

---

Electronic Thesis and Dissertation Repository

---

12-17-2012 12:00 AM

## Estimating Subsurface Anisotropy: Combining Waveform Tomography and Simulated Annealing

Michael Victor Afanasiev  
*The University of Western Ontario*

Supervisor  
Gerhard Pratt  
*The University of Western Ontario*

Graduate Program in Geophysics  
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science  
© Michael Victor Afanasiev 2012

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Geophysics and Seismology Commons](#)

---

### Recommended Citation

Afanasiev, Michael Victor, "Estimating Subsurface Anisotropy: Combining Waveform Tomography and Simulated Annealing" (2012). *Electronic Thesis and Dissertation Repository*. 1000.  
<https://ir.lib.uwo.ca/etd/1000>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

ESTIMATING SUBSURFACE ANISOTROPY: COMBINING  
WAVEFORM TOMOGRAPHY AND SIMULATED ANNEALING

(Spine title: Estimating Subsurface Anisotropy using Simulated  
Annealing)

(Thesis format: Monograph)

by

Michael Victor Afanasiev

Graduate Program in Geophysics

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Master of Science

The School of Graduate and Postdoctoral Studies  
The University of Western Ontario  
London, Ontario, Canada

© Michael Victor Afanasiev 2012



THE UNIVERSITY OF WESTERN ONTARIO  
School of Graduate and Postdoctoral Studies

**CERTIFICATE OF EXAMINATION**

Examiners:

Supervisor:

.....  
Dr. G. Atkinson

.....  
Dr. Gerhard Pratt

.....  
Dr. J. Gerhard

.....  
Dr. K. Tiampo

The thesis by

**Michael Victor Afanasiev**

entitled:

**Estimating Subsurface Anisotropy: Combining Waveform Tomography and  
Simulated Annealing**

is accepted in partial fulfillment of the  
requirements for the degree of  
Master of Science

.....  
Date

.....  
Chair of the Thesis Examination Board

# Abstract

Crosshole seismic tomography has been used by Vale to provide geophysical images of mineralized massive sulphides in the Eastern Deeps deposit at Voisey's Bay. High resolution seismic images are presented by applying acoustic waveform tomography to these data. In waveform tomography, an initial model is required which predicts the first arrival traveltimes to within a half-cycle of the lowest frequency in the data. Because seismic velocity anisotropy can be significant, the initial model must quantify the anisotropy, as well as the velocity, in order to meet the half-cycle criterion.

In our case study, difficulties were encountered in generating an accurate anisotropy model through traveltime tomography, and the starting model for waveform tomography failed the half-cycle criterion. We formulate a new, semi-global approach for finding the best-fit 1-D elliptical anisotropy model using simulated annealing, and successfully apply this technique to the Voisey's Bay dataset, as well as synthetically generated datasets.

**Keywords:** Waveform Tomography, Simulated Annealing, Anisotropy, Inverse Methods

# Acknowledgements

From my days as an undergraduate in the physics department at McMaster, I would like to thank Greg Stinson, Hugh Couchman, and James Wadsley for instilling within me an interest in high performance computing, and for drilling home the importance of becoming a good programmer. As well, I mustn't forget the influence of a group of outstanding gentlemen, and our late-night 'yes, but why' gatherings in Cootes Paradise. Those answers certainly kept the interest alive.

In London, thanks go to Laura Sanchez, Dave Edey, Tim Officer, Nicolle Bellisimo, and Jessica Flynn, for making time spent here such a pleasure. As well, I need to acknowledge my fellow members of the Western cover band A-league team 'Aerial Knife': Tom Ulanowski, Brian Branfireun, and Bart Nowack. No matter how busy things got, there was always time for a jam night with the boys. Thanks very much to the helpful administrative staff at UWO, especially Kevin Jordan and Marie Schell. Discussions with the ever-knowledgeable Norm Duke were invaluable in understanding the geology at Voisey's Bay. Thanks to Barry Price for his technical support, and also to Bernie Dunn, for sharing his vast knowledge of Linux systems, and for always finding a way out of whatever computational pickle I got myself into. Never again will I write gigabytes of data to /tmp.

At Vale, I'd like to thank Glenn McDowell and Krystal O'Neil for providing the dataset, and answering any questions I had regarding equipment or acquisition. As well, thanks to Jennifer MacLean from MITACS, and both Vale and MITACS for their financial support.

I would like to extend a heartfelt thanks to my mentor, lab-mate, and friend, (soon-to-be-Dr.) Rie Kamei. Her influence is felt throughout this thesis, and her extensive geophysical knowledge and expertise were consistent motivators for my own self-improvement. It was from her that I was first introduced to the topic of global inverse methods. As well, I must thank Brendan Smithyman for great conversations on all things related to computers and geophysics. The python SEG-Y library that he wrote helped immensely in producing the images in this thesis.

I'd like to especially thank Dr. Gerhard Pratt for his mentorship and supervision over the past year and a half. Scientific role models don't come better, and his (bottomless) knowledge, encouragement, and the freedom he offered my research, helped turn waveform tomography from project to passion. His scientific reputation also allowed me to travel and collaborate with researchers from around the world, and of those I'd like to especially thank Jean Virieux, Calin Cosma, and Heiner Igel for their specific suggestions and influence.

Finally, thanks to my friends, my family, and most of all, to my parents. Without your love and support, none of this would be possible.

# Statement of Co-Authorship

The source code for the computational implementations of traveltime tomography and waveform tomography was provided by Dr. Gerhard Pratt. Significant modifications to this code (such as the implementation of simulated annealing) were developed entirely by myself. The processing of the real data described in Chapter 3 was also performed by myself, with suggestions and advice from Dr. Gerhard Pratt, and Rie Kamei. The creation of the synthetic models in Chapter 4 was assisted by code written by Rie Kamei. The interpretations of the failures of anisotropic traveltime tomography, and the success of waveform tomography, are my own.

The real data described in Chapter 3 were collected by Vale, with the supervision of Krystal O'Neil and Glenn McDowell.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Statement of Co-Authorship</b>	<b>v</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Elastic Wave Equation . . . . .	2
1.2 Thomsen’s Parameters . . . . .	7
1.3 Seismic Tomography . . . . .	8
1.4 Objective of Thesis . . . . .	10
<b>2 Theory</b>	<b>14</b>
2.1 Traveltime Tomography . . . . .	17
2.2 Waveform Tomography . . . . .	21
2.3 Simulated Annealing . . . . .	26
<b>3 Results from a Real Crosshole Dataset</b>	<b>30</b>
3.1 Geologic Background . . . . .	30
3.2 Survey Details . . . . .	32
3.3 Traveltime Tomography . . . . .	37
3.4 Waveform Tomography . . . . .	42
3.4.1 Anisotropy . . . . .	47
3.4.2 Velocity . . . . .	53
3.5 Future Survey Design . . . . .	57
<b>4 Results from a Synthetic Crosshole Dataset</b>	<b>59</b>

4.1	Model Details . . . . .	59
4.2	Traveltime Tomography . . . . .	60
4.3	Waveform Tomography . . . . .	62
4.3.1	Anisotropy . . . . .	62
4.3.2	Velocity . . . . .	64
4.3.3	An-elliptic VTI Media ( $\delta = 0, \phi = 0^\circ$ ) . . . . .	69
4.3.4	An-elliptic TTI Media ( $\delta = 0, \phi = 30^\circ$ ) . . . . .	71
4.4	Discussion . . . . .	71
<b>5</b>	<b>Conclusions</b>	<b>76</b>
5.1	Conclusions and Discussion . . . . .	76
5.2	Final Remarks . . . . .	79
	<b>Bibliography</b>	<b>81</b>
	<b>Curriculum Vitae</b>	<b>87</b>

# List of Figures

3.1	Geologic cross section of the Eastern Deeps deposit. . . . .	31
3.2	Crosshole survey geometry . . . . .	33
3.3	Selected raw receiver gathers . . . . .	36
3.4	Frequency-wavenumber spectrum of a typical receiver gather. . . . .	38
3.5	Traveltime velocity and anisotropy models . . . . .	40
3.6	1-D anisotropy profiles for real data, and objective function behavior during simulated annealing . . . . .	42
3.7	Real, frequency domain data files: before pre-processing . . . . .	44
3.8	Real, frequency domain data files: after pre-processing . . . . .	46
3.9	Real data source inversions . . . . .	47
3.10	Flowchart illustrating the steps in simulated annealing. . . . .	49
3.11	Real and synthetic receiver gathers, showing the effects of simulated annealing and waveform tomography . . . . .	51
3.12	True and estimated frequency domain data files, showing the effect of simulated annealing and waveform tomography . . . . .	52
3.13	Examination of wavefield residuals at 1300 Hz for real data . . . . .	54
3.14	Final velocity model obtained from waveform tomography . . . . .	56
4.1	True models used in synthetic tests . . . . .	61
4.2	Velocity and anisotropy models obtained from traveltime tomography using synthetic data . . . . .	63
4.3	Examination of wavefield residuals at 1300 Hz for synthetic data . . . . .	64
4.4	True and estimated frequency domain data files, showing the effect of simulated annealing and waveform tomography . . . . .	65
4.5	1-D anisotropy profiles for synthetic data, and objective function behavior during simulated annealing . . . . .	66
4.6	Results from applying waveform tomography to elliptic VTI synthetic data	67
4.7	Real and inverted velocity profiles for synthetic data . . . . .	68
4.8	Inverted 1-D anisotropy profiles for an-elliptic TI synthetic data . . . . .	69



4.9	Results from applying waveform tomography to an-elliptic VTI synthetic data . . . . .	70
4.10	Results from applying waveform tomography to an-elliptic TTI synthetic data . . . . .	72
4.11	Wavepath and raypath overlaid on an image of traveltime velocity error .	75

# List of Tables

1.1	List of elastic properties discussed in this thesis. . . . .	12
1.2	List of some mathematical symbols . . . . .	13
3.1	Parameters for travelttime tomography. . . . .	39

# Chapter 1

## Introduction

We experience life as waves. Imagine sitting on the ocean's edge, at sunset. Immediately before us, waves are lapping at the shore. They were set in motion a great distance away, and arrive here carrying energy, some to be transferred kinetically to the continental bedrock, and some to be dissipated as heat, and as sound. These new forms of energy are again, waves: minute vibrations are set up on land, the sound wave propagates through the atmosphere, heat increases the kinetic energy of surrounding particles, altering their wavefunctions; a wave of probabilities. We see the ocean itself as the vibrations of a sea of electromagnetic waves: the dark blue of the water coming from high frequencies, with the deep red of the sunset at the low frequencies. Outside this limited, visible spectrum, low frequency radio waves are racing across the Earth, high frequency X and gamma rays are being created as particles interact at the edge of our atmosphere, and, with a frequency of one hundred and sixty billion cycles per second, the electromagnetic field is vibrating more or less uniformly throughout the universe; the echo of the big bang. Reality seems to be built on the influence of a multitude of waves, of different types and frequencies, which are integrated over change (time) to become experience.

One of the most important properties of these waves is the fact that they carry *information*. From the astronomer, studying the electromagnetic spectrum of galaxies billions of light years away, to the physicist, studying quantum wavefunctions with a pen and paper, to the geophysicist, studying the propagations of elastic waves through the Earth, one of the key jobs of a scientist is to *extract* information from waves in a meaningful manner, and use that information to parameterize a *model* of the system under study. Fortunately, many types of waves follow a relatively simple set of rules, which, if understood, allow the model parameters to be estimated. The focus of this thesis is the parameterization of an Earth model through the observation of seismic waves, which propagate elastically through the Earth.

## 1.1 The Elastic Wave Equation

When discussing a given system, there are generally two properties that are used to describe any inherent symmetries: Spatial variation, and directional variation. The term ‘isotropy’ refers to systems which are directionally symmetric, and look identical no matter what rotation is applied. For example, if a system is isotropic, movement in a certain direction will be indistinguishable to movement in any other direction, and an isotropic system is said to have infinitely many symmetry planes and symmetry axes. Homogeneity refers to uniformity in composition. For example, if a system is homogenous, it will appear the same at any position. The opposites of these properties, anisotropy, and heterogeneity, refer to systems with a countable number of symmetry planes, and variable systems, respectively.

When applying the concepts of anisotropy and heterogeneity to the Earth, heterogeneity is likely the easiest to imagine. The uneven appearance of rocks on a beach, or that of an exposed rock face, are good examples. As well, we know the density of rocks ranges from  $0.2 \text{ g/cc}$  for Pumice (Whitham and Sparks, 1986) to  $12.9 \text{ g/cc}$  in the Earth’s inner core (Smylie, 1992), and seismic wave velocities can vary from less than  $1 \text{ km/s}$  in soils (Milsom, 2003) to over  $13 \text{ km/s}$  near the core-mantle boundary (Dziewonski and Anderson, 1981). The identification of anisotropy is more subtle. If we are interested in the anisotropic propagation of seismic waves through the Earth, we are looking to describe the variation of wave velocity with angle. Thomsen (1986) gives four main factors which influence the anisotropic propagation of seismic waves:

1. Intrinsic anisotropy, due to symmetric or preferred orientations of mineral grains;
2. The thin bedding of isotropic layers on a scale which is smaller than the seismic wavelength;
3. The presence of fractures or microcracks, which have a preferred orientation;
4. Nonhydrostatic stresses.

Anisotropy resulting from factors 2-4 is known as *extrinsic* anisotropy, from the fact that it is external to any single mineral property. It is readily apparent that extrinsic anisotropy and heterogeneity are interchangeable in some cases; specifically, their relative importance depends on the scale under consideration. If we consider a subsurface region which is composed of finely bedded, isotropic layers, at short wavelengths this region will appear isotropic and heterogeneous, but at large wavelengths, the region will appear

anisotropic and homogeneous (Backus, 1962). In the following discussion, we will see how seismic anisotropy can arise from any extrinsic symmetries in the elastic tensor.

The wave equation for a general, elastic, anisotropic, heterogeneous medium is

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \tau_{ij}}{\partial x_j} = f_i \quad (1.1)$$

(Aki and Richards, 2002; Tsvankin, 2012), where  $\tau_{ij}$  is the stress tensor,  $\rho$  is the density,  $\mathbf{u} = (u_1, u_2, u_3)$  is the displacement vector,  $\mathbf{f} = (f_1, f_2, f_3)$  is the external force per unit volume,  $t$  is the time, and  $x_i$  are the Cartesian coordinates. Here, the Einstein summation convention is used, i.e. summation over repeated indices is implied. We are usually interested in relating the displacement field  $\mathbf{u}$  to the elastic properties of rocks. To do this, we first relate the stress tensor to strain, and then the strain to displacement. If the stress-strain relation is linear, as it is for a small strain (such as that induced by the passage of a seismic wave (Aki and Richards, 2002)), the generalized Hooke's law gives

$$\tau_{ij} = c_{ijkl} e_{kl} \quad (1.2)$$

where  $c_{ijkl}$  is the general fourth-order elastic tensor, and  $e_{kl}$  is the strain tensor

$$e_{kl} = \frac{1}{2} \left( \frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right). \quad (1.3)$$

Substituting equations (1.2) and (1.3) into equation (1.1) gives

$$\rho \frac{\partial^2 u_i}{\partial t^2} - c_{ijkl} \frac{\partial^2 u_k}{\partial x_j \partial x_l} = f_i. \quad (1.4)$$

This is the general displacement formulation of the elastic wave equation, which is valid for anisotropic, heterogeneous media. If a trial solution of the form

$$u_k = U_k e^{(i\omega(n_j x_j / v - t))} \quad (1.5)$$

is chosen, and we attempt to solve the homogeneous (source free) form of equation (1.4)

$$\rho \frac{\partial^2 u_i}{\partial t^2} - c_{ijkl} \frac{\partial^2 u_k}{\partial x_j \partial x_l} = 0, \quad (1.6)$$

we obtain three solutions (Tsvankin, 2012), which correspond to three modes of wave propagation: Pressure (P-wave), horizontal shear (SH-wave), and vertical shear (SV-wave). For the majority of this thesis, we will focus solely on P-waves. Note that so far

we have made a simplifying assumption, which is that the stress-strain relation is linear. This is important, as higher order terms in equation (1.3) would lead to a nonlinear wave equation, which is very difficult to solve. Fortunately, this linear relationship is appropriate for the study of seismic wave propagation (Tsvankin, 2012).

Examining equation (1.4), we see that if the source  $\mathbf{f}$  is known, we must know the values of  $c_{ijkl}$  before solving for the displacement field  $\mathbf{u}$ . The general tensor has  $3^4 = 81$  components, but inherent symmetries reduce the number of independent components. First, because the stress and strain tensors are symmetric, the indices  $i, j, k,$  and  $l$  can be interchanged in the following way

$$c_{ijkl} = c_{jikl}, c_{ijkl} = c_{ijlk}. \quad (1.7)$$

As well, from thermodynamic considerations

$$c_{ijkl} = c_{klij} \quad (1.8)$$

(Aki and Richards, 2002). These symmetries reduce the number of independent components from 81 to 21. At this stage, it is common to write the elastic tensor as a symmetric  $6 \times 6$  matrix by following the Voigt recipe for index pairs  $(i, j)$  and  $(k, l)$ :

$$\mathbf{C}_{ij} \equiv \mathbf{C}_{ji} \equiv c_{(ij), (kl)} : 11 \mapsto 1, 22 \mapsto 2, 33 \mapsto 3, 23 \mapsto 4, 13 \mapsto 5, 12 \mapsto 6. \quad (1.9)$$

For the full, 21 independent elastic parameters, the Voigt-transformed elastic tensor can be written as:

$$\mathbf{C}_{ij} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} \\ c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{12} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{13} & c_{23} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{14} & c_{24} & c_{34} & c_{44} & c_{45} & c_{46} \\ c_{15} & c_{25} & c_{35} & c_{45} & c_{55} & c_{56} \\ c_{16} & c_{26} & c_{36} & c_{46} & c_{56} & c_{66} \end{pmatrix}, \quad (1.10)$$

where the symmetries expressed in equation (1.8) have been exploited, resulting in  $\mathbf{C}_{ij} = \mathbf{C}_{ji}$ . Materials possessing this most general of symmetries are known as ‘triclinic’. Triclinic symmetry can result from, among other things, multiple fracture axes, with each axis at an arbitrary angle to the others. Unfortunately, the large parameter space of triclinic models makes any estimations of the parameters severely underdetermined, and usually additional symmetries are assumed which allow us to decrease the number

of independent quantities.

Assuming monoclinic symmetry reduces the number of independent elastic parameters to 13. This stems from the fact that monoclinic media possess a plane of mirror symmetry. This behavior can result from three or more sets of identical fractures, with each fracture set possessing a symmetry axis orientated arbitrarily with respect to other fracture sets. Further reductions to the number of parameters are obtained in the presence of orthorhombic symmetry, which reduces the required number of parameters to 9. Orthorhombic symmetry can result from three mutually orthogonal fracture systems, or two identical fracture systems at an arbitrary angle to each other. One further parameter space contraction can be made, while still respecting anisotropic behavior, and this is by assuming transversely isotropic media.

Transversely isotropic (TI) media possess a single axis of symmetry. If the symmetry axis is tilted at an arbitrary angle, we use the term ‘tilted transversely isotropic’ (TTI), and if the symmetry axis is aligned with vertical, we say the material is ‘vertically transversely isotropic’ (VTI). The Voigt-transformed elastic tensor for VTI media is

$$\mathbf{C}_{ij} = \begin{pmatrix} c_{11} & c_{11} - 2c_{66} & c_{13} & 0 & 0 & 0 \\ c_{11} - 2c_{66} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & c_{66} \end{pmatrix}, \quad (1.11)$$

and the total number of independent parameters is now 5. For more general TTI media, there are still only 5 independent elastic parameters, but additional information on the orientation of the symmetry axis is needed. General TI media can result from the presence of a single fracture system. As well, Backus (1962) showed that finely layered media can appear as effectively TI in the long-wavelength limit (as mentioned above).

If the medium is assumed to be isotropic, which corresponds to the medium possessing an infinite number of symmetry planes, the number of independent elastic parameters reduces to 2:  $\eta$  (the modulus of fluid incompressibility), and  $\mu$  (the shear modulus), known as the ‘Lamé parameters’. The Voigt-transformed elastic tensor for isotropic

media is

$$\mathbf{C}_{ij} = \begin{pmatrix} \eta + 2\mu & \eta & \eta & 0 & 0 & 0 \\ \eta & \eta + 2\mu & \eta & 0 & 0 & 0 \\ \eta & \eta & \eta + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}, \quad (1.12)$$

and the complete isotropic elastic wave equation can be written as

$$\rho \frac{\partial^2 u_i}{\partial t^2} - (\eta + \mu) \frac{\partial^2 u_j}{\partial x_i \partial x_j} - \mu \frac{\partial^2 \mu_i}{\partial x_j \partial x_j} = f_i. \quad (1.13)$$

Another way to further simplify the wave equation is to consider wave propagation within a fluid. In this case, there is no support for shear stresses, so the shear modulus ( $\mu$ ) is zero. As a result, only P-waves exist in an acoustic medium. Although the Earth's crust is a fully elastic medium, if one is only trying to model P-waves (as is common in exploration seismology), then using the acoustic wave equation is a useful approximation (due to the reduction of computational effort required to model acoustic wave propagation, when compared to the elastic case). In the isotropic approximation, the acoustic stress tensor first introduced in equation (1.1) is simply

$$\tau_{ij} = -P\delta_{ij}, \quad (1.14)$$

where  $P$  is the hydrostatic pressure, and  $\delta_{ij}$  is the Kronecker delta. Substituting this into equation (1.1) gives

$$\rho \frac{\partial^2 u_i}{\partial t^2} + \frac{\partial P}{\partial x_i} = f_i. \quad (1.15)$$

Dividing this relationship by the density  $\rho$ , and taking the divergence, gives

$$\nabla \cdot \frac{\partial^2 \mathbf{u}_i}{\partial t^2} + \nabla \cdot \left( \frac{1}{\rho} \right) \nabla P = \nabla \cdot \left( \frac{1}{\rho} \right) f_i. \quad (1.16)$$

Using Hooke's Law for a fluid

$$P \equiv -\kappa \nabla \cdot \mathbf{u} \quad (1.17)$$

we define the pressure  $P$  in terms of the displacement ( $\mathbf{u}$ ) and the bulk modulus ( $\kappa$ ), resulting in

$$\left[ \frac{1}{\kappa} \frac{\partial^2}{\partial t^2} - \nabla \cdot \left( \frac{1}{\rho} \right) \nabla \right] P = -\nabla \cdot \left( \frac{1}{\rho} \right) \mathbf{f}. \quad (1.18)$$



This is the acoustic wave equation, which describes the propagation of a pressure wave in medium with an (isotropic) velocity of

$$c = \sqrt{\frac{\kappa}{\rho}}. \quad (1.19)$$

## 1.2 Thomsen's Parameters

To assist in the interpretation of seismic anisotropy in TI media, Thomsen (1986) introduced a set of parameters which are now commonly known as ‘Thomsen’s parameters’. The purpose of these parameters is to express the influence of anisotropy in terms of the velocities along the symmetry axes. For pressure waves in TI media, within the symmetry plane the following parameters are defined:

$$\begin{aligned} V_{P0} &\equiv \sqrt{\frac{c_{33}}{\rho}}; \\ \varepsilon &\equiv \frac{c_{11} - c_{33}}{2c_{33}}; \\ \delta &\equiv \frac{(c_{13} + c_{55})^2 - (c_{33} - c_{55})^2}{2c_{33}(c_{33} - c_{55})}. \end{aligned} \quad (1.20)$$

Here,  $V_{P0}$  refers to the P-wave velocity along the symmetry axis,  $\varepsilon$  refers to the fractional difference in velocities between the fast and slow axes, and  $\delta$  defines the behavior of  $V_p$  close to the slow (symmetry) axis. The dependence of velocity on angle, in terms of  $\varepsilon$  and  $\delta$ , can be approximated to first order with the equation

$$V_P(\theta) = V_{P0}(1 + \delta \sin^2 \theta + (\varepsilon - \delta) \sin^4 \theta) \quad (1.21)$$

(Thomsen, 1986). By using Thomsen notation, and if the orientation of the symmetry plane is known, the number of independent parameters is reduced to 3. Of special relevance to this report is the case where  $\varepsilon = \delta$ , which is known as elliptical anisotropy. This case is especially important, as the anisotropic wavefield can be completely characterized (kinematically) by a simple geometrical stretching along the slow axis (Helbig, 1983; Dellinger, 1991; Pratt et al., 2004), and the number of independent parameters is reduced once again to 2.

### 1.3 Seismic Tomography

In all but the simplest cases, the elastic and acoustic wave equations cannot be solved analytically. As such, computers are used to numerically estimate solutions. These solutions, in an abstract sense, can be considered as representing *forward models*. That is, given the parameters describing a system (i.e. the elastic constants  $c_{ijkl}$ , the density  $\rho$ , or the bulk modulus,  $\kappa$ ), and a source  $\mathbf{f}$ , they describe a physical result (i.e. the displacement field  $\mathbf{u}$ , or the pressure field  $P$ ). The topic of this thesis is just the opposite, in that given a physical result, the goal is to infer information on a set of parameters that led to that result. When the physical result is a collection of transmitted seismic waveforms, the term ‘seismic tomography’ is used, the mechanics of which are under the umbrella of *inverse theory*, which will be reviewed in Chapter 2. Before any formal discussion of inverse theory, I will quickly introduce ‘waveform tomography’, a subset of seismic tomography, which is the subject of this thesis.

Historically, seismic tomography has used asymptotic ray theory (ART) to approximate the propagation of elastic and acoustic waves. This approach assumes that seismic energy propagates along infinitely thin raypaths between source and receiver locations, and that the raypaths refract and reflect through heterogeneous media in accordance with Fermat’s law (asymptotically correct for waves of an infinite frequency) (Červeny, 2001). ART is attractive due to the low computational cost of modelling the raypaths, and the approximation is valid as long as any heterogeneities present are large compared to the dominant wavelength (Červeny, 2001). The tomographic reconstruction of Earth parameters (usually seismic velocity) through the use of ART is commonly known as traveltime tomography. In traveltime tomography, the relative arrival times of seismic energy, emanating from a source at time zero, are identified at a series of receivers. This is known as ‘first arrival picking’, or ‘first break picking’. Within the assumptions of ray theory, this traveltime  $T$  can be related to an integration over a raypath  $\mathcal{L}$  through the equation (Červeny, 1972; Červeny and Jech, 1982; Chapman and Pratt, 1992)

$$T = \int_{\mathcal{L}} \frac{1}{\xi(\mathbf{x})} d\mathcal{L}, \quad (1.22)$$

where  $\xi(\mathbf{x})$  is the spatially variable velocity. For a small perturbation to the traveltime  $\delta T$ , we can write

$$\delta T = \int_{\mathcal{L}} \frac{1}{\delta\xi(\mathbf{x})} d\mathcal{L}. \quad (1.23)$$

It is important to note that equation (1.23) is linear for small perturbations. For large perturbations, the raypath  $\mathcal{L}$  will itself change, and equation (1.23) will no longer be

linear. The goal of traveltimes tomography is to produce a velocity distribution  $\xi(\mathbf{x})$  that accurately predicts the traveltimes  $T$ , and we call the sum of the squared differences between the true and predicted traveltimes the ‘data misfit’, or ‘residuals’. Common practice, which will be expanded upon in Chapter 2, is to iteratively update  $\xi(\mathbf{x})$  until the data misfit is minimized: Iterations where  $\mathfrak{L}$  is kept constant are called *linear* iterations; whenever  $\mathfrak{L}$  is recalculated, we call this a *nonlinear* iteration.

Due to the phenomena of wavefront healing, the spatial resolution of traveltimes tomography is limited to the Fresnel zone (Williamson and Worthington, 1993). Regardless of this limitation, and the condition on validity discussed above, traveltimes tomography has enjoyed much success in both exploration and academic contexts (Zelt and Barton, 1998).

As the power of computers has increased, it has become feasible to move beyond the ray approximation, and instead to numerically model wave propagation through the Earth using the elastic or acoustic wave equations. To accomplish this, Earth properties are usually discretized onto a grid or mesh, and the wave equations are solved through finite-difference (Virieux, 1986; Pratt, 1990), finite-element (Shin and Cha, 2008), or spectral-element (Komatitsch and Tromp, 2002) methods. Tomographic methods using these techniques are commonly referred to as ‘full waveform inversion’, or ‘waveform tomography’, and were first introduced by Lailly (1983) and Tarantola (1984). As in traveltimes tomography, waveform tomography seeks to iteratively update Earth parameters with the goal of producing a match between synthetically generated and observed data. However, in waveform tomography the data are seismic waveforms, instead of first arrival times. Unlike traveltimes tomography, wave behavior in complex, heterogeneous media can be modelled accurately, and the resolution limit of waveform tomography is on the order of the seismic wavelengths used in the tomographic reconstruction (Wu and Toksoz, 1987). As well, the use of the wave equation negates the need to approximate wave propagation along an infinitely thin raypath. Waveform tomography allows the true multipath nature of the wave equation to be realized, and the result is simulated propagation along a spatially broad ‘wavepath’ (Woodward, 1992).

