions*, **53**, 370–418. doi:10.1098/rstl.1763.0053.

Bodin, T., and Sambridge, M., 2009: Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, **178**(3), 1411–1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N., 2012: Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research*, **117**(B2).

Brodie, R. C., and Sambridge, M., 2012: Transdimensional Monte Carlo Inversion of AEM Data. In *22nd International Geophysical Conference and Exhibition, Brisbane, Australia*, 1.

Brooks, B. A., and Neil Frazer, L., 2005: Importance reweighting reduces dependence on temperature in Gibbs samplers: an application to the coseismic geodetic inverse problem. *Geophysical Journal International*, **161**(1), 12–20. ISSN 0956540X. doi: 10.1111/j.1365-246X.2005.02573.x.

Buland, A., and Kolbjornsen, O., 2012: Bayesian inversion of CSEM and magnetotelluric data. *Geophysics*, **77**(1), E33–E42.

Chen, J., Hoversten, G. M., Vasco, D., Rubin, Y., and Hou, Z., 2007: A Bayesian model for gas saturation estimation using marine seismic AVA and CSEM data. *Geophysics*, **72**(2), WA85–WA95.

Constable, S., 2010: Ten years of marine CSEM for hydrocarbon exploration. *Geophysics*, **75**(5), 75A67–75A81.

Constable, S. C., Parker, R. L., and Constable, C. G., 1987: Occam's inversion - A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(03), 289–300.

Dettmer, J., and Dosso, S. E., 2012: Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains. *The Journal of the Acoustical Society of America*, **132**(4), 2239–2250. ISSN 1520-8524. doi: 10.1121/1.4746016.

Dettmer, J., Dosso, S. E., and Holland, C. W., 2010: Trans-dimensional geoacoustic inversion. *The Journal of the Acoustical Society of America*, **128**(6), 3393–3405.

Dosso, S. E., Holland, C. W., and Sambridge, M., 2012: Parallel tempering for strongly nonlinear geoacoustic inversion. *The Journal of the Acoustical Society of America*, **132**(5), 3030–3040. ISSN 1520-8524. doi:10.1121/1.4757639.

Earl, D. J., and Deem, M. W., 2005: Parallel tempering: theory, applications, and new perspectives. *Physical chemistry chemical physics*, **7**(23), 3910–3916. ISSN 1463-9076.

Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M. C., MacGregor, L. M., and Constable, S., 2002: Remote sensing of hydrocarbon layers by seabed logging (SBL): Results from a cruise offshore Angola. *The Leading Edge*, **21**, 972–982.

Geyer, C. J., 1991: Markov chain Monte Carlo maximum likelihood. In *Proceedings of the 23rd Symposium on the Interface, New York*, 156. American Statistical Association.

Green, P. J., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.

Gunning, J., Glinsky, M. E., and Hedditch, J., 2010: Resolution and uncertainty in 1D CSEM inversion: A Bayesian approach and open-source implementation. *Geophysics*, **75**(6), F151–F171.

Hastings, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.

Key, K., 2009: 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, **74**(2), F9–F20.

Key, K., and Ovall, J., 2011: A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling. *Geophysical Journal International*, **186**(1), 137–154.

Liang, F., Liu, C., and Carroll, R. J., 2010: *Bayesian Inference and Markov Chain Monte Carlo.* John Wiley & Sons, Ltd. ISBN 9780470669723. doi: 10.1002/9780470669723.ch1.

Loseth, L. O., Pedersen, H. M., Ursin, B., Amundsen, L., and Ellingsrud, S., 2006: Low-frequency electromagnetic fields in applied geophysics: Waves or diffusion? *Geophysics*, **71**(4), W29–W40.

MacKay, D. J. C., 2003: *Information theory, inference and learning algorithms.* Cambridge university press.

Malinverno, A., 2002: Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, **151**(3), 675–688.

Minsley, B. J., 2011: A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. *Geophysical Journal International*, **187**(1), 252–272. ISSN 0956540X. doi:10.1111/j.1365-246X.2011.05165.x.

Newman, G. A., and Alumbaugh, D. L., 2000: Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, **140**(2), 410–424.

Ray, A., and Key, K., 2012: Bayesian inversion of marine CSEM data with a trans-dimensional self parametrizing algorithm. *Geophysical Journal International*, **191**, 1135–1151.

Sambridge, M., Bodin, T., Gallagher, K., and Tkalčić, H., 2013: Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A*, (December 2012).

Sasaki, Y., 2013: 3D inversion of marine CSEM and MT data : An approach to shallow-water problem. *Geophysics*, **78**(1), E59–E65.

Scales, J. A., and Sneider, R., 1997: To Bayes or not to Bayes? *Geophysics*, **62**(4), 1045–1046.

Swendsen, R. H., and Wang, J. S., 1987: Nonuniversal Critical Dynamics in Monte Carlo Simulations. *Physical Review Letters*, **58**(2), 86–88.

Tarantola, A., and Valette, B., 1982: Inverse problems= quest for information. *J geophys*, **50**, 159–170.

Tompkins, M. J., Fernandez Martinez, J. L., Alumbaugh, D. L., and Mukerji, T., 2011: Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: Marine controlled-source electromagnetic examples. *Geophysics*, **76**(4), F263–F281.

Trainor, W., and Hoversten, G. M., 2009: Practical challenges of stochastic inversion implementation for geophysical problems. In *2009 SEG Annual Meeting*.

Trainor-Guitton, W., and Hoversten, G. M., 2011: Stochastic inversion for electromagnetic geophysics: Practical challenges and improving convergence efficiency. *Geophysics*, **76**(6), F373–F386.

Weiss, C., 2007: The fallacy of the "shallow-water problem" in marine CSEM exploration. *Geophysics*, **72**(6), A93–A97.

Young, P. D., and Cox, C. S., 1981: Electromagnetic active source sounding near the East Pacific Rise. *Geophysical Research Letters*, **8**, 1043–1046.

Chapter 4

Bayesian 2D trans-dimensional inversion of marine CSEM data from the Scarborough gas field

Anandaroop Ray[1], Kerry Key[1], Thomas Bodin[2], David Myer[3] and Steven
Constable[1]

[1] Scripps Institution of Oceanography, La Jolla, CA, USA
[2] University of California Berkeley, Berkeley, CA, USA
[3] BlueGreen Geophysics, Encinitas, CA, USA

Abstract. We apply a reversible jump Markov chain Monte Carlo method to sample the Bayesian posterior model distribution of 2D seafloor resistivity as constrained by marine controlled source electromagnetic data. This posterior distribution of earth models conveys information on which parts of the model space are illuminated by the data. Whereas conventional gradient based inversion approaches require subjective regularization choices to stabilize this highly non-linear and non-unique inverse problem and provide only a single solution with no model uncertainty information, the method we use entirely avoids model regularization. The result of our approach is an ensemble of models that can be visualized and queried to provide meaningful information about

the sensitivity of the data to the subsurface, and the level of resolution of model parameters. We represent models in 2D using a Voronoi cell parameterization. To make the 2D problem practical we use a source-receiver common midpoint approximation with 1D forward modelling. Our algorithm is trans-dimensional and self parameterizing where the number of resistivity cells within a 2D depth section is variable, as are their positions and geometries. Two synthetic studies demonstrate the algorithm's use in the appraisal of a thin, segmented, resistive reservoir which makes for a challenging exploration target. As a demonstration example we apply our method to survey data collected over the Scarborough gas field on the Northwest Australian shelf.

4.1 Introduction

The marine controlled source electromagnetic (CSEM) method is an active source sounding method that has been in use for over three decades for the detection of geology with high resistivity contrasts (Young and Cox, 1981; Chave and Cox, 1982). Industry funded research and extensive commercialization of this technology over the last decade has led to CSEM being added to the standard suite of seismic methods in an exploration scenario (Ellingsrud et al., 2002; Constable, 2010). Conductive media such as sea-water or brine filled sediments have a characteristic electromagnetic scale length (skin depth) $\delta = \sqrt{\frac{2}{\mu\omega\sigma}}$ that is dependent on both the medium conductivity σ and the frequency of propagation ω, where μ is the permeability of the medium. Owing to the fact that δ is smaller in conductive media, marine geophysical EM methods almost always operate in the lower frequency quasi-static regime. This allows for deeper penetration of the CSEM fields into the earth, but as a consequence it is more a diffusive process than wave like (Loseth et al., 2006). To first order, it is this diffusive decay which helps characterize the conductivity of a given medium. For hydrocarbon bearing geology, it is the high resistivity of the hydrocarbon accumulation with respect to its surroundings that produces a signature quite different from what would have been observed in the absence of hydrocarbons (e.g., Constable, 2006).

Given this diffusive nature, the stratigraphic resolution of CSEM is much lower than that of the seismic method. However, the value of CSEM lies in its sensitivity to resistivity, which may be indicative of hydrocarbon saturation, and not acoustic impedance, which is more indicative of geological structure. As a consequence of this diffusive nature, robust inferences made from a CSEM survey are necessarily from inversion of the data, and not merely from examination of the data itself (Weiss, 2007). Typically, regularized and linearized gradient based inversion methods have been used to arrive at models that minimize data misfit and are also 'optimal' in some user-defined sense. For instance, models can be pre-determinedly smooth or prejudiced to be close to a reference model. By means of regularization, highly oscillatory features in the model that are thought to be outside the resolution of CSEM are suppressed (e.g., Constable et al., 1987; Newman and Alumbaugh, 2000; MacGregor and Sinha, 2000; Abubakar et al., 2008; Key, 2009; Sasaki, 2013; Mittet and Gabrielsen, 2013). Though gradient based inversion methods

methods are highly efficient and well understood, they provide only a single smooth model as a result, or a suite of smooth models. These models provide a limited insight into the various classes of models that are compatible with the observed data given the noise. Furthermore, a clear understanding of the resolvability of subsurface resistivity and non-uniqueness of the final solution does not emerge from a linearized treatment of the non-linear CSEM problem.

To quantify the uncertainty inherent in the inversion of CSEM data, one can utilize a Bayesian framework where information is expressed as probability density functions or PDFs. Since Bayesian probability (Bayes and Price, 1763) is a measure of information and since it is the aim of geophysical inversion to provide information about the earth's subsurface, it is natural to postulate geophysical inverse problems in a Bayesian framework (Tarantola and Valette, 1982; Scales and Sneider, 1997). In such a framework, model parameters are treated as random variables, and their fit to the observed data given the observed statistical noise allows one to formulate a model likelihood. To make the connection with deterministic inversion methods, to first order, models with low misfit possess a higher likelihood. After incorporating prior knowledge of the models that is independent of the data, the product of the prior model probability and the likelihood is known as the posterior model probability. This posterior PDF describes the full solution to the inverse problem — it represents the probability of the model, given the observed data. Those parts of the model space that are more frequently required by the data than other parts manifest with greater posterior probability, and hence are more certain to be properties of the earth (Backus, 1988).

Our study is not the first to apply Bayesian methods for inversion of marine CSEM data. One of the earliest applications focused on joint inversion of CSEM and seismic data in order to improve estimates of reservoir properties (Hou et al., 2006; Chen et al., 2007). Gunning et al. (2010) use a hierarchical Bayesianized bootstrap scheme for CSEM inversion. Trainor-Guitton and Hoversten (2011) use a sampling scheme which involves both the Metropolis-Hastings algorithm (Hastings, 1970) and slice sampling (Neal, 2003) in order to improve convergence upon the distribution of solution models. Buland and Kolbjornsen (2012) apply the Metropolis-Hastings algorithm to invert marine CSEM data together with magnetotelluric (MT) data in order to constrain

the range of likely resistivities as a function of depth. In all these studies, with the exception of Gunning et al. (2010), the parameterization is fixed at the outset by the user. However, fixing a particular model parameterization (e.g., fixing the number of layers or cells) for the inversion is known to produce posterior distributions, *only* for the given parametrization (Dettmer et al., 2010). This is where the 'trans-dimensional' or 'reversible jump' (Green, 1995) Markov Chain Monte Carlo (RJ-MCMC) differs from traditional MCMC methods, in sampling from a posterior distribution where the number of unknowns (i.e., the parameterization) is also treated as part of the inverse problem. In other words, the parameterization is also inferred from the observed data. A review of applications which use trans-dimensional MCMC can be found in Sisson (2005). Such algorithms have a 'parsimony' property (Malinverno, 2002), which refers to the fact that Bayes' Theorem deems models that explain the data with simpler parameterizations more probable. MacKay (2003) discusses this aspect of Bayes' Theorem in some detail. Malinverno (2002) was the first to use this method in a geophysical application for DC resistivity inversion. Agostinetti and Malinverno (2010) have used this method for receiver function inversion, as have Bodin et al. (2012). Recent applications of the trans-dimensional method to solve geophysical EM methods can be found in Minsley (2011) and Brodie and Sambridge (2012), who apply it the airborne EM problem, and in Ray and Key (2012), who tackle the marine CSEM problem. An introduction to geophysical trans-dimensional Bayesian inversion can be found in Sambridge et al. (2013).

For probabilistic inversion, the final solution is a large ensemble of models which are statistically distributed according to a posterior model PDF. As this involves a direct parameter search, hundreds of thousands of models must be evaluated and sampled before convergence to the desired posterior model probability PDF. This computational expense has largely limited the application of Bayesian methods for highly non-linear problems to those with a computationally efficient 1D model parameterization. Notable exceptions can be found in the work of Chen et al. (2012) and Rosas-Carbajal et al. (2013). However, though both of these works invert 2D MT and plane-wave electrical resistivity tomography (ERT) data respectively, they use a fixed number of parameters. JafarGandomi and Binley (2013) use a trans-dimensional approach to invert multiple datasets within a 2D depth section, but only the vertical parametrization is

trans-dimensional while the lateral parameterization is fixed. Fully 2-dimensionally pa-
rameterized, trans-dimensional inversions have been carried out only within the last five
years. For example, Bodin and Sambridge (2009) perform seismic surface wave tomogra-
phy using Voronoi cells in a trans-dimensional formulation. Luo (2010) finds the shapes
of bodies which cause a gravity anomaly (given a fixed density contrast). Young et al.
(2013a,b) use trans-dimensional inversion for P wave tomography and seismic ambient
noise inversion. Dettmer and Dosso (2013) use this approach for geo-acoustic inversion.
To the best of our knowledge this work is the first trans-dimensional Bayesian inver-
sion that uses a true 2D model parameterization with Voronoi cells for the inversion of
geophysical EM data.

After validating our methodology using synthetic examples, we invert CSEM
data from over the Scarborough gas field on the Exmouth Plateau, off the North West
Australian shelf (Myer et al., 2010, 2012).

4.2 Theory

4.2.1 Bayesian inversion with the reversible jump Markov chain Monte Carlo algorithm

Bayesian information is contained in PDFs represented by $p(\cdot)$. Using Bayes'
theorem, (Bayes and Price, 1763) we write

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m})}{p(\mathbf{d})}. \tag{4.1}$$

This can be interpreted as

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior assumptions}}{\text{evidence}}. \tag{4.2}$$

For Bayesian geophysical inversion, the observed data vector $\mathbf{d}$ is a constant.
All PDFs with a model dependence are functions of the random variable $\mathbf{m}$. The term
$p(\mathbf{d}|\mathbf{m})$ can then be interpreted as the model likelihood, the functional form of which
depends on the statistics of the noise distribution, and the value of which depends on

the model $\mathbf{m}$ being sampled and its fit to observed data. For Gaussian noise, the model likelihood is given as:

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(- \frac{[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1} [\mathbf{d} - f(\mathbf{m})]}{2} \right). \tag{4.3}$$

Here $f(\mathbf{m})$ corresponds to the modeled data and $\mathbf{C}_d$ is the covariance matrix of data errors and $[\mathbf{d} - f(\mathbf{m})]^T \mathbf{C_d}^{-1} [\mathbf{d} - f(\mathbf{m})]$ is the χ^2 misfit for the evaluated model $\mathbf{m}$. Care must be taken to ensure that for complex observed data (such as frequency domain data) where the total variance equals twice that of the real or imaginary parts, the factor of 2 should be removed from equation (4.3) and that the vector transpose should be Hermitian (see the chapter on circularly symmetric complex random variables in Vaidyanathan et al., 2010). The prior model distribution $p(\mathbf{m})$ represents our state of knowledge *independent* of the survey data. The evidence term $p(\mathbf{d})$ corresponds to a constant PDF normalizing factor equal to the integral over all models of the numerator in equation (4.1). The evidence can help in 'model selection,' the process of deciding which model parameterization is more probable than the other — e.g., should we use two cells as opposed to three to represent the subsurface? However, evidence is very challenging to compute as it requires evaluation of a multi-dimensional integral over different models, evaluated for different model parameterizations. Another means of performing model selection is to calculate the full posterior model probability distribution $p(\mathbf{m}|\mathbf{d})$ by allowing the problem to be trans-dimensional, i.e., have a varying number of model parameters (Dettmer et al., 2010). This is the RJ-MCMC approach that we have used in this paper, which is different from the usual MCMC approach in the following manner: Treating the evidence as a proportionality constant, it follows from equation 4.1 that

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m}). \tag{4.4}$$

In the trans-dimensional method, for a given model $\mathbf{m}$, we split the prior distribution $p(\mathbf{m})$ into 2 parts. One part contains information about the number of cells k in the model, $p(k)$. The other part $p(\mathbf{m_k}|k)$ contains information about their physical properties such as the locations of the cells and what the resistivities of these cells are, given the

number of cells k. Thus it follows from equation 4.4 that

$$p(\mathbf{m}|\mathbf{d}) \propto p(\mathbf{d}|\mathbf{m}) \times p(\mathbf{m_k}|k) \times p(k). \tag{4.5}$$

This formulation allows the parameterization to be an unknown in the problem. Our task is to evaluate uncertainty in the models inverted from the observed data. To this end we must arrive at the posterior distribution of models, most of which fit the data well, by evaluating their misfit and sampling models according to equation 4.5. However, it is nearly impossible to exhaustively sample the model space for more than a few parameters owing to the 'curse of dimensionality,' hence we resort to probing this highly non-linear distribution using various Markov chain Monte Carlo methods (e.g., Liang et al., 2011) and focus on the RJ-MCMC or trans-dimensional method.

4.2.2 Parametrization with Voronoi cells

Voronoi cells are an efficient topology for parameterizing a 2D space (Voronoi, 1908). They can represent various complicated geometries, and their tessellations are simple to compute (Okabe et al., 2009). We can specify a set of points (Voronoi nodes) in a bounded plane, and for each node there will be a corresponding region consisting of all points closer to that node than to any other node. These regions are called Voronoi cells. For geophysical problems, this parameterization presents itself as a convenient means to assign Voronoi cells with properties such as velocity (Bodin and Sambridge, 2009; Dettmer and Dosso, 2013) or as we propose to do in this work, resistivity. Only the Voronoi nodes and their locations need to be kept track of. This is shown in Fig. 4.1, where a tabular body of resistivity 100 ohm-m in a 1 ohm-m background has been approximated almost perfectly by a sparse Voronoi representation with 5 cells. It should be noted that this Voronoi diagram is not the solution to a geophysical inverse problem but merely an example illustrating its use in simple representations of common shapes.

4.2.3 Forward modelling and parameterization

For a given transmitter-receiver (Tx-Rx) separation, we use the approximation that observed data are primarily sensitive to the the vertical profile of seafloor resistivity
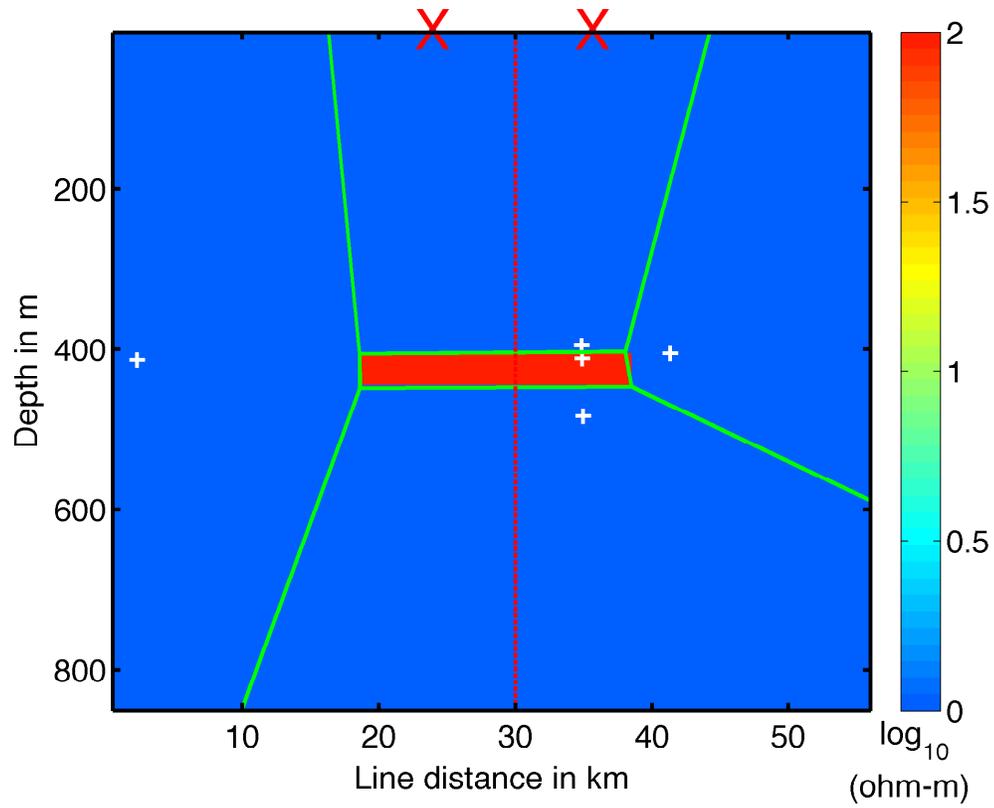
Figure 4.1: Five Voronoi cells being used to approximate a rectangular reservoir in a uniform background. The nodes are shown with white plus signs. The cells are entirely defined by their node positions, as the edges (green lines) run through the perpendicular bisectors between neighboring nodes. A profile through the reservoir (dashed red) is located at a source-receiver midpoint with source and receiver locations each marked by an X.

at the Tx-Rx mid point (Mittet et al., 2008; Gunning et al., 2010; Silva Crepaldi et al., 2011). This is similar to a common mid-point formulation (CMP) used in reflection seismology. This 1D approximation is numerically far less expensive than modelling the full 2D problem. A 2D model forward evaluation, depending on the geological complexity of the model, the number of frequencies and parallel cores used, may take from a few seconds to minutes (Key and Ovall, 2011). However, 1D evaluations across all CMPs for a 50 km long 2D line takes little more than half a second on average. We use 2D Voronoi cells to parameterize the vertical cross section of the earth in a profile along the survey line (e.g., Dettmer and Dosso, 2013). For each CMP, a vertical profile is extracted through this Voronoi cell parameterization as shown in Fig. 4.1. The response due to this 1D profile, at the given Tx-Rx offset and data frequencies is calculated using Dipole1D (Key, 2009). Since we use Voronoi cells to parameterize a 2D depth section, a vertical profile can be extracted anywhere within this section. This ensures that there is no restriction on locations of sources or receivers along a linear profile.

4.2.4 Trans-dimensional Bayesian inversion

In the simplest terms, the objective of trans-dimensional Bayesian inversion is to sample the distribution given by (4.5). This is achieved with the reversible jump algorithm by drawing candidate models from a proposal distribution. These models are then examined to see if they fall within geophysically sensible uniform bounds of resistivity and depth. If a proposed model falls outside the prior distribution, it is rejected and the Markov chain retains the previous model as the next model. If a proposed model is within the prior bounds, an acceptance probability is calculated using a ratio of the proposal probability, the prior probability and the likelihood of the candidate model with respect to the previous model. The proposed model is either accepted with the calculated probability and it becomes the next model in the chain, or it is rejected and the previous model is retained as the next in the chain. Complete details of this process are given in the appendix for the interested reader. As the algorithm proceeds, hundreds of thousands of models are sampled, with a data-driven addition or deletion of Voronoi nodes ('birth/death' in RJ-MCMC parlance), such that a chain of models, most of which fit the data well within the noise, are retained at the end. To ensure thorough

sampling of this multi-dimensional parameter space, various independent Markov chains with different starting models are run in parallel and finally concatenated into a model ensemble that is representative of the sought after posterior model distribution (4.5).

4.3   2D Segmented reservoir example

In order to validate our methodology, we created a 60 km long synthetic model, with a 1 km deep sea water layer and a 30 km long, 30 m thick segmented reservoir buried at 1 km below the seafloor (Fig. 4.2). The resistivity segments for the reservoir are 30, 10 and 5 ohm-m, tapering away from the centre. The background resistivity is 1 ohm-m. Given that the reservoir is thin and that the resistivity contrasts are not very high, this presents a challenging exploration target. This geometry is motivated by the size and shape of the Scarborough reservoir in the NW Australian shelf (Myer et al., 2010), which we consider later in this work. Synthetic data were forward calculated using the 2.5D adaptive finite element code of Key and Ovall (2011). Gaussian noise at 5% of the signal amplitude was added to the data at four different frequencies of 0.1, 0.3, 0.7 and 1.1 Hz. Receivers recording the inline electric field were spaced every 500 m along the line.

4.3.1   1D modeling and inversion through the central segment

Before embarking on the inversion of 2D data, we examine a 1D Bayesian inversion of 1D responses from the central part of the segmented reservoir model to highlight some salient features of Bayesian inversion. We use the 1D CSEM code and inversion algorithm described in Ray and Key (2012). 96 independent, parallel MCMC chains with 500,000 models in each chain were sampled to form a posterior model distribution. This inversion required four hours on 96 CPU cores (dual socket @2.6GHz). 50 random models from the sampled posterior, along with their data fits in phase and amplitude are shown in Fig. 4.3. The true model is shown in red in the left most panel.