Waveform tomography is a popular tomographic method in both industrial and academic contexts. The technique has been employed successfully across a large range of scales with both synthetically generated data (Brenders and Pratt, 2007a,b) and real data (Pratt et al., 2004, 2008; Prioux et al., 2011; Fichtner et al., 2009; Kamei et al., 2012). A major difficulty in waveform tomography stems from the nonlinearity of the inverse problem, coupled with the high computational cost of numerically modelling the wave equation. These difficulties can be mitigated by ensuring that the starting model

for waveform tomography, commonly found through travelttime tomography (Pratt et al., 2004; Brenders and Pratt, 2007b; Virieux and Operto, 2009), is very similar to the true model, and is close to what we call the ‘global minimum’ of the inverse problem (Sirgue and Pratt, 2004). As we will see in the upcoming chapters, the degree of similarity required is directly linked to the frequency bandwidth of the data. As well, instead of iteratively improving the model by locally minimizing some metric of error (i.e. the difference between observed and modelled waveforms), global inverse methods may be used (Sen and Stoffa, 1995, 1991; Tarantola, 2005), although these methods only mitigate the effects of nonlinearity, and in fact greatly increase the computational burden. As such, they are not widely used in the context of waveform inversion, although some examples do exist (such as Sen and Stoffa, 1991).

In an exploration context, waveform tomography has been applied extensively in the oil and gas industry, but its adoption rate has been significantly slower in mineral exploration. Although examples do exist (Xu and Greenhalgh, 2010), this thesis represents one of the first applications of waveform tomography to a seismic dataset acquired for the purposes of mineral exploration.

## 1.4 Objective of Thesis

In this thesis, I investigate the application of waveform tomography, based on the acoustic approximation given in equation (1.18) to both a real and synthetically generated crosshole seismic dataset. In Chapter 2, a review of relevant topics in geophysical inverse theory is given. Chapter 3 describes how waveform tomography was applied to a real crosshole seismic dataset, collected by Vale in Voisey’s Bay, Canada. A walkthrough of the workflow, from first-arrival picking to the completion of waveform tomography, is given. Major challenges were faced during the processing of this dataset, and these challenges, along with their solutions, are described. In Chapter 4, I describe the creation and processing of a synthetic crosshole seismic dataset, which was created to understand the cause of the challenges described in Chapter 3. A physical interpretation of the cause of these challenges is given. Finally, Chapter 5 presents some conclusions, and suggestions for future work.

As we shall see, the generation of an appropriate starting model for use in waveform tomography is one of the most important, and difficult, aspects of the technique. The original work I present here concerns a *semi-global* method for discovering an appropriate starting anisotropy model; a method which requires little-to-no human intervention, and one which blazes the trail towards a stable, automatic, and simple workflow for waveform

tomography.

Symbol	Description
$c_{ijkl}$	Elastic tensor
$\rho$	Density
$\tau_{ij}$	Stress tensor
$\mathbf{f}$	Source field
$e_{kl}$	Strain tensor
$\mathbf{C}_{ij}$	Voigt-transformed elastic tensor
$\eta$	Lamé's modulus of fluid incompressibility
$\mu$	Lamé's shear modulus
$P$	Pressure
$\delta_{ij}$	Kronecker delta
$V_{P0}$	Thomsen's velocity along a symmetry axis for TTI media
$\varepsilon$	Thomsen's epsilon, the fractional velocity difference for TTI media
$\delta$	Thomsen's delta, controlling the behavior of velocity close to the symmetry axis
$\kappa$	Bulk modulus
$c$	Pressure wave velocity

Table 1.1: List of elastic properties discussed in this thesis.

Symbol	Description
$\mathfrak{M}$	<i>Model space</i> , inhabited by elements of $\mathbf{m}$
$\mathfrak{D}$	<i>Data space</i> , inhabited by elements of $\mathbf{d}$
$\mathbf{d}$	Vector of true data points
$\bar{\mathbf{d}}$	Vector of estimated data points
$\mathbf{m}$	True model parameters
$\bar{\mathbf{m}}$	Estimated model parameters
$g$	Function performing the mapping $\mathfrak{M} \mapsto \mathfrak{D}$
$g^{-1}$	Function performing the mapping $\mathfrak{D} \mapsto \mathfrak{M}$
$P(\mathbf{m})$	Probability distribution over $\mathfrak{M}$ , the solution to the inverse problem
$E(\mathbf{m})$	Function measuring the misfit between $\bar{\mathbf{m}}$ and $\mathbf{m}$ , and/or $\bar{\mathbf{d}}$ and $\mathbf{d}$
$\nabla_m E$	Gradient of $E(\mathbf{m})$ at a point $\mathbf{m}$
$\mathbf{H}$	Hessian matrix
$\partial T$	Small, traveltime perturbation
$\mathcal{L}$	Raypath
$\mathbf{p}$	Traveltime slowness vector
$a_{ijkl}$	Density normalized elastic tensor
$F_{kij}$	Fréchet derivatives for traveltime tomography
$\mathbf{C}_d$	Data covariance matrix
$\mathbf{C}_m$	Model covariance matrix
$\sigma_d$	<i>a priori</i> data variance
$\sigma_m$	<i>a priori</i> model variance
$\lambda$	Relative weight of $\sigma_d$ and $\sigma_m$
$\kappa$	Relative weight of isotropic vs. anisotropic parameter perturbations
$\epsilon$	Weight given to roughness penalization
$\mathbf{S}(\mathbf{x}, \omega)$	Impedance matrix
$\mathbf{u}(\mathbf{x}, \omega)$	Wavefield
$\mathbf{f}(\mathbf{x}, \omega)$	Source field
$\delta \mathbf{d}$	Wavefield residuals at the receiver locations
$\mathbf{G}$	Derivative of $E(\mathbf{m})$ at a point $\mathbf{m}$
$\mathbf{F}$	Virtual source field
$\alpha$	Gradient steplength
$T$	‘Temperature’ for use in simulated annealing
$\mathbb{A}(\mathbf{m}, T)$	Metropolis criterion for a model $\mathbf{m}$ at a ‘temperature’ $T$
$\mathbb{G}(\mathbf{m})$	Probability of generating model $\mathbf{m}$
$K$	Boltzmann’s constant

Table 1.2: List of some mathematical symbols

# Chapter 2

## Theory

As stated in Chapter 1, one of the goals of geophysicists is to extract meaningful information from seismic waves, and use the information to parameterize a model of the Earth. This is the subject of *inverse theory*. Here I provide a review of some general concepts in inverse theory, before performing a quick review of (1) anisotropic travelttime tomography, (2) waveform tomography, and (3) simulated annealing. Common to all methods is the desire to find a physical model  $\mathbf{m}$  which is related to a set of data points  $\mathbf{d}$  by an operator  $g$ , which can be expressed by the general equation

$$\mathbf{d} = g(\mathbf{m}). \tag{2.1}$$

The concept of a model space is defined as a set  $\mathfrak{M}$  spanned by the elements of  $\mathbf{m}$ . As well, we say that elements of  $\mathbf{d}$  span the data space  $\mathfrak{D}$ . Note here that  $\mathfrak{D}$  is usually a much different (and smaller) space than  $\mathfrak{M}$ , as in most inverse problems (such as those described in this thesis) there are usually far fewer data points than there are model parameters. We say then that the action of  $g$  is to perform the mapping  $\mathfrak{M} \mapsto \mathfrak{D}$ , known as the forward problem. Since the model  $\mathbf{m}$  and operator  $g$  will likely not be perfect representations of reality, the actual solution of equation (2.1) will be an estimate of the true data, or  $\bar{\mathbf{d}}$ , and we need to relax the equality.

The purpose of the methods described in this chapter will be to investigate the inverse relationship

$$\mathbf{m} = g^{-1}(\mathbf{d}), \tag{2.2}$$

which estimates a model from a set of data, or performs the mapping  $\mathfrak{D} \mapsto \mathfrak{M}$ . This type of mapping is known as the inverse problem. Unfortunately, the forward mappings discussed in this thesis are often singular or extremely ill-conditioned, so the operator  $g^{-1}$  is impossible to calculate directly. Even if we were to find  $g^{-1}$ , the solution of the



inverse problem would in general not capture all the details of the true model, as the forward operator  $g$  is itself built on simplifying assumptions. This leads to a need to relax the equality in equation (2.2) as well, and only allows us to find an estimate of the true model,  $\bar{\mathbf{m}}$ . With this in mind, we must take care in our interpretation of what it means to find a ‘solution’ to the inverse problem.

Tarantola (2005) defines the solution of the inverse problem as a probability distribution over  $\mathfrak{M}$ ,

$$P(\bar{\mathbf{m}} \in \mathfrak{M}) = \exp\left(-\frac{1}{2}E(\bar{\mathbf{m}})\right), \quad (2.3)$$

where the ‘misfit function’  $E(\bar{\mathbf{m}})$  is a quantitative measure of the quality of  $\bar{\mathbf{m}}$ . A common way to define the misfit function is by the  $L_2$  (least squares) norm of the differences between observed and calculated data, combined with appropriate *a priori* information, or

$$E(\mathbf{m}) = (\bar{\mathbf{d}} - \mathbf{d})^2 + (\bar{\mathbf{m}} - \mathbf{m}_0)^2 = (g(\bar{\mathbf{m}}) - \mathbf{d})^T \mathbf{C}_d^{-1} (g(\bar{\mathbf{m}}) - \mathbf{d}) + (\bar{\mathbf{m}} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\bar{\mathbf{m}} - \mathbf{m}_0). \quad (2.4)$$

Here we have included a measure of uncertainty by the incorporation of data and model covariance matrices,  $\mathbf{C}_d$  and  $\mathbf{C}_m$  respectively. Note that this adoption of the least squares function assumes a Gaussian probability distribution over the data and model spaces. Pausing here to consider the *justification* of assuming Gaussian statistics, it is admitted that there is no sound justification (Pratt and Chapman, 1992) other than historical use and success, and the simplification that it provides to the mathematics. The probabilistic interpretation of inverse problems implies that our goal is to find a model  $\bar{\mathbf{m}}$  which is most likely given the data, and given the prior model. To accomplish this, we can try to maximize equation (2.3), which corresponds to minimizing  $E(\mathbf{m})$ .

We can find the minimum of the misfit function by first expanding equation (2.4) in a Taylor series about  $\bar{\mathbf{m}}$ :

$$E(\bar{\mathbf{m}} + \delta\mathbf{m}) = E(\bar{\mathbf{m}}) + \delta\mathbf{m}^t \nabla_m E(\bar{\mathbf{m}}) + \frac{1}{2} \delta\mathbf{m}^t \mathbf{H} \delta\mathbf{m} + O(\|\delta\mathbf{m}\|^3), \quad (2.5)$$

where  $\delta\mathbf{m}$  is a perturbation about the estimated model  $\bar{\mathbf{m}}$ ,  $\nabla_m E(\bar{\mathbf{m}})$  is the gradient of the misfit function with respect to the model parameters, and  $\mathbf{H}$  is the Hessian, defined as  $\mathbf{H}_{ij} = \frac{\partial^2 E(\mathbf{m})}{\partial m_i \partial m_j}$ . We then differentiate this series with respect to the model perturbations  $\delta\mathbf{m}$ , and set the result to zero, obtaining

$$\delta\mathbf{m} = -\mathbf{H}^{-1} \nabla_m E(\bar{\mathbf{m}}). \quad (2.6)$$

Note that if  $g$  is a linear operator  $\mathbf{G}$ , and all model parameters are independent, the above equation reduces to the normal equation for the least squares solution of a linear system  $\bar{\mathbf{m}} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{d}$  (if  $\mathbf{m}_0 = 0$ ). Unfortunately, this reduction is not valid for the problem presented in this thesis, so many of the powerful theorems of linear inverse theory do not apply. In addition, for many realistic problems  $\mathbf{H}$  is an extremely large matrix, and is very expensive to calculate, let alone invert.

Instead of using equation (2.6), perturbations to the model are usually found by using information on the local gradient of the misfit function, and model updates are iteratively estimated using

$$\mathbf{m}^{(k)} = \mathbf{m}^{(k-1)} - \alpha^{(k-1)} \nabla_m E(\mathbf{m})^{(k-1)}, \quad (2.7)$$

where  $k$  refers to the iteration number. Let us carefully interpret equation (2.7). The location in model space of  $\mathbf{m}$  at iteration  $k$  is equal to the location of  $\mathbf{m}$  at the previous iteration, minus some constant  $\alpha$  times the local gradient direction of the misfit function. What this means is that each successive model is found by moving a distance  $\alpha$  *downhill* in the misfit function. Algorithms which update models in this fashion, by exploiting information on the local derivatives of  $E(\mathbf{m})$ , are known as *local* inverse methods, or descent methods. It is obvious that if  $E(\mathbf{m})$  has a definite single minimum, and  $\nabla E_m(\mathbf{m})$  points downhill for all  $\mathbf{m} \in \mathfrak{M}$ , then a local inverse method will converge to the best-fit model every time. As we will see in the upcoming sections, for problems we are interested in this condition is never satisfied. While there are many clever variations on equation (2.7), all local inverse methods rely on the same concept: That of successive minimizations to the misfit function along the gradient direction, with the assumption that the initial model is strictly ‘uphill’ from the global minimum of  $E(\mathbf{m})$ .

An alternative to local inverse methods are *global* inverse methods. Global inverse methods place little to no weight on the structure of the misfit function, and instead try to find  $\bar{\mathbf{m}}$  by ‘exploring’ the model space  $\mathfrak{M}$ . The simplest example of a global optimization method is a grid search, which tests discrete models  $\mathbf{m}_i$  in  $\mathfrak{M}$ , and which returns an ensemble of models which best fit the data (as evaluated by  $E(\mathbf{m})$ ). A subtle, but very important, point must be made here. In defining  $P(\mathbf{m})$  in equation (2.3), we interpreted the solution of the inverse problem as a probability distribution over  $\mathfrak{M}$ . In the local optimization routines discussed thus far, this interpretation has been suspiciously neglected, and presented a single solution  $\mathbf{m}^k$  which best fits the data. This is because of the deterministic nature of gradient algorithms, which in most cases sample  $\mathfrak{M}$  much too sparsely to make any rigorous claims about  $P(\mathbf{m})$ . Global optimization methods, which by their nature sample the entire model space, overcome this problem, and their solutions lead to the natural interpretation of equation (2.3) as a probability distribution.

Similar to global inverse methods are ‘semi-global’ inverse methods, which sample a large portion of the model space, but as well exploit information on the structure of the misfit function. While extremely robust, global and semi-global inverse methods have not been widely adopted in traveltime or waveform tomography. This is mainly due to the extreme size of the parameter space, and the significant amount of time needed to perform the forward modelling.

I have pointed out some of the strengths and weaknesses of local and global inverse methods, and will now review three specific examples of inverse methods relevant to the subject of this paper.

## 2.1 Traveltime Tomography

Chapman and Pratt (1992) and Pratt and Chapman (1992) showed that a traveltime perturbation  $\delta T$ , for a qP-wave due to a small variation in background elastic properties, can be related to an integration along the corresponding seismic raypath  $\mathfrak{L}$  by the expression

$$\delta T = -\frac{1}{2} \int_{\mathfrak{L}} \frac{1}{\alpha^3} \hat{p}_i \hat{p}_l \hat{p}_j \hat{p}_k \delta a_{ijkl} dl, \quad (2.8)$$

where  $\alpha$  represents the isotropic background velocity,  $\mathbf{p}$  is the slowness vector in an unperturbed, isotropic medium, and  $\delta a_{ijkl}$  is a small perturbation to the general, density-normalized elastic tensor. Note that this is very similar to equation (1.23), although it is more general, as it contains contributions from the full (density-normalized) elastic tensor  $a_{ijkl}$ , not just the isotropic velocity. To examine how we might invert for the parameter perturbations  $\mathbf{p}$ , given a traveltime error  $\delta T$ , I follow the derivation of Pratt and Chapman (1992). If, as is common in 2-D cross-borehole traveltime tomography, the rays are assumed to be confined to the 2-D plane  $(x_1, x_3)$ ,  $\hat{p}_2$  is always 0, and the number of independent terms in the above integral can be reduced to 5:

$$\delta T = -\frac{1}{2\alpha^3} \int_{\mathfrak{L}} (\hat{p}_1^4 \delta q_1 - \hat{p}_1^3 \hat{p}_3 \delta q_2 - \hat{p}_1^2 \hat{p}_3^2 \delta q_3 - \hat{p}_1 \hat{p}_3^3 \delta q_4 - \hat{p}_3^4 \delta q_5) dl, \quad (2.9)$$

where

$$\delta q_1 = \delta a_{1111}; \quad \delta q_2 = 4\delta a_{1131}; \quad \delta q_3 = 2\delta a_{1133} + 4\delta a_{3131}; \quad \delta q_4 = 4\delta a_{3331}; \quad \delta q_5 = \delta a_{3333}. \quad (2.10)$$

We discretize equation (2.9) in order to solve it in a computationally tractable manner,

leaving us with the linear system of equations

$$\delta T_k = F_{kij} \delta q_{ij}, \text{ or } \delta \mathbf{T} = \mathbf{F} \delta \mathbf{q}, \quad (2.11)$$

which relate the parameter perturbations  $\delta q_{ij}$  (the  $j$ th elastic parameter ( $j = 1, \dots, 5$ ) in the  $i$ th region (grid point)) to the travelttime perturbation of the  $k$ th ray  $\delta T_k$  through a multiplication by the Fréchet derivatives  $F_{kij}$ . Through discretization, we replace the integral in equation (2.9) by a summation, and replace the variable of integration with  $\Delta l_{ki}$ , which is the distance that ray  $k$  travels in cell  $i$ . This leaves us with a general expression for the Fréchet derivatives (Pratt and Chapman, 1992)

$$F_{kij} = \frac{\partial T_k}{\partial q_{ij}} = -\frac{1}{2\alpha^3} \hat{p}_1^{5-j} \hat{p}_3^{j-1}, \quad j = 1, \dots, 5. \quad (2.12)$$

Of course, in any realistic tomography experiment, it is unlikely that the isotropic component of velocity  $\alpha$  will be exactly correct. Pratt and Chapman (1992) add an additional parameter  $\delta q_0 = \delta \alpha^2$  into this system of equations to represent perturbations to the isotropic background velocity. This results in the now non-linear system of equations

$$\delta \mathbf{T} = \mathbf{F} \delta \hat{\mathbf{q}}, \quad (2.13)$$

as the parameter update may contain a perturbation to the isotropic velocity, which equation (2.12) is itself dependent on. As discussed in Chapter 1, the distribution of raypaths throughout the model will change with velocity, and as such the term  $\Delta l_{ki}$  will also change. We handle this limitation by iteratively updating the model: after a new  $\alpha$  distribution is obtained, we relinearize equation 2.12 (by re-calculating raypaths) and estimate a new set of parameters  $\delta \hat{\mathbf{q}}$ .

We can relate equation (2.13) to the general relationship given in equation (2.1), with  $\delta \mathbf{T}$  taking to role of the data,  $\delta \hat{\mathbf{q}}$  taking the role of the model, and  $\mathbf{F}$  representing the map from  $\mathfrak{M} \mapsto \mathfrak{D}$ . Of course, we are interested in the inverse mapping  $\mathfrak{D} \mapsto \mathfrak{M}$ . This physically describes how differences in the travelttime residuals  $\delta \mathbf{T}$  are related to changes in the model. To find the best fit model, we define a misfit function similar to equation (2.4) to quantify the data misfit

$$E(\mathbf{m}) = \sigma_d^{-2} (\mathbf{F} \delta \hat{\mathbf{q}} - \delta \mathbf{T})^T (\mathbf{F} \delta \hat{\mathbf{q}} - \delta \mathbf{T}) + \delta \hat{\mathbf{q}}^T \mathbf{C}_p^{-1} \delta \hat{\mathbf{q}}, \quad (2.14)$$

in which we have assumed that the data are uncorrelated and have equal variance ( $\mathbf{C}_d = \sigma_d^2 \mathbf{I}$ ), and that all *a priori* information on the model is contained completely within the

model covariance matrix  $\mathbf{C}_p$  ( $\delta\hat{\mathbf{q}}_0$  is 0). If we substitute this into equation (2.6), we obtain

$$\delta\hat{\mathbf{q}} = (\mathbf{F}^T\mathbf{F} + \sigma_d^2\mathbf{C}_p^{-1})^{-1}\mathbf{F}^T\delta\mathbf{T}, \quad (2.15)$$

where the quantity within the brackets takes the place of the Hessian. As in equation (2.6), the quantity  $\mathbf{F}^T\mathbf{F}$  is difficult to invert by itself, and the addition of the term  $\sigma_d^2\mathbf{C}_p^{-1}$  serves to better condition the system of equations. Pratt and Chapman (1992) construct the model covariance matrix by making the following assumptions:

1. Each of the model parameters are uncorrelated and have an uncertain variance;
2. Variances of the anisotropic parameters are smaller than those representing isotropic velocity;
3. The parameter perturbations are spatially smooth.

Assumption (1.) is equivalent to setting  $\mathbf{C}_p = \sigma_p^2\mathbf{I}$ , which transforms equation (2.15) to  $\delta\hat{\mathbf{q}} = (\mathbf{F}^T\mathbf{F} + \lambda^2\mathbf{I})^{-1}\mathbf{F}^T\delta\mathbf{T}$ , where  $\lambda = \frac{\sigma_d}{\sigma_p}$ . Assumption (2.) implies that more weight will be given to the isotropic velocity than to the anisotropic components, and requires a careful ordering of the parameters. Here we now define  $\delta\hat{q}_0$  to represent the isotropic velocity, and  $\delta\hat{q}_j$ ,  $j = 1, \dots, 5$  to represent the purely anisotropic parameters. Augmenting  $\mathbf{C}_p$ , we add  $\mathbf{C}_p = \sigma_p^2 \frac{\lambda^2}{\lambda^2 + \kappa^2}$  for parameters  $\delta q_j$ ,  $j > 0$ . This results in the system

$$\delta\mathbf{q} = (\mathbf{F}^T\mathbf{F} + \lambda^2\mathbf{I} + \kappa^2\mathbf{I}')^{-1}\mathbf{F}^T\delta\mathbf{T}, \quad (2.16)$$

where  $\mathbf{I}'$  is a modified identity matrix where the elements corresponding to  $\delta q_0$  are equivalent to zero.

Let us interpret the roles of  $\lambda$  and  $\kappa$ :

- If  $\lambda$  is large, this implies that  $\sigma_d > \sigma_p$ . This means that the variance in the data is larger than the variance in the model parameters, and results in the final model being similar to the *a priori* model, at the expense of a better fit to the data.
- If  $\lambda$  is small, this implies that  $\sigma_d < \sigma_p$ . This means that the variance in the data is small (i.e. the first arrival picks are accurate), and the variance in the model parameters is large (i.e. can be updated to a position far from the *a priori* model).

The parameter  $\kappa$  takes the role of assigning a relative variance between the isotropic velocity and anisotropic parameters:

- If  $\kappa$  is small, the variance of the anisotropic parameters will be almost completely determined by the *a priori* model variance. This allows the introduction of anisotropy into the model without significant increases to  $E(\mathbf{m})$ .
- If  $\kappa$  is large, the variance of the anisotropic parameters will be much smaller than the *a priori* model variance, and perturbations to any anisotropic parameters will significantly affect the value of  $E(\mathbf{m})$ . This penalizes any introduction of anisotropy into the model.

The assumption (3.) that the parameter perturbations are spatially smooth is an interesting one, and will be very important to the body of this thesis. Its incorporation stems partly from the seemingly incompatible concepts of the coarse resolution of traveltime tomography, which is theoretically limited by the size of the Fresnel zone, due to wavefront healing, and the ray theoretical approximation, which approximates the passage of seismic energy as being limited to infinitely thin ray paths. What we are left with is a method which can only update the low wavenumber parameter features, but which itself is based on a high frequency (and wavenumber) forward model. Although this approximation is valid in the limit where the heterogeneities are much larger than the seismic wavelength (Červeny (2001)), it will break down in areas of strong, rapidly varying heterogeneity. Some consequences of this breakdown will be presented in Chapter 4.

To incorporate the assumption of smoothness, we penalize rough solutions by modifying equation (2.16):

$$\delta\mathbf{q} = (\mathbf{F}^T\mathbf{F} + \lambda^2\mathbf{I} + \kappa^2\mathbf{I}' + \epsilon^2\mathbf{R}^T\mathbf{R})^{-1}\mathbf{F}^T\delta\mathbf{T} \quad (2.17)$$

(Pratt and Chapman, 1992), where  $\epsilon$  gives weight to the entries in the roughness matrix  $\mathbf{R}$ . In this implementation, 3 separate roughness matrices are used:  $\mathbf{R}_x$ ,  $\mathbf{R}_z$ , and  $\mathbf{R}_{\nabla^2}$ .  $\mathbf{R}_x$  and  $\mathbf{R}_z$  take the form of finite difference stars, and represent the gradient of the model in the  $x$  and  $z$  directions respectively.  $\mathbf{R}_{\nabla^2}$  is a 2-D laplacian finite difference star, which represents the magnitude of the curvature of the model.

Finally, we can incorporate all the regularization terms into the system

$$\delta\hat{\mathbf{q}} = (\hat{\mathbf{F}}^T\hat{\mathbf{F}})^{-1}\hat{\mathbf{F}}^T\delta\hat{\mathbf{T}} \quad (2.18)$$

(Pratt et al., 1993), where  $\hat{\mathbf{F}} = (\mathbf{F} \ \epsilon\mathbf{R}_x \ \epsilon\mathbf{R}_z \ \epsilon\mathbf{R}_{\nabla^2} \ \kappa\mathbf{I}' \ \lambda\mathbf{I})^T$  and  $\delta\hat{\mathbf{T}} = (\delta\mathbf{T} \ 0 \ 0 \ 0 \ 0 \ 0)^T$ , which gives an expression for the parameter perturbations  $\delta\hat{\mathbf{q}}$ . In practice, however, the matrix inversion in (2.18) is almost never calculated, due to the computational expense

of inverting the large matrix ( $\hat{\mathbf{F}}^T \hat{\mathbf{F}}$ ). Instead, Pratt and Chapman (1992) use the LSQR solver (Lanczos, 1950; Paige and Saunders, 1982), which is similar in form to the gradient algorithm given in equation (2.7), in that it searches for a best fit solution by exploiting information from the local derivatives of the misfit function. Once the values of  $\delta \hat{\mathbf{q}}$  are found, it is possible to map them into the equivalent Thomsen’s parameters  $\varepsilon$  and  $\delta$  via a coordinate transform (Chapman and Pratt, 1992; Pratt and Chapman, 1992).

The degrees of freedom in the above regularization parameters scheme suggest that a proper strategy is needed to choose values for  $\lambda$ ,  $\kappa$ , and  $\epsilon$ . Pratt and Chapman (1992) suggested that the regularization parameters be chosen according to the appearance of a suite of tomograms, and relaxing the constraints until a model is generated which fits the data to within an acceptable margin of error while remaining geologically plausible. Deciding what an ‘acceptable margin of error’ is can be somewhat ambiguous, and Pratt and Chapman’s (1992) suggestion is to try and fit the data to within the estimated picking error. We shall see in the next section that if the goal of traveltime tomography is the generation of a starting model for waveform tomography, a much more quantitative and less subjective estimate of the acceptable margin of error is provided: That the error in the traveltime residuals be less than one-half of the period of the lowest frequency waves present in the survey data (Sirgue and Pratt, 2004). This is known as the ‘half-cycle criterion’.

## 2.2 Waveform Tomography

Waveform tomography is a form of high resolution seismic tomography (Tarantola, 1984; Pratt et al., 1998; Virieux and Operto, 2009). In waveform tomography, we are interested in inferring information on geological parameters by exploiting the ‘full wave equation’, rather than basing inversions on the restrictive ray-theoretical approximation. Before proceeding, I must clarify what is meant by the ‘full wave equation’. In reality, seismic waves propagating through the Earth exist in several modes (pressure, shear, surface, etc.), and dissipate energy as heat. To completely characterize seismic wave propagation, we then need a *visco-elastic* wave equation. If we attempt to simplify the picture by setting the shear modulus to zero, we are left with the *visco-acoustic* wave equation. Furthermore, while the true Earth is three dimensional (3-D), a common simplification in exploration seismic tomography is made by restricting wave propagation to two dimensional (2-D) slices. For the above reasons, these simplifications serve to void the use of the ‘full wave equation’ in waveform tomography. For this thesis, the wave equation that is adopted is the 1-D isotropic acoustic wave equation, with modifications to approximate

elliptical VTI anisotropy.