To visualize the full posterior distribution, we separately bin the resistivities and interfaces at 10 m depth intervals. The resulting marginal distributions on resistivity and interfaces at depth are shown in Fig. 4.4. The 5% and 95% quantiles on resistivity
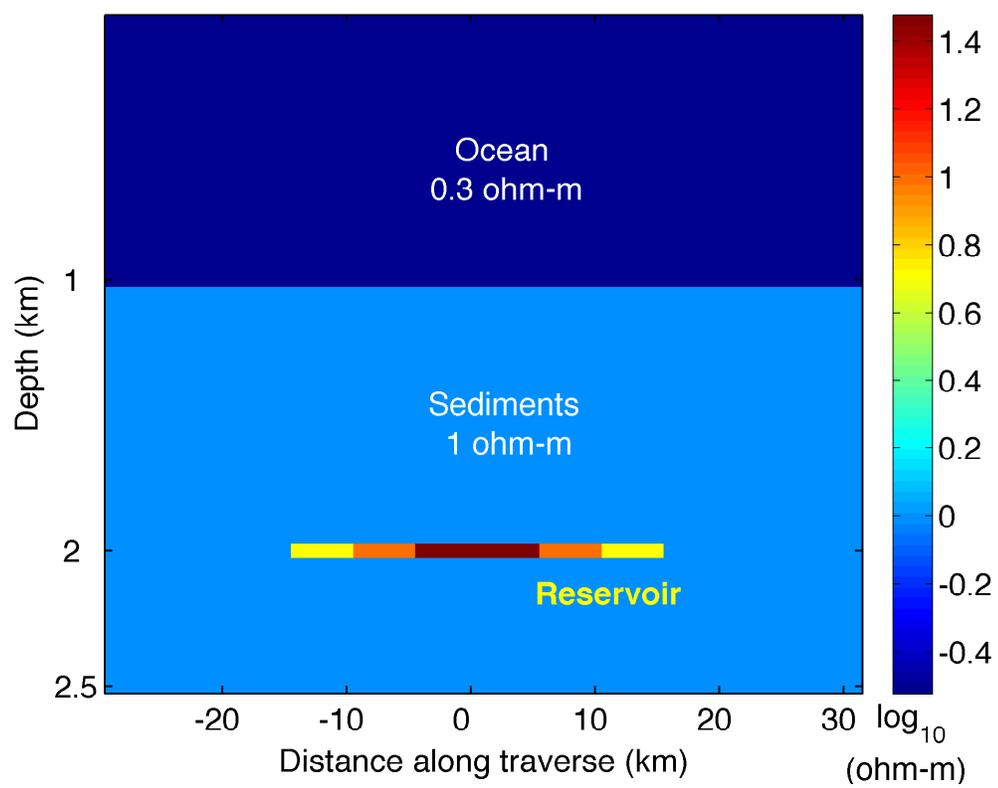
Figure 4.2: A 30 km long, 30 m thin synthetic segmented resistor model representative of the Scarborough reservoir.
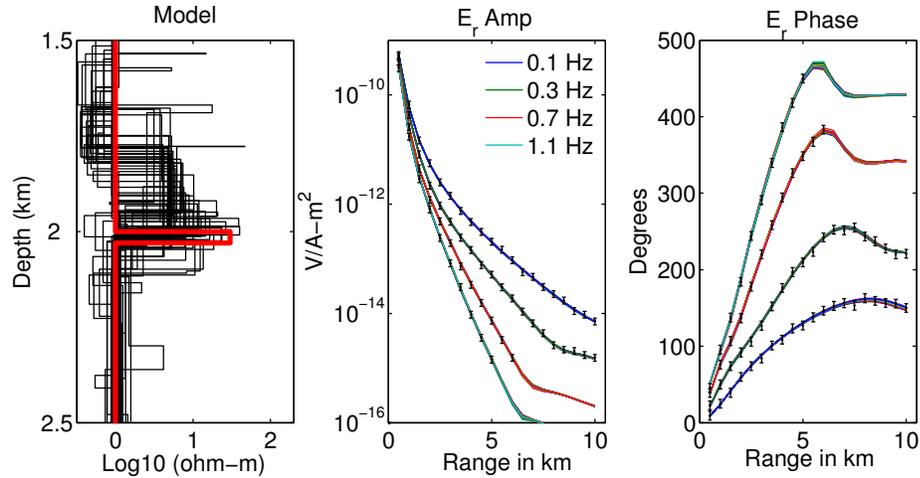
Figure 4.3: Inversion of noisy, synthetic radial electric field data E_r calculated from the 1D reservoir corresponding to the centre of the 2D segmented reservoir model. The true 1D model is shown in red in the left panel, with 50 randomly selected models from the posterior distribution $p(\mathbf{m}|\mathbf{d})$ shown in black. The responses due to these models, along with the observed data are shown as E field amplitude (centre) and phase (right) plots with range.

with depth are shown in the left panel as thin black lines. Hotter colors correspond to higher probabilities in the left panel. It can be seen that the true reservoir resistivity is not the most probable, that it is even outside the 95% quantile, but we do see a clear trend in sampled resistivities increasing with depth and then returning to a background value. The fact that the true values are not the most probable, though not intuitive, is not new as has been shown by Ray et al. (2013). This should serve as a note of caution to prevent researchers from picking the mode of any posterior distribution as the 'truth.' This seems to be a statement of the fact that there are many more different ways to fit the data within the noise, than with the very delta function like true model. In fact, one can clearly see the well known CSEM trade-off between thicker, less resistive layers and thinner, more resistive layers (e.g., Constable and Weiss, 2006). It is worthwhile to point out that such trade-offs cannot be quantified by any single result from a conventional gradient based deterministic inversion scheme. The interface probabilities in the right panel do show a pronounced peak in the probability of interfaces at the right depths. For a thin reservoir, this seems to imply good sensitivity to the bottom.
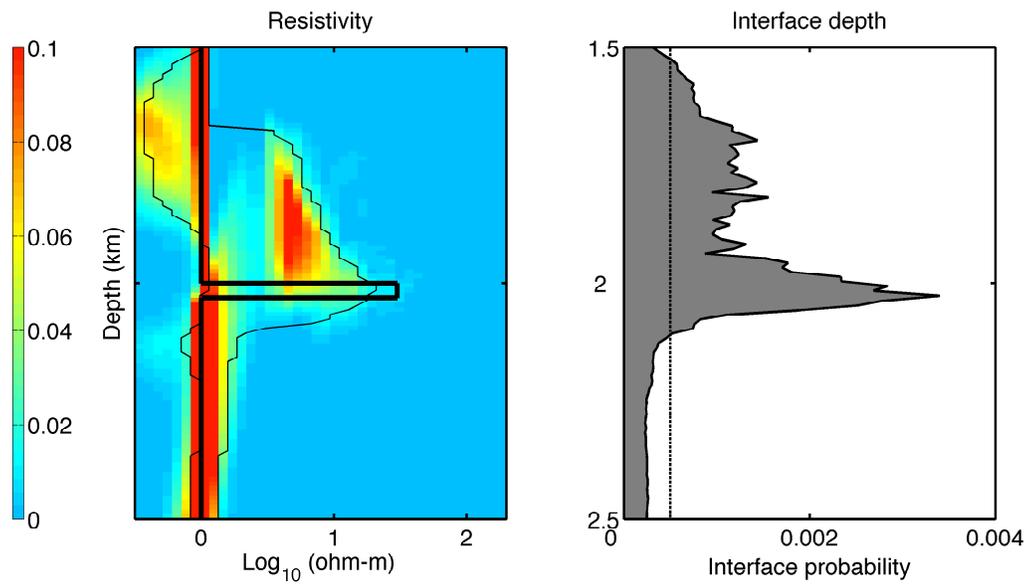
Figure 4.4: Marginal distributions on resistivity (left) and interface depth (right) from the 1D synthetic data inversion with true model corresponding to the 1D segment through the centre of the 2D synthetic model. The true model is shown with a thick black line in the left panel. The 5% and 95% quantiles on resistivity with depth are shown as thin black lines on the left panel. Hotter colours are more probable.

4.3.2 Inversion of 2D data with the CMP approach

In this section we present the inversion of the noisy synthetic 2D data from the segmented reservoir model using the trans-dimensional Voronoi parameterization and the CMP approach. During the inversion we calculate the forward responses for 1D profiles through the Voronoi tessellation at every CMP location. 96 independent, parallel RJ-MCMC chains were run on 96 CPU cores (dual socket @2.6GHz), out of which 92 chains converged to acceptable RMS values. We would like to point out here that not all Markov chains are able to escape local probability maxima within the given run time. Each chain sampled 10^6 models but the first half of the chain is thrown away to achieve 'burn-in' to RMS (root mean square) values between 1.2 and 1.3. Given that the model physics is approximate and parameterization not the same as was used to create the forward model, we do not expect the mean RMS misfit of the sampled models to be exactly 1.0. Uniform prior bounds require models to possess resistivities between 0.3 and 200 ohm-m, and from 10 to 150 nodes placed between 1 and 3 km depth anywhere within the 60 km long 2D line.

To illustrate the process of forming the posterior model ensemble, one MCMC chain is shown in Fig. 4.5. All converged chains can be concatenated to form a single, long chain like the one shown in the figure. One slice shows an arbitrary 2D model parameterized by Voronoi cells. The other perpendicular slice, near the beginning of the line, shows resistivities in vertical section across all sampled models. A histogram of resistivities with depth can be made from the vertical section to produce a display akin to the left pane in Fig. 4.4. Further, such histograms can be made from all vertical slices along the 2D line to form a probability cube, with axes being resistivity, line distance and depth, as shown in Fig. 4.6. A vertical slice of the probability cube through the centre of the 2D line (the central segment of the reservoir) shows that there is a clear rise in the resistivity of the subsurface with depth and then a return to the background. A horizontal slice at 2000 m depth shows a clear increase in resistivities from a 1 ohm-m background as we move along the survey line, increasing in steps and symmetrically decreasing as we would expect for the segmented reservoir. Also shown in this figure is the PDF on the number of Voronoi cells (nodes) required to form the displayed probability cube. Note how the algorithm does not cluster around the maximum or minimum possible number

of nodes (10 and 150) required to fit the observed data.

Based on the experience of our 1D modeling study showing the posterior trade-off between resistivity and reservoir thickness, we would not expect the most probable resistivities sampled to be the true values, and this is indeed the case. The maximum sampled resistivity (Fig. 4.6) is around 10 ohm-m at the reservoir depth in the central segment, which is quite distant from the true value of 30 ohm-m. However, the probability of the integrated resistivity thickness product $\tau = \int \rho(z)dz$ between 1500 m and 2500 m depth, containing the reservoir interval, yields some valuable insights, as can be seen in Fig. 4.7. True values are shown with black lines, and the background and segmented reservoir τ values are clearly visible as falling within the probable parts of the marginal distribution. Previous studies have shown that τ seems to be a more robust indicator of reservoir presence than resistivity alone (Constable and Weiss, 2006; Myer et al., 2012; Connell and Key, 2013). The posterior probabilities on this quantity bolster this idea. For instance, even if we were to pick the mode of this distribution on τ in Fig. 4.7 near the centre of the line, though 1700 ohm-m^2 is less than the true value of 1870 ohm-m^2, it would not lead to a bad estimate of the reservoir resistivity. Assuming a 30 m thick resistive layer and a 1 ohm-m background, this τ value is consistent with a 24 ohm-m reservoir resistivity. Often such information on thickness is available from an external source such as seismic imaging. Conversely, if the true resistivity of 30 ohm-m was known, for example from well-logs, we could then estimate a probable thickness of 24 m which is not too far from the true value of 30 m.