Any discretized wave equation in the frequency domain can be represented in matrix form as (Marfurt, 1984; Pratt et al., 1998)

$$\mathbf{S}(\mathbf{x}, \omega)\mathbf{u}(\mathbf{x}, \omega) = \mathbf{f}(\mathbf{x}, \omega), \quad (2.19)$$

where  $\mathbf{S}(\mathbf{x}, \omega)$  is a (complex-valued) impedance matrix,  $\mathbf{u}(\mathbf{x}, \omega)$  is the wavefield, and  $\mathbf{f}(\mathbf{x}, \omega)$  is the source term. Note that each term is dependent on both the spatial location  $\mathbf{x}$  ( $x_i \in \mathbb{R}^3$ ) and the frequency  $\omega$ . We are interested in estimating the wavefield  $\mathbf{u}$  at any point in the model. Accomplishing this would require us to invert the impedance matrix, giving

$$\mathbf{u} = \mathbf{S}^{-1}\mathbf{f}, \quad (2.20)$$

dropping the spatial and frequency dependence for sake of clarity. Comparing equation (2.20) to equation (2.1), we see that equation (2.20) can be viewed as a forward modelling equation, where  $\mathbf{S}^{-1}\mathbf{f}$  takes the place of  $g(\mathbf{m})$ , and  $\mathbf{u}$  (at the receiver positions) takes the place of  $\mathbf{d}$ . In practice,  $\mathbf{S}^{-1}$  is not calculated; instead its factorization into lower and upper triangular matrices ( $LU$  factorization) is calculated. The  $LU$  factors can be saved, making the solution of equation (2.20) for multiple sources very efficient (Pratt and Worthington, 1990).

To see how we might solve the inverse problem, let us examine the gradient of the misfit function more closely. We choose equation (2.4), and for simplicity we do not include *a priori* model constraints. As well, we set the data covariance matrix  $\mathbf{C}_d = \mathbf{I}$ . This gives:

$$E(\mathbf{m}) = \frac{1}{2} \times (g(\mathbf{m}) - \mathbf{d})^t (g(\mathbf{m}) - \mathbf{d})^*, \quad (2.21)$$

where the complex conjugate (\*) is used to ensure that the misfit function is real-valued for general, complex-valued data. Introducing the linear operator  $\mathbf{G}$  as the derivative of  $g$  at a point  $\mathbf{m}$  ( $\mathbf{G} = \frac{\partial g(\mathbf{m})}{\partial \mathbf{m}}$ )

$$g(\mathbf{m} + \delta\mathbf{m}) = g(\mathbf{m}) + \mathbf{G}\delta\mathbf{m} + (\|\delta\mathbf{m}\|^2), \quad (2.22)$$

we can calculate the gradient via a formal differentiation

$$\nabla E_m(\mathbf{m})\delta\mathbf{m} = \lim_{\nu \rightarrow 0} \frac{1}{\nu} \{E(\mathbf{m} + \nu\delta\mathbf{m}) - E(\mathbf{m})\}. \quad (2.23)$$



This gives the gradient in each direction  $\delta \mathbf{m}$  as

$$\nabla_m E(\mathbf{m}) = \Re \left\{ \mathbf{G}^t (g(\mathbf{m})^* - \mathbf{d}^*) \right\} = \Re \left\{ \left( \frac{\partial g(\mathbf{m})}{\partial \mathbf{m}} \right)^t (g(\mathbf{m}) - \mathbf{d})^* \right\}. \quad (2.24)$$

Let us attempt to physically interpret equation (2.24). Remember that  $g(\mathbf{m})$  represents our forward modeled wavefield  $\mathbf{u}_{est}$ , and that our data  $\mathbf{d}$  are the observed wavefields, or  $\mathbf{u}_{obs}$ . Substituting this into equation (2.24) gives

$$\nabla_m E(\mathbf{m}) = \Re \left\{ \left( \frac{\partial \mathbf{u}_{est}}{\partial \mathbf{m}} \right) (\mathbf{u}_{est} - \mathbf{u}_{obs})^* \right\}. \quad (2.25)$$

We now see that the gradient of the model parameters is simply the wavefield residuals  $(\mathbf{u}_{est} - \mathbf{u}_{obs})$  pre-multiplied by the partial derivatives of the estimated wavefield with respect to the model parameters  $\mathbf{m}$ , or the ‘partial derivative wavefields’ (Pratt et al., 1998).

Let us examine how we might calculate the gradient using equation (2.25). First, we must calculate  $\mathbf{u}_{est}$  by the application of forward modelling operator to the set of model parameters, which corresponds to one evaluation of  $g(\mathbf{m})$ , and is equivalent to the evaluation of equation (2.20). We then calculate the wavefield residuals  $\delta \mathbf{d}$  at the ( $n$ ) receiver positions, giving (Pratt et al., 1998)

$$\delta d_i = u_i - d_i, \quad i = (1, \dots, n), \quad (2.26)$$

where  $i$  is the receiver number. The wavefield residuals are then pre-multiplied by the partial derivative wavefields to obtain the gradient. Unfortunately, the calculation of the partial derivative wavefields in the form shown in equation (2.24) is prohibitively expensive. We must perform one forward model for a perturbation in each direction  $\delta \mathbf{m} \in \mathfrak{M}$ , where each time we evaluate  $g(\mathbf{m})$ . Fortunately, there is a much more efficient way to calculate the gradient.

To arrive at another expression for the partial derivative wavefield from equation (2.24), we can take the derivative with respect to the model parameters of equation (2.19). This gives

$$\frac{\partial \mathbf{u}_{est}}{\partial \mathbf{m}} = -\mathbf{S}^{-1} \left( \frac{\partial \mathbf{S}}{\partial \mathbf{m}} \mathbf{u}_{est} \right). \quad (2.27)$$

Here, we note that this equation is actually very similar to equation (2.20), with the

source term  $\mathbf{f}$  replaced with  $-\frac{\partial \mathbf{S}}{\partial \mathbf{m}} \mathbf{u}_{est}$ . In fact, we can then rewrite equation (2.27) as

$$\frac{\partial \mathbf{u}_{est}}{\partial \mathbf{m}} = -\mathbf{S}^{-1} \mathbf{F} \quad (2.28)$$

where  $F_i = -\frac{\partial \mathbf{S}}{\partial m_i} \mathbf{u}_{est}$ , and each  $F_i$  is known as a ‘virtual source’ (Pratt et al., 1998). Now, we can evaluate the required partial derivative wavefields by propagating each virtual source throughout the model. This has some efficiency gains over the direct implementation of equation (2.25), as  $\mathbf{S}^{-1}$  only needs to be evaluated once. However, this form still requires the propagation of  $m$  virtual sources, and is still much too computationally inefficient. Fortunately, we do not need to calculate the partial derivative wavefield explicitly, as we are interested only in the gradient of the misfit function. This means that what we are really interested in is the action that the partial derivative wavefield has on the wavefield residuals.

Substituting equation (2.28) into equation (2.25) gives

$$\nabla_m E(\mathbf{m}) = \Re \{ (\mathbf{S}^{-1} \mathbf{F})^t \delta \mathbf{d}^* \} = \Re \{ \mathbf{F}^t (\mathbf{S}^{-1})^t \delta \mathbf{d}^* \} \quad (2.29)$$

This is an important result. Now, instead of  $m + 1$  wavefield propagations (1 to initially calculate  $\mathbf{u}_{est}$ , and  $m$  to calculate the response of each virtual source), we now require only one, with the complex-conjugated wavefield residuals  $\delta \mathbf{d}$  taking the place of the source term.

The gradient of  $E(\mathbf{m})$  can now be interpreted as follows: The time-reversed data residuals are propagated through the model, by means of the transpose of the forward operator. Note that for reciprocal problems,  $(\mathbf{S}^{-1})^t = \mathbf{S}^{-1}$ . At each point in the model, the back-propagated residual wavefield is then multiplied by the corresponding virtual source, and the complete gradient has been obtained. Note here that we have saved a significant amount of computational effort: The full calculation of the gradient direction requires only two propagations, one to calculate  $\mathbf{u}_{est}$ , and one to calculate the residual wavefield. With this information, we are now closer to being able to use equation (2.7).

Next we need the steplength  $\alpha$ , and our initial model  $\mathbf{m}^0$ . To calculate the step length, we use the standard linear step length estimation

$$\alpha = \frac{\| \nabla_m E(\mathbf{m}) \|}{\| \mathbf{J} \nabla_m E(\mathbf{m}) \|} \quad (2.30)$$

(Pratt et al., 1998). To evaluate this expression, we perturb the model parameters by a small amount in the gradient direction (say, 1%), and calculate  $\mathbf{J}$  through a finite differ-

ence approximation. This requires the estimation of a new wavefield,  $\mathbf{u}^\alpha$ , bringing the total number of modelling steps to three. Other search techniques, such as a line search, or a golden section search, which remain true to the nonlinear nature of the forward problem, are applicable, although these methods require additional forward models, and often the linear estimation is sufficient (Pratt, 1999). If desired, at each iteration the estimated source,  $\mathbf{f}(\mathbf{x}, \omega)$ , can also be updated according to the equation

$$\mathbf{f}(\omega) = \frac{\mathbf{u}^* \mathbf{d}}{\mathbf{u}^* \mathbf{u}}, \quad (2.31)$$

where in this case  $\mathbf{u}$  is a delta-function source. This results in one additional modelling step (to estimate the wavefield in the new, updated model).

We now have all the ingredients of a gradient-based inversion, save for the starting model. In waveform tomography, it is common to obtain a starting model through travelttime tomography as described in section 2.1. Inserting a starting model into equation (2.7) as  $\mathbf{m}^0$ , we seek to iteratively improve  $\mathbf{m}^0$  until we have decreased the wavefield residuals by a sufficient amount, at which point we stop iterating. A quick look at what the residual wavefield actually represents, though, gives some pause. The complex-valued frequency domain acoustic wave equation has two defining properties at any given frequency: amplitude and phase. For this project we will ignore amplitudes, and only include the wavefield phase in our definition of the residuals due to the robustness of phase information in highly heterogeneous media. If we consider phase, we can quickly see that the behavior of the residuals is oscillatory. They will be at a maximum when the real and observed wavefield are  $\pm \frac{\pi}{2}$  out of phase, and will decrease monotonically as the phase error approaches either 0 or  $\pi$ .

When coupled with a gradient-based inversion scheme, these facts are disturbing. If the estimated wavefield through the initial model is out of phase with the observed wavefield by more than  $\frac{\pi}{2}$ , the only way to decrease the residuals will be to move the phase towards  $\pi$ . This would result in an erroneous final model, which has converged to the *wrong* cycle, as gradient-based optimizers will never produce model updates which *increase* the value of the misfit function. This pathological behavior is the reason we identified the half-cycle criterion in the previous section. If the initial model predicts the first arrivals waveforms to within a half cycle of their correct arrival time, then the gradient-based updates will perturb the model in the correct direction. Usually, we test the half cycle criterion by visually inspecting time-domain waveforms modeled using the starting model. If the real first break picks match the predicted waveforms to within one half cycle, we can proceed confidently with gradient-based waveform tomography. This

also exemplifies the importance of low frequencies in waveform tomography, as lower frequencies will have longer cycle times, which will allow the half cycle criterion to be more easily met (Sirgue and Pratt, 2004).

## 2.3 Simulated Annealing

In the previous two sections, I described the application of local inverse methods to tomographic problems. As discussed above, an alternative to this approach is the use of global inverse methods. Instead of trying to approximate the general inverse operator  $g^{-1}(\mathbf{d}) : \mathfrak{D} \mapsto \mathfrak{M}$ , global inverse methods rely on repeated applications of the the forward operator  $g(\mathbf{m}) : \mathfrak{M} \mapsto \mathfrak{D}$ , and attempt to build a statistical picture of how the model space maps into the misfit function. In this section, I give a general overview of simulated annealing, a type of semi-global inverse method. I use the term semi-global, as simulated annealing (SA) does not search the entire model space, but is designed to exploit some information on the structure of the misfit function.

The concepts behind SA are taken from statistical mechanics, specifically from the analysis of a solid cooling in a heat bath, which we refer to as the system. If the system is in thermal equilibrium at a temperature  $T$ , the probability that the system will be in a state  $i$  with energy  $E_i$  is given by the Gibbs, or Boltzmann probability distribution function (PDF)

$$P(E_i) = \frac{\exp(-E_i/KT)}{\sum_{j \in S} \exp(-E_j/KT)} = \frac{1}{Z(T)} \exp(-E_i/KT), \quad (2.32)$$

where  $S$  consists of all possible system configurations,  $K$  is Boltzmann's constant,  $T$  is the temperature, and  $Z(T)$  is the partition function

$$Z(T) = \sum_{j \in S} \exp\left(-\frac{E_j}{KT}\right). \quad (2.33)$$

The temperature is then reduced slowly, such that the system remains near equilibrium, and as  $T \rightarrow 0$ , the probability that the system is in the ground state (in a state  $i$  where  $E_i$  is as low as possible, defined as  $E_g$ ) approaches 100% (Sen and Stoffa, 1995). If the system is cooled quickly (known as quenching), then the material may freeze in a state where  $E_i \neq E_g$ , forming a glass.

Re-casting the above concept using the terminology of inverse theory we have developed in the previous sections, we associate  $E_i$  with the value of the misfit function  $E(\mathbf{m})$

at a point  $\mathbf{m} \in \mathfrak{M}$ .  $T$  is a control parameter with same dimensions as the misfit function, and the constant  $K$  is set equal to 1, yielding

$$P(E(\mathbf{m})) = \frac{1}{Z(T)} \exp(-E(\mathbf{m})/T) \quad (2.34)$$

(Sen and Stoffa, 1995). It can be seen that if we explore the model space randomly at a given temperature, after a suitable number of iterations, we will obtain a Gibbs PDF over  $\mathfrak{M}$ . However, equation (2.34) does not look like a suitable inverse method yet. In fact, it only suggests a random grid search, which we discussed earlier as being too computationally expensive for realistic problems. In addition, it relies on the computation of the partition function given in equation (2.33), which *itself* relies on the computation of  $E(\mathbf{m})$  for all  $\mathbf{m} \in \mathfrak{M}$ . If SA is formulated as in equation (2.34), we must *begin* by mapping the entirety of the misfit function, which would negate the need to use SA, or any other inverse method.

There are several ways to simulate the evolution of a heat bath without explicitly calculating the partition function (Sen and Stoffa, 1995). The one we adopt is the Metropolis algorithm (Metropolis et al., 1953). The algorithm proceeds as follows. First, a random perturbation is made to a model  $\mathbf{m}^{(k)}$ , resulting in  $\mathbf{m}^{(k+1)} \{ \mathbf{m}^{(k+1)} = \mathbf{m}^{(k)} + \Delta \mathbf{m}(\mathbf{m}^{(k)}), \mathbf{m}^{(k+1)} \in \mathfrak{M} \}$ . The misfit function  $E(\mathbf{m})$  is evaluated for  $\mathbf{m}^{(k)}$  and  $\mathbf{m}^{(k+1)}$ . Depending on the values of  $E(\mathbf{m}^{(k)})$  and  $E(\mathbf{m}^{(k+1)})$ ,  $\mathbf{m}^{(k+1)}$  is accepted with a probability  $\mathbb{A}(\mathbf{m}^{(k+1)}, T^{(k+1)})$  where

$$\mathbb{A}(\mathbf{m}^{(k+1)}, T^{(k+1)}) = \begin{cases} \exp\left(-\frac{E(\mathbf{m}^{(k+1)}) - E(\mathbf{m}^{(k)})}{T^{(k+1)}}\right) & \text{if } E(\mathbf{m}^{(k+1)}) > E(\mathbf{m}^{(k)}) \\ 1 & \text{if } E(\mathbf{m}^{(k+1)}) \leq E(\mathbf{m}^{(k)}) \end{cases} \quad (2.35)$$

This formula is known as the Metropolis criterion (Metropolis et al., 1953). We can write the probability  $\mathbb{P}(\mathbf{m}_j)$ , of generating any model  $\mathbf{m}_j \in \mathfrak{M}$  as

$$\begin{aligned} \mathbb{P}(\mathbf{m}_j) &= \mathbb{G}(\mathbf{m}_j) \times \mathbb{A}(\mathbf{m}_j, T^{(k)}) \\ \mathbb{G}(\mathbf{m}_j) &= \frac{1}{M}, \end{aligned} \quad (2.36)$$

where  $\mathbb{G}(\mathbf{m}_j)$  is the generation probability of  $\mathbf{m}_j$ , depending inversely on the number of elements  $M$  in  $\mathfrak{M}$ . Note that here we are assuming a finite size for  $\mathfrak{M}$ , and that the generation probability is uniform.

Let us interpret the Metropolis criterion in terms of inverse theory. Assume that we start from a randomly generated model  $\mathbf{m}^{(k)}$ , and subsequently select a random model

$$\mathbf{m}^{(k+1)}, E(\mathbf{m}^{(k+1)}) \leq E(\mathbf{m}_i^{(k)}).$$

Equation (2.35) tells us that this particular  $\mathbf{m}^{(k+1)}$  would be accepted every time. If this were the entire algorithm, we would essentially be using a local inverse method (albeit a very inefficient one), as only models which decreased the value of the misfit function would be accepted. Topographically, this is akin to only moving downhill in the misfit function. The semi-global characteristics of simulated annealing are brought into play when  $E(\mathbf{m}^{(k+1)}) > E(\mathbf{m}^{(k)})$ , and the Metropolis criterion is tested. There is now a finite probability that  $\mathbf{m}^{(k+1)}$  will be accepted, even though it has increased the total misfit, and has climbed ‘uphill’ in the misfit function.

Equation (2.35) shows that this probability is a function of the difference in misfit values between  $\mathbf{m}^{(k+1)}$  and  $\mathbf{m}^{(k)}$ , as well as the parameter  $T^{(k+1)}$ . For a large  $T$ , the probability that  $\mathbf{m}^{(k+1)}$  will be accepted ( $E(\mathbf{m}^{(k+1)}) > E(\mathbf{m}^{(k)})$ ) will be greater compared with the same  $\mathbf{m}^{(k+1)}$ ,  $\mathbf{m}^{(k)}$  pair at a smaller value of  $T$ . At  $T = 0$  the probability that this  $\mathbf{m}^{(k+1)}$  is accepted drops to 0. Again, we see that SA behaves like a local inverse method in the limit  $T \rightarrow 0$ .

To use simulated annealing with the Metropolis algorithm, we begin at iteration 0 with a high value for  $T^0$ , and slowly decrease  $T^k$  over the course of the inversion. While  $T^k$  is large, most model perturbations will be accepted. As  $T^k$  decreases, only perturbations which decrease  $E(\mathbf{m})$ , or which only increase  $E(\mathbf{m})$  by a small amount, are likely to be accepted. Eventually, once  $T^k$  drops to a low enough value, it will be overwhelmingly unlikely that any model perturbations that increase the misfit will be accepted, and we assume convergence.

Note the following limitations to simulated annealing: If any finite-time cooling schedule is used, there is still a chance that the the global minimum of  $E(\mathbf{m})$  will not be found. To overcome this, we should set the cooling schedule to be as slow as possible. However, we still require that the algorithm will finish in a reasonable amount of time. As a compromise, the cooling schedule is usually set so that SA will be complete after a certain amount of computer time, or after a certain number of iterations (Sen and Stoffa, 1995). By doing this, we trade a guarantee of convergence for the understanding that convergence is overwhelmingly likely. Unfortunately, setting cooling schedules in this manner is largely ad-hoc and problem dependent (Tarantola, 2005; Sen and Stoffa, 1995), although Nulton and Salamon (1988) and Mosegaard and Vestergaarad (1991) develop implementations which set the cooling schedule according to some statistical information gained during the model search. Also, we require an estimate of the initial temperature  $T_0$ . We must ensure that at this temperature the Metropolis criterion will be passed for a large variety of models; otherwise, the initial model will have an impact on the final result.

The choice of the initial temperature also is problem dependent, although I suggest a strategy in Chapter 3.

A comprehensive review of global and semi-global inverse methods within a geophysical context is given by Sambridge and Mosegaard, 2002.

# Chapter 3

## Results from a Real Crosshole Dataset

Here I describe the application of the methodologies introduced in Chapter 2 to a real, crosshole seismic dataset, which was provided to us by Vale. I will refer to this data set as the 540-542 dataset (Perozzi et al., 2012).

### 3.1 Geologic Background

The Voisey’s Bay deposit, located on the East Coast of Canada in Labrador, contains significant economic quantities of sulphide-hosted nickel, copper, and cobalt ores. Through an exploration program funded by Diamond Field Resources, the deposit was first discovered by Archean Resources in 1993. The property was purchased by Inco in August, 1996, and Inco was subsequently purchased by Vale in 2006, to form Vale-Inco. Recently, Vale-Inco has been re-branded as simply ‘Vale’, and as such Vale is the current owner of the Voisey’s Bay deposit (Evans-Lamswood et al., 2000).

The mineralized deposits are part of the Voisey’s Bay Intrusive Suite, which consists primarily of layered troctolite-gabbros intruding into Archean quartz-feldspar-biotite gneisses of the Nain province, and Proterozoic sulfidic garnetiferous paragneisses of the Churchill province (Ryan et al., 1995). Part of the Nain Plutonic Suite, the intrusion is dated between 1.34 and 1.29 Ga (Ryan, 1996). There are two magma chambers associated with the Voisey’s Bay deposit, known as the Eastern Deeps and Western Deeps chambers. As this project deals solely with the Eastern Deeps chamber, details of its geology will be expanded upon.

The Eastern Deeps magmatic system consists of a feeder tube (the Eastern Deeps Feeder), which leads into the large Easter Deeps Chamber. A geologic cross section of



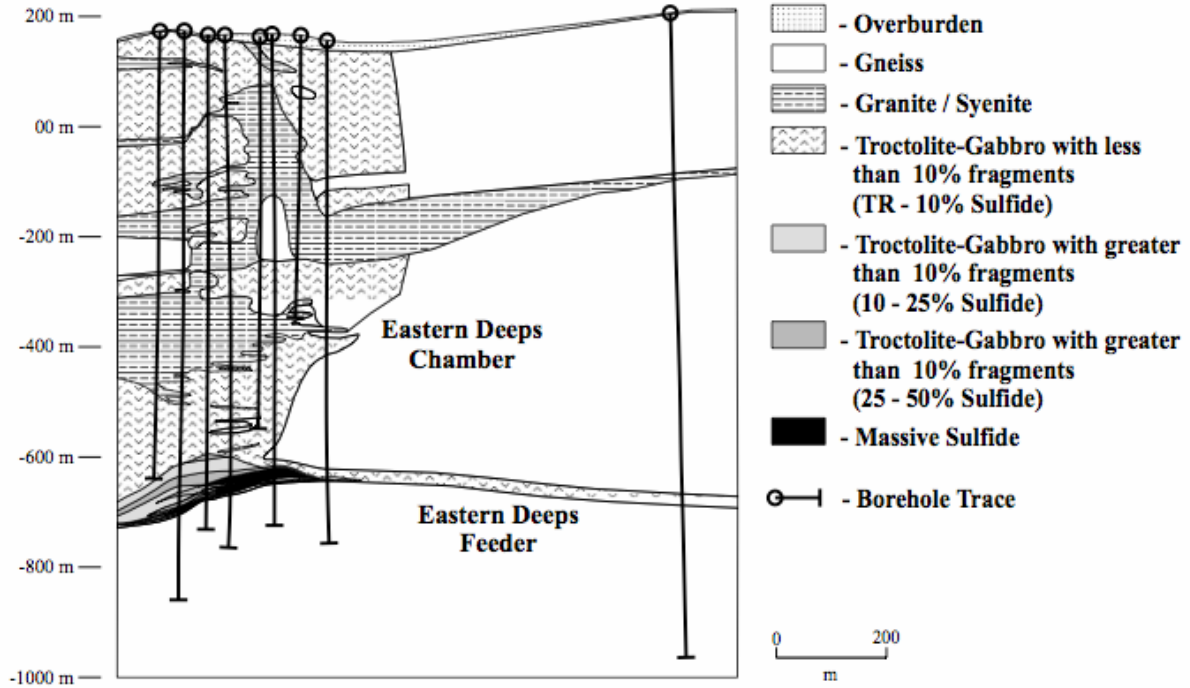


Figure 3.1: Geological cross-section of the Eastern Deeps deposit. From Evans-Lamswood, 2000.

the area is given in Figure 3.1. The main control on the location of mineralization is thought to be structural, which implies that deposition is controlled by morphological changes in the conduit geometry, and not by gravitational settling within the chamber itself. Evidence for this can be seen in Figure 3.1, where the majority of the massive sulphide mineralization occurs where the Eastern Deeps Feeder enters the Eastern Deeps Chamber. Outside this area of massive mineralization, the chamber is dominated by troctolite-gabbros, containing a variable amount of sulphide fragments, and granite/syenite. We were provided with data from the Eastern Deeps Chamber, in the form of seismic waveforms obtained from a crosshole seismic survey designed to map the continuity of the sulphide mineralization as it exited the Eastern Deeps Feeder.

A crosshole seismic survey refers to a seismic tomography experiment in which seismic sources and receivers are inserted into two, roughly parallel boreholes (the source and receiver holes, respectively). Seismic energy generated at the source hole through some mechanism (either impulsively, as with a hammer source, or over time, as with a swept-impact source), then propagates and is recorded in the receiver hole by an array (or ‘fan’) of receivers, which may consist of geophones or hydrophones. The purpose of such a survey is to use the information carried by the received waveforms, along with the geometry of the sources and receivers, to infer information on the geology between the

boreholes. This information is usually presented in the form of a tomographic velocity image of the between-hole velocity structure.

The specific boreholes in the crosshole survey of interest were drilled in 2000 by Boart Longyear (Inco, 2000). After drilling, the borehole geometries were investigated with the use of a north-seeking gyro, which recorded the azimuthal orientation and dip of the holes at 30 m intervals. As expected from the geological cross section, a large portion of the boreholes intersected troctolites containing variable amounts of sulphide, but also intersected an approximately 60 m thick zone of dipping mineralized massive sulphides at depth, corresponding to the location of the Eastern Deeps Feeder.

## 3.2 Survey Details

Since the main goal of the survey was to investigate the continuity of the massive sulphide zone, the survey geometry was designed accordingly. In the receiver hole, a string of 29 hydrophones, spaced 2 m apart, was used. Three receiver fans were used, each spaced 60 m apart. This resulted in a total of 87 unique receiver locations. Sources were spaced at 1 m intervals in the source hole, and 121 shots were recorded for each receiver layout. The combination of a static receiver layout and the corresponding shots is called a panel, and thus in total 3 panels were shot. As there was some overlap in the shot locations after a layout change, the total number of unique shot locations was 236. The shots thus extended for a total of 236 m, beginning approximately 100 m above the top of the massive sulphide zone, and ending approximately 50 m below its bottom. Due to the dip of the target zone, the first receiver was also approximately 100 m above the top of the massive sulphide zone. Unfortunately, due to the truncation of the receiver borehole, the final receiver was located only a few meters below the target zone. The hole-to-hole offset at the survey depth was approximately 30 m, although this narrowed to approximately 27 m near the bottom of the holes. Both the source and the hydrophones were manufactured by Vibrometric Ltd. A schematic of the survey geometry, overlaid on an estimated geological map (derived from Figure 3.1), can be seen in Figure 3.2.

The source was based on the Swept Impact Source Technique (SIST) (Park et al., 1996), as opposed to a traditional impact source. In SIST, a single shot consists of a series of relatively low amplitude pulses, generated by applying a voltage to a stack of piezoelectric crystals. Each pulse excites an identical wavelet, in which the frequency of excitation monotonically increases with time. After a shot is complete, the received waveform is decoded, and represented as a single pulse. This allows an acceptable signal-to-noise ratio to be achieved without using high energy impulsive sources (which can

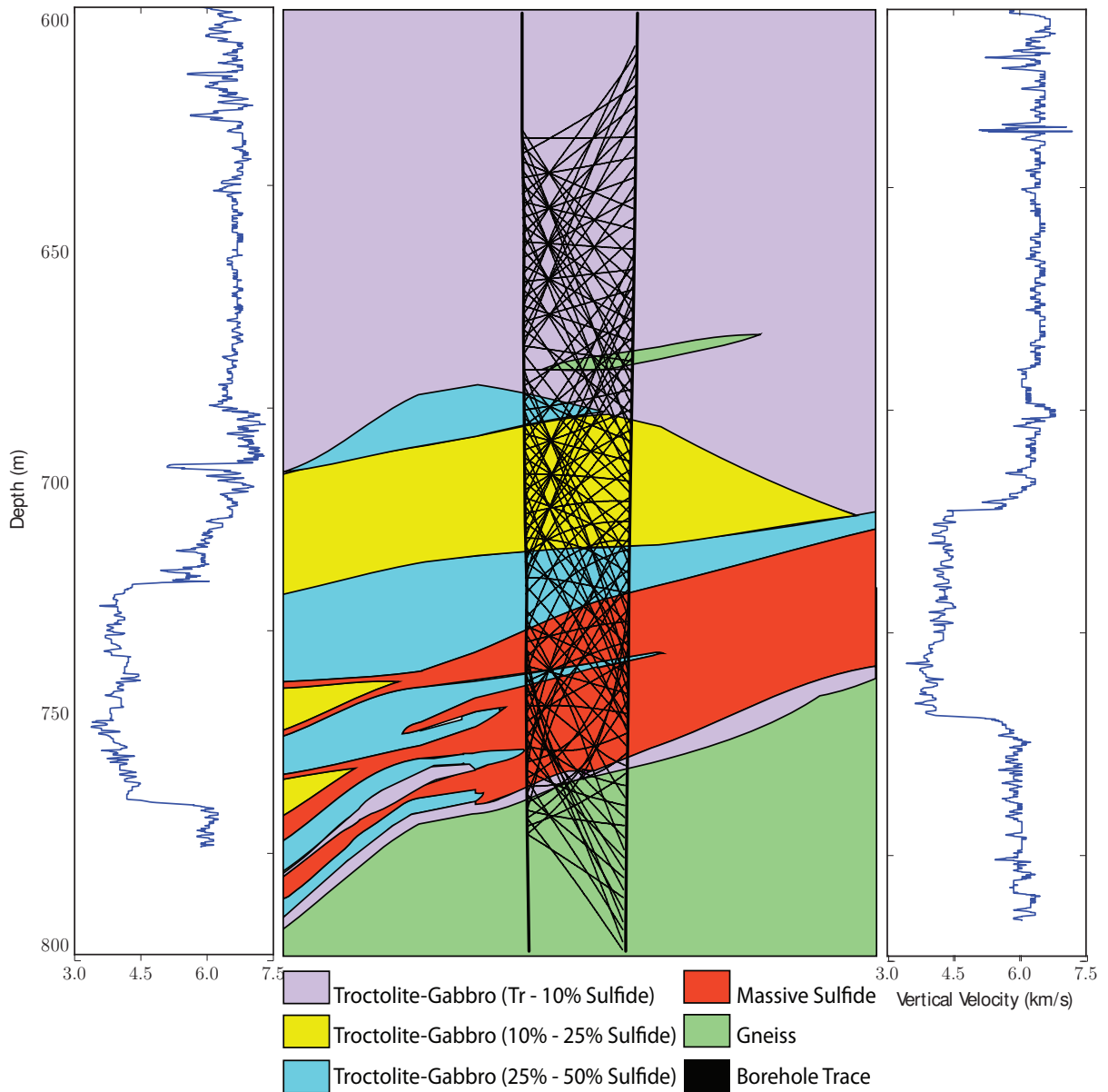


Figure 3.2: Hypothetical geological cross-section illustrating the expected geology at survey depths. The majority of the country rock is troctolite-gabbro, which surrounds a mineralized massive sulphide zone. Two boreholes, and an example of some seismic raypaths are overlain. Also present are the sonic velocity logs in the receiver (left) and source (right) holes.

damage, and sometimes collapse, the host boreholes).

To understand how the SIST works, consider that a conventional seismic trace, recording a high energy pulse source, can be written as

$$r_s(t) = s(t) * e(t) + n(t) \quad (3.1)$$

(Park et al., 1996), where  $s(t)$  represents the pulse source,  $e(t)$  is the impulse response of the Earth, and  $n(t)$  is ambient seismic noise. In SIST,  $s(t)$  is not a single pulse, and in fact we have

$$\Psi(t) = y(t) * s(t). \quad (3.2)$$

Now,  $\Psi(t)$  is the complete seismic source, which is the *convolution* of a series of impulses (or impacts) controlled by a binary coding sequence  $y(t)$ , and the low energy wavelet  $s(t)$ . Of course, in this form the received data  $r_s(t)$  will be very hard to interpret, as it will consist of a long record of very low energy seismic pulses. To decode the data, Park et al., 1996 wrote

$$\begin{aligned} r_d(t) &= y(t) \otimes r_c(t); \\ r_d(t) &= y(t) \otimes \{y(t) * s(t) * e(t) + n(t)\}; \\ r_d(t) &= \{y(t) \star y(t)\} * s(t) * e(t) + y(t) \otimes n(t), \end{aligned} \quad (3.3)$$

where  $\otimes$  represents cross-correlation, and  $\star$  represents auto-correlation. From the second equation, we can see that  $r_d(t)$ , the decoded seismic signal, can be found by cross-correlating the recorded coding sequence with the received trace. This is equivalent to auto-correlating the code sequence  $y(t)$ , convolving this sequence with the impulse earth response  $s(t)*e(t)$ , and adding the cross-correlation of the code sequence with the random seismic noise. This serves to boost the amplitude of the source signal, and, if the seismic noise is assumed to be random, increase the signal to noise ratio (as  $y(t) \otimes n(t) \rightarrow 0$  as  $t \rightarrow \infty$ , if  $n(t)$  is a zero-mean Gaussian distribution). Of course, the perfect application of equation (3.3) requires that the auto-correlation function is exactly the same as the effective source signature, something that is unlikely in practice. Nevertheless, Park et al. (1996) and Cosma and Enescu (2003) reported a very high signal to noise ratio with a relatively small number of impacts, compared to other coded impact techniques such as Mini-Soise (Barbier et al., 1976).

Examples of receiver gathers from the Voisey's Bay survey, showing the received waveforms from 121 unique source positions, are given in Figure 3.3. For these data, I picked the first arrival times (also known as first break times) by hand, using the

proprietary software package PROMAX (licensed from Landmark). Where present in Figure 3.3, first arrival times are highlighted in red. Figures 3.3a and 3.3b shows receivers located within the high velocity troctolite zone. Here, the waveforms were well behaved, and the data were relatively easy to pick. Figure 3.3c shows an example from within the low velocity zone. In this case, many of the sources were also located within the low velocity massive sulphide zone. The waveforms here were significantly more complicated than those in the troctolite zone, due mainly to the influence of the strong velocity contrast at the troctolite-sulphide boundary.

As well, in Figure 3.3a precursor component can be seen, visible as oscillations appearing prior to the first arrival picks. These precursors, which do not represent physical data, made picking even more difficult, and are believed to be the result of an acausal filter being applied to the data before we received them. Due to evidence presented in the next few paragraphs, the filter appears to have been a high-pass filter, which removed most of the data below 1000 Hz. As the processing software package provided by Vibrometric for use with their swept-impact sources includes bandpass filtering capabilities (Vibrometric, 2009b), it is likely that the filtering occurred at acquisition time. Figure 3.3d shows an example of a corrupted receiver gather. At time 0, coherent high-amplitude noise can be seen. So as not to corrupt any further processing, data from receivers exhibiting this signature was not included in any inversions.

Also evident in Figure 3.3 (especially in Figure 3.3c) are high amplitude arrivals which appear to have a constant, steep slope. These are caused by the presence of tube waves, which refer to waves propagating vertically within the water-filled boreholes (Pratt et al., 2004). In the source hole, when these waves encounter geological discontinuities in the borehole, such as those at the edges of the massive sulphide zone, their energy is converted to body waves, which eventually reach the receiver hole. In addition to vibrating the hydrophones, a portion of the transmitted energy will be transformed into tube waves in the receiver hole as well. These receiver hole tube waves are visible in Figure 3.3, traveling through the water column at approximately 1482 m/s. As our waveform tomography routine does not model tube waves, their presence in the data can lead to artifacts in the final image.

Since tube waves appear at a near constant slope, it is common to remove them from the data through frequency-wavenumber filtering ( $f$ - $k$  filtering). A plot of the  $f$ - $k$  spectrum for a typical receiver gather is given in Figure 3.4. Immediately obvious are high amplitude, steeply dipping (slow) signals. These correspond to tube waves. What is also obvious is the complete absence of signal below 1000 Hz, a sign that the data have been bandpass filtered. This is troublesome, because waveform tomography relies heavily

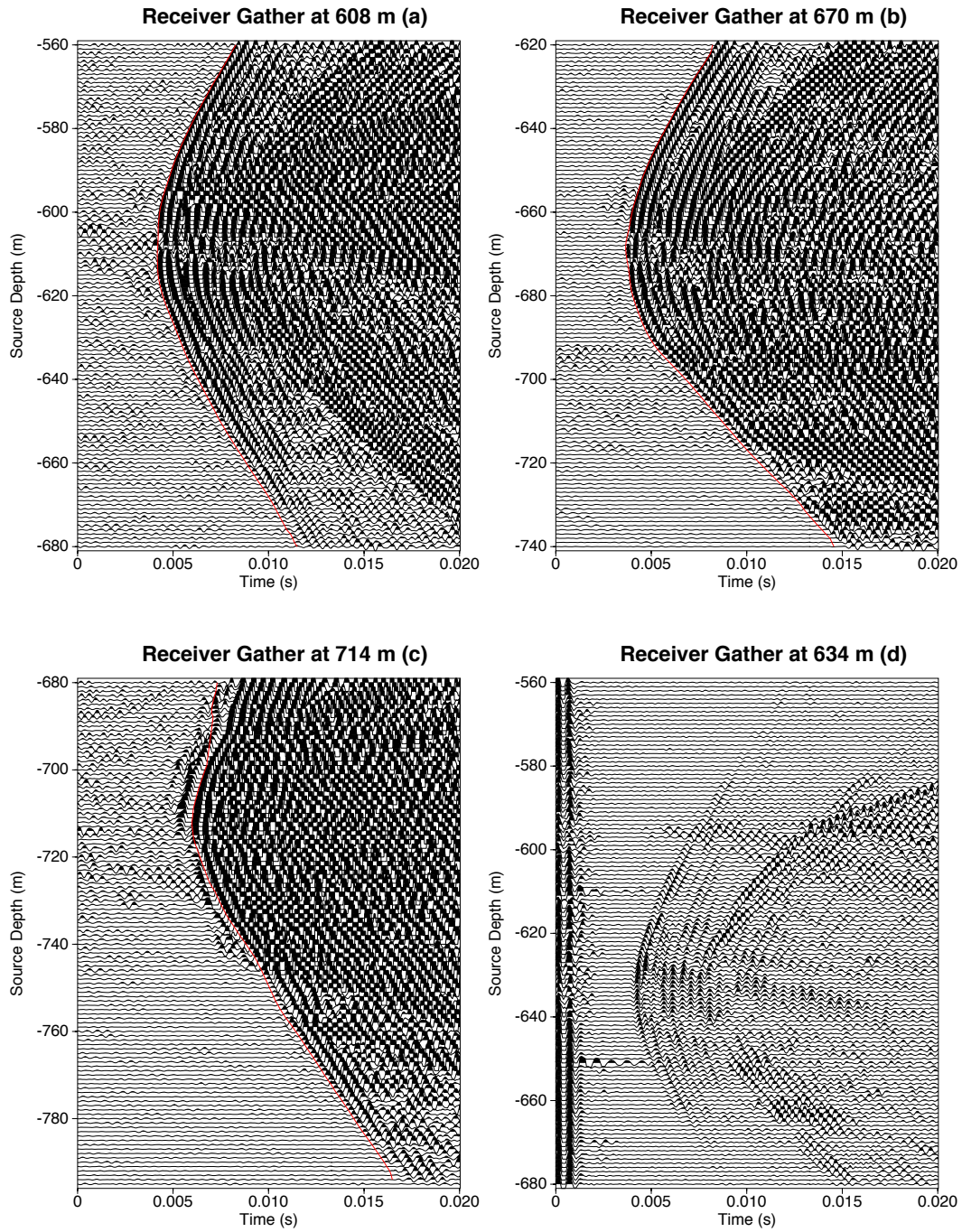


Figure 3.3: Selected raw receiver gathers; (a) and (b) are representative examples from the high velocity zone; (c) shows a receiver located within the low velocity massive sulphide zone; (d) shows an example of a corrupted receiver gather (which was excluded from the inversions).

on the existence of coherent, low frequency data, for reasons discussed in section 2.2. In addition, this means that the tube waves present in Figure 3.3 can only be removed at higher frequencies, above the Nyquist frequency, where they are aliased. Looking closely at Figure 3.3, we see that the first aliased mode of the tube waves has an apparent dip similar to some of the first arrivals. This echoes what we see in Figure 3.4, where the majority of the tube wave energy is mixed in with the bulk of the data.

Due to the band limited appearance of the data, and the aliasing of the tubes waves, they could not be removed cleanly from the data. We might therefore expect this to result in artifacts in our final waveform tomography images. Unfortunately, we also expect the lack of low frequency data to increase the chances of cycle skipping, as at 1000 Hz the starting model must predict the observed waveforms to within an accuracy of 0.5 milliseconds, a level of accuracy that is hard to achieve in practice.

### 3.3 Traveltime Tomography

In this section, the application of traveltime tomography to the 540-542 crosshole dataset is described. Previous examples of traveltime tomography in the Voisey’s Bay area can be found in Enescu et al. (2002), Cosma and Enescu (2003), and Perozzi et al. (2012).

The first step in preparing the data for traveltime tomography was the assignment of the survey geometry. Since both our traveltime and waveform tomography routines are 2-D, the inherently 3-D survey geometry was projected onto a best-fit 2-D plane. First, the measurements from the north seeking gyro were transformed into 3-D  $(x, y, z)$  coordinates, resulting in a cloud of points. A plane was then fit to the coordinates, based on the  $L_2$  norm of the distance residuals. The 3-D coordinates were then projected onto the 2-D null space of a vector normal to the plane, resulting in a set of 2-D coordinates  $(x, z)$ .

First break picking was then performed manually on the data. This was a very labour intensive process, and one that was repeated many times. Due to the small source-receiver distance, picking errors were very significant, and great care had to be taken. As mentioned above, the data had been pre-filtered, resulting in an apparent precursor in the waveforms, requiring repeated efforts to pick the data consistently. In Chapter 4, synthetic data, which were generated in a model similar to the real survey, are presented. In some areas, such as those surrounding the low velocity zone, the behavior of the synthetic waveforms were invaluable in decision making during the picking of real data.

Our traveltime tomography approach is based on the work of Chapman and Pratt,



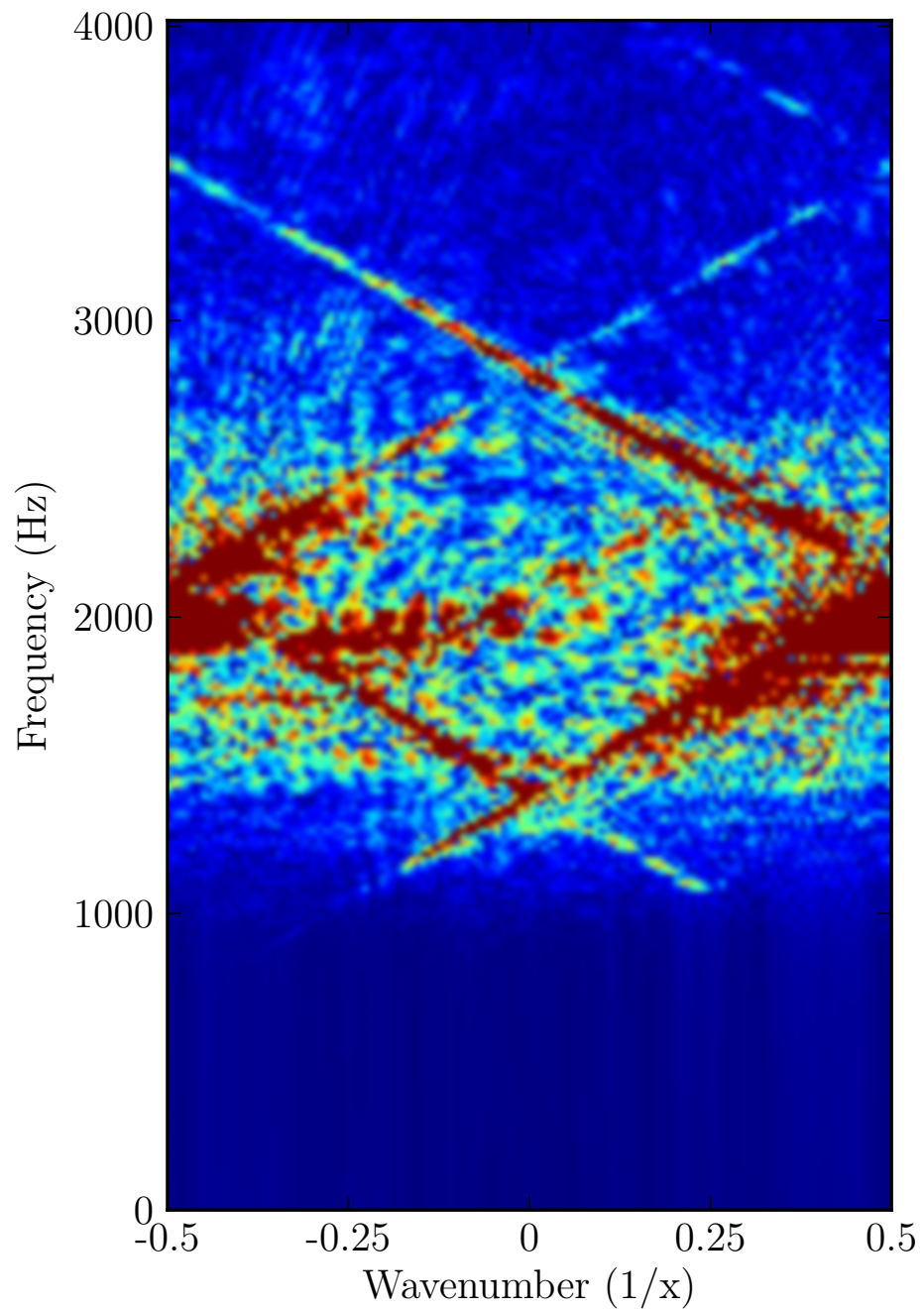


Figure 3.4: Frequency-wavenumber spectrum of a typical receiver gather.



Iteration	Constrained Model				Unconstrained Model			
	$\lambda$	$\kappa$	$\epsilon$	$E(\mathbf{m})$ (ms)	$\lambda$	$\kappa$	$\epsilon$	$E(\mathbf{m})$ (ms)
1	0.1	0.05	0.5	0.447	0.1	0.00	0.5	0.400
2	0.05	0.05	0.1	0.140	0.05	0.00	0.1	0.112
3	0.01	0.05	0.1	0.103	0.01	0.00	0.1	0.069
4	0.005	0.05	0.05	0.080	0.005	0.00	0.05	0.050
5	0.001	0.05	0.05	0.078	0.001	0.00	0.05	0.049

Table 3.1: Parameters for traveltime tomography.

(1992), and Pratt and Chapman, (1992); some of the theory was presented in Chapter 2. The initial results from TTI traveltime tomography are given in Figure 3.5. Two sets of models are shown, with a velocity model (Figure 3.5(a, c)) and a corresponding  $\epsilon$  model (Figure 3.5(b, d)) exhibited for each set. These results were obtained using five outer (nonlinear) iterations, between each of which new raypaths were calculated by the method of Um and Thurber (1987), and the system was re-linearized. The number of linearized iterations was variable; they were stopped when the rate of change of the misfit function dropped below a predefined value, dependent on the constraints used. If perturbations to the model were heavily constrained, the reduction of the misfit function plateaued quickly.

The major difference between these two sets of models is the magnitude of the constraints on the perturbations to the anisotropic parameters (the magnitude of  $\kappa$  in equation (2.18)). The models in Figure 3.5a and Figure 3.5b apply a positive  $\kappa$ , constraining  $\epsilon$  to be close to zero, while those in Figure 3.5c and 3.5d apply no such constraint. I refer to these pairs of models as the constrained and unconstrained inversions respectively. The values of  $\lambda$ ,  $\kappa$ , and  $\epsilon$  are given in Table 3.1.

Comparing the  $\epsilon$  models in the constrained and unconstrained inversions in the vicinity of the low velocity zone (between approximately 690 m and 760 m in depth) we see a curious effect. If anisotropy is constrained, this zone appears to be (almost) filled with a uniform  $\epsilon$ . However, if anisotropy is left unconstrained, the anisotropy in this zone resolves into several lobes around the edges of the zone, suggesting the presence of artifacts (in section 4.4 I provide a discussion and a possible explanation). Due to the highly stressed and fractured nature of the hard rock environment, a 2-D TTI symmetry system was assumed. As our waveform tomography routine currently assumes that the anisotropy possesses VTI symmetry, and is elliptic ( $\epsilon = \delta$ ), the information gained from images of  $\delta$  was not included in further processing.

The values of the anisotropic coefficients in the troctolites (Figure 3.5,  $\epsilon = 4 - 5\%$ ) coincides well with previous studies. In Voisey’s Bay, there have been reports by

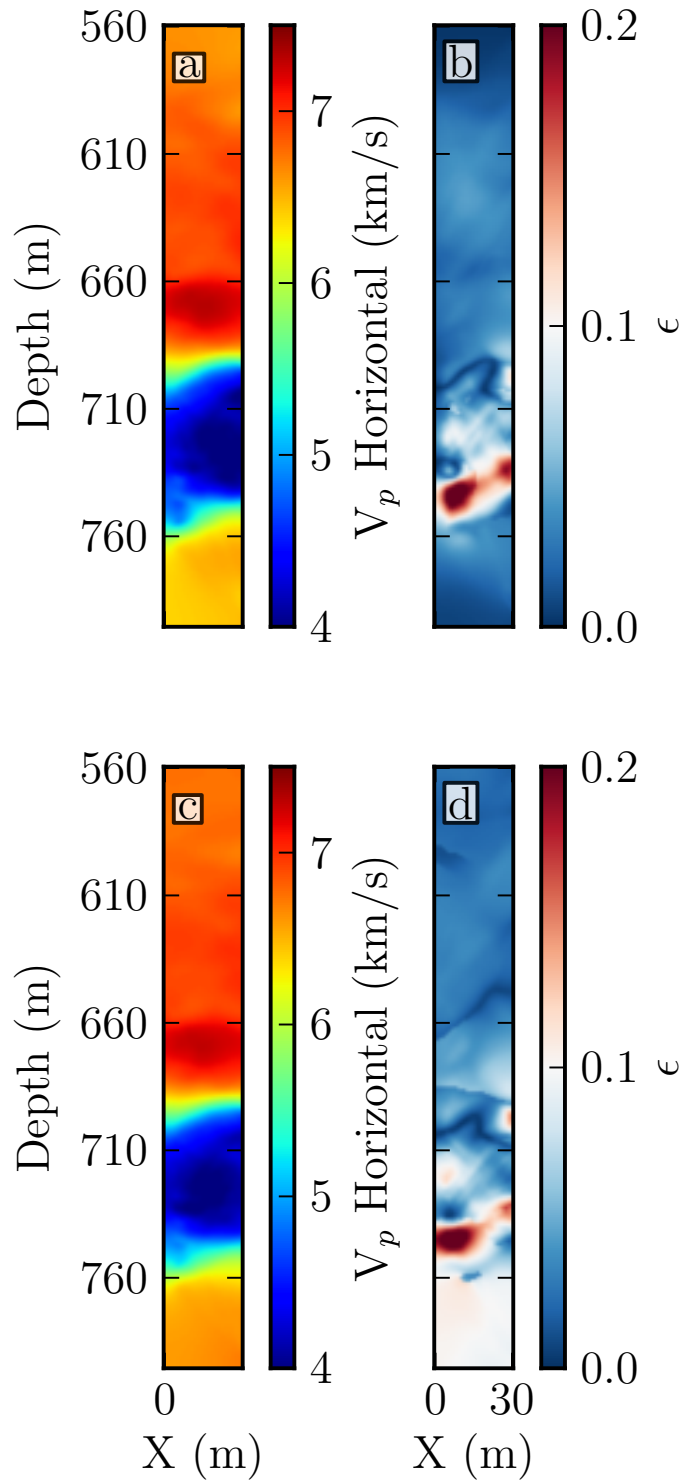


Figure 3.5: Traveltime velocity and anisotropy models. (a) Velocity, with a strong constraint on anisotropy; (b)  $\epsilon$ , with a strong constraint on anisotropy; (c) Velocity, with a weak constraint on anisotropy; (d)  $\epsilon$ , with a weak constraint on anisotropy.

Vibrometric (Cosma and Enescu, 2003) that place the anisotropy at approximately 3%. Furthermore, laboratory measurements have reported on troctolite anisotropy being as high as 12% (Babuska, 1968), with results from Iturrino et al. (1991) reporting values as high as 8%. While these laboratory measurements are very accurate, their relevance to true, in-situ anisotropy parameters are not known, as in-situ pressures may close any fracture planes which are present at atmospheric pressure (Tsvankin, 2012). Our estimated value of  $\varepsilon$  at 4-5% in the troctolites is therefore not surprising, and agrees well with previous results.

We should also note that although we assume a transversely isotropic model (with an arbitrary axis of symmetry) in traveltime tomography, this is likely a simplification of reality. The borehole core logs (Inco, 2000) report on highly fractured rock, with multiple fracture symmetry axes, overprinted on local flow and shear fabrics. As well, post emplacement serpentine metamorphism, known to be highly anisotropic (Kern, 1993), along some cracks may add additional symmetry. To completely characterize the anisotropy would likely require orthorhombic, monoclinic, or even triclinic symmetry systems. These systems require the inversion of 9 elastic parameters for the orthorhombic model, 13 for monoclinic, and the full 21 elastic parameters for triclinic media. The capability of inverting a system of equations which is so underdetermined is currently the subject of leading-edge research (Grechka and Mateeva, 2007), and is beyond the scope of this thesis. Note, however, that these complex symmetry systems may eventually yield better velocity inversions in the Voisey's Bay area.

To prepare the model for waveform tomography, we created a 1-D, vertically varying anisotropy profile by averaging the epsilon model in Figure 2.18d over horizontal layers. The resultant profile can be seen in Figure 3.6. Synthetic waveforms were generated through this model, to determine if the half-cycle criterion was satisfied. The resulting synthetic receiver gathers from the traveltime model (Figure 3.11b) show a mismatch in the first arrival times at higher offsets. Specifically, in Figure 3.11b, we see this for shots deeper than 720 m. At close offsets, the first arrivals were consistently predicted to within one half wavelength of the dominant frequency. The pattern of a good fit at near offsets, and of cycle skipping at higher offsets, was observed everywhere within or near the low velocity zone. Since an increase in offset corresponds to an increase in angle away from the horizontal, it was suspected that the anisotropy model found through traveltime tomography was insufficient for use in waveform tomography.

A brief aside: In an attempt to manually overcome the cycle skip problem, the picking was repeated many, many times. Early on, it was not known that the data had been filtered (this fact came to light after a meeting with Calin Cosma, owner of Vibrometric,

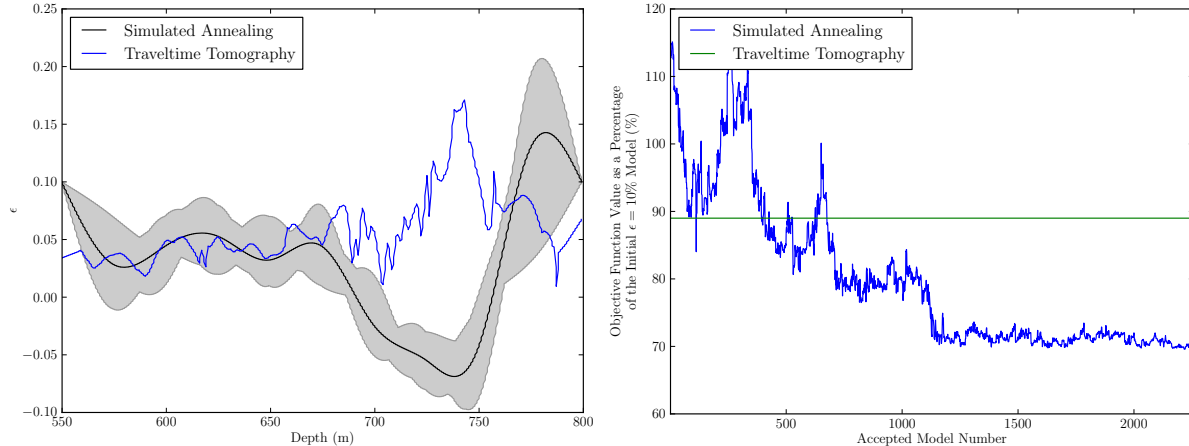


Figure 3.6: Left: Epsilon profiles for use with waveform tomography. Blue: profile obtained by averaging the  $\epsilon$  profile from traveltime tomography; black: ideal profile obtained through simulated annealing. The grey area represents the range of epsilon models, found through simulated annealing, which lie within one standard deviation of the absolute best fit value. Right: Objective function behavior during simulated annealing, as a percentage of the seed model.

in Copenhagen during July, 2012), and it was ambiguous as to whether the precursor component was signal or artifact. Many different interpretations of the first arrival times were followed, in an effort to produce a reasonable-looking anisotropy model which fit the data, but to no avail.

The philosophy behind the traveltime inversion strategy I eventually developed was as follows: Since an anisotropy model which fit the data to the degree required for waveform tomography could not be found through manual picking and traveltime tomography, a novel method was developed to discover a ‘best fit’ 1-D anisotropy model. The actual values of  $\epsilon$  and  $\delta$  imaged through traveltime tomography were ignored, and attention was given to reducing the traveltime residuals while maintaining a geologically reasonable horizontal velocity model. As such, I did not penalize perturbations to the anisotropy models. Instead, I discarded the traveltime anisotropy models, and developed an approach based on simulated annealing, to determine an  $\epsilon$  model by examining the waveform data (see section 3.4.1 below).