4.4 Real Data: Scarborough Gas Field on the Exmouth Plateau, NW Australian Shelf

4.4.1 Regional geology and reservoir setting

In this section we apply our methodology to data from the Scarborough gas field, which lies inside the Exmouth Plateau. The plateau (Fig. 4.8) is a passive margin between continental and oceanic crust which remains after the break-up of Australia and India, and is surrounded on three sides by oceanic crust at abyssal depths. The plateau,
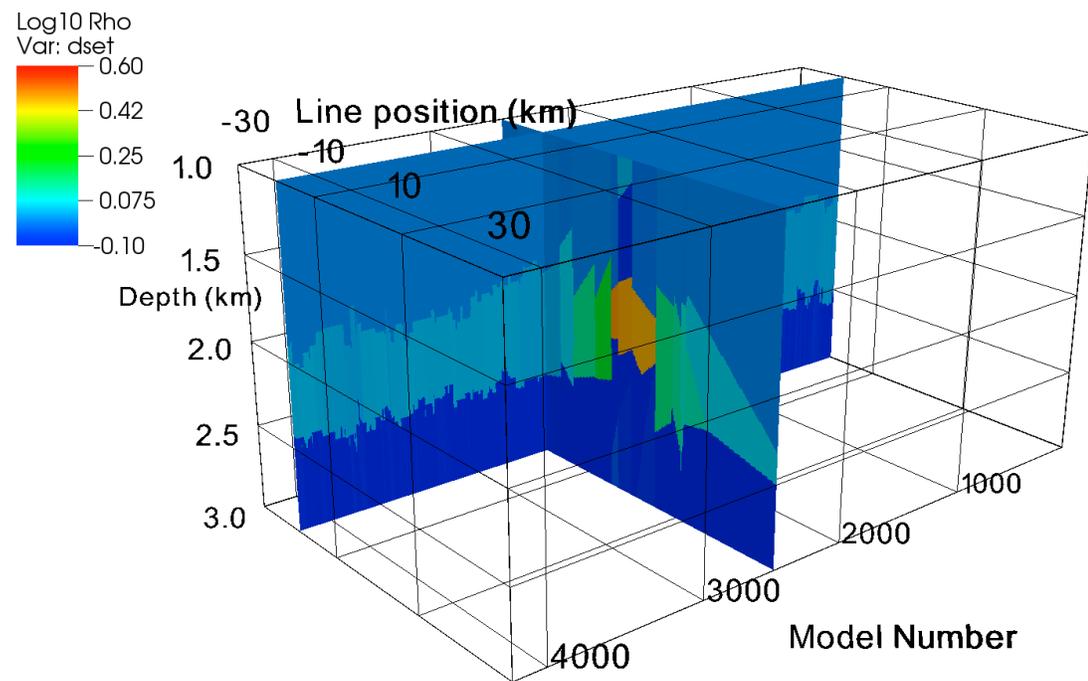
Figure 4.5: Example MCMC chain from the segmented reservoir inversion. Here we show both a particular Voronoi cell model #2500 and a vertical slice across all models at -20 km along the line. Colours correspond to $\log_{10}$ of resistivity.
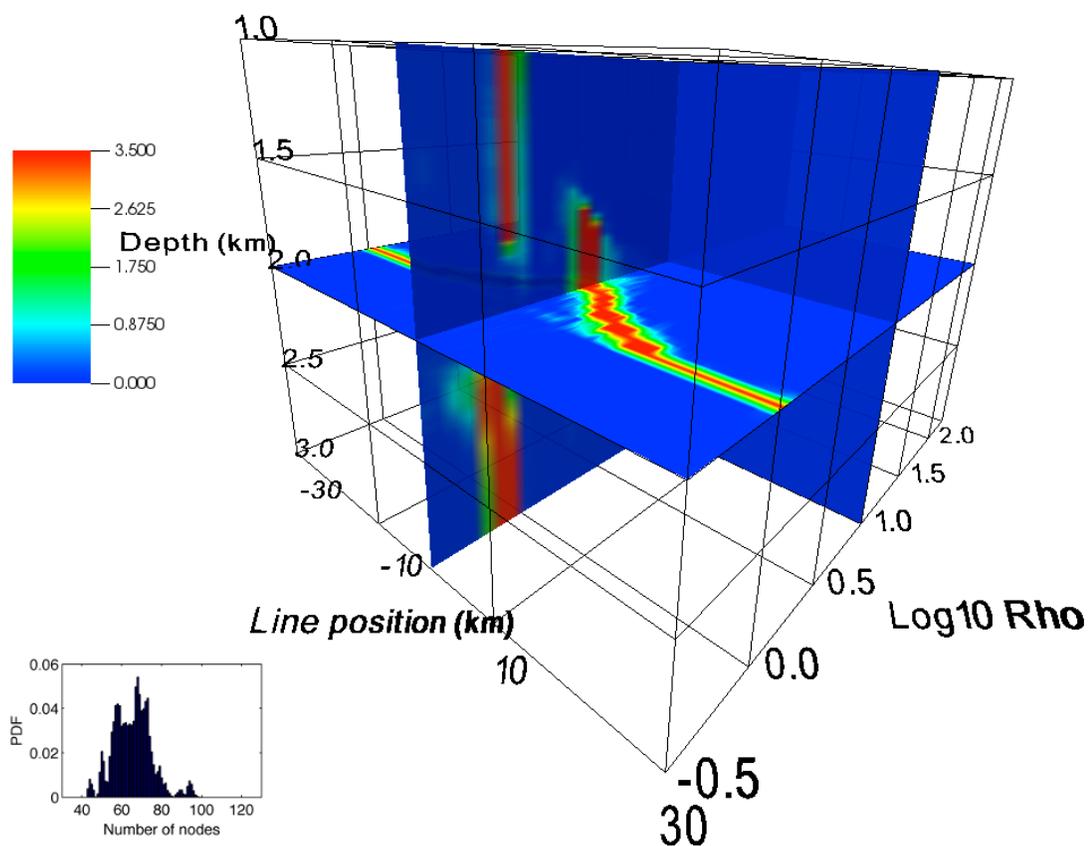
Figure 4.6: Slices through a probability cube from the segmented reservoir inversion, with axes representing $\log_{10}$ resistivity, position along the 2D line and depth. The probability cube is made from histograms of all MCMC chains such as the one shown in Fig. 4.5. Hotter colours correspond to higher probabilities. The inset box shows the PDF on the number of Voronoi nodes required to form the probability cube.
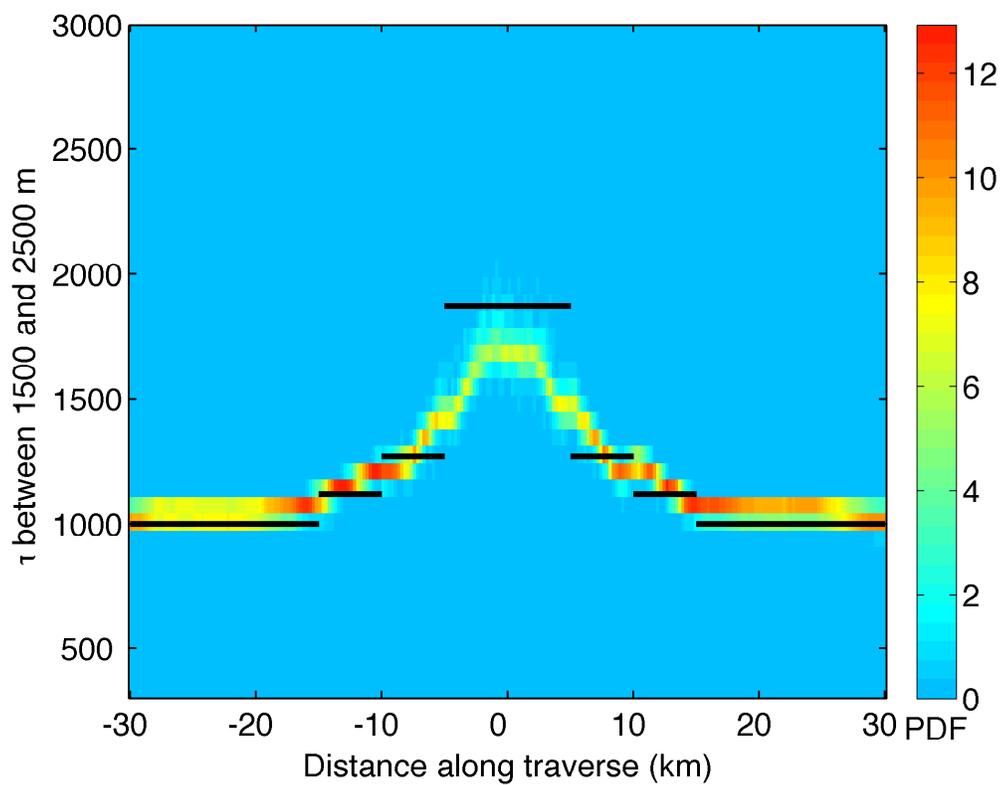
Figure 4.7: Probability of the integrated resistivity thickness product τ between 1500 and 2500 m depth from the segmented reservoir inversion. Black lines show the true values corresponding to the background and reservoir segments. Hotter colours are more probable.

which is $\sim$ 400 km $\times$ 600 km, is bounded to the northeast and southwest by transform faults. The transition between continental and oceanic crust to the northwest is thought to be bounded by a subhorizontal detachment fault that undercuts the plateau at about 10 km depth, dipping toward the Australian continent (Driscoll and Karner, 1998). Since the Mesozoic era, the plateau has undergone a complex sequence of fracture, extension, uplift, truncation, and subsidence. The result is that the plateau is covered by a number of mostly horizontal sedimentary layers of resistivities varying between 1 and 10 ohm-m (Myer et al., 2012).

Five exploration wells have been drilled in the Scarborough gas field and their data, combined with 3D seismic coverage, were used to define the areal extent and section profile of the reservoir. The white contour in Fig. 4.8 is the 50% gas saturation line. The reservoir itself (Fig. 4.9) is a 20–30 m layer residing between 1900 and 2000 m below sea level (mbsl) in about 900–950 m of water. It has a moderate resistivity of 25 ohm-m and is overlain by several thin layers of lower gas saturation with resistivities of 5–10 ohm-m.

4.4.2   Scripps 2009 survey and previous work in 1D

CSEM and MT data were collected during a month long research cruise using the Scripps Institution of Oceanography's R/V Roger Revelle (Fig. 4.8). Details of the CSEM acquisition and MT mapping of a deep sub-horizontal conductive layer can be found in Myer et al. (2011, 2012) and Myer et al. (2013), respectively. In this work we have concentrated our efforts on inverting the inline electric field CSEM data at frequencies of 0.25, 0.75, 1.75 and 3.25 Hz acquired over the $\sim$50 km long Line 2 to the south.

The reservoir is not a large CSEM target, as the τ value is only $\sim$900 ohm-m^2 above the background levels of $\sim$200 ohm-m^2 for the reservoir section. This difference is much less than typical transverse resistances of 10^4 considered in many past marine CSEM model studies (Constable and Weiss, 2006; Myer et al., 2012). Further, there is a confounding overlying layer in the form of the resistive Gearle siltstones between 1650 and 1750 mbsl (Veevers and Johnstone, 1974), which may make it difficult to tease apart the reservoir layer and the Gearle siltstones without *a priori* information, as reported by Myer et al. (2012) and illustrated in Fig. 4.10. Without introducing cuts in

the inversion to separately delineate the siltstones and the reservoir, the Occam1DCSEM inversion (thick black line) was unable to distinguish these two layers from a single layer of moderately increased resistivity near reservoir depth. For 1D comparison, we performed purely 1D Bayesian inversions using the methodology of Ray and Key (2012) at both the on and off reservoir sites (Figure 4.10). These inversions suggest that the probability distribution of resistivities at reservoir depth on-reservoir (bottom, left panel)) is more resistive than the distribution at the same depth off-reservoir (top, left panel). Note that the 1D RJ-MCMC Bayesian inversions did not utilize any prejudice models or roughness penalty cuts.

The separate trans-dimensional Bayesian results shown in Figure 4.10 do indeed indicate that it is more probable to have a resistive reservoir at the on reservoir site than at the off reservoir site. However, these results use a 1D parameterization, and in no way can distributions of resistivity at the two sites be correlated *a posteriori*. No inference can be made about how the PDFs of resistivity vary laterally across the line. We could perform numerous independent 1D RJ-MCMC inversions at all sites across the 2D line, but they would not need to agree between adjacent sites, and such an exercise would require significant computational expense. These considerations motivate the next part of this section on inverting the Scarborough CSEM data along Line 2 using a fully 2D model parametrization and the RJ-MCMC method.

4.4.3   Results from 2D trans-dimensional inversion using the CMP approach

An inversion with 160 independent parallel chains on 160 CPU cores (dual socket @2.6GHz) was run on CSEM data acquired along Line 2. Four frequencies at 0.25, 0.75, 1.75 and 3.25 Hz were used, but Rx-Tx offsets were limited to 4 km for the observed data to be compatible with the 1D forward model physics. Given that variations in bathymetry along the line are minimal (see Fig. 4.8), the 2D models used a flat seafloor. The true Rx-Tx elevation and relative seawater conductivity stratification was maintained. Each chain sampled 2×10^6 models, with the first half of the chain being thrown away to achieve sampled RMS values between 1.28 and 2.08, with a total of 90 chains converging to acceptable values. The total run time was 192 hours. Uniform prior
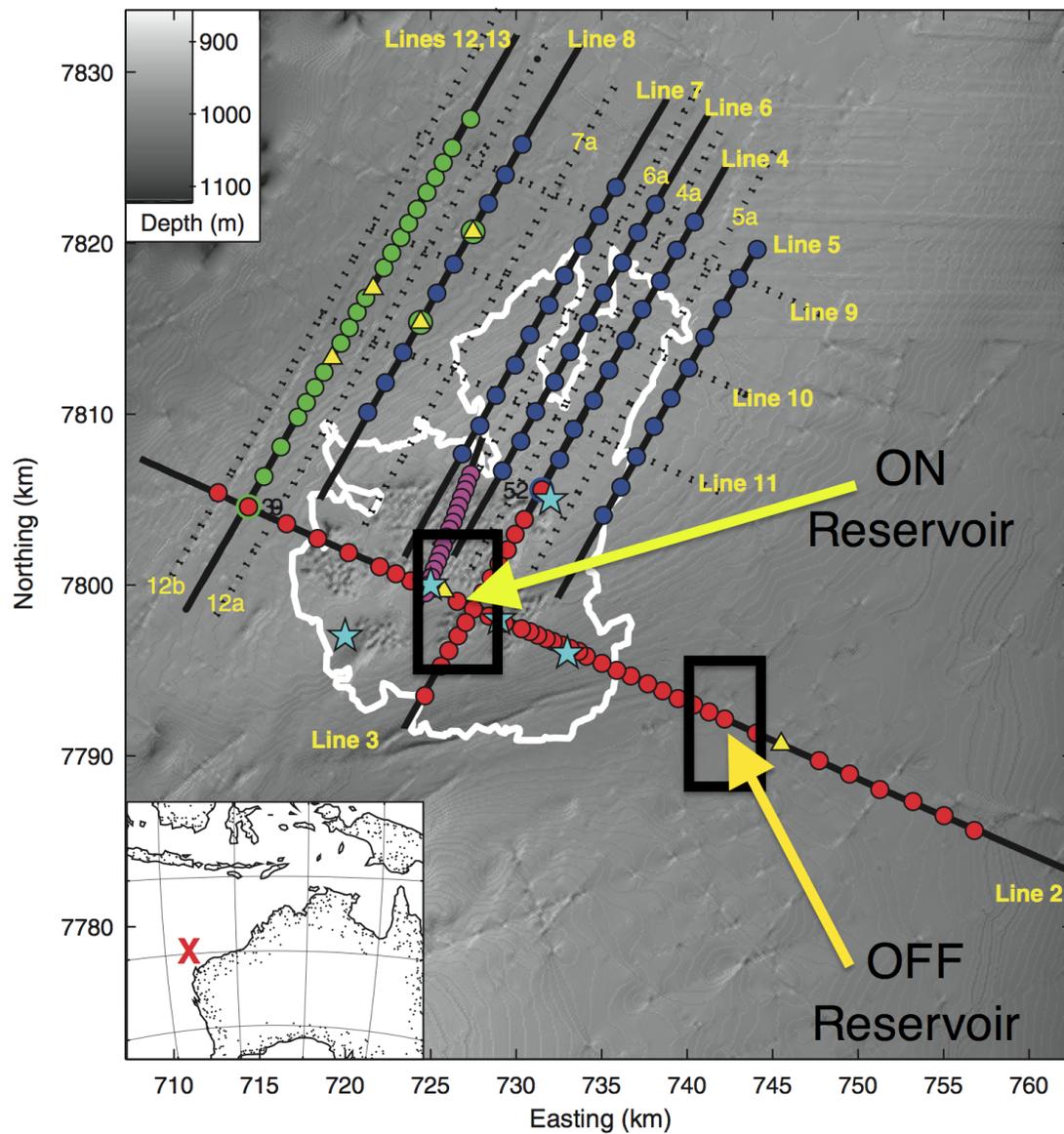
Figure 4.8: A map of the Scarborough reservoir outline (white) overlaid on the bathymetry. Instruments were deployed at points marked by filled circles in 4 phases colored red, blue, green and magenta. In this work, we have focused on the flagship Line 2 towards the south. Modified from Myer et al. (2012).
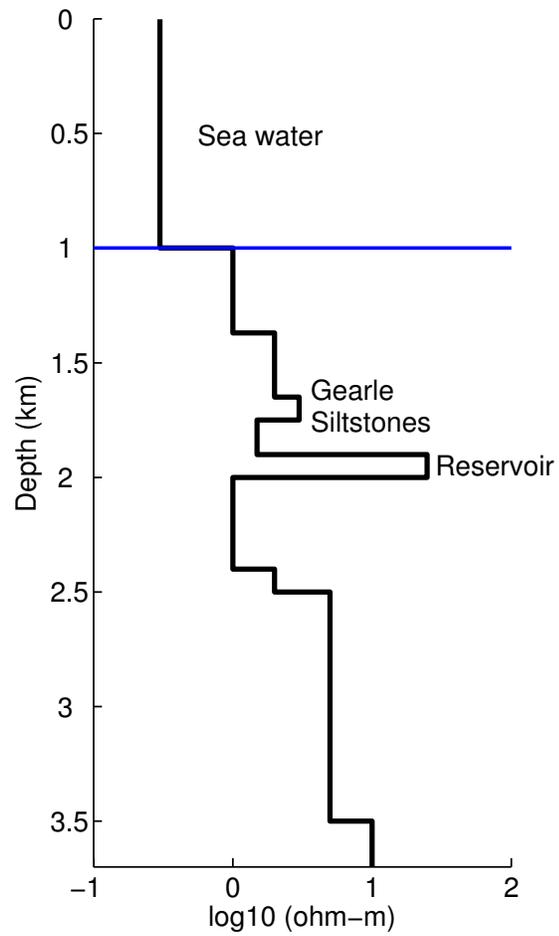
Figure 4.9: Simplified cross section of Scarborough reservoir electrical resistivity structure, reproduced from Myer et al. (2012)
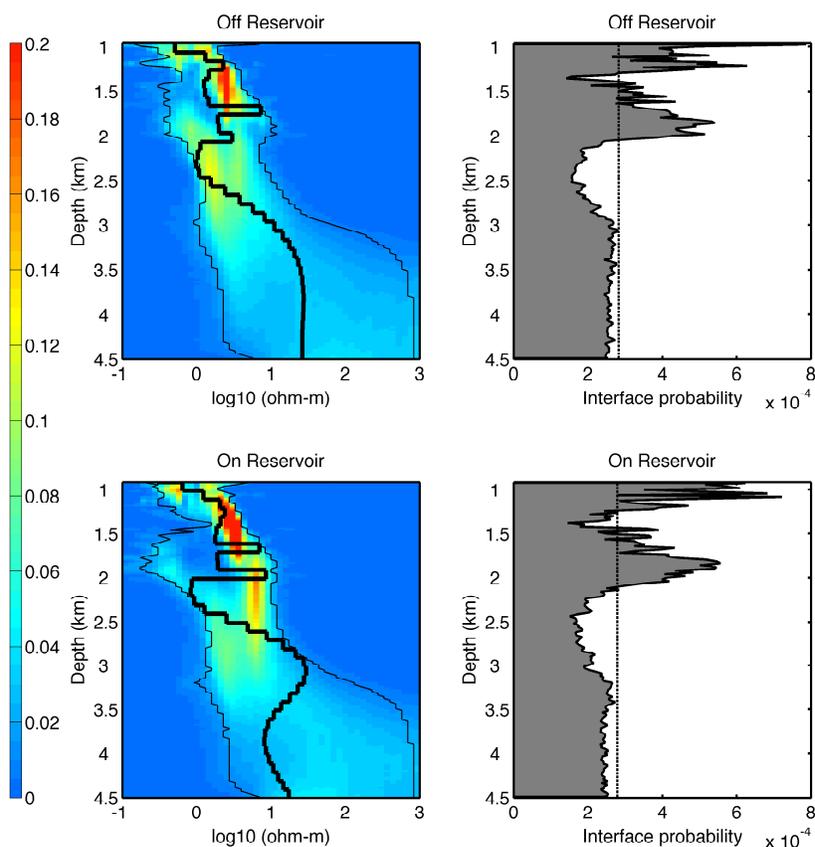
Figure 4.10: 1D Occam's inversion results (thick black line) from Myer et al. (2012) with roughness penalty cuts at the reservoir and Gearle siltstone sections at sites off reservoir (top) and directly on reservoir (bottom) overlain on purely 1D Bayesian inversion probabilities for resistivity with depth using the same data. The 5% and 95% quantiles on resistivity with depth are shown as thin black lines on the left panels. The 1D Bayesian inversions followed the methodology of Ray and Key (2012) and were performed separately.