### 3.4 Waveform Tomography

An extremely effective approach to acoustic waveform tomography may be formulated in the Laplace-Fourier domain (Shin and Cha, 2009; Kamei et al., 2012). Each trace is transformed into the Laplace-Fourier domain via a (discrete version) of the complex-

valued Laplace Transform:

$$\hat{F}(s) = \int_{-\infty}^{\infty} f(t)e^{-st} dt, \quad (3.4)$$

where  $f(t)$  is the (discretely sampled) time-domain trace, and  $\hat{F}(s)$  is the discrete Laplace domain representation of  $f(t)$  at a single complex-valued Laplace parameter  $s$ . The parameter  $s$  has the form

$$s = \sigma + i\omega, \quad (3.5)$$

so equation (3.4) can be expanded to give

$$\hat{F}(\sigma, \omega) = \int_{-\infty}^{\infty} f(t)e^{-\sigma t} e^{-i\omega t} dt. \quad (3.6)$$

The parameter  $\sigma$  represents a characteristic decay time

$$\tau = \frac{1}{\sigma}, \quad (3.7)$$

so what equation (3.6) *really* represents is the time-damped Fourier transform of each trace

$$\hat{F}(\tau, \omega) = \int_{-\infty}^{\infty} f(t)e^{-(t/\tau)} e^{-i\omega t} dt. \quad (3.8)$$

The purpose of using the Laplace-Fourier domain, instead of just the Fourier domain, is to use the Laplace transform to preferentially weight data corresponding to early arrivals. Although the data used for waveform tomography were time-damped with  $\tau = 0.01$ , the wavefield images presented here were created by setting  $\tau = \infty$ . This results in equation (3.6) reducing to the classic Fourier transform. To simulate anisotropy, a geometrical stretching approach along the vertical is used to mimic the elliptic case (where  $\varepsilon = \delta$ ) (Dellinger, 1991).

If the raw traces are transformed into the Laplace-Fourier domain without any pre-processing, the resulting data appear as illustrated in Figure 3.7, where examples of 500 Hz, 800 Hz, 1300 Hz, and 1600 Hz frequency domain data files are presented. These images represent the Fourier transformed waveforms from each trace. In these images, sources increase from 1 to 236 along the x-axis, and receivers increase from 1 to 87 along the y-axis. Three distinct blocks are visible, and each block corresponds to one survey panel. Note that some sources are present in more than one panel, illustrating that source positions were repeated. Due to the geometry of the survey, the source-receiver pairs with the closest offsets appear approximately along the main diagonal in the images.

These data are unusable in their current form, as we expect to see some form of

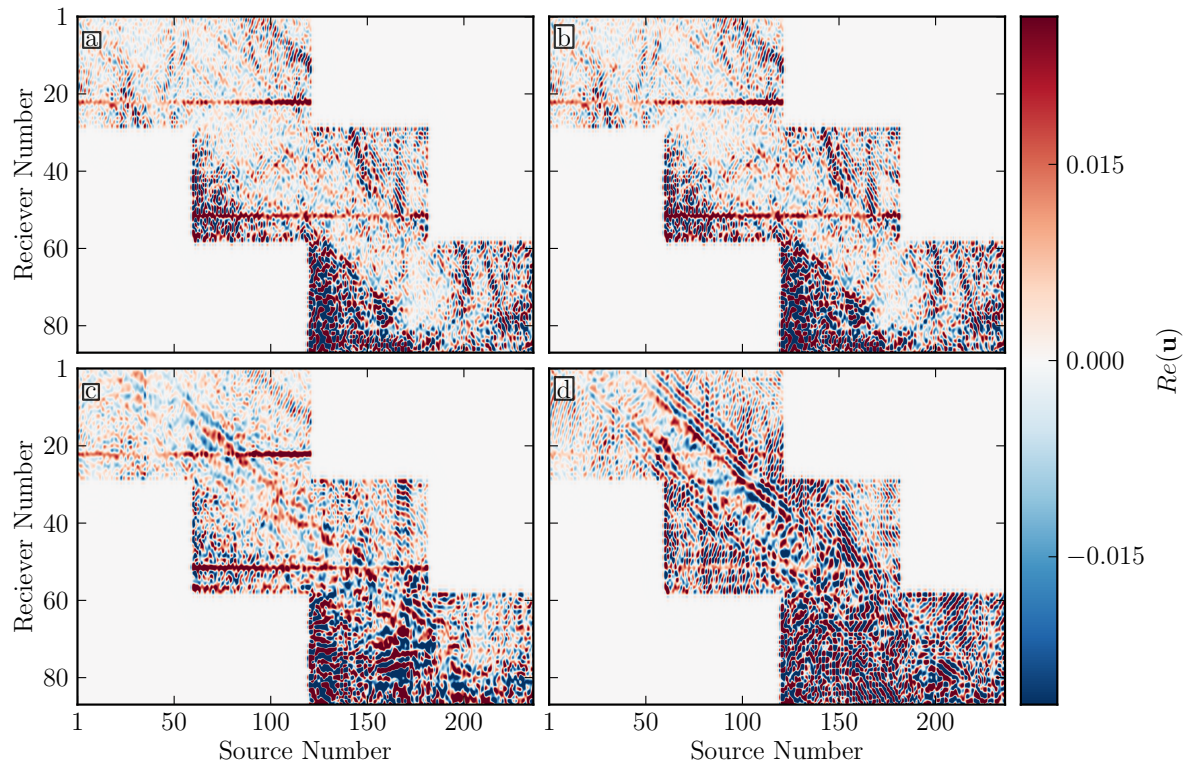


Figure 3.7: Real component of the complex-valued frequency domain wavefields ( $Re(\mathbf{u})$ ), before bandpass filtering and time windowing. (a) 500 Hz; (b) 800 Hz; (c) 1300 Hz; (d) 1600 Hz.

coherent pattern. In the 500 Hz and 800 Hz datafiles, no semblance of coherency can be extracted at all. This echoes our previous observation that the data were bandpass filtered to above 1000 Hz, and below this value we have just random noise. At 1300 Hz and 1600 Hz, a pattern does start to emerge: that of repeating stripes parallel to the main diagonal, at close offsets. However, as we move away from the main diagonal, the pattern disappears. To properly use these data in waveform tomography, the waveforms must be pre-processed.

In our case, pre-processing involved bandpass filtering the data to less than 2000 Hz (to remove any high frequency data), and windowing the waveforms in the time domain. This windowing served to mute any noise present before the first arrivals, and also to exclude (as much as possible) late-arriving shear waves (which are not modeled by the acoustic wave equation). We built our time window around the first break picks, choosing to begin the window 1.5 ms before, and end the window 3.5 ms after, the picked time.

Images of selected frequency domain data files for several frequencies, after bandpass filtering and time windowing, are given in Figure 3.8. As explained in section 3.2, the data appear to have been filtered and bandpassed above 1000 Hz before they were passed along to us. While our own time-windowing of the data introduced a low frequency component, it is obvious from Figure 3.8 that most of these data are corrupted and noisy. While the waveforms at 500 Hz seem coherent (Figure 3.8a), those between 600 Hz and 1000 Hz are very incoherent (a representative example at 800 Hz is given in Figure 3.8b). Above 1000 Hz, the data are of good quality and coherent. Representative examples at 1300 Hz and 1600 Hz can be seen in Figures 3.8c and 3.8d respectively. As stated in section 3.2, the inclusion of low frequency data is essential to the success of waveform tomography, and determines the required accuracy of the starting model.

In addition to data quality control, other relevant information can be visually extracted from Figure 3.8. For those sources and receivers in the low velocity zone (sources 130 - 180, receivers 58 - 87), we see that the apparent dip of the data changes. Outside of this zone, a cyclical pattern is seen, parallel to the main diagonal, while inside this zone the cyclical pattern appears perpendicular to the main diagonal. This is a sign that the data in this region are aliased, which strongly suggests a corresponding decrease in seismic velocity (shortening the wavelengths to below the Nyquist criterion). Discontinuities parallel to the receiver axis are also apparent at several locations. As the x-axis represents source location, these discontinuities may be related to source-borehole coupling issues, as well as source misallocation issues. Finally, discontinuities parallel to the source axis are also present. These represent deleted receivers, which were excluded due to excessive noise (see Figure 3.3d for an example). Specifically, the 23rd receiver in the

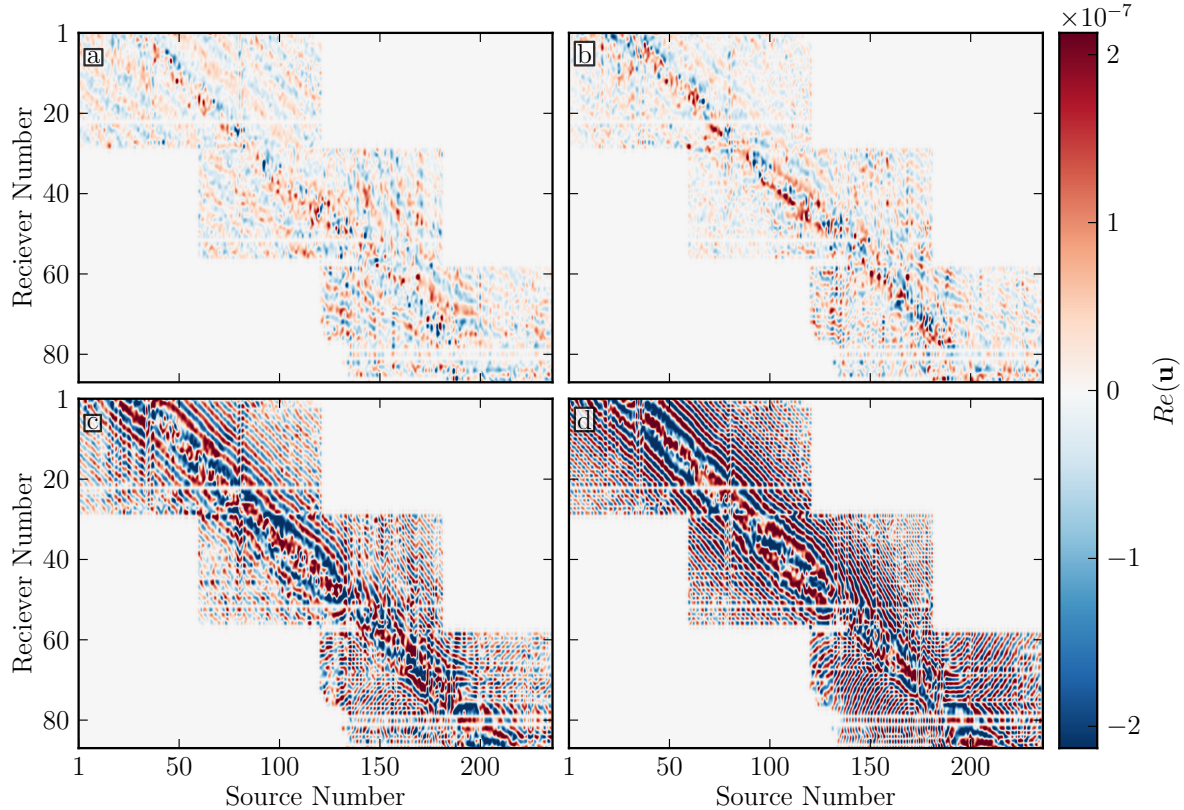


Figure 3.8: Real, frequency domain wavefields ( $Re(\mathbf{u})$ ), after bandpass filtering and time windowing. (a) 500 Hz; (b) 800 Hz; (c) 1300 Hz; (d) 1600 Hz. Note that here the amplitudes have been normalized between source and receiver gathers.

chain was consistently corrupted, and is absent from each panel.

From equation (2.19), we can see that the wavefield  $\mathbf{u}$  is linear with respect to the source  $f(\mathbf{x}, \omega)$ . Assuming the velocity model is exact, inverting for the source is then a linear inverse problem. The true, complex valued source can be calculated :

$$f(\mathbf{x}, \omega) = \frac{\mathbf{u}^t \mathbf{d}^*}{\mathbf{u}^t \mathbf{u}^*}. \quad (3.9)$$

(Pratt et al., 1998). Equation (3.9) is exact, and requires only one forward modelling run (to estimate  $\mathbf{u}$ ). Equation (3.9) can also be used as a convenient way to measure the quality of the velocity model: If the source signature is identical for all shots, we should expect to recover the same source signature for all shots (again, if the velocity model is correct). In fact, the consistency of the estimated source signature is a good metric by which to judge the quality of the inversions (Pratt and Symes, 2002).

An image of the estimated sources for all 236 source locations, using the traveltimes velocity and anisotropy models, is given in Figure 3.9a. We see that between source 1



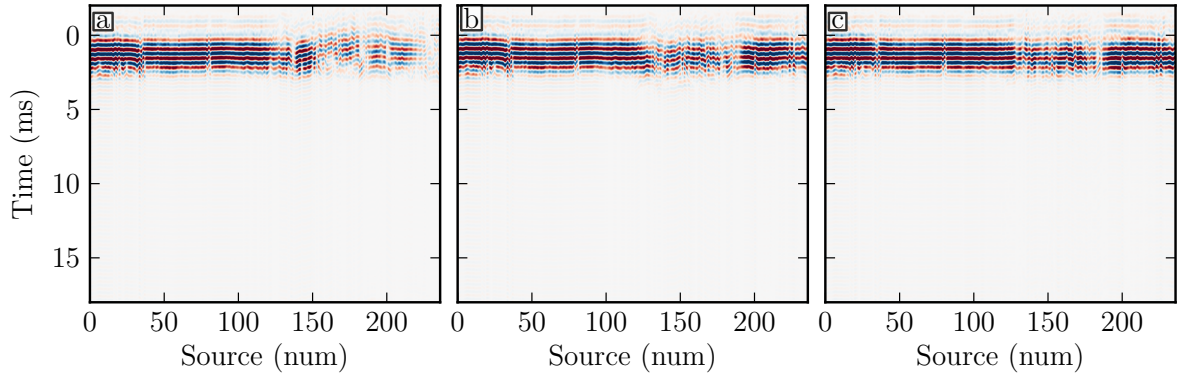


Figure 3.9: Estimated source signatures: (a) using the travelttime velocity and anisotropy models; (b) using the travelttime velocity model, and the anisotropy model found from simulated annealing; (c) using the waveform tomography velocity model, and the anisotropy model found from simulated annealing.

and 130, the signature is relatively consistent. This suggests that the travelttime velocity model is reliable in the high velocity troctolites. At source numbers above 130, where the sources are located within the low velocity zone, the signature is much less coherent. We attribute this incoherency to errors in the travelttime velocity and anisotropy models. In order to extract a single, useful source estimate, I isolated sources 1-130, and averaged them. This average source signature was used as the initial source for waveform tomography.

### 3.4.1 Anisotropy

To overcome the cycle skipping problem described above, and to allow for the proper convergence of waveform tomography, a novel approach was developed to determine the best-fit 1-D VTI elliptical anisotropy profile. The approach sought to minimize the logarithm of the phase residuals (Shin and Min, 2006; Kamei et al., 2012) using the semi-global method of simulated annealing. The setup was as follows:

- The model space  $\mathfrak{M}$  was defined as the space of all  $\varepsilon(z)$  values between  $-0.1$  and  $1.0$ ;
- $\mathfrak{M}$  was discretized in intervals of  $\Delta\varepsilon = 0.01$ , and assumed to be smoothly varying. The  $\varepsilon(z)$  profile was interpolated via cubic spline interpolation between tie points spaced every 25 m;
- The seed model consisted of a constant  $\varepsilon$  value of  $0.1$ .

At each iteration, a random model parameter, defined as a cubic spline tie point, was perturbed to a random value in model space. Synthetic waveforms at 500 Hz, 1000 Hz, 1300 Hz, and 1600 Hz were then calculated. The  $L_2$  norm of the phase residuals was evaluated, and this value was compared to the  $L_2$  norm of the phase residuals in the previously accepted model. If the residuals in the current model were less than those in the previous model, the new model was automatically accepted. If the residuals were greater than those in the previous model, the probability of this new model being accepted ( $P_{accept}$ ) was then calculated from the Metropolis criterion (equation (2.35)). A random number  $R$  was then drawn on the interval  $[0, 1]$ . If  $P_{accept}$  was greater than  $R$ , the Metropolis criterion was satisfied, and the new model was accepted. If  $P_{accept} < R$  the new model is discarded. A flowchart illustrating this procedure is given in Figure 3.10.

The choice of a starting temperature in simulated annealing has traditionally been problem dependent (Sen and Stoffa, 1995). A useful suggestion is to choose a temperature that allows 30% - 50% of randomly proposed models to be accepted (Tarantola, 2005). To determine our starting temperature  $T_0$ , a ‘warm up’ phase was added to the simulated annealing algorithm. Throughout this phase, all perturbations to the model were accepted. Assuming that a random perturbation would be accepted 50% of the time (i.e.  $\mathbb{A}(\mathbf{m}^{(k+1)}, T^{(k+1)})$  from equation (2.35) was set to 0.5), the Metropolis criterion was used to determine the temperature of an equivalent system. This procedure was time-invariant, i.e. the absolute values of  $\Delta E(\mathbf{m})$  were used. After 100 iterations, the average equivalent temperature of all these perturbations was calculated, and this was set as  $T_0$ .

This warm up phase also ensured that the starting model used for the subsequent simulated annealing, varied randomly from run to run. The cooling schedule was set to reduce the temperature to 95% of its current value after 50 iterations, determined by trial and error. Longer warm up periods and cooling schedules were tested, although they were found to have a negligible effect on the starting temperature. Longer cooling schedules were also tested, although this resulted in no significant change to the final models, and vastly increased the time needed for convergence. Shorter periods resulted in the inversions sometimes ‘freezing’ at local minima, and converging to unrealistic results.

The results of using simulated annealing to find the best fit 1-D elliptical anisotropy profile for the Voisey’s Bay dataset is shown in Figure 3.6. The velocity model used is given in Figure 3.5c. In Figure 3.6a, the blue line represents the anisotropy profile obtained by averaging the  $\varepsilon$  model shown in Figure 3.5d over horizontal layers. The black line is the ideal 1-D elliptical anisotropy profile found through simulated annealing. Figure 3.6b shows the behavior of the value of the misfit function over the course of the

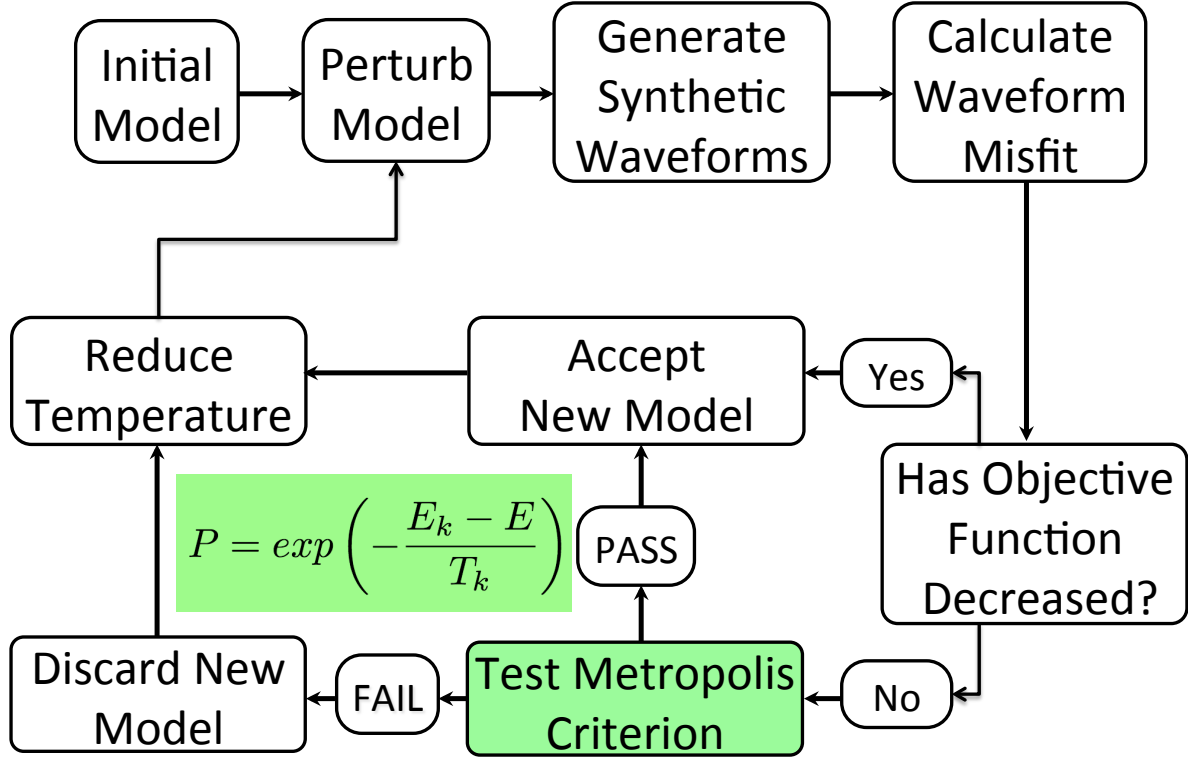


Figure 3.10: Flowchart illustrating the steps in simulated annealing.

inversion, as a percent of the misfit evaluated for the seed model. Also shown, in grey, is the value of the misfit function obtained by using the average  $\varepsilon$  model from travelttime tomography. The x-axis represents all accepted models: those which either pass the Metropolis criterion, or decrease  $E(\mathbf{m})$ . This behavior of  $E(\mathbf{m})$  is exactly as expected. Early on in the routine, the value of the objective function varies wildly with subsequent iterations. As the number of iterations increases, and the temperature  $T_k$  decreases, the variations become more stable.

The ideal 1-D anisotropy profile was extracted by first examining the convergence of simulated annealing via Figure 3.6b. The dip in the error function at model 1200 was interpreted as the point where the algorithm discovers the global minimum, and I refer to this as the ‘phase change’ (as it is where the model ‘freezes’ close to its final state). An ensemble of all models accepted after this phase change was then taken, and those which lay within one standard deviation of the absolute best-fit model were averaged to give the ideal model.

The range of models lying within one standard deviation of the absolute best fit model is given by the shaded area in Figure 3.6a. Note that near the edges of the model, the variance becomes very small. This is due to the fact that limited data are available at these locations, and the model at the final spline point was tied down to that found in the seed model.

There are several physical explanations consistent with the values of  $\epsilon$  depicted in Figure 3.6a. From the borehole logs, there are reports of multiple fracture symmetries within the low velocity massive sulphide zone, specifically major fractures at  $30^\circ$ ,  $50^\circ$ , and  $60^\circ$  to the core axis. As the fast axis of crack induced anisotropy is usually parallel to the fracture axis (Tsvankin, 2012), we see here that the VTI assumption may be violated, and in fact the fast axis may be closer to the horizontal. Of course, the assumption of a single fast axis already assumes that the anisotropy is at least TTI, which is likely suspect. Instead, the negative values obtained for  $\epsilon$  are indicative of a symmetry axis that is far from vertical. It should also be noted that it is very likely that  $\epsilon \neq \delta$ . If this is the case, the  $\epsilon$  profile found with simulated annealing has the additional responsibility of accounting for the effect of  $\delta$  on the slowness, and as such will differ from the true value of  $\epsilon$ .

Figure 3.11c illustrates the effect of the new anisotropy model derived from simulated annealing on the synthetic time domain waveforms. Much of the misfit at larger offsets has now been explained, and cycle skipping is no longer apparent. Improvements can also be seen in the frequency domain wavefields. Figure 3.12c shows the estimated frequency domain wavefield at 1200 Hz after simulated annealing. Comparing this with the real data in Figure 3.12a, and that modeled with the traveltime anisotropy model in Figure 3.12b, we see some significant improvements. First, at near offsets (between sources 120 and 130, and receivers 40 and 50) we see that the wavefield post-simulated annealing matches the real data more closely than it did with the traveltime anisotropy model. As well, just above the low velocity zone (shots 120 - 140, and receivers 30 - 40) the phase in Figure 3.12c is closer to the real data, and does not contain the rapid variations seen in Figure 3.12b.

Figure 3.9b shows the estimated source signature, calculated using the traveltime velocity model and the anisotropy profile found through simulated annealing. We can see that this signature is much more coherent than that shown in Figure 3.9a.

The elimination of the majority of the cycle skipping seen prior to simulated annealing suggested that the new anisotropy model was far more appropriate for use in waveform tomography.

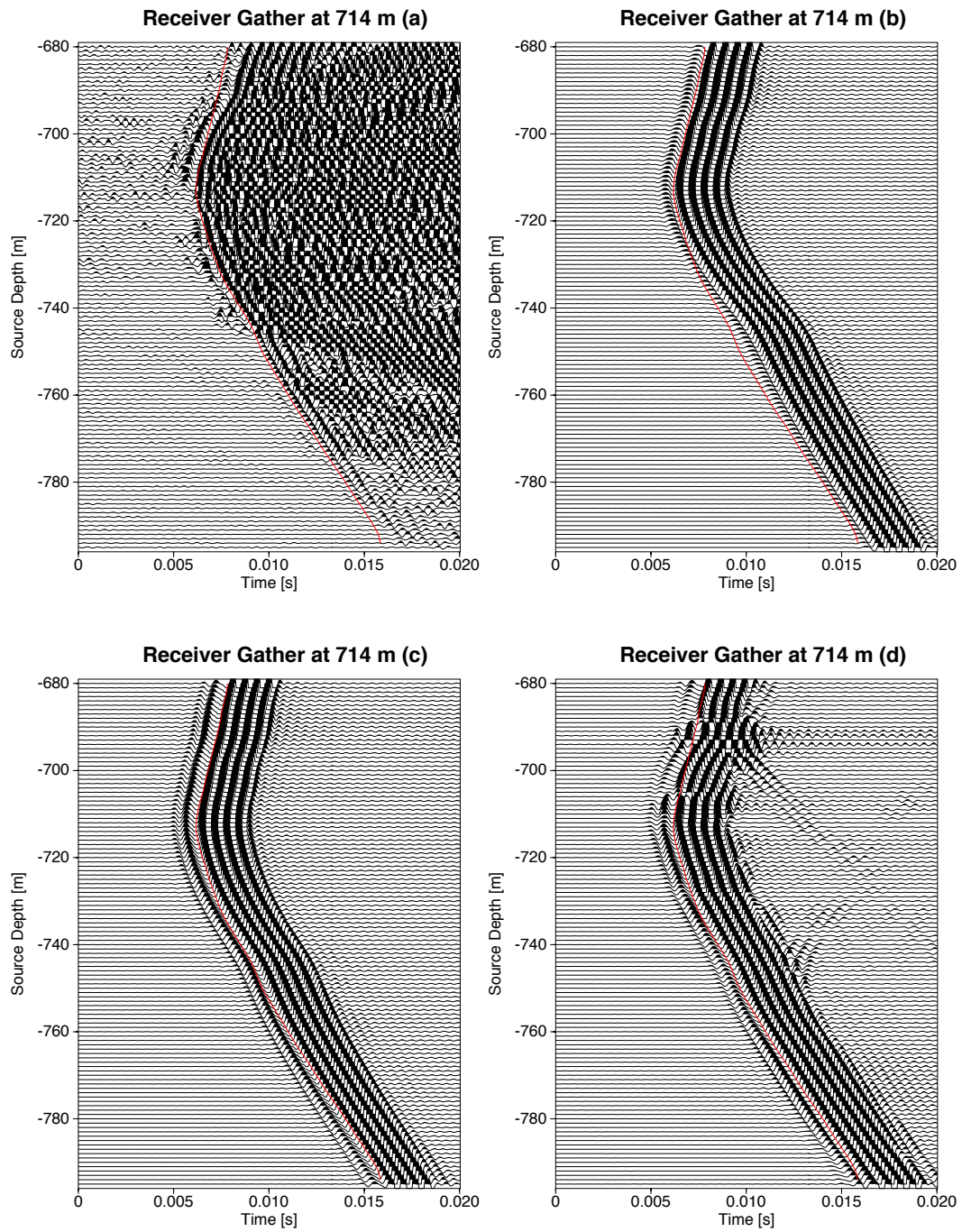


Figure 3.11: Receiver gathers. (a) Real, unfiltered data; (b) synthetic data, modeled using TT velocity model and TT anisotropy; (c) synthetic data, modeled using TT velocity model and SA anisotropy; (d) synthetic data, modeled using WT velocity model and SA anisotropy.