bounds require that models possess resistivities between 0.2 and 200 ohm-m, and from 20 to 400 nodes placed between 940.6 and 3500 m depth anywhere within the 50 km long 2D line.

A probability cube with axes representing $\log_{10}$ resistivity, Line 2 distance and depth was formed as is shown in Fig. 4.11. The top figure shows a horizontal slice at 1950 m depth (reservoir level). The background resistivity is well sampled in this slice, at around 1 ohm-m. The Gearle is also visible everywhere at 3.2 ohm-m in this horizontal slice (though it should be shallower, between 1650 and 1750 mbsl and not visible in the 1950 m slice). Most interestingly, between 6 and 24 km along Line 2, where the gas saturation was deemed to be greater than 50% from seismic and petrophysical analyses, there is a gradual increase in resistivities such that they lie largely between 6.3 to 15.8 ohm-m. Given that we have not regularized or guided our inversion in any manner besides use very broad uniform parameters to sample within, this is a clear indication of a more resistive body in that section of line at ~1950 m depth. The middle figure shows a vertical section through 11 km position along the line, which lies in the middle of the reservoir outline on Fig. 4.8 (on reservoir). There is a clear increase in resistivity at 2 km depth to values of 10 ohm-m and above, with a return to background values at deeper depths, with suggestions of a very resistive basement at 3 km depth. The bottom figure shows a vertical section through 50 km along Line 2, at the extreme east end of the line, 25 km outside the reservoir outline (off reservoir area). If we follow the distribution of resistivities along the line at 2 km depth from the middle figure to the bottom figure, the bulk of the models in the distributions of resistivity at 2 km depth change from values above 10 ohm-m down to values close to 3.16 ohm-m. Thus, looking at all three parts of Fig. 4.11, we are able to infer that there is indeed a resistor at 1950–2000 m depth, the lateral extents of which are limited roughly between 6 and 24 km position along the line. We emphasize here that looking at the spatial changes in the distributions of resistivity — as made possible by a Bayesian inversion, is a more robust method of interpreting geology, than looking at the changes in one single, inverted model from a regularized inversion method. The PDF on the number of Voronoi nodes required by the trans-dimensional inversion is shown in Fig. 4.12. Again, this PDF does not cluster either around the maximum or minimum possible number of nodes (20 and 400) and

shows that the data require 50–150 nodes to be fit.

The probability of τ between 1500 and 2500 m depth, illustrated in Fig. 4.13 (top), shows three modes corresponding to the background (1000-2000 ohm-m^2), Gearle (2500-3000 ohm-m^2) and the basement ($> 10^4$ ohm-m^2) resistivities. In addition there is a clear departure from the Gearle resistivities in the same lateral section of line as the reservoir is known to be in. This is highlighted by the boxed area of Fig. 4.13 (top) and in the close view given in Fig. 4.13 (bottom). Further, this anomaly is $\sim$900 ohm-m^2 above the background modal value of 3000 ohm-m^2, as expected for the reservoir.

Similar to Myer et al. (2012), we have not been able to separate the Gearle and reservoir levels. However, unlike Myer et al. (2012) the Bayesian posterior distribution *does* concentrate models with resistive anomalies at the reservoir depth. This in itself, we think is a noteworthy aspect of our result.

4.4.4   Comparison with MARE2DEM deterministic 2D inversion

An inversion over Line 2 utilizing the full range of Rx-Tx offsets ($\sim$5.5 km) using the deterministic MARE2DEM code (Key and Ovall, 2011; Key, 2012) is shown in Fig. 4.14. This inversion uses 2D forward model physics and fits the data to RMS 1.0. For comparison with the 2D Bayesian inversion, the MARE2DEM inversion result is shown intersected by two probability cube slices at 11 km and 50 km along Line 2 (Fig. 4.15). The resulting intersections are shown in thick black as profiles along the two slice planes. The most dominant feature from the MARE2DEM inversion at 11 km along the line is high basement resistivity below 2500 m depth. Interestingly, the Bayesian inversion probability slice at 11 km along the line also shows a large change in resistivity distribution near this depth. Both the Bayesian probability distribution and the deterministic inversion show an increase in resistivity at reservoir depth ($\sim$1950 m) in this slice. In the slice at 50 km along Line 2, both the deterministic and Bayesian inversions show a much reduced resistivity at reservoir depth, and both are unable to localize the basement resistivity.

The MARE2DEM inversion shows a single model that is more resistive than the background in the reservoir area. Depending on the model regularization philosophy and choices of regularization parameter used, a different result may have been obtained.
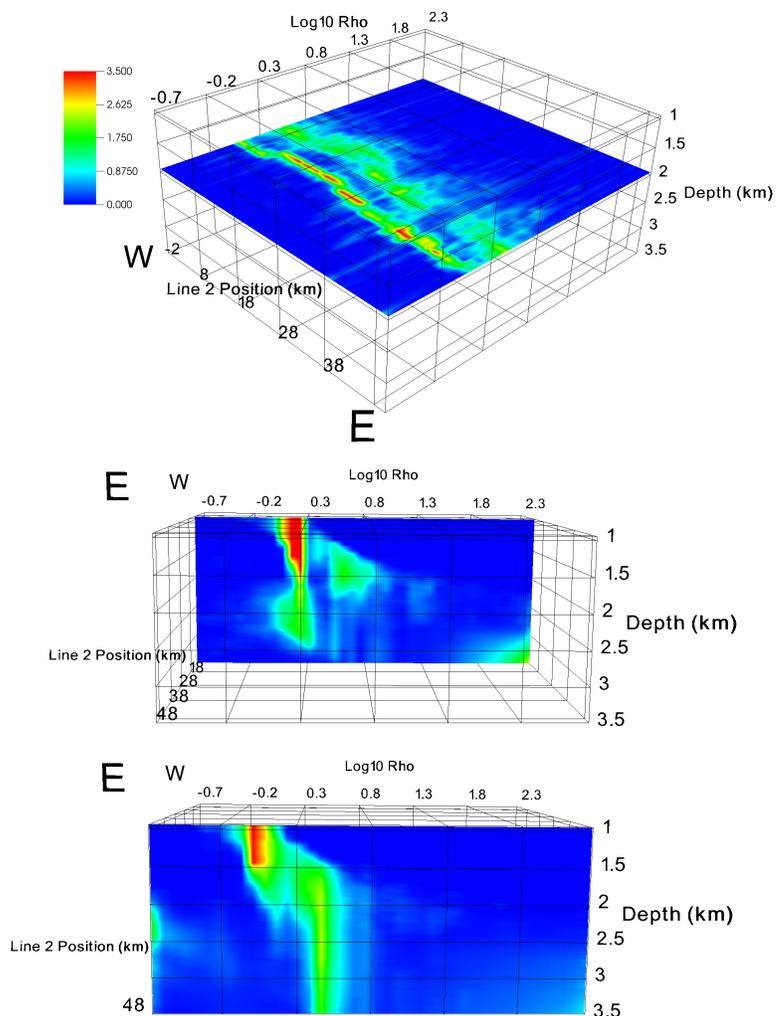
Figure 4.11: Top: A slice through the probability cube for the Scarborough data inversion at 1950 m depth. Hotter colours are more probable. Note the probability of increasing resistivity between 6 and 24 km along Line 2, indicative of reservoir. Middle: A slice through 11 km along the line (on reservoir), within the known reservoir outline shown head on. The probability of a 10 ohm-m resistive anomaly is seen clearly at 2000 m depth. Bottom: A slice through 50 km along the line (off reservoir area), 25 km east and outside of the known reservoir outline. The probability of a resistive anomaly at 2000 m depth has now all but disappeared.
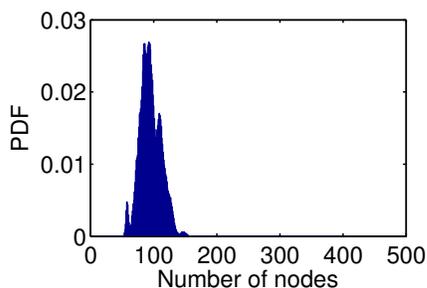
Figure 4.12: PDF on the number of Voronoi nodes required by the trans-dimensional inversion to fit the Scarborough CSEM data.
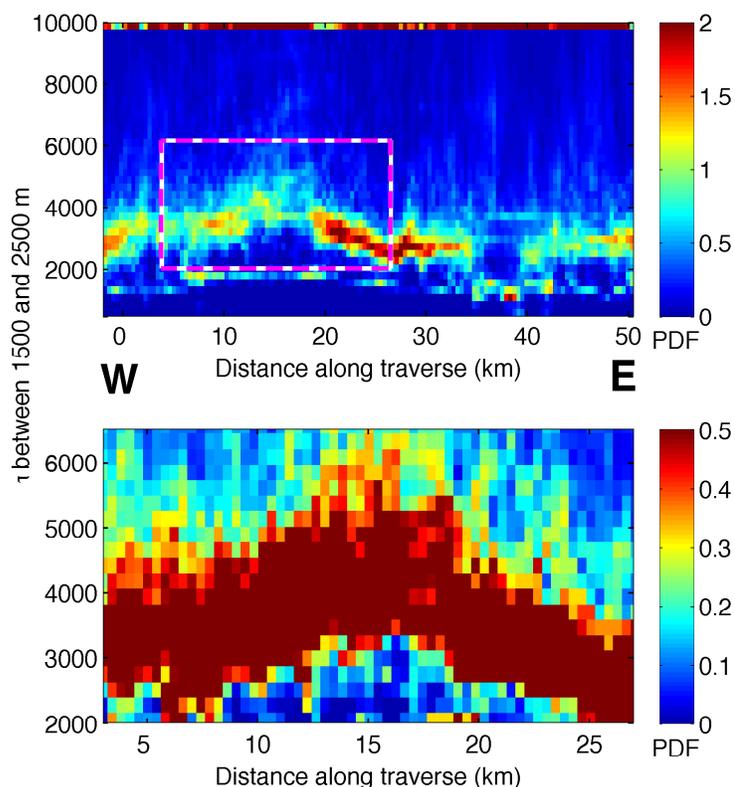
Figure 4.13: Top: Probability of integrated resistivity thickness product τ between 1500 and 2500 m depth for the Scarborough data inversion. Hotter colours are more probable. The reservoir is seen at roughly 900 ohm-m^2 above the background between 6 and 24 km along the line (dashed box). Throughout the line, the basement τ consistently shows up as being the most probable in this depth window. Bottom: Same as top, but zoomed in to the reservoir area.
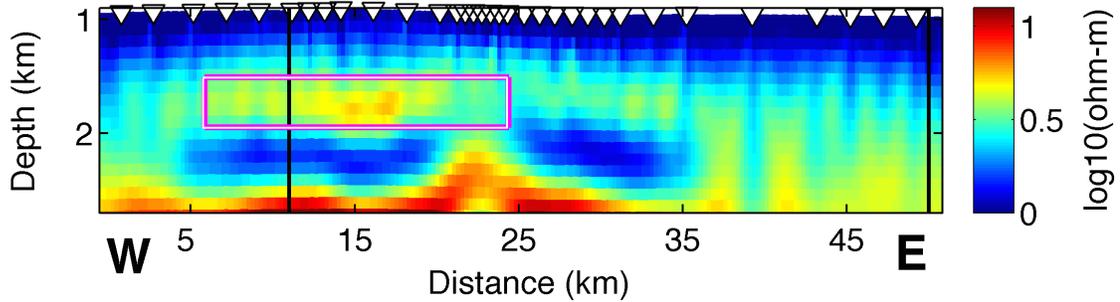
Figure 4.14: The MARE2DEM deterministic inversion result using 2.5D forward modelling for Scarborough CSEM data over Line 2. Seafloor receivers are shown as white inverted triangles. The most prominent feature is high basement resistivity below 2.5 km depth. The reservoir section (boxed) at ∼1.9 km depth is more resistive than its surroundings. Resistivity at 1.9 km depth falls off to lower values towards the east. For purposes of comparison with the probabilistic inversion, two locations for vertical slices have been shown in black at 11 and 50 km. These are the same sections shown in Fig. 4.11 middle and bottom.

Further, a single model does not yield any information on the uncertainty with which the model is associated. The Bayesian ensembles do indeed show how the probability distributions of resistivity vary in the earth. However, the Bayesian inversion, though it is parameterized in 2D, uses a 1D forward modeling engine and thus approximates the true physics. Both of these results taken jointly, reinforce each other to make a stronger argument for the presence of a reservoir, with an associated uncertainty.

4.5 Conclusions

We have successfully implemented a flexibly parameterized inversion scheme using 2D Voronoi cells for CSEM data. In order to accomplish this, we have used a 'birth/death' RJ-MCMC algorithm (Green, 1995; Bodin and Sambridge, 2009; Ray and Key, 2012). Synthetic studies were carried out to understand the model space associated with thin, mildly resistive reservoirs. Our results show that the true values of resistivity for these reservoirs are not the most probable, *a posteriori*, and that the integrated resistivity thickness product τ is a more robust quantity to examine in these cases. Following the spatial changes in posterior distributions on both resistivity as well as τ yields valuable information on the presence and possible geometries of resistive
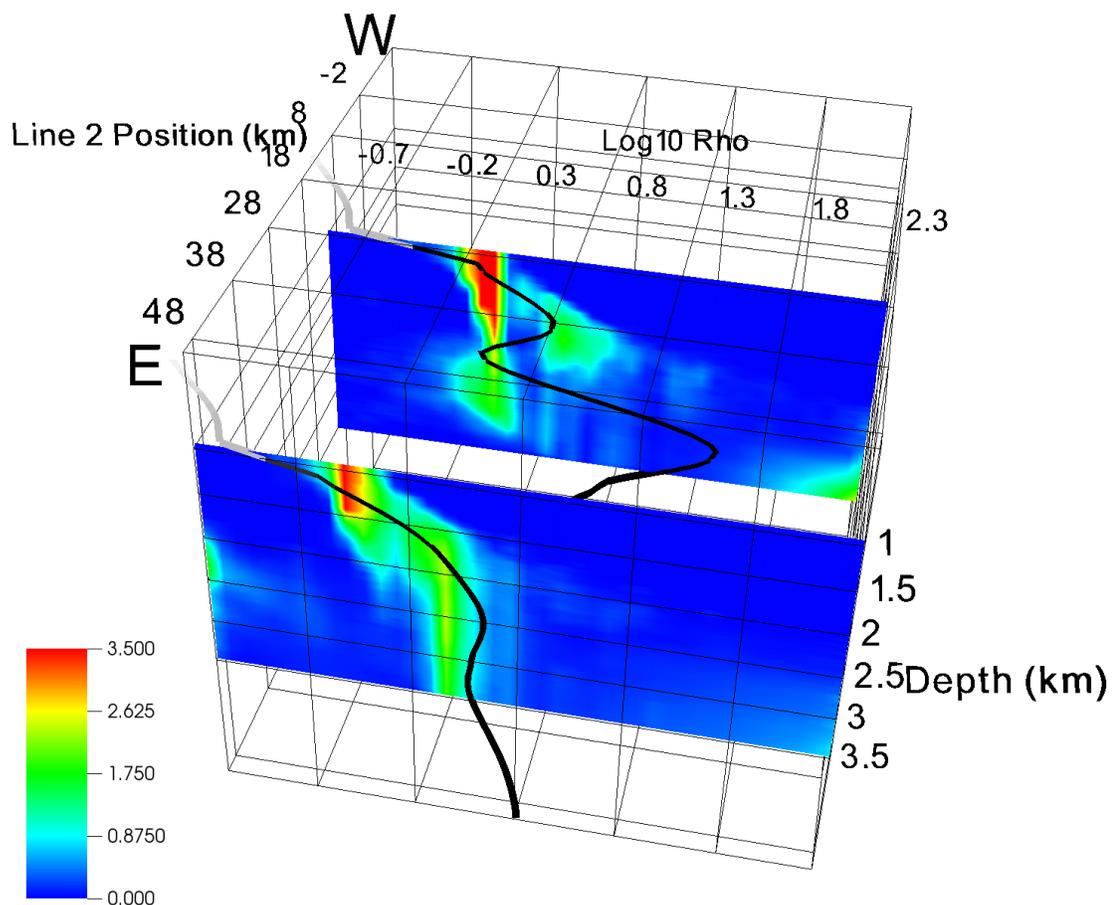
Figure 4.15: Two slices through the Scarborough inversion probability cube shown at 11 km (on reservoir) and 50 km (off reservoir area) along the line, comparing the Bayesian inversion distributions (PDF proportional to colour hotness) and the MARE2DEM inversion result (thick black line intersecting the 2 slice planes). The 11 km slice runs through the middle of the reservoir whereas the 50 km slice is 25 km to the east of the 50% gas saturation line. These are at the same locations marked in Fig. 4.14 and show the same slices as Fig. 4.11 middle and bottom.

bodies. Our Bayesian inversion of field data from the Scarborough CSEM survey was successfully able to delineate the reservoir laterally, with a good indication of its depth but not thickness. Comparison with a MARE2DEM deterministic inversion showed good agreement between the Bayesian and deterministic results with the added benefit of uncertainty about the deterministic result provided by the Bayesian posterior model ensemble. There is a degree of re-assurance in the presence of a reservoir provided by the Bayesian posterior model distribution, when *en masse*, un-regularized inverted models at certain locations within the earth tend to be more resistive.

The obvious next step is to move to full 2D forward modelling of the Voronoi cell models used by our trans-dimensional algorithm. While 2D model responses are computationally more expensive than their 1D counterparts, with the use of accelerated sampling methods such as parallel tempering (Dettmer et al., 2011; Dosso et al., 2012; Sambridge, 2013; Ray et al., 2013) and the advent of highly parallel, cluster-computing oriented 2D forward codes (Key and Ovall, 2011), this step is not far away. Lastly, we would like to point out that our algorithm is quite flexible and can be applied to various classes of geophysical problems such as surface wave dispersion, seismic receiver functions, etc.

4.A Appendix: Mathematical details for the trans-dimensional algorithm

4.A.1 The prior probability

The prior PDF contains information on our knowledge about the subsurface independent of the survey data. This can be based on data from well logs, seismic surveys, etc. In the trans-dimensional formulation, we split the prior into 2 parts. One part contains information about the number of cells k in the model, $p(k)$. The other part $p(\mathbf{m_k}|k)$ in our particular case, contains information about the physical parameters associated with a model $\mathbf{m_k}$ of dimension k, such as where these cells are, and what the resistivities of these cells are. Using the chain rule, we can thus write

$$p(\mathbf{m}) = p(\mathbf{m_k}, k) = p(\mathbf{m_k}|k) \times p(k). \tag{4.6}$$

where

$$\mathbf{m_k} = [\mathbf{z}, \mathbf{x}, \boldsymbol{\rho}], \tag{4.7}$$

$$\mathbf{z} = [z_1, z_2, ..., z_k], \tag{4.8}$$

$$\mathbf{x} = [x_1, x_2, ..., x_k], \tag{4.9}$$

$$\boldsymbol{\rho} = [\rho_1, \rho_2, ..., \rho_k], \tag{4.10}$$

and

$$z \in \mathbf{z}, x \in \mathbf{x}, \rho \in \boldsymbol{\rho}. \tag{4.11}$$

We use a uniform prior on k, given by

$$p(k) = \begin{cases} \frac{1}{k_{max} - k_{min} + 1} & \text{if } k_{min} \leq k \leq k_{max} \\ 0 & \text{else} \end{cases}. \tag{4.12}$$

If we restrict our region of interest to be a rectangle within the vertical plane running through the CSEM survey line, we can define a uniform prior on the position of Voronoi nodes within this rectangle $p(\mathbf{z}, \mathbf{x})$. We assume no *a priori* knowledge between the locations of nodes in the model and the Voronoi cell resistivities $\boldsymbol{\rho}$. These random variables are independent and therefore their PDFs can be separated in the following product form,

$$p(\mathbf{m_k}|k) = p(\mathbf{z}, \mathbf{x}|k)p(\boldsymbol{\rho}|k). \tag{4.13}$$

Voronoi nodes can be located anywhere in the rectangular subsurface area defined by $[z_{min}, z_{max}]$ and $[x_{min}, x_{max}]$. A given node can be at any of K (temporarily discrete) points within this rectangle. For k nodes, we can arrange them without paying attention to their ordering in $\frac{K!}{k!(K-k)!}$ ways. Note that this unspecified variable K will cancel out of the expressions we need to use in the algorithm and is only introduced for ease of mathematical derivation. Thus,

$$p(\mathbf{z}, \mathbf{x}|k) = \begin{cases} \left[\frac{K!}{k!(K-k)!}\right]^{-1} & \text{if } z_{min} \leq z \leq z_{max} \text{ and } x_{min} \leq x \leq x_{max} \\ 0 & \text{else} \end{cases}. \tag{4.14}$$

Assuming that all k Voronoi cell resistivities within a given model lie uniformly between ρ_{min} and ρ_{max}, independent of each other, we write

$$p(\boldsymbol{\rho}|k) = \begin{cases} \left[\frac{1}{\rho_{max}-\rho_{min}} \right]^k & \text{if } \rho_{min} \le \rho \le \rho_{max} \\ 0 & \text{else} \end{cases} . \qquad (4.15)$$

To obtain the explicit expression for the prior model probability, we write $\Delta\rho = \rho_{max} - \rho_{min}$ and $\Delta k = k_{max} - k_{min} + 1$ and substitute equations $(4.12) - (4.15)$ into (4.6) to get

$$p(\mathbf{m}) = \begin{cases} \frac{k!(K-k)!}{K!\Delta k(\Delta\rho)^k} & \text{if } z \in [z_{min}, z_{max}], \ x \in [x_{min}, x_{max}], \\ & \rho \in [\rho_{min}, \rho_{max}], \ \forall k \in [k_{min}, k_{max}] \\ 0 & \text{else} \end{cases} . \qquad (4.16)$$

We must mention here that it is natural in geophysical EM to parameterize models using $\log_{10}(\text{resistivity})$ instead of linear resistivity (e.g., Ray and Key, 2012), a practice we have followed in our implementation.

4.A.2 MH algorithms and the acceptance probability

What guides an MCMC sampler like the MH algorithm to convergence upon the posterior distribution is the acceptance probability α (e.g., Liang et al., 2011). At every step of the Markov Chain, a candidate model is sampled by perturbing the current model from a known distribution (the proposal distribution q) and the acceptance α is calculated. A random number r is then sampled uniformly from the interval $[0,1]$. If $r < \alpha$ the proposed perturbation is accepted, else the old model is retained. The rationale behind this algorithm can be explained by examining in more detail the expression for α (Bodin and Sambridge, 2009), where

$$\alpha(\mathbf{m}'|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{m}')}{p(\mathbf{m})} \times \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})} \times |\mathbf{J}| \right]. \qquad (4.17)$$

Here $\mathbf{m}'$ is the new proposed model and $\mathbf{m}$ is the old model (throughout this paper, primes will denote new model values). Specifically, $\frac{p(\mathbf{m}')}{p(\mathbf{m})}$ is the prior ratio, $\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}$ is the likelihood ratio and $\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}$ is the proposal ratio. The Jacobian term $|\mathbf{J}|$ is not to be

confused with the model Jacobian needed for gradient based inversions (e.g., Constable et al., 1987), but is a matrix that incorporates changes in model dimension when moving from $\mathbf{m}$ to $\mathbf{m}'$. In a classic MH algorithm with a fixed number of dimensions, the prior ratio (for uniform priors), proposal ratio (for symmetric proposals), and Jacobian term are all 1 (Dettmer et al., 2010). Hence the algorithm always moves towards areas of higher posterior probability if the data misfit improves (likelihood ratio > 1). However, it can also move to areas of lower posterior probability with a probability α if the misfit does not improve (likelihood ratio < 1).