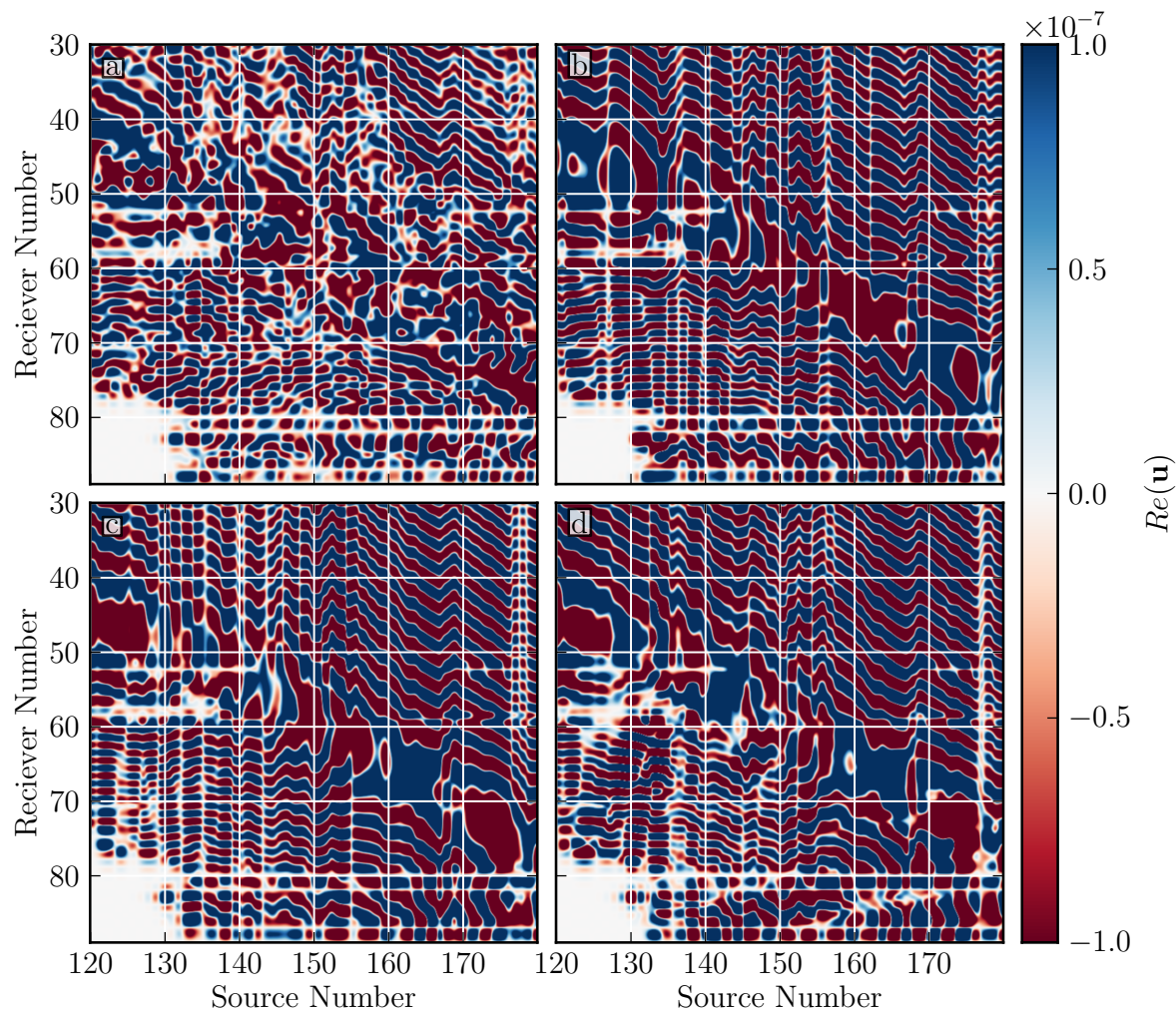


Figure 3.12: A sample of the real frequency domain wavefields ( $Re(\mathbf{u})$ ) at 1200 Hz for the real crosshole data, zoomed in on the area surrounding the low velocity zone. (a) Observed wavefield; (b) estimated wavefield using the traveltime velocity model, and the anisotropy model found from traveltime tomography; (c) estimated wavefield using the traveltime velocity model, and the anisotropy model found through simulated annealing; (d) estimated wavefield after waveform tomography, using the anisotropy model found through simulated annealing.

### 3.4.2 Velocity

Once the initial model satisfied the half cycle criterion, I was able to proceed confidently with waveform tomography. In order to take advantage of the robustness of phase in the presence of elastic effects, and due to the suspected low cut filtering, the residuals were defined as the difference between the synthetic and observed waveform phase, and wave amplitudes were ignored. A sequential frequency selection strategy was employed (Sirgue and Pratt, 2004; Kamei et al., 2012), in which blocks of four frequencies were inverted at a time, beginning with 1000 Hz (the lowest, clean frequency), and increasing in 100 Hz intervals. After five iterations, a new block of frequencies was inverted, retaining the highest three frequencies from the preceding block. This was continued until a block containing 1600 Hz, the highest frequency used in the inversions, was inverted. At each iteration, the phase and amplitude of the (averaged) source was updated according to equation (3.9). The resultant velocity model from waveform tomography can be seen in Figure 3.14. A synthetic receiver gather computed using this velocity model can be seen in Figure 3.11.

When solving problems as undetermined as those explored by waveform tomography, a correlation of results with any relevant information is essential in determining the quality of a model. Here we compare our inversion results with borehole sonic logs, and our predicted waveforms with waveform data in the time and frequency domain.

A comparison of the real and estimated wavefields, calculated using the final velocity model from waveform tomography, is given in Figure 3.13. Figure 3.13a shows the real data, Figure 3.13b gives the synthetic data modeled through the final velocity model from waveform tomography, and Figure 3.13c gives the final wavefield residuals. Through layouts 1 and 2, the waveform residuals are acceptable.

A significant phase error is present along the main diagonal, representing the near offsets. I attribute this to the presence of tubes waves in the data, which are not modeled in the inversions. At near offsets, the tubes waves are consistently mixed in with the first arrivals. As the offsets increase, the slow-moving tube waves are windowed out from the data. Also apparent in Figure 3.13c are some vertical stripes, which I attribute to misfit introduced by source-borehole coupling, or possible source mislocations.

Finally, in these two panels, a trend in phase mismatch is present. For source receiver pairs where the source is located below the receiver, the phase error is dominantly negative, while for pairs where the receiver is below the source, the error is dominantly positive. This suggests that the error is dependent on angle, and that our anisotropy model may not be entirely correct. This is expected due to the 1-D elliptical approximation of what is likely a 3-D orthorhombic or triclinic media.

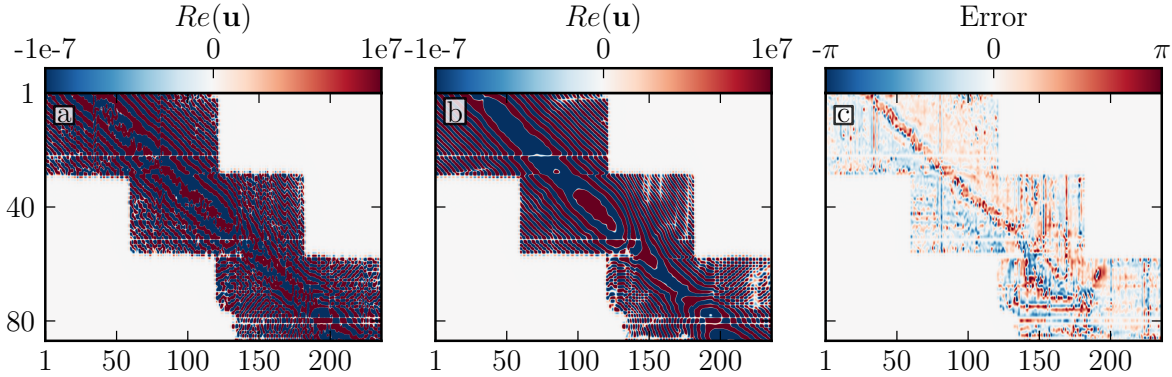


Figure 3.13: (a) Real, frequency domain wavefield ( $Re(\mathbf{u})$ ); (b) Synthetic real-valued, frequency domain wavefield calculated using the final velocity model from waveform tomography; (c) wavefield phase residuals in the final model.

In the third (lower) panel, the results are less encouraging. Semi-circular rings of constant phase mismatch are apparent, and these are indicative of cycle skipping at 1300 Hz. This is not entirely unexpected, as these particular data (shots 130 - 180, shooting to receivers 60 - 87) are likely the most difficult; they represent those wavepaths which spend the longest time within the dipping low velocity zone. Since this zone is where the 1-D anisotropy assumption is most likely to be invalid, and is the region most corrupted by tube waves, we expect these data to be the most difficult to fit. Further work in improving the starting model may be needed to remedy this cycle skipping issue. In Chapter 4, I will present synthetic noise-free data, in which a similar residual pattern emerges. Figure 3.12d shows a zoomed-in view of the estimated 1200 Hz wavefield in the low velocity zone, calculated using the velocity model from waveform tomography.

Borehole sonic velocities were available from each of the boreholes, and are presented along with the equivalent vertical velocities extracted from the waveform tomography velocity model in Figure 3.14. Since the sonic velocities measure the seismic velocity along the near vertical boreholes, a correction for anisotropy is needed to compare these data to the horizontal velocities in Figure 3.14. Since we are assuming elliptical anisotropy, an estimate of the vertical velocity can be obtained by multiplying the horizontal velocities by  $(1+\varepsilon)^{-1}$ . These equivalent vertical velocities are plotted in Figure 3.14. The waveform tomography results are in blue, and the sonic velocities are plotted in green. The match is not exact throughout the model, although this is expected. Borehole sonic velocities are usually measured at approximately 20 kHz, while our inversions were for much lower frequencies. As well, the 1-D elliptic VTI assumption is likely invalidated by the complex shearing and fracturing present. This may lead to not only image distortion, but also



to errors in calculating the equivalent vertical velocities (as the symmetry axis may be at some other, arbitrary angle). In the source hole, an apparent static shift in vertical velocities is apparent, where the inverted velocities are consistently faster than the sonic velocities, although they show the same major features. This may be a result of the true, 2-D anisotropy structure. Nevertheless, some promising observations emerge.

In the receiver hole, a sharp dip is seen in the sonic velocity at approximately 675 m. This is echoed closely by a corresponding dip in the inverted velocity. By examining the borehole core logs, it is possible to tie these velocity anomalies to a 1 m thick massive sulphide veinlet. Further examination of the image allows us to trace the passage of this zone of relatively low velocity: it traverses the panel, and appears to split near the source hole. This is also in agreement with the core logs, on which brecciated host rock in this zone give way to local irregular massive sulphide veins.

The high velocities on either side of the vein at the receiver well exhibit a good match between the sonic and inverted velocities, and appear to be geologically reasonable structures, which however do not appear to extend to the source well (as was expected from the sonic velocities). Just below 650 m, a zone of relatively low velocity can be traced across the panel. While this agrees well with the sonic velocities, there is no corresponding geologic signature noted in the core logs.

Finally, at the top edge of the low velocity zone, it is very apparent that the velocity gradient in the source well is much stronger than that in the receiver well. This matches with the sonic velocities, and also has geologic significance. In the receiver well, significant quantities (40% - 75%) of massive sulphide are present in the troctolites beginning at 689 m, which results in a decrease in velocity. Also, between 697 m and 700 m, the existence of a granite intrusion is noted, and this may be represented as a slight leveling off of the velocity gradient in both the sonic and inverted velocities. While the receiver well intersects a zone where the massive sulphide content increases gradationally, the source well experiences a much sharper geological contrast. Between 683 m and 686 m, the sulphide content increases to 40%, with the massive sulphide zone beginning directly below.

The match between the sonic velocities and inverted velocities within the massive sulphide zone is slightly less promising than those outside. While the overall structure and velocities match well, there are some discrepancies when examining fine structure. We attribute this partially to the breakdown of the 1-D anisotropy assumption, when combined with the 2-D massive sulphide zone. The borehole logs report on significant off-axis fracture systems (30, 50, and 60 degrees to the core axis) which also likely leads to the failure of the VTI elliptic approximation. We also expect artifacts due to the

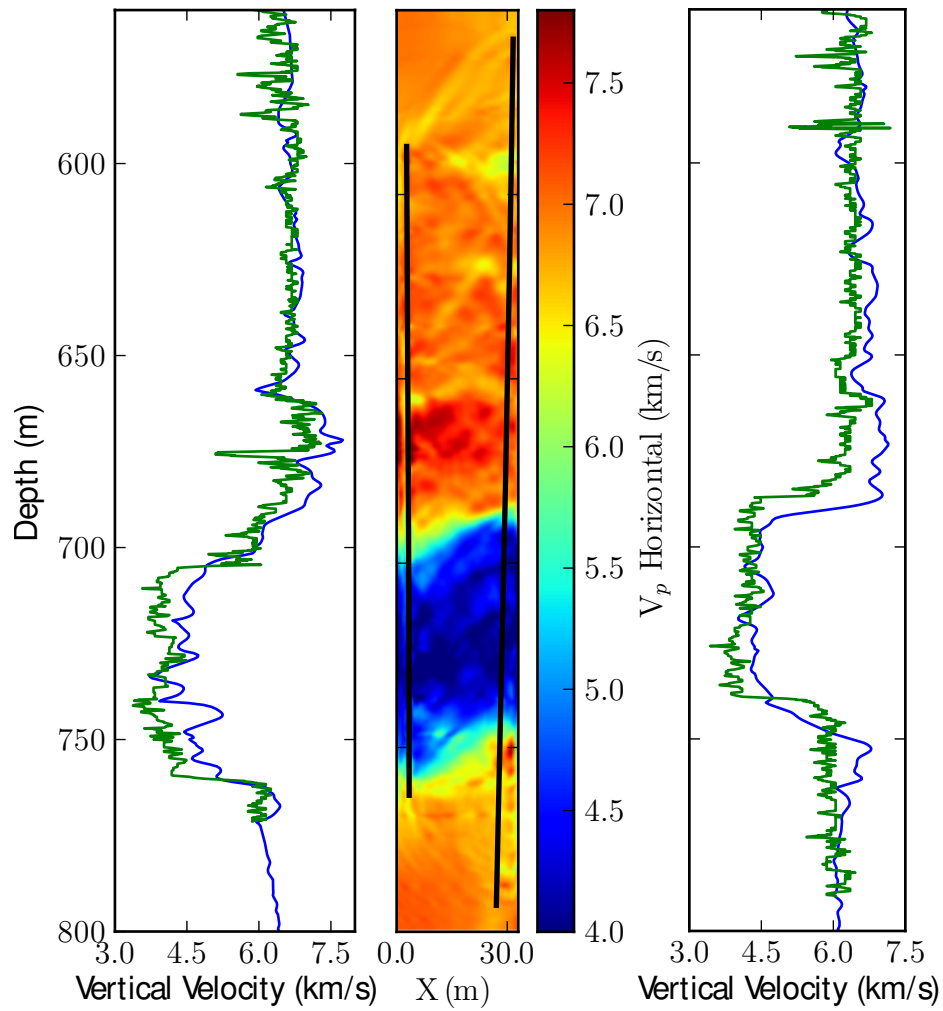


Figure 3.14: Final velocity model obtained from waveform tomography, using the anisotropy profile found from simulated annealing.

presence of tube waves in the data, which are especially apparent within the low velocity zone.

### 3.5 Future Survey Design

The experience with the Voisey's Bay crosshole dataset led to a number of suggestions that should be considered when attempting any future seismic surveys, specifically those designed for waveform tomography.

1. Access to the complete, unfiltered waveform data is essential. The acausal band-pass filtering to greater than 1000 Hz may have had detrimental effects on the final inversion products, as there were no coherent low frequency data to stabilize the inversion. In future surveys, data should be handed to those responsible for processing in a raw, unfiltered state. Ideally, the processing geophysicists should be involved in the data acquisition themselves, so that the processing flow of the waveform data is completely controlled.
2. The use of a source rich in low frequencies is important for the stability of waveform tomography. Whenever possible, a seismic source capable of producing a strong, low frequency signal should be used. Current viable options are (a), a powerful, impulsive source such as a borehole hammer (Vibrometric, 2009a), or (preferably) (b), a modified swept impact source capable of producing low frequencies. A modified swept impact source is preferred, as impulsive sources may damage and collapse unstable boreholes. As well, when choosing a source, the method of borehole coupling (water, clamping) and radiation pattern should be noted, as these factors may require special attention to be paid to the estimated source signature. Finally, as important as low frequencies are for waveform tomography, their inclusion should not take priority over obtaining a good signal to noise ratio.
3. The source and receivers should be designed with the target in mind. In the survey presented here, there was a discontinuity in the hydrophone chain near the top of the low velocity massive sulphide zone. As this boundary is very challenging to image, good data in this region is especially essential. Future survey geometries should be designed so that the sources and receivers are continuous across discontinuities.
4. Accurate borehole location techniques should be used. Ideally, a north-seeking instrument, as well as a gyroscopic borehole measurement tool, should be used, and should report on the borehole location with the finest detail that is feasible.

Accurate location estimates for sources and receivers are essential, although the magnitude of location effects should decrease as survey offsets are increased. If possible, additional instrumentation (such as gamma ray detectors) should be attached to the source and receiver chain, as they may help to constrain locations.

5. Larger offset surveys should be attempted. At small offsets, source-receiver mislocation can be a significant issue. As well, the level of detail gained from extremely small offset surveys may not be worth the effort put into acquisition, as an accurate geologic map can be constructed from the core. At larger offsets, the chances of gaining valuable information through waveform tomography are increased.
6. Source and receiver intervals should be as small as possible, to prevent aliasing. In previous sections, we remarked that tube waves were present in the data, and due to aliasing these could not be removed cleanly. If the source and receiver locations were close enough so that the tube waves were not aliased in the frequency range of interest, their removal would be trivial with frequency-wavenumber filtering. The optimal spacing will be situation-specific, as it depends on local rock velocities, but for the Voisey's Bay dataset, halving the receiver spacing would have allowed us to remove tube waves cleanly from frequencies up to 1800 Hz.

Of course, in undertaking any survey, there are practical considerations to be made (i.e. the expense of shipping equipment, or the length of time needed to complete the survey). The first priority should be the complete coverage of the target zone.

# Chapter 4

## Results from a Synthetic Crosshole Dataset

To test the effectiveness of simulated annealing, and to understand the failure of anisotropic traveltimes tomography described in Chapter 3, synthetic velocity and anisotropy models were created, and seismic data were numerically simulated in the resulting models. The velocity model was based on the Voisey’s Bay 540-542 crosshole dataset processed in Chapter 3, and the anisotropy models were designed to simulate the essential elements of the inverted field data. Three 2-D anisotropy models were tested: (1) a VTI elliptical model ( $\varepsilon = \delta$ ), (2) a VTI an-elliptic model ( $\varepsilon \neq \delta$ ), and (3) a TTI an-elliptic model ( $\varepsilon \neq \delta$ ,  $\phi = 30^\circ$ ).

### 4.1 Model Details

The velocity model was constructed to be similar to that obtained from the field data in Chapter 3. The true models for  $V_p$  and  $\varepsilon$  are given in Figure 4.1. The source-receiver geometry was the same as that for the field data, with boreholes approximately 30 m apart, sources spaced 1 m apart, and receivers spaced 2 m apart. The background velocity was set to be 7500 km/s. A tilted, low velocity layer was introduced, simulating the actual location and orientation of the massive sulphide zone, estimated from the borehole sonic logs (see Figure 3.14). As well, a layer of slightly higher velocity was placed just above the low velocity layer, corresponding to the slight increase in borehole sonic velocities seen above the massive sulphide zone. Random heterogeneities were then added to the model using the methods of Kamei et al. (2005); the magnitudes of these heterogeneities were estimated from observed fluctuations in the borehole logs. In the high velocity zone, the heterogeneities had a standard deviation of 300 m/s, while in the low velocity

layer the heterogeneities had a standard deviation of  $210 \text{ m/s}$ . I used a vertical:horizontal correlation length of 1:4 for these heterogeneities, to simulate the impact of horizontal layering on the anisotropy inversions. Finally, the long axis of the heterogeneities added to the low velocity zone was rotated to coincide with the dip of the zone itself. Density values for the high and low velocity zones were chosen to correspond to the average density of troctolite ( $2.7 \text{ g/cc}$ ) and pyrrhotite ( $4.6 \text{ g/cc}$ ) respectively. The poisson's ratio was set everywhere equal to 1.87.

Values for  $\varepsilon$  and  $\delta$  were also included. For the VTI elliptic test, in the high velocity zone,  $\varepsilon$  was set to a constant 0.15, while in the low velocity layer it dropped to 0.05. For the VTI an-elliptic test,  $\varepsilon$  remained the same, but  $\delta$  was set to 0. The TTI an-elliptic test also retained the value of these parameters, but the axis of symmetry was tilted  $30^\circ$  from vertical.

For all 3 models, elastic time domain TTI finite difference modelling was performed using the freely available open source software package Madagascar (Yan and Sava, 2011). The elastic parameters were discretized onto a  $600 \times 3125$  grid, with each grid point representing an  $8 \text{ cm} \times 8 \text{ cm}$  square. An explosive, Keuper-wavelet source was modeled, with a peak frequency of 666 Hz. The scalar curl-free (pressure) component of the wavefield was extracted at each receiver position; this was to eliminate shear waves. The main purpose of this experiment was to determine whether similar behavior in the traveltimes anisotropy models could be generated, and not to determine the efficacy of acoustic waveform tomography in an elastic environment, as explored by Brenders and Pratt (2007b,a). Madagascar was chosen so that a realistic, anelliptic 2-D anisotropy model could be used, and as well to avoid the pitfalls of performing an ‘inverse crime’ (Wirgin, 2004; Colton and Kress, 1998). An inverse crime refers to creating synthetically generated data with the same forward modelling operator used in inverse modelling (i.e. I would have committed an inverse crime had I used the forward propagator from our waveform tomography routine to generate the synthetic data).

## 4.2 Traveltimes Tomography

To prepare the synthetic data for traveltimes tomography, little effort was required. Due to the noise-free nature of the synthetic data, the majority of the first break picking was very easy, and could mostly be carried out semi-automatically. This being said, there were still significant difficulties at the edges of the low velocity zone, where manual intervention was required.

Traveltimes tomography was performed following the same philosophy given in section

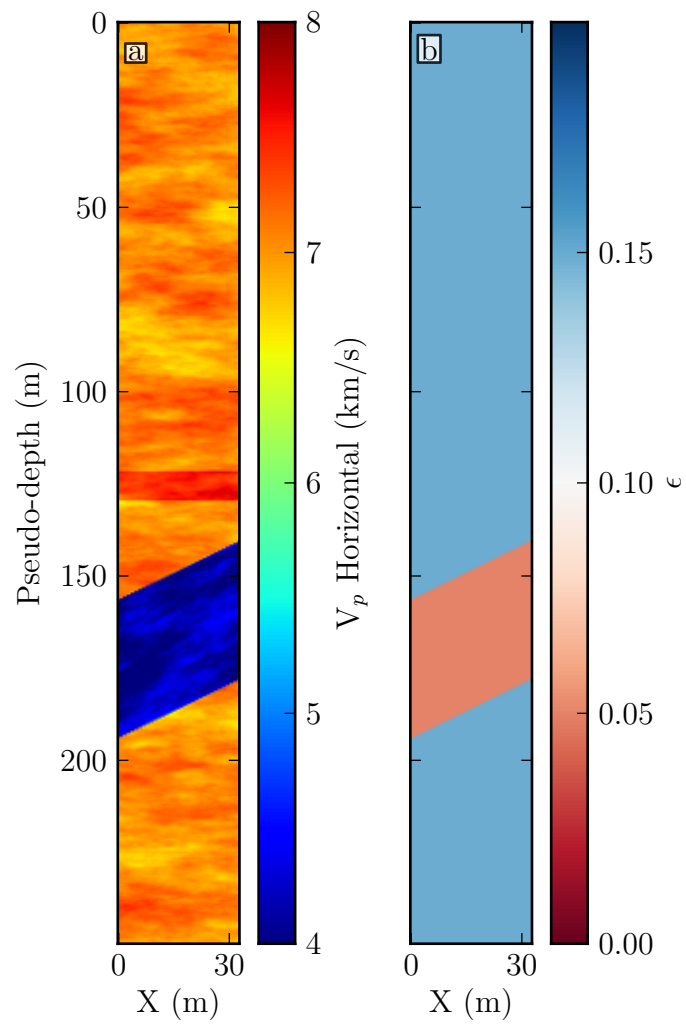


Figure 4.1: True models used in the synthetic tests: (a)  $V_p$  horizontal; (b)  $\epsilon$ .

3.3. Five nonlinear inversions were carried out, during which the constraint on model roughness was gradually relaxed. Traveltime velocity and anisotropy models can be seen in Figure 4.2. A similar pattern to that observed in the real data can be seen in these synthetic examples, where the  $\varepsilon$  values in the low velocity zone become larger as the constraints on anisotropy are increased. In the case of the unconstrained inversions, the values of  $\varepsilon$  in the high velocity zone are overestimated by approximately 5%. This is attributed to the fine scale of the horizontal heterogeneities that lie outside the resolution capability of ray tomography.

## 4.3 Waveform Tomography

Unlike the real data, the synthetic data were noise-free, so minimal pre-processing was required. The time domain waveforms were simply windowed and bandpass filtered before being transformed into the frequency domain, as input for waveform tomography. An example of a frequency domain wavefield is given in Figure 4.3a. We can see a significant increase in waveform complexity in the low velocity zone (sources 120 - 180; receivers 59 - 70). Also apparent is aliasing in this zone, which, as in the real data, suggests that the velocity here is dramatically lower here than in the rest of the model. Figure 4.4a shows the original, or ‘true’ 1600 Hz wavefield, zoomed into the vicinity of the low velocity zone.

### 4.3.1 Anisotropy

To prepare the anisotropy models for waveform tomography, I extracted a 1-D vertically varying anisotropy profile by averaging the traveltime  $\varepsilon$  model in Figure 4.2e over horizontal layers. The resultant profile can be seen in Figure 4.5. As in the real data example from Chapter 3, simulated annealing was run with a warm up phase consisting of 100 iterations, during which the Metropolis criterion was used to calculate the effective temperature which would allow 50% of proposed models to be accepted. The behavior of the objective function after the warm up phase is given in Figure 4.5. Here we see a similar pattern to the real data, with the value of the objective function decreasing over time. As before, a ‘phase change’ occurs around iteration 1000. As before, the best-fit 1-D  $\varepsilon$  profile was calculated by examining all the models accepted after this final phase change, and taking the average of those which lay within one standard deviation of the absolute best fit profile. The value of the objective function decreases significantly over the course of simulated annealing, much more so than in the real data example. This is attributed to the noise free characteristics of this synthetic test.



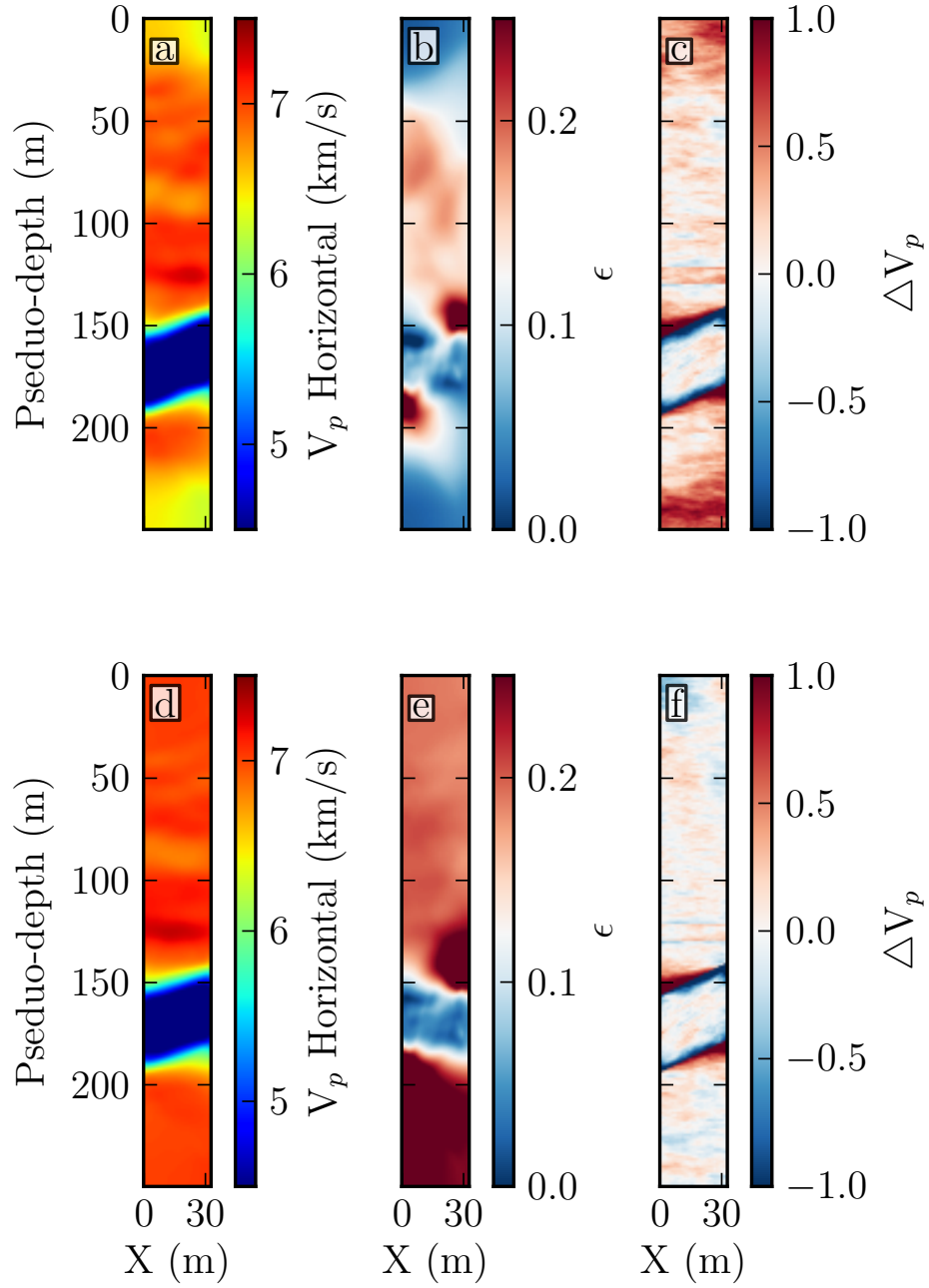


Figure 4.2: (a) Velocity model found through traveltime tomography; (b)  $\epsilon$  model found through traveltime tomography; (c) velocity error ( $\Delta V = V_{true} - V_{est}$ ); (a - c) are inverted with strong constraints on anisotropy; (d) velocity model found through traveltime tomography; (e)  $\epsilon$  model found through traveltime tomography; (f) velocity error ( $\Delta V = V_{true} - V_{est}$ ); (d - f) are inverted with no constraints on anisotropy.