To be able to compare likelihoods between models with different numbers of parameters (i.e., with different dimensions), the Jacobian in the acceptance term in equation (4.17) needs to be evaluated. There are various implementations of RJ-MCMC, and in all the examples cited so far, a 'birth-death' scheme has been used. As shown in Bodin and Sambridge (2009) and Dettmer et al. (2010) for the 'birth-death' RJ-MCMC scheme, this Jacobian term is unity. We have adopted the 'birth-death' algorithm in this paper and shall not concern ourselves with this Jacobian term any further.

As to why the algorithm should not always look to improve the data fit by simply increasing the number of parameters (Voronoi cells in the seabed), if we examine equation (4.17) we find that even if the likelihood ratio times the proposal ratio is greater than one for a proposed move that inserts a new cell into the model, the prior ratio will be less than one owing to the fact that the new prior PDF $p(\mathbf{m}')$ needs to integrate over a larger number of parameters to equal 1. Hence, there is an opposition to the 'birth' of a new layer (which may lead to improvement of data fit) by the prior ratio.

4.A.3   Outline of our algorithm

We start the algorithm with a very simple model, with $k = k_{min}$. We then allow the algorithm to iteratively add Voronoi nodes ('birth') or remove them ('death'), perturbing the Voronoi cell resistivities, as the data may demand via the acceptance probability α in (4.17). In brief, this is how we proceed:

Initialization

Start the algorithm with $k = k_{min}$ and all resistivities set to that of a uniform 1 ohm-m half space. The Voronoi nodes are randomly distributed within the prior bounds rectangle.

Choose one of 4 moves

1) Update: Perturb a randomly chosen Voronoi cell resistivity about its current value using a Gaussian proposal $q(\mathbf{m}'|\mathbf{m})$ with a standard deviation Σ_ρ, where

$$q(\mathbf{m}'|\mathbf{m}) = \frac{1}{\sqrt{2\pi}\Sigma_\rho}\exp\left[-\frac{1}{2\Sigma_\rho^2}(\rho' - \rho)^2\right]. \tag{4.18}$$

Note that this update move does not involve a change in the number of cells.

2) Birth of a new node: $k' = k + 1$. In the rectangular area defined by $[z_{min}, z_{max}]$ and $[x_{min}, x_{max}]$, randomly and with uniform probability we select an unoccupied point and insert a node. This node forms the nucleus for a new Voronoi cell, the resistivity of which is assigned by perturbing the old resistivity value at that location according to a Gaussian proposal with standard deviation Σ_{bd}.

3) Death of an interface: $k' = k - 1$. An existing node is selected at random and deleted.

4) Move a node location: An existing node is selected at random and its position is perturbed by 2 independent 1D Gaussian proposals with standard deviation Σ_{mz} and Σ_{mx} for the z and x directions, respectively. Note that this step does not involve a change in the number of cells either.

At each step, one of these moves is chosen with a certain probability, such that the move probabilities sum to unity. In addition, the birth and death probabilities must be set equal. We set the probabilities as follows:

$$[\text{update}, \text{birth}, \text{death}, \text{move}] \equiv \left[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right].$$

At each step of the Markov chain, the proposed model is evaluated for acceptance. If it is accepted, it becomes the current model. If it is rejected, the current model is preserved and the algorithm moves on to the next step. In order to compute the

acceptance, one needs to evaluate equation (4.17), for which we explicitly describe the proposal distributions and their ratios in the next section.

4.A.4 Proposal distributions and acceptance probabilities

Fixed dimension moves

For all moves that are neither birth nor death, the number of cells remain fixed. In these moves, we have elected to use Gaussian proposals to suggest the new model parameters by centering the proposals on the old parameters and drawing a random number from a normal distribution with a given standard deviation (step size). We can see from equation (4.18) these kinds of moves are symmetric, implying that the probability to go from the old state to the new state is the same as it would be in going from the new state to the old state:

$$\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{fixed}} = 1. \tag{4.19}$$

Since the number of dimensions remains constant, the prior ratio in equation (4.17) is 1. Hence for fixed dimension moves, we find that the acceptance probability is simply the ratio of the likelihoods:

$$\alpha_{\text{f}} = \begin{cases} \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \text{if } z \in [z_{min}, z_{max}],\, x \in [x_{min}, x_{max}], \\ & \\ & \rho \in [\rho_{min}, \rho_{max}], \\ 0 & \text{else} \end{cases} . \tag{4.20}$$

Birth move

For a birth move, one can select from out of $K - k$ unoccupied spaces. The perturbation for the birthed cell's resistivity is drawn from a Gaussian with standard deviation Σ_{bd}, centered about the old value in the cell. Since the selection of a position

and the perturbations are independent, we can write

$$
\begin{aligned}
q(\mathbf{m}'|\mathbf{m}) &= q(\mathbf{z}', \mathbf{x}'|\mathbf{m})q(\boldsymbol{\rho}'|\mathbf{m}) && (4.21)\\
&= \frac{1}{(K-k)}\frac{1}{\sqrt{2\pi}\Sigma_{bd}}\exp\left[-\frac{(\rho'-\rho)^2}{2\Sigma_{bd}^2}\right]. && (4.22)
\end{aligned}
$$

For the reverse move in a birth, keeping in mind that the current state has k cells, there were $k+1$ cells to delete from, and the probability of removing resistivities in a cell in the reverse move is 1. Thus we have

$$
\begin{aligned}
q(\mathbf{m}|\mathbf{m}') &= q(\mathbf{z}, \mathbf{x}|\mathbf{m}')q(\boldsymbol{\rho}|\mathbf{m}') && (4.23)\\
&= \frac{1}{(k+1)}\times 1. && (4.24)
\end{aligned}
$$

It follows in a birth move, from equations (4.22) and (4.24), that the proposal ratio can be written as

$$
\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{birth}} = \frac{(K-k)\sqrt{2\pi}\Sigma_{bd}}{k+1}\exp\left[\frac{(\rho'-\rho)^2}{2\Sigma_{bd}^2}\right]. \qquad (4.25)
$$

Finally from equations (4.16), (4.17) and (4.25) we get for the birth moves, the following acceptance probability

$$
\alpha_{\mathrm{b}} = \begin{cases} \min\left[1, \frac{\sqrt{2\pi}\Sigma_{bd}}{\Delta\rho}\exp\left[\frac{(\rho'-\rho)^2}{2\Sigma_{bd}^2}\right]\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \rho \in [\rho_{min}, \rho_{max}], \ \forall k \in [k_{min}, k_{max}] \\ 0 & \text{else} \end{cases}.
$$

$$(4.26)$$

Death move

In a death move, one can select one of k places for deletion. Further, the probability of removing resistivities in a cell is certain. Thus,

$$
\begin{aligned}
q(\mathbf{m}'|\mathbf{m}) &= q(\mathbf{z}', \mathbf{x}'|\mathbf{m})q(\boldsymbol{\rho}'|\mathbf{m}) && (4.27)\\
&= \frac{1}{k}\times 1. && (4.28)
\end{aligned}
$$

In the reverse move for death, since the reference state has k interfaces, there are $K - (k-1)$ sites at which to add an interface. Further, the resistivity perturbations are

proposed using a Gaussian centered around the current value. Hence

$$q(\mathbf{m}|\mathbf{m}') \quad = \quad q(\mathbf{z}, \mathbf{x}|\mathbf{m}')q(\boldsymbol{\rho}|\mathbf{m}') \tag{4.29}$$

$$= \quad \frac{1}{K-k+1} \times \frac{1}{\sqrt{2\pi}\Sigma_{bd}}\exp\left[-\frac{(\rho-\rho')^2}{2\Sigma_{bd}^2}\right]. \tag{4.30}$$

Thus we can see from equations (4.28) and (4.30) that the proposal ratio for death can be written as:

$$\left[\frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right]_{\text{death}} = \frac{k}{(K-k+1)\sqrt{2\pi}\Sigma_{bd}}\exp\left[-\frac{(\rho-\rho')^2}{2\Sigma_{bd}^2}\right]. \tag{4.31}$$

Again from equations (4.16), (4.17) and (4.31) we get for the death moves, the following acceptance probability

$$\alpha_{\mathrm{d}} = \begin{cases} \min\left[1, \frac{\Delta\rho}{\sqrt{2\pi}\Sigma_{bd}}\exp\left[-\frac{(\rho'-\rho)^2}{2\Sigma_{bd}^2}\right]\frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})}\right] & \forall k \in [k_{min}, k_{max}] \\ 0 & \text{else} \end{cases} \tag{4.32}$$

It should be noted that the derived expressions for α in equations (4.20), (4.26) and (4.32) do not involve the variable K (as promised) and are identical in form to the expressions derived in Bodin and Sambridge (2009). This demonstrates how flexible the algorithm is when solving completely different kinds of geophysical problems.

Further, we mention here that in practice, the log of the acceptance probabilities are calculated and compared against the log of a uniform random number between 0+ and 1. Taking log avoids many problems of numerical stability in the evaluation of equation 4.17.

4.A.5 Synergy between birth and death

At this juncture, we point out that the birth and death moves operate 'in concert'. Examining equations 4.26 and 4.32, we see that for the same change in the likelihood ratio, the birth and death acceptance probabilities are inverses of each other. The birth move encourages large steps to be taken in the model space, while the death move encourages very small steps in the model space. Thus birth ensures that there should

be an increase in dimension, only when the current model is quite different from the proposed model (within the prior bounds, of course). On the other hand, the death move ensures that when the proposed model is very similar to the current one, there should be a decrease in dimensions and an unnecessary cell is then removed.

This synergy between birth and death also explains why the RJ-MCMC algorithm does not keep adding or deleting cells in order to explain the observed data. This is another aspect of Bayesian natural parsimony.

4.A.6   Convergence to the posterior distribution

The algorithm is run for a given number of steps until it is deemed to have collected enough samples to provide a reasonable estimate of the posterior model PDF. There are a couple of caveats in this regard, as there are with any MCMC sampler (Liang et al., 2011). If the algorithm is seeded with an initial model that is in a low posterior probability region, it may take quite a few steps till it reaches a region of high posterior probability, such that it begins to sample models, most of which fit the data within the given data error. The number of such required steps (which are subsequently discarded in the final chain) is known in MCMC parlance as the 'burn-in' period, which depends on how well the proposal distributions have been scaled (Chib and Greenberg, 1995). This brings us to the step sizes (scaling) in the proposal distributions in the form of the standard deviations $\Sigma_\rho, \Sigma_{mz}, \Sigma_{mx}$ and Σ_{bd} required in the various proposals to generate a new candidate model. The form of the proposal distributions should 'emulate' the posterior for efficient sampling, but since the posterior distribution may be complicated (and unknown *a priori*), any kind of simple distribution, symmetric where possible, can be used. The exact form of the proposal does not affect the final solution, at least in theory. For the Scarborough problem, we used the following step sizes with Gaussian proposals:

$$[\Sigma_\rho, \Sigma_{mz}, \Sigma_{mx}, \Sigma_{bd}] \equiv \left[0.1, 10, 150, 0.6\right],$$

with the resistivity step sizes Σ_ρ and Σ_{bd} specified in $\log_{10}$ units and the move node step sizes Σ_{mz} and Σ_{mx} in metres.

The suitability of the step size for the problem at hand can be examined by

looking at the number of samples accepted in a large interval of steps, referred to as the acceptance rate. If the acceptance rate is too low, it means that the step sizes are too large as lots of steps are falling outside the prior bounds or are being rejected as they land in low probability (high misfit) areas. If the acceptance rate is too high, then it implies that the algorithm isn't exploring the model space enough and will again be slow to converge upon the posterior distribution.

While the sampled posterior should not depend on the size of the steps taken, one has to factor in the optimality of the step size as otherwise convergence will be very slow (Bodin and Sambridge, 2009). For an illuminating discussion on this matter, one can refer to Chib and Greenberg (1995) or Trainor-Guitton and Hoversten (2011) for a more recent discussion relevant to marine CSEM. For further discussions on convergence diagnostics and the practicality of their application, one can refer to Liang et al. (2011). A heuristic method to examine speed of convergence is to monitor the 'distance' travelled from one sampling step to the next over a window of a certain number of steps. We can project model resistivity values on to an underlying grid, and then look at the norm of the difference between grid values for two successive models. We can then find the average distance travelled over a given step size window.

The algorithm should be run long enough at the lowest acceptable range of RMS values (achieved after the burn-in period) such that there is at least stationarity achieved in the square misfit with iteration number. Further, to ensure that the inferred posterior is not biased due to being trapped in local maxima (of the posterior probability), we recommend that the algorithm be run from many different starting points, ideally in parallel for computational efficiency. The final ensemble for posterior inference can be constructed by concatenating the various parallel chains (e.g., Dettmer et al., 2010; Bodin et al., 2012; Ray and Key, 2012). Another means of efficient convergence to the posterior distribution can be to use parallel interacting Markov chains as described in Dettmer and Dosso (2012) and Ray et al. (2013).

Acknowledgments

References

Abubakar, A., Habashy, T. M., Druskin, V. L., Knizhnerman, L., and Alumbaugh, D., 2008: 2.5D forward and inverse modeling for interpreting low-frequency electromagnetic measurements. *Geophysics*, **73**(4), F165–F177.

Agostinetti, N. P., and Malinverno, A., 2010: Receiver function inversion by transdimensional Monte Carlo sampling. *Geophysical Journal International*, **181**, 858–872.

Backus, G. E., 1988: Bayesian inference in geomagnetism. *Geophysical Journal International*, **92**(1), 125–142.

Bayes, T., and Price, R., 1763: An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S. *Philosophical Transactions*, **53**, 370–418. doi:10.1098/rstl.1763.0053.

Bodin, T., and Sambridge, M., 2009: Seismic tomography with the reversible jump algorithm. *Geophysical Journal International*, **178**(3), 1411–1436.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N., 2012: Transdimensional inversion of receiver functions and surface wave dispersion. *Journal of Geophysical Research*, **117**(B2).

Brodie, R. C., and Sambridge, M., 2012: Transdimensional Monte Carlo Inversion of AEM Data. In *22nd International Geophysical Conference and Exhibition, Brisbane, Australia*, 1.