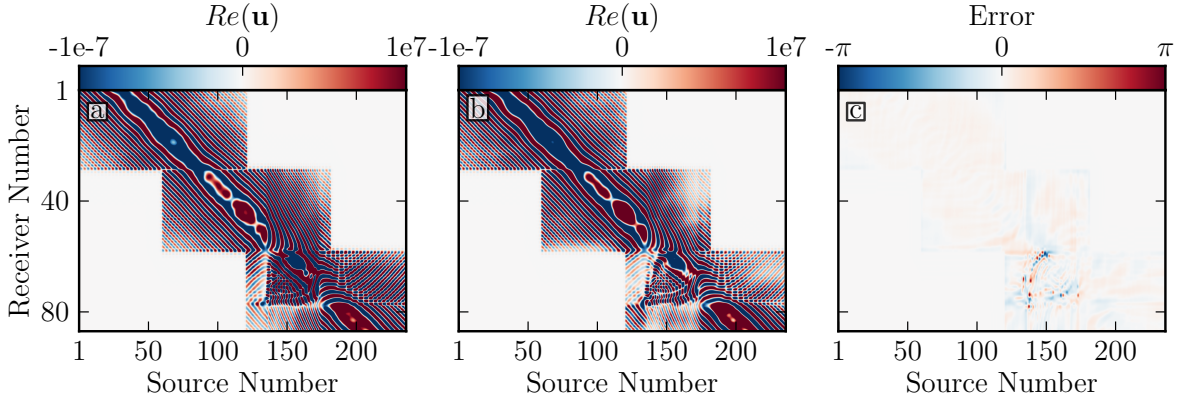


Figure 4.3: A sample of the real frequency domain wavefield ( $Re(\mathbf{u})$ ) at 1300 Hz for synthetic elastic data. (a) True modeled wavefield; (b) estimated wavefield after waveform tomography, using the anisotropy profile found through simulated annealing; (c) wavefield phase residuals after waveform tomography, using the anisotropy profile found through simulated annealing.

Figure 4.4b and and Figure 4.4c compare the estimated frequency domain wavefields in the vicinity of the low velocity zone before and after simulated annealing, respectively. As in the real data, we see an improvement in waveform fit. Specifically, when comparing with the true data in Figure 4.4a, we see that cycle skipping has been eliminated at sources 120 - 140, between receivers 78 and 87. As well, a discontinuity at source 137 has been partially remedied by the new anisotropy model. Note finally the dramatic change in waveform phase around source 130 and receiver 50.

Figure 4.5a shows the  $\varepsilon$  profile obtained from simulated annealing, compared to that found with travelttime tomography, and to the true (averaged) model. We see a good agreement between the (averaged) actual values of  $\varepsilon$  (15% in the high velocity zone, and 5% in the low velocity zone) and those found from simulated annealing. As well, in the low velocity zone, the value of 0.05 is also recovered quite well. There are some discrepancies across the boundary, this is expected due to the 2-D nature of the anisotropy in this location. This  $\varepsilon$  profile was used for subsequent waveform tomography, as discussed in the next section.

### 4.3.2 Velocity

Following the extraction of the ideal  $\varepsilon$  profile from simulated annealing, waveform tomography was performed on the synthetic models, using the same strategy as for the real data. A sequential frequency strategy was used, with each block of five iterations using four frequencies, spaced at 100 Hz intervals. After five iterations, a new frequency was

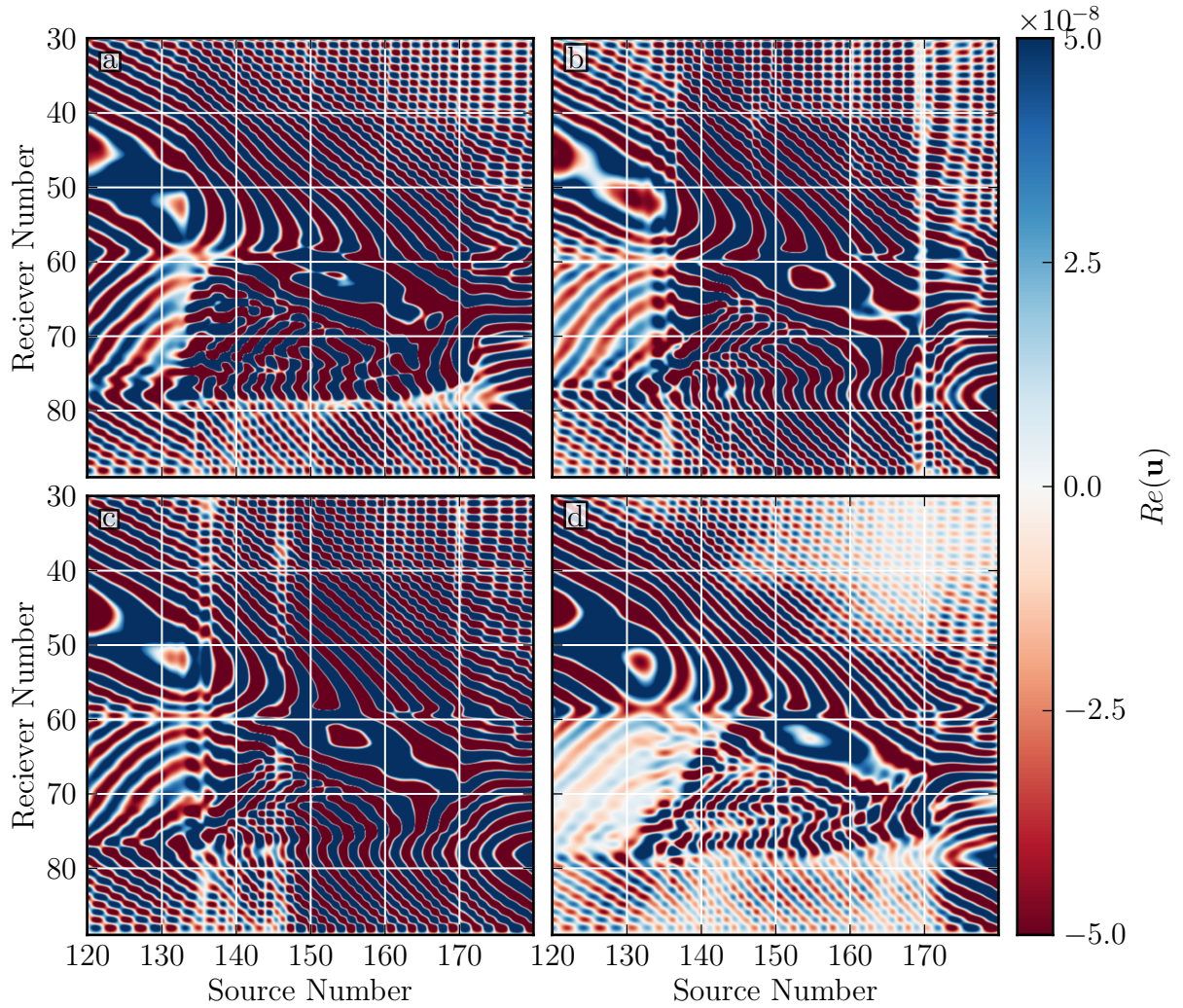


Figure 4.4: A sample of the real frequency domain wavefield ( $Re(\mathbf{u})$ ) at 1600 Hz for synthetic elastic data, zoomed in on the area surrounding the low velocity zone. (a) True modeled wavefield; (b) estimated wavefield using the traveltime velocity model, and the anisotropy model found from traveltime tomography; (c) estimated wavefield using the traveltime velocity model, and the anisotropy model found through simulated annealing; (d) estimated wavefield after waveform tomography, using the anisotropy model found through simulated annealing.

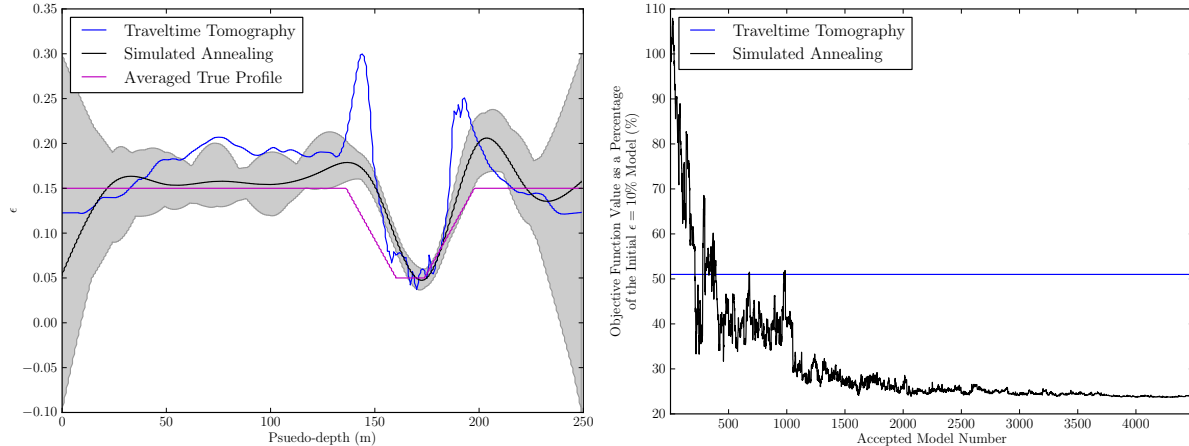


Figure 4.5: Left:  $\epsilon$  profile obtained from traveltime tomography (blue) and the best fit profile obtained from simulated annealing (black). The shaded area represents the range of anisotropy profiles that lay within one standard deviation of the absolute best fit profile. Right: objective function behavior over the course of simulated annealing (black), compared to the value obtained from traveltime tomography (blue). The best fit profile from simulated annealing was constructed by considering all models after the ‘phase change’ around iteration 1000.

added, and the lowest frequency from the previous block was dropped. The lowest frequency used was 500 Hz, and the highest was 1600 Hz. At each iteration, the amplitude and phase of the estimated source signature was updated.

Velocity models from traveltime and waveform tomography are given in Figure 4.6. We see that the final model from waveform tomography (Figure 4.6d) matches the true model (Figure 4.6a) very well. What is also apparent is that the waveform tomography result using the traveltime anisotropy model (Figure 4.6c) is significantly fast. 1-D velocity profiles extracted from the velocity models are shown in Figure 4.7, and further support these observations. The velocities found from waveform tomography using the  $\epsilon$  profile from traveltime tomography are significantly higher compared to the true model, while the waveform tomography results with the anisotropy model from simulated annealing matches the true model very well.

The geometry of the survey results in the majority of the data representing wave propagation at non-zero horizontal angles. Since our goal in waveform tomography is the minimization of the misfit function  $E(\mathbf{m}) = \| \mathbf{u}_{est} - \mathbf{u}_{obs} \|$ , it is immediately apparent that the value of the misfit function will be significantly affected by anisotropy (as the majority of the estimated wavefield is susceptible to anisotropy). As such, the horizontal velocity plays a minimal role. If the estimated anisotropy is too strong, and as a result the non-horizontal velocities are too slow, this can be compensated for by increasing the

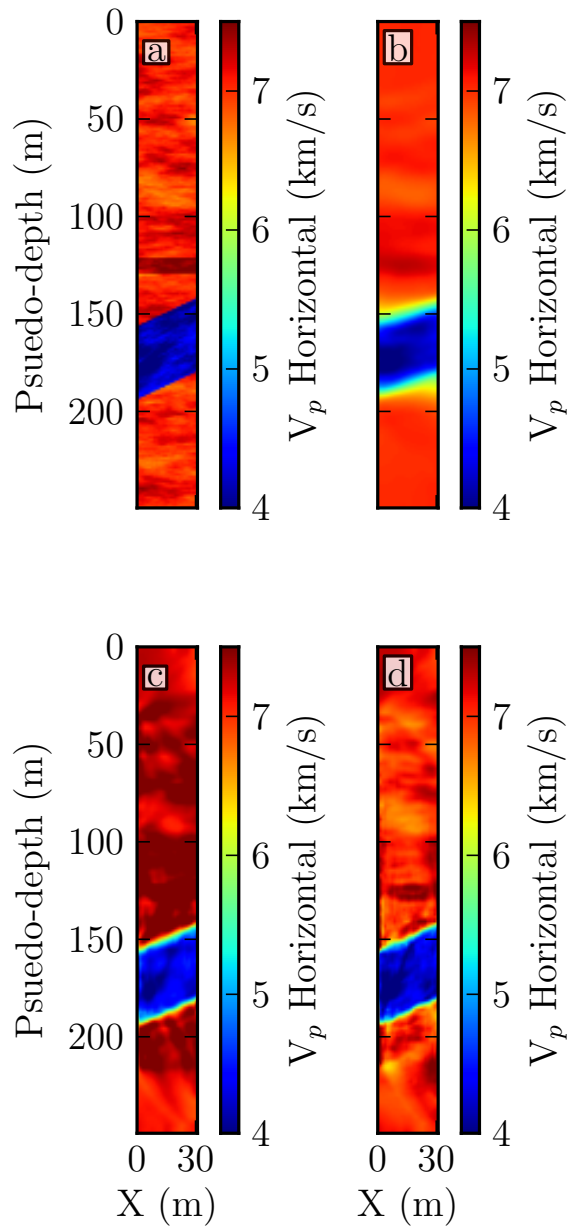


Figure 4.6: (a) True, horizontal velocity (as in Figure 4.1); (b) velocity model found through anisotropic traveltime tomography; (c) velocity model found through waveform tomography, using the anisotropy profile from traveltime tomography; (d) velocity model found through waveform tomography, using the anisotropy profile from simulated annealing.

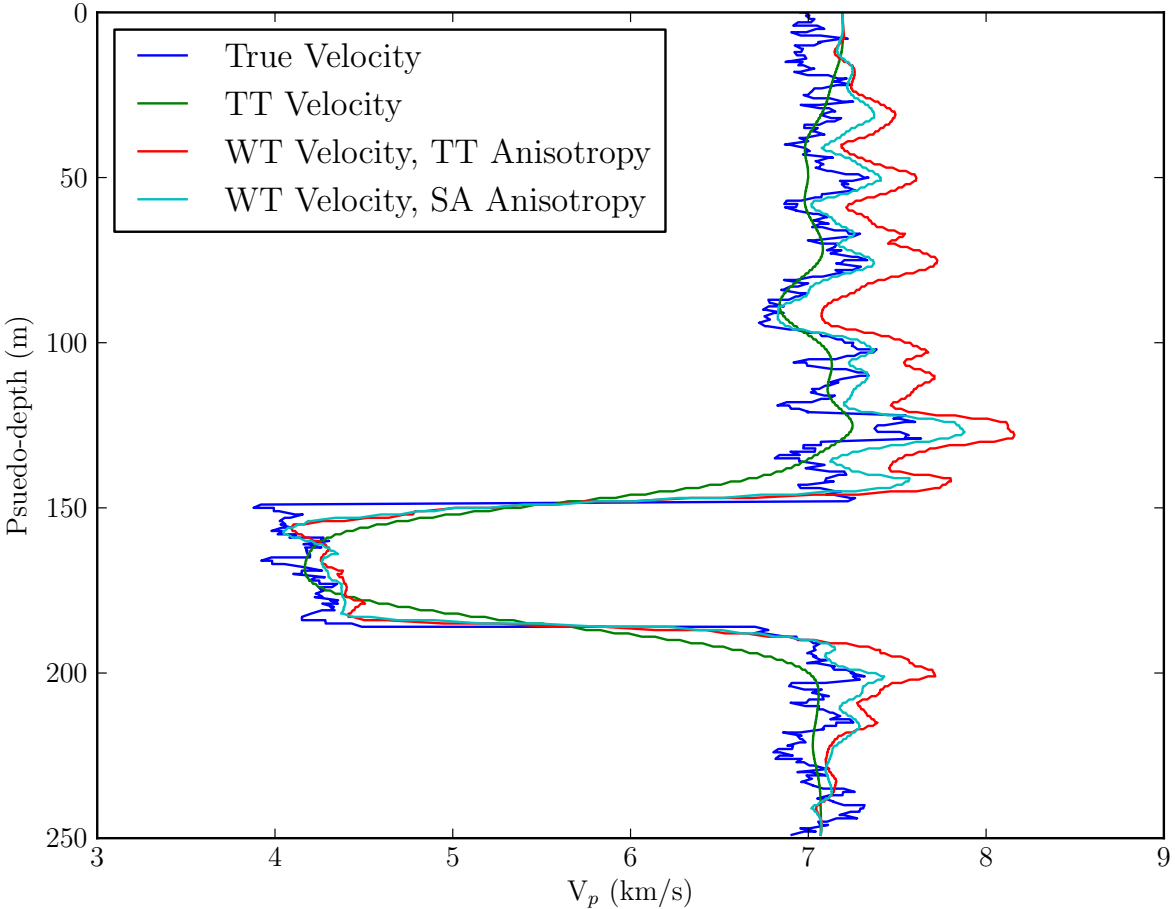


Figure 4.7: Horizontal velocity profiles taken through the middle of the velocity model.

horizontal velocity. Since truly horizontally propagating waves contribute negligibly to the misfit function, there is not much penalty in this incorrect adjustment. I believe this is why the horizontal velocity is significantly faster when simulated annealing is not used. Even when the anisotropy is estimated from simulated annealing, the horizontal velocities are slightly too fast on average (Figure 4.7). This is likely due to the fact that the ideal anisotropy profile, while good, is not perfect, and in fact is slightly too strong. Nevertheless, the velocity model created using simulated annealing is much more accurate than the inversion using traveltime anisotropy.

Figure 4.3b shows the final frequency domain wavefield, modeled using the results from waveform tomography. By comparing this to Figure 4.3a, we see a good match almost everywhere, i.e. at most sources, receivers, and offsets. Figure 4.3c shows the residual phase differences between the modeled and true wavefield. Again, the phase differences throughout most of the model are negligible. However, for sources and receivers near the edges of the low velocity zone, we see a significant increase in wavefield



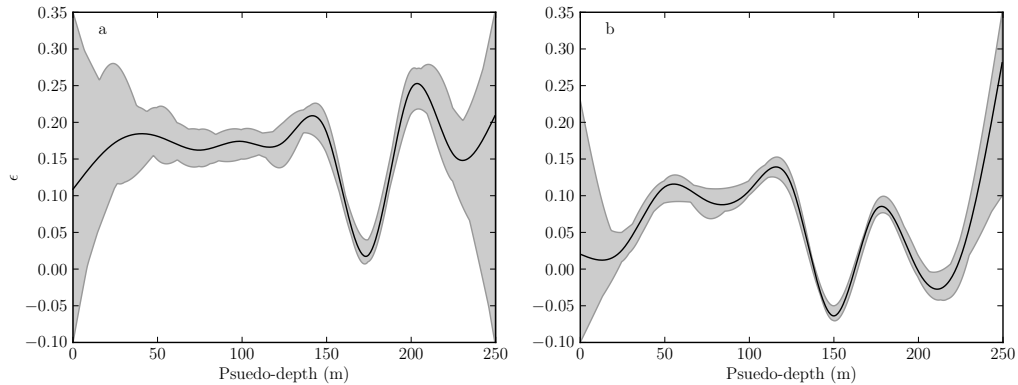


Figure 4.8:  $\varepsilon$  profile obtained from simulated annealing for an-elliptic models: (a)  $\phi = 0^\circ$ ; (b)  $\phi = 30^\circ$ . The shaded area represents the range of anisotropy profiles that lay within one standard deviation of the absolute best fit profile.

phase residuals, similar to the pattern seen in the real data. This is expected, as the velocity contrast in this area is extremely strong, and has spatial components beyond the resolution limit of the frequencies used. As a result, it is also in this area where the estimated anisotropy is furthest from its true value.

### 4.3.3 An-elliptic VTI Media ( $\delta = 0$ , $\phi = 0^\circ$ )

To test the performance of simulated annealing on models with an-elliptic anisotropies,  $\delta$  was set to 0 throughout the model, while  $\varepsilon$  remained as in the earlier example (Figure 4.1b). As described above, the curl-free component of the elastic waveforms were numerically modeled using Madagascar. First break picking was performed using PROMAX, and the data were time-windowed and bandpass filtered before being transformed into the Laplace-Fourier domain for use in waveform tomography.

The inverted 1-D  $\varepsilon$  profile found using simulated annealing is shown in Figure 4.8a. Note that the profile is not an exact match with the true  $\varepsilon$  profile. This is expected, as  $\varepsilon \neq \delta$ , and the elliptical assumption made by waveform tomography is invalidated. As such, the estimated values of  $\varepsilon$  from simulated annealing will also approximate the effects of  $\delta$  as much as possible. The final horizontal velocity model found with waveform tomography, using the  $\varepsilon$  profile from simulated annealing, is given in Figure 4.9. We see a relatively good match with the true model, and conclude that simple an-elliptic VTI velocity models are resolvable using the 1-D elliptic assumption.

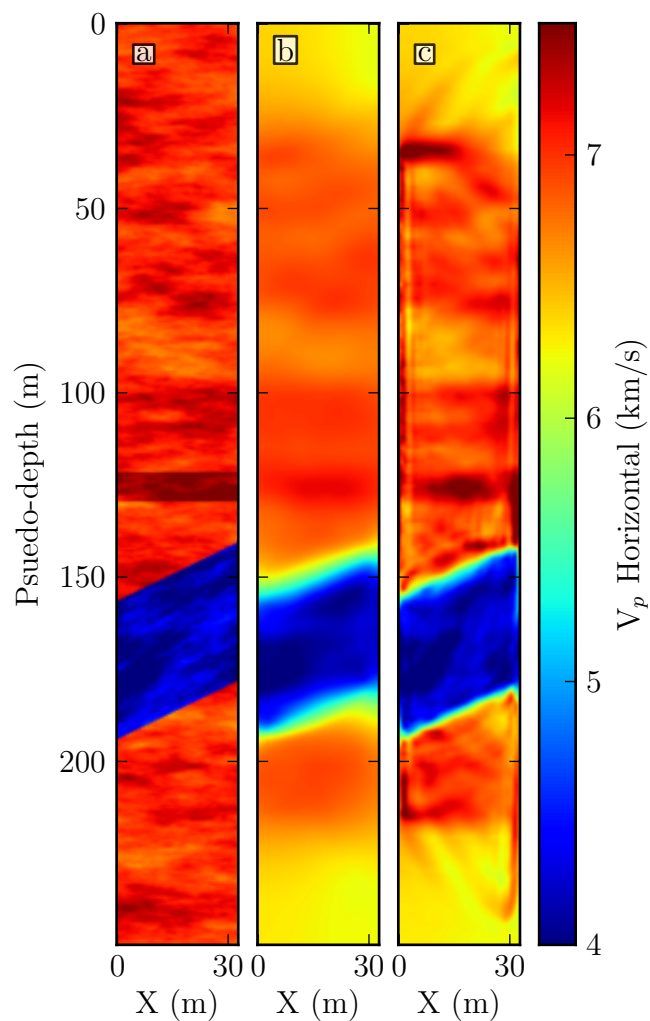


Figure 4.9: An-elliptic VTI synthetic test. (a) True velocity model (as shown in Figure (4.1)), (b) Velocity model obtained from traveltime tomography; (c) Velocity model obtained from waveform tomography, using anisotropy estimated from simulated annealing. The anisotropy was an-elliptic and VTI.



#### 4.3.4 An-elliptic TTI Media ( $\delta = 0$ , $\phi = 30^\circ$ )

To test the performance of simulated annealing on an-elliptic, TTI anisotropy systems,  $\delta$  was set equal to 0 throughout the model, while the distribution of  $\varepsilon$  was retained (i. e. as in Figure 4.1). The dip angle  $\phi$  was set at a constant  $30^\circ$  from the vertical. Again, the curl-free component of elastic waveforms were numerically modeled using Madagascar, and first break picking was performed using PROMAX. The data were time-windowed and bandpass filtered before being transformed into the Laplace-Fourier domain for use in waveform tomography.

The inverted 1-D epsilon profile is shown in Figure 4.8b. The profile is significantly different than the true value of  $\varepsilon$ . Interestingly, in the low velocity zone, the value of  $\varepsilon$  is slightly negative. This is comparable to  $\varepsilon$  values obtained from the real data, in a similar environment. This lends credence to the interpretation that negative  $\varepsilon$  values recovered for the real data represent a large deviation of the symmetry axis from vertical. The final horizontal velocity model found through waveform tomography is given in Figure 4.10. Although the velocity model from traveltimes tomography succeeds in generating a reasonable, albeit smooth, velocity model, we see that the waveform tomography result represents a poor match with the true model, with significant image distortion present.

## 4.4 Discussion

It is worth delving into the question of why waveform tomography, combined with simulated annealing, succeeds in generating suitable anisotropy models, while traveltimes tomography fails. To answer this question, let us look closely at the synthetic dataset. As seen in Figure 4.2, changing the constraints on anisotropy produce completely different  $\varepsilon$  models, but the resultant velocity models are qualitatively similar. This suggests the well understood notion that some of the anisotropy parameters are within the null space of the traveltimes tomography problem, and do not have much of an effect on the traveltimes fit. Nevertheless, the slight reduction in traveltimes residuals, and the consistent manner in which the anisotropy models behave in the presence of a strong velocity contrast, lead us to believe that certain situations squeeze the normally vastly under-determined anisotropy parameters into the range of the traveltimes tomography inverse problem. Pratt and Chapman (1992) achieve this squeezing by enforcing smoothness constraints, but here I make the case that the fundamental physical assumptions of classic traveltimes tomography make the anisotropy parameters very difficult to estimate under the conditions encountered with the real data.

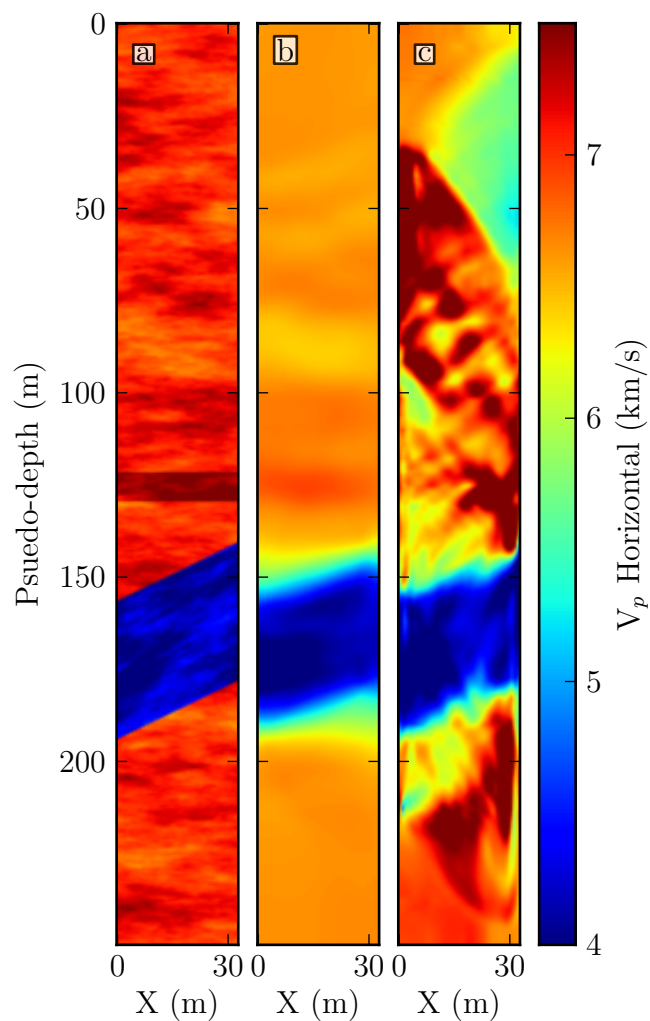


Figure 4.10: An-elliptic TTI synthetic test. (a) True velocity model (as shown in Figure 4.1); (b) Velocity model obtained from traveltime tomography; (c) Velocity model obtained from waveform tomography, using anisotropy estimated from simulated annealing. The anisotropy was an-elliptic and TTI, with  $\phi = 30^\circ$ .

I believe Figure 4.11 demonstrates the reason for the success of simulated annealing. The base figure in red-white-blue is the  $V_{true} - V_{image}$  from Figure 4.2c, overlain with a single, curved seismic raypath used in traveltime tomography (yellow) and the corresponding wavepath from 500 Hz waveform tomography (greyscale). Note the magnitude and sign of the velocity error sampled by either the raypath or wavepaths. Between this specific source-receiver pair, the ray is only sensitive to velocities which are on average too fast. This is a result of the unresolvability of the sharp velocity boundary by the ray-theoretical approximation, which stems from the fundamental resolution limit of ray-based inversion (discussed in Chapter 2). As well, we have forced  $\partial^2 \mathbf{m} / \partial \mathbf{x}^2 \leq \kappa$  where  $\kappa$  represents the highest wavenumber allowed, through the smoothness constraint, and features with a wavenumber higher than  $\kappa$  (appearing at the boundaries) go unresolved. These two effects result in the velocity residual pattern seen in Figure 4.11, as the velocity discontinuity is smoothed out over the boundary.

Infinitely thin rays, existing within erroneous velocity models, will experience a corresponding increase or decrease in traveltime. In most of the model, we see that this is not a major issue, as positive and negative errors tend to average out by the time the ray arrives at the receiver. The ray in Figure 4.11 however, is part of a small subset of rays which travel mostly within areas where the cumulative velocity error along the raypath is positive, and which as a result carry large positive traveltime residuals. To remedy this mismatch in traveltime error, either the velocity or anisotropy must be perturbed. Since the inverted velocity in the majority of the low velocity zone is roughly correct, perturbing velocities along these troublesome raypaths will not solve the problem, as it will increase the residuals between other source-receiver pairs. Instead, it appears that the traveltime tomographic algorithm resolves these errors through perturbations to the anisotropy model.

To see what is happening, let us examine the two inverted anisotropy models closely. In Figure 4.2b, we see that if anisotropy is constrained, parts of the low velocity zone erroneously become anisotropic, and additional lobes of strong anisotropy appear at the top and bottom boundaries. In Figure 4.2e, we see that if we relax the constraint on anisotropy, the low velocity zone becomes relatively isotropic (albeit with an inaccurate shape), but the boundary lobes acquire significantly more anisotropy.

These phenomena arise as follows: If the anisotropy is heavily constrained (i. e. if  $\kappa$  is large in equation 2.18), there is a significant penalty on any anisotropic model parameters which deviate from zero (this is equivalent to assuming *a priori* that the region is isotropic, and heavily weighting this information). As such, the magnitude of the lobes do not increase to the values seen in for the unconstrained case (Figure 4.2e).

Instead, increased anisotropy appears in the low velocity zone, corresponding with those highly-angled raypaths. If there is *no* constraint on anisotropy, the lobe anisotropy is allowed to grow arbitrarily large, and the traveltimes errors can be accounted for without invoking anisotropy in the low velocity zone.

There is another interesting feature in the traveltimes anisotropy models, and that is the difference in anisotropy in the high velocity zone between Figure 4.2b and e; specifically the anisotropy in Figure 4.2b is significantly lower than that in Figure 4.2e. This is likely due to the presence of horizontal layering in the true velocity models, which is not resolvable at the resolutions of traveltimes tomography. Backus (1962) showed that a horizontally layered medium is indistinguishable from a homogenous anisotropic medium in the long wavelength limit. Since the horizontal layering in the true model is beyond the resolution of traveltimes tomography, the inverted velocity model is relatively homogeneous compared to the true model, and thus exhibits *extrinsic*, or additional, anisotropy beyond the *intrinsic* value of  $\varepsilon = 15\%$ . Pratt and Sams (1996) showed the effects of extrinsic anisotropy on traveltimes tomography, and they conclude that fine layering beyond the resolution limit may indeed appear as extrinsically anisotropic when inverted using the ray-theoretical approach.

The question remains as to why simulated annealing, which uses a finite-difference implementation of the acoustic wave equation, succeeds in determining the correct (1-D) anisotropy model, whereas traveltimes tomography does not. On Figure 4.11, the wavepaths for a 500 Hz wave is depicted in greyscale. Comparing this to the equivalent raypath, we see a significant difference: The multipath nature of the wave equation allows velocities external to the raypath to be sampled. In this particular case, the wave propagates outside the low velocity zone, experiencing areas of both positive and negative velocity errors. I propose that by the time the wave arrives at the receiver, the cumulative error along the wavepaths results in the predicted phase to be roughly equivalent to the observed phase, thus negating the need to perturb any model parameters. As well, in the high velocity zone where the anisotropy was overestimated by traveltimes tomography, information contained in the waveforms supports the fact that the medium is finely layered, instead of being extrinsically anisotropic.

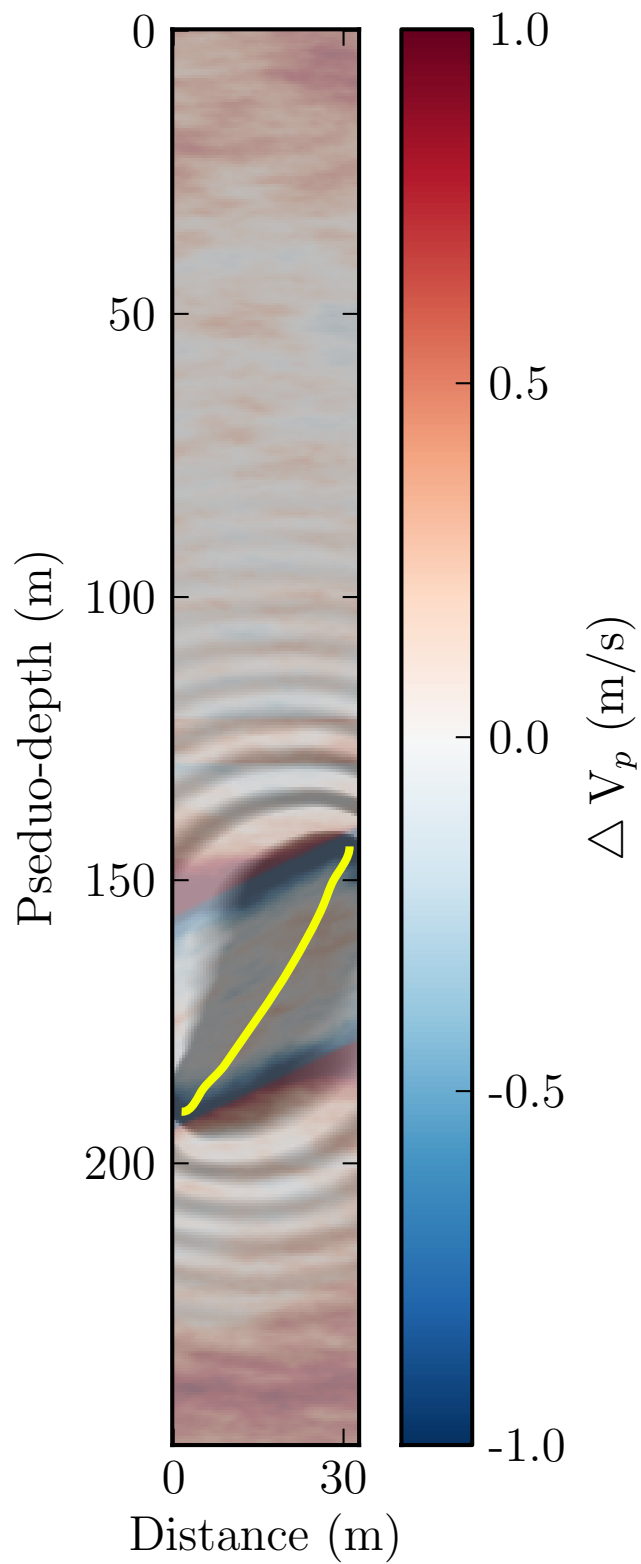


Figure 4.11: Raypath (yellow) and 500 Hz wavepath (greyscale), for a single source-receiver pair, superimposed on the travelttime inversion velocity error ( $\Delta V = V_{true} - V_{est}$ ).

# Chapter 5

## Conclusions

This thesis described the application of waveform tomography to a real, crosshole seismic dataset, as well as a synthetically generated crosshole dataset. The bulk of the original work presented involved the implementation of the semi-global inverse method of simulated annealing to solve a serious problem relating to the anisotropy model.

### 5.1 Conclusions and Discussion

In Chapter 1, the basics of anisotropic wave propagation were discussed. We saw that seismic anisotropy can be related to any symmetries present in rocks, and that symmetry and anisotropy are scale dependent. Thomsen's parameters (Thomsen, 1986) were introduced, which parameterize transversely isotropic P-wave propagation in terms of three parameters:  $V_{p0}$ , the P-wave velocity along the symmetry axis,  $\varepsilon$ , the fractional difference between P-wave velocity along and perpendicular to the symmetry axis, and  $\delta$ , a measure of how the velocity behaves at angles close to the symmetry axis. Traveltime and waveform tomography were also introduced.

Chapter 2 presented a basic discussion of inverse theory, and the mechanics behind traveltime and waveform tomography were explored. We saw that the traveltime inverse problem requires much *regularization* to stabilize the problem, which was introduced in the form of a priori data and model variances, as well as assuming that the parameters are spatially smooth. Waveform tomography was also discussed, and we learned that by using the back-propagation of the residual wavefield, the gradient of the misfit function could be computed with a vast reduction in computational effort. Finally, simulated annealing was introduced as an efficient, Monte-Carlo, semi-global inverse method, which searched for the global minimum of a function using concepts borrowed from statistical mechanics.

In Chapter 3, the application of the three inverse methods described in Chapter 2,

to a real, crosshole seismic dataset, was presented. We saw that, unfortunately, the data seemed to have been pre-filtered. As well, it was clear that the anisotropy models obtained from traveltimes tomography were not sufficient for use in waveform tomography, as the models failed the half-cycle criterion. The successful use of simulated annealing to determine a new anisotropy model was described, and the final velocity inversions from waveform tomography were interpreted, with promising results. Specifically, the identification of a 1 m thick massive sulphide vein, along with many other geological features missing from the traveltimes inversions, serve to promote the efficacy of waveform tomography in mineral exploration. It is hoped that these results will lead to future seismic surveys, specifically tuned for waveform tomography, which will allow Vale to produce high resolution wireframe models, while reducing the number of (expensive) boreholes drilled. Finally, to help prevent many of the issues that were found to detrimentally affect waveform tomography, a ‘checklist’ of recommendations for any future crosshole surveys was given.

Chapter 4 described the creation, modelling, and inversion of a series of synthetic crosshole datasets, modeled after the 540-542 dataset processed in Chapter 3. For the elliptic case, similar patterns in the traveltimes anisotropy models were seen, and again simulated annealing was successful in finding an anisotropy model which better fit the data (and the true model). These patterns, and the general failure of traveltimes tomography to accurately resolve anisotropy models in regions of strong heterogeneity, were explained in terms of the limitations of the ray-theoretical approximation, and we saw how the finite-difference wavepath used in waveform tomography mitigated these problems. Finally, an-elliptic VTI media were tested with the same processing flow (with good results), as well as an-elliptic TTI media (with poor results).

The failure of waveform tomography to resolve the correct velocity structure in severely tilted TTI media is not unexpected, but deserves some comment. The symmetry axis was tilted at  $30^\circ$  to the vertical, and throughout the majority of the model  $\varepsilon$  was set equal to 0.15 (which is large). Since, at the moment, our waveform tomography routine can only handle elliptic, VTI anisotropy, any deviations of the symmetry axis from vertical will corrupt the inversions. As stated in the borehole logs (Inco, 2000), multiple fracture axes present in the 542-540 dataset indicate that the symmetry axis may indeed not be vertical.

For several reasons, I do not think the deviation from the VTI assumption catastrophically corrupted the velocity inversion shown in Chapter 3. First, the traveltimes (starting) velocity model shown in Figure 3.5 is similar in structure to the final model from waveform tomography (Figure 3.9). This is *not* what we see when we compare the

initial and final models for the synthetic TTI case in Figure 4.10. The models here are much different, which is to be expected, as the travelttime model can handle TTI media. As well, the magnitude of the anisotropy in the real data is likely much less than that in the synthetic data. This being said, it is obvious that the incorporation of TTI into the next generation of our waveform tomography code is essential. Unfortunately, the Earth is not so kind as to provide us with the simple approximations which are present in *all* discrete, computational representations of reality. With each approximation that is dropped, we inch closer to the truth.

Another point that requires attention is the state of global and semi-global inverse methods in geophysical inversion. As I mentioned in Chapter 2, these types of methods have not been widely adopted, and with good reason. For many interesting problems, their computational cost is simply too high. The models presented in this thesis are *significantly* smaller than most exploration scale problems, and it is this fact which greatly assisted in the successful implementation of simulated annealing. While coding the simulated annealing subroutine, I parallelized the forward modelling code over seismic sources, and this cut the time per iteration by approximately half.

All simulations were run on the SharcNet (Shared Hierarchical Academic Research Computing Network) cluster *Orca*, and were parallelized over 16 compute cores. *Orca* has recently been expanded to include new Intel Xeon Sandy Bridge processors, and on these processors a single iteration takes approximately four seconds. For 3-D, industrial scale problems, the time taken for single forward modelling run is measured in terms of *hours*. There is obviously much more work to be done before simulated annealing can be feasibly used for problems of this size. If we are aiming to increase the dimensions of our parameter space and include *velocity* in our parameterization, the situation seems quite hopeless.

I would like to speculate, though, on what it might take to make these problems feasible. First, for large-scale problems, I don't believe it is necessary to use a global method to determine the final distribution of  $P(\mathbf{m})$ . A large body of very successful work on gradient-based optimization methods has shown that this approach works well, with some limitations. In waveform tomography, a major limitation deals with the choice of  $\mathbf{m}^0$ , or the initial model. If, though,  $\mathbf{m}^0$  is chosen properly, the robustness of the gradient algorithm negates the need to include any other *a priori* information (Tarantola, 2005).

Following the example of Tarantola (2005) (pp. 156), I believe a global approach may (at least currently) have a role to play in determining  $\mathbf{m}^0$ . If we are basing our misfit function on the goodness of waveform fit ( $\mathbf{u}_{est} - \mathbf{u}_{obs}$ ), then via the half-cycle criterion we realize that our  $\mathbf{m}^0$  needs to predict  $\mathbf{u}_{obs}$  to within one half cycle of the lowest



usable frequency, but needs to be no better. In this case, perhaps we should design each parameter perturbation to change  $\mathbf{u}_{est}$  by more than a half cycle at the lowest frequency. This saves time spent searching within the vicinity of the global minimum, as our efficient gradient-based search will perform this task later. To accomplish this, we could choose a random velocity parameter (and perturb a fresnel-zone sized volume centered on this parameter) with a magnitude that ensures an incident low-frequency wave will be delayed or advanced by at least one half cycle.

## 5.2 Final Remarks

Here I would like to expand on one more point. In Chapters 3 and 4, I referred to the anisotropy profile found through simulated annealing as the ‘ideal model’. This choice was deliberate, as it is ‘ideal’ within the restrictions introduced by all the approximations made thus far (i.e. acoustic wave propagation, 1-D anisotropy, elliptical anisotropy, etc.). It does *not* correspond to the true values of  $\varepsilon$  in the Voisey’s Bay region, as it must also account for the influence of  $\delta$ . Remember how this model was generated; it involved the consideration of *many* models after the ‘phase change’. We did not choose the model which simply decreased the residuals by the greatest amount, as we recognize our forward operator is not a perfect representation of reality. In this sense, it was not the final model obtained at the last iteration, nor was it the best fit model: it was the ideal model, under our assumption of Gaussian statistics.

When inverting for velocity we did not have this luxury of browsing a selection of models at leisure, due to the computational expense of waveform tomography. Here, our ideal and final models are the same. Given enough time, it would be desirable to explore the global minimum of *this* model space as well, and to construct a separate ideal model that may help mitigate some effects of the error in our forward problem. Of course, this is not a new concept, all models are based on approximations.

I make this distinction because there is sometimes hesitation when people talk of expensive, high resolution imaging techniques such as 3-D, elastic, or semi-global, full waveform inversion. Why should all this computational time and effort be spent on new techniques, when the old techniques (such as travelttime tomography) give very good results?

Sagan (1985) tells the story of Johannes Kepler, as he spent time in Prague in the early 1600s. Working with the theory that planets orbited the sun along circular paths, he struggled to find a model accurate enough to explain all the apparent ‘eccentricities’ of planetary motions. Using this theory, he was relatively successful, incorporating pro-

grade and retrograde motion about the orbital axes. Nevertheless, by examining the observational data with the utmost accuracy, he realized that some data just *did not* fit his model. In the end, the theory of circular orbits was incorrect. His complex retrograde orbits were just an *ideal* model, given incorrect assumptions. The advent of more powerful telescopes allowed more accurate data to be collected, and when these data were analyzed Kepler realized that planetary orbits were, in fact, ellipses. This is perhaps one of the earliest examples of the successes of high resolution numerical modelling. In a geophysical context, high resolution imaging techniques, as well as the study of non-local misfit structure, allow our ideal models to creep towards the true Earth parameters. In the example of Kepler, we see that increasing the numerical accuracy led to a wholly new model.

There is much more work that needs to be done, but one should not be too wary about the current computational cost of semi-global, or highly accurate, modelling techniques. Advances in computational power will allow the currently inefficient to become the norm, and the currently impossible to become... perhaps at least inefficient. And all the while, the Universe will hum along.

# Bibliography

- Aki, K. and Richards, P. G. (2002). *Quantitative siesmology*. University Science Books.
- Babuska, V. (1968). Elastic anisotropy of igneous and metamorphic rocks. *Studia Geophysica et Geodaetica*, 12(3):291–303.
- Backus, G. E. (1962). Long-wave elastic anisotropy produced by horizontal layering. *J. Geophys. Res.*, 67.
- Barbier, M. G., Bondon, P., Mellinger, P., and Viallix, J. R. (1976). Mini-SOISE for land seismology. *Geophys. Prosp.*, 24:518–527.
- Brenders, A. J. and Pratt, R. G. (2007a). Efficient waveform tomography for lithospheric imaging: implications for realistic, two-dimensional acquisition geometries and low-frequency data. *Geophysical Journal International*, 168:152–170.
- Brenders, A. J. and Pratt, R. G. (2007b). Full waveform tomography for lithospheric imaging: results from a blind test in a realistic crustal model. *Geophysical Journal International*, 168:133–151.
- Červený, V. (1972). Seismic rays and ray intensities in inhomogeneous anisotropic media. *Geophys. J. R. astr. Soc.*, 29:1–13.
- Červený, V. (2001). *Seismic ray theory*. Cambridge University Press.
- Červený, V. and Jech, J. (1982). Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic meda. *Journal of Geophysics*, 51:96–104.
- Chapman, C. H. and Pratt, R. G. (1992). Traveltime tomography in anisotropic media I. Theory. *Geophysical Journal International*, 109(1):1–19.
- Colton, D. and Kress, R. (1998). *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer.

- Cosma, C. and Enescu, N. (2003). Ore delineation by crosshole seismics, transmission and reflection imaging at Voisey's Bay, Canada. In *Extended Abstracts*. 7th International Symposium on Recent Advances in Exploration Geophysics.
- Dellinger, J. A. (1991). *Anisotropic seismic wave propagation*. PhD thesis, Stanford University.
- Dziewonski, A. M. and Anderson, D. L. (1981). Preliminary reference Earth model. *Physics of the Earth and Planetary Interiors*, 25(4):297–356.
- Enescu, N., McDowell, G., Cosma, C., and Bell, C. (2002). Crosshole seismic investigations at Voisey's Bay, Canada. In *Extended Abstracts*. 72nd Annual SEG Meeting.
- Evans-Lamswood, D. M., Butt, D. P., Jackson, R. S., Lee, D. V., Muggridge, M. G., and Wheeler, R. I. (2000). Physical controls associated with the distribution of sulfides in the Voisey's Bay Ni-Cu-Co deposit, Labrador. *Economic Geology*, 95:749–769.
- Fichtner, A., Kennet, B. L. N., Igel, H., and Bunge, H. P. (2009). Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International*, 179:1703–1725.
- Grechka, V. and Mateeva, A. (2007). Inversion of P-wave VSP data for local anisotropy: Theory and a case study. *Geophysics*, 72:D69–D79.
- Helbig, K. (1983). Elliptical anisotropy; its significance and meaning. *Geophysics*, 28(7):825–832.
- Inco (2000). INCO Technical Services Limited: Borehole core log.
- Iturrino, G. J., Christensen, N. I., Kirby, S., and Salisbury, M. H. (1991). 11. seismic velocities and elastic properties of oceanic gabbroic rocks from hole 735b. In *Proceedings of the Ocean Drilling Program, Scientific Results*, volume 118.
- Kamei, R., Hato, M., and Matsuoka, T. (2005). Random heterogeneous model with bimodal velocity distribution for methane hydrate exploration. *Exploration Geophysics*, 36:41–49.
- Kamei, R., Pratt, R. G., and Tsuji, T. (2012). Waveform tomography imaging of a megasplay fault system in the seismogenic Nankai subduction zone. *Earth and Planetary Science Letters*, 317-318:343–353.

- Kern, H. (1993). P- and S-wave anisotropy and shear-wave splitting at pressure and temperature in possible mantle rocks and their relation to rock fabric. *Physics of the Earth and Planetary Interiors*, 78:245–256.
- Komatitsch, D. and Tromp, J. (2002). Spectral-element solutions of global seismic wave propagation-1. Validation. *Geophysical Journal International*, 149:390–412.
- Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In Bednar, J. B., Redner, R., Robinson, E., and Weglein, A., editors, *Conference on Inverse Scattering: Theory and Application*, pages 206–220. SIAM Philadelphia.
- Lanczos, C. (1950). *Linear differential operators*. Van Nostrand.
- Marfurt, K. (1984). Accuracy of finite-difference and finite-element modelling of the scalar and elastic wave equation. *Geophysics*, 49:533–549.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., and E., T. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21.
- Milson, J. (2003). *Field Geophysics*. Wiley, 3 edition.
- Mosegaard, K. and Vestergaard, P. D. (1991). A simulated annealing approach to seismic model optimization with sparse prior information. *Geophysical Prospecting*, 39:599–611.
- Nulton, J. and Salamon, P. (1988). Statistical mechanics of combinatorial optimization. *Physical Review A*, 37(4).
- Paige, C. C. and Saunders, M. A. (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71.
- Park, C. B., Miller, R. D., Steeples, D. W., and Black, R. A. (1996). Swept impact seismic technique (SIST). *Geophysics*, 61(6):1789–1803.
- Perozzi, L., Gloaguen, E., Rondenay, S., and McDowell, G. (2012). Using stochastic crosshole seismic velocity tomography and Bayesian simulation to estimate Ni grades: Case study from Voisey’s Bay, Canada. *Journal of Applied Geophysics*, 78:85–93.
- Pratt, R. G. (1990). Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging. *Geophysics*, 55(5):626–632.
- Pratt, R. G. (1999). Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model. *Geophysics*, 64(3):888–901.

- Pratt, R. G. and Chapman, C. H. (1992). Traveltime tomography in anisotropic media II. Application. *Geophysical Journal International*, 109(1):20–37.
- Pratt, R. G., Hou, F., Bauer, K., and Weber, M. H. (2004). Waveform tomography images of velocity and inelastic attenuation from the Malik 2002 Crosshole Seismic Surveys. Technical report, Malik 2002 Gas Hydrate Production Research Programme.
- Pratt, R. G., McGaughey, W. J., and Chapman, C. H. (1993). Anisotropic velocity tomography: A case study in near-surface rock mass. *Geophysics*, 58(12):1748–1764.
- Pratt, R. G. and Sams, M. S. (1996). Reconciliation of crosshole seismic velocities with well information in a layered sedimentary environment. *Geophysics*, 61(2):549–560.
- Pratt, R. G., Shin, C., and Hicks, G. J. (1998). Gauss Newton and full Newton methods in frequency space seismic waveform inversion. *Geophysical Journal International*.
- Pratt, R. G., Sirgue, L., Hornby, B., and Wolfe, J. (2008). Crosswell waveform tomography in fine-layered sediments: meeting the challenges of anisotropy. In *Extended Abstracts*. 70th EAGE Conference and Exhibition.
- Pratt, R. G. and Symes, W. (2002). Semblance and differential semblance optimisation for waveform tomography: a frequency domain implementation. In *Journal of Conference Abstracts*, volume 7. Cambridge Publications.
- Pratt, R. G. and Worthington, M. H. (1990). Inverse theory applied to multi-source cross-hole tomography. Part 1: Acoustic wave-equation method. *Geophysical Prospecting*, 38:287–310.
- Prieux, V., Brossier, R., Gholami, Y., Operto, S., Virieux, J., Barkved, O., and Kommedal, J. (2011). On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International*, 187:1495–1515.
- Ryan, B. (1996). Commentary on the location of the Nain-Churchill boundary of the Nain area. Report 96-1, Newfoundland Department of Natural Resources.
- Ryan, B., Wardle, R. J., Gower, C. F., and Nunn, G. A. G. (1995). Nickel-copper sulphide mineralization in Labrador: The Voisey’s Bay discovery and its exploration implications. Report 95-1, Newfoundland Department of Natural Resources.
- Sagan, C. (1985). *Cosmos*. Ballantine Books.

- Sambridge, M. and Mosegaard, K. (2002). Monte Carlo Methods in Geophysical Inverse Problems. *Review of Geophysics*, 40(3).
- Sen, M. and Stoffa, P. L. (1995). *Global optimization methods in geophysical inversion*. Elsevier.
- Sen, M. K. and Stoffa, P. L. (1991). Nonlinear one-dimensional seismic waveform inversion using simulated annealing. *Geophysics*, 56(10):1624–1638.
- Shin, C. and Cha, Y. H. (2008). Waveform inversion in the Laplace domain. *Geophysical Journal International*, 173(3):922–931.
- Shin, C. and Cha, Y. J. (2009). Waveform inversion in the Laplace-Fourier domains. *Geophysical Journal International*, 177(3):1067–1079.
- Shin, C. and Min, D.-J. (2006). Waveform inversion using a logarithmic wavefield. *Geophysics*, 71(3):R31–R42.
- Sirgue, L. and Pratt, R. G. (2004). Efficient waveform inversion and imaging: a strategy for selecting temporal frequencies. *Geophysics*, 69(1):231–248.
- Smylie, D. E. (1992). The inner core translational triplet and the density near Earth’s center. *Science*, 255(5052):1678–1682.
- Tarantola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49(8):1259–1266.
- Tarantola, A. (2005). *Inverse Problem Theory*. Society for Industrial and Applied Mathematics.
- Thomsen, L. (1986). Weak elastic anisotropy. *Geophysics*, 51:1954–1966.
- Tsvankin, I. (2012). *Seismic signatures and analysis of reflection data in anisotropic media*. Society of Exploration Geophysicists, 3 edition.
- Um, J. and Thurber, C. (1987). A fast algorithm for two-point seismic ray tracing. *Bulletin of the Seismological Society of America*, 77(3):972–986.
- Vibrometric (2009a). *Vibrometric MH DW70 Mechanical Borehole Hammer*.
- Vibrometric (2009b). *Vibrometric SPH 6454 Swept Impact Piezoelectric Borehole Source*.

- Virieux, J. (1986). P-SV wave propagation in heterogeneous media: Velocity-stress finite-difference method. *Geophysics*, 51(4):889–901.
- Virieux, J. and Operto, S. (2009). An overview of full waveform inversion in exploration geophysics. *Geophysics*, 74(6):WCC1–WCC26.
- Whitham, A. G. and Sparks, R. S. J. (1986). Pumice. *Bulletin of Volcanology*, 48:209–223.
- Williamson, P. R. and Worthington, M. H. (1993). Resolution limits in ray tomography due to wave behavior: Numerical experiments. *Geophysics*, 58(3):727–735.
- Wirgin, A. (2004). The inverse crime. *arXiv:math-ph/0401050*.
- Woodward, M. J. (1992). Wave-equation tomography. *Geophysics*, 57(1):15–26.
- Wu, R. S. and Toksoz, M. N. (1987). Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, 52:11–25.
- Xu, K. and Greenhalgh, S. (2010). Ore-body imaging by crosswell seismic waveform inversion: A case study from Kambalda, Western Australia. *Journal of Applied Geophysics*, 70:38–45.
- Yan, J. and Sava, P. (2011). Elastic wave-mode separation for TTI media. *Geophysics*, 76(4):65–78.
- Zelt, C. and Barton, P. (1998). Three-dimensional seismic refraction tomography: A comparison of two methods applied to data from the Faeroe Basin. *J. Geophys. Res.*, 103(B4):9187–7210.



## Curriculum Vitae

**Name:** Michael Victor Afanasiev

**Post-Secondary Education and Degrees:** McMaster University  
Hamilton, ON, Canada  
2010 B. Sc. (Honours) Physics

University of Western Ontario  
London, ON, Canada  
2012 M. Sc. Geophysics

**Honours and Awards:** Dean's List  
McMaster University  
2010

MITACS scholarship  
University of Western Ontario  
2011

CSEG award for best geophysics talk  
AESRC Meeting, Kingston, ON  
2012

**Related Work Experience:** Field Geologist/Geophysicist  
Apex Geoscience  
Summer 2010

### Refereed Conference Publications:

Afanasiev, M., Pratt, R. G., Kamei, R. and McDowell, G. Crosshole tomography, waveform inversion, and anisotropy: A combined approach using simulated annealing: 74th EAGE Conference & Exhibition, Copenhagen, Denmark. *Expanded Abstracts*, W016.

Afanasiev, M., Pratt, R. G., Kamei, R. and McDowell, G. Crosshole waveform inversion: Quantifying anisotropy through simulated annealing: 82nd SEG Annual Meeting, Las Vegas, USA. *Expanded Abstracts*, SI 4.3.