Buland, A., and Kolbjornsen, O., 2012: Bayesian inversion of CSEM and magnetotelluric data. *Geophysics*, **77**(1), E33–E42.

Chave, A. D., and Cox, C. S., 1982: Controlled electromagnetic sources for measuring electrical-conductivity beneath the oceans.~1.~Forward problem and model study. *Journal Of Geophysical Research-Solid Earth*, **87**, 5327–5338.

Chen, J., Hoversten, G. M., Key, K., Nordquist, G., and Cumming, W., 2012: Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site. *Geophysics*, **77**(4), E265.

Chen, J., Hoversten, G. M., Vasco, D., Rubin, Y., and Hou, Z., 2007: A Bayesian model for gas saturation estimation using marine seismic AVA and CSEM data. *Geophysics*, **72**(2), WA85.

Chib, S., and Greenberg, E., 1995: Understanding the metropolis-hastings algorithm. *American Statistician*, 327–335.

Connell, D., and Key, K., 2013: A numerical comparison of time and frequency-domain marine electromagnetic methods for hydrocarbon exploration in shallow water. *Geophysical Prospecting*, **61**, 187–199.

Constable, S., 2010: Ten years of marine CSEM for hydrocarbon exploration. *Geophysics*, **75**(5), 75A67–75A81.

Constable, S., and Weiss, C. J., 2006: Mapping thin resistors and hydrocarbons with marine EM methods: Insights from 1D modeling. *Geophysics*, **71**(2), G43.

Constable, S. C., 2006: Marine electromagnetic methods A new tool for offshore exploration. *The Leading Edge*, **25**, 438–444.

Constable, S. C., Parker, R. L., and Constable, C. G., 1987: Occam's inversion - A practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics*, **52**(03), 289–300.

Dettmer, J., and Dosso, S. E., 2012: Trans-dimensional matched-field geoacoustic inversion with hierarchical error models and interacting Markov chains. *The Journal of the Acoustical Society of America*, **132**(4), 2239–2250. ISSN 1520-8524. doi: 10.1121/1.4746016.

Dettmer, J., and Dosso, S. E., 2013: Probabilistic two-dimensional water-column and seabed inversion with self-adapting parameterizations. *The Journal of the Acoustical Society of America*, **133**(5), 2612–23. ISSN 1520-8524. doi:10.1121/1.4795804.

Dettmer, J., Dosso, S. E., and Holland, C. W., 2010: Trans-dimensional geoacoustic inversion. *The Journal of the Acoustical Society of America*, **128**(6), 3393–3405.

Dettmer, J., Dosso, S. E., and Holland, C. W., 2011: Sequential trans-dimensional Monte Carlo for range-dependent geoacoustic inversion. *The Journal of the Acoustical Society of America*, **129**(4), 1794–806. ISSN 1520-8524. doi:10.1121/1.3557052.

Dosso, S. E., Holland, C. W., and Sambridge, M., 2012: Parallel tempering for strongly nonlinear geoacoustic inversion. *The Journal of the Acoustical Society of America*, **132**(5), 3030–3040. ISSN 1520-8524. doi:10.1121/1.4757639.

Driscoll, N. W., and Karner, G. D., 1998: Lower crustal extension across the Northern Carnarvon basin, Australia: Evidence for an eastward dipping detachment. *Journal Of Geophysical Research-Solid Earth*, **103**(B3), 4975–4991.

Ellingsrud, S., Eidesmo, T., Johansen, S., Sinha, M. C., MacGregor, L. M., and Constable, S., 2002: Remote sensing of hydrocarbon layers by seabed logging (SBL): Results from a cruise offshore Angola. *The Leading Edge*, **21**, 972–982.

Green, P. J., 1995: Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**(4), 711–732.

Gunning, J., Glinsky, M. E., and Hedditch, J., 2010: Resolution and uncertainty in 1D CSEM inversion: A Bayesian approach and open-source implementation. *Geophysics*, **75**(6), F151–F171.

Hastings, W. K., 1970: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.

Hou, Z., Rubin, Y., Hoversten, G. M., Vasco, D., and Chen, J., 2006: Reservoir-parameter identification using minimum relative entropy-based Bayesian inversion of seismic AVA and marine CSEM data. *Geophysics*, **71**(6), O77—-O88.

JafarGandomi, A., and Binley, A., 2013: A Bayesian trans-dimensional approach for the fusion of multiple geophysical datasets. *Journal of Applied Geophysics*, **96**, 38–54. ISSN 09269851. doi:10.1016/j.jappgeo.2013.06.004.

Key, K., 2009: 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, **74**(2), F9.

Key, K., 2012: Marine EM inversion using unstructured grids: a 2D parallel adaptive finite element algorithm. In *SEG Technical Program Expanded Abstracts*, 9, 1–5. Las Vegas.

Key, K., and Ovall, J., 2011: A parallel goal-oriented adaptive finite element method for 2.5-D electromagnetic modelling. *Geophysical Journal International*, **186**(1), 137–154. ISSN 0956540X. doi:10.1111/j.1365-246X.2011.05025.x.

Liang, F., Liu, C., and Carroll, R., 2011: *Advanced Markov chain Monte Carlo methods: learning from past samples*, volume 714. Wiley. com.

Loseth, L. O., Pedersen, H. M., Ursin, B., Amundsen, L., and Ellingsrud, S., 2006: Low-frequency electromagnetic fields in applied geophysics: Waves or diffusion? *Geophysics*, **71**(4), W29–W40.

Luo, X., 2010: Constraining the shape of a gravity anomalous body using reversible jump Markov chain Monte Carlo. *Geophysical Journal International*, **180**(3), 1067–1079. ISSN 0956540X. doi:10.1111/j.1365-246X.2009.04479.x.

MacGregor, L., and Sinha, M., 2000: Use of marine controlled-source electromagnetic sounding for sub-basalt exploration. *Geophysical Prospecting*, **48**(6), 1091–1106.

MacKay, D. J. C., 2003: *Information theory, inference and learning algorithms.* Cambridge university press.

Malinverno, A., 2002: Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem. *Geophysical Journal International*, **151**(3), 675–688.

Minsley, B. J., 2011: A trans-dimensional Bayesian Markov chain Monte Carlo algorithm for model assessment using frequency-domain electromagnetic data. *Geophysical Journal International*, **187**(1), 252–272. ISSN 0956540X. doi:10.1111/j.1365-246X.2011.05165.x.

Mittet, R., Brauti, K., Maulana, H., and Wicklund, T. A., 2008: CMP inversion and post-inversion modelling for marine CSEM data. *First Break*, **26**(August), 59–67.

Mittet, R., and Gabrielsen, P. l. T., 2013: Decomposition in upgoing and downgoing fields and inversion of marine CSEM data. *Geophysics*, **78**(1), E1–E17. ISSN 0016-8033. doi:10.1190/geo2011-0520.1.

Myer, D., Constable, S., and Key, K., 2010: A marine EM survey of the Scarborough gas field, Northwest Shelf of Australia. *First Break*, **28**, 77–82.

Myer, D., Constable, S., and Key, K., 2011: Broad-band waveforms and robust processing for marine CSEM surveys. *Geophysical Journal International*, **184**, 689–698.

Myer, D., Constable, S., and Key, K., 2013: Magnetotelluric evidence for layered mafic intrusions beneath the Vøring and Exmouth rifted margins. *Physics of the Earth and Planetary Interiors*, **220**, 1–10. ISSN 00319201. doi:10.1016/j.pepi.2013.04.007.

Myer, D., Constable, S., Key, K., Glinsky, M. E., and Liu, G., 2012: Marine CSEM of the Scarborough gas field, Part 1: Experimental design and data uncertainty. *Geophysics*, **77**(4), E281–E299.

Neal, R. M., 2003: Slice sampling. *The Annals of Statistics*, **31**(3), 705–767.

Newman, G. A., and Alumbaugh, D. L., 2000: Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophysical Journal International*, **140**(2), 410–424.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S. N., 2009: *Spatial tessellations: concepts and applications of Voronoi diagrams*, volume 501. John Wiley & Sons.

Ray, A., Alumbaugh, D. L., Hoversten, G. M., and Key, K., 2013: Robust and accelerated Bayesian inversion of marine controlled-source electromagnetic data using parallel tempering. *Geophysics*, **78**(6), E271–E280.

Ray, A., and Key, K., 2012: Bayesian inversion of marine CSEM data with a trans-dimensional self parametrizing algorithm. *Geophysical Journal International*, **191**, 1135–1151.

Rosas-Carbajal, M., Linde, N., Kalscheuer, T., and Vrugt, J. a., 2013: Two-dimensional probabilistic inversion of plane-wave electromagnetic data: Methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*. ISSN 0956-540X. doi:10.1093/gji/ggt482.

Sambridge, M., 2013: A Parallel Tempering algorithm for probabilistic sampling and multimodal optimization. *Geophysical Journal International*, **196**(1), 357–374. ISSN 0956-540X. doi:10.1093/gji/ggt342.

Sambridge, M., Bodin, T., Gallagher, K., and Tkalčić, H., 2013: Transdimensional inference in the geosciences. *Philosophical Transactions of the Royal Society A*, (December 2012).

Sasaki, Y., 2013: 3D inversion of marine CSEM and MT data : An approach to shallow-water problem. *Geophysics*, **78**(1), E59–E65.

Scales, J. A., and Sneider, R., 1997: To Bayes or not to Bayes? *Geophysics*, **62**(4), 1045–1046.

Silva Crepaldi, J. a. L., Pereira Buonora, M. P., and Figueiredo, I., 2011: Fast marine CSEM inversion in the CMP domain using analytical derivatives. *Geophysics*, **76**(5), F303–F313. ISSN 0016-8033. doi:10.1190/geo2010-0237.1.

Sisson, S. a., 2005: Transdimensional Markov Chains. *Journal of the American Statistical Association*, **100**(471), 1077–1089. ISSN 0162-1459. doi:10.1198/016214505000000664.

Tarantola, A., and Valette, B., 1982: Inverse problems= quest for information. *J geophys*, **50**, 159–170.

Trainor-Guitton, W., and Hoversten, G. M., 2011: Stochastic inversion for electromagnetic geophysics: Practical challenges and improving convergence efficiency. *Geophysics*, **76**(6), F373–F386.

Vaidyanathan, P. P., Phoong, S.-M., and Lin, Y.-P., 2010: *Signal processing and optimization for transceiver systems*. Cambridge University Press.

Veevers, J. J., and Johnstone, M. H., 1974: Comparative stratigraphy and structure of the western Australian margin and the adjacent deep ocean floor. *Initial Reports of the DSDP*, **27**, 571–585.

Voronoi, G., 1908: Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites. *Journal für die reine und angewandte Mathematik*, **133**, 97–102. doi:10.1515/crll.1908.133.97.

Weiss, C., 2007: The fallacy of the "shallow-water problem" in marine CSEM exploration. *Geophysics*, **72**(6), A93–A97.

Young, M. K., Rawlinson, N., and Bodin, T., 2013a: Transdimensional inversion of ambient seismic noise for 3D shear velocity structure of the Tasmanian crust. *Geophysics*, **78**(3), WB49–WB62.

Young, M. K., Tkalcić, H., Bodin, T., and Sambridge, M., 2013b: Global P wave tomography of Earth's lowermost mantle from partition modeling. *Journal of Geophysical Research: Solid Earth*, **118**(10), 5467–5486. ISSN 21699313. doi:10.1002/jgrb.50391.

Young, P. D., and Cox, C. S., 1981: Electromagnetic active source sounding near the East Pacific Rise. *Geophysical Research Letters*, **8**, 1043–1046.

Chapter 5

Conclusions

This dissertation has attempted to weave together the worlds of marine CSEM and Bayesian inversion, which have so far been fairly disparate topics, at least in the opinion of the author. Further, the method of inversion we have applied is trans-dimensional and utilizes the Reversible Jump MCMC scheme. It is a powerful tool that allows us to automatically carry out a second level of inference - namely what is a good parameterization to model the data with?

In Chapter 2 we have detailed the mathematical innards of the RJ-MCMC algorithm to the point that any reasonably capable geophysicist will be able to code the method up and apply it to their problem of interest. To address the computational demands made by RJ-MCMC, in Chapter 3 we implemented a technique in which parallel MCMC chains with different annealing factors are allowed to interact and exchange information, drastically cutting down the amount of sampling required. In a recent development (pers. comm. with J. Dettmer), we have removed the restriction on allowing only adjacent chains to interact and have allowed all unique pairs of chains to exchange models, which allows for even faster convergence to the posterior ensemble. In fact, we have been able to adapt this modified algorithm with relative ease for solving the 1D MT inverse problem (pers. comm. with D. Alumbaugh). In Chapter 4 we have successfully solved a CSEM inverse problem which requires a 2D model parameterization for data acquired over the Scarborough gas field in NW Australia. We have been able to delineate an existing hydrocarbon reservoir which has consistently proved to be a difficult CSEM inversion target.

I would like to take this opportunity to stress the fact that marine CSEM unlike its high frequency counterparts such as ground penetrating radar, is not 'just like the seismic method, only lower frequency.' As explained in Chapter 1, the physics of diffusive propagation offer intrinsically low resolution, but may yet be used to indicate presence of resistive structure as opposed to stratigraphic information.

Uncertainty is an inescapable aspect of our lives that we need to embrace, especially in geophysical inversion. All exploration geophysics is essentially decision support – and decisions made with insufficient information can have very negative consequences. With the rapid rate at which computational power is increasing, the application of data assimilation and inference techniques to complicated real world problems such as speech recognition and machine learning is now routine. These techniques are intrinsically probabilistic and I expect to see their application to geophysical problems in the near future.

To be honest we are still novices at interpreting uncertainty and making decisions based on our knowledge of uncertainty. This is true in the geophysical world, both in the realm of academia as well as industry. The challenge will be to integrate sufficient knowledge of physics, geology and machine learning into a framework that provides interpretable uncertainty information on which to base real world decisions. Although I cannot claim to have much knowledge of any of the three fields recently alluded to, this dissertation is what I hope will be a step in that direction